LUND UNIVERSITY

Multi-Meson Dynamics in Chiral Perturbation Theory

Sjö, Mattias

2023

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Multi-Meson Dynamics in Chiral Perturbation Theory

MATTIAS SJÖ
DEPARTMENT OF PHYSICS | FACULTY OF SCIENCE | LUND UNIVERSITY

Multi-Meson Dynamics in Chiral Perturbation Theory

Multi-Meson Dynamics in Chiral Perturbation Theory

by Mattias Sjö

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisor: Johan Bijnens
Faculty opponent: Bastian Kubis

To be presented, with the permission of the Faculty of Science of Lund University, for public criticism in Rydbergsalen at the Department of Physics on Friday the 1st of December 2023 at 13:00.

Organization	Document name
LUND UNIVERSITY Department of Physics Professorsgatan 1 SE–223 63 LUND Sweden	**DOCTORAL DISSERTATION**
	Date of disputation 2023-12-01
	Sponsoring organization
Author(s) Mattias Sjö	

Title and subtitle
Multi-Meson Dynamics in Chiral Perturbation Theory

Abstract

This thesis concerns scattering of light mesons in the case where the number of external particles exceeds four (that is, three or more particles in the initial or final state), calculated using chiral perturbation theory. This is both of theoretical interest in the study of scattering amplitudes, and relevant for systems of three or more pions simulated using lattice quantum chromodynamics.

The introduction derives chiral perturbation theory from a background of quantum chromodynamics, which itself is motivated from general principles for the benefit of non-expert readers. Both theories are connected to the phenomenology of mesons and other hadrons, and related frameworks (including lattice quantum chromodynamics) are surveyed. The various computational techniques employed in the papers are summarized, related, and slightly expanded upon.

Paper I derives a generalized diagrammatic method for computing tree-level scattering amplitudes with a large number of external particles and including higher-order vertices, which serve as counterterms to loop diagrams. The treatment of flavor and kinematics is formalized, and some novel scattering amplitudes are computed explicitly.

Paper II applies the principles of analyticity, unitarity and crossing symmetry to the general two-to-two meson scattering amplitude in order to derive bounds on the otherwise free parameters that govern chiral perturbation theory. Previous results are generalized, both methodologically and in incorporating also the next-to-next-to-leading order of the amplitude. A mathematical framework is derived for organizing, improving and visualizing these bounds.

Paper III describes the calculation of the next-to-leading-order amplitude with six external particles, using a family of variants of chiral perturbation theory and mesons of nonzero but equal mass. Some concepts from paper I are used and improved.

Paper IV applies the amplitude considered in paper III to derive the three-pion K-matrix, which parametrizes finite-volume effects and pion mass dependence in three-pion systems in lattice quantum chromodynamics. This is done for the maximum-isospin channel, for which lattice results are available. A significant discrepancy is resolved between these results and the K-matrix obtained using only the leading-order amplitude.

Paper V generalizes the results of paper IV to all isospin channels, in anticipation of future lattice results and in preparation for further theoretical developments.

Key words
Scattering amplitudes, Effective Field Theories, Spontaneous Symmetry Breaking Nonlinear sigma model, Chiral Lagrangians, Hadronic Spectroscopy, Structure and Interactions, Lattice QCD

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language English
ISSN and key title	ISBN 978-91-8039-865-7 (print) 978-91-8039-866-4 (pdf)

Recipient's notes	Number of pages 399	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature _____

Date _____2023-10-17_____

Multi-Meson Dynamics in Chiral Perturbation Theory

by Mattias Sjö

LUND

UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover illustration front: Stylized representation of the three-to-three scattering that is central to papers III-V. Each set of three lines represents three mesons moving from one side to the other, interacting multiple times via Feynman diagrams. The interactions increase in complexity from top to bottom: simple two-particle subprocesses (compare figs. III.1 and III.2), then three-particle scattering at tree and one-loop level (compare figs. III.3 and III.4), and finally a few two-loop diagrams, foreshadowing a very complicated calculation that might happen some day. The backdrop is a misty forest captured by my grandfather Lars Johansson; with a bit of imagination, beautiful trees emerging from the mist can be taken as a metaphor for paper I, where beautiful tree diagram emerge from abstract considerations of symmetry.

Cover illustration back: Picture of me, wearing the traditional student cap among the blossoming magnolias in Lundagård, after taking part in the traditional welcoming of Spring on 1st of May, 2023.

Till farfar Rune
och morfar Lars

Contents

Publications

Paper IV: The isospin-3 three-particle K-matrix at NLO in ChPT   **279**

Acknowledgements

There are so many people who have, in one way or another, helped make this thesis reality. Not only that, but they have made it a pleasure to write! The place of honor naturally goes to my supervisor, Hans Bijnens,[1] who set the whole thing in motion. His dependability, no-nonsense guidance, and profound concern for the well-being of everyone in his little research group has been an inexhaustible resource to my growth as a scientist. He is, of course, also an absolute *wizard* of ChPT and everything related to it.

I am also greatly indebted to my other collaborators: Karol, who along with Hans introduced me to this topic, and who kindled my love for amplitudes; Benjamin, who taught me the noble way of the mathematician; Tomáš, with whom I worked side by side on some of the best parts of my Ph.D., battling the mighty bull's head and taking manuscript-writing to new levels—paper III, what a beauty!—and Steve, Fernando and Jorge, who made papers IV-V a joy and who helped me open the door to the lattice. I also wish to thank my ever-inspiring "academic older brother", Nils, and also Laurent, Antoine and the others who are bringing me through that lattice door into the next exciting chapter of my career.

Research aside, I have found sublime joy in teaching, for which I thank all my excellent students, as well as the lecturers—Malin, Torbjörn S, Leif L, Rikkert, Roman—who have given me these opportunities and lots of support. Speaking of teaching, holding word-of-the-day sessions has been a highlight of my day, and I thank everyone who took part in them with such enthusiasm. And speaking of support, I don't know how I could learn and teach without Rathsman, or use computers without Bosse and Christian, or where I'd even be without Caroline, Louise and the other grown-ups who ensure we confused scientists don't get lost. I also give my heartfelt thanks to Andrew, Timea, Leif (both of them), Jasmina, Tetiana, Emil, Robin and Harsh, who have helped me write and proof-read the thesis (or allowed me to help myself by helping them).

My office-mates—Hugo, Astrid, Jonas, Tomáš, Joan, Andrecia, Anca (sorry about the cold)—have all been awesome, and the same goes for the rest of the division: research is fun, but it wouldn't be the same were it not punctuated by all the activities kept going by Chiara, Jasmina and the other juniors, or the fascinating stories by Gösta and the other elders, or the endless discussions about obscure mathematics or whatever with Torbjörn L or Robin, or boardgames with Hannah and the gang, or cheerful coffee-room chats with Rikkert, Korinna or whoever else happens to be the agent of our great atmosphere at any given moment. I truly love the atmosphere at the division, and I love it all the more for how it is enriched by all the visitors and diploma students who pass through.

But life is so much more than work, no matter how enjoyable. I therefore thank everyone in Naturvetarkören and Ostrochorus for making my free time so melodious, and föreningen Rydbergs Minne for running my beloved ground state, pub Rydbergs. I also thank Anna-Maria, Johanne, Jennifer, Malin, Hanno, Tindra, Simon, Smita (both of them), Bibi, Brian, Ellika, Ellen and all my other siblings-in-arms at LDK, NDR, LUNA, TLTH and

---

[1]The author lists and bibliographies use his name, *Johan* Bijnens, but here I will consistently refer to him as *Hans*, which is what he is actually called.

List of publications

This thesis is based on the following publications:

I   **Higher-order tree-level amplitudes in the nonlinear sigma model**

Johan Bijnens, Karol Kampf, **Mattias Sjö**
JHEP **11** 074 (2019) [Erratum: *JHEP* **03**, 066 (2021)]
e-print: `arXiv:1909.13684 [hep-th]`

II   **NNLO positivity bounds on chiral perturbation theory for a general number of flavours**
Benjamin Alvarez, Johan Bijnens, **Mattias Sjö**
JHEP **03** 159 (2022)
e-print: `arXiv:2112.04253 [hep-ph]`

III   **Six-meson amplitude in QCD-like theories**

Johan Bijnens, Tomáš Husek, **Mattias Sjö**
Phys. Rev. D **106** (2022) 054021
e-print: `arXiv:2206.14212 [hep-ph]`

IV   **The isospin-3 three-particle K-matrix at NLO in ChPT**

Jorge Baeza-Ballesteros, Johan Bijnens, Tomáš Husek, Fernando Romero-López, Stephen Sharpe, **Mattias Sjö**
JHEP **05** 187 (2023)
e-print: `arXiv:2303.13206 [hep-ph]`

V   **The three-particle K-matrix at NLO in ChPT**

Jorge Baeza-Ballesteros, Johan Bijnens, Tomáš Husek, Fernando Romero-López, Stephen Sharpe, **Mattias Sjö**
In preparation for submission to *JHEP*

All papers are reproduced with the permission of their respective publisher, with minor stylistic changes in the layout and wording.

During my Ph.D. I have also published the following works, which are not included in this thesis but are listed here for the sake of completeness:

VI   **CELLoGeNe – an energy landscape framework for logical networks controlling cell decisions**
Emil Andersson, **Mattias Sjö**, Keisuke Kaji, Victor Olariu
iScience **25** (2022) 8, 104743
e-print: `bioRxiv:2022.02.09.479734`

VII   **NNLO positivity bounds on χPT for a general number of flavors**
Benjamin Alvarez, Johan Bijnens, **Mattias Sjö**
Contribution to the 10th international workshop on chiral dynamics (CD21)
e-print: `arXiv:2209.03208 [hep-ph]`

VIII   **Three-pion scattering: From the chiral Lagrangian to the lattice**
Jorge Baeza-Ballesteros, Johan Bijnens, Tomáš Husek, Fernando Romero-López, Stephen Sharpe, **Mattias Sjö**
Contribution to the 26th high-energy physics international conference in quantum chromodynamics (QCD23)
e-print: `arXiv:2309.17107 [hep-ph]`

IX   **The three-particle K-matrix at NLO in ChPT**
Jorge Baeza-Ballesteros, Johan Bijnens, Tomáš Husek, Fernando Romero-López, Stephen Sharpe, **Mattias Sjö**
Contribution to the 40th international symposium on lattice field theory (Lattice 2023)

Note that paper VI uses traditional (non-alphabetic) author ordering.

Populärvetenskaplig sammanfattning på svenska

Partikelfysik är läran om materiens allra innersta inre. Dock har det, gång efter annan, visat sig att det vi trodde var det allra innersta inte alls var det: ur atomer skakade vi fram kärnor, och ur dessa, *protoner* och *neutroner*; snart fann vi även deras släktingar och kallade dem *mesoner* och *baryoner*; därefter insågs det att dessa på svårbegripliga sätt var uppbyggda av vad vi fantasifullt kom att kalla *kvarkar* och *gluoner*, och som är det innersta vi har nått hittills; och gradvis sammanställde vi detta och mycket mer i den betydligt mindre fantasifullt namngivna *Standardmodellen*, en ganska elegant och mestadels komplett teori om nästan allt,[2] och som i sin mest kortfattade form går att skriva ner på en t-shirt eller kaffekopp. Standardmodellen inbegriper många delar med fantasieggande namn såsom *elektrosvag växelverkan* och *Higgsmekanismen*, men vill du veta mer om dem får du läsa en annan avhandling, för jag kommer ägna min odelade uppmärksamhet åt delen med det mest fantasieggande namnet av dem alla: *kvantkromodynamik* (eng. *quantum chromodynamics*, förkortat QCD), som beskriver just kvarkarna, gluonerna och deras växelverkan.

Att teoretiskt förutsäga partiklars exakta beteende utifrån QCD (eller, för den delen, resten av Standardmodellen) är tyvärr förkrossande svårt. I princip det enda fall vi kan lösa är detta: en ensam partikel som sitter i tomrum utan att växelverka med andra (och även där är "lösa" en sanning med modifikation). Räddningen finns dock i form av *störningsräkning* (eng. *perturbation theory*), där principen är att du börjar med något du faktiskt kan hantera—ett fåtal isolerade partiklar—och sedan tillåter du ett, två, kanske tre ögonblick av växelverkan, bara precis så att den process du vill studera över huvud taget kan hända. Mellan dessa ögonblick får partiklarna återgå till sina enkla, isolerade jag. Vanligtvis ger detta en grov men helt okej beskrivning av verkligheten, samtidigt som dessa få växelverkningar är så begränsade—bara en flyktig *störning* ovanpå den trygga tillvaron hos isolerade partiklar—att beräkningarna förblir enkla. Räcker inte detta kan en andra störning göras: tillåts det *näst minsta* antalet växelverkningar blir beräkningarna förvisso krångligare, men resultatet korrigeras och hamnar förhoppningsvis ytterligare lite närmre verkligheten. I denna avhandling stannar jag där, men det går såklart att fortsätta; i vissa extremt precisa beräkningar kan hela sex störningar, den ena krångligare än den andra, komma på fråga.

Dessvärre fungerar störningsräkning inte alltid. Typexemplet är QCD vid låg energi, där kvarkar och gluoner binds samman till baryoner och mesoner. Dessa nya partiklar är så fullkomligt olika en samling av isolerade kvarkar och gluoner att det är lönlöst att ta sådana samlingar som utgångspunkt; försöker du likt förbannat arbeta dig igenom störningsräkningen finner du att varje steg ändrar resultatet till oigenkännlighet istället för att vara den finjustering det borde vara. Det finns alltså en till synes oöverbryggelig avgrund mellan hur QCD ter sig i Standardmodellen och hur den yttrar sig i vår lågenergetiska vardag.

Kiral störningsräkning (eng. *chiral perturbation theory*, förkortat ChPT) är en väg över den-

---

[2]Med detta menas, enligt kosmologernas uppskattningar, ungefär 5% av Universums innehåll. Fast resten är mörk materia, som vi inte har någon aning om vad det är, och mörk energi, som vi vet ännu mindre om. Så Standardmodellen beskriver i alla fall allt som har någon form av påverkan på vår vardag. Eller, just det— gravitation är en grej, och lyser med sin frånvaro i Standardmodellen. Attans! Verkar som att vi fysiker har en del kvar att lista ut trots allt, och då har jag ändå inte kommit till vad jag försöker lista ut i den här avhandlingen. Vad sägs om att vi avslutar den här fotnoten och går vidare? Bra—tillbaka till texten med dig nu.

na avgrund. Istället för att utgå från isolerade kvarkar och gluoner utgår vi från isolerade mesoner, vars inre struktur vi helt sonika sopar under mattan, och eftersom utgångspunkten då ligger i samma härad som verkligheten, fungerar störningsräkning igen! Det kan tyckas att vi nu fuskar och tappar all teoretisk grund, men icke! Det finns grundläggande symmetrier hos QCD som, oavsett hur mycket krångel som gömmer sig under mattan, måste finnas kvar hos mesonerna, och andra symmetrier bryts på mycket regelbundna sätt: medan kvarkar förekommer i två former som är varandras spegelbilder, precis som vänster och höger hand,[3] är en meson sin egen spegelbild, som en näsa.[4] Alla dessa symmetrier sätter stränga matematiska villkor på hur mesoner kan växelverka. Till första störningen tillåts bara två enkla sätt att växelverka, till andra störningen ytterligare 12, och så vidare. Till varje växelverkan hör ett tal, en parameter, som mäter dess styrka, och i våra drömmars land går dessa att teoretiskt härleda från QCD. I praktiken måste de tillföras från experiment eller liknande, men likväl kan ChPT alltså ta oss från närmast total okunnighet till en beskrivning av verkligheten som är entydig sånär som till en handfull parametrar att justera. Mycket bättre än så går inte att säga ens om självaste Standardmodellen!

Med detta inte sagt att allt är solsken och fågelkvitter när det kommer till ChPT. Jämfört med QCD:s rena form är den klumpig, och även enkla beräkningar med den tenderar att skena iväg till oöverskådliga drivor av formler. Om det introduceras så mycket energi att mesonerna börjar delas upp i sina beståndsdelar bryter ChPT fullkomligt ihop, precis som QCD gör i det motsatta fallet när mesoner börjar formas. Redan när du har tre eller fyra mesoner är denna sammanbrottsenergi föga större än vad som redan finns till hands genom $E = mc^2$, så ChPT fungerar bara i ett snävt spann av partikelfysikens breda energiskala. Inom detta spann sker dock många intressanta och viktiga processer, och där verkar även *gitter-QCD* (eng. *lattice QCD*), ett helt annat sätt att korsa QCD-avgrunden.

I gitter-QCD begränsas rummet till en liten låda, stor nog att på sin höjd rymma en handfull atomkärnor, och delas sedan upp i ett rutnät, ett gitter, ungefär som när en digitalkamera pixelerar vår i övrigt högst opixliga verklighet. Detta, plus några matematiska konstfärdigheter, förenklar problemet med lågenergi-QCD så att en väldigt, *väldigt* kraftfull dator kan hantera det. På så vis simulerar man från grunden hur kvarkar och gluoner beter sig på detta gitter, och kan studera saker som är otillgängliga både för teori och experiment, som till exempel protonens och neutronens inre struktur. Gitter-QCD lider dock av flera problem kopplade till hur rymden (visar det sig visst) inte består av ett gitter i en låda, och det är här ChPT, som lever i vår vanliga, oändliga rymd, kommer in. Med viss ansträngning kan dess förutsägelser översättas till ett språk som gittret förstår och därmed kompensera för dessa problem. I gengäld bistår gittret med beräkning av ChPT:s parametrar.

Det är i denna symbios mellan ChPT och gitter-QCD som jag kommer in. I den avhandling du nu håller i handen gör jag diverse framsteg, både i beräkningar med ChPT och i översättningen till gitter-språk. Därmed drar jag med stor glädje mitt strå till den stora, underbara, ständigt växande (men alltjämt ofullständiga) stack som är vår gemensamma förståelse av materiens allra innersta inre.

---

[3] Ordet "kiral" kommer av grekiskans χείρ, "hand". Det hela är besläktat med hur "kiral" används inom kemin.

[4] Ordet "meson" kommer dock *inte* av grekiskans ord för "näsa", utan istället från μέσος, "mellan", eftersom mesoner är lite sådär mellantunga med partikelmått mätt. Som jämförelse kommer "baryon" av βαρύς, "tung".

Popular summary in English

Particle Physics is the study of the innermost nature of matter. Or, well, time and again we have found that what we thought to be innermost wasn't so at all: from atoms we pulled nuclei, and from those, *protons* and *neutrons*; soon, we also found their cousins and called them *mesons* and *baryons*; then we realized that these were composed, in confusing ways, of what we imaginatively called *quarks* and *gluons*, and which are the deepest level we've reached so far; and gradually we compiled all of this and much more into the much less imaginatively named *Standard Model*, a somewhat elegant and mostly complete theory of almost everything,[5] and which, in its briefest form, can be written on a t-shirt or a coffee mug. The Standard Model contains many parts with intriguing names such as the *electroweak interaction* and the *Higgs mechanism*, but you'll have to read another thesis if you wish to know more about them, for I will give my undivided attention to the part with the most intriguing name of them all: *quantum chromodynamics* (abbreviated QCD), which describes the quarks, the gluons, and their interactions.

Unfortunately, it is mind-numbingly difficult to theoretically predict the exact behavior of particles based on QCD (or, for that matter, the rest of the Standard Model). In principle, the only case we can solve is that of a single particle sitting in a void without interacting with any others (and even then, "solve" is a bit of an overstatement). Salvation comes in the form of *perturbation theory*, in which you start with something you can actually handle— a few isolated particles—and then allow one, two, maybe three instant interactions, just enough that whatever process you wish to study can happen at all. Between these instances, the particles resume their simple, isolated existences. This usually gives a rough but quite decent description of reality, but at the same time, the few interactions are so limited— just a fleeting *perturbation* on top of the safety of isolated particles—that the calculations remain simple. If this isn't enough, a second perturbation can be used: allowing the *second smallest* number of interactions does complicate the calculations, but the result gets refined and hopefully ends up even closer to reality. That is where I stop in this thesis, but one can of course continue; in certain extremely precise calculations, as many as six perturbations— each more complicated than the last—may be on the agenda.

Unfortunately, perturbation theory doesn't always work. The classic example is QCD at low energy, where quarks and gluons fuse into baryons and mesons. These new particles are so fundamentally different from a collection of isolated quarks and gluons that it is futile to use that as your starting point; if you nevertheless try to work through the perturbations, you'll find that each step changes the result beyond recognition, rather than being the slight refinement it's supposed to be. Thus, there is a seemingly insurmountable chasm between QCD as it appears in the Standard Model and in our low-energy lives.

Chiral perturbation theory (abbreviated ChPT) is a way across this chasm. Instead of start-

---

[5]By this I mean, according to the cosmologists' estimates, about 5% of the contents of the Universe. But the rest is dark matter, of which we know hardly anything, and dark energy, of which we know even less. So at least the Standard Model describes everything that has any effect on our everyday lives. Or, right—gravity is a thing, and is blatantly absent from the Standard Model. Damn it! It seems us physicists still have a thing or two to figure out, and I still haven't got around to what I'm trying to figure out in this thesis. How about we end this footnote and move on? Good—get yourself back to the text, now.

ing from isolated quarks and gluons, we start from isolated mesons, promptly sweeping their inner structures under the carpet, and since the starting point now lies in the same ballpark as reality, perturbation theory works again! You may think that in doing this, we cheat and lose our theoretical footing, but that's not the case! There are fundamental symmetries of QCD that, regardless of the mess that hides under the carpet, must remain with the mesons, and other symmetries are broken in highly regular ways: while quarks come in two varieties, which are mirror images just like your left and right hand,[6] a meson is its own mirror image, like your nose.[7] All these symmetries impose strict mathematical conditions on how mesons may interact. For the first perturbation, only two simple kinds of interaction are allowed, for the second one 12, and so on. Each interaction is associated with a number, a parameter, determining how strong it is, and in our dreams these can be derived from QCD. In practice, they must be taken from experiments or the like, but still, ChPT takes us from near-complete ignorance to a description of reality that is unambiguous except for a few parameters that need adjustment. One couldn't say much better of the Standard Model itself!

This is not to say that all is fun and games with ChPT. It is cumbersome compared to QCD, and even simple calculations tend to snowball into incomprehensible piles of formulae. It breaks down completely if one introduces enough energy that the mesons start to separate into their constituents, much like how QCD does in the opposite situation where mesons start to form. Already with three or four mesons, this breakdown energy is hardly greater than what is already at hand through $E = mc^2$, so ChPT only works within a narrow segment of the wide energy range of particle physics. Within that range, though, live many interesting and important processes, and it is also the realm of *lattice QCD*, an entirely different way across the QCD chasm.

In lattice QCD, you limit space to a tiny box, big enough to host a handful of atomic nuclei at most, and then divide it into a grid, a lattice, like how a digital camera pixelates our otherwise rather un-pixely reality. This, plus some esoteric mathematics, simplifies the problem of low-energy QCD so that a very, *very* powerful computer can handle it. Thus, you simulate from first principles how quarks and gluons behave on this lattice, and can study things that are inaccessible both to theory and experiment, such as the inner structures of protons and neutrons. However, lattice QCD suffers from several issues connected to how space (so it turns out) does not consist of a lattice in a box, and it is here that ChPT, which lives in our normal, infinite space, enters. With some effort, its predictions can be translated into a language that the lattice understands, thereby compensating for these issues. In return, the lattice aids ChPT by calculating its parameters.

It is in this symbiosis between ChPT and lattice QCD that I enter. In the thesis you now hold in your hand, I make an assortment of contributions, both to ChPT calculations and to the translation into lattice-speak. Thereby, I gladly do my small part in the great, glorious, steadily progressing (yet still incomplete) quest that is our common understanding of the innermost nature of matter.

---

[6]The word "chiral" comes from the Greeek χείρ, "hand". This is related to the use of "chiral" in chemistry.

[7]The word "meson" does *not* come from the Greek word for "nose" however, but from μέσος, "middle", since mesons are middleweights as far as particles go. For comparison, "baryon" comes from βαρύς, "heavy".

Introduction

On a few hundred pages in a book there lived a thesis. Not a massive, dense monograph thesis, filled with groundbreaking results and an arcane technicality, nor yet a dry, bare, minimalist compilation thesis with nothing in it but a summary and a few papers: it was Mattias' thesis, and that means comfort.

— J. R. R. Tolkien, in "The Hobbit" (1937), somewhat paraphrased

You are holding in your hand the culmination of four and a half years of research, itself the culmination of almost ten years of studying at Lund University. It is a mighty tome, thicker than it maybe should be. This is partly because it holds five research papers, none of which are short and one of which—paper II—is staggeringly long. But it is also because I have poured a lot of love and effort into this introduction, seeking not to just summarize and contextualize the papers, but to provide more background and explanation than one can normally indulge in when writing a paper, and to make it just a little bit less arcane than theoretical particle physics tends to be.[8]

Writing this thesis, and doing the work behind it, has been great fun, and I hope that the bit of Feynmannian flair and Bengtssonesque quirks I've tried put in some explanations can help it be fun to read as well. Where it fails to be fun, I of course hope it can at least be informative; I have often looked to previous theses (especially Johan Thorén's [1] and Nils Hermansson-Truedsson's [2]) for good explanations of things that research papers gloss over and textbooks never get to, and I dream that for some future doctoral students, this thesis will be that kind of resource.

I therefore hope that you, dear reader—whether you are a member of my defense committee who knows more about most things than I do, or one of my young relatives who still has almost all of the fun learning ahead of you—will read on, at least for a bit. I will start gently, assuming my reader has very little background knowledge (but hopefully keeping it interesting also for those who do). I will gradually make things more technical, so that—like the apocryphal frog in a slowly heated pot which doesn't notice that it's boiling until it's too late—you find yourself at page 89, feeling like a theoretical particle physicist and eager to at least taste the contents of the papers.

---

[8]And maybe—just maybe—because I like using many words and almost as many footnotes.

1   The bigger picture

We can't see dark matter, but it doesn't matter. Matter matters.

— *Stefania Xella (2019)*

When looking for needles, you need to learn a lot about hay.

— *Giulia Repellino (2023)*

Particle physics is a broad, sprawling topic, and holds the honor of being the most fundamental of the physical sciences. It is highly mature as a field of research, and all the low-hanging fruit has been picked; progress happens through the collaboration of hundreds or thousands of people, or in tiny increments by many small groups of theorists. Given this, along with the often obtusely technical nature of the progress, it is easy to lose sight of the broader context of one's work.

Particle physics has moved through several stages in its existence. The first few subatomic particles—the electron, proton, neutron, photon, muon, and so on—were discovered around the first few decades of the 20th century, and around the same time, the union of quantum physics and special relativity into *quantum field theory* (QFT) laid a sound theoretical groundwork that stands to this day. One specific instance of QFT, *quantum electrodynamics* (QED) soon found great success in describing photons and electric charges.

After the world wars, particle physics entered a stage of rapid discovery. A confusing mess of new particles, which I attempt to survey in section 2.1, was found, and QFT fell short of explaining their properties and the *strong force* that governed them. However, theory gradually caught up during the 1960s, and during a short and exciting time in the 1970s, several theoretical developments and experimental discoveries came together to unlock the puzzle. A theory not too unlike QED, *quantum chromodynamics* (QCD, covered in section 2.2) was at last confirmed to explain the strong force. Thus, QED and QCD, along with an equally satisfying description of the *weak force*, formed our current paradigm for particle physics, the *Standard Model* (SM). Soon thereafter, lattice QCD (described in section 5) was born, and helped clarify less easy-to-calculate aspects of QCD.

The subsequent decades were, if less exciting, all the more satisfying. The SM predicted a number of hitherto unseen particles, and they were all discovered one after the other, as our ability to search for them progressed at a steady pace. But while our scientific capabilities grew, the stream of discoveries began to dry up. Within my lifetime, the number of truly groundbreaking discoveries can be counted on my fingers, and the number of newly discovered fundamental particles (the tau neutrino and the Higgs boson) can be counted on my thumbs. After passing much scrutiny and having all its key predictions confirmed, the SM is left only with its glaring inability to explain gravity or dark matter. Gravity is more of a formal issue—we have a perfectly good understanding of it, theoretically inconsistent with QFT but with no clashes that we can actually measure—but dark matter poses a worse problem, one that can hopefully be solved through a suitable extension of the SM

without razing its QFT foundations. But without clear hints as to how such an extension should look, theorists are left formulating a vast variety of tentative *beyond-the-SM* (BSM) theories,[9] with experiments blindly searching for signals that may confirm one of them.

I do not wish to imply that particle physics has entered a period of stagnation, though. Even though the SM looks almost exactly the same as it did when first written down in the 1970s, the precision and confidence we have regarding its details are at a completely different level. Just look at the "history plots" section of the 2022 PDG review [3, p. 19],[10] which summarizes how several important measurements have evolved since they were first performed. The oldest error bars are as wide as the plots, while the newest ones can barely be seen at all! Particle physics has truly entered an era of *precision*.

Particle physics may not look that precise at first—look ahead to page 56, and you will see that the parameters that govern my work are seldom known to better precision than a percent or so—but one has to bear in mind that particle processes are incredibly complicated and noisy, so extracting the signal one is looking for is like finding a needle in a haystack. Modern particle experiments are marvels of data collection and analysis, and the capability of the biggest detectors to filter out extremely weak signals can be compared to a scale so robust yet sensitive that you can drive a truck onto it and, engine still running, weigh a single grain of sand on its trailer with the accuracy of a chemist's analytical balance.[11]

The haystack analogy is an apt one, since I am very much studying the hay, not looking for the needle. This is not a BSM thesis, where I propose some elegant extension of the SM and extract its most salient predictions, hoping that it may someday be confirmed or at least confidently excluded. It is also not a thesis studying the "cool" parts of the SM, with heavy, relatively recently discovered particles at extremely high energy (for that, see Timea Vitos' thesis [4]). No, this is a thesis studying some of the most mundane SM particles, which are so amply produced in high-energy collisions that they are seldom even thought of as particles, but as anonymous components of particle jets dumped en masse into the detectors. This is a thesis about light *mesons*, and especially about the lightest and most mundane among them, the *pions*.

Light mesons are rather peculiar manifestations of QCD, and are best studied with a reformulation thereof called *chiral perturbation theory* (ChPT), which I describe in great detail in section 3. ChPT only applies in the low-energy regime,[12] but there is plenty of science to be had there. Take the famous *anomalous magnetic moment of the muon*, for example. This aspect of how a certain charged particle interacts with a photon is almost entirely the domain of QED, and has been both measured and calculated to some of the highest

---

[9]Here, as in "string theory", the word "theory" is used more in the mathematical sense of "self-consistent system constructed by theorists", and not in the scientific sense of "well-tested description of reality". The SM, which by all accounts is a proper scientific theory, humbly calls itself a model.

[10]The PDG review authoritatively summarizes virtually all of particle physics every few years. Its contents are also freely available online at `https://pdg.lbl.gov/`.

[11]See `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/index.html` and `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined` for visualizations of the vast range in cross-section (i.e., signal strength) measured at the Large Hadron Collider.

[12]"Low-energy" here means that the total energy involved is significantly smaller than 1 GeV, which is small by particle physics standards but still vastly larger than what is typical for nuclear (a few MeV) or atomic/chemical (a few eV) processes. (Readers who are unfamiliar with eV should look at page 6.)

precisions ever achieved in any science. Yet, theoretical predictions do not quite match experimental results [5],[13] even when the minute disturbances due to the rest of the SM are taken into account. Of these, the effect of mesons is among the largest and most uncertain, bringing these particles to the forefront of a very exciting corner of the haystack. That is the subject of Nils Hermansson-Truedsson's thesis [2] though, not mine.

Most ChPT calculations concern few-body processes like decays of a single meson, inserting a meson pair into another process, or the scattering of two mesons. I instead study processes involving at least three mesons, which is what the "multi-meson dynamics" in the title refers to. Such processes are much less well-studied, partly because they have fewer applications (nobody is colliding more than two beams at a time), and partly because they are difficult: the kinematics of few-body systems are very limited, but there is much more freedom in larger systems, so everything becomes more complicated.[14]

Something being difficult is, of course, its own reward, and multi-meson dynamics in simplified versions of ChPT has long been of interest in purer theoretical areas. Paper I operates in that relatively active realm, but when Hans and Tomáš sat down with more realistic ChPT to calculate the six-pion amplitude [6] that we subsequently generalized in paper III, they improved on something that had not been substantially worked on for 50 years [7–9]. That is not to say that their effort was uncalled for: six-meson processes have recently garnered much interest because three-meson systems have become an active topic in lattice QCD. These are, in turn, particularly relevant for modeling heavy, highly unstable particles (that is, *resonances*) that predominantly decay into three lighter ones, of which there are several interesting examples. As I will explain in section 5.3, the properties of such a system of three particles can be related to the properties of those same particles scattering elastically against each other, which is described by a six-particle amplitude (three in, three out). Such a comparison builds on the strengths of both ChPT and lattice QCD, and is worked out to great success in papers IV and V.

Well, that just about summarizes it. Gone are the days when we stepped on needles whenever we went for a walk; now, the haystack of particle physics is combed ever finer, granting insights and honing our collective skills as we go. I do not actively search for the needle of dark matter or whatever comes next, but I do find great joy in digging ever deeper into whatever my mathematical tools find purchase on. Every development gives the next big breakthrough fewer places to hide.

---

[13]For the non-academics reading this: a nice thing about particle physics is that virtually everything published in my lifetime is made openly available at the arXiv, `https://arxiv.org/`, with links given under each reference. The arXiv contains preprints, not actual papers, but these tend to be nearly identical.

[14]Some terminology for the non-experts: *kinematics* concerns the universal rules of how things *can* move, while *dynamics* concerns the specific rules of how they *do* move. For example, if a particle beam collides with something stationary, most of the beam's hard-earned energy has to go into maintaining the momentum of whatever comes out, since momentum cannot be created or destroyed; this is why we collide beams head on, so that the opposite momenta cancel out and all the hard-earned energy can be spent on interesting things. All of that is kinematics, and dynamics only enters when thinking about what those interesting things actually are.

2   Introduction for the non-experts

> *Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which are there.*
>
> — *Richard Feynman, in "The Character of Physical Law" (1964)*

Theoretical particle physics is a notoriously difficult topic to familiarize oneself with, and sadly, my work isn't of that elegant kind where the results can be understood without much insight into the technical details. But I'll be damned if I resort to hiding behind such a thin excuse as "it's complicated"! In the following pages, I will do my very best to convey to you, not the general idea of the concepts as in the popular introduction, but the heart and soul of the actual machinery. It will be quite a mouthful to digest, but it should give a bright and curious mind at least a fighting chance against the arcane lore of the main text. (Experts of that lore are of course free to read on for entertainment, or skip ahead to page 40.)

Before going about it, let me spend the remainder of this page discussing *how* I'm going about it. Particle physicists prefer condensing a lot of information into a few symbols, with much meaning left unstated in the assumption that the reader already knows it. I could try to make things accessible by being more explicit, writing out all the details and carefully working through concrete examples. But then I'd be writing a textbook, which there certainly isn't room for in this already-overlong thesis, and textbooks are for students hoping to *learn* how to do this themselves, rather than readers hoping to just *understand* what it means. Writing out all the details, I'd likely just cause sensory overload. I could dumb things down instead, and explain the rough ideas in words without formulae, but then I'd be very far from the true nature of this very mathematical field, and I'd be doing a disservice to anyone hoping to push a bit further beyond page 40. (So be ye warned: beyond this page, there be mathematics!)

What I *have* ended up doing is taking the particle physicist's preference to the extreme: I condense the formulae down to their most minimalist forms, eschewing detail for a bird's-eye view of their core meaning. With words, I try to highlight the take-home messages while at the same time exposing the mathematical bones of the arguments. In short, I try to make statements that are structurally simple and conceptually abstract.

Lacking detail, I advise you to lean on the *duck principle*: if it walks and swims and quacks like a duck, you might as well call it a duck. Likewise, if it adds and multiplies and transforms like a vector, you call it a vector, even if what is "under the hood" is vastly more complicated than the $\vec{x} = (x, y, z)$ represented as an arrow on a blackboard. If you hear me say "vector" while pointing at some mysterious infinite-dimensional thing, you don't think "this is a mysterious infinite-dimensional thing, and I'm scared"—you think "ah, I can manipulate it just like I could (x, y, z), and I shall fear no evil". Asking what things *are* is a dangerous, mind-bending, sometimes unanswerable question, but asking what things *do* is easy, especially when you just need them to do some of the things they can. A duck is a duck, and a superintelligent AI whose sound card only supports quacking is, for many purposes, also a duck. With that, dear reader, I wish you the best of luck.

Throughout the introduction, I will place these boxes containing information on basic topics. For the non-experts, they should alleviate the problem of not having all the background knowledge that the main text takes for granted; most things that are still left unexplained ought to be within Wikipedia's reach. For the experts, they should help clarify my notations and conventions.

For a warm-up, in this first box I will remind that ∂ stands for a *partial derivative*, $\int$ for an *integral*, and $\sum$ for *summation*.[i] The symbol $\equiv$ means "equal by definition", and the *complex conjugate* of z is written z^*, not $\bar{z}$ as some silly mathematicians will tell you.

I would also like to declare my use of *natural units* throughout.[ii] That is, I choose the *electronvolt*, eV, as the unit of energy, and then set the speed of light and the reduced Planck's constant to unity: $c = \hbar = 1$. This equates energy and mass (since $E = mc^2$ is now $E = m$) as well as energy and frequency (since $E = \hbar\omega$ is now $E = \omega$). Setting $c = 1$ also removes the distinction between distance and time,[iii] and since time is inverse frequency, all of spacetime can be measured in inverse eV. The conversions to "ordinary" units are

$$1\,000\,000\,000 \text{ eV} = 1 \text{ GeV} = 0.1602176634 \text{ nJ} = 1.78266192 \text{ yg},$$
$$1 \text{ GeV}^{-1} = 0.197327 \text{ fm} \qquad = 0.6582119 \text{ ys}. \tag{2.1}$$

Since the 2019 SI redefinition, these relations are *exact*. Furthermore, I express charges in units of the elementary charge, i.e. the charge of the proton.

---

[i]Particle physics plays rather fast and loose with all three: $\partial_x f$ means the same as $\partial f/\partial x$; $\int \mathrm{d}^4 x$ the same as $\int \mathrm{d}t \int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z$ (each from $-\infty$ to ∞); and $\sum_i$ means "sum over whichever i are relevant", because something like $\sum_{i=1}^n$ or $\sum_{i\in\mathcal{I}}$ is too much of a hassle. Note also the convention to write the integration measure (the $\mathrm{d}x$ bit) immediately after the $\int$ rather than at the end of the integral, which makes sense once they grow complicated enough.

[ii]At least, I use what most particle physicists call "natural units", where "natural" means that some fundamental constants are fixed to 1 so that the number of basic units is reduced from the cumbersome set of seven that the SI prescribes. In other contexts, different constants may be set to unity, and the decision to express things in terms of eV's might be considered a bit queer.

[iii]A popular example of this is the unit *light-year*. If you measure time in years and distance in light-years, then you automatically have $c = 1$: light travels one light-year per year! The same happens if you measure using minutes and light-minutes, seconds and light-seconds, and so on. At some point, you realize that you can just drop the "light-" and express both time and distance in terms of your chosen time unit. That c is normally some big, weird number is just an artifact of how our unit of distance (meter) is much shorter than what light travels in our unit of time (second), which in turn is an artifact of how *slow* humans are.

2.1 Invitation: the Particle Zoo

> …ἐνεῖναι πολλά τη καὶ παντοῖα ἐν πᾶσι τοῖς συγκρινομένοις καὶ σπέρματα πάντων χρημάτων καὶ ἰδέας παντοίας ἔχοντα καὶ χροιὰς καὶ ἡδονάς.
>
> …there exist many things of all kinds, the seeds of all objects, having all sorts of forms and colors and flavors.
>
> — *Anaxagoras (d. 428 BCE), fragment B4*

Open any thesis from my division in Lund, and you'll most likely find a thorough description of the SM near the beginning. I have decided not to do one of those, since frankly, the physics I work on only cares about a very small part of the SM: only half the quarks come into play, few leptons are mentioned even in passing, and the W, Z and Higgs bosons could just as well not exist. Therefore, I'll leave all those particles alone and instead describe the particles that *do* figure in my work, and some of their close relatives. In the more confusing times that were the 1950s, roughly this collection of particles was known as the *particle zoo*, before early versions of the SM began to clear things up.

One important thing to note is that none of the zoo particles are *fundamental*, unlike the electrons, photons, etc., which (as far as we know) don't consist of any smaller parts. But they are very tightly knit together, much more so than atoms or their nuclei, so tightly that it's fundamentally impossible to break them apart and study their components in isolation. Even so, we can glean a bit about the subcomponents by looking at what properties the bigger particles have. And since this thesis is about studying the properties of the smallest of these bigger particles—the *pseudoscalar mesons*—I figured it'd be worth a few pages to really wrap our minds around them before moving on.

Splitting atoms. Let me begin with an analogy about the biggest particles any particle physicist would dare to look at, namely *atoms*. These are electrically neutral, but with a slight nudge (hitting them with ionizing radiation, or doing whatever it is chemists do all day) they can be turned into electrically charged *ions*. Importantly, electric charge is *quantized*: an ion can be charged $+1$ or -3 or whatever, but never $-\frac{1}{7}$ or $+\pi$ or anything else that'sn't an integer.[15] Also, electric charge can't be created or destroyed, so these nudgings must be adding or removing some kind of subcomponent that carries the charge with it. An ion is also pretty much equal in mass to an atom, so whatever it is we are adding and removing must be very, very light.

With heavier nudges (whatever it is nuclear physicists do all day), both the charge and the mass of the atom can be changed in big steps, so it seems we are dislodging some heavy, charged subcomponent. It must be oppositely charged to whatever that light thing is, which allows the overall atom to be neutral. Those big mass steps can also be taken without changing the charge, so there seems to be some heavy, neutral subcomponent as well. And what do you know: atoms are made from light, negatively charged *electrons* (weighing 0.511 MeV), heavy, positively charged *protons* (938 MeV) and neutral *neutrons* (940 MeV)![16] The mass of an atom never precisely matches the sum of its parts, but that's not that strange—an awful lot of energy goes into binding those components together, and as we know, energy and mass are one and the same.

All of this we could have figured out without ever disassembling an atom to look at the parts in isolation (although of course we have). For the particles we are about to study, however, we have no choice but to go about things in that non-disassembly way. I will actually start with the protons and neutrons, the tamest beasts in the zoo.

---

[15]Remember this whenever quantum physics becomes too scary: at the bottom of it all, the word "quantized" doesn't mean anything more than "only available steps", and a *quantum* (plural: *quanta*) is just the size of one of those steps. I once heard it said (in hard-to-translate Swedish) that "*klassisk fysik köps i lösvikt, men för kvantfysik är det styckpriser som gäller!*"

[16]Following standard SI prefixes, an MeV is 1 000 000 eV or 1/1000 of a GeV.

Quantized spin

Spinning things carry *angular momentum*, and like so much else, this gets quantized down in the quantum world: the orbit of an electron around a nucleus can have angular momentum 1 or 2 or even 0, but never anything inbetween.[i]

Elementary particles have a kind of built-in angular momentum called *spin*, which holds a fixed value for each kind of particle. You should not think that the particles are actually spinning—they are so tiny that they would have to rotate faster than the speed of light to achieve enough angular momentum—but rather that they have something mysterious that, for all intents and purposes, *acts* like angular momentum. Remember the duck principle!

Things get weirder when considering the fact that angular momentum (or spin) has a direction, dictated by where the axis of rotation is pointing.[ii] It is quantized in each direction separately, but not independently: if you set things up so that the entire spin J points in the x-direction and then measure in the z-direction, you won't find 0 as you'd expect—you'd randomly find any value between $+J$ and $-J$, in steps of 1. For the lowest couple of total spins J, the possible directions look something like

$$(2.2)$$
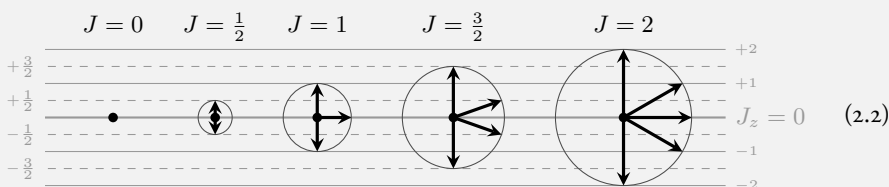
where the z-direction is vertical, and the horizontal represents some unknowable combination of x and y. Note that this leaves room for half-integer J-values like $\frac{1}{2}$ or $\frac{3}{2}$, and indeed we find that many common particles—protons, neutrons, and electrons—have spin $\frac{1}{2}$.

Spin gets added up when smaller particles form larger ones, and different combinations can be formed. For instance, two spin-$\frac{1}{2}$ particles make a spin-1 particle if their spins align, or a spin-0 particle if they don't—measured along one particle's spin axis, the other one will contribute either $+\frac{1}{2}$ (aligned) or $-\frac{1}{2}$ (not aligned).

---

[i]We don't notice this in the everyday world, because in everyday units those quantized values are separated by steps of $\hbar$, which is absurdly tiny. Even the minuscule and extremely slowly revolving month hand found on some wristwatches has an angular momentum well in excess of a trillion $\hbar$.

[ii]Specifically, the axis of *counterclockwise* rotation—if it goes clockwise, point the axis the other way.

Inventing isospin. Electric charge aside, protons and neutrons (collectively called *nucleons*, since they form atomic nuclei) are remarkably similar—they have the same spin and almost exactly the same mass. They also readily transform into one another, which is the mechanism behind radioactive beta decay. Normally, neutrons turn into protons, since this allows the mass difference to be liberated as energy, but the context of an atomic nucleus can tip this energy balance so that protons turn into neutrons.[17,18] Then it makes

---

[17]When protons and neutrons turn into each other, they also emit positrons or electrons (seen as $\beta^{\pm}$ particles) and neutrinos or antineutrinos (not seen at all unless you try *very* hard). These extra particles ensure that total charge is conserved, and that the liberated energy has somewhere to go.

[18]In one nucleus, Thorium-229m, the energy balance is so delicate—a mere 8.3 eV difference [10]—that even the chemical context can tip the scale. (Normally, chemical forces are far too feeble to affect nuclei.)

sense to guess that maybe the proton and neutron are mostly the same inside, differing only by altering one subcomponent. That alteration should account for their difference in charge, but clearly, charge can't be the important thing here—whatever force causes nucleons and nuclei to form must be vastly stronger than the electromagnetic force, or else all those equally-charged protons would fly apart.[19] Instead, there must be some new, deeper property. But what?

While we are trying to be clever, the duck principle waddles up and quacks at us: what if it is like spin? What if the proton and neutron are just the $+\frac{1}{2}$ and $-\frac{1}{2}$ states of a common particle, the nucleon, which has a total "spin" of $\frac{1}{2}$? It can't be spinning in regular space, not even in the not-really-spinning way of spin—spins flip if you flip space, but protons don't turn into neutrons just because you stand on your head[20]—but if things *behave* as if they are spinning in a more-or-less made-up space, then surely we can describe it like spin? The main forces involved would then be insensitive to the "spin" direction, but less important ones could give rise to differences in mass and charge.[21] Thus quacks the duck principle, and this idea turns out to be a solid one. To distinguish it from spin, we call this *(strong) isospin*, I;[22] the component that distinguishes protons and neutrons is labelled I_3 (similar to the z component of ordinary spin).

The baryons. The isospin-$\frac{1}{2}$ nucleon with components (p^+, n^0) has some close relatives in the zoo, and by analogy, they can also be organized by isospin. There is the isospin-1 Σ with components $(\Sigma^-, \Sigma^0, \Sigma^+)$, the isospin-$\frac{1}{2}$ Ξ with components (Ξ^-, Ξ^0), and the lonely isospin-0 Λ^0. But all of these particles, besides being a bit heavier than the nucleon (see table 1), are also somehow… strange. Given their large mass, they should happily decay into nucleons, and indeed they do, but much slower than they should.[23] It is as if they possessed some quantity not present in nucleons, one that Nature can create and destroy, but only reluctantly. This puzzling quantity is—I kid you not—called *strangeness*, S, and Λ and Σ have $S = -1$.[24] Ξ is *doubly* strange, $S = -2$, and it takes two strangely slow decays

---

[19]That these forces (the nucleon-forming and the nucleus-forming) are one and the same is, perhaps, not obvious, but sure as eggs are eggs, they are none other than the unimaginatively named *strong force*. Its purer, stronger manifestation forms nucleons, and a weaker, indirect manifestation binds them together into nuclei. This is similar to how purer, stronger electromagnetism forms atoms and assembles them into molecules, while weaker, indirect forms like the van der Waals force binds them together into liquids and solids.

[20]This might sound silly, but one of the most profound discoveries ever, the *Wu experiment* that proved that Nature fundamentally distinguishes left and right [11], was in essence performed by doing a very precise measurement, and then doing it again after flipping suitable parts of the setup.

[21]In everyday life, a similar thing happens when you wear polaroid sunglasses: the millions of kilometers of vacuum and dozens of kilometers of air between you and the Sun treat all photons the same regardless of where their spin points (for the record, photons have spin 1), but the thin sheet of plastic in front of your eyes blocks some spin directions and lets other ones through.

[22]There is also the physically unrelated *weak isospin*, which goes to show that the idea is good enough to apply twice. Weak isospin is incredibly important to the SM, but not to this thesis, so I won't mention it again.

[23]A particle decay needs two things to happen: energy to liberate, and a force to do the job. The stronger the force, and the more energy is liberated, the faster the decay happens. But forces follow rules, and the strong force, while strong, is forbidden from changing isospin. The electromagnetic force is similar but weaker—it can *sense* differences in isospin, which is how protons and neutrons differ in mass and charge, but it can't *change* isospin. The weak force is even weaker, but has fewer restrictions, so it allows protons and neutrons to decay by changing their isospin. This is a slow process—left alone, a neutron lasts about 20 minutes on average.

[24]Never mind the minus—it's a historical artifact from before we had strangeness fully figured out, much like the minus in the charge of the electron. As you may guess, the weak force is the only one that can change strangeness, but it's even worse at it than it is at changing isospin.

(a) Spin-$\frac{1}{2}$ baryons **(b)** Spin-$\frac{3}{2}$ baryons
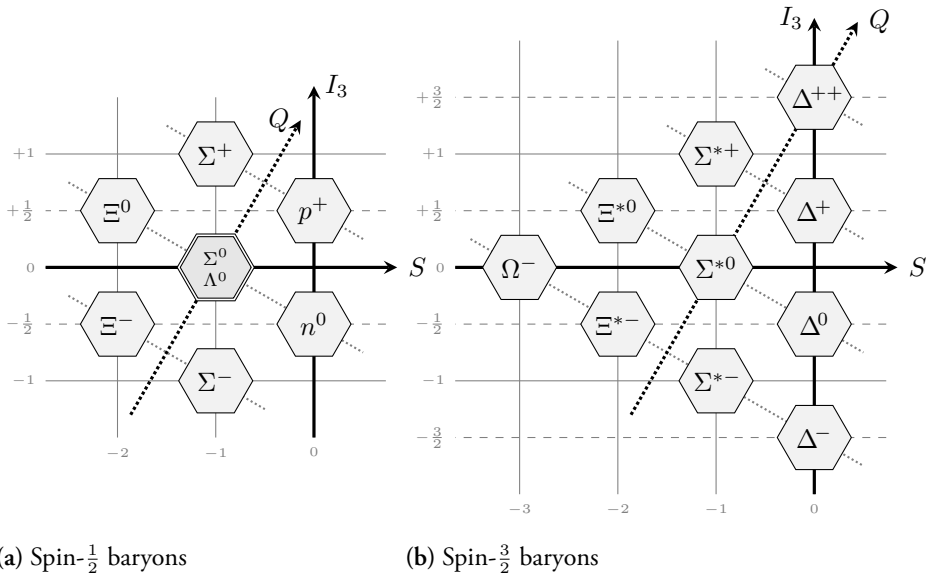
Figure 1: The baryons, arranged by isospin I_3, strangeness S, and electric charge Q.

decays, via an intermediate $S = -1$ particle, before it becomes an $S = 0$ nucleon.

These close relatives all have spin $\frac{1}{2}$ (the regular spin, not the iso kind), but there is also a family of spin-$\frac{3}{2}$ cousins with a similar layout: the Ξ and Σ are still there, but with stars to remind us that these are the spinnier versions; there is no Λ^0, but instead a *triply* strange Ω^-, and the nucleon is replaced by an isospin-$\frac{3}{2}$ Δ with components $(\Delta^-, \Delta^0, \Delta^+, \Delta^{++})$.[25]

Arranging all these *baryons* (as they are collectively called) by isospin and strangeness, the beautiful hexagonal pattern of fig. 1a emerges, and the delightful triangle of fig. 1b. Not only is it pleasingly symmetric, but also electric charge, Q, appears in an orderly fashion, increasing diagonally through the array.

The quarks. So, can we figure out the subcomponents of the baryons based solely on this information? Yes, we can! Reasonably, one kind of subcomponent takes care of the strangeness: Ξ has two of those, Σ/Λ has one, and the nucleon lacks it. Judging from the similarity between the Ξ and the nucleon, the strange subcomponent doesn't affect the isospin at all. Catching a whimsical mood, this subcomponent (purely hypothetical at the time) was given the name *strange quark*, s.

Next comes isospin, which is reasonably also carried by quarks, an isospin-$\frac{1}{2}$ pair (u, d) whose $I_3 = +\frac{1}{2}$ part is called *up* and $I_3 = -\frac{1}{2}$ part *down*, following how the I_3 axis points in fig. 1. That solves Ξ: if, in addition to the two s quarks, we give Ξ^0 an u and Ξ^- a d, both S and I_3 look correct. The same u/d difference could be what distinguishes proton

---

[25]The spin-$\frac{3}{2}$ particles decay extremely quickly, since the strong or electromagnetic forces are perfectly capable of jettisoning the extra spin, charge and I_3. Only Ω^- is relatively long-lived, since it takes the weak force to reduce the strangeness.

from neutron, but we can't simply identify $p^+ = u$ and $n^0 = d$, since the Ξ's are vastly different from what you'd get from a nucleon with two s quarks on top; no, there must be something more. Reluctant to introduce more kinds of particles, an elegant solution is to add one u and one d to each nucleon, giving them more substance while keeping the isospin where we want it.

And there we have it: if each baryon consists of three quarks drawn from u, d and s, everything falls into place. With one s, the remaining choices of uu, ud and dd perfectly explain the isospins of the Σ's. The same exercise can be repeated for the spin-$\frac{3}{2}$ family, and with the addition of $\Delta^- = ddd$, $\Delta^{++} = uuu$ and $\Omega^- = sss$, everything checks out. The electric charges also add up if we suppose that u has $Q = +\frac{2}{3}$ while d and s have $Q = -\frac{1}{3}$, and the spins add up if all quarks have spin $\frac{1}{2}$: they can be combined either as $\frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2}$ or $\frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2}$. The whole layout is summarized in table 1.

Some questions are reasonable to ask. Why is there no Λ^{*0}, and no spin-$\frac{1}{2}$ versions of Δ^-, Δ^{++} and Ω^-? The answer comes down to *symmetry*: quarks and their spins can be arranged in ways that are either symmetric or antisymmetric, and the subtle mathematics of spin-combining dictate different symmetries for spin-$\frac{1}{2}$ and spin-$\frac{3}{2}$ combinations. Nature likes symmetry and firmly requires that any antisymmetry is cancelled by another antisymmetry. Λ^0 is an antisymmetric arrangement of quarks, and there is an antisymmetric spin-$\frac{1}{2}$ combination to counter it, but no spin-$\frac{3}{2}$ combination. Likewise, arrangements of three identical quarks are too symmetric to be matched by a spin-$\frac{1}{2}$ combination.[26]

SYMMETRY AND ANTISYMMETRY

The word "symmetry" carries a broad and important meaning in particle physics. Given any transformation—mirroring, rotation, rearrangement, or any of the more abstract ones we will encounter later on—we say that an object is *symmetric under that transformation* if it's not changed by it. A symmetric thing (in the everyday sense) is symmetric under mirroring, and a sphere is symmetric under both mirroring and rotation.

A thing is *antisymmetric* under a transformation if it changes sign, but is otherwise unaffected (this only makes sense for things that have mathematical signs, of course). For example, $A + B$ is symmetric under the swap $A \leftrightarrow B$, but $A - B$ is antisymmetric.

Anything that can be combined with addition can be *symmetrized*, i.e., summed over some transformation, which creates something symmetric. Inserting suitable minus signs results in something antisymmetric instead. For instance, Σ^0 and Λ^0 are the symmetrization and antisymmetrization of uds, respectively, over rearrangement of the quarks:[i]

$$\Sigma^0 = (uds + sdu) + (dsu + usd) + (sud + dus),$$
$$\Lambda^0 = (uds - sdu) + (dsu - usd) + (sud - dus).$$
(2.3)

[i]Never mind what uds actually means—essentially, it's shorthand for a three-quark quantum state.

Having established that there are quarks, there are of course also *antiquarks*, which form *antibaryons* that look just like the baryons except that they are equal and opposite in charge,

---

[26]The details of this explanation, which are fiddly but not too difficult, are given in Griffiths [12, sec. 5.9].

Table 1: Masses and quark contents of the baryons shown in fig. 1 (values taken from the 2022 PDG review [3, pp. 90-99]). Numbers in parentheses indicate the uncertainty in the last digits, so 1314.86(20) means that Ξ^0 probably weighs between 1314.66 MeV and 1315.06 MeV. The Δ mass isn't known separately for each charge/isospin, but only as an average.

Content	Spin $\frac{1}{2}$	Mass [MeV]	Spin $\frac{3}{2}$	Mass [MeV]
uuu			Δ^{++}	$1\,232(2)$
uud	p^+	$938.272\,088\,16(29)$	Δ^+	$1\,232(2)$
udd	n^0	$939.565\,420\,5(5)$	Δ^0	$1\,232(2)$
ddd			Δ^-	$1\,232(2)$
uus	Σ^+	$1\,189.37(7)$	Σ^{*+}	$1\,382.83(34)$
$uds\ (I=1)$	Σ^0	$1\,192.642(24)$	Σ^{*0}	$1\,383.7(1\,0)$
$uds\ (I=0)$	Λ^0	$1\,115.683(6)$		
dds	Σ^-	$1\,197.449(30)$	Σ^{*-}	$1\,387.2(5)$
uss	Ξ^0	$1\,314.86(20)$	Ξ^{*0}	$1\,531.80(32)$
dss	Ξ^-	$1\,321.71(7)$	Ξ^{*-}	$1\,535.0(6)$
sss			Ω^-	$1\,672.45(29)$

strangeness and isospin. For instance, anti-Λ^0 or $\bar{\Lambda}^0$ (a bar over a particle symbol indicates its antiparticle) has $S = 1$ and contains an antistrange quark, $\bar{s}$.

The mesons. More interestingly, combinations of a quark and an antiquark give rise to a new family of particles, the *mesons*, which exist in spin-0 and spin-1 versions (based on the spin alignment of the quark and antiquark) as shown in fig. 2. Despite the completely different recipe, the mesons arrange themselves very similarly to the spin-$\frac{1}{2}$ baryons! Sure,

(a) Spin-0 (pseudoscalar) mesons   (b) Spin-1 (vector) mesons
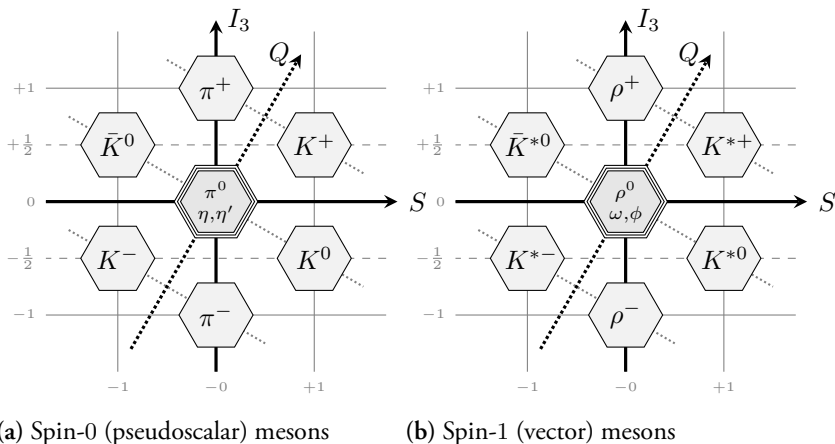
Figure 2: The mesons, arranged by isospin I_3, strangeness S, and electric charge Q. Note that this covers both mesons and anti-mesons: anti-π^+ is π^-, etc.

the axes are shifted, and there are now *three* different particles at the center, but this all makes sense based on the quarks—after all, we have both $u\bar{u}$, $d\bar{d}$ and $s\bar{s}$ to play with. Mesons are the main characters in this thesis—the spin-0 or *pseudoscalar* mesons, to be precise.[27] In fact, almost all my physically applicable work concerns the isospin-1 triplet of *pions*, (π^-, π^0, π^+), with some limited attention given to the two isospin-$\frac{1}{2}$ pairs of *kaons*, (K^0, K^+) and $(K^-, \bar{K}^0)$, and the isospin-0 *eta*, η.[28]

Vectors and their friends

Staying true to the duck principle, anything that behaves remotely like an arrow on a blackboard is considered a *vector*. Specifically, they must add and subtract in a sensible way, and accept rescaling by numbers called *scalars*. Vectors of the same kind collectively form a *(vector) space*, and that space is subject to *transformations*, which are called *linear* if they respect the way vectors add and scale. Vectors may be written in many ways: $\boldsymbol{v}$, $|v\rangle$, etc.

Every space can be given a *basis*,[i] a set of vectors $\boldsymbol{b}_i$ such that any vector can be expressed (uniquely) as a *linear combination* thereof,

$$\boldsymbol{v} = \sum_i v_i \boldsymbol{b}_i ,\tag{2.4}$$

where v_i are called the *coordinates* of $\boldsymbol{v}$. (Using *Einstein's summation convention*, a repeated index is understood to be summed without $\sum$, so $\boldsymbol{v} = v_i \boldsymbol{b}_i$.) Listing the components allows any abstract vector to look just as sensible as (x, y, z), although the length of the list—the *dimension* of the space—might be infinite.

Every space worth having has an *inner product*, a way to take two vectors and get a scalar.[ii] These are often expressed as combining a vector and a *conjugate vector*, like in $\boldsymbol{v}^\mathsf{T}\boldsymbol{u}$ or $\boldsymbol{v}^\dagger\boldsymbol{u}$ or $\langle v|u\rangle$ (vectors on the right, conjugates on the left). The components of the conjugate are usually obtained through complex conjugation: $\boldsymbol{u}^\dagger\boldsymbol{v} = u_i^* v_i$.

If you switch to a new basis, all coordinates must change accordingly so that the vector remains the same; this change is a linear transformation, written something like $v_i \to R_{ij}v_j$. (In this indexed form, R is a *matrix* and is conveniently written as a square array.) This transformation property also applies to the spin components of spin-1 particles, which is why they—duck goes "quack"—are called "vectors". Scalars don't change with the basis, so if $\boldsymbol{v} \to R\boldsymbol{v}$ then $\boldsymbol{v}^\dagger \to \boldsymbol{v}^\dagger R^\dagger$ with $R^\dagger R = 1$, so that scalar products are unchanged. Spin-0 particles also remain unchanged, so they—duck quacks again—are called "scalars".

Of special note is the *parity transformation*, multiplying the basis by -1 to mirror the space. This should also multiply vectors by -1, but leave scalars unchanged. Something that looks like a scalar but also gets a -1 is called a *pseudoscalar*, and something that looks like a vector but *doesn't* get its -1 is called a *pseudovector*, or more commonly an *axial vector*.

---

[i]Strictly speaking, this assertion relies on (and implies) the axiom of choice [13], for those of you who enjoy such mathematical nitpickings.

[ii]A product that results in another vector isn't necessarily a feature of vector spaces, though. The cross product, for example, is a bonus feature specific to three-dimensional space.

---

[27]The reason why they are pseudoscalars instead of scalars has to do with the finer details of how spins are combined. There are also scalar mesons, but they are heavier and highly unstable.

[28]If you wonder why I included baryons in the first place, it was because they do a better job of explaining quarks, and because protons and neutrons are closer to home for most people. Likewise, I include vector mesons for completeness and context.

Table 2: Masses and quark contents of the mesons shown in fig. 2 (values taken from the 2022 PDG review [3, pp. 33-42]). As in table 1, numbers in parentheses indicate the uncertainty in the last digit(s). The contents of π^0, η, η' and their spin-1 counterparts are only approximate due to mixing (see section 3.5).

Content	Spin 0	Mass [MeV]	Spin 1	Mass [MeV]
$u\bar{u} - d\bar{d}$	π^0	134.976 8(5)	ρ^0	775.26(23)
$u\bar{d}/d\bar{u}$	$\pi^\pm$	139.570 39(18)	$\rho^\pm$	775.11(34)
$u\bar{s}/s\bar{u}$	$K^\pm$	493.677(16)	$K^{*\pm}$	891.67(26)
$d\bar{s}/s\bar{d}$	$K^0/\bar{K}^0$	497.611(13)	$K^{*0}/\bar{K}^{*0}$	895.55(20)
$u\bar{u} + d\bar{d} - 2s\bar{s}$	η	547.862(17)	ω	782.66(13)
$u\bar{u} + d\bar{d} + s\bar{s}$	η'	957.78(6)	ϕ	1 019.461(16)

Now, have yourself a good, long stare at tables 1 and 2, especially the masses. There is no way to assign masses to the quarks that even remotely adds up to the masses we see, but this is expected, since there is probably an overwhelmingly large amount of energy bound up in such solid particles that the masses of the constituents hardly matter. Adding s quarks seems to add some mass, so maybe s is a tad heavier than the others, but overall, there just seems to be a general mass difference between baryons and mesons, a general mass increase at higher spin, and some smaller variations between individual particles.

But then focus on the pions. *They are so light*! Barely a sixth of the nucleon mass, yet they are supposed to have two thirds the quark content. They are also much lighter than the kaons, but we don't see a similar difference between the ρ and K^*. There must be something special with the pions, something that relaxes the forces and allows them to get away with less bound-up energy. Come to think of it, the kaons and η also seem to have some of that mysterios lightness to them: they weigh just over half that of η', but K^*, ω and ϕ differ much less dramatically.

The reason for this lightness is a deep, subtle and beautiful one, and one that forms the very foundation of everything this thesis is built on. The explanation, which I state in section 3, spontaneously gives birth to a very powerful toolbox for computing the behavior of these lightest animals in the zoo, and they in turn have a multitude of physical consequences.

With that, I conclude this warm-up section of baryons and mesons, collectively called *hadrons* (generally, a hadron is anything that can be built from quarks and antiquarks). It is only the tip of the iceberg—leaf through the PDG review [3, pp. 33-107] hadron lists, and you will see that there are many more than the ones I've shown, with quarks excited to higher energy levels and orbiting to add to the spin. There are also three more kinds of quarks, but they are heavy enough that even the lightest hadrons they form are too heavy to care about. Then there is the elephant in the room—how on Earth the quarks form hadrons, and why they can't be disassembled—but that isn't healthy to think about now (I'll get back to it on page 32). Instead, we should press on to gain a much better understanding of what quarks are and what they do, and pick up a whole load of mathematics on the way. Trust me, *that's* healthy.

Relativity and Lorentz indices

Things that are vectors in spacetime are written using Greek *Lorentz indices* running from 0 to 3, like the spacetime coordinate vector or the partial derivative:[i]

$$x^\mu \equiv (x^0, x^1, x^2, x^3) \equiv (t, x, y, z) \qquad \equiv (t, \boldsymbol{x}), \tag{2.5}$$

$$\partial_\mu \equiv (\partial_0, \partial_1, \partial_2, \partial_3) \equiv (\partial_t, \partial_x, \partial_y, \partial_z) \equiv (\partial_t, \boldsymbol{\nabla}). \tag{2.6}$$

Note that x^μ can mean both "the entire x-vector" or "one of its components, labelled by μ". The space part of a vector is written as a boldface version of the full spacetime vector, with the sole exception of $\boldsymbol{\nabla}$ which gets a special symbol.

Unlike other kinds of indices, which can be written either up or down based on convenience, Lorentz indices strictly distinguish up and down, and Einstein summation requires one index to be up and one down, which is called a *contraction*. For example,

$$\partial_\mu x^\mu \equiv \partial_t t + \boldsymbol{\nabla} \cdot \boldsymbol{x} \equiv \partial_t t + \partial_x x + \partial_y y + \partial_z z = 4. \tag{2.7}$$

Upper and lower indices can be exchanged using the *metric*, g: $x_\mu = g_{\mu\nu} x^\nu$ and $x^\mu = g^{\mu\nu} x_\nu$. In my sign convention (which may differ from yours), $g^{\mu\nu} = g_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$, where diag indicates a matrix with the listed elements on the diagonal and zeroes elsewhere.

Going to a new reference frame changes upper-index and lower-index things oppositely: $x^\mu \to \Lambda^\mu_\nu x^\nu$ and $\partial_\mu \to (\Lambda^{-1})^\nu_\mu \partial_\nu$ with $\Lambda^{-1}\Lambda = 1$. This also works for things with multiple indices (called *tensors*), which take one Λ for each upper index, and one Λ^{-1} for each lower one. An expression where all Lorentz indices have been contracted is the same in *all* reference frames; in other words, it's *Lorentz-invariant*. That's important: Nature herself is Lorentz-invariant, so whatever physics we try to do better have all its indices in order.

---

[i]Take care not to confuse the indices with exponents. Luckily, I will almost never write out components such as x^2 explicitly (instead, x^2 is short for $x^\mu x_\mu$) and will almost never use something that could be a Lorentz index (such as μ) as an exponent.

All that quantum stuff

Possibly to your surprise, dear reader, there is very little actual quantum mechanics in this thesis. Suffice it to say, then, that quantum states, written $|\phi\rangle$ for some descriptive label ϕ, are just vectors in an infinite-dimensional space, and that they are acted on by linear operators, such as those seen in eq. (2.9). If for some operator $\hat{O}$ we have $\hat{O}|\phi\rangle = \lambda|\phi\rangle$, then $|\phi\rangle$ is an *eigenstate* of $\hat{O}$ with *eigenvalue* λ, and can be said to possess a well-defined quantity of whatever λ represents; if not, then $|\phi\rangle$ is a *superposition* of those states and doesn't have a well-defined λ of its own.

In QFT, as the name suggests, much of the role of states is taken over by *fields*, which are operators that are also functions of spacetime. As operators, fields add or remove one particle from the state it acts upon.[i] Any operators that would act on a one-particle state can therefore equally well act on the field that would create it, which is why eq. (2.9) is written in terms of a field ϕ instead of a state.

---

[i]For details, see the first few chapters of Peskin & Schroeder [14], which treat the operator nature of quantum fields in great detail. Zee [15], on the other hand, manages perfectly fine without discussing states and operators at all.

2.2 Re-inventing Quantum Chromodynamics

Free quark searches: All searches since 1977 have had negative results.

<div align="right">

— *The Particle Data Group [3, p. 32],*
for once needing only one sentence
to say everything there is to say.

</div>

Why QCD, you may ask? After all, my work isn't in QCD, but in ChPT! Well, look up any introduction to ChPT—the lecture notes of Pich [16] and Scherer & Schindler [17], or the book by Donoghue [18, chapter VII], for instance—and they open cold with QCD and its properties laid out, deriving the rest from there. Therefore, QCD is where we must get before we're ready for ChPT, and incidentally, deriving QCD from scratch happens to make for a wonderful crash course in many useful things.

Quantum field theory from nothing. Imagine, then, that almost all of particle physics has been lost in some kind of cataclysm, and that all that remains is a well-catalogued set of experimental observations, and maybe a few charred pages of my favorite textbooks. From the ashes, we are now tasked with creating something to describe the innards of a hadron.

We can start with a few things we remember from *outside* particle physics. Any child knows that $E = mc^2$, and the cool kids who use natural units write it as $E = m$. The even cooler kids know that $E = m$ is a special case for stationary things, and that the general form includes the momentum $\boldsymbol{p}$ like[29]

$$m^2 = E^2 - \boldsymbol{p}^2 = p^\mu p_\mu \,. \tag{2.8}$$

Some of the cool natural-units kids may know that if ϕ is some quantum thing (see the box for what it can be), the operators for extracting its energy and momentum are

$$-i\partial_t \phi = E\phi \quad \text{and} \quad -i\boldsymbol{\nabla}\phi = \boldsymbol{p}\phi \,, \tag{2.9}$$

respectively. Putting eqs. (2.8) and (2.9) together we can conclude that

$$m^2\phi = -(E^2 - \boldsymbol{p}^2) = -(\partial_t^2 - \boldsymbol{\nabla}^2)\phi = \partial_\mu \partial^\mu \phi \tag{2.10}$$

or, more elegantly,

$$(\partial_\mu \partial^\mu + m^2)\phi = 0 \,. \tag{2.11}$$

Just like that, we have found the wonderful *Klein–Gordon equation*. As we will see, it's a cornerstone of QFT: loosely speaking, satisfying the Klein–Gordon equation is a basic requirement for any aspiring quantum field ϕ. Put another way, if $(\partial_\mu \partial^\mu + m^2)\phi$ equals anything other than 0, there better be a *very* good motivation for it.

We just casually unified quantum physics with relativity, and the Klein–Gordon equation simply fell out of that. We can complete this trinity with the holiest of concepts from classical mechanics, the *principle of least action*. The *action*, S, is composed of all physical

---

[29]This also handles cases where "momentum" is something less straightforward than "mass times velocity", which is good for things like photons, which have momentum but no mass.

things in such a way that, out of all the ways those things *could* behave, the way they *actually* behave is the way that makes S as small as possible—the way with the least action. (Of course, in quantum mechanics one can never be sure about precise behaviors, but the *most probable* behavior should be the least-action one.) If one lets S be a spacetime integral of some other thing, $S = \int \mathrm{d}^4 x\, \mathcal{L}$, then some very pretty math (left to the reader's imagination) reformulates this least-action requirement into the *Euler–Lagrange equation*,[30]

$$0 = \sum_i \left[\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right], \tag{2.12}$$

where ϕ_i serve as a stand-ins for those physical things; $\mathcal{L}$, which is called the *Lagrangian*,[31] is built out of all of these.

If you want classical physics, it's not too hard to find the Lagrangian such that the Euler–Lagrange equation is equivalent to Newton's laws of motion. But for our purposes, some playing around reveals that if the Lagrangian is

$$\mathcal{L} = \tfrac{1}{2} \left[\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2 \right], \tag{2.13}$$

then the Euler–Lagrange equation becomes the Klein–Gordon equation! This is the basic Lagrangian of a quantum field theory, with ϕ being a basic quantum field.

Fancier fields. Dissatisfied with "basic"—QCD is, after all, anything but—we start playing around. Our ϕ is a quantum operator and a function of spacetime, but that's it, and there is relatively little we can do without breaking the Klein–Gordon equation. Sure, nothing stops us from inventing some space and making ϕ a vector in that, promoting the Lagrangian to

$$\mathcal{L} = \tfrac{1}{2} \left[\partial_\mu \boldsymbol{\phi}^\mathsf{T} \partial^\mu \boldsymbol{\phi} - m^2 \boldsymbol{\phi}^\mathsf{T} \boldsymbol{\phi} \right], \tag{2.14}$$

but all that does is create an array of identical Klein-Gordon equations. Making ϕ complex, with $\phi^\dagger$ independent of ϕ, doesn't do much either, besides introducing the concept of an antiparticle. (Okay, maybe that isn't such a small deal after all. Let's move on, though.)

For more excitement, we could make ϕ a vector (or even a tensor!) in spacetime, equipping it with one or more Lorentz indices and carefully contracting them. What happens then? Well, here comes something profound: a field with one Lorentz index makes particles with spin 1, since the vector ϕ^μ and the spin axis transforms in precisely the same way (the duck principle quacks happily). This generalizes elegantly: a field with n Lorentz indices gives spin-n particles.

The thought of spin draws our eyes to one of those charred textbook pages, which carries a fragmented record of the *spin-statistics theorem*.[32] To our dismay, that theorem states that any such integer-spin particles will be *bosons*, with the defining feature that one can pack

---

[30]Even though it looks off, the Lorentz indices are *not* placed wrong in the equation! The derivative with respect to an upper-index thing behaves like a lower-index thing, and vice versa: just think about $\partial / \partial x^\mu = \partial_\mu$.

[31]Strictly speaking, it's the Lagrangian *density*. The actual Lagrangian is the spatial integral $\int \mathrm{d}^3 \boldsymbol{x}\, \mathcal{L}$, but since that non-Lorentz-invariant object isn't used in particle physics, the "density" is dropped for convenience.

[32]Zee [15, chapter II.4] has a nice exposition on the spin-statistics theorem and how catastrophically physics would fail if it weren't true. To me, that's almost more convincing than an actual proof.

any number of them into the same place and quantum state. That's certainly not what we want—experimental data (and common sense) clearly shows that matter takes great offense at being packed too tightly. The same theorem states that fields with half-integer spin are *fermions*, which display a more matter-like refusal to be packed. But how on Earth does one get a field with *half* a Lorentz index? It simply doesn't make sense!

Fancier equations. Disappointed with what playing around with ϕ gave us, we turn back to the Klein–Gordon equation. Perhaps we can peel it back a layer, pull something simpler out of it—yes, what if we can find $\Delta_\pm$ such that $\Delta_+\Delta_- = \partial_\mu\partial^\mu + m^2$? Then if $\Delta_-\phi = 0$, the Klein–Gordon equation will automatically be satisfied, since clearly $\Delta_+ 0 = 0$, but perhaps $\Delta_-\phi = 0$ will impose some neat properties on ϕ. The solution that immediately jumps out is $\Delta_\pm = m \pm i\partial_\mu$, but that naked Lorentz index means that $\Delta_-\phi = 0$ can't be a physical, Lorentz-invariant equation.

To fix that Lorentz leak, we make up a vector: $\Delta_\pm = m \pm i\gamma^\mu\partial_\mu$. Our plan immediately springs a new leak: how do we avoid breaking $\Delta_+\Delta_- = \partial_\mu\partial^\mu + m^2$, now that we get a messy $\gamma^\mu\partial_\mu\gamma^\nu\partial_\nu$? A little fiddling reveals that the mess unravels if we say

$$\{\gamma^\mu, \gamma^\nu\} = -2g^{\mu\nu} \,, \tag{2.16}$$

but this means that the components of γ^μ can't simply be numbers, but instead something that multiplies in a fancier way—matrices immediately spring to mind.[33]

Frustratingly, this results in yet another leak: eq. (2.16) isn't compatible with the way γ^μ should transform based on its Lorentz index. Desperately, we make ϕ a vector in the same space that γ^μ is a matrix (because that makes sense) and guess that maybe spacetime transformations shuffle its components around a bit? Yes! There is a unique way to do

---

[33]Arguably the simplest objects that have the required behaviour are the 4×4 matrices

$$\gamma^0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \gamma^i = \begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix},$$

where σ_i are the Pauli matrices [see eq. (3.50)] and "0" and "1" are actually 2×2 matrices. But following the duck principle, any objects that satisfy eq. (2.16) can play the role of γ^μ.

this—never mind the details[34]—in such a way that $i\gamma^\mu\partial_\mu\phi$ transforms just the same way as $m\phi$, paving the way for something consistent with relativity: an equation can transform whatever weird way it wants, as long as all parts of it transform the same way. This gives us the equation

$$(i\gamma^\mu\partial_\mu - m)\psi = 0\,, \tag{2.17}$$

which happens to be called the *Dirac equation*. I have renamed the field to ψ in its honor, leaving the name ϕ for fields wihtout these extra properties.

To find the Lagrangian for eq. (2.17), which must be invariant under all transformations, it makes sense to form inner products like $\psi^\dagger\psi$ and $\psi^\dagger\gamma^\mu\psi$. Doing this *almost* works, but there are some signs that don't come out right. A glance at eq. (2.16) gives the inspiration to introduce $\bar{\psi} \equiv \psi^\dagger\gamma^0$, which fixes the signs and enables the Lagrangian

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \tag{2.18}$$

to give eq. (2.17) as its Euler–Lagrange equation.

Even though this was all terribly *ad hoc*, in the end the pieces fell into place in such a neat way that maybe there is some deeper meaning to it. And yep, there is: the weird transformation properties imposed on ψ by the Dirac equation happen to be just the way that a spin-$\frac{1}{2}$ object transforms, so we have our fermions! Come to think of it, ψ comes uncannily close to having "half a Lorentz index", but in a way that makes sense. Anyhow, we are now closer to our goal: quarks must be described by some variation of eq. (2.18)!

Fancier transformations. Taking a moment to adore eq. (2.18), we suddenly realize that ψ has some freedom to it. If we redefine $\psi \to e^{i\alpha}\psi$ for some real number α, the Lagrangian (and, therefore, the physics) will be completely unchanged since the change is cancelled by $\bar{\psi} \to \bar{\psi}e^{-i\alpha}$. This is of course obvious—multiplying ψ by a constant does absolutely nothing to the Dirac equation. What if we make α a bit more interesting than a constant, then? What if we make it a function, $\alpha(x)$, of space and time? But no, that can't work, since the derivative will bite into it and produce

$$\mathcal{L} \to \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \bar{\psi}(\gamma^\mu\partial_\mu\alpha)\psi\,, \tag{2.19}$$

which isn't the same as eq. (2.18) unless $\partial_\mu\alpha = 0$. Even worse, such a transformation doesn't even make sense! Just look at the basic definition of a derivative,

$$\partial_\mu\psi = \lim_{\epsilon\to 0} \tfrac{1}{\epsilon}\left[\psi(x + \epsilon\hat{\mu}) - \psi(x)\right]\,, \tag{2.20}$$

where $\hat{\mu}$ is a unit vector (vector of length 1) in the x^μ direction. If α depends on x, then $\psi(x + \epsilon\hat{\mu})$ and $\psi(x)$ don't transform the same way, and the transformation of $\partial_\mu\psi$ is ill-defined. If we ever hope to introduce such a ridiculous transformation, we would need

---

[34]If you *do* mind the details, write the transformation of a Lorentz vector like $v^\mu \to \exp\left[\frac{1}{2}(\omega^\mu{}_\nu - \omega_\nu{}^\mu)\right]v^\nu$, where $\omega_{\mu\nu} = -\omega_{\nu\mu}$ is an antisymmetric tensor defined by this relation, and the exponential of a tensor is interpreted as $\exp(T^\mu_\nu) \equiv \delta^\mu_\nu + T^\mu_\nu + \frac{1}{2}T^\mu_\rho T^\rho_\nu + \ldots$. Then the desired transformation is $\phi \to \exp\left(\frac{1}{8}\omega_{\mu\nu}[\gamma^\mu, \gamma^\nu]\right)\phi$. Note that γ^μ itself doesn't transform at all, despite the index, but the way it interacts with $[\gamma^\mu, \gamma^\nu]$ ensures the correct overall behavior. All of this is explained in detail in Peskin & Schroeder [14, chapter 3.2]. See also the thesis of Andrew Lifson [19] for an elegant exposition, including the deeper group-theoretical motivation.

some kind of modified derivative,

$$D_\mu \psi \equiv \lim_{\epsilon \to 0} \tfrac{1}{\epsilon} \left[\psi(x + \epsilon \hat{\mu}) - U(x + \epsilon \hat{\mu}, x) \psi(x) \right], \tag{2.21}$$

where $U(x, y)$ is defined to smoothly reach across the gap between x and y and fix the transformation by itself transforming like

$$U(x, y) \to e^{i\alpha(x)} U(x, y) e^{-i\alpha(y)}. \tag{2.22}$$

We then get the nice and consistent $D_\mu \psi \to e^{i\alpha(x+\epsilon\hat{\mu})} D_\mu \psi$, and the world has a hope of making sense again. We can *Taylor expand* U around $\epsilon = 0$, giving

$$U(x + \epsilon\hat{\mu}, x) = U(x, x) - i\epsilon\hat{\mu} g G_\mu(x) + \mathcal{O}(\epsilon^2) \qquad \text{(no sum on } \mu), \tag{2.23}$$

where $g G_\mu(x)$ is defined by this, and I snuck in the constant g for later convenience. Combining this with eq. (2.22), we can tell that G_μ transforms like

$$G_\mu \to G_\mu - \tfrac{i}{g} \partial_\mu \alpha, \tag{2.24}$$

and, letting $\epsilon \to 0$ in eq. (2.21), that $D_\mu \psi = \partial_\mu \psi + i g G_\mu \psi$ if we make the reasonable definition $U(x, x) = 1$. Swapping our old derivative for the new, we get

$$\mathcal{L} = \bar{\psi} \left[i\gamma^\mu (\partial_\mu + g G_\mu) - m \right] \psi, \tag{2.25}$$

which *is* unchanged under $\psi \to e^{i\alpha(x)} \psi$ even when $\alpha(x)$ is a function of spacetime.

BIG-O NOTATION

The notation $\mathcal{O}(x)$ captures something's rough behavior while ignoring unimportant details. $\mathcal{O}(x)$ may stand for x or $500x$ or $x + \tfrac{1}{3}x^3$ or x^{100}—it doesn't matter, but what *does* matter is that $\mathcal{O}(x)$ goes to zero when $x \to 0$, but $\tfrac{1}{x}\mathcal{O}(x)$ has a chance of being finite, and $\tfrac{1}{x^2}\mathcal{O}(x)$ most likely explodes.

In other contexts, $\mathcal{O}(x)$ may instead capture rough behavior as $x \to \infty$ in a similar way. The only reliable convention is that something called ϵ is always small, not large.

The transformation $\psi \to e^{i\alpha(x)}\psi$, which by now has had quite a redemption arc, is called a *gauge transformation*,[35] and as we will see, it ends up being possibly the most important transformation in the history of transformations. The field G_μ just fell out of our search for consistency, but to really motivate its existence, we should promote it to a proper quantum field by having it obey the Klein–Gordon equation. A neat way to write the Lagrangian for such a spin-1 field is $\mathcal{L}_G \equiv G^{\mu\nu} G_{\mu\nu}$, where $G_{\mu\nu} \equiv \partial_\mu G_\nu - \partial_\nu G_\mu$ is invariant under eq. (2.24).[36] Our combined Lagrangian is then

$$\mathcal{L} = \bar{\psi} \left[i\gamma^\mu (\partial_\mu + g G_\mu) - m \right] \psi + G_{\mu\nu} G^{\mu\nu}. \tag{2.26}$$

---

[35]The word "gauge", specifically, refers to the fact that the symmetry transformation is allowed to vary across space and time, rather than being the same everywhere. For the history of this concept, and the reason why a word that normally means "measuring device" or "spacing of railroad tracks" came to be used this way, see ref. [20].

[36]That the Euler–Lagrange equation of $\mathcal{L}_G$ is the Klein–Gordon equation is easy to see when $\partial_\mu G^\mu = 0$, and no matter what $\partial_\mu G^\mu$ happens to be, making it zero just requires a gauge transformation, which (as we have already concluded) doesn't change $\mathcal{L}_G$.

If we let $g \to 0$, this reduces to ψ and G_μ being individual fields that satisfy the Klein–Gordon equation in their own way. When $g \neq 0$, we get something new and beautiful, where ψ and G_μ alter each other and deviate from the Klein–Gordon equation in a well motivated way—maybe this is QCD?

Well, if we sit down and carefully derive the physical consequences of eq. (2.26), it turns out we actually discovered QCD's cousin, QED: if G_μ is the photon, ψ the electron and g its electric charge, eq. (2.26) perfectly describes the electromagnetic force. (Any other particles with charge gQ can be introduced if they transform with $e^{iQ\alpha}$.)

MORE MATRIX PROPERTIES

Given a matrix $\mathbf{M}$ with elements M_{ij}, one can take the *transpose* $(\mathbf{M}^\mathsf{T})_{ij} \equiv M_{ij}$, *conjugate* $(\mathbf{M}^\dagger)_{ij} \equiv M_{ji}^*$ and *trace* $\langle \mathbf{M} \rangle \equiv M_{ii}$; since traces feature so prominently in this thesis, I typically use this compact angle-bracket notation for them. Note that traces are *cyclic*: $\langle \mathbf{AB} \cdots \mathbf{YZ} \rangle = \langle \mathbf{ZAB} \cdots \mathbf{Y} \rangle = \langle \mathbf{YZAB} \cdots \rangle = \ldots$, as is easily proven by looking at the indices. A matrix with zero trace is called *traceless*.

A matrix is *hermitian* if $\mathbf{M}^\dagger = \mathbf{M}$ and *unitary* if $\mathbf{M}^\dagger \mathbf{M} = \mathbf{M}\mathbf{M}^\dagger = 1$. Note that the complex exponential of a hermitian matrix is unitary: $e^{i\mathbf{M}}\left[e^{i\mathbf{M}}\right]^\dagger = e^{i\mathbf{M} - i\mathbf{M}^\dagger} = e^0 = 1$. [Matrix exponentials work like regular exponentials, but with a few complications if the matrices don't commute; see section 3.4.]

QCD, at last!   Okay, let's try one more thing: $\psi \to e^{i\alpha}\psi$ and $\bar{\psi} \to \bar{\psi}e^{-i\alpha}$ also works if α is promoted from a number to an $n \times n$ hermitian matrix, as long as ψ is promoted to a matching vector so that $\mathcal{L}$ remains a scalar. In fact, α has to be a *traceless* hermitian matrix, because otherwise we can write it as $\alpha = \alpha' + \frac{1}{n}\langle\alpha\rangle$ where α' is traceless and $\frac{1}{n}\langle\alpha\rangle$, which is again just a number, results in another unwanted copy of QED. We can still let α be a function of spacetime, but following eq. (2.24), we must then also promote G_μ to a traceless hermitian matrix. This, in turn, complicates the $G_{\mu\nu}G^{\mu\nu}$, which now fails to be gauge-invariant since the G_μ don't commute. This is resolved by redefining

$$G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu - ig[G_\mu, G_\nu]. \tag{2.27}$$

With these modifications, a Lagrangian like eq. (2.26) works also for these more general matrix transformations.

Now, we are close! Then comes the crucial question: which n gives QCD? All we can say from first principles is that it can't be 1, since the only traceless 1×1 matrix is 0. Well, a hint comes from section 2.1. Look at Δ^{++}, Δ^- or Ω^-, for example. As we have established, they consist of three identical quarks, all with their spins pointing the same way and with no orbital motion or anything. That is: three *fermions* are sitting in the *same* quantum state! Ah! Nature cries out in disgust!

The solution comes from the space we introduced, where α and G_μ are matrices and ψ is a vector, since it offers another degree of freedom to separate the otherwise identical states. If $n = 2$, there are only two independent directions, which isn't enough, so n better be at least 3. By Occam's razor (and, it turns out, quite a few experimentally measurable things), we can drop the "at least" and conclude that $n = 3$ for QCD. Inspired by the "C" in

"QCD" (from Greek χρῶμἄ, "color"), we call this space *color space* and name its three axes **red**, **green** and **blue**, which breeds many useful analogies (just peek at fig. 3 on page 33).

Now, we have finally found QCD. It should be clear by now that ψ represents the quarks, which come in several different *flavors* named up, down, strange, etc.; these must all look the same in color space, although they can have different mass.[37] In light of this, let me polish up the Lagrangian into

$$\mathcal{L}_{\text{QCD}} = \sum_f \bar{q}_f \left[i\gamma^\mu (\partial_\mu + gG_\mu) - m_f \right] q_f + G_{\mu\nu} G^{\mu\nu} , \tag{2.28}$$

where q_f is the quark of flavor f, and m_f is its mass. The field G_μ represents the *gluons*, which have no mass [proof: try to find a "m_G" term in eq. (2.28)] and serve to bind quarks together; since a traceless 3×3 matrix has 8 degrees of freedom (3 times 3, minus 1 for the trace), there are really eight gluons with different color combinations. *Which* color combinations, you ask? Well, the gauge transformation is a rotation in color space, so you can't really give a gauge-invariant answer. Similarly, you can never claim that a certain quark is **blue**—a gauge transformation can make it **red**—but one *can* say that all quarks in the same baryon have different colors.

At this point, the particle zoo makes a bit more sense, but not a lot of it. Why are there no hadrons that consist of, say, two quarks, or four, or two quarks and an antiquark? Why, still, are quarks impossible to take out of the hadrons that contain them? To make an answer that isn't utter hand-waving, I have to take quite a mathematically involved detour, but don't worry—the journey and the destination are their own, separate rewards. (If that's too much reward for you, hold your breath and skip to page 32.)

2.3 Getting physics out of QCD

> *Feynman diagrams make it possible for anybody to do quantum field theory!*
>
> — *attributed to a supposedly annoyed Julian Schwinger*

In the previous section, I may have uttered the weasely words "sit down and work out the physical consequences".[38] As hinted at in the popular science introduction, doing so in practice is no mean feat—in fact, almost every single scientific contribution I present in this thesis is some subtle aspect of "sitting down and working out the physical consequences"! Therefore, here is a crash course in just that.

The path to physics. Early in the previous section, I talked about the principle of least action, how the way things behave is whatever way makes the action, S, as small as possible. There exists a thing, one of the cleverest things any physicist has ever thought of, which

---

[37]It can't be stressed enough that this use of "color" and "flavor" has nothing to do with their everyday meanings. Particle physics parlance is just silly like that sometimes.

[38]A close relative to "it is easy to see that…", which I may also be guilty of using sometimes. I solemnly swear that I only use it when it's actually easy!

captures that principle as a mathematical object, one that can be manipulated, calculated and thrown about, and which is much better than the Euler–Lagrange equation at surviving the weirdness that QFT brings to the table. That glorious thing is the *Feynman path integral*, and it looks like this:

$$\langle 0|\hat{O}|0\rangle = \frac{1}{Z} \int \mathcal{D}\phi \; \hat{O} \; e^{iS[\phi]} \,. \tag{2.29}$$

Here, $\hat{O}$ is an *operator*: any collection of quantum fields and other stuff that you can possibly dream up. In quantum notation, $\langle 0|\hat{O}|0\rangle$ roughly means "$\hat{O}$ sets up some physical system (out of the vacuum, $|0\rangle$), lets physics take its course, and then takes down the system again (back to the vacuum, $\langle 0|$)". That's an extremely general way to describe any physical process, and $\langle 0|\hat{O}|0\rangle$ is a number describing the probability that the process will actually take place.[39]

Okay, so the left-hand side of eq. (2.29) is just physics. The right-hand side, then, is the clever bit! The notation $\int \mathcal{D}\phi$ tells us to integrate—that is, sum—over all possible ways the field ϕ can be; if there are multiple fields, we just integrate over all of them in turn. The factor $e^{iS[\phi]} = \cos(S[\phi]) + i\sin(S[\phi])$ tells how much each way contributes to the total probability. If $S[\phi]$ changes as the integral scans over similar ways, the sine and cosine will oscillate rapidly between ± 1, and those contributions will more or less cancel out. The only way that doesn't suffer any cancellation is the one where $S[\phi]$ remains perfectly still— and that, my friend, is the one with the least action: a valley is always perfectly flat at its bottom.[40] This is how the path integral makes physics obey the principle of least action.[41]

That leaves one piece of eq. (2.29): what is Z? Well, if we let $\hat{O} = 1$ and assume that the vacuum (like any good quantum state) is normalized, $\langle 0|0\rangle = 1$, then clearly[42]

$$Z = \int \mathcal{D}\phi \; e^{iS[\phi]} \,. \tag{2.30}$$

If something is weird with the normalization of the vaccum, Z will adjust to hide that. This is good—the QFT vacuum is not the kind of void you want to stare into for too long.

Particles going from here to there. Suppose we're after $\langle 0|\phi(x)|0\rangle$, the probability that we'll find ϕ if we look around in the vacuum at x (never mind exactly what that would mean, physically). For calculations like this, there is an extra bit of cleverness that will let Z handle everything for us, so that we never have to deal with the path integrals themselves.

---

[39]If you're wondering why we can't make that probability as large as we please—even larger than 1—by rescaling $\hat{O} \to \Lambda\hat{O}$ for some large number Λ, it's because a fundamental law of quantum mechanics—one that exists precisely so that the idea of probability makes sense—says that any operator that can actually be measured must be unitary: $\hat{O}^\dagger\hat{O} = 1$. This leaves no room for Λ shenanigans.

[40]Well, a mountain is also perfectly flat on its peak, so this would point us towards the way with the *most* action. In general, any way that produces *stationary* action—that is, $S[\phi]$ doesn't change when the way is adjusted slightly—will be favored in the integral. Sensible, physical actions don't permit this situation, though.

[41]In the everyday world, any action will be a very large number, so those oscillations will be *extremely* rapid and the cancellation nearly complete (this is backed by a formal mathematical result known as the *stationary phase approximation*). Therefore, only the least-action way contributes anything at all, and physics becomes a clear-cut chain of cause and effect dictated by the Euler–Lagrange equation. In the quantum world, $S[\phi]$ may be small, so the cancellation is less strict, and this perfectly captures all quantum weirdness there can be.

[42]Z is properly called the *generating functional*, but I think that name does more harm than good to our odds of understanding things, so let's just keep calling it Z.

FUNCTION SPACES AND FOURIER TRANSFORMS

In one of the greatest triumphs of the duck principle, functions (of some given variables) are vectors! They can be added, subtracted and rescaled, and a perfectly valid inner product between two functions $f(x)$ and $g(x)$ is the following integral (x can stand for a list of several variables):

$$f \cdot g \equiv \int \mathrm{d}x \, \left[f(x)\right]^* g(x) \,. \tag{2.31}$$

An important kind of linear transformation on this space of functions is *(linear) differential operators* such as ∂_μ, $[\partial^2 - m^2]$ or $[\gamma^\mu \partial_\mu - m]$. More generally, function transformations can be with an integral and a two-argument function,[i]

$$\Delta f(x) \equiv \int \mathrm{d}y \, \Delta(x, y) f(y) \,, \tag{2.32}$$

which looks very much like the more familiar $(\mathbf{M}\boldsymbol{v})_i = M_{ij} v_j$ if you imagine x and y as continuous versions of the indices i and j. In function space, the role of 1 is played by the *Dirac delta function*,[ii] defined by

$$f(x) = \int \mathrm{d}y \, \delta(x - y) f(y) \,. \tag{2.33}$$

One can take derivatives with respect to functions; these are called *functional derivatives* and denoted by δ to distinguish them. For instance, $\delta(f \cdot g)/\delta f(x) = g(x)$ and $\delta f(x)/\delta f(y) = \delta(x - y)$. The inverse of a differential operator Δ, i.e. the object $\Delta^{-1}(x, y)$ such that $\Delta\Delta^{-1}(x, y) = \delta(x - y)$, is known as a *Green's function*.

Function space is infinite-dimensional, but has a useful basis in the form of the functions $e^{ipx}/\sqrt{2\pi}$ for all p (which in general are vectors conjugate to x). This basis is orthonormal in the sense that $e^{ipx} \cdot e^{ikx} = 2\pi\delta(p - k)$. The coordinates in this basis are

$$\tilde{f}(p) \equiv \frac{1}{\sqrt{2\pi}} \int \mathrm{d}x \, e^{-ipx} f(x) \,, \tag{2.34}$$

which incidentally is another function, this time of p. This is known as the *Fourier transform* of $f(x)$ and is immensely useful; for instance, the Fourier transform of $\partial_x f(x)$ is simply $ip\tilde{f}(p)$, as follows from eq. (2.34) through integration by parts.[iii] The transform is also conveniently its own inverse:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}p \, e^{-ipx} \tilde{f}(p) \,. \tag{2.35}$$

When x describes position in spacetime, the Fourier transform gains a physical meaning, with p being the momentum. One often works in this momentum space instead, since many things (like differential operators) are simpler there.

---

[i]Strictly speaking, $\Delta(x, y)$ may be a more general kind of entity known as a *distribution*.

[ii]Case in point: it's actually the Dirac delta *distribution*, but physicists who care about that distinction are few, far between, and not very popular at physicist parties.

[iii]This requires that $f(x)$ is sufficiently well-behaved, although being physicists, we assume that it is. Readers with a penchant for mathematical rigor are kindly asked to check that all functions we deal with are Schwartz functions.

Let's pretend for a moment that the Lagrangian contains an extra term that involves just ϕ, times some function f, and write this made-up Lagrangian as $\mathcal{L}[\phi, f] = \mathcal{L} + \phi(x)f(x)$. Since $S = \int d^4 x \, \mathcal{L}$, this carries over to S and Z, so we write $S[\phi, f]$ and $Z[f]$. This is harmless: to get our actual $\mathcal{L}$, S and Z back, all we have to do is set $f = 0$ everywhere. But look:

$$\langle 0|\phi(x)|0\rangle = \frac{1}{Z}\int \mathcal{D}\phi \, \phi(x) \, e^{iS[\phi, f]} = \frac{1}{Z}\frac{-i\delta}{\delta f(x)}\int \mathcal{D}\phi \, e^{iS[\phi, f]} = \frac{-i\delta}{\delta f(x)}\log Z[f]\Big|_{f\to 0}. \tag{2.36}$$

There's no integral in the last step, and $f \to 0$ means we undo that made-up change in the end, so we haven't broken physics, either. All we needed was $\delta/\delta f(x)$, the *functional derivative*. Never mind exactly how it works—what's important is that on one hand, it can reach inside an integral and pull out a function, like

$$\frac{\delta S[\phi, f]}{\delta f(x)} = \phi(x), \tag{2.37}$$

and on the other other hand, it behaves just like an ordinary derivative, and we know that ordinary derivatives do stuff like $\frac{d}{dx}e^{ax} = ae^{ax}$ and $\frac{d}{dx}\log z = \frac{1}{z}\frac{dz}{dx}$. Just squint and pretend δ is d! That's how eq. (2.36) does its magic.

The neat thing is that now we can build any operator we want by doing the right made-up changes and taking the right functional derivatives. Let's look at

$$D(x, y) \equiv \langle 0|\phi(y)\phi(x)|0\rangle = \frac{-i\delta}{\delta f(y)}\frac{-i\delta}{\delta f(x)}\log Z[f]\Big|_{f\to 0}, \tag{2.38}$$

which *has* a clear physical interpretation: it's the probability that, if we set up a ϕ-particle at x, it will have moved (propagated) to y by the time we take it down. This earns $D(x, y)$ the name of *propagator*. Remember that word—I will say "propagator" a lot from now on.

Let's try calculating a propagator using a simple action whose Euler–Lagrange equation is the Klein–Gordon equation. You can check that

$$\mathcal{L} = -\tfrac{1}{2}\phi[\partial^2 + m^2]\phi + f\phi \equiv -\tfrac{1}{2}\phi\Delta\phi + f\phi \quad \text{where} \quad \Delta \equiv \partial^2 + m^2 \tag{2.39}$$

does the job just as well as eq. (2.13) when $f = 0$. Since we've established a solid duck-principle relation between functional and ordinary derivatives, we can go through the following algebra by squinting and pretending that ϕ, Δ and f are numbers instead of scarier mathematical objects. Let's start by completing the square:

$$\mathcal{L} = -\tfrac{1}{2}\big(\phi - \tfrac{f}{\Delta}\big)\Delta\big(\phi - \tfrac{f}{\Delta}\big) + \tfrac{f^2}{2\Delta} \equiv -\tfrac{1}{2}\phi'\Delta\phi' + \tfrac{f^2}{2\Delta}, \tag{2.40}$$

where $\phi' \equiv \phi - f/\Delta$. Path integrals allow shifting integration variables just like regular integrals do, so $\int \mathcal{D}\phi' = \int \mathcal{D}\phi$.[43] But thanks to our manipulations, ϕ' and f no longer

---

[43]Okay, that's a risky statement: there are subtle cases where the path integral doesn't take well to shifting the fields like that, which I will get back to on page 38. Those subtleties don't show up in calculations like the one we're doing at the moment, though.

show up together, so we can pull one outside the integral:

$$D(x,y) = \frac{-i\delta}{\delta f}\frac{-i\delta}{\delta f} \log\left[e^{\frac{if^2}{2\Delta}} \int \mathcal{D}\phi' e^{iS[\phi',f]} \right]\Bigg|_{f\to 0}, \tag{2.41}$$

The path integral doesn't care that ϕ has been renamed ϕ' everywhere inside it, so it simply gives a Z. Then, thanks to the properties of logarithms,

$$D(x,y) = \frac{-i\delta}{\delta f}\frac{-i\delta}{\delta f}\left[\frac{if^2}{2\Delta} + \log Z\right]\Bigg|_{f\to 0} = \frac{-i\delta}{\delta f}\left[\frac{f}{\Delta} + 0\right]\Bigg|_{f\to 0} = -\frac{i}{\Delta}. \tag{2.42}$$

How simple! Surely, we can now un-squint and see things for what they actually are: $1/\Delta$ is of course the Green's function $\Delta^{-1}(x,y)$.[44] Okay, maybe we shouldn't have un-squinted so fast: what horrible function could possibly be the inverse of an operator like $\Delta \equiv \partial^2 + m^2$?

This is where we must seek refuge in *momentum space*, taking the Fourier transform of absolutely everything. This turns $D(x,y)$ into $\tilde{D}(p)$, or roughly the probability that a particle will move with momentum p. No information is lost; momentum space and spacetime are just two sides of the same coin. But in momentum space, ∂_μ becomes ip_μ, so $\tilde{\Delta}$ is just $m^2 - p^2$, and that *is* just a number! Or, well, relativity would tell us that $p^2 = m^2$, and if $\tilde{\Delta} = 0$ we certainly can't just take $1/\tilde{\Delta}$. But are we sure that $p^2 = m^2$ holds? No! Just look at[45]

$$D(x,y) = \frac{1}{\sqrt{\pi}} \int \mathrm{d}^4 p \; e^{-ip_\mu(x-y)^\mu} \tilde{D}(p), \tag{2.43}$$

which clearly involves *all* possible p. Therefore, we can safely conclude that

$$\tilde{D}(p) = \frac{i}{p^2 - m^2} \tag{2.44}$$

and trust that someone more careful will work out a way to handle those cases where we do hit $p^2 = m^2$. This turns out to be both easy and elegant, but we don't have to bother with it—you get very far in practical calculations before that detail matters.

More particles. Will you be happy or disappointed if I tell you that the above was the hardest part? There's less math from here on, because eq. (2.44) embodies a big portion of what we have to know in order to calculate *anything*. Generalizing the propagator to more complicated fields seldom involves more than putting an identity matrix of whatever space ϕ now lives in in the numerator, and some γ^μ stuff in the case of fermions.

If you want to describe the situation of n particles moving about, all you have to do is make a version of eq. (2.38) with n derivatives. And if you work through those derivatives, you find there is little new to do: pair up the fields two by two, make a propagator for each pair,

---

[44]The function-space notation presented in the box on page 24 makes it not too hard at all to retrace all these calculations without squinting: just carefully write Δ^{-1} instead of $1/\Delta$, express things as inner products of functions, and keep track of where x and y go. The duck will quack mostly the same.

[45]The reason why $D(x,y)$ depends only on $(x - y)$ is *translation invariance*: physics is the same everywhere in the Universe, so $D(x,y)$ can only depend on the separation between x and y, not on their absolute position.

and sum over all ways of forming the pairs. For instance, with four points x_1, x_2, x_3, x_4, the result is

$$D(x_1, x_2)D(x_3, x_4) + D(x_1, x_3)D(x_2, x_4) + D(x_1, x_4)D(x_2, x_3) \,. \qquad (2.45)$$

We can represent this in pictures using our first example of *Feynman diagrams*, where a line between two points represents a propagator connecting them:

$$
\begin{array}{c}
x_4 \\
x_3
\end{array}
\Bigg)
\Bigg(
\begin{array}{c}
x_2 \\
x_1
\end{array}
\;\; + \;\;
\begin{array}{cc}
x_4 \!\!\!\!\rule[0.5ex]{3em}{0.4pt}\!\!\!\! x_2 \\
x_3 \!\!\!\!\rule[0.5ex]{3em}{0.4pt}\!\!\!\! x_1
\end{array}
\;\; + \;\;
\begin{array}{c}
x_4 \quad x_2 \\
\times \\
x_3 \quad x_1
\end{array}
\qquad (2.46)
$$

Here, two particles start at x_1, x_2 on the right and go to x_3, x_4 on the left.[46] The last two diagrams are easy to interpret: the particles move from right to left as expected, possibly swapping places. The first diagram, though, is a bit weird: do the particles get annihilated into nothing and then pop back into existence later? We'll have to think about that.

Interacting particles. Now, a natural instinct is to try our hands on something more savory than plain Klein–Gordon theory. What about QCD, for example? Unfortunately, it's practically impossible to redo pages 25 and 26 with Lagrangians that don't have the Klein-Gordon (or Dirac) equation as their Euler–Lagrange equation—just try, and you will be stuck trying to recover anything resembling eq. (2.39).

Very well, but what if we *insist* on being able to work with any Lagrangian? A good start is to split the action (and, similarly, the Lagrangian) like

$$S = S_0 + gS' \,, \qquad (2.47)$$

where S_0 is the easy part [the one that looks more or less like eq. (2.11)], and gS' is the rest. I have separated out a number g, called a *coupling*, because you usually find one multiplying the complicated bits of a Lagrangian. QCD is a good example: there, gS' is just the part containing the coupling g, namely $\bar{q}[g\gamma^\mu G_\mu]q$ and so on. In general, there may be many S', each with their own coupling.

Now, assume that g is small, so that gS' too is, in some sense, small. (Exactly *how* small depends on what we hope to achieve, but generally, g should be significantly less than 1.) This allows us to *Taylor expand* the path integral in g:

$$
\begin{aligned}
\int \mathcal{D}\phi \; e^{iS_0 + igS'} &= \sum_{k=0}^{\infty} \int \mathcal{D}\phi \; \frac{(igS')^k}{k!} e^{iS_0} \\
&= \int \mathcal{D}\phi \; e^{iS_0} + i \int \mathcal{D}\phi \; gS' e^{iS_0} - \frac{1}{2} \int \mathcal{D}\phi \; gS' \, gS' e^{iS_0} + \mathcal{O}(g^3) \,.
\end{aligned}
\qquad (2.48)
$$

---

[46]I adhere to the convention of time flowing right-to-left in each Feynman diagram. This is how we draw diagrams in papers IV and V, and reflects how the initial state is on the right and the final state on the left when defining the amplitude [see eq. (2.54)]. Other authors may have time flowing in other directions, with left-to-right probably being the most common. But in most of this thesis, including papers I and III, I draw diagrams with no time direction at all, as I explain on page 60.

In each term, we are back to the situation we can handle: an $\int \mathcal{D}\phi$, an e^{iS_0}, and a couple of fields that look as if they fell down through the application of derivatives like those in eq. (2.36). With g small enough, we can ignore everything except the first few terms in the expansion. (What "enough" means is again a bit vague, as is "few".) The first term that contributes is called *leading-order* (LO), the next is called the *next-to-leading-order* (NLO), then *next-to-next-to-leading-order* (NNLO) and so on, with N^kLO having k *next-to*'s.[47]

This is the gist of *perturbation theory*, most accessible and versatile tool of QFT (and many other parts of physics): for each term in eq. (2.48), we can keep applying our established method of pairing points and connecting them with propagators. But there is an important difference: all fields present in each S' are at the *same* point x.[48] When viewed like eq. (2.46), there will be multiple lines connected to one point, and it will look as if multiple particles met, touched, and then spread out again. Actually, that's pretty much what happens: gS' is how you introduce *interactions* between particles. There is also an integral over the interaction point, since $S' = \int d^4x \, \mathcal{L}'$, so it seems like the interaction gets summed over all the possible locations (in both time and space) where it could take place—let's not think too hard about that.

Things become clearer in everyone's favorite place, momentum space. Fourier transforming all the fields leaves nothing but a couple of e^{-ipx}'s inside the x integral, leading to

$$\int d^4x \, e^{-ip_1 x} e^{-ip_2 x} \cdots = \delta(p_1 + p_2 + \dots). \tag{2.49}$$

This Dirac delta is zero unless all the momenta add up to zero, which is easy to interpret: the momentum flowing into an interaction point must equal the momentum flowing out of it, which makes heaps of sense! In momentum space, there are no more sums over locations, but there is flow of momentum along all the lines [which are propagators like eq. (2.44)], and there is *conservation of momentum* wherever lines meet. Borrowing a term from graph theory, such an interaction point is called a *vertex*.

Momentum space really paints a new and better picture of particle processes. External particles [those fields that we have control over and insert using derivatives like in eq. (2.38)] come in from we-don't-really-care-where carrying momenta that flow through the diagram much like current in a circuit. The theory gives us the building blocks: S_0 gives the propagators, one for each type of particle, and S' gives the vertices, what kinds of propagators they connect, and any other details. This all comes together in a list called the *Feynman rules*.[49] With that in place, drawing the diagrams is a fun and intuitive activity.

Let's look at some examples, where the Feynman rules are a single kind of propagator and a vertex that joins four of those. Among other things, these Feynman rules are found in the

---

[47]A bit of a subtlety: normally, LO is defined as the first term that contributes to any given thing, but in some contexts (including ChPT; see page 46) it has a fixed meaning, and then you will see situations where the first order that contributes is called NLO or even NNLO.

[48]In principle, one can write an action with fields at different points. Such actions are called *nonlocal*, and they have a very hard time getting along with the principles of relativity. The deeper portions of string theory get away with nonlocality, but good old QFT doesn't.

[49]Some examples of important Feynman rules can be found in Peskin & Schroeder [14, appendix B]. A comprehensive, convention-independent list for the entire SM is given in ref. [21].

gluon-only part of QCD, and (as we will see later) in ChPT. Some of the simplest diagrams we can draw are

$$\text{and} \qquad \text{and} \qquad . \tag{2.50}$$

The *legs* pointing right are where particles come in, and the ones pointing left are where they go out; lines joining two vertices are propagators. Note the difference between the first diagram here and the last one in eq. (2.46): the particles don't just cross here, but meet and exchange momentum.

Let's take a closer look at the momentum flow in the middle diagram, with three particles going in (call their momenta p_1, p_2, p_3) and three particles coming out (call their momenta p'_1, p'_2, p'_3):

$$
\begin{aligned}
&\leftarrow p'_1 \qquad\qquad\quad \leftarrow p_1 \\
&\leftarrow p'_2 \quad \leftarrow q \quad \leftarrow p_2 \\
&\leftarrow p'_3 \qquad\qquad\quad \leftarrow p_3
\end{aligned}
\tag{2.51}
$$

The "particle" in the middle has momentum $q = p_1 + p_2 + p_3 = p'_1 + p'_2 + p'_3$ due to conservation. Yes, I put quotes around "particle": the things that move around in Feynman diagrams aren't really particles, but they have a lot of particle-like properties, so with the duck giving a careful half-quack, we call them *virtual particles*. One un-particle-y thing about virtual particles is that they don't obey $q^2 = m^2$, like we discussed around eq. (2.43). In eq. (2.51), you'll find that $q^2 \geq 3m^2$ no matter what p_1, p_2, p_3 we plug in.

Come to think of it, the last diagram in eq. (2.50) might be a bit upsetting: does the virtual particle go forward or backward in time? Well, neither! We're in momentum space, so there's no time, just the vague "before" and "after" of in- and outgoing particles. All that matters is momentum flow, and we can move vertices and bend lines to our hearts' content. I draw my diagrams the way I do for aesthetic reasons, but I could just as well have drawn the last one like

$$\text{or} \qquad \text{or} \qquad \text{or} \qquad \tag{2.52}$$

or like the heavily stylized ones on the front cover. Remember that in the end, everything is just a product of propagators and whatever other factors the theory drops in along them. The rest is just a picture.

Scattering and amplitudes. Let's not lose sight of what it's all about—where's the physics? Very well, then, let's put it like this: given some state $|X\rangle$ describing a handful of particles we've set up, what is the probability that some other handful $|Y\rangle$ comes out? If we hand the responsibility of implementing all the physics to some operator $\mathcal{S}$, imaginatively called the *S-matrix*,[50] $|X\rangle$ will turn into $\mathcal{S}|X\rangle$. There will be a certain "nothing happens" part to

---

[50]Between this, strangeness, the action, and various other things later on in the thesis, the letter S gets used for a lot of important things, so be sure to keep them apart. I hope it helps at least a little that I use calligraphic $\mathcal{S}$ for the S-matrix.

$\mathcal{S}$, which can be taken out and discarded,

$$\mathcal{S} = 1 + i\mathcal{T}, \tag{2.53}$$

where the *T-matrix* $\mathcal{T}$ is the "something happens" part. That is, $i\mathcal{T}|X\rangle$ combines all the new states $|X\rangle$ can become, and we pick out the one we are interested in with $\langle Y|$. Let me rephrase this as

$$\langle Y|i\mathcal{T}|X\rangle \equiv [\ldots]\,\mathcal{M}(X \to Y), \tag{2.54}$$

where $\mathcal{M}$ is the *(scattering) amplitude* for the process $X \to Y$,[51] and the "..." contains some normalizations and stuff that I ignore for simplicity.

Given that "scattering" and "amplitude" appears a lot in the title of my thesis and papers, you could guess that $\mathcal{M}$ is important. Then you'd be correct: it's the key to several worlds. On the one hand, it gives the *(scattering) cross-section* σ through

$$\sigma(X \to Y) = \int \mathrm{d}\Phi \, \mathcal{M}^*(X \to Y)\mathcal{M}(X \to Y). \tag{2.55}$$

The cross-section is, finally, something tangible: multiply it by the number of X's you put in, and you get the expected number of Y's as a result. It is typically measured in terms of a few parameters, like how much energy there is in total, while $\mathcal{M}$ depends on every single detail of all particles; $\int \mathrm{d}\Phi$ is the integral over all those less relevant details, hiding them from σ.

On the other hand, the amplitude is closely related to Feynman diagrams. There is just this rule: *each diagram must be a single, connected piece, where no leg can be separated from the rest of the diagram by removing a single propagator.* It ensures we ignore things that go into the "nothing happens" part of $\mathcal{S}$, or cancel against $1/Z$, or can be absorbed back into $|X\rangle$ and $|Y\rangle$; Peskin & Schroeder [14, chapter 4] explains the details.[52] If this rule is followed, $\mathcal{M}$ is the sum of all $X \to Y$ Feynman diagrams, and is often approximated well by cutting of the *perturbative series*, eq. (2.48), at some point.

Amplitudes are also great labor-saving devices, thanks to something called *crossing symmetry*. Note how back in eq. (2.46), I just said that the particles start at x_1, x_2 and end at x_3, x_4, but eq. (2.45) doesn't care which x_i is which. That extends to all the more complicated cases: you can pretty freely switch particles between "incoming" and "outgoing". For example, two diagrams in eq. (2.50) are related by such particle-switching, or *crossing*:

$$\tag{2.56}$$

Crossing a leg like this doesn't affect $\mathcal{M}$ very much: the momentum carried by the leg gets reversed ($p_i \to -p_i$) and the particle it represents gets replaced by its antiparticle (if it

---

[51]Sometimes called a *matrix element*, and sometimes written $\mathcal{A}$.

[52]In particular, all three diagrams in eq. (2.46) are disqualified from the amplitude, since they are not a single, connected piece. They are "nothing happens" diagrams—turns out we didn't have to think about them at all!

has one). Therefore, calculating a single diagram automatically gives you the contributions from many other diagrams. I use this a lot papers I and III, and extract some more clever relations from it in paper II.

Loops. We're done now, right? You just draw the Feynman diagrams and get all the scattering you ever wanted? No: *QFT is not that merciful*. Feynman diagrams get *weird* later on in the perturbative series, where there are too many vertices for the number of legs. Then we get *loops*,[53] like

$$\text{(diagrams)} \qquad \text{or} \qquad \text{(diagrams)} \qquad \text{or} \qquad \text{(diagrams)}, \qquad (2.57)$$

where conservation of momentum tells us nothing about the momenta going through the halves of the loop, except their sum. We have an undetermined momentum—or even two in the last diagram[54]—and a closer look at the Fourier transforms reveals that there is an integral over each such undetermined momentum—all four dimensions of it!

Evaluating such *loop integrals* is a deep and dark art, and its difficulty is usually what prevents us from going too far in the perturbative series. This thesis is limited to one loop for a good reason, although two [22] and three [23] loops have been handled in ChPT. The state of the art for physically relevant calculations is five [5, sec. 6.3.4], and for certain highly symmetric theories it has been pushed to eight [24]. I usually tell you not to be afraid of whatever mathematics QFT throws at you, but please *do* be afraid of loop integrals!

But maybe you are foolhardy, and sneak off to calculate a loop integral, just a simple one [eq. (4.5) is a good example]. Then you'll find that the result blows up in your face—the integral is infinite! That's perhaps not so strange: in most of momentum space, we are providing something particle-like with more energy and momentum than has ever been given to a real particle by even the most powerful particle accelerators or cosmic rays. Such high energies allow the particle to produce innumerable particle-antiparticle pairs out of the vacuum, or even touch on physics beyond what is currently known—QFT itself is expected to break down around 10^{19} GeV, so all of momentum space beyond that (which is most of it) is a realm of cosmic horror.

To avoid these problems, you should stop integrating at some point before things get bad.[55] A feature of Fourier transforms is that *large* momenta in momentum space correspond to *small* distances in regular space, so in effect, this means that you leave a tiny zone around

---

[53]This is also borrowed from graph theory, but not quite correctly: what we call "loops" are more properly known as *cycles*, with the word "loop" reserved for loops that attach to a single vertex (for whimsical reasons, these are known as *tadpoles* in Feynman diagrams).

[54]To count the number of loops, or equivalently the number of undetermined momenta, just see how many propagators can be removed from the diagram without separating it into multiple pieces.

[55]I should mention that just cutting off the loop integrals at some point is a clumsy (but valid) way to go about it, and there are other ways that are less easy to understand but more practical to use. Arguably the coolest one (and, incidentally, the one I normally use) is *dimensional regularization*, where one sets the number of spacetime dimensions to be *slightly less than four*, $d = 4 - 2\epsilon$. After taking appropriate care so that this makes sense (trust me, it does), one finds that the loop integrals are no longer infinite, but instead contain factors of $1/\epsilon$. One can then redefine the coupling to have a piece that also contains $1/\epsilon$, carefully adjusting its size so that all $1/\epsilon$'s cancel and ϵ can be safely brought to zero to restore the real-world dimension. Just like that, the infinities go away, leaving nothing but a modified coupling!

each particle unexplored. What we view as a particle is then not just the "bare" particle in the middle, but everything else inside the zone too, whatever it may be.

Of course, there is some freedom in exactly where we stop, and if we venture just a little further into the zone, the outer layers will just contain more of the physics we know how to handle; the weirdness is buried deeper. Take an electron in QED, for example. The outer parts of the zone will just contain more photons and electron-positron pairs. Photons do little, but the pairs reduce the electric field of the electron, in a way that is strikingly similar to how certain materials, *dielectrics*, reduce electric fields inside them. Therefore, as you peel away more and more of the pairs surrounding an electron by exploring higher energies, the strength of QED will seem to gradually increase because there is less of that reduction. That's right: weird as it may seem, the QED coupling increases when you go to higher energy.

QCD is similar, with quark-antiquark pairs reducing the strength of the strong force, but there is a crucial difference between gluons and photons: with no charged particles around, photons are non-interacting particles that simply satisfy the Klein–Gordon equation, but with no quarks around, the gluon Lagrangian still contains interactions due to the extra $-ig[G_\mu, G_\nu]$ term in eq. (2.27). Therefore, gluons form their own pairs, but due to a subtle difference in their Feynman rules, they *amplify* the force instead of reducing it. The effect of gluons beats that of quarks, so QCD will *weaken* dramatically when you go to higher energies.

QCD, unravelled. As a result of this weakening, called *asymptotic freedom*, high-energy QCD is a perfectly nice theory where you can draw Feynman diagrams to your heart's content. But g gets larger at lower energies, and below a few GeV, it's so large that each additional term in the perturbative series, eq. (2.48), contributes *more* than the last. Then perturbation theory breaks down, and with it our hopes of understanding things by drawing nice diagrams where it looks like free quarks and gluons are bumping into each other.

So, what is QCD like below this point? Well, the formal word for it is *confined*, which does a good job of describing the situation of quarks. Any color-charged object is surrounded by an immensely strong force field, courtesy of the *strong* force; this field consists of gluons, which are color-charged too, continuing the cycle and leading to a QCD soup where the concept of individual particles loses its meaning. Anything with color charge will be ruthlessly bullied by QCD, and anything that QCD allows to exist for any appreciable time must be color-neutral.

Electrons and their kin are of course color-neutral, since they have nothing to do with QCD, but quarks and gluons can (or rather, *must*) also form color-neutral combinations: these are *hadrons*, where the QCD soup is confined to the inside and leaves a more sensible situation outside. Color-neutrality makes plenty of sense for mesons: the color of the quark cancels the anticolor of the antiquark, leaving nothing. More formally, a quark is a vector q^i in color space, and an antiquark a conjugate vector $\bar{q}_i$, so their product $\bar{q}_i q^i$ is a scalar and therefore no different for color purposes than, say, an electron. There is another way, though: using the Levi-Civita symbol, three vectors can also form a (pseudo)scalar, $\epsilon_{ijk} q^i q^j q^k$, and three quarks is precisely what makes a baryon! Similarly, $\epsilon^{ijk} \bar{q}_i \bar{q}_j \bar{q}_k$ makes an antibaryon. A more intuitive picture is painted in fig. 3a.

(a) Illustration of how quark colors combine to form colorless states (black or white) in mesons (top), baryons (bottom left) and antibaryons (bottom right). Complementary colors are used for anticolor: **antired**, **antigreen** and **antiblue**.

(b) Illustration of how (top to bottom) a quark is knocked out of a baryon and forms a gluon string, which breaks and leaves newly formed mesons behind. Colored dots represent (anti)quarks, and the gluon field is represented in **gray**.

Figure 3: Schematic illustrations of some features of QCD, using real-world colors.

THE LEVI-CIVITA SYMBOL

An object with d indices in d-dimensional space that is *totally antisymmetric* (i.e., antisymmetric under swapping any two indices) has only one degree of freedom, since any element is either zero because it has two equal indices, or equal to ± 1 times the element with indices $123 \cdots d$. Therefore, it can be written $C\epsilon_{ij\ldots}$ where the *Levi-Civita symbol* ϵ is defined by $\epsilon_{123\ldots d} = 1$ and the constant C can be extracted with $\epsilon_{ij\ldots}\epsilon^{ij\cdots} = 1$.[i]

The Levi-Civita symbol is intimately tied to the determinant of a matrix:

$$\det(M) = \epsilon^{ijk\cdots} \epsilon_{i'j'k'\ldots} M_i^{i'} M_j^{j'} M_k^{k'} \cdots . \tag{2.58}$$

This can be proven by noting that the right-hand side has all the defining properties of the determinant, and recalling that the determinant is the unique function with these properties.

Despite its tensor-like appearance, the Levi-Civita symbol is a constant, *not* a tensor, but it is close to one: if it was a tensor, a transformation R would act like

$$R_i^{i'} R_j^{j'} R_k^{k'} \epsilon_{i'j'k'\ldots} = \det(R)\epsilon_{ijk\ldots} , \tag{2.59}$$

so it only differs from a tensor by a factor of $\det(R)$. This doesn't matter for rotations, which have $\det(R) = 1$, but it does for parity transformations, which have $\det(R) = -1$ in odd dimensions. Therefore, the word "pseudo" tends to appear when ϵ is around.

---

[i]Note that $\epsilon_{ij\ldots}$ is named after Tullio Levi-Civita, not two people named Levi and Civita.

So now that we understand why hadrons are what they are, let's restate the original question: why can't they be disassembled? Can't a quark be broken free from its hadronic prison, perhaps by using a particle collider to give it a hard kick so it flies out? Nope—it will still be connected with a force field carried by gluons, and unlike electromagnetic fields which weaken with distance as the field lines thin out, gluonic field lines "stick together" due to gluons interacting with each other, resulting in a taut "string" of force that never

weakens even as it's pulled longer and longer by the escaping quark.[56] The string holds immense tension—about 1 GeV/fm, or the weight of a fully-loaded truck—so after only a tiny distance, it has either pulled the quark back or built up enough energy to form several quark-antiquark pairs, each connected to another by new pieces of string. Once all available energy is spent on breaking the string this way, all quarks find themselves confined to new hadrons, never gaining freedom. This tyranny of QCD is roughly illustrated in fig. 3b.

2.4   Symmetries and how to break them

> *Break a vase, and the love that reassembles the fragments is stronger than that love which took its symmetry for granted when it was whole.*
>
> — *Derek Walcott*

With QCD behind us, we are now in principle ready to talk about ChPT, at least to the extent that I could pull its Lagrangian out of a hat and get to work drawing Feynman diagrams and calculating amplitudes. But that would leave the question of *why* the ChPT Lagrangian looks the way it does—it's pretty peculiar, much more so than those we've seen so far—and why ChPT should be a theory of light mesons at all. The answer to those questions lies in a remarkably elegant symmetry argument, so it's well worth spending some time looking at symmetries.

Noether and her currents.   Generally, a *current* is something that flows around, so we can describe its flow by a vector field, $J(x)$. All currents worth having (electric current, water current, etc.) are *conserved*, meaning that if more J flows into some region than flows out, some charge Q (electric charge, water, etc.) must build up there, and if Q decreases somewhere, more J must flow out than flows in. Writing $J^\mu \equiv (Q, J)$, the conservation criterion is neatly and Lorentz-invariantly captured by $\partial_\mu J^\mu = 0$.

Consider, then, a transformation. A *continuous* one, that can be made as small as you like, like a rotation; not like a permutation, which only happens in steps. If made small enough, such a transformation will look like $\phi \to \phi + \delta\phi$ for some small $\delta\phi$. If the Lagrangian is symmetric under this (that is, unchanged by it), then

$$0 = \frac{\delta\mathcal{L}}{\delta\phi}\delta\phi + \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\delta(\partial_\mu\phi) = \partial_\mu\left[\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\right]\delta\phi + \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\delta(\partial_\mu\phi) = \partial_\mu\left[\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\delta\phi\right], \quad (2.60)$$

where I used the Euler–Lagrange equation in the second equality. And would you look at that! We just found ourselves a current, $J^\mu \equiv \left[\delta\mathcal{L}/(\partial_\mu\phi)\right]\delta\phi$, that's conserved: $\partial_\mu J^\mu = 0$.

What we just found is *Noether's theorem*, a strong contender for the most profound and beautiful result in all of physics: *whenever there is a continuous symmetry, there is a conserved*

---

[56]This idea of a string can be put in a mathematically precise form called the *Lund string model* [25], which is greatly successful at predicting how hadrons form after particle collisions (see ref. [3, sec. 43] for an overview). It is not to be confused with string *theory*, which isn't successful at predicting anything.

current. For example, if we take the $\psi_i \to e^{ieQ_i\alpha}\psi_i$ symmetry of the QED Lagrangian, eq. (2.26), with multiple fields ψ_i of charge Q_i, the current we get is

$$J^\mu_{\text{QED}} = \sum_i ieQ_i\bar\psi_i\gamma^\mu\psi_i \,. \qquad (2.61)$$

This is a current, carried by matter particles (fermions), each proportional to their charge. Of course: this is the plain old electric current! According to Noether, electric charge can't be created or destroyed, just moved around, because eq. (2.26) happens to have a symmetry. If you ever look at an electronic device and wonder how it works, the answer is always "symmetry of eq. (2.26) under $\psi_i \to e^{ieQ_i\alpha}\psi_i$" (any extra details are just an exercise for the engineers). A similar look at QCD shows that each color is conserved in the same way.

THE FIFTH GAMMA MATRIX

There is a fifth gamma matrix,[i]
$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 \,, \qquad (2.62)$$
which has the easily verifiable properties $(\gamma^5)^2 = 1$ and $\{\gamma^5, \gamma^\mu\} = 0$. A key consequence of this is that γ^5 changes the behavior of parity transformations: $\bar q q$ is a scalar, but $\bar q\gamma^5 q$ is a pseudoscalar, and $\bar q\gamma^\mu q$ is a vector, but $\bar q\gamma^\mu\gamma^5 q$ is an axial-vector.[ii]

---

[i]The reason it's called γ^5 and not γ^4 is that, in less enlightened times, the time component of a Lorentz vector had index 4 instead of 0, so γ^0 was called γ^4 and was naturally followed by γ^5.

[ii]These four combinations, along with the tensor $\bar q[\gamma^\mu, \gamma^\nu]q$, completes the list of possible ways to combine $\bar q$ and q into something that transforms sensibly. This fact is known as *Fierz' theorem* [26].

Chiral symmetry. Perhaps the coolest word in the title of this thesis, *chiral*, makes its entry in the form of yet another symmetry of QCD. (Well, of *massless* QCD, strictly speaking, but the u, d and s quarks are so light, and the strong force so strong, that it makes sense to ignore masses for a while.)

Recall from before that "chiral" means two identical but mirrored things. The idea here is that quarks are chiral in that they meaningfully consist of a left and a right half:

$$q = q_L + q_R \,, \qquad \text{where} \qquad q_L = P_L q \,, \quad q_R = P_R q \,. \qquad (2.63)$$

You can verify yourself that if $P_L = \frac{1}{2}(1 + \gamma^5)$ and $P_R = \frac{1}{2}(1 - \gamma^5)$, they have all the necessary properties: $P_L + P_R = 1$ (the left and right are two halves of a whole), $P_L P_R = P_R P_L = 0$ (the left half has no right half, and vice versa) and so on.

Because of the $\pm\gamma^5$, P_L and P_R turn into each other when commuting past γ matrices:

$$P_L\gamma^\mu = \gamma^\mu P_R \,, \qquad P_R\gamma^\mu = \gamma^\mu P_L \,, \qquad \bar q_L = \bar q P_R \,, \qquad \bar q_R = \bar q P_L \,. \qquad (2.64)$$

All of this combined means that the structure $\bar q\gamma^\mu q$, which is the only form in which massless quarks enter the QCD Lagrangian, splits neatly into a left half and a right half:

$$\bar q\gamma^\mu q = \bar q(P_L + P_R)\gamma^\mu(P_L + P_R)q = \bar q_L\gamma^\mu q_L + \bar q_R\gamma^\mu q_R \,. \qquad (2.65)$$

Quarks are not only chiral, but those left and right halves are *completely separate* as far as gluons (and photons) are concerned! The could just as well be separate particles.

Another thing to note is that gluons treat all quark flavors the same, so we are free to rotate the quarks in flavor space: $q_f \rightarrow R_{ff'}q_{f'}$. And since the left and right halves are separated, we can do different rotations for q_L and q_R, giving the fabled *chiral symmetry* on which ChPT is built: *symmetry under independent left and right flavor rotations.*

There is, of course, some rain on this parade: the mass term in the Lagrangian reintroduces the left and right half to each other, since it contains a pesky

$$\bar{q}q = \bar{q}(P_L + P_R)(P_L + P_R)q = \bar{q}_L q_R + \bar{q}_R q_L \,, \tag{2.66}$$

and since different quarks have different masses, you can't really swap the flavors around all willy-nilly: Nature forces you to write $q^{\mathsf{T}} = (u, d, s, \ldots)$ and forbids any rotation ("$\ldots$" are those heavier quarks I keep ignoring). Still, Nature is pretty relaxed when forbidding this, and as long as you limit yourself to u, d and s quarks, you get a decent description of reality if you just pretend that chiral symmetry actually works. As I describe in section 3.3, ChPT is built by pretending this in the most proper and systematic way possible.

Spontaneous symmetry breaking. Noether worked in classical physics, and her theorem is a classical truth. There are two ways, one subtle and one *very* subtle, in which QFT can circumvent it. To see how, let's recall the path integral from eq. (2.29):

$$\langle 0|\hat{O}|0\rangle = \frac{1}{Z} \int \mathcal{D}\phi \, \hat{O} \, e^{iS} \,, \tag{2.67}$$

where $\hat{O}$ is any combination of fields and other things, such as the pair $\phi(x)\phi(y)$ that gave us the propagator, or really anything else; eq. (2.67) tells us the probability that, against the background of the vacuum $|0\rangle$, $\hat{O}$ will do whatever it is it does.

Noether's theorem gave a prediction based on the assumption that $\mathcal{L}$, and therefore S, has a symmetry. It doesn't really matter if $\hat{O}$ has the symmetry or not—we're free to poke at whatever non-symmetric nonsense we want—but we have quietly assumed that $|0\rangle$ is also symmetric. What if it isn't? To see how a vacuum can be non-symmetric, consider a field with n components, $\Phi^{\mathsf{T}} = (\phi_1, \phi_2, \ldots, \phi_n)$, and give it the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}\partial_\mu \Phi^{\mathsf{T}} \partial^\mu \Phi - m^2 \Phi^{\mathsf{T}}\Phi - g(\Phi^{\mathsf{T}}\Phi)^2 \,. \tag{2.68}$$

If $g = 0$, this is just eq. (2.14). The expression $m^2\Phi^{\mathsf{T}}\Phi + g(\Phi^{\mathsf{T}}\Phi)^2$ plays the role that the potential energy plays in classical physics, so reasonably, the lowest-energy state—the vacuum—is the one that minimizes $m^2\Phi^{\mathsf{T}}\Phi + g(\Phi^{\mathsf{T}}\Phi)^2$. If $g \geq 0$, this is simply at $\Phi = 0$, and $\mathcal{L}$ just describes particles of mass m on top of that vacuum.[57] Equation (2.68) is clearly symmetric under rotation of the n-dimensional space where Φ lives. This symmetry carries over to the vacuum and to the particles on top of it, since $\Phi = 0$ is the center of those rotations.

---

[57]This is known as *phi-four theory*, which (especially for $n = 1$) is one of the easiest QFTs. It is, incidentally, another theory whose Feynman diagrams look like those in eqs. (2.50), (2.52) and (2.57).

But let's say $g < 0$. Then some quick algebra reveals that the potential minimum is now at $\Phi^{\mathsf{T}}\Phi = -m^2/g$, which isn't a single point, but an entire sphere of radius $v \equiv m/\sqrt{-2g}$ in Φ-space. If we tried to calculate $\langle 0|\Phi|0\rangle$, it would be a vector in Φ-space, pointing who knows where. That's just as headache-inducing as it sounds.

Our salvation comes from the symmetry, which lets us *choose* where it points—after all, all directions in Φ-space are the same as far as the Lagrangian is concerned. A choice as good as any is $\langle 0|\Phi|0\rangle^{\mathsf{T}} = (0, \ldots, 0, v)$, so let me write

$$\Phi = (\phi_1, \ldots, \phi_{n-1}, v + \sigma) = (\boldsymbol{\phi}, v + \sigma)\,, \tag{2.69}$$

which introduces a new field σ (not to be confused with the cross-section!). Thanks to the v I pulled out of the ϕ_n component, both σ and the remaining $\phi_i's$ are zero in the vacuum. This v, which is the magnitude of $\langle 0|\Phi|0\rangle$, is called the *vacuum expectation value* (VEV) of Φ. Formulating eq. (2.68) in terms of these fields gives

$$\mathcal{L} = \tfrac{1}{2}\left[\partial_\mu\sigma\partial^\mu\sigma + \partial_\mu\boldsymbol{\phi}^{\mathsf{T}}\partial^\mu\boldsymbol{\phi}\right] - m^2\sigma^2 - 4\lambda v\sigma(\sigma^2 + \boldsymbol{\phi}^{\mathsf{T}}\boldsymbol{\phi}) - \lambda(\sigma^2 + \boldsymbol{\phi}^{\mathsf{T}}\boldsymbol{\phi})^2\,. \tag{2.70}$$

Two strange things happened here. Firstly, σ gets mass m, but $\phi_1, \ldots, \phi_{n-1}$ get no mass at all, since there is no term with $m\boldsymbol{\phi}^{\mathsf{T}}\boldsymbol{\phi}$. Secondly, our n-dimensional rotation symmetry is nowhere to be seen, since $\boldsymbol{\phi}$ and σ are completely different. There only remains a $(n-1)$-dimensional symmetry when rotating $\boldsymbol{\phi}$. I promise that this is not because we did something wrong: the symmetry has actually *spontaneously broken* because of the vacuum, from n-dimensional to $(n-1)$-dimensional rotations.

This is an example of a broader concept, *Goldstone's theorem*: whenever a symmetry spontaneously breaks, a number of new, massless scalar particles, *Nambu–Goldstone bosons* (NGBs), appear. The number of NGBs is equal to the number of symmetries lost, and in our case, the difference in degrees of freedom between n-dimensional and $(n-1)$-dimensional rotations is indeed $n - 1$.[58]

Spontaneous symmetry breaking in ChPT. When confinement sets in, the so-called *quark condensate* appears:

$$0 \neq \langle 0|\bar{q}q|0\rangle = \langle 0|\bar{q}_L q_R|0\rangle + \langle 0|\bar{q}_R q_L|0\rangle\,. \tag{2.71}$$

Like $\langle 0|\Phi|0\rangle$, this has destructive consequences: it removes the separation between left and right, breaking chiral symmetry and leaving only symmetry under simultaneous rotations in both halves of flavor space. Goldstone tells us that there will be some NGBs as a result, and as I will show more convincingly in the next section, these NGBs are none other than the pseudoscalar mesons: pions, kaons, and η!

But wait, you say—these can't be NGBs, since NGBs don't have mass! Well, remember that since quarks already have a little mass, chiral symmetry was never a perfect symmetry in the first place, so the symmetry breaking will also not be perfect. But as I said before, reality can be closely *approximated* by one where chiral symmetry is perfect and breaks perfectly, so we will get particles that are *approximately* NGBs (called *pseudo-NGBs*). This

---

[58]For fans of the fancier parts of the SM, I should mention that this is similar to how the Higgs mechanism works, with the Higgs boson playing a role similar to σ and the NGBs giving the weak force its structure.

gives the pions, kaons and η the excuse to be closer to massless (that is, lighter) than the other hadrons, and the excuse works best for the pions, which don't involve the heavier strange quark.

But wait, you say again—what about η'? It's in the same list as the pions, kaons and η, so why isn't it a pseudo-NGB too? Well, that's a very good question. To answer it, I have to summon some even deeper, darker parts of QFT—please hold on while I fetch my candles, pentagrams and sacrificial goats.

Quantum anomalies. Spontaneous symmetry breaking was the subtle way to circumvent Noether; the *very* subtle way has to do with the last part of eq. (2.67) we haven't looked at yet: $\mathcal{D}\phi$. When working with eq. (2.39), I said in passing that my change $\phi \to \phi'$ just changed $\mathcal{D}\phi \to \mathcal{D}\phi'$. But what if our transformation, the one that we hope is a symmetry, doesn't do this? What if it actually changes $\mathcal{D}\phi$? Then the "symmetry" will actually change any $\langle 0|\hat{O}|0\rangle$ we try to calculate, so it won't be a symmetry at all!

This most subtle of subtleties is known as an *anomaly*, because when you accidentally run into it when trying to calculate Feynman diagrams, it looks like the result of devious calculation mistakes rather than something deeper. Sadly, I don't know of any anomaly that's simple enough to derive here, so you'll have to take my word for what I'm about to say.[59]

As you can verify, the QCD Lagrangian, eq. (2.28), is symmetric under the *axial transformation* $q \to e^{-i\theta\gamma^5}q$, so there should be a current, the *axial current* (so called because it's an axial-vector),

$$J_5^\mu \equiv \sum_f \frac{\delta\mathcal{L}}{\delta(\partial_\mu q_f)}[-i\gamma^5 q_f] \quad \Rightarrow \quad J_5^\mu = \bar{u}\gamma^\mu\gamma^5 u + \bar{d}\gamma^\mu\gamma^5 d + \bar{s}\gamma^\mu\gamma^5 s + \dots . \quad (2.72)$$

(Again, "..." would include those heavier quarks we are consistently ignoring.) According to Noether, J_5^μ should be conserved, but if you very, *very* carefully check how the axial transformation affects $\mathcal{D}\bar{q}\mathcal{D}q$, you find that it changes. This change—this anomaly—breaks Noether's theorem, so the axial current isn't conserved at all. Instead, one finds

$$\partial_\mu J_5^\mu \equiv \frac{3g^2}{32\pi^2}\epsilon_{\mu\nu\alpha\beta}G^{\mu\nu}G^{\alpha\beta} . \quad (2.73)$$

Note how the structure of J_5^μ matches the quark content of the η' meson (see table 2). And yes, it's as simple as that: this current is carried by η' mesons, and the anomaly essentially tells us that η' is free to transform into gluons, $G^{\mu\nu}$, and back, which affects its composition and therefore its mass. Seen from another angle, the anomaly sets a particular combination of flavors apart from the others, destroying part of the chiral symmetry of QCD. With less symmetry to break, there is one NGB fewer than there would otherwise have been, and the excluded one is of course η'. This explains its large mass: since η' isn't even *approximately* a NGB, it has no excuse for being light.

---

[59]Zee [15, chapter IV.7] has an exposition that nicely captures the frustrated confusion physicists must've felt the first time they ran into anomalies. Donoghue [18, chapter III-3] has a more sober, thorough treatment, and Peskin & Schroeder [14, chapter 19] has one with lots of details and examples; both of these are quite advanced. Pich [16, sec. 7] has a succinct description of what is needed in ChPT.

The currents carried by η and π^0 also have anomalies, but here it's the *photon* field that shows up, not the gluon field. Therefore, η and π^0 are still non-anomalous as far as pure QCD is concerned, but in the full SM, they can decay into photons. This is very important—if it wasn't for the anomaly, π^0 would be much stabler than we find it to be in experiments, since becoming two photons is one of the few possible ways for it to decay. That's how important symmetries are for particle physics: even when they fail, they tell us something about the way reality behaves.

Alright, I think that's plenty of material for even the most diligent non-expert to digest. It's about time we let the experts back in.

3 Chiral Perturbation Theory

Pions! Pions! We don't understand the electron. Why do you bother with pions?

— *Albert Einstein, as quoted by Leon Lederman (2014)*

In this section, I present the framework in which essentially all my work has been performed: Chiral Perturbation Theory (ChPT). Unlike the previous section, I assume that you, the reader, are some kind of particle physicist, or something close enough, or at least that you took the previous section to heart and are hungry for more.[60] Nevertheless, I will keep the pace gentle for the benefit of those who come from completely different corners of the field (including my friends, the experimentalists), or those who are just a bit rusty.

3.1 Sigma models

To illustrate the structure and emergence of ChPT, let us revisit the basic spontaneous symmetry breaking discussed in section 2.4; in particular, consider eq. (2.70), the so-called *linear sigma model*. In the case $n = 4$, it can be recast in a more interesting form. Rather than using the 4-component vector Φ, we use the 2×2 matrix

$$\Sigma = \tfrac{1}{2}\big[(v + \sigma) + i\sigma^i \phi_i\big], \tag{3.1}$$

where σ^i (not to be confused with σ) are the Pauli matrices [see eq. (3.50)]. This is equivalent to the Φ representation: it contains the same four fields, and due to the trace properties of σ^i,

$$\langle \Sigma^\dagger \Sigma \rangle = \Phi^\mathsf{T} \Phi. \tag{3.2}$$

The unbroken Lagrangian eq. (2.68) can be compactly written

$$\mathcal{L} = \tfrac{1}{2}\langle \partial_\mu \Sigma^\dagger \partial^\mu \Sigma \rangle - \lambda\big[\langle \Sigma^\dagger \Sigma \rangle - v^2\big]^2. \tag{3.3}$$

The most general transformation that leaves this Lagrangian invariant is $\Sigma \to g_L \Sigma g_R^\dagger$, where $g_{L,R}$ are 2×2 unitary matrices. We thus have a $U(2)_L \times U(2)_R$ symmetry, which is quite chiral-looking. This symmetry is broken by the VEV,

$$\langle 0|\Sigma|0\rangle = v \quad \longrightarrow \quad \langle 0|g_L \Sigma g_R^\dagger|0\rangle = v g_L g_R^\dagger, \tag{3.4}$$

so only the subgroup where $g_L = g_R$, which we call $SU(2)_V$, remains a symmetry.

In the treatment of section 2.4, $SO(4)$ symmetry was broken to $SO(3)$ symmetry.[61] It should therefore not come as a surprise that $SO(4) \cong SU(2) \times SU(2)$ and $SO(3) \cong SU(2)$,

---

[60]If that is the case: Amazing! Are you sure you don't want to *become* a particle physicist?

[61]Strictly speaking, eqs. (2.68) and (2.70) have $O(n)$ and $O(n-1)$ symmetry, respectively, but $O(n) \cong SO(n) \times \mathbb{Z}_2$, where $\mathbb{Z}_2$ captures the determinant, ± 1, of the transformation; only $SO(n)$ is continuous and relevant to Goldstone's theorem. [Compare $U(n) \cong SU(n) \times U(1)$, where $U(1)$ captures the now continuous set of possible determinant values.]

so that $SO(4) \to SO(3)$ and $SU(2)_L \times SU(2)_R \to SU(2)_V$ are equivalent as symmetry breaking patterns.

The form of Σ unfortunately makes it inconvenient to write a more explicit Lagrangian like eq. (2.70). Instead, we once again re-express the fields, this time in a *non-linear* fashion, so that

$$\Sigma = \tfrac{1}{2}[v + \hat{\sigma}]U(\hat{\phi}), \qquad U(\hat{\phi}) = \exp\left[i\sigma^i \hat{\phi}_i/v\right], \qquad (3.5)$$

where $\hat{\sigma}, \hat{\phi}$ are defined in terms of σ, ϕ. The transformation properties under $SU(2)_L \times SU(2)_R$ must be $\hat{\sigma} \to \hat{\sigma}$ and $U(\hat{\phi}) \to g_L U(\hat{\phi})g_R$; the complicated transformations of $\hat{\phi}$ themselves are not needed explicitly. With this, the broken Lagrangian, eq. (2.70), becomes

$$\mathcal{L} = \frac{v^2}{4}\left[1 + \frac{\hat{\sigma}}{v}\right]^2 \langle\partial_\mu U^\dagger \partial^\mu U\rangle + \tfrac{1}{2}\partial_\mu\hat{\sigma}\partial^\mu\hat{\sigma} - m^2\hat{\sigma}^2 - \frac{m^2}{2v}\hat{\sigma}^3 - \frac{m^2}{8v^2}\hat{\sigma}^4. \qquad (3.6)$$

This neatly separates the massive and massless fields, and sets the stage for the next step.

3.2 Effective field theory

You can write things much more nicely if you replace formulations like "all kinds of crap" with the word "effective".

— *Tanja Hinderer (2019)*

Although physicists seek to describe nature in as fundamental a way as possible, it is seldom practical to always have in mind everything all the way down to the fundaments. It would be prohibitively complicated to do nuclear physics with quarks and gluons in mind, biochemistry with nuclear structures in mind, or atmospheric science with individual molecules in mind. Omitting details not accessible at one's working resolution is, on one hand, usually an extremely good approximation, and on the other, absolutely crucial for obtaining a manageable theory.

This principle is somewhat formalized by the concept of an *effective field theory* (EFT): a theory which ostensibly is not fundamental, but which accurately describes the effective behavior of things in some limited context. The canonical way to obtain an EFT is by *integrating out* massive degrees of freedom. Take, for example, the Lagrangian eq. (3.6) and consider processes involving energies much smaller than the $\hat{\sigma}$ mass m. Then $\hat{\sigma}$ clearly cannot appear in the initial or final state, and its presence in intermediate states will be hidden from view; the propagators $i/(q^2 - m^2)$ will be indistinguishable from constant factors i/m^2 when $q^2 \ll m^2$. Thus, it will look as if only $\hat{\phi}$ fields are involved, although they will posses additional modes of interactions whose couplings involve factors of $1/m^2$. The formal process of integrating out $\hat{\sigma}$ is nicely described in Donoghue [18, pp. 112-4], but here I am content with stating the result: the *effective Lagrangian* describing only $\hat{\phi}$ is

$$\mathcal{L}_{\text{eff}} = \frac{v^2}{4}\langle\partial_\mu U^\dagger \partial^\mu U\rangle + \frac{v^2}{8m^2}\langle\partial_\mu U^\dagger \partial^\mu U\rangle^2 + \mathcal{O}(m^{-4}), \qquad (3.7)$$

where each additional term in $\mathcal{O}(m^{-4})$ is easy to determine from eq. (3.6), but since m is assumed to be very large compared to other scales involved, it is safe to truncate the series rather early. However, one must keep in mind that as the energy approaches m, the EFT ultimately breaks down and the exact Lagrangian, eq. (3.6), must be restored.

An EFT for QCD. Equation (3.7) is an example of the *nonlinear sigma model* (NLSM),[62] which in a generalized form is the foundation of ChPT. However, ChPT is different in some very important ways: there is no heavy degree of freedom to integrate out (rather, the degrees of freedom we wish to hide are quarks and gluons, which are light), and the degrees of freedom that are supposed to remain are mesons, which are not present in the QCD Lagrangian. Furthermore, eq. (3.6) [and even more so, eq. (2.70)] are quite easy to work with, so any results obtained from eq. (3.7) can be checked against the exact results. With low-energy QCD, Nature has no such mercy, and therefore ChPT must be capable of sweeping much complexity under the carpet.

The way forward is to put all our trust into spontaneous symmetry breaking. As briefly covered in section 2.4, massless n_f-flavor QCD is symmetric under chiral $\mathrm{U}(n_f)_L \times \mathrm{U}(n_f)_R$ rotations in flavor space, with $\mathrm{U}(n_f) \cong \mathrm{SU}(n_f) \times \mathrm{U}(1)$, but the axial anomaly, eq. (2.73), destroys one $\mathrm{U}(1)$ symmetry, and the other $\mathrm{U}(1)$, which causes baryon number conservation, is irrelevant for our purposes. Thus, our original symmetry is the chiral $\mathrm{SU}(n_f)_L \times \mathrm{SU}(n_f)_R$ symmetry, which we can express as

$$q_L \to g_L q_L\,, \qquad q_R \to g_R q_R\,, \qquad g_{L,R} \in \mathrm{SU}(n_f)_{L,R}\,, \tag{3.8}$$

where the left- and right-handed quarks $q_{L,R}$ form vectors in flavor space: $q^\mathsf{T} = (u, d, s, \ldots)$. The emergence of the quark condensate, eq. (2.71), breaks the symmetry into the subgroup $\mathrm{SU}(n_f)_V$ where $g_L = g_R$: that is, $\mathrm{SU}(n_f)_L \times \mathrm{SU}(n_f)_R \to \mathrm{SU}(n_f)_V$. *A priori*, we know very little about this breaking other than the fact that it *happens*—its details are a matter of intractable low-energy QCD—so we must now consider the effects of such a symmetry breaking on a general, abstract plane.

General symmetry breaking. Consider a symmetry breaking $\mathcal{G} \to \mathcal{H}$, where $\mathcal{H}$ is a subgroup of $\mathcal{G}$. By Goldstone's theorem, this will give rise to $\dim(\mathcal{G}/\mathcal{H})$ NGBs, which we label ϕ. I write the action of an element $g \in \mathcal{G}$ on these NGBs as $g\phi$, but bear in mind that they may act in an arbitrarily complicated representation. Furthermore, recalling that before a specific choice of vacuum was made, the entire manifold containing the NGBs was symmetric under $\mathcal{G}$, it must be possible to reach any ϕ from the vacuum, ϕ_0, with an appropriate transformation g_ϕ:

$$\phi = g_\phi \phi_0 = g_\phi h \phi_0, \qquad h \in \mathcal{H}\,, \tag{3.9}$$

where in the second step I used the fact that, by definition, ϕ_0 is invariant under $\mathcal{H}$. This associates ϕ with the entire *(left) coset* $g_\phi \mathcal{H} \equiv \{g_\phi h \mid h \in \mathcal{H}\}$. This association is unique, for if two different cosets $g\mathcal{H}, g'\mathcal{H}$ are associated with the same ϕ, then

$$g\phi_0 = g'\phi_0 \quad \Rightarrow \quad \phi_0 = g'g^{-1}\phi_0 \quad \Rightarrow \quad g'g^{-1} \in \mathcal{H} \quad \Rightarrow \quad g' \in g\mathcal{H} \quad \Rightarrow \quad g'\mathcal{H} = g\mathcal{H} \tag{3.10}$$

---

[62]Note that, despite having "sigma" in its name, it no longer has any field called σ as a degree of freedom, since it has been integrated out.

due to the closedness of $\mathcal{H}$.

We have now established that the field ϕ and the coset $g_\phi \mathcal{H}$ are entirely equivalent. Based on this, we say that ϕ lives in the *coset space* $\mathcal{G}/\mathcal{H} \equiv \{g\mathcal{H} \mid g \in \mathcal{G}\}$, which also motivates why the number of NGBs is $\dim(\mathcal{G}/\mathcal{H})$, as Goldstone's theorem states. Since cosets are sets rather than single group elements, it is simpler to take one arbitrary *coset representative* $\xi(\phi) \in g_\phi \mathcal{H}$ to represent ϕ. This runs into consistently issues, though: if $\xi(\phi)$ is the representative from $g_\phi \mathcal{H}$, there is no guarantee that $g\xi(\phi) = \xi(g\phi)$, where $\xi(g\phi)$ is the representative from $gg_\phi \mathcal{H}$ chosen to represent $g\phi$. However, there exists some $h[\xi(\phi), g] \in \mathcal{H}$ such that $g\xi(\phi) = \xi(g\phi)h[\xi(\phi), g]$, and since the ξ-choice was arbitrary to start with, we have the leeway to simply include this h in the transformation properties of $\xi(\phi)$:

$$\xi(\phi) \to g\xi(\phi)h^\dagger\big[\xi(\phi), g\big] = \xi(g\phi), \tag{3.11}$$

Now, let us return to the case relevant for QCD, where $\mathcal{G} = \mathrm{SU}(n_f)_L \times \mathrm{SU}(n_f)_R$ and $\mathcal{H} = \mathrm{SU}(n_f)_V$. Then $g = (g_L, g_R)$ and we correspondingly write $\xi(\phi) = \big(\xi_L(\phi), \xi_R(\phi)\big)$, with transformation properties

$$\xi_L(\phi) \to g_L\xi_L(\phi)h^\dagger, \qquad \xi_R(\phi) \to g_R\xi_R(\phi)h^\dagger, \tag{3.12}$$

where $h \equiv h\big[\xi(\phi), g\big]$ is the same for both L and R since $\mathrm{SU}(n_f)_V$ is invariant under the swap $L \leftrightarrow R$.

We can now make a curious observation: since $\mathrm{SU}(n_f)_L \cong \mathrm{SU}(n_f)_R \cong \mathrm{SU}(n_f)_V$, the cosets $g_{L,R}\,\mathrm{SU}(n_f)_V$ are also isomorphic to $\mathrm{SU}(n_f)$.[63] Thus, ξ may in fact choose *any* element of $\mathrm{SU}(n_f)$ as a representative for each ϕ.[64] We are therefore free to make a choice such as

$$\xi_R(\phi) = \xi_L^\dagger(\phi) \equiv u(\phi), \qquad u(g\phi) = g_R u(\phi)h^\dagger = hu(\phi)g_L^\dagger, \tag{3.13}$$

or—to get rid of that pesky h—use

$$U(\phi) \equiv \big[u(\phi)\big]^2, \qquad U(g\phi) = g_R U(\phi)g_L^\dagger. \tag{3.14}$$

This is deliberately reminiscent of the U introduced in eq. (3.5). Indeed, we can [possibly after a non-linear rearrangement of the NGBs similar to that which led to eq. (3.5)] parametrize $U(\phi)$ as[65]

$$U(\phi) = \exp\left[\frac{i\phi_a t^a}{F}\right]. \tag{3.15}$$

Here, t^a are the generators of $\mathrm{SU}(n_f)$ (equal to the Pauli matrices for $n_f = 2$, and the Gell-Mann matrices for $n_f = 3$; see section 3.4), and F is a constant introduced for future convenience.

---

[63]In paper III, we deal with a number of theories (explained in ref. [27]) with a similar structure to ChPT but different symmetry breaking patterns . These patterns invalidate some parts of this derivation—some are not even chiral—but rest assured that the end result remains valid.

[64]Counter-intuitively, this does not invalidate the above statements about the ϕ-to-$g_\phi \mathcal{H}$ relation being unique. The reason for this is essentially the oft-neglected difference between "isomorphic" and "equal".

[65]This is the most common parametrization, and the one that makes it the most manifest that $U \in \mathrm{SU}(n_f)$, but it is far from the only one; generally, any symmetry-consistent reparametrization $\phi \to \phi + \mathcal{O}(\phi^2)$ will leave the physics intact. Appendix B of paper III discusses the most general parametrization of U.

Terms and coefficients. What we have seen so far is that, based entirely on the chiral symmetry breaking pattern, our theory must contain NGBs in the shape of $U(\phi)$. No mention is made of other fields, so a readily available (and, it turns out, well-motivated) assumption is that there *are* no other fields in the theory! Conveniently, this avoids the issue of integrating anything out. However, we do not yet have a Lagrangian, and without any further information, the best we can do is to write down *the most general Lagrangian built from $U(\phi)$ and consistent with its transformation properties*. This most general Lagrangian has an infinite number of terms, but the simplest of them are

$$\mathcal{L} = \frac{F^2}{4}\langle\partial_\mu U^\dagger \partial^\mu U\rangle + \hat{L}_0\langle\partial_\mu U^\dagger\partial_\nu U\partial^\mu U^\dagger\partial^\nu U\rangle + \hat{L}_1\langle\partial_\mu U^\dagger\partial^\mu U\rangle\langle\partial_\nu U^\dagger\partial^\nu U\rangle$$
$$+ \hat{L}_2\langle\partial_\mu U^\dagger\partial_\nu U\rangle\langle\partial^\mu U^\dagger\partial^\nu U\rangle + \hat{L}_3\langle\partial_\mu U^\dagger\partial^\mu U\partial_\nu U^\dagger\partial^\nu U\rangle + \dots, \quad \text{(3.16)}$$

where $\hat{L}_i$ are unknown constants (more about them shortly) and "..." consists of terms with more than four derivatives. Anything else one can write down is either inconsistent, part of the "...", or can be reduced to these five terms using identities such as

$$U^\dagger U = 1, \qquad 0 = \partial_\mu(U^\dagger U) = \partial_\mu U^\dagger U + U^\dagger\partial_\mu U, \qquad \text{(3.17)}$$

etc. Equation (3.16) brings us rather close to our goal: it is, in fact, a basic variant of the ChPT Lagrangian!

While the symmetry considerations are very powerful in determining the structure of the possible interactions, they cannot decide their strength, which is why the $\hat{L}_i$ appear in eq. (3.16). In a general EFT context, such constants are commonly called *Wilson coefficients*, but in ChPT, their canonical name is *low-energy constants* (LECs). The LEC of the first term is related to F by the requirement that, after expanding U in terms of ϕ, the $\frac{1}{2}\partial_\mu\phi^a\partial^\mu\phi^a$ term shows up with the correct normalization; but the other LECs are only determined by the details of the underlying theory.

We have already encountered one example of LEC determination: taking eq. (2.70) as the underlying theory results in eq. (3.7), which is eq. (3.16) with $n_f = 2$, $F = v^2$, $\hat{L}_1 = v^2/8m^2$ and $\hat{L}_0 = \hat{L}_2 = \hat{L}_3 = 0$. But for QCD, obtaining the LECs that way is hopeless. Paper II deals with how self-consistency conditions on ChPT forbid some LEC values, but in general, the only way to determine them is through experiments or lattice simulations.

Despite containing this infinite sequence of unknown parameters, ChPT is still quite a predictive theory: as we will see below, only a limited set of LECs is relevant at any given level of precision, and to that precision, the *same* set of LECs must match all measurable processes. A nice example of this is seen in paper IV, where fitting the LECs directly to the process at hand agrees reasonably well with the standard set of LECs determined from a wide array of other processes.

3.3 The ChPT Lagrangian

Having now arrived at the gist of ChPT through a lengthy, mostly-from-scratch derivation, it is about time I gather all the pieces and present it in the full glory of its modern form.

Before doing that, let me quickly summarize its history.

The linear and non-linear sigma models were introduced by Gell-Mann and Lévy [28] before QCD was conceived of, and the latter was taken much of the way toward ChPT by Weinberg [29] as QCD was gaining acceptance, but it was in the work of Gasser & Leutwyler [30, 31] that ChPT as such was born. Importantly, they solved an issue I have avoided ever since first describing chiral symmetry: *chiral flavor symmetry is not a symmetry of QCD*. It is broken by by quark masses, and by most non-QCD interactions that the quarks partake in (for instance, electromagnetism breaks it since the u, c, t and d, s, b quarks have different charges). This is not an anomaly or spontaneous symmetry breaking giving rise to interesting physics—it is simply the absence of symmetry.

However, as I argued already in section 2.4, chiral flavor symmetry is an *approximate* symmetry of QCD. Non-QCD interactions are very weak compared to QCD, and can be mostly ignored. The masses of the u and d quarks are very small, and s is fairly light as well, so chiral symmetry is an excellent approximation for $n_f = 2$, and a decent one for $n_f = 3$. For larger n_f it is sadly out of the question, at least in our universe's version of QCD.

What Gasser & Leutwyler did was to systematically incorporate this slight breaking of chiral symmetry into the framework built up under the assumption that QCD is chirally symmetric, using the *external field method*. It consists in writing down all consistent ways for quarks to interact with something outside pure, massless QCD,

$$\mathcal{L} = \mathcal{L}_{\mathrm{QCD}}^{[m=0]} + \bar{q}\{\gamma^\mu(v_\mu + \gamma^5 a_\mu) - (s - i\gamma^5 p)\}q\,, \tag{3.18}$$

where v_μ, a_μ, s, p (so called because they are *v*ector, *a*xial-vector, *s*calar and *p*seudoscalar fields, respectively) are matrices in flavor space.[66] They may simply be fixed objects inserted into the theory, or may be fully dynamic quantum fields; electroweak interactions can be incorporated through v_μ and a_μ, and Higgs interactions through s, which is also where the quark masses enter.

Lagrangian building blocks. In order to build ChPT with these external fields added, they must be given transformation properties consistent with chiral symmetry. How to do this is most easily seen by forming the linear combinations

$$\ell_\mu \equiv v_\mu - a_\mu\,, \qquad r_\mu \equiv v_\mu + a_\mu\,, \qquad \chi \equiv 2B(s + ip)\,, \tag{3.19}$$

(I will get back to the meaning of the constant B later) so that

$$\mathcal{L} = \mathcal{L}_{\mathrm{QCD}}^{[m=0]} + \bar{q}_L\gamma^\mu\ell_\mu q_L + \bar{q}_R\gamma^\mu r_\mu q_R - \tfrac{1}{2B}\left[\bar{q}_R\chi q_L - \bar{q}_L\chi^\dagger q_R\right]. \tag{3.20}$$

Thus, under $g = (g_L, g_R)$ they must transform as

$$\ell_\mu \to g_L\ell_\mu g_L^\dagger + ig_L\partial_\mu g_L^\dagger\,, \qquad r_\mu \to g_R r_\mu g_R^\dagger + ig_R\partial_\mu g_R^\dagger\,, \qquad \chi \to g_R\chi g_L^\dagger\,. \tag{3.21}$$

---

[66]This is not an exhaustive list, but it covers all SM interactions and includes all fields normally used in Lagrangian-building. The consequences of including some further fields are studied in refs. [32, 33].

The field $U(\phi)$ remains with the same structure and transformation as above, although the ϕ^a are now demoted to *pseudo*-NGBs and may obtain masses.

These fields are still not entirely convenient for putting together the most general Lagrangian, since their transformation properties—and even more so those of their derivatives—are not easy to keep track of. One improvement, borrowing forms and terminology from gauge theory, is the introduction of the *field strength tensors*

$$
\begin{aligned}
F_L^{\mu\nu} &\equiv \partial^\mu \ell^\nu - \partial^\nu \ell^\mu - i[\ell^\mu, \ell^\nu] &\rightarrow\quad & g_L F_L^{\mu\nu} g_L^\dagger \,, \\
F_R^{\mu\nu} &\equiv \partial^\mu r^\nu - \partial^\nu r^\mu - i[r^\mu, r^\nu] &\rightarrow\quad & g_R F_R^{\mu\nu} g_R^\dagger \,.
\end{aligned}
\tag{3.22}
$$

Originally [30, 31], the ChPT Lagrangian was written in terms of U, χ, $F_{L,R}^{\mu\nu}$ and suitable derivatives. However, I will present the Lagrangian in terms of the alternative "building blocks" introduced in ref. [34] , which make the Lagrangian both easier to construct and to look at. The building blocks all transform like $X \rightarrow hXh^\dagger$, where $h \in \mathrm{SU}(n_f)_V$ is the compensatory transformation introduced in eq. (3.11). To achieve this, we bring back the field u introduced in eq. (3.13), and work out that

$$
\begin{aligned}
u_\mu &\equiv i\big[u^\dagger(\partial_\mu - ir_\mu)u - u(\partial_\mu - i\ell_\mu)u^\dagger\big] \,, \\
\chi_\pm &\equiv u^\dagger \chi u^\dagger \pm u \chi^\dagger u \,, \qquad f_\pm^{\mu\nu} \equiv u F_L^{\mu\nu} u^\dagger \pm u^\dagger F_R^{\mu\nu} u
\end{aligned}
\tag{3.23}
$$

are the simplest reformulations of u, χ, $F_{L,R}^{\mu\nu}$ with the desired transformation properties. Furthermore, one can construct the *covariant derivative* ∇_μ such that if $X \rightarrow hXh^\dagger$, then $\nabla_\mu X \rightarrow h\nabla_\mu X h^\dagger$ as well:

$$
\nabla_\mu X \equiv \partial_\mu X + [\Gamma_\mu, X] \,, \qquad \Gamma_\mu \equiv \tfrac{1}{2}\big[u^\dagger(\partial_\mu - ir_\mu)u + u(\partial_\mu - i\ell_\mu)u^\dagger\big] \,.
\tag{3.24}
$$

Now for the important topic of *power counting*, which will allow us to systematically organize the Lagrangian into LO, NLO, NNLO, etc., sections, such that the $\mathrm{N}^k\mathrm{LO}$ Lagrangian is only needed in $\mathrm{N}^K\mathrm{LO}$ calculations with $K \geq k$. The orders are decided by counting the powers of some small parameter, and since ChPT is a low-energy theory, the relevant small parameter is the amount of energy involved, which we will express in terms of p^2, the square of some typical momentum. As can be verified through explicit calculations, LO is at $\mathcal{O}(p^2)$, NLO at $\mathcal{O}(p^4)$, NNLO at $\mathcal{O}(p^6)$, and so on, and one can also verify that[67]

$$
u_\mu = \mathcal{O}(p) \,, \quad \nabla_\mu = \mathcal{O}(p) \,, \quad \chi_\pm = \mathcal{O}(p^2) \,, \quad f_\pm^{\mu\nu} = \mathcal{O}(p^2) \,.
\tag{3.25}
$$

To build the $\mathrm{N}^k\mathrm{LO}$ Lagrangian, one simply has to take enough building blocks that they add up to $\mathcal{O}(p^{2k})$, assemble them with traces and index contractions in all possible ways, and (less simply) find all possible relations between them so that only the minimal number of terms is retained. The general procedure is described in ref. [35].

The Lagrangian. The LO Lagrangian is very simple:

$$
\mathcal{L}_{\mathrm{LO}} = \frac{F^2}{4}\langle u_\mu u^\mu + \chi_+ \rangle \,,
\tag{3.26}
$$

---

[67]This assignment is not entirely unique, and one can have consistent power counting with, e.g., $\chi_\pm = \mathcal{O}(p)$.

where $u_\mu u^\mu$ is equivalent to the $\partial_\mu U^\dagger \partial^\mu U$ term in eq. (3.16). The LEC of the χ_+ term, B, is conventionally absorbed into the definition of χ, so it is not visible here.

The NLO Lagrangian, due to Gasser & Leutwyler, is

$$\mathcal{L}_{\text{NLO}} = \hat{L}_0 \langle u_\mu u_\nu u^\mu u^\nu \rangle + \hat{L}_1 \langle u_\mu u^\mu \rangle \langle u_\nu u^\nu \rangle + \hat{L}_2 \langle u_\mu u_\nu \rangle \langle u^\mu u^\nu \rangle + \hat{L}_3 \langle u_\mu u^\mu u_\nu u^\nu \rangle$$
$$+ \hat{L}_4 \langle u_\mu u^\mu \rangle \langle \chi_+ \rangle + \hat{L}_5 \langle u_\mu u^\mu \chi_+ \rangle + \hat{L}_6 \langle \chi_+ \rangle \langle \chi_+ \rangle + \hat{L}_7 \langle \chi_- \rangle \langle \chi_- \rangle$$
$$+ \hat{L}_8 \langle \chi_+^2 + \chi_-^2 \rangle - i\hat{L}_9 \langle u_\mu u_\nu f_+^{\mu\nu} \rangle + \hat{L}_{10} \langle f_+^{\mu\nu} f_{\mu\nu}^+ - f_-^{\mu\nu} f_{\mu\nu}^- \rangle, \quad (3.27)$$

where the first four terms are recognizable from eq. (3.16). In addition to this, one can write down a few more terms, which (after expansion) only involve the external fields; the two such *contact terms* at NLO are

$$H_1 \langle \chi_+^2 - \chi_-^2 \rangle + H_2 \langle f_+^{\mu\nu} f_{\mu\nu}^+ + f_-^{\mu\nu} f_{\mu\nu}^- \rangle. \quad (3.28)$$

These do not contribute to physical processes, but are nevertheless important in certain analyses beyond the scope of this thesis.

The NNLO Lagrangian was derived by Bijnens, Colangelo & Ecker [36], refining earlier work by Fearing & Scherer [37]. Its general form consists of 112 terms and 3 contact terms, each with an LEC K_i, which is too much to show here.[68] However, to get a feel for the Lagrangian, I will present the subset of it that is relevant for the analyses of paper II:[69]

$$\mathcal{L}_{\text{NNLO}} \supset 4K_1 \langle u_\mu u^\mu \nabla_\nu u_\rho \nabla^\nu u^\rho \rangle + 4K_2 \langle u_\mu u^\mu \rangle \langle \nabla_\nu u_\rho \nabla^\nu u^\rho \rangle + 4K_3 \langle \nabla_\mu u_\nu u_\rho \nabla^\mu u^\nu u^\rho \rangle$$
$$+ 4K_4 \langle \nabla_\mu u_\nu u_\rho \rangle \langle \nabla^\mu u^\nu u^\rho \rangle + 8K_5 \langle \nabla_\mu u_\nu u_\rho \nabla^\mu u^\rho u^\nu \rangle + 4K_6 \langle \nabla_\mu u_\nu u_\rho \rangle \langle \nabla^\mu u^\rho u^\nu \rangle$$
$$+ K_7 \langle u_\mu u^\mu u_\nu u^\nu \chi_+ \rangle + K_8 \langle u_\mu u^\mu u_\nu u^\nu \rangle \langle \chi_+ \rangle + K_9 \langle u_\mu u^\mu \rangle \langle u_\nu u^\nu \chi_+ \rangle$$
$$+ K_{11} \langle u_\mu u_\nu u^\nu u^\mu \chi_+ \rangle + K_{13} \langle u_\mu u_\nu u^\mu u^\nu \chi_+ \rangle + K_{14} \langle u_\mu u_\nu u^\mu u^\nu \rangle \langle \chi_+ \rangle$$
$$+ K_{16} \langle \chi_+ \rangle \langle u_\mu u_\nu \rangle \langle u^\mu u^\nu \rangle + 4K_{17} \langle \chi_+ \nabla_\mu u_\nu \nabla^\mu u^\nu \rangle + 4K_{18} \langle \chi_+ \rangle \langle \nabla_\mu u_\nu \nabla^\mu u^\nu \rangle$$
$$+ K_{19} \langle u_\mu u^\mu \chi_+^2 \rangle + K_{20} \langle u_\mu u^\mu \chi_+ \rangle \langle \chi_+ \rangle + K_{21} \langle u_\mu u^\mu \rangle \langle \chi_+^2 \rangle + K_{22} \langle u_\mu u^\mu \rangle \langle \chi_+ \rangle^2$$
$$+ K_{23} \langle \chi_+ u_\mu \chi_+ u^\mu \rangle + iK_{28} \big[\langle \chi_- \nabla_\mu u_\nu u^\mu u^\nu \rangle + \langle \chi_- u_\mu u_\nu \nabla^\mu u^\nu \rangle \big]$$
$$+ iK_{29} \langle \chi_- \nabla_\mu u_\nu \rangle \langle u^\mu u^\nu \rangle + iK_{31} \langle \chi_- u_\mu \nabla_\nu u^\mu u^\nu \rangle + iK_{32} \langle \chi_- u_\mu \rangle \langle \nabla_\nu u^\mu u^\nu \rangle$$
$$+ K_{33} \langle u_\mu u^\mu \chi_-^2 \rangle + K_{35} \langle u_\mu u^\mu \rangle \langle \chi_-^2 \rangle + K_{37} \langle \chi_- u_\mu \chi_- u^\mu \rangle + K_{38} \langle \chi_- u_\mu \rangle \langle \chi_- u^\mu \rangle. \quad (3.29)$$

Here and below, I abuse the $\supset$ (superset) symbol as "contains these and other terms".

The N³LO Lagrangian, derived by Bijnens, Hermansson-Truedsson & Wang [35], has 1840 terms and 21 contact terms. The number of terms (but not their explicit form) is known to much higher order using Hilbert series [38], although phenomenological relevance decreases sharply beyond NNLO.

---

[68]Note that by "term", I mean the expression that multiplies exactly one LEC, even if it, like the $\hat{L}_8$ term in eq. (3.27), strictly speaking consists of more than one term.

[69]In ref. [36], terms containing $\nabla_\mu u_\nu$ are instead formulated in terms of $h_{\mu\nu} \equiv \nabla_\mu u_\nu + \nabla_\nu u_\mu$. Using the identity $f_{\mu\nu}^- = \nabla_\mu u_\nu - \nabla_\nu u_\mu$, easily verifiable from eqs. (3.23) and (3.24), I have reformulated these in terms of $\nabla_\mu u_\nu$, leaving behind some factors of 2 and contributions to terms containing $f_{\mu\nu}^-$, which are not shown here. In the case of K_5, there is an additional factor of 2 since its original form consisted of two terms which are equivalent after applying the identity.

The Cayley–Hamilton theorem [see eqs. (3.43) and (3.47)] reduces the number of terms in the $n_f = 2$ and $n_f = 3$ Lagrangians. The reduced $\mathcal{L}_{\text{NLO}}$ can be found in ref. [39], $\mathcal{L}_{\text{NNLO}}$ in ref. [36], and $\mathcal{L}_{\text{N}^3\text{LO}}$ in ref. [35]; $\mathcal{L}_{\text{LO}}$ does not reduce. The LECs are conventionally given different symbols: $\hat{L}_i$ are replaced by L_i at $n_f = 3$ and by ℓ_i (or l_i) at $n_f = 2$. Paper I uses a different, nonstandard notation for the LECs.

Lastly, the Lagrangian contains the so-called *anomalous sector*, which I mention for completeness but do not use. There is a single $\mathcal{O}(p^4)$ anomalous term, the Wess–Zumino–Witten (WZW) term [40, 41], which can be derived geometrically and whose LEC is fixed. It is not invariant under the chiral symmetry, but is nevertheless permitted in the Lagrangian since its change under chiral transformations has the form of a total derivative. It mediates π^0 decay and other processes that are contingent on the axial anomaly.[70] The $\mathcal{O}(p^6)$ [42, 43] and $\mathcal{O}(p^8)$ [44] anomalous Lagrangians are also known; these are chirally invariant and not directly connected to the axial anomaly, but they are so called because they, like the WZW term but unlike the "regular" Lagrangian, contain $\epsilon_{\mu\nu\alpha\beta}$.

3.4 Lie algebra for ChPT

The power of the SUN, in the palm of my hand! — *Hugo Serôdio (2019)*

Practical ChPT calculations in this thesis involve a lot of $\mathrm{SU}(n_f)$ Lie algebra, i.e., manipulations of the generators t^a that appear in eq. (3.15). This algebra is usually introduced on the fly in introductory ChPT texts, and comprehensive Lie algebra texts are not written with ChPT in mind. Here, I will therefore afford the luxury of a comfprehensive, ChPT-tailored derivation of the necessary algebra once and for all.

Definitions. The starting point is $U \equiv e^{i\alpha} \in \mathrm{SU}(n_f)$ for some $n_f \times n_f$ matrix α. As noted before, the defining $\mathrm{SU}(n_f)$ properties

$$1 = U^\dagger U = e^{i[\alpha - \alpha^\dagger]} \qquad \text{and} \qquad 1 = \det(U) = e^{i\langle \alpha \rangle} \qquad (3.30)$$

imply that α is traceless and Hermitian.[71] Conversely, every traceless and Hermitian $n \times n$ matrix generates an element of $\mathrm{SU}(n)$.

Now, let $\{t^a\}$ be a basis for the space of traceless Hermitian matrices; with complex coefficients, this is a basis for the space of *all* traceless matrices. These t^a are the *generators* of the $\mathrm{SU}(n_f)$ Lie algebra, and we write $\alpha = \alpha_a t^a$, etc.

The t^a must be traceless, but not necessarily Hermitian. In fact, the Hermitian matrix $\phi_a t^a$ in eq. (3.15) must use non-Hermitian t^a if ϕ_a are to represent the mesons, since some of

---

[70]See Donoghue [18, chapter VII-5 & 6] or Pich [16, sec. 7] for a detailed description of the WZW term and its use in calculating anomalous processes.

[71]As introduced on page 21, $\langle \alpha \rangle$ denotes the trace of α. Strictly speaking, $\langle \alpha \rangle$ and $\alpha - \alpha^\dagger$ could be allowed to be integer multiples of 2π, but unlike the traceless Hermitian matrices, matrices with these more lenient properties do not form a vector space (simply consider the effect of the scalar $\frac{1}{2}$) and are therefore worthless for all algebraic purposes.

these are complex-valued fields [e.g., $(\pi^+)^\dagger = \pi^-$]. I will therefore deviate from convention and assume t^a are non-Hermitian, and adopt the upper-lower index distinction

$$(t^a)^\dagger \equiv t_a\,, \qquad (t_a)^\dagger \equiv t^a\,. \tag{3.31}$$

As we will see, all algebraic properties of Hermitian t^a carry over to the non-Hermitian case if one enforces the Lorentz-like rule that all contractions involve one upper and one lower index. For Hermitian generators, this can of course be ignored.

The generators can be chosen to be orthonormal under the trace-product

$$\langle t^a t_b \rangle = \tau \delta^a_b\,, \tag{3.32}$$

where τ is a normalization factor (see below). This is what necessitates the upper-lower convention, since δ^a_b is real. The symbols $\delta^{ab} \equiv \frac{1}{\tau}\langle t^a t^b\rangle$ and $\delta_{ab} \equiv \frac{1}{\tau}\langle t_a t_b\rangle$, which may be neither real nor diagonal in a and b, act as a metric for raising and lowering indices.

(Anti)commutation. To see that the generators have well-defined commutation relations, note that with $e^{i\alpha}, e^{i\beta} \in \mathrm{SU}(n_f)$, closedness requires that $e^{i\alpha}e^{i\beta} = e^{i\gamma} \in \mathrm{SU}(n_f)$. The Baker–Campbell–Hausdorff formula then states that[72]

$$\gamma = \alpha + \beta + \tfrac{1}{2}[\alpha, \beta] + \tfrac{1}{12}\big[\alpha - \beta, [\alpha, \beta]\big] + \ldots\,, \tag{3.33}$$

where "$\ldots$" consists of more nested commutators of α and β. Since γ is traceless and Hermitian, $[\alpha, \beta]$ must be traceless and Hermitian too, which makes all the nested commutators traceless and Hermitian as well.[73] This implies that the commutator of two generators is a linear combination of the generators, i.e.,

$$[t^a, t^b] = i\kappa f^{abc} t_c\,, \tag{3.34}$$

where the coefficients f^{abc} are called *structure constants*, and κ is another normalization factor. The f^{abc} are totally antisymmetric in abc, as follows from

$$i\kappa\tau f^{abc} = \langle [t^a, t^b]\, t^c\rangle = \langle [t^b, t^c]\, t^a\rangle = \langle [t^c, t^a]\, t^b\rangle\,, \tag{3.35}$$

and in the case of Hermitian generators (but not in general), they are real. The structure constants also obey the *Jacobi identity*,

$$\big[f^{abe}f^{cde'} + f^{bce}f^{ade'} + f^{cae}f^{bde'}\big]\delta_{ee'} = 0\,, \tag{3.36}$$

which follows from the fact that

$$\big[[t^a, t^b], t^c\big] + \big[[t^b, t^c], t^a\big] + \big[[t^b, t^c], t^a\big] = \sum_{\mathcal{S}(abc)} (t^a t^b t^c - t^a t^b t^c) = 0\,, \tag{3.37}$$

where $\mathcal{S}(abc)$ symbolizes all permutations of abc. The Jacobi identity is a defining feature of a Lie algebra, so the above essentially shows that what we are working with is, indeed, a Lie algebra.

---

[72]A short, accessible proof is given by Eichler [45]. See also ref. [46] for the historical development of this formula, which involved many people beyond its three namesakes.

[73]To see that there are no subtle cancellations spoiling this argument, note that the terms scale differently under the operation $\alpha \to x\alpha$, $\beta \to y\beta$.

Unlike general Lie algebras, $SU(n_f)$ also has a well-defined anticommutation relation: since $\langle\{t^a, t^b\}\rangle = 2\langle t^a t^b\rangle = 2\tau\delta^{ab}$, the anticommutator must be proportional to δ^{ab} plus a traceless matrix. Therefore,

$$\{t^a, t^b\} = \tfrac{2\tau}{n_f}\delta^{ab} + \kappa d^{abc}t_c \,, \tag{3.38}$$

where d^{abc} are similar to the structure constants except that they are totally symmetric in abc and do not satisfy the Jacobi identity.

Normalization. There are multiple normalization conventions in common use. I generally follow Hans Bijnens' convention $(\tau, \kappa) = (1, 1)$, although some of my sources (e.g., ref. [47]) use $(\tau, \kappa) = (1, \sqrt{2})$. Zee [15] also uses $(\tau, \kappa) = (1, 1)$, but Peskin & Schroeder [14] and Donoghue [18] use $(\tau, \kappa) = (\tfrac{1}{2}, 1)$, and the Pauli and Gell-Mann matrices, eqs. (3.50) and (3.51), conventionally have $(\tau, \kappa) = (2, 2)$. I keep the normalization generic throughout this introduction for easier comparison, but be aware that the papers themselves use $(\tau, \kappa) = (1, 1)$ and assume Hermitian generators.

The Fierz identity. Another consequence of the completeness and trace-orthogonality of $\{t^a\}$ is that *any* matrix M can be written as

$$M = M_a t^a + M_0 \,, \qquad \text{where} \quad M_a \equiv \tfrac{1}{\tau}\langle M t_a\rangle \quad \text{and} \quad M_0 \equiv \tfrac{1}{n}\langle M\rangle \tag{3.39}$$

Writing this with explicit matrix indices,

$$M_{ij} = \tfrac{1}{\tau}M_{\ell k}(t_a)_{k\ell}(t^a)_{ij} + \tfrac{1}{n}M_{kk}\delta_{ij}$$
$$\Rightarrow \quad 0 = M_{\ell k}\left[\tfrac{1}{\tau}(t_a)_{k\ell}(t^a)_{ij} + \tfrac{1}{n}\delta_{\ell k}\delta_{ij} - \delta_{i\ell}\delta_{jk}\right]. \tag{3.40}$$

Since M_{ij} is arbitrary, the quantity in square brackets must be zero in general, so

$$\tfrac{1}{\tau}(t_a)_{k\ell}(t^a)_{ij} = \delta_{i\ell}\delta_{jk} - \tfrac{1}{n}\delta_{ij}\delta_{k\ell} \,, \tag{3.41}$$

which is the *Fierz identity*.[74] Inserted into traces along with arbitrary matrices X and Y,

$$\tfrac{1}{\tau}\langle X t_a\rangle\langle Y t^a\rangle = \langle XY\rangle - \tfrac{1}{n}\langle X\rangle\langle Y\rangle \,, \qquad \tfrac{1}{\tau}\langle X t_a Y t^a\rangle = \langle X\rangle\langle Y\rangle - \tfrac{1}{n}\langle XY\rangle \,, \tag{3.42}$$

it serves as a cornerstone of section 4.4, and of papers I and III.

The Cayley–Hamilton theorem. At fixed n_f, further identities follow from the *Cayley–Hamilton theorem*, which states that the characteristic polynomial $p_M(\lambda) \equiv \det(\lambda - M)$, for any $n_f \times n_f$ matrix M, is solved by M itself: $p_M(M) = 0$.[75] The coefficients of $p_M(\lambda)$ can in general be written in terms of traces of powers of M. For $n_f = 1$, we get the trivial statement $0 = M - \langle M\rangle$, and for $n_f \geq 4$, the theorem is impractical, so let us turn our attention to $n_f = 2$ or 3.

With M a traceless 2×2 matrix, the theorem gives

$$0 = M^2 - \tfrac{1}{2}\langle M^2\rangle \,. \tag{3.43}$$

---

[74]Markus Fierz' original work [26] concerned spinor bilinears (recall page 35), but his name has since been applied to a wide variety of analogous relations such as eq. (3.42).

[75]Note that this holds when M is substituted for λ in the polynomial $p_M(\lambda)$, not in the determinant $\det(\lambda - M)$, which would be a trivial result. Cayley and Hamilton only proved special cases, with the general statement being due to Frobenius [48].

Introducing $M = A + B$ with A and B traceless, we get

$$0 = A^2 + \{A, B\} + B^2 - \tfrac{1}{2}\langle A^2 + 2AB + B^2\rangle = \{A, B\} - \langle AB\rangle, \qquad (3.44)$$

where we invoked eq. (3.43) on A and B to get rid of the quadratic terms in the last step. This identity allows for great simplification of the $n_f = 2$ ChPT Lagrangian, and by setting $A = t^a, B = t^b$ and comparing to eq. (3.38), it implies that $d^{abc} = 0$.

With M a traceless 3×3 matrix, the theorem gives

$$0 = M^3 - \tfrac{1}{2}M\langle M^2\rangle - \tfrac{1}{3}\langle M^3\rangle. \qquad (3.45)$$

Introducing $M = A + B + C$, all traceless, and applying eq. (3.45) repeatedly, we are left with

$$0 = \sum_{\mathcal{S}(ABC)} ABC - \sum_{\mathbb{Z}(ABC)} A\langle BC\rangle - \langle A\{B, C\}\rangle, \qquad (3.46)$$

where $\mathcal{S}$ and $\mathbb{Z}$ symbolize permutations and cyclic permutations, respectively. Tracing eq. (3.46) with a fourth traceless matrix D gives

$$\sum_{\mathcal{S}(ABC)} \langle ABCD\rangle = \sum_{\mathbb{Z}(ABC)} \langle AB\rangle\langle CD\rangle. \qquad (3.47)$$

While not as powerful as its $n_f = 2$ counterpart, it allows for simplification of $n_f = 3$ ChPT. Furthermore, combining eq. (3.47) with

$$\{\{t^a, t^b\}, t^c\} + \{\{t^b, t^c\}, t^a\} + \{\{t^b, t^c\}, t^a\} = 2 \sum_{\mathcal{S}(abc)} t^a t^b t^c, \qquad (3.48)$$

[compare eq. (3.37)] and applying eq. (3.38) yields the useful, Jacobi-like identity

$$\left[d^{abe}d^{cde'} + d^{bce}d^{ade'} + d^{cae}d^{bde'}\right]\delta_{ee'} = \frac{2\tau}{3\kappa^2}\left(\delta^{ab}\delta^{cd} + \delta^{bc}\delta^{ad} + \delta^{ca}\delta^{bc}\right). \qquad (3.49)$$

Explicit matrices. At $n_f = 2$, the canonical set of Hermitian generators is the Pauli matrices $\{\sigma^a\}$,

$$\sigma^1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad (3.50)$$

for which $d^{abc} = 0$ and $f^{abc} = \epsilon^{abc}$, assuming $(\tau, \kappa) = (2, 2)$. At $n_f = 3$, with the same normalization, the canonical choice is the *Gell-Mann matrices* $\{\lambda^a\}$,

$$\lambda^{1,2,3} \equiv \begin{pmatrix} \sigma^{1,2,3} & 0 \\ 0 & 0 \end{pmatrix}, \qquad \lambda^4 \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda^5 \equiv \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

$$\lambda^6 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \lambda^7 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \lambda^8 \equiv \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \qquad (3.51)$$

Note how the matrices form what I call *off-diagonal pairs* [e.g., (σ^1, σ^2) or (λ^4, λ^5)], consisting of one real symmetric and one imaginary antisymmetric matrix with nonzero elements in the same position.

The pattern for arbitrary n_f is simple: when going to $n_f + 1$, keep all previous matrices, padded with an extra row and column of zeroes, then add the n_f off-diagonal pairs with elements in that row and column, and finally add the unique (up to a sign) traceless diagonal matrix that is orthogonal to all previous ones.

A particularly useful set of non-Hermitian generators, which I here denote with the addition of a tilde, is obtained by combining the off-diagonal pairs so that all generators are real. This generalizes the common combination $\sigma^\pm \equiv \frac{1}{2}(\sigma^1 \pm i\sigma^2)$: $\lambda^{1,2}$ are replaced by $\tilde{\lambda}^1 = \tilde{\lambda}_2 = (\lambda^1 + i\lambda^2)/2$, etc., yielding

$$\tilde{\lambda}^1 = \tilde{\lambda}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tilde{\lambda}^4 = \tilde{\lambda}_5 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tilde{\lambda}^6 = \tilde{\lambda}_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.52)$$

[recall eq. (3.31)]. The most natural normalization for these is $\tau = 1$. The diagonal generators remain the same up to normalization: $\tilde{\lambda}^3 = \tilde{\lambda}_3 = \lambda^3/\sqrt{2}$, etc.

Each off-diagonal $\tilde{\lambda}^a$ commutes with all other off-diagonal $\tilde{\lambda}^b$ except its transpose, $\tilde{\lambda}_a$. They are also behave more simply when combined inside traces with the diagonal generators, as we shall see below. These properties hold also for $\tilde{t}^a$ at general n_f.

3.5 Masses and phenomenology

Having established ChPT on theoretical grounds, let me summarize the most salient parts of its connection to real-world phenomena.

Looking back at eq. (3.18), quark masses are introduced by adding $M \equiv \mathrm{diag}(m_u, m_d, m_s, \ldots)$ to s. Then, via eqs. (3.15), (3.19), (3.23) and (3.26), the leading mass terms for the NGBs come from

$$\mathcal{L} \supset \frac{F^2}{4}\langle\chi_+\rangle = \frac{BF^2}{2}\langle M(U + U^\dagger)\rangle = BF^2\langle M\rangle - \frac{B}{2}\phi_a\langle M\{t^a, t_b\}\rangle\phi^b + \mathcal{O}(\phi^4). \quad (3.53)$$

Only $-B\phi^a\phi^b\langle Mt^at^b\rangle$ is of interest for leading-order masses. Likewise, charges are introduced by adding $-eA_\mu Q$ to v_μ, where A_μ is the photon field and $Q \equiv \mathrm{diag}(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}, \ldots)$, so the leading electromagnetic interaction is[76]

$$\mathcal{L} \supset \frac{F^2}{4}\langle u_\mu u^\mu\rangle = \phi_a\left\{ieA^\mu\langle Q[t^a, t_b]\rangle\partial_\mu - \frac{e^2}{2}A^\mu A_\mu\langle[Q, t^a][Q, t_b]\rangle\right\}\phi^b + \mathcal{O}(\phi^4). \quad (3.54)$$

Ideally, the mass and electromagnetic terms should simplify to $\phi_a(M_b^a + Q_b^a)\phi^b$, where M_b^a and Q_b^a are diagonal matrices so that the NGBs are mass and charge eigenstates. This is not

---

[76]Note the similarity to scalar QED, which provides the simplest gauge-invariant coupling between A_μ and a complex scalar field: $\mathcal{L}_{\text{sQED}} \supset -ieA_\mu(\phi\partial_\mu\phi^* - \phi^*\partial_\mu\phi) + e^2A_\mu\phi A^\mu\phi^*$.

the case if t^a are Pauli or Gell-Mann matrices: $\langle Q[\sigma^1, \sigma^2]\rangle$ and $\langle M\{\lambda^3, \lambda^8\}\rangle$ are nonzero, for instance.

The non-Hermitian generators $\tilde{\lambda}^a$ introduced in the previous section mostly solve this. Due to the properties discussed at the end of that section, all traces in eqs. (3.53) and (3.54) vanish when $\tilde{\lambda}^a$ or $\tilde{\lambda}_b$ is off-diagonal, except when $a = b$. Therefore, for $n_f = 3$ we can identify three particle-antiparticle pairs associated with these generators, and from the charges and the quark content implied by the masses, the identification with real-world mesons is obvious:

$$\pi^{\pm} = \frac{\phi_1 \mp i\phi_2}{\sqrt{2}}, \qquad M_{\pi^{\pm}}^2 = B(m_u + m_d),$$

$$K^{\pm} = \frac{\phi_4 \mp i\phi_5}{\sqrt{2}}, \qquad M_{K^{\pm}}^2 = B(m_u + m_s), \qquad (3.55)$$

$$K^0, \bar{K}^0 = \frac{\phi_6 \mp i\phi_7}{\sqrt{2}}, \qquad M_{K^0}^2 = B(m_d + m_s).$$

Here and below, ϕ_a are the real non-eigenstate fields used when writing $U = \exp(\phi_a \lambda^a / F)$ in terms of Gell-Mann matrices. Equation (3.55) reproduces observed meson mass ratios well, implying the quark mass ratios $m_u : m_d : m_s = 0.55 : 1 : 20.3$. Note how the strange quark is much heavier than the others, as expected.

When $n_f = 2$, ϕ_3 is a zero-charge mass eigenstate with the same mass as $\pi^{\pm}$, so it is natural to identify it with π^0. When $n_f \geq 3$, however, $\langle M\{t^a, t_b\}\rangle$ mixes all fields associated with diagonal generators. Unlike the off-diagonal ones, there is no "clean" change of basis that resolves it; instead, some mixing angle must be introduced. With the choice that π^0 is "mostly" ϕ_3 and η is "mostly" ϕ_8,[77] the $n_f = 3$ mixing is

$$\pi^0 = \phi^3 \cos\theta_{\pi\eta} - \phi^8 \sin\theta_{\pi\eta}, \qquad \eta = \phi^3 \sin\theta_{\pi\eta} + \phi^8 \cos\theta_{\pi\eta}. \qquad (3.56)$$

The charge terms trivially vanish for diagonal generators, so the small angle $\theta_{\pi\eta}$ is, to leading order, determined entirely by the masses. With the aforementioned quark mass ratio, the mass matrix is diagonalized by

$$\tfrac{1}{2}\tan(2\theta_{\pi\eta}) = \frac{\sqrt{3}}{2}\frac{m_u - m_d}{2m_s - (m_u + m_d)} \approx 0.01 \approx \theta_{\pi\eta}, \qquad (3.57)$$

and the eigenstates obtain the masses

$$M_{\pi^0}^2 = B(m_u + m_d) - \delta_{\pi\eta} + \mathcal{O}(\delta_{\pi\eta}^2), \qquad M_\eta^2 = \tfrac{1}{3}B(m_u + m_d + 4m_s) + \delta_{\pi\eta} + \mathcal{O}(\delta_{\pi\eta}^2), \quad (3.58)$$

where

$$\delta_{\pi\eta} \equiv \frac{B}{2}\frac{(m_u - m_d)^2}{2m_s - (m_u + m_d)} \approx 0.0016 M_{\pi^{\pm}}^2. \qquad (3.59)$$

Again, this gives a fairly accurate value for the masses, up to percent-level discrepancies; for instance, $\delta_{\pi\eta}$ is smaller than the actual $\pi^{\pm}$-π^0 mass-squared difference of about $0.065 M_{\pi^{\pm}}^2$. There are higher-order corrections to the masses, which also involve LECs such as $\hat{L}_6$,

---

[77]This also motivates the approximate quark content in table 2, which corresponds to the diagonal elements of λ^3 and λ^8.

$\hat{L}_7$ and $\hat{L}_8$, as well as electromagnetic and other effects; these allow more exact matching between ChPT and measured masses (see ref. [39] and references therein).

Condensates and decays. At this point, it is relatively straightforward to connect B, and therefore the masses, to the quark condensate, whose appearance is what breaks chiral symmetry in the first place. Formulating QCD using eq. (3.18), the exact path-integral treatment [recall eq. (2.36)] gives

$$\langle 0|\bar{q}q|0\rangle = \frac{-i\delta \ln Z[s, p, \ldots]}{\delta s} , \tag{3.60}$$

restoring $s, p, \ldots$ to their physical values after taking the derivative. Assuming that ChPT models the same low-energy physics, this relation must remain if we replace the QCD Lagrangian by the ChPT one. Then the first term in eq. (3.53) simply gives

$$\langle 0|\bar{q}q|0\rangle = BF^2 + [\text{corrections}] . \tag{3.61}$$

A similar argument applies to the pion decay rate, $\langle 0|J_L^\mu|\pi\rangle$, where the weak current J_L^μ mediating the decay is obtained with $-i\delta \ln Z/\delta\ell^\mu$. Such a derivative applied to $\langle u_\mu u^\mu\rangle$ shows that F is precisely the leading-order decay constant:

$$\langle 0|J_L^\mu|\pi\rangle = k^\mu F + [\text{corrections}] , \tag{3.62}$$

where k^μ is the momentum of the pion.[78] The NLO corrections to F and the pion mass can be seen in eq. (III.2.17), and are known up to N^{3}LO [23].

These results finally put to rest the questions about light pseudoscalar mesons discussed in section 2. If chiral symmetry were exact, the ϕ^a would be NGBs, and so these mesons would be massless. Since chiral symmetry is almost exact for $n_f = 2$, the pions are in some sense close to the NGBs they would be in a chirally symmetric world, and are therefore rather close to massless; *how* close is parametrized within ChPT via the fields s, p, etc. The same applies to the kaons and η, but due to the higher mass of the strange quark, they are less close to being NGBs and are correspondingly heavier. The other mesons have no relation to any NGBs, and therefore obtain their masses on the same conditions as the baryons, which do not rely on chiral symmetry breaking to obtain mass.[79] In particular, η' is non-light due to the axial anomaly [recall section 2.4].[80]

---

[78]The proportionality to k^μ follows from Lorentz invariance; the only thing that needs showing is that the proportionality constant is F.

[79]This is made clearer by the extensions of ChPT that include baryons, which are briefly reviewed in section 3.7. See also ref. [49] and references therein for lattice calculations that explicitly demonstrate the properties of baryons in worlds with unbroken chiral symmetry.

[80]Without the anomaly, chiral symmetry would be $U(n_f)_L \times U(n_f)_R$ and would break to $U(n_f)_V$. Then η' would be a ninth pseudo-NGB corresponding to the unit matrix, which along with the $SU(n_f)$ generators generate $U(n_f)$. Following the calculation outlined in section 3.5, it would have the pion-like mass $M_{\eta'}^2 = B(m_u + m_d) + \delta_{\pi\eta} + \mathcal{O}(\delta_{\pi\eta}^2)$. It would exist also at $n_f = 2$, serving as the isospin-0 state, as opposed to π^0 which is the $I_3 = 0$ component of an isospin-1 triplet. [η appears first at $n_f = 3$; the doublet (π^0, η) and singlet η' then form an arrangement under flavor rotations that is analogous to the $I = 1$ states described in eq. (4.24) and in paper V.]

3.6 The low-energy constants

Absolute truths are expressed by dimensionless numbers.

— Bo "Bosse" Söderberg (2016)

Renormalization. ChPT is, strictly speaking, not a renormalizable theory, in which the renormalization of a finite number of parameters is enough to cancel loop divergences to all orders. It is, however, order-by-order renormalizable, where the new LECs entering at each order cancel any additional divergences and leave renormalized LECs that are physically observable and different from the "bare" LECs present in the Lagrangian. The LO LECs F and B likewise get replaced in physical results by the mass and decay constants F_π, M_π. See ref. [50, 51] and references therein for more information on ChPT renormalization, and page 59 for a more explicit look at loop divergences.

Using *dimensional regularization*, divergences are captured by the $d \to 4$ singularity of

$$\Lambda \equiv \frac{\kappa}{d-4}\,, \qquad \text{where} \quad \kappa \equiv \frac{1}{16\pi^2}\,, \tag{3.63}$$

so the NLO and NNLO renormalization is [30, 31, 51][81]

$$L_i = (c\mu)^{d-4}\left(L_i^{\mathrm{r}}(\mu,d) + \hat{\Gamma}_i\Lambda\right), \tag{3.64}$$

$$K_i = \frac{(c\mu)^{2(d-4)}}{F^2}\left(K_i^{\mathrm{r}}(\mu,d) - \hat{\Gamma}_i^{(2)}\Lambda^2 - \left[\hat{\Gamma}_i^{(1)} + \hat{\Gamma}_i^{(L)}(\mu,d)\right]\Lambda\right), \tag{3.65}$$

where μ is the renormalization scale, conventionally set to 770 MeV $\approx M_\rho$; $L_i^{\mathrm{r}} \equiv L_i^{\mathrm{r}}(\mu,4)$ and $K_i^{\mathrm{r}} \equiv K_i^{\mathrm{r}}(\mu,4)$ are renormalized, physically observable LECs; $\hat{\Gamma}_i^X$ are coefficients listed in ref. [51, eq. (3.14) & tables 3-4]; and c specifies the renormalization scheme: following Gasser & Leutwyler's original work [30], the ChPT-specific convention is

$$\ln c = -\tfrac{1}{2}\left(\ln 4\pi + \gamma_E + 1\right), \tag{3.66}$$

where γ_E is the Euler–Mascherino constant; this differs from the standard $\overline{\mathrm{MS}}$ scheme by the extra $+1$. With $\epsilon \equiv 2 - d/2$ and $1/\tilde{\epsilon} \equiv 1/\epsilon - 2\ln(c\mu)$, eq. (3.64) becomes

$$\hat{L}_i^{\mathrm{r}} = \hat{L}_i + \frac{\kappa\hat{\Gamma}_i}{2\tilde{\epsilon}}\,. \tag{3.67}$$

The $n_f = 2, 3$ LECs left by the Cayley–Hamilton theorem are renormalized similarly. The counterparts of $\hat{\Gamma}_i$ are called Γ_i for $n_f = 3$ and are listed in eq. (III.2.13); at $n_f = 2$ they are called γ_i and can be found in eq. (IV.3.5) or ref. [30, eq. (9.2)]. All NNLO coefficients are listed in ref. [51, tables 1-2].

Different renormalization scales μ, μ' are related by (at NLO)

$$\hat{L}_i^{\mathrm{r}}(\mu') = \hat{L}_i^{\mathrm{r}}(\mu) + \frac{\kappa\hat{\Gamma}_i}{2}\log\frac{\mu^2}{\mu'^2}\,, \tag{3.68}$$

---

[81]The analogous renormalization of the N^{3}LO LECs [35] has not yet been carried out.

and by convention, the $n_f = 2$ LECs are re-expressed as scale-independent $\bar{\ell}_i$ by fixing $\mu = M_\pi$ through

$$\ell_i^{\mathrm{r}}(\mu) = \tfrac{1}{2}\gamma_i\big(\kappa\bar{\ell}_i + L\big), \tag{3.69}$$

where $L \equiv \kappa \ln(M_\pi^2/\mu^2)$.

LEC values. The most recent comprehensive review of the LECs is that of Bijnens & Ecker [39], and I use values drawn from there whenever needed. They adopt

$$F_\pi = 92.2(1) \text{ MeV} \tag{3.70}$$

and the two-flavor LECs

$$\begin{aligned}
&\bar{l}_1 = -0.4(6)\,, &&\bar{l}_2 = 4.3(1)\,, &&\bar{l}_3 = 2.9(2\,4)\,, \\
&\bar{l}_4 = 4.4(2)\,, &&\bar{l}_5 = 12.24(21)\,, &&\bar{l}_6 = 15.24(39)\,.
\end{aligned} \tag{3.71}$$

Note that the uncertainties range from $\mathcal{O}(1\%)$ to $\mathcal{O}(100\%)$; the less uncertain $\bar{l}_{5,6}$ can be measured through electromagnetic interactions, whereas the others govern purely hadronic processes.

The thee-flavor case is more complicated, and ref. [39] presents several different fits. The one used in paper II gives

$$\begin{aligned}
&10^3 L_1^{\mathrm{r}} = 1.11(10)\,, &&10^3 L_2^{\mathrm{r}} = 1.05(17)\,, &&10^3 L_3^{\mathrm{r}} = -3.82(30)\,, &&10^3 L_4^{\mathrm{r}} = 1.87(53)\,, \\
&10^3 L_5^{\mathrm{r}} = 1.22(6)\,, &&10^3 L_6^{\mathrm{r}} = 1.46(46)\,, &&10^3 L_7^{\mathrm{r}} = -0.39(8)\,, &&10^3 L_8^{\mathrm{r}} = 0.65(7)\,.
\end{aligned} \tag{3.72}$$

Reference [39] also presents estimates of the NNLO LECs; these are used in paper II, but I do not reproduce them here.

3.7 Limits, limitations and extensions of ChPT

Limits. ChPT calculations are very difficult unless some simplifications are made. Common ones include the *chiral limit*, where $s = p = v_\mu = a_\mu = 0$ in eq. (3.18) so that unbroken chrial symmetry is exact and the mesons are actual massless NGBs, not pseudo-NGBs; the *equal-mass limit*, where the pseudoscalar mesons are identical in all regards except flavor; the *isospin limit*, where isospin symmetry is made exact so that all members of the same isospin multiplet has the same mass (equivalent at $n_f = 2$ to the equal-mass limit); and the *planar limit*, which takes $n_f \to \infty$, some effects of which are discussed on page 69. Papers II-V (and refs. [6, 27], etc.) work in the isospin or equal-mass limits, and paper I (and ref. [47], etc.) works in the chiral limit. Notable results using non-equal masses include the NLO $2 \to 2$ scattering amplitude at $n_f = 3$ [52] and one of the progenitors of paper II [53].

Limitations. Even without the above simplifications, the EFT nature of ChPT makes it inherently limited in its applicability. Naïvely, it should be expected to break down around the energy scale where the lightest state not covered by the theory (typically the ρ meson) becomes available. It can also be shown [54] that even self-contained ChPT should cease to

be perturbatively valid around the scale $4\pi F_\pi / \sqrt{n_f}$, which is $\approx 6 M_\pi$ at $n_f = 2$ and $\approx 5 M_\pi$ at $n_f = 3$. It is not clear exactly where breakdown happens or how sharp it is, and part of my work presses close to the region where the convergence could be questioned. Paper II discusses this, and in paper IV we find NLO corrections that are large compared to the LO results, and indications that there may be some large NNLO corrections; see also ref. [39, sec. 4.7]. It remains an open question how far ChPT can be pushed in these directions.

Extensions. There are multiple ways to extend ChPT; I use none of them here, but I will briefly survey them for completeness. They are typically quite successful at describing masses, decays, etc., but constant care has to be taken to assure the validity of the assumptions and expansions involved.

There are chiral Lagrangians including the vector meson octet in addition to the pseudoscalar one [55], and also *resonance chiral theory* [34, 56, 57], which systematically extends ChPT to higher energies by coupling the pseudo-NGBs to mesonic resonances of various spin, parity and mass, including the pseudoscalar η' and the broad scalar resonance σ. Likewise, chiral Lagrangians have been constructed to include the decuplet of spin-$\frac{3}{2}$ baryons up to NLO [58], and the octet of spin-$\frac{1}{2}$ baryons up to NNLO [59] and, in limited cases, N³LO [60]. The vector mesons and spin-$\frac{1}{2}$ baryons are arranged similarly to the pseudoscalar mesons, while the spin-$\frac{3}{2}$ baryons form a totally symmetric three-index tensor in flavor space.

There are even chiral Lagrangians describing baryons and mesons containing heavy (i.e., charm or bottom) quarks [61, 62]; these are otherwise the domain of the *heavy quark effective theory* [63], which is quite different from ChPT.

4 ChPT amplitude calculations

> *The idea […] that I want to describe now is a positive thing. It's a way that we actually use to make calculations and understand nature. Excuse me, to make calculations! […] Understanding real nature, we are unable to do.*
>
> — *Richard Feynman (1983)*

Having firmly established ChPT, I will now cover the methods used for calculating its scattering amplitudes, including the technical development that form the core of my scientific output. For simplicity, the equal-mass limit (see page 56) is assumed throughout this section.

4.1 Feynman diagrams in ChPT

The standard Feynman diagram approach needs some modification in ChPT, since it produces an infinite number of interaction vertices. Only a finite number are needed for a given process at fixed order, though, as determined by Weinberg's power-counting formula [29]: a diagram with L loops and n_k vertices drawn from the $\mathcal{O}(p^k)$ Lagrangian will be $\mathcal{O}(p^m)$, where[82]

$$m = 2 + 2L + \sum_k n_k(k - 2). \tag{4.1}$$

Thus, LO diagrams are trees constructed using only the LO Lagrangian; NLO diagrams have one loop *or* one NLO vertex; NNLO diagrams have two loops, one loop and one NLO vertex, two NLO vertices, or one NNLO vertex; and so on.[83]

Vertices with any even number of legs (odd numbers are forbidden by parity) are available at all orders, with the exception that there is no LO two-leg vertex. The order of legs around a vertex does not matter; along with eq. (4.1), this makes it straightforward to draw ChPT Feynman diagrams, some beautiful examples of which are found in paper III. However, the vertex factors themselves are unwieldy enough that tables of Feynman rules are not made; even the four-point vertex provided by the LO Lagrangian, eq. (3.26), is messy:

$$\frac{1}{6F^2}\langle t^a t^b t^c t^d\rangle\Big[\phi^a\phi^b\partial_\mu\phi^c\partial^\mu\phi^d - \tfrac{1}{2}\phi^a\partial_\mu\phi^b\phi^c\partial^\mu\phi^d\Big] + \frac{B}{6F^2}\langle t^a t^b t^c t^d(s + ip)\rangle\Big[\phi^a\phi^b\phi^c\phi^d\Big]. \tag{4.2}$$

---

[82]This is due to Euler's formula $V - E + L = 1$ for a connected graph with V vertices, E edges and L loops: assign $\mathcal{O}(p^{-2})$ to each edge (propagator) and $\mathcal{O}(p^d)$ to each loop (integral); eq. (4.1) then follows in the specific case $d = 4$. Euler's formula, in turn, is proven by starting from the simplest nonempty graph, consisting of a single vertex and no edges or loops, and noting that all ways to extend it (adding a new vertex connected by a new edge, splitting an edge in two with a new vertex, or adding an edge between two existing vertices to form a loop) leave $V - E + L$ invariant. Note that this graph does not include the external legs of the corresponding Feynman diagram.

[83]More informally, the number of N's in a $\mathrm{N}^k\mathrm{LO}$ diagram is the sum of all N's in the vertices, plus the number of loops.

It is more practical to extract the vertices "on the fly", expanding the Lagrangian terms to the desired order as needed using computer algebra. Hans Bijnens and I invariably use FORM [64] for this due to its suitability for performing relatively straightforward manipulations of enormously complicated expressions.

Being a scalar theory, ChPT has a simple set of $i/(p^2 - M_a^2)$ propagators for the various particles. In the equal-mass limit, where there is a single mass, M_π, the propagator further simplifies to a unit matrix in flavor space:

$$\frac{i\delta_b^a}{p^2 - M_\pi^2}.\tag{4.3}$$

When connecting internal and external legs, care must be taken to sum over all distinct index contractions on each vertex. Furthermore, a derivative ∂_μ acting on a leg becomes a factor of $-ip_\mu$, where p_μ is the leg momentum flowing *into* the vertex.

Loop integrals. A general P-propagator one-loop integral has the d-dimensional form

$$\frac{1}{i}\int\frac{\mathrm{d}^d l}{(2\pi)^d}\frac{(l\cdot r_1)(l\cdot r_2)\cdots(l\cdot r_N)(l^2)^R}{D(q_1)D(q_2)\cdots D(q_P)}\,,\qquad N,R\geq 0\,,\tag{4.4}$$

where r_i and q_i are fixed momenta and $D(q)\equiv(l-q)^2 - M_\pi^2$; by shifting l, one of the q_i can be set to zero. Due to the derivative couplings, ChPT typically results in more complicated numerators (larger N and R) than are usually encountered in renormalizable theories. Through Passarino–Veltman reduction [65], the numerators can be systematically expressed in terms of $D(q_i)$, $r_i\cdot r_j$ and things that vanish by symmetry, reducing the integrals to ones with trivial numerators. With one propagator, this is[84]

$$A\equiv\frac{1}{i}\int\frac{\mathrm{d}^d l}{(2\pi)^d}\frac{1}{l^2 - M_\pi^2} = -M_\pi^{d-2}\frac{\Gamma\left(1-\frac{d}{2}\right)}{(4\pi)^{d/2}} = M_\pi^2\left(\frac{\kappa}{\tilde{\epsilon}} - L\right) + \mathcal{O}(\tilde{\epsilon})\,,\tag{4.5}$$

using ChPT conventions (recall section 3.6). With two propagators, one gets

$$B(q^2)\equiv\frac{1}{i}\int\frac{\mathrm{d}^d l}{(2\pi)^d}\frac{1}{D(0)D(q^2)} = \frac{\kappa}{\tilde{\epsilon}} - \kappa - L + \bar{J}(q^2) + \mathcal{O}(\tilde{\epsilon})\,.\tag{4.6}$$

See appendix C.3 of paper II for details on the function $\bar{J}$. Two-loop order is also well-explored for ChPT; see, e.g., ref. [22, 27].

In principle, Passarino–Veltman reduction reduces all three-or-more-propagator one-loop integrals in terms of a few functions, which in turn can be written in terms of $\bar{J}(q_i^2)$, but we find in paper III (originally, ref. [6]) that these reduced expressions are prohibitively long for the integrals that appear in multi-meson one-loop amplitudes. See appendix A of paper III for the formulation used instead.

---

[84]Readers who may be a bit rusty on the details of one-loop calculations are kindly directed to Peskin & Schroeder [14, pp. 249-51 & appendix A.4].

4.2 Stripped amplitudes

Ἀπεκδυσάμενοι τὸν παλαιὸν ἄνθρωπον σὺν ταῖς πράξεσιν αὐτοῦ, καὶ ἐνδυσάμενοι τὸν νέον τὸν ἀνακαινούμενον εἰς ἐπίγνωσιν...

You have stripped off the old self with its practices and have clothed yourselves with the new self, which is being renewed in knowledge...

— *Colossians 3:9-10*

Consider any vertex that can be obtained from the ChPT lagrangian. Like in eq. (4.2), the fields and derivatives can be pulled out, leaving a product of traces containing $SU(n_f)$ generators, connected to the fields via *flavor indices* $a, b, c, \ldots$. In general, the traces will also contain the source fields s, p, v_μ and a_μ, but in the equal-mass limit, these are proportional to the unit matrix and can be pulled out of the traces as well. These traces, the *flavor structure*, can be leveraged to significantly simplify the presentation and calculation of ChPT amplitudes. Much of this is done in analogy with the use of color structures in perturbative QCD (see Johan Thorén's thesis [1] for an overview).

Throughout the rest of this section, I will follow papers I and III (and refs. [27, 47], etc.) in adopting the *all-ingoing convention*, where n-point amplitudes are expressed with n ingoing and zero outgoing particles. This puts all momenta on the same footing, making symmetries more manifest and simplifying conservation of momentum to $\sum_i p_i = 0$. It also gives more freedom in drawing Feynman diagrams, since the direction of time is irrelevant. Physically relevant scattering amplitudes are easily obtained from an all-ingoing amplitude through crossing, as is done in papers IV-V.

Flavor structures. I will now introduce some concepts first used in paper I, although I will stay closer to the notation of paper III. I denote a trace of generators carrying r flavor indices b_i by

$$\mathcal{F}_{\{r\}}(b_1, b_2, \ldots, b_r) \equiv \langle t^{b_1} t^{b_2} \cdots t^{b_r} \rangle, \tag{4.7}$$

and more generally, I define a *flavor structure* $\mathcal{F}_R(b_1, \ldots, b_n)$ by

$$\mathcal{F}_{\{r_1, r_2, \ldots\}}(b_1, \ldots, b_n) \equiv \langle t^{b_1} t^{b_2} \cdots t^{b_{r_1}} \rangle \, \mathcal{F}_{\{r_2, \ldots\}}(b_{r_1+1}, \ldots, b_n), \tag{4.8}$$

where $R \equiv \{r_1, \ldots\}$ is the *flavor split*; the r_i must add up to n, and by convention, they are sorted in ascending order. For example, $\mathcal{F}_{\{2,3\}}(a, b, c, d, e) = \langle t^a t^b \rangle \langle t^c t^d t^e \rangle$, and $\mathcal{F}_{\{1,2\}}(a, b, c) = \mathcal{F}_{\{1\}}(a) = 0$ due to tracelessness.

Given a permutation $\sigma \in \mathcal{S}_n$ that maps $i \to \sigma(i)$, I define the permuted flavor structure $\mathcal{F}_R^\sigma(\ldots)$ by

$$\mathcal{F}_R^\sigma(b_1, b_2, \ldots, b_n) \equiv \mathcal{F}_R(b_{\sigma(1)}, b_{\sigma(2)}, \ldots, b_{\sigma(n)}). \tag{4.9}$$

Since traces are cyclic and commute, $\mathcal{F}_R(\ldots)$ is invariant when one cycles the indices within individual traces, or swaps two traces containing the same number of indices. I denote the group of all such operations by $\mathbb{Z}_R$, so named because $\mathbb{Z}_{\{r\}}$ is the cyclic group $\mathbb{Z}_r$. $\mathbb{Z}_R$ is non-abelian if two r_i are equal, and can in general be written as a semi-direct product, as shown in paper I (see page 101).

Stripping. After connecting all legs of a N^kLO vertex $\mathcal{V}$, it will be a function of the momenta p_i and flavors b_i of the legs. Taking the order of the momenta as fixed, the vertex will be a sum over all ways to contract the flavor indices into the flavor structure,

$$\mathcal{V}(p_1, b_1; p_2, b_2; \ldots; p_n, b_n) = \sum_R \sum_{\sigma \in \mathcal{S}_n/\mathbb{Z}_R} \mathcal{F}_R^\sigma(b_1, \ldots, b_n) \mathcal{V}_R^\sigma(p_1, \ldots, p_n), \qquad (4.10)$$

which defines the purely kinematic object $\mathcal{V}_R^\sigma$: it consists of whatever factors happen to multiply $\mathcal{F}_R^\sigma$. Here, we sum over all flavor splits R imposed by the N^kLO Lagrangian (for LO, this will just be $\{n\}$), and the $\mathbb{Z}_R$-symmetry of $\mathcal{F}_R$ ensures we only have to sum over $\mathcal{S}_n/\mathbb{Z}_R$ rather than $\mathcal{S}_n$. This implies that $\mathcal{V}_R^\sigma(p_1, \ldots)$ is also $\mathbb{Z}_R$-symmetric.[85]

Due to Bose symmetry, the vertex is totally symmetric in its legs, so simultaneous permutations of the b_i and p_i may be freely applied. By definition [eq. (4.9)],

$$\mathcal{F}_R^\sigma(b_{\rho(1)}, \ldots, b_{\rho(n)}) = \mathcal{F}_R^{\sigma \circ \rho}(b_1, \ldots, b_n), \qquad (4.11)$$

where $\sigma \circ \rho$ denotes the composition of the permutations ρ and σ. Consequently, the same holds for the kinematics:

$$\mathcal{V}_R^\sigma(p_{\rho(1)}, \ldots, p_{\rho(n)}) = \mathcal{V}_R^{\sigma \rho}(p_1, \ldots, p_n). \qquad (4.12)$$

In particular,

$$\mathcal{V}_R^\sigma(p_1, \ldots, p_n) = \mathcal{V}_R^{\mathrm{id}}(p_{\sigma(1)}, \ldots, p_{\sigma(n)}) \equiv \mathcal{V}_R(p_{\sigma(1)}, \ldots, p_{\sigma(n)}) \qquad (4.13)$$

where $\mathrm{id}(i) = i$ is the identity permutation. This object $\mathcal{V}_R$, called the *stripped vertex*, encodes all relevant information about $\mathcal{V}$, since "dressing" it back to a full vertex using eqs. (4.10) and (4.13) is a trivial operation. It is also called a *flavor-ordered* vertex, since it is obtained from the full vertex by keeping those factors that multiply $\mathcal{F}_R(b_1, b_2, \ldots, b_n)$, the flavor structure whose flavor indices are in order.

Amplitudes. In the equal-mass limit, the ChPT propagator is a unit matrix in flavor space, so using the Fierz identity, eq. (3.42), any number of vertices connected with any number of propagators retains the form of eq. (4.10). Thus, the above arguments for $\mathcal{V}$ apply also to equal-mass ChPT scattering amplitudes, so

$$\mathcal{M}(p_1, b_1; p_2, b_2; \ldots; p_n, b_n) = \sum_R \sum_{\sigma \in \mathcal{S}_n/\mathbb{Z}_R} \mathcal{F}_R^\sigma(b_1, \ldots, b_n) \mathcal{M}_R(p_{\sigma(1)}, \ldots, p_{\sigma(n)}), \quad (4.14)$$

where $\mathcal{M}_R$ is the *stripped amplitude*. This is typically a more manageable object than the full amplitude, with roughly a factor $|\mathcal{S}_n/\mathbb{Z}_R|$ fewer terms; this ranges from $n!/2^n$ for $R = \{2, 2, \ldots, 2\}$ to $(n-1)!$ for $R = \{n\}$.

The general idea described here is not new; in perturbative QCD, it is common to strip away the color structure, leaving a color-ordered object called the *primitive amplitude* (see,

---

[85]As explained in paper III, objects like $\mathcal{V}_R^\sigma(\ldots)$ derived from a charge-conjugation-invariant Lagrangian actually have a larger symmetry, denoted $\mathbb{Z}_R^{+\mathrm{TR}}$, which also includes trace-reversal, i.e., reversing the order of the indices in each trace. For simplicity, I only use $\mathbb{Z}_R$ here.

e.g., refs. [66–69]), and similar concepts are found in older literature (e.g., refs. [8, 70]) using the term "duality". Stripped amplitudes for massless LO ChPT were introduced in ref. [47] and generalized beyond LO in paper I, although special cases are much older; for instance, the standard form of the 4-pion ChPT amplitude is

$$
\begin{aligned}
\mathcal{M}_{4\pi}(p_1, b_1; \ldots; p_4, b_4) = {}& \left(\langle t^{b_1} t^{b_2} t^{b_3} t^{b_4} \rangle + \langle t^{b_4} t^{b_3} t^{b_2} t^{b_1} \rangle \right) B(s, t, u) \\
& + \left(\langle t^{b_1} t^{b_3} t^{b_4} t^{b_2} \rangle + \langle t^{b_2} t^{b_4} t^{b_3} t^{b_1} \rangle \right) B(t, u, s) \\
& + \left(\langle t^{b_1} t^{b_4} t^{b_2} t^{b_3} \rangle + \langle t^{b_3} t^{b_2} t^{b_4} t^{b_1} \rangle \right) B(u, s, t) \\
& + \langle t^{b_1} t^{b_2} \rangle \langle t^{b_3} t^{b_4} \rangle C(s, t, u) + \langle t^{b_1} t^{b_3} \rangle \langle t^{b_2} t^{b_4} \rangle C(t, u, s) \\
& + \langle t^{b_1} t^{b_4} \rangle \langle t^{b_2} t^{b_3} \rangle C(u, s, t)
\end{aligned}
\tag{4.15}
$$

(compare eq. (III.3.2) and ref. [27]), where s, t, u are the *Mandelstam variables*[86]

$$
s \equiv (p_1 + p_2)^2, \quad t \equiv (p_1 + p_3)^2, \quad u \equiv (p_1 + p_4)^2; \qquad s + t + u = 4M_\pi^2. \tag{4.16}
$$

We can clearly identify the stripped amplitudes as $\mathcal{M}_{\{4\}} = B$ and $\mathcal{M}_{\{2,2\}} = C$. The functions $B(s, t, u)$ and $C(s, t, u)$ are known explicitly to NNLO [27], but note that no diagrammatic calculations are needed to write down the general form of eq. (4.15).

The Cayley–Hamilton theorem, eqs. (3.43) and (3.47), can reduce the number of distinct flavor splits, and therefore the number of stripped amplitudes. For instance, the $n_f = 2$ version of eq. (4.15), originally due to Weinberg [72], is

$$
\begin{aligned}
\mathcal{M}_{4\pi}(p_1, b_1; \ldots; p_4, b_4) = {}& \langle t^{b_1} t^{b_2} \rangle \langle t^{b_3} t^{b_4} \rangle A(s, t, u) + \langle t^{b_1} t^{b_3} \rangle \langle t^{b_2} t^{b_4} \rangle A(t, u, s) \\
& + \langle t^{b_1} t^{b_4} \rangle \langle t^{b_2} t^{b_3} \rangle A(u, s, t),
\end{aligned}
\tag{4.17}
$$

where $A(s, t, u) = C(s, t, u) + B(s, t, u) + B(t, u, s) - B(u, s, t)$. It is of course possible to reformat this so that there is a $\mathcal{M}_{\{4\}}$ but no $\mathcal{M}_{\{2,2\}}$ instead.

Papers I and III contain many more examples of flavor splits and stripped amplitudes. Paper I also proves (in appendix B) that the stripped amplitude is unique, as long as n_f is kept variable; however, fixed n_f can lead to relations like that seen above.

4.3 Unitarity, analyticity, and all that

> *[E]ntre deux vérités du domaine réel, le chemin le plus facile et le plus court passe bien souvent par le domaine complexe.*
>
> *[B]etween two truths of the real domain, the easiest and shortest path quite often passes through the complex domain.*
>
> — *Paul Painlevé (1900)*

A large number of facts can be derived about the scattering amplitude without appealing to perturbation theory, and these can in turn be used as an aid or complement to perturbative

---

[86]Introduced by Stanley Mandelstam [71] in essentially their modern form. Also commonly called Mandelstam *invariants*, reflecting their Lorentz-invariant nature.
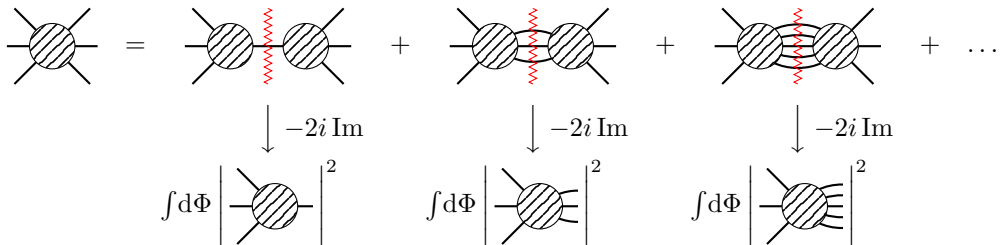
Figure 4: Schematic representation of cuts (red zigzag lines) through 3-to-3 Feynman diagrams, employing the optical theorem, eq. (4.19). Shaded blobs represents the sum of all diagrams with the appropriate number of legs. Odd numbers of legs are assumed to be forbidden.

calculations. Firstly, due to conservation of probability, the S-matrix must be a unitary operator, $\mathcal{S}^{\dagger}\mathcal{S} = 1$, which implies for the T-matrix that [recall eq. (2.53)]

$$-i(\mathcal{T} - \mathcal{T}^{\dagger}) = \mathcal{T}^{\dagger}\mathcal{T}. \tag{4.18}$$

Surrounding this by $\langle X| \dots |X\rangle$ and inserting between $\mathcal{T}^{\dagger}\mathcal{T}$ the unit matrix in the form of the sum over intermediate states, $\sum_{Y} \int d\Phi_Y |Y\rangle\langle Y| = 1$, this results in the *optical theorem* [recall eq. (2.55)]:

$$-2i\,\mathrm{Im}\,\mathcal{M}(X \to X) = \sum_{Y} \int d\Phi_Y \,\mathcal{M}^*(X \to Y)\mathcal{M}(X \to Y) \sim \sum_{Y} \sigma(X \to Y), \tag{4.19}$$

where "$\sim$" omits kinematic details. Only states $|Y\rangle$ that are kinematically allowed to be produced on-shell contribute to the right-hand side.

On the diagram level, this is is naturally interpreted in terms of *cuts* through diagrams, separating the inital and final state by cutting some internal lines. Whenever the kinematics permit the cut lines to go on-shell, eq. (4.19) produces a corresponding contribution to the imaginary part, as shown schematically in fig. 4.

Analyticity. Another fundamental property is that the amplitude, seen as a complex function of the momenta, is *analytic* (this easily follows from the properties of Feynman diagrams, but holds also non-perturbatively). Let us therefore write $\mathcal{M}$ as a function of s, the center-of-mass energy squared, while keeping all other kinematic parameters fixed. If $\mathcal{M}$ is real—i.e., $\mathcal{M}^*(s) = \mathcal{M}(s^*)$—in any region of the complex plane, then by analytic continuation this relation holds everywhere. Thus, for real s and $\epsilon > 0$,

$$\mathrm{Re}\,\mathcal{M}(s + i\epsilon) = \mathrm{Re}\,\mathcal{M}(s - i\epsilon), \qquad \mathrm{Im}\,\mathcal{M}(s + i\epsilon) = -\,\mathrm{Im}\,\mathcal{M}(s - i\epsilon), \tag{4.20}$$

but if s is sufficient to produce some state $|Y\rangle$, then eq. (4.19) implies that $\mathrm{Im}\,\mathcal{M}(s)$ is nonzero, so $\mathrm{Im}\,\mathcal{M}(s)$ is discontinuous across the real line. Consequently, $\mathcal{M}(s)$ has a *branch cut*, starting at the threshold for producing the lightest state $|Y\rangle$ and continuing along the real line.
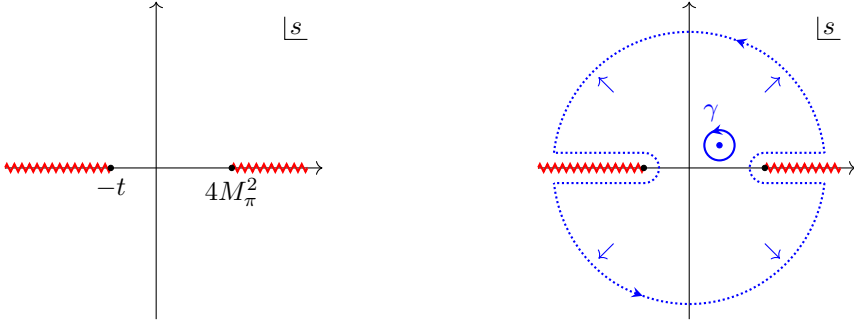
(a) Branch cuts (**red**): the s-channel starts at $s = (2M_\pi)^2$ and continues to the right; the u-channel starts at $u = (2M_\pi)^2$ (i.e., $s = -t$) and continues to the left.

(b) The contour (**blue**, solid) used in the dispersion relation around a point (**blue**), and the deformed contour (**blue**, dotted).

Figure 5: The analytic structure of the $2\pi \to 2\pi$ amplitude, shown in the s complex plane for fixed t (real and positive here, but that is not necessary as long as the cuts do not overlap). The lightest possible state is that of two pions at rest.

Specifically, the above describes the s-channel cut. There will be an additional cut in each channel that can be related to the s-channel through crossing symmetry, so for the $2\pi \to 2\pi$ scattering amplitude, there will also be a t-channel and a u-channel cut. This can be seen directly in eq. (4.15): for instance, if $B(s,t,u)$ has a cut in s, it must also have one in u, since $B(s,t,u) = B(u,t,s)$ due to $\mathbb{Z}_{\{4\}}$ symmetry. At fixed t, eq. (4.16) relates s and t via $u = 4M_\pi - s - t$, giving two branch cuts as shown in fig. 5a.[87] The branch-cut structure for $3\pi \to 3\pi$ and above is of course much more intricate.

Dispersion relations. The analyticity of $\mathcal{M}$ allows the use of *dispersion relations*, which follow directly from Cauchy's integral formula,

$$\frac{\mathrm{d}^k \mathcal{M}(s)}{\mathrm{d}s^k} = \frac{k!}{2\pi i} \oint_\gamma \mathrm{d}z \, \frac{\mathcal{M}(s)}{(z-s)^{k+1}} \,. \tag{4.21}$$

Here, γ is a small circular contour around $z = s$, inside of which $\mathcal{M}(s)$ must lack singularities. This contour can be deformed as illustrated for $2 \to 2$ scattering in fig. 5b, hugging the branch cuts at a distance $\epsilon \to 0$ and sending the circular part to infinity; assuming k is large enough, this portion of the contour does not contribute due to suppression by $1/z^{k+1}$. The real part of the cut-hugging contour cancels due to eq. (4.20), but due to the discontinuity, the imaginary part remains, leaving

$$\frac{\mathrm{d}^k \mathcal{M}(s)}{\mathrm{d}s^k} = \frac{k!}{\pi} \int_{\mathrm{thr.}}^\infty \mathrm{d}z \, \mathrm{Im}\,\mathcal{M}(z + i\epsilon) + \sum [\text{crossings}] \,, \tag{4.22}$$

where "thr." is the threshold at which the s-channel cut starts, and "crossings" represents similar contributions from the other cuts, such as the u-channel in fig. 5b.

---

[87] The one-loop integral function $\bar{J}(q^2)$ [recall eq. (4.6)] has a branch cut starting at $q^2 = (2M_\pi)^2$, so the general branch-cut structure of $\mathcal{M}_{2\pi}$ is manifest already at one-loop level.

The dispersion relations allow $\mathcal{M}(s)$ to be related to $\operatorname{Im}\mathcal{M}$ integrated over a completely different kinematic region than the one s is chosen from. In paper III, this is leveraged along with the optical theorem, which relates $\operatorname{Im}\mathcal{M}$ to the real, positive quantity $\sigma(2\pi \to 2\pi)$, in order to obtain non-perturbatively valid relations stating that the left-hand side of eq. (4.22) is positive. Choosing s, t in a region where fixed-order ChPT is accurate, this places constraints on the possible values of $\mathcal{M}$, and therefore on the LECs that appear in it.

Isospin channels. Many things are simplified by decomposing the amplitude into channels distinguished by some conserved quantity. For example, the two-flavor ChPT used in papers III-V conserves charge and isospin, so for each total charge, the $k\pi \to k\pi$ amplitude can be decomposed over a basis of k-pion states with definite isospin; since pions have isospin 1, the total isospin ranges from 0 to k. The details can be worked out either through group theory or by explicitly constructing states; see ref. [27] for a detailed treatment of $k = 2$, and paper V for one of $k = 3$. Each isospin channel may have multiple components, which can further be categorized by how they transform under permutations of the initial or final pions; for instance, the $k = 2$ and $k = 3$ cases decompose as

$$\mathbf{3} \otimes \mathbf{3} = \underbrace{\mathbf{1}}_{T} \oplus \underbrace{\mathbf{3}}_{A} \oplus \underbrace{\mathbf{5}}_{T}, \tag{4.23}$$

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \underbrace{\mathbf{1}}_{A} \oplus \underbrace{\mathbf{3}}_{T} \oplus \underbrace{\mathbf{3} \oplus \mathbf{3}}_{S} \oplus \underbrace{\mathbf{5} \oplus \mathbf{5}}_{S} \oplus \underbrace{\mathbf{7}}_{T}, \tag{4.24}$$

where $(I^2 + 1)$ in boldface represents a multiplet of isospin I. The permutation behavior is indicated on the right-hand side: T stands for "trivial" (totally symmetric in the initial or final pions), A for "alternating" (totally antisymmetric), and S for the two-dimensional standard representation of $\mathcal{S}_3$; see ref. [73, appendix C].

With n_f flavors, the natural generalization is that mesons transform under the *adjoint* representation of $\mathrm{SU}(n_f)$, which for $n_f = 3$ gives the familiar meson octet, $\mathbf{8}$. This generally results in far more channels, with five or six appearing in the 2-to-2 amplitude used in paper II; the details are again given in ref. [27]. A 3-to-3 analysis like that in paper V has not yet been performed for $n_f > 2$.

After decomposing the amplitude over (generalized) isospin channels, crossing the amplitude becomes a simple linear transformation in the space spanned by these channels. This allows eq. (4.22) to be put in a useful form in paper II.

Factorization and recursion. As a final application of the topics in this section, consider the single-particle cut in fig. 4. Since the kinematics of a single on-shell particle is entirely constrained and has smaller s than any multi-particle state, this produces an isolated pole in $\mathcal{M}$ rather than an extended branch cut. Generally, for *any* tree-level propagator with momentum P, the amplitude in the vicinity of $P^2 = M_\pi^2$ must go like

$$\mathcal{M}(p_1, \ldots, p_m; k_1, \ldots, k_n) \sim \frac{\mathcal{M}_L(p_1, \ldots, p_m, +P)\mathcal{M}_R(k_1, \ldots, k_n, -P)}{P^2 - M_\pi^2}, \tag{4.25}$$

where $\mathcal{M}_{L,R}$ are the halves on the left- and right-hand side of the cut, respectively (corresponding to the blobs in fig. 4), and p_i, k_i are the the corresponding momenta, with

$-\sum_{i=1}^{m} p_i = +\sum_{i=1}^{n} k_i = P$ in the all-ingoing convention. $\mathcal{M}_L$ and $\mathcal{M}_R$ are valid $(m+1)$- and $(n+1)$-point amplitudes, respectively.

The factorization property, eq. (4.25), can be used directly to express the amplitude also away from the poles by extending the definition of $\mathcal{M}_{L,R}$ to the case where P is off-shell. There is no unique way of doing this, and there will in general be a remainder that lacks poles and does not factorize, but it is nevertheless useful—a significant part of the six-point amplitude in paper III is expressed in terms of the four-point amplitude by consistently choosing its off-shell form and summing over all ways of distributing the six momenta across $\{p_1, p_2, p_3\}$ and $\{k_1, k_2, k_3\}$, with each way producing one pole. The non-facorizing remainder is still complicated [just look at eq. (III.4.2)!], but simpler than the full amplitude.

However, being analytic, $\mathcal{M}$ is *entirely* described by the residues at its poles, and following eq. (4.25), those residues are simply $\mathcal{M}_L \mathcal{M}_R$. Therefore, it must be possible to recursively state $\mathcal{M}$ in terms of lower-point amplitudes, without the need for off-shell forms or non-factorizing remainders. This is done in Britto–Cachazo–Feng–Witten (BCFW) recursion [74, 75], where one deforms all momenta q_i into $q_i(z)$ such that $q_i(0)$ is the original, un-deformed q_i, and each pole is obtained at some $z = z_i$. Assuming $\mathcal{M}(z)$ falls off sufficiently fast as $z \to \infty$, Cauchy's theorem then gives

$$0 = \operatorname*{Res}_{z=0} \frac{\mathcal{M}(z)}{z} + \sum_i \operatorname*{Res}_{z=z_i} \frac{\mathcal{M}(z)}{z} \,, \qquad (4.26)$$

where the $z = 0$ residue is simply $\mathcal{M}$, and the remaining residues are equal to the right-hand side of eq. (4.25), times a z_i-dependent factor depending on the scheme used for the deformation. This can be applied recursively to $\mathcal{M}_{L,R}$ until one arrives at simple amplitudes without poles.

A large class of theories, including QCD, are *on-shell constructible*, meaning that their tree-level amplitudes are completely determined by the BCFW procedure. Thus, they only require knowledge of one or a few simple amplitudes—essentially on-shell equivalents of Feynman rules—and the pattern of poles that can appear. ChPT is *not* on-shell constructible in the conventional sense, though, since its amplitudes do not fall off sufficiently fast as $z \to \infty$. Nevertheless, recursion relations have been found for chiral-limit ChPT [47, 76, 77] by taking advantage of a property of a large class of EFTs, namely the *Adler zero* [78, 79] or more generally the *soft behavior*: as any external momentum q_i is sent to zero, $\mathcal{M}$ goes to zero as $\mathcal{O}(|q_i|^\sigma)$, where $\sigma = 1$ for ChPT (see ref. [80] for a classification of EFTs in terms of σ). This allows for enhanced momentum deformations that suppress the $z \to \infty$ behavior; were it not for the soft behavior, these would instead cause problems as $z \to 0$.

Recursion is particularly convenient for constructing stripped amplitudes with $R = \{n\}$, due to their flavor-ordered property: poles can only appear where P is a sum of *consecutively* indexed momenta, and $\mathcal{M}_{L,R}$ will also be stripped, flavor-ordered amplitudes. This connects intimately with paper I, which generalizes a complementary, non-recursive approach developed in ref. [47] to easily handle $R \neq \{n\}$; this is described in the next section. It is worth mentioning, though, that recursive methods have successfully handled $R = \{2, 4\}$ [81] as well as planar-limit one-loop amplitudes [82]; the $z \to \infty$ suppression

of ref. [77] works at NLO, but (unlike paper I) breaks down at NNLO.

4.4 Diagrammatic flavor-ordering

Well represented, half solved. — *Mathematicians' proverb*
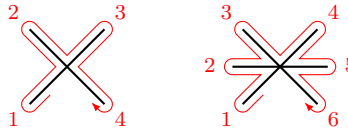
In section 4.2, we took the step from stripped vertex to stripped amplitude by noting that the equal-mass propagator preserves the general structure via the Fierz identity, eq. (3.42), which I will reproduce here for convenience:

$$\frac{1}{\tau}\langle Xt_a\rangle\langle Yt^a\rangle = \langle XY\rangle - \frac{1}{n_f}\langle X\rangle\langle Y\rangle\,, \tag{4.27a}$$

$$\frac{1}{\tau}\langle Xt_aYt^a\rangle = \langle X\rangle\langle Y\rangle - \frac{1}{n_f}\langle XY\rangle\,. \tag{4.27b}$$
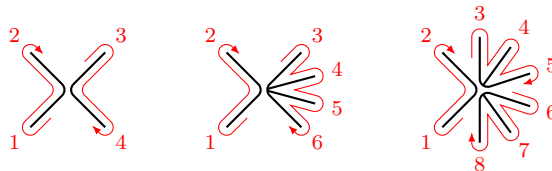
Recalling that the stripped vertex is defined by being flavor-ordered, i.e., attached to $\mathcal{F}_R^{\mathrm{id}}$, we note that eq. (4.27) preserves the ordering of the indices inside X and Y. This makes it possible to carefully connect together a flavor-ordered amplitude from flavor-ordered vertices, obtaining the stripped amplitude without going through the much more unwieldy and computationally expensive full amplitude. The method of doing so is called *diagrammatic flavor-ordering* (DFO), which was introduced in paper I based on the massless LO case treated in ref. [47], and named and refined in paper III.

DFO uses modified Feynman diagrams, where the order of legs around a vertex is fixed to reflect the flavor-ordering. Legs are indexed clockwise around the vertex, with leg i having momentum p_i and flavor b_i:[88]

$$\tag{4.28}$$

The starting point is arbitrary due to $\mathbb{Z}_R$ symmetry. Vertices with multi-trace flavor structures are drawn in several pieces, separated by a small gap in the middle, so that b_i and b_j are in the same trace precisely when legs i and j are in the same piece. Pieces are indexed individually, starting with the smaller ones:[89]

$$\tag{4.29}$$

---

[88] The **red** lines and numbers are included for illustration only. The diagram itself is drawn in **black**.

[89] The graphical form of a vertex should also reflect its power-counting order, but for simplicity, I omit that here. Two different conventions for indicating orders are used in papers I and III.

The ordering of same-size pieces is arbitrary due to $\mathbb{Z}_R$ symmetry. I stress that the gaps only indicate the flavor structure and do not hinder momentum flow.

Now, consider joining two vertices with an equal-mass propagator, which in flavor space is just δ_b^a. Setting $n_f \to \infty$ for the moment, eq. (4.27a) merges the flavor structures by directly joining one trace from each, preserving their ordering: $\langle X t_a \rangle \delta_b^a \langle Y t^b \rangle = \langle XY \rangle$. Repeating this with more vertices, we find that tree diagrams can be indexed just like single vertices, following the outline of each part in turn, and the result is guaranteed to be flavor-ordered:

$$(4.30)$$

For traces that span several vertices, the starting point for the indexing does matter—just consider which momenta flow through each propagator—so an un-indexed flavor-ordered diagram represents the sum of all valid ways to index it, divided by a symmetry factor to account for not all indexings being distinct. Pieces contained in a single vertex are not summed due to the symmetry of the vertex, so single-vertex diagrams always have a single indexing and a symmetry factor of 1. The issue of distinct indexings is discussed more in paper I; some examples of symmetry factors are[90]

$$(4.31)$$

Re-instating $n_f < \infty$, we see that the last term in eq. (4.27a) behaves like a propagator that does not merge the traces. In DFO, this term is represented as a separate propagator, drawn with a dashed line (- - - - -) and containing the factor of $-1/n_f$. Dashed lines are ignored when indexing a diagram, keeping to the rule that only contiguous pieces of a diagram (as they appear graphically) belong to the same trace:

$$(4.32)$$

These propagators can be interpreted by noting that the $\mathrm{U}(n_f)$ version of eq. (4.27) lacks the $-1/n_f$ terms, as can be derived by including the extra generator $t^0 \equiv \sqrt{\tau/n_f}$. Thus, they represent how a $\mathrm{SU}(n_f)$ theory is obtained from a $\mathrm{U}(n_f) \cong \mathrm{SU}(n_f) \times \mathrm{U}(1)$ theory by

---

[90]The symmetry factors shown here assume that all vertices are of the same order; different-order vertices can break diagram symmetries. The diagrams in eq. (4.31) have some historical significance for the development of DFO. From left to right, they are the simplest diagram to feature a nontrivial symmetry factor, one that breaks some naïve assumptions about symmetry factors (which would give it a factor of $1/4$), and a pathological example that broke my first attempt at automated diagram generation. Note that the second and third diagrams are at least $\mathrm{N}^3\mathrm{LO}$ and $\mathrm{N}^4\mathrm{LO}$, respectively.

subtracting the contributions of the the U(1) singlet field ϕ^0.[91] Therefore, these dashed propagators are called *singlet propagators*. The LO interaction $\langle u_\mu u^\mu \rangle$ does not couple ϕ^0 to ϕ^a [47], but the LO mass term $\langle \chi_+ \rangle$ does, as do the higher-order interactions. Thus, singlets are normally needed at all orders, but only at NNLO and above in the chiral limit.

When forming loops, a propagator might connect two indices in the same trace, in which case eq. (4.27b) is used. The non-singlet term introduces a flavor split, separating the "inside" and "outside" of the loop, while the singlet does not. All cases are consistently covered when indexing diagrams by following their outline and ignoring singlet lines, even when non-planar diagrams are introduced at NNLO:

$$(4.33)$$
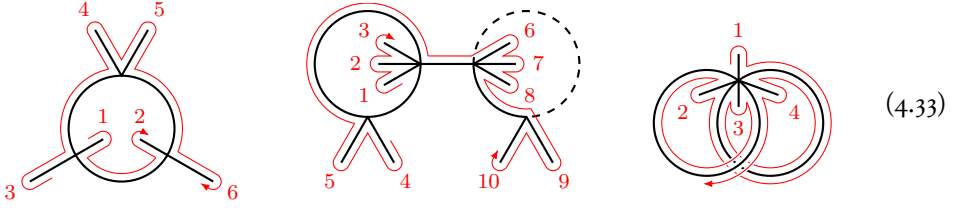
Note that empty traces, giving factors of $\langle 1 \rangle = n_f$, may appear due to loops or singlets:

$$(4.34)$$
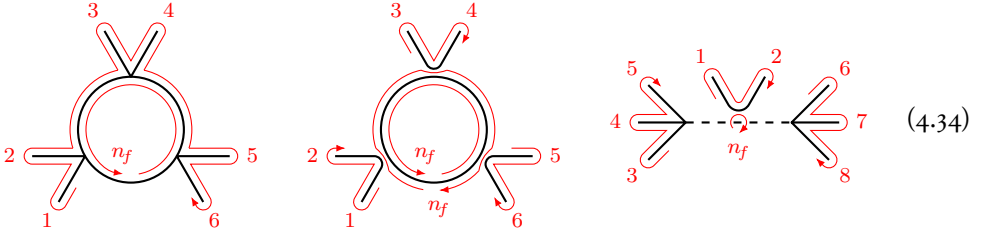
The *planar limit*, obtained by discarding subleading terms in a large-n_f expansion, removes singlets, non-planar diagrams, and loops that introduce flavor splits rather than empty traces. Since multi-trace vertices act as counterterms to loops that introduce flavor splits, they are removed as well. Calculations simplify significantly in the planar limit; see, e.g., ref. [82].

That summarizes DFO. Coupled with a simplified version of the computer algebra for calculating amplitudes and a suitable diagram generator,[92] it is extremely powerful at tree-level in the massless limit, which is covered in paper I; its utility is diminished by loops and masses, which lead to a large number of diagrams, but it speeds up the algebra and clarifies some structures, as described in paper III. Note that it is entirely reliant on the equal-mass limit, and breaks down if the propagator is not proportional to a unit matrix.

---

[91]Naïvely, this singlet field would correspond to the η', but trying to include η' in ChPT by just not subtracting the singlet fails due to the axial anomaly. It is possible to include η' in the limit of infinitely many QCD colors [83] or using resonance chiral theory [34, 56, 57].

[92]Paper I describes a DFO diagram generator called FODGE (Flavor-Ordered Diagram GEnerator). In connection with paper III, I developed a version, LODGE, that supports LOops.
Both are available from `https://github.com/mssjo/fodge`.

4.5 Kinematics and deorbiting

Kinematics is a simple matter for 4 or fewer particles, but requires more attention in the larger systems studied throughout this thesis (except in paper II). Here, I describe the systematic treatment of many-particle kinematics and how it allows amplitudes to be rewritten in even simpler forms than the stripped amplitude described above.

The kinematics of n spinless particles is described by n momenta $\{p_i\}_{i=1}^n$. These have nd components in d dimensions, but are not all independent: the on-shellness condition $p_i^2 = m_i^2$ means each momentum has only $(d-1)$ degrees of freedom, and conservation eliminates one momentum through $\sum_i p_i = 0$, bringing the total degrees of freedom, N_{dof}, down to $d(n-1) - n$. Furthermore, Lorentz invariance must be taken into account, which can be done in two ways. Subtracting the dimension of the Lorentz group, consisting of $d-1$ boosts and $(d-1)(d-2)/2$ rotations, results in

$$N_{\text{dof}} = (d-1)n - \frac{d(d+1)}{2} \,, \tag{4.35}$$

or $3n - 10$ in $d = 4$. Observing instead that there are $n(n+1)/2$ ways to form Lorentz-invariant products $p_i^\mu p_{j\mu}$ and accounting for on-shellness and conservation of momentum, which eliminate n products each, gives

$$N_{\text{dof}} = \frac{n(n-3)}{2} \,, \tag{4.36}$$

independent of d.[93] These two ways of counting degrees of freedom disagree except at $n = d + 1$. When $n > d + 1$, eq. (4.36) overestimates the number by assuming that all remaining products are independent, which they are not. When $n < d+1$, the up to $n-1$ linearly independent momenta only span a subspace, so only a part of the full Lorentz group actually applies to them and eq. (4.35) underestimates the number. Thus, eq. (4.36) applies for $n \leq d + 1$ and eq. (4.35) for $n \geq d + 1$.

Even though eq. (4.35) gives the number of degrees of freedom for all but the lowest n, it turns out to be less than useful for general treatment of the kinematics, since it is difficult to give a minimal set of Lorentz-invariant kinematic parameters; a minimal but frame-dependent set is given for $n = 6$, $d = 4$ in paper IV. The number of parameters stated by eq. (4.36), on the other hand, is easy to work with, and even though they will in general not be independent, they are useful for purposes such as papers I and III.

Mandelstams. In the spirit of eq. (4.35), kinematical parameters are conveniently expressed in terms of *generalized Mandelstam variables* (or *Mandelstams* for short),

$$\hat{s}_{ij\ldots} \equiv (p_i + p_j + \ldots)^2 \,, \tag{4.37}$$

which generalize the $n = 4$ Mandelstam variables, eq. (4.16),

$$s \equiv \hat{s}_{12} = \hat{s}_{34} \,, \qquad t \equiv \hat{s}_{13} = \hat{s}_{24} \,, \qquad u \equiv \hat{s}_{14} = \hat{s}_{23} \,, \tag{4.38}$$

---

[93]I ignore other Lorentz-invariant contractions such as $\epsilon_{\mu\nu\rho\sigma} p_i^\mu p_j^\nu p_k^\rho p_l^\sigma$, which cannot arise from the (non-anomalous) ChPT Lagrangian.

which obey $s + t + u = \sum_i m_i^2$, leaving two independent parameters in accordance with eq. (4.35). Note that Mandelstams are totally symmetric in their indices, and that each can be represented in two complementary ways, since $\left(\sum_{i \in I} p_i\right)^2 = \left(\sum_{i \notin I} p_i\right)^2$ for any list of indices I due to conservation of momentum.

It is not obvious which Mandelstams to use in order to express an amplitude. Reasonably, the choice should include the square of any momentum that appears in an on-shell propagator—this is s, t and u for 4-point amplitudes in theories with three-particle vertices, such as QED and QCD—but more subtle conditions arise when considering *deorbiting*, which I will cover next.

Deorbiting. Consider an expression $\mathcal{X}$ that is invariant under a group $\mathcal{G}$ (in practice, $\mathcal{X} = \mathcal{M}_R$ and $\mathcal{G} = \mathbb{Z}_R$). There always exists some expression $\mathcal{X}'$ such that

$$\mathcal{X} = \mathcal{G}\mathcal{X}' \equiv \sum_{g \in \mathcal{G}} g\mathcal{X}', \qquad (4.39)$$

since this is trivially true for $\mathcal{X}' = \mathcal{X}/|\mathcal{G}|$, where $|\mathcal{G}|$ is the number of distinct $g \in \mathcal{G}$. One can often find an $\mathcal{X}'$ that is much shorter than $\mathcal{X}$—this was done by hand for the massless LO ChPT amplitude with $n \leq 10$ in ref. [47], and $n \leq 12$ in paper I—but in general, it is not feasible to do this without a more systematic method. Such a method, called *deorbiting* for reasons that will be explained shortly, was developed in paper III and used to condense the main results. Here, I will give a more comprehensive description and a novel proof of its optimality.

First, I must specify what I mean by "shorter expression", since the same expression can be written in a multitude of ways. We must assume that we have chosen a set $\mathcal{T}$ of *terms*, which can consist of masses, LECs, Mandelstams, loop integrals, propagators, and all possible products thereof. Importantly, no term may be a linear combination of other terms. Having chosen $\mathcal{T}$, any expression $\mathcal{X}$ that can be written in terms of $t \in \mathcal{T}$ will, after all products have been expanded out, take the well-defined form of a vector in the space spanned by $\mathcal{T}$. The *length* of the expression, $|\mathcal{X}|$, is then simply the number of nonzero components in this vector, and serves as a proxy for the actual length of the printed expression.

Now, we must make an important assumption: that our choice of terms is *closed* under $\mathcal{G}$. By this, I mean that for any $t \in \mathcal{T}$ and $g \in \mathcal{G}$, we have $gt = \lambda(g, t)t'$ where $t' \in \mathcal{T}$ and $\lambda(g, t) \in \mathbb{C}$. This is not true in general; for instance, $\{s, t\}$ is closed under $\mathbb{Z}_{\{2,2\}}$ and $\{s, u\}$ under $\mathbb{Z}_{\{4\}}$, but not vice versa: t goes to $u = -s - t$ (neglecting masses) under $(p_3 \leftrightarrow p_4) \in \mathbb{Z}_{\{2,2\}}$. This assumption of closedness is crucial in order for deorbiting to work, as will be made clear below.

Let me then introduce the *orbit* of some term t under $\mathcal{G}$:

$$\mathcal{G}[t] \equiv \left\{ \frac{gt}{\lambda(g, t)} \,\middle|\, g \in \mathcal{G} \right\} \subseteq \mathcal{T}. \qquad (4.40)$$

By closedness, $\mathcal{G}[u] = \mathcal{G}[t]$ for any $u \in \mathcal{G}[t]$. Therefore, orbits do not overlap, and $\mathcal{T}$ decomposes into a union of disjoint orbits. Since all terms are linearly independent, each orbit spans a separate linear subspace, so $\mathcal{X}$ splits into a sum of vectors, one for each such subspace, and these do not mix under $\mathcal{G}$.

This brings us to *deorbiting*, which, given $\mathcal{X}$, produces an expression $\tilde{\mathcal{X}}$ such that

$$\mathcal{G}\tilde{\mathcal{X}} = \mathcal{X}\,, \qquad \text{and} \qquad |\tilde{\mathcal{X}}| \le |\mathcal{X}'| \quad \text{for all } \mathcal{X}' \text{ with } \mathcal{G}\mathcal{X}' = \mathcal{X}\,. \tag{4.41}$$

The procedure is simple:

1. Write down $\mathcal{X}$ with the terms ordered in an arbitrary yet consistent fashion.[94] Let $\tilde{\mathcal{X}} = 0$ initially.

2. Select the first term $\rho_t t$ in $\mathcal{X}$, where $t \in \mathcal{T}$ and $\rho_t \neq 0$. Compute its *orbit multiplicity* $\mu_t \equiv |\mathcal{G}|/|\mathcal{G}[t]|$, which counts the number of distinct $g \in \mathcal{G}$ that give the same gt.

3. Modify the expressions:

$$\tilde{\mathcal{X}} \to \tilde{\mathcal{X}} + \rho_t t/\mu_t\,, \qquad \mathcal{X} \to \mathcal{X} - \rho_t(\mathcal{G}t)/\mu_t\,, \tag{4.42}$$

which removes from $\mathcal{X}$ the term $\rho_t t$ and all other terms in the same orbit; meanwhile, the same terms become part of $\mathcal{G}\tilde{\mathcal{X}}$. Now, $\mathcal{X}$ projected onto the subspace spanned by $\mathcal{G}[t]$ is zero.

4. Repeat from 2 until $\mathcal{X} = 0$; we will always reach this, since we remove $\mathcal{X}$ from one subspace each iteration while leaving the others unchanged. Now $\tilde{\mathcal{X}}$ has the property eq. (4.41).

Effectively, we collapse each orbit contained in $\mathcal{X}$ into a single representative and put it in $\tilde{\mathcal{X}}$, which is why the method is called "deorbiting". It is clear that $\mathcal{G}\tilde{\mathcal{X}} = \mathcal{X}$, but to see that $\tilde{\mathcal{X}}$ is the shortest such expression, assume there is some $\mathcal{X}'$ with $|\mathcal{X}'| < |\tilde{\mathcal{X}}|$. Then there must be at least one term $t \in \tilde{\mathcal{X}}$ that is not present in $\mathcal{X}'$, but then $\mathcal{G}\mathcal{X}'$ must have zero projection onto the subspace spanned by $\mathcal{G}[t]$. But $\mathcal{X}$ has nonzero projection onto this subspace, or else it would not have contained $\rho_t t$, so $\mathcal{G}\mathcal{X}' \neq \mathcal{X}$. This completes the proof.

Deorbiting is fast and systematic, and is optimal within its scope, but there are a few caveats. As mentioned above, all products must be expanded out, so deorbiting cannot account for factoring—the four-term expression $(a + b)(c + d)^2$ must be treated as the six-term expression $ac^2 + 2acd + ad^2 + bc^2 + 2bcd + bd^2$, which might throw off the notion of "shortest expression". This is the case in paper III, where a seemingly better choice of $\mathcal{G}$ ($\mathbb{Z}_R^{+\mathrm{TR}}$ instead of $\mathbb{Z}_R$) resulted in longer expressions after deorbiting and factoring. Furthermore, there is no guarantee that there is no different choice of $\mathcal{T}$ that allows even shorter expressions.

The most restrictive caveat, however, is that deorbiting requires $\mathcal{X}$ to be closed under $\mathcal{G}$. To see that this requirement is indeed necessary, consider the case where gt may be a linear combination of multiple terms. Then the eq. (4.40) fails, and we must resort to a definition like $\mathcal{G}[t] \equiv \{gt \mid g \in \mathcal{G}\}$, which is not necessarily a subset of $\mathcal{T}$. Then the separation into subspaces breaks down, and everything that follows from it.

As a concrete example of the failure not just of the *proof* of deorbiting, but also of its *applicability* (and that of any straightforward improvement of it), consider a simple alternating

---

[94] One way is to, for whatever textual representation one chooses, sort the terms alphabetically. In practice, we implement deorbiting using FORM [64], so we use FORM's internal ordering of the terms.

group $\mathcal{G} = \{1, g\}$ with $g^2 = 1$ and four terms $\mathcal{T} = \{s, t, u, v\}$ that transform like

$$gs = \frac{t + u}{\sqrt{2}}, \quad gt = \frac{s + v}{\sqrt{2}}, \quad gu = \frac{s - v}{\sqrt{2}}, \quad gv = \frac{t - u}{\sqrt{2}}. \qquad (4.43)$$

The subspaces they span are no longer separate: $\mathcal{G}[t]$ is contained in the span of $\mathcal{G}[s] \cup \mathcal{G}[v]$. And even with a simple expression like $\mathcal{X} = \mathcal{G}v$, where $\tilde{\mathcal{X}} = v$ obviously satisfies eq. (4.41), how do we find it? After all, $\mathcal{X} = (t - u)/\sqrt{2} + v$ (assuming alphabetic ordering), and if we apply the algorithm above starting at t, it never terminates. One can imagine trying different starting terms one by one until something works, but that is foiled by even simple linear combinations like $\alpha(\mathcal{G}t) + \beta(\mathcal{G}v)$, where the coefficient in front of t will not be equal to α. Thus, something radically different (and unknown to me) would be needed to achieve deorbiting-like results without the assumption of closedness.

Closed bases. Let me then conclude this section with a consideration of closedness. For $R = \{n\}$, the set of all Mandelstams with $n/2$ or fewer consecutive indices (e.g., $\hat{s}_{123}$, $\hat{s}_{456}$ and $\hat{s}_{n12}$ but not $\hat{s}_{135}$ or $\hat{s}_{12\cdots n}$) forms a basis (assuming eq. (4.35) holds) that is closed under $\mathbb{Z}_R$. For $n = 4$, this basis is $\{s, u\}$.

It remains an open question how to find a closed Mandelstam basis for arbitrary $\mathbb{Z}_R$, or if one even exists. Appendix C of paper III presents closed bases for all $\mathbb{Z}_R$ at $n = 6$ (as does appendix D of paper I, but with errors). Using brute-force searches, I have also found closed bases (not published) for $\mathbb{Z}_{\{2,6\}}$, $\mathbb{Z}_{\{2,2,4\}}$ and $\mathbb{Z}_{\{2,2,3,3\}}$, among others, but not for, e.g., $\mathbb{Z}_{\{2,3,3\}}$. No basis is closed under $\mathcal{S}_n = \mathbb{Z}_{\{1,\dots,1\}}$ [both $\{s, t, u\}$ for $n = 4$ and the analogous set $\{\hat{s}_{ijk} \mid i < j < k\}$ for $n = 6$ have one more element than the N_{dof} given by eq. (4.35), and the gap grows with n], so unless some relevant property can be found that distinguishes those $\mathbb{Z}_R$ for which R does not contain 1, a general proof of existence is out of the question. For $\mathbb{Z}_{\{2,2,2,2\}}$ and many $n > 8$ cases, N_{dof} is smaller than the number of distinct squared momenta that can appear in propagators; this does not preclude the existence of a closed basis, but makes it less convenient.

In conclusion, the kinematic situation is trivial for $n = 4$ and essentially solved for $n = 6$, while there are unsolved issues for $n \geq 8$. However, as papers III-V show, plenty of interesting amplitude work remains for $n = 6$.

5 The lattice and its ChPT connections

> *In a universe that is all gradations of matter, from gross to fine to finer, so that we
> end up with everything we are composed of in a lattice, a grid, a mesh, a mist,
> where particles or movements so small we cannot observe them are held in a strict
> and accurate web, that is nevertheless nonexistent to the eyes we use for ordinary
> living…*
>
> — *Doris Lessing, from "Canopus in Argos" (1979-1983)*

ChPT is a fine framework for low-energy QCD, and the bulk of this thesis rests on its
shoulders. However, there is another one, which is perhaps even more powerful and which
lurks in the background of papers I-III before approaching center stage in papers IV and V:
lattice QCD. In a nutshell, lattice QCD is simply the QCD version of a basic idea that oc-
curs all over physics: when some system cannot be solved exactly and cannot be reasonably
approximated by something that can (e.g., through perturbation theory), formulate a dis-
crete version of it and simulate it numerically on a computer. This is easier said than done,
of course; here, I will only give a rather shallow introduction of lattice QCD before turning
to the aspects most relevant to papers IV-V.

Lattice QCD was first proposed by Wilson in 1974 [84], shortly after QCD was shown to
be a viable theory for the strong interaction.[95] It took some theoretical and algorithmic
developments before actual computations could be made, with initial breakthroughs in
1979 by Creutz and others: first a minimalistic proof-of-concept gauge theory [87] and then
something resembling QCD [88].[96] Even the earliest results were qualitatively successful in
explaining hard-to-study phenomena such as confinement, and quantitatively reasonable
lattice measurements of hadron masses were achieved in 1981 [89, 90], but for a long time,
insufficiently sophisticated methods and computers meant lattice measurements were less
precise and reliable than those performed through experiments. In recent years, however,
lattice QCD has begun to be consistently competitive with experimental measurements
over an ever-wider range of observables (see, for instance, the recent measurement of the
muon anomalous magnetic moment [91]). Much of the state of the art is covered in the
FLAG review [92].

5.1 Overview of Lattice QCD

Here, I will quickly summarize the basics of lattice QCD, which is more comprehensively
covered in notes such as those of Gupta [93] and Hoelbling [94]. Like so much else, lattice
QCD has as its central goal to compute expectation values of observables $\hat{O}$, for which the

---

[95]Although Wilson's paper is recognized as seminal, similar ideas were arrived at independently by others. See
ref. [85] for a comprehensive history of lattice QCD, as well as Wilson's own account [86].

[96]In particular, ref. [88] remains a very good read: it gives a clear summary of the theory, and then a detailed
description of lattice simulations that are simple enough to easily reproduce on a modern laptop, yet sufficiently
sophisticated to demonstrate confinement.

path-integral formalism is adopted (recall section 2.3):

$$\langle 0|\hat{O}|0\rangle = \frac{1}{Z} \int \mathcal{D}G_\mu \mathcal{D}\bar{q}\mathcal{D}q \; \hat{O} \; \exp\left[i \int \mathrm{d}^4x \; \mathcal{L}_{\mathrm{QCD}}\right]. \tag{5.1}$$

(I omit quark flavors for simplicity.) To make this integral computationally feasible, we discretize space into a four-dimensional grid (the eponymous lattice) with spacing a, and confine it to a finite volume, typically a hypercube of side L but sometimes with a longer time dimension. To avoid boundary effects, periodic boundary conditions are typically imposed. If everything is done correctly, the discretized result will be a good approximation of the *continuum limit* $a \to 0$.

The finite lattice spacing automatically provides an ultraviolet (UV) cutoff, and the finite volume an infrared (IR) cutoff. Imposing a lattice breaks Lorentz invariance, but that is unavoidable. Additionally, a provides a unit in terms of which all other dimensionful quantities can be measured; this is called *lattice units*.

The fields must be replaced by discretized versions obtained by averaging over the lattice cells, the details of which I omit here. The discretized (anti)quark fields have the same properties as in the continuum, but live only on the lattice points. For gauge invariance, the gluon field G_μ must be discretized in the same way as ∂_μ, which becomes a finite difference operator between neighboring lattice points. Thus, discretized gluons live on the links joining neighboring lattice points, i.e.,

$$G_\mu(x) = U(x, x + a\hat{\mu}) = U^\dagger(x + a\hat{\mu}, x) \in \mathrm{SU}(3), \tag{5.2}$$

borrowing notation from the introduction to gauge theory on pages 19 and 20. With the transformation properties of eq. (2.22), the discretized Lagrangian becomes gauge-invariant. Gauge-invariant objects also include operators like

$$\bar{q}(x)G_\mu(x)q(x + a\hat{\mu}) \quad \text{and} \quad P_{\mu\nu}(x) \equiv \left\langle G_\mu(x)G_\nu(x + a\hat{\mu})G_\mu^\dagger(x + a\hat{\nu})G_\nu^\dagger(x)\right\rangle, \tag{5.3}$$

or in general any chain of gluons that is either terminated by a $\bar{q}q$ pair or forms a closed loop. Ordinary loops are called *Wilson loops*, with the *plaquette* $P_{\mu\nu}$ being the simplest example, while loops that wrap around the periodic boundary conditions are called *Polyakov loops*; see fig. 6.

Consider now the QCD Lagrangian in the form

$$\mathcal{L}_{\mathrm{QCD}} = G_{\mu\nu}G^{\mu\nu} + \bar{q}Fq, \qquad F \equiv i\gamma^\mu[\partial_\mu + gG_\mu] - m, \tag{5.4}$$

which makes it manifest that the $\bar{q}q$ integral is just a Gaussian integral, leaving[97]

$$\langle 0|\hat{O}|0\rangle = \frac{1}{Z} \int \mathcal{D}G_\mu \mathcal{D}\bar{q}\mathcal{D}q \; \hat{O} \; \det(F) \; \exp\left[i \int \mathrm{d}^4x \; G_{\mu\nu}G^{\mu\nu}\right]. \tag{5.5}$$

Despite this simplification, the extremely non-local $\det(F)$, which depends on all lattice points simultaneously, complicates further progress. A simple solution called *quenching*

---

[97]This works also in the continuum (see, e.g., Peskin & Schroeder [14, chapter 9] for an introduction) but is significantly less subtle on the lattice: due to the finite number of degrees of freedom, F is just a (potentially very large) finite matrix that depends of G_μ. Note that some manipulation will be needed first if $\hat{O}$ contains $\bar{q}q$.
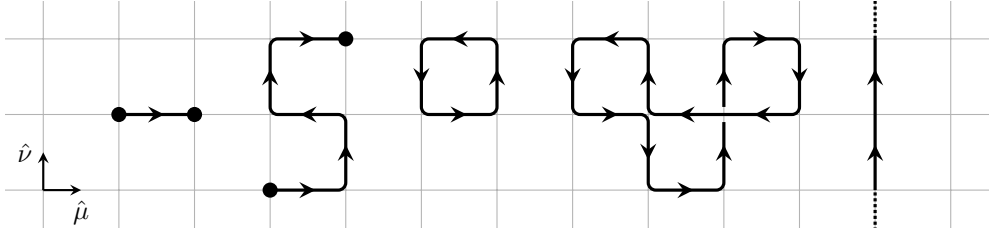
Figure 6: Examples of gauge-invariant operators on a two-dimensional lattice, with (anti)quarks represented as dots and links (gluons) as directed line segments. From left to right: $\bar{q}G_\mu q$, a more complicated $\bar{q}\ldots q$ chain, a plaquette $P_{\mu\nu}$, a more complicated Wilson loop, and part of a Polyakov loop wrapping around in the $\hat{\nu}$ direction.

is to take $\det(F)$ to be constant, making the quarks non-dynamic. More sophisticated approaches are frustrated by the *Nielsen–Ninomiya no-go theorem* [95], which states that any way of putting fermions on a lattice will have at least one of the following undesirable properties:

1. the incorrect continuum limit,
2. a non-local fermion operator,
3. violation of chiral symmetry,
4. spurious extra fermion modes.

The last two are survivable if handled correctly, as is done by the several different methods used in modern lattice QCD; I will not cover them here.

Assuming the quarks are taken care of some way or other, the gluon integral remains challenging; as remarked by Creutz [96, sec. 8], even tiny lattices like those he used in ref. [88] would take much longer than the age of the Universe to integrate over naïvely, and furthermore, the complex exponential e^{iS} causes oscillations that would be numerically hard even with few degrees of freedom. The solution is to perform a Wick rotation, $t \to it$, so that e^{iS} becomes the more well-behaved e^{-S}. This prevents the lattice from directly accessing real-time physics, but it is a good trade-off nonetheless: it still allows masses and may other things to be measured, and it means that

$$\frac{1}{Z} \int \mathcal{D}G_\mu \mathcal{D}\bar{q}\mathcal{D}q \; \hat{O} \; e^{-S} \tag{5.6}$$

has exactly the same form as a correlation function in thermodynamics, with S playing the role of βE. Thus, the Wick rotation puts lattice QCD in a class of numerical problems that also include the Ising model and its variations, and which are amenable to *Monte Carlo integration*: the integral is approximated by an average over a set of configurations of the gluon field randomly drawn from the probability distribution proportional to e^{-S}. Such a set is straightforward to obtain by means such as the *Metropolis algorithm*, although obtaining a sufficiently large set of configurations on a sufficiently large lattice is extremely computationally intensive.

The computational difficulty tends to rise dramatically the lighter the lightest stable degree of freedom is (typically, the pion), and while calculations using the physical pion mass (i.e.,

the real-world value) are now commonly done, they remain something of a luxury; it is often easier to perform a series of calculations with heavier masses and then extrapolate to the physical one. This is one of many places where ChPT synergizes with the lattice: relating a lattice result to one calculated in ChPT, where M_π is just a parameter, gives a more reliable connection to the physical mass than that provided by fitting a curve to the data. In the context of this thesis, this forms a central part of papers IV-V, along with the topic of the following sections.

5.2   The finite-volume quantization condition

This area is the portal between lattice QCD and my ChPT work, and forms the background for papers IV and V. Interestingly, it is a variant of the system that is often used as a "Hello, world!" calculation in introductory quantum physics courses: the particle in a box. Here, the box is a lattice of side L, assumed to be fine enough that discretization effects are of no relevance to the analysis, and it contains two or more stable hadrons. (With electroweak interactions are turned off, the pions, kaons, nucleons, Σ, Λ and Ξ are stable.) In the infinite-volume ($L \to \infty$) limit, the ground-state energy of such a system will simply be the sum of the masses of the particles, but the finite volume will induce a shift in the ground-state energy and give rise to a discrete spectrum of excited states.

The key idea here is that this finite-volume n-particle spectrum is related to the elastic n-to-n scattering amplitude in the infinite volume. The simplest case, namely the ground-state energy $E_0(L)$ for two identical particles, was derived long before the establishment of lattice QCD [97]:

$$E_0(L) = 2M + \frac{4\pi a_{\pi\pi}}{(2\pi L)^3} + \mathcal{O}(L^{-4}) \,, \tag{5.7}$$

where M is the particle mass and a the *scattering length*, related to the elastic 2-to-2 scattering cross-section in the limit of zero momentum transfer through $4\pi a_{\pi\pi}^2 = \lim_{q\to 0} \sigma(q^2)$.

Lüscher [98–100] gave a more general formula for the entire two-particle spectrum at all orders in L, the *finite-volume quantization condition*, which has subsequently been extended to a large number of more general cases; see ref. [101] for a comprehensive review, and ref. [102] for one focusing on cases closer to papers IV-V. Importantly, finite-volume spectra can be measured on the lattice, while the inputs to quantization conditions are infinite-volume amplitudes that can be calculated using continuum theories like ChPT, granted that they cover the same physical regime as the lattice measurements.

I will now give a greatly simplified derivation of Lüscher's formula, in the case of a theory where all particles have the same mass M, based mostly on the more thorough one given in ref. [102, sec. 2.1.3]. Let us first consider $\mathcal{M}_2$, the elastic 2-to-2 scattering amplitude. In the center-of-momentum (CM) frame, it depends only on the total energy E^* (in general, a star indicates a quantity taken in the CM frame) and the angles between the external momenta, which can be decomposed over spherical harmonics:

$$\mathcal{M}_2 = \sum_{\ell'm',\ell m} 4\pi Y^*_{\ell'm'}(\hat{\boldsymbol{p}}'^*) \mathcal{M}_2(E^*)_{\ell'm',\ell m} Y_{\ell m}(\hat{\boldsymbol{p}}^*) \,, \tag{5.8}$$

Here, $\pm\boldsymbol{p}^*$ and $\pm\boldsymbol{p}'^*$ are the in- and outgoing momenta respectively, which are back-to-back in the CM frame. In the following, everything will be decomposed in a similar way, and I will leave the ℓm indices implicit.

Now, consider $\mathcal{M}_2$ to be computed using Feynman diagrams, to all orders in perturbation theory, and let $\mathcal{M}_{2,L}$ be the same diagrams evaluated in the finite volume. $\mathcal{M}_{2,L}$ remains physically valid, and its poles will give the finite-volume spectrum, so the goal here is to express the pole positions of $\mathcal{M}_{2,L}$ in terms of $\mathcal{M}_2$.

Since all finite-volume momenta k^μ are constrained by the periodic boundary conditions to $k^\mu = \frac{1}{2\pi L}n^\mu$, where n^μ is a vector of integers, all loop integrals in $\mathcal{M}_2$ will be discrete sums in $\mathcal{M}_{2,L}$. For continuous integrands, the Poisson integration formula shows that the sum and integrand differ only by exponentially suppressed terms, $\mathcal{O}(e^{-ML})$, which I discard throughout this derivation.[98] Thus, the non-suppressed finite-volume effects can only come from singular integrands, i.e., loops containing propagators that can go on-shell. Assuming a center-of-mass energy of $2M \leq E^* < 4M$, so that no more than two particles can simultaneously go on-shell (even-to-odd transitions are assumed to be forbidden), we can therefore express either amplitude diagramatically as

$$\mathcal{M}_{2(,L)} \equiv \text{[diagram]} = \text{[diagram]} \cdots \text{[diagram]} = \text{[diagram]} + \text{[diagram]} \,, \quad (5.9)$$

where [diagram] represents the amplitude and [diagram] represents a *Bethe–Salpeter* (BS) *kernel*, B_2, defined as the sum of all 2-to-2 diagrams (including no-scattering diagrams) that have no two-particle cuts. In other words, no propagator inside a BS kernel can go on-shell when $2M \leq E^* < 4M$, so its loop integrands are free of singularities. By the arguments above, all integrals involved in calculating the BS kernels can be ignored; only the two-particle loops connecting the kernels have singular integrands and give rise to differences between $\mathcal{M}_2$ and $\mathcal{M}_{2,L}$ that are not $\mathcal{O}(e^{-ML})$.

In the last equality of eq. (5.9), I have rephrased it as a recursive relation; symbolically,

$$\mathcal{M}_2 = B_2 + B_2 \otimes_I \mathcal{M}_2 \,, \qquad \mathcal{M}_{2,L} = B_2 + B_2 \otimes_S \mathcal{M}_{2,L} \,, \qquad (5.10)$$

where $\otimes_I$ and $\otimes_S$ represent the integral or sum, respectively, over the two-particle loop. (Bear in mind that all products involve contraction of the ℓm indices.) The sum-integral

---

[98]The Poisson integration formula states that $\sum_{\boldsymbol{k} \in \mathbb{Z}^d} f(\boldsymbol{k}) = \sum_{\boldsymbol{x} \in \mathbb{Z}^d} \tilde{f}(\boldsymbol{x})$ for any continuous function f with Fourier transform $\tilde{f}$, where $\mathbb{Z}^d$ is a d-dimensional lattice of integer points. The Fourier transform involves an integral over an exponential $e^{-2\pi i \boldsymbol{x} \cdot \boldsymbol{k}}$, which after Wick rotation is dominated by a falling real exponential; only the $\boldsymbol{x} = 0$ term is unaffected. Describing position and momentum coordinates as multiples of L and M and normalizing appropriately, we thus obtain the well-known formula

$$\frac{1}{L^3} \sum_{\boldsymbol{k}} f(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 \boldsymbol{k}}{(2\pi)^3} f(\boldsymbol{k}) + \mathcal{O}(e^{-ML}) \,.$$

difference can be expressed in terms of some $F_2^{i\epsilon}$,[99] which is a function of L and the kinematics, but (importantly) does not depend on the field content of the theory, and fixes all momenta connected to it on-shell (see refs. [102, 104] for derivations of the latter fact):

$$B_2 \otimes_S \mathcal{M}_{2,L} = B_2 \otimes_I \mathcal{M}_{2,L} + B_2 F_2^{i\epsilon} \mathcal{M}_{2,L} \,, \tag{5.11}$$

We can now insert eq. (5.11) into eq. (5.10) and then eq. (5.10) into itself,

$$\mathcal{M}_{2,L} = B_2 + B_2[\otimes_I + F_2^{i\epsilon}]\mathcal{M}_{2,L} = \sum_{a=0}^{\infty} B_2\Big([\otimes_I + F_2^{i\epsilon}]B_2\Big)^a$$

$$= \sum_{a=0}^{\infty} \left(\sum_{b=0}^{\infty} B_2[\otimes_I B_2]^b\right)\left(F_2^{i\epsilon} \sum_{b=0}^{\infty} B_2[\otimes_I B_2]^b\right)^a \,, \tag{5.12}$$

where we identify the two b-sums as $\mathcal{M}_2$. This leaves a geometric series that can be summed, giving

$$\mathcal{M}_{2,L} = \mathcal{M}_2 \sum_{a=0}^{\infty}[F_2^{i\epsilon}\mathcal{M}_2]^a = \mathcal{M}_2\big[1 - F_2^{i\epsilon}\mathcal{M}_2\big]^{-1} \,. \tag{5.13}$$

The poles of $\mathcal{M}_{2,L}$—that is, the finite-volume spectrum—are obtained when $[1 - F_2^{i\epsilon}\mathcal{M}_2]$ is singular. This leads to Lüscher's quantization condition:

$$\det\left[F_2^{i\epsilon}(E, \boldsymbol{P}, L) + \mathcal{M}_2^{-1}(E^*)\right] = 0 \,, \tag{5.14}$$

with the determinant taken over the ℓm indices. Note that $F_2^{i\epsilon}$, unlike $\mathcal{M}_2$, is not Lorentz invariant, so it also depends on the total three-momentum, $\boldsymbol{P}$.

The above form is somewhat problematic, since $\mathcal{M}_2$ possesses branch cuts. However, it can be shown based on unitarity that, above threshold,

$$\mathcal{M}_2(E^*)^{-1} = \mathcal{K}_2(E^*)^{-1} - \frac{iq^*}{16\pi E^*} \,, \tag{5.15}$$

where the *K-matrix* $\mathcal{K}_2$ is real and lacks branch cuts, and $q^* = \sqrt{\frac{1}{4}E^{*2} - M^2}$ is the magnitude of the external momenta in the CM frame. This last factor can be absorbed into $F_2^{i\epsilon}$, yielding another cut-free quantity called F_2, and giving the final form of the two-particle quantization condition:

$$\det\left[F_2^{-1}(E, \boldsymbol{P}, L) + \mathcal{K}_2(E^*)\right] = 0 \,. \tag{5.16}$$

5.3 Three-particle quantization and the K-matrix

Equation (5.16) has a direct generalization as the *three-particle* quantization condition,

$$\det\left[F_3^{-1}(E, \boldsymbol{P}, L) + \mathcal{K}_3(E^*)\right] = 0 \,. \tag{5.17}$$

---

[99] The "$i\epsilon$" refers to the regulator found in the Feynman propagator. The explicit form of $F_2^{i\epsilon}$ is stated in ref. [102, eq. (26)], but is not particularly illuminating for the purposes of the derivation. Reference [103] contains (using a simpler theory) an explicit calculation of the ground-state energy shift, and therefore deals with sum-integral differences explicitly, which can be helpful for building an intuition.

However, the structures of F_3 and $\mathcal{K}_3$ are vastly more complicated than those of F_2 and $\mathcal{K}_2$. There are several formalisms in use for defining and evaluating the three-particle quantization condition, the main ones being relativistic field theory (RFT) [105, 106], nonrelativistic field theory (NRFT) [107, 108] and finite-volume unitarity (FVU) [109, 110]. Papers IV-V use only RFT, and I will do the same here; see ref. [102, sec. 3] for a RFT-centric overview of NRFT and FVU.

Importantly, the three-particle K-matrix is (unlike the amplitude) a smooth function free from poles and cuts; to signify this, it is written $\mathcal{K}_{\mathrm{df},3}$ for "divergence-free". This makes it feasible to extract it from finite-volume spectra measured on the lattice, and on the practical side, the divergence-free condition provides guidance in the calculation of $\mathcal{K}_{\mathrm{df},3}$ performed in papers IV-V. The derivation of F_3 and $\mathcal{K}_{\mathrm{df},3}$ is unfortunately too complicated for me to describe here; see ref. [102, sec. 2.2] for a summary of the full derivation [105, 106]. The conceptual approach is similar to that in section 5.2, but involves both 2-to-2 and 3-to-3 BS kernels. Unavoidably, F_3 depends on the 2-to-2 (but not 3-to-3) amplitude, unlike F_2 which is a purely geometric quantity. Likewise, the relation between $\mathcal{M}_3$ and $\mathcal{K}_{\mathrm{df},3}$ is not an algebraic relationship like eq. (5.15), but is an integral equation.

As shown in paper IV, the $\mathcal{M}_3$-$\mathcal{K}_{\mathrm{df},3}$ relation simplifies when all quantities are restricted to NLO in the ChPT power counting. This allows $\mathcal{K}_{\mathrm{df},3}$ to be calculated by subtracting from $\mathcal{M}_3$ a small number of expressions, which depend on $\mathcal{M}_2$ and cancel the different divergences of $\mathcal{M}_3$. This procedure is well described in sections 2.2 to 2.5 of paper IV, and I will avoid repeating myself here; papers VIII and IX contain brief summaries of the calculation.

Let me instead conclude by mentioning the main bigger-picture motivation behind papers IV-V: the study of hadron resonances. Unlike light pseudoscalar mesons, which are stable in the absence of weak interactions, resonances retain their extremely large decay widths on the lattice. Therefore, like in ordinary particle phenomenology, they are necessarily studied in terms of their decay products. Scattering theory assumes these to be well-separated in space; the lattice equivalent requires determining the volume-dependence of the finite-volume spectrum so that the $L \to \infty$ limit can be reconstructed.

Several of the most interesting resonances have prominent three-particle decay channels, so accurate lattice simulations thereof require a reliable three-particle finite-volume formalism. These resonances include the relatively narrow vector meson $\omega(782)$ seen in section 2.1, the broad axial-vector meson $a_1(1260)$, and the *Roper resonance* $N(1440)$ [111], a doubly excited nucleon state that is (somewhat puzzlingly) lighter than the singly excited state, $N(1535)$. The 3π decay modes of $\omega(782)$ and $a_1(1260)$ are in principle covered by paper V, although the results of that paper only apply to energies lower than $5M_\pi$, whereas the $\pi\pi N$ modes (N being any nucleon) of the Roper resonance lie significantly further down the road. Nevertheless, the work I present in papers IV-V is an important step towards gaining the ability to understand these states from first principles.

6 Summary and outlook

We do not yet know how to get correct answers, but we begin to understand which are the right questions to ask.

— *Giorgio Parisi (1977)*

This concludes the introduction to this thesis. I hope that by now, the connection between my papers and the wider context has been clarified a little. Paper IV and its sequel, paper V, unites the RFT formalism of ref. [105, 106] with the results of paper III and its predecessor, ref. [6], to produce results that are applicable to the study of three-pion systems, and therefore interesting resonances, in lattice QCD. Paper III, in turn, extends the six-pion amplitude [6] in the style of earlier four-meson work [27] while incorporating methods from paper I (itself stemming from amplitude methods via ref. [47]) to alleviate the difficulties of multi-meson scattering. Lastly, paper II stands a bit to the side, but has numerous underlying connections with the other work, as seen at various points throughout the introduction.

None of the topics covered by the papers are by any means done. Karol Kampf and collaborators [82] have continued to develop methods adjacent to DFO, and I maintain a sporadic exchange of ideas with them. There have been numerous recent developments (e.g., refs. [112–114]) in the area of paper II, with varying scopes and methods, some of which could be combined with those of the paper for further improvements. As mentioned at the end of the previous section, papers IV and V represent two steps on a ladder with may rungs left, with the inclusion of kaons, η and nucleons being possible extensions. Most of these have a corresponding extension of paper III as a prerequisite, and there is of course the separate possibility to push the amplitude to NNLO or to even more external particles. While the phenomenological relevance of an eight-meson amplitude is far smaller than that of a six-meson one, the tour-de-force calculation it would require would be interesting in itself.

I have learned incredibly much during the writing of this thesis and the papers it contains, and along the way I have gathered many useful tools—DFO, linear constraints, deorbiting, the numerous techniques in papers IV-V, and so on—that I am sure will find purchase on many problems beyond those they were originally constructed for. Following the conclusion of this Ph.D., I have every intent to carry this toolbox along to help me with the new, exciting challenges I will face.

References

[1] Thorén, Johan. *Multiplet bases, recursion relations and full color parton showers* PhD thesis (Lund University, oct 2018). ISBN: 978-91-7753-854-7.

[2] Hermansson Truedsson, Nils. *Higher Order Calculations for Low Energy Precision Physics* PhD thesis (Lund University, 2019). ISBN: 978-91-7895-214-4.

[3] Particle Data Group, Workman, R. L. *et al.* "Review of Particle Physics". *PTEP* **2022,** 083C01 (2022).

[4] Vitos, Timea. *Precision Standard Model Phenomenology for High Energy Processes* PhD thesis (Lund University, may 2023). ISBN: 978-91-8039-667-7.

[5] Aoyama, T. *et al.* "The anomalous magnetic moment of the muon in the Standard Model". *Phys. Rept.* **887,** 1–166. arXiv: 2006.04822 [hep-ph] (2020).

[6] Bijnens, J. & Husek, T. "Six-pion amplitude". *Phys. Rev. D* **104,** 054046. arXiv: 2107.06291 [hep-ph] (2021).

[7] Osborn, H. "Implications of Adler zeros for multipion processes". *Lett. Nuovo Cim.* **2,** 717–723 (1969).

[8] Susskind, L. & Frye, G. "Algebraic aspects of pionic duality diagrams". *Phys. Rev. D* **1,** 1682–1686 (1970).

[9] Cornwell, D. T. "The six pion amplitude to fourth order in momenta". *Nucl. Phys.* **B34,** 125–135 (1971).

[10] Seiferle, B. *et al.* "Energy of the ^{229}Th nuclear clock transition". *Nature* **573,** 243–246 (2019).

[11] Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. & Hudson, R. P. "Experimental Test of Parity Conservation in β Decay". *Phys. Rev.* **105,** 1413–1414 (1957).

[12] Griffiths, D. *Introduction to elementary particles* ISBN: 0-471-60386-4 (John Wiley & Sons, 1987).

[13] Blass, A. "Existence of bases implies the axiom of choice". *Contemporary mathematics* **31,** 31–33 (1984).

[14] Peskin, M. E. & Schroeder, D. V. *An Introduction to quantum field theory* ISBN: 978-0-201-50397-5 (Addison-Wesley, Reading, USA, 1995).

[15] Zee, A. *Quantum field theory in a nutshell* ISBN: 978-0-691-14034-6 (Princeton University Press, 2003).

[16] Pich, A. *Effective Field Theory with Nambu-Goldstone Modes* in *Les Houches summer school: EFT in Particle Physics and Cosmology* (2018). arXiv: 1804.05664 [hep-ph].

[17] Scherer, S. & Schindler, M. R. *A Primer for Chiral Perturbation Theory* 1st ed. ISBN: 978-3-642-19253-1 (Springer-Verlag, 2012).

[18] Donoghue, J. F., Golowich, E. & Holstein, B. R. *Dynamics of the Standard Model: Second edition* ISBN: 978-1-00-929103-3 (Cambridge University Press, 2014).

[19] Lifson, Andrew. *The Chirality-Flow Formalism and Optimising Scattering Amplitudes* PhD thesis (Lund University, mar 2023). ISBN: 978-91-8039-589-2.

[20] Jackson, J. D. & Okun, L. B. "Historical roots of gauge invariance". *Rev. Mod. Phys.* **73**, 663–680. arXiv: hep-ph/0012061 (2001).

[21] Romão, J. C. & Silva, J. P. "A resource for signs and Feynman diagrams of the Standard Model". *Int. J. Mod. Phys. A* **27**, 1230025. arXiv: 1209.6213 [hep-ph] (2012).

[22] Bijnens, J. "Chiral perturbation theory beyond one loop". *Prog. Part. Nucl. Phys.* **58**, 521–586. arXiv: hep-ph/0604043 (2007).

[23] Bijnens, J. & Hermansson Truedsson, N. "The Pion Mass and Decay Constant at Three Loops in Two-Flavour Chiral Perturbation Theory". *JHEP* **11**, 181. arXiv: 1710.01901 [hep-ph] (2017).

[24] Dixon, L. J., Gurdogan, O., McLeod, A. J. & Wilhelm, M. "Bootstrapping a stress-tensor form factor through eight loops". *JHEP* **07**, 153. arXiv: 2204.11901 [hep-th] (2022).

[25] Andersson, B. *The Lund model* ISBN: 978-0-521-01734-3 (Cambridge University Press, July 2005).

[26] Fierz, M. "Zur fermischen theorie des β-zerfalls". *Zeitschrift für Physik* **104**, 553–565 (1937).

[27] Bijnens, J. & Lu, J. "Meson-meson Scattering in QCD-like Theories". *JHEP* **03**, 028. arXiv: 1102.0172 [hep-ph] (2011).

[28] Gell-Mann, M. & Lévy, M. "The axial vector current in beta decay". *Nuovo Cim.* **16**, 705 (1960).

[29] Weinberg, S. "Phenomenological Lagrangians". *Physica A* **96**, 327–340 (1979).

[30] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory to One Loop". *Annals Phys.* **158**, 142 (1984).

[31] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark". *Nucl. Phys. B* **250**, 465–516 (1985).

[32] Catà, O. & Mateu, V. "Chiral perturbation theory with tensor sources". *JHEP* **09**, 078. arXiv: 0705.2948 [hep-ph] (2007).

[33] Donoghue, J. F. & Leutwyler, H. "Energy and momentum in chiral theories". *Z. Phys. C* **52**, 343–351 (1991).

[34] Ecker, G., Gasser, J., Pich, A. & de Rafael, E. "The Role of Resonances in Chiral Perturbation Theory". *Nucl. Phys. B* **321**, 311–342 (1989).

[35] Bijnens, J., Hermansson-Truedsson, N. & Wang, S. "The order p^8 mesonic chiral Lagrangian". *JHEP* **01**, 102. arXiv: 1810.06834 [hep-ph] (2019).

[36] Bijnens, J., Colangelo, G. & Ecker, G. "The Mesonic chiral Lagrangian of order p^6". *JHEP* **02**, 020. arXiv: hep-ph/9902437 (1999).

[37] Fearing, H. W. & Scherer, S. "Extension of the chiral perturbation theory meson Lagrangian to order p^6". *Phys. Rev. D* **53**, 315–348. arXiv: hep-ph/9408346 (1996).

[38] Gráf, L., Henning, B., Lu, X., Melia, T. & Murayama, H. "2, 12, 117, 1959, 45171, 1170086,…: a Hilbert series for the QCD chiral Lagrangian". *JHEP* **01**, 142. arXiv: 2009.01239 [hep-ph] (2021).

[39] Bijnens, J. & Ecker, G. "Mesonic low-energy constants". *Ann. Rev. Nucl. Part. Sci.* **64,** 149–174. arXiv: 1405.6488 [hep-ph] (2014).

[40] Wess, J. & Zumino, B. "Consequences of anomalous Ward identities". *Phys. Lett. B* **37,** 95–97 (1971).

[41] Witten, E. "Global Aspects of Current Algebra". *Nucl. Phys. B* **223,** 422–432 (1983).

[42] Bijnens, J., Girlanda, L. & Talavera, P. "The Anomalous chiral Lagrangian of order p^6". *Eur. Phys. J. C* **23,** 539–544. arXiv: hep-ph/0110400 (2002).

[43] Ebertshauser, T., Fearing, H. W. & Scherer, S. "The Anomalous chiral perturbation theory meson Lagrangian to order p^6 revisited". *Phys. Rev. D* **65,** 054033. arXiv: hep-ph/0110261 (2002).

[44] Bijnens, J., Hermansson-Truedsson, N. & Ruíz-Vidal, J. "The order p^8 anomalous mesonic chiral Lagrangian". *(In preparation).*

[45] Eichler, M. "A new proof of the Baker-Campbell-Hausdorff formula". *Journal of the Mathematical Society of Japan* **20,** 23–25 (1968).

[46] Achilles, R. & Bonfiglioli, A. "The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin". *Archive for history of exact sciences* **66,** 295–358 (2012).

[47] Kampf, K., Novotný, J. & Trnka, J. "Tree-level Amplitudes in the Nonlinear Sigma Model". *JHEP* **05,** 032. arXiv: 1304.3048 [hep-th] (2013).

[48] Frobenius, H. "Über lineare Substitutionen und bilineare Formen". *Journal für die reine und angewandte Mathematik (Crelles Journal)* **84,** 1–63 (1878).

[49] Sasaki, C. "Parity doubling of baryons in a chiral approach with three flavors". *Nucl. Phys. A* **970,** 388–397. arXiv: 1707.05081 [hep-ph] (2018).

[50] Bijnens, J., Colangelo, G., Ecker, G., Gasser, J. & Sainio, M. E. "Pion-pion scattering at low energy". *Nucl. Phys. B* **508.** [Erratum: Nucl.Phys.B **517,** 639–639 (1998)], 263–310. arXiv: hep-ph/9707291 (1997).

[51] Bijnens, J., Colangelo, G. & Ecker, G. "Renormalization of chiral perturbation theory to order p^6". *Annals Phys.* **280,** 100–139. arXiv: hep-ph/9907333 (2000).

[52] Gomez Nicola, A. & Pelaez, J. R. "Meson meson scattering within one loop chiral perturbation theory and its unitarization". *Phys. Rev. D* **65,** 054009. arXiv: hep-ph/0109056 (2002).

[53] Mateu, V. "Universal Bounds for SU(3) Low Energy Constants". *Phys. Rev. D* **77,** 094020. arXiv: 0801.3627 [hep-ph] (2008).

[54] Chivukula, R. S., Dugan, M. J. & Golden, M. "Analyticity, crossing symmetry and the limits of chiral perturbation theory". *Phys. Rev. D* **47,** 2930–2939. arXiv: hep-ph/9206222 (1993).

[55] Jenkins, E. E., Manohar, A. V. & Wise, M. B. "Chiral perturbation theory for vector mesons". *Phys. Rev. Lett.* **75,** 2272–2275. arXiv: hep-ph/9506356 (1995).

[56] Ecker, G., Gasser, J., Leutwyler, H., Pich, A. & de Rafael, E. "Chiral Lagrangians for Massive Spin 1 Fields". *Phys. Lett. B* **223,** 425–432 (1989).

[57] Cirigliano, V. *et al.* "Towards a consistent estimate of the chiral low-energy constants". *Nucl. Phys. B* **753,** 139–177. arXiv: hep-ph/0603205 (2006).

[58] Holmberg, M. & Leupold, S. "The relativistic chiral Lagrangian for decuplet and octet baryons at next-to-leading order". *Eur. Phys. J. A* **54**, 103. arXiv: `1802.05168 [hep-ph]` (2018).

[59] Fettes, N., Meissner, U.-G. & Steininger, S. "Pion-nucleon scattering in chiral perturbation theory. 1. Isospin symmetric case". *Nucl. Phys. A* **640**, 199–234. arXiv: `hep-ph/9803266` (1998).

[60] Fettes, N. & Meissner, U.-G. "Pion-nucleon scattering in chiral perturbation theory. 2. Fourth order calculation". *Nucl. Phys. A* **676**, 311. arXiv: `hep-ph/0002162` (2000).

[61] Zou, Y.-H., Liu, H., Liu, Y.-R. & Jiang, S.-Z. "Chiral Lagrangians for singly heavy baryons to $\mathcal{O}(p^4)$ order". *Phys. Rev. D* **108**, 014027. arXiv: `2304.09183 [hep-ph]` (2023).

[62] Liu, H., Zou, Y.-H., Liu, Y.-R. & Jiang, S.-Z. "Chiral Lagrangians for spin-$\frac{1}{2}$ and spin-$\frac{3}{2}$ doubly charmed baryons". *Phys. Rev. D* **108**, 014032. arXiv: `2304.04575 [hep-ph]` (2023).

[63] Manohar, A. & Wise, M. *Heavy Quark Physics* ISBN: 978-05-2164-241-5 (Cambridge University Press, 2000).

[64] Vermaseren, J. A. M. "New features of FORM". arXiv: `math-ph/0010025` (2000).

[65] Passarino, G. & Veltman, M. J. G. "One Loop Corrections for e^+e^- Annihilation Into $\mu^+\mu^-$ in the Weinberg Model". *Nucl. Phys. B* **160**, 151–207 (1979).

[66] Mangano, M. L. "The Color Structure of Gluon Emission". *Nucl. Phys. B* **309**, 461–475 (1988).

[67] Bern, Z. & Kosower, D. A. "Color decomposition of one loop amplitudes in gauge theories". *Nucl. Phys. B* **362**, 389–448 (1991).

[68] Schuster, T. "Color ordering in QCD". *Phys. Rev. D* **89**, 105022. arXiv: `1311.6296 [hep-ph]` (2014).

[69] Reuschle, C. & Weinzierl, S. "Decomposition of one-loop QCD amplitudes into primitive amplitudes based on shuffle relations". *Phys. Rev. D* **88**, 105020. arXiv: `1310.0413 [hep-ph]` (2013).

[70] Ellis, J. R. & Renner, B. "On the relationship between chiral and dual models". *Nucl. Phys. B* **21**, 205–216 (1970).

[71] Mandelstam, S. "Determination of the Pion-Nucleon Scattering Amplitude from Dispersion Relations and Unitarity. General Theory". *Phys. Rev.* **112**, 1344–1360 (4 1958).

[72] Weinberg, S. "Pion scattering lengths". *Phys. Rev. Lett.* **17**, 616–621 (1966).

[73] Hansen, M. T., Romero-López, F. & Sharpe, S. R. "Generalizing the relativistic quantization condition to include all three-pion isospin channels". *JHEP* **07**, 047. arXiv: `2003.10974 [hep-lat]` (2020).

[74] Britto, R., Cachazo, F. & Feng, B. "New recursion relations for tree amplitudes of gluons". *Nucl. Phys. B* **715**, 499–522. arXiv: `hep-th/0412308` (2005).

[75] Britto, R., Cachazo, F., Feng, B. & Witten, E. "Direct proof of tree-level recursion relation in Yang-Mills theory". *Phys. Rev. Lett.* **94,** 181602. arXiv: `hep-th/0501052` (2005).

[76] Kampf, K., Novotny, J. & Trnka, J. "Recursion relations for tree-level amplitudes in the $SU(N)$ nonlinear sigma model". *Phys. Rev.* **D87,** 081701. arXiv: `1212.5224` `[hep-th]` (2013).

[77] Cheung, C., Kampf, K., Novotný, J., Shen, C.-H. & Trnka, J. "On-Shell Recursion Relations for Effective Field Theories". *Phys. Rev. Lett.* **116,** 041601. arXiv: `1509.` `03309` `[hep-th]` (2016).

[78] Adler, S. L. "Consistency conditions on the strong interactions implied by a partially conserved axial vector current". *Phys. Rev.* **137,** B1022–B1033 (1965).

[79] Adler, S. L. "Consistency conditions on the strong interactions implied by a partially conserved axial-vector current. II". *Phys. Rev.* **139,** B1638–B1643 (1965).

[80] Cheung, C., Kampf, K., Novotný, J., Shen, C.-H. & Trnka, J. "A Periodic Table of Effective Field Theories". *JHEP* **02,** 020. arXiv: `1611.03137` `[hep-th]` (2017).

[81] Low, I. "Double Soft Theorems and Shift Symmetry in Nonlinear Sigma Models". *Phys. Rev.* **D93,** 045032. arXiv: `1512.01232` `[hep-th]` (2016).

[82] Bartsch, C., Kampf, K. & Trnka, J. "Recursion Relations for One-Loop Goldstone Boson Amplitudes". arXiv: `2206.04694` `[hep-th]` (2022).

[83] Kaiser, R. & Leutwyler, H. "Large N_c in chiral perturbation theory". *Eur. Phys. J. C* **17,** 623–649. arXiv: `hep-ph/0007101` (2000).

[84] Wilson, K. G. "Confinement of Quarks". *Phys. Rev. D* **10** (ed Taylor, J. C.) 2445–2459 (1974).

[85] Ukawa, A. "Kenneth Wilson and lattice QCD". *J. Statist. Phys.* **160,** 1081. arXiv: `1501.04215` `[hep-lat]` (2015).

[86] Wilson, K. G. "The Origins of lattice gauge theory". *Nucl. Phys. B Proc. Suppl.* **140** (eds Bodwin, G. T. *et al.*) 3–19. arXiv: `hep-lat/0412043` (2005).

[87] Creutz, M., Jacobs, L. & Rebbi, C. "Experiments with a Gauge Invariant Ising System". *Phys. Rev. Lett.* **42,** 1390 (1979).

[88] Creutz, M. "Monte Carlo Study of Quantized SU(2) Gauge Theory". *Phys. Rev. D* **21,** 2308–2315 (1980).

[89] Weingarten, D. "Monte Carlo Evaluation of Hadron Masses in Lattice Gauge Theories with Fermions". *Phys. Lett. B* **109** (eds Julve, J. & Ramón-Medrano, M.) 57 (1982).

[90] Hamber, H. & Parisi, G. "Numerical Estimates of Hadronic Masses in a Pure SU(3) Gauge Theory". *Phys. Rev. Lett.* **47** (eds Julve, J. & Ramón-Medrano, M.) 1792 (1981).

[91] Borsanyi, S. *et al.* "Leading hadronic contribution to the muon magnetic moment from lattice QCD". *Nature* **593,** 51–55. arXiv: `2002.12347` `[hep-lat]` (2021).

[92] Flavour Lattice Averaging Group (FLAG), Aoki, Y. *et al.* "FLAG Review 2021". *Eur. Phys. J. C* **82,** 869. arXiv: `2111.09849` `[hep-lat]` (2022).

[93] Gupta, R. *Introduction to lattice QCD: Course* in *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions* (1997), 83–219. arXiv: `hep-lat/9807028`.

[94] Hoelbling, C. "Lattice QCD: concepts, techniques and some results". *Acta Phys. Polon. B* **45** (ed Praszalowicz, M.) 2143. arXiv: 1410.3403 [hep-lat] (2014).

[95] Nielsen, H. B. & Ninomiya, M. "No Go Theorem for Regularizing Chiral Fermions". *Phys. Lett. B* **105**, 219–223 (1981).

[96] Creutz, M. "Lattice gauge theory: A Retrospective". *Nucl. Phys. B Proc. Suppl.* **94** (eds Bhattacharya, T., Gupta, R. & Patel, A.) 219–226. arXiv: `hep-lat/0010047` (2001).

[97] Huang, K. & Yang, C. N. "Quantum-mechanical many-body problem with hard-sphere interaction". *Phys. Rev.* **105**, 767–775 (1957).

[98] Lüscher, M. "Volume Dependence of the Energy Spectrum in Massive Quantum Field Theories. 1. Stable Particle States". *Commun. Math. Phys.* **104**, 177 (1986).

[99] Lüscher, M. "Volume Dependence of the Energy Spectrum in Massive Quantum Field Theories. 2. Scattering States". *Commun. Math. Phys.* **105**, 153–188 (1986).

[100] Lüscher, M. "Two particle states on a torus and their relation to the scattering matrix". *Nucl. Phys. B* **354**, 531–578 (1991).

[101] Briceño, R. A., Dudek, J. J. & Young, R. D. "Scattering processes and resonances from lattice QCD". *Rev. Mod. Phys.* **90**, 025001. arXiv: 1706.06223 [hep-lat] (2018).

[102] Hansen, M. T. & Sharpe, S. R. "Lattice QCD and Three-particle Decays of Resonances". *Ann. Rev. Nucl. Part. Sci.* **69**, 65–107. arXiv: 1901.00483 [hep-lat] (2019).

[103] Hansen, M. T. & Sharpe, S. R. "Perturbative results for two and three particle threshold energies in finite volume". *Phys. Rev. D* **93**, 014506. arXiv: 1509.07929 [hep-lat] (2016).

[104] Kim, C. h., Sachrajda, C. T. & Sharpe, S. R. "Finite-volume effects for two-hadron states in moving frames". *Nucl. Phys. B* **727**, 218–243. arXiv: `hep-lat/0507006` (2005).

[105] Hansen, M. T. & Sharpe, S. R. "Relativistic, model-independent, three-particle quantization condition". *Phys. Rev. D* **90**, 116003. arXiv: 1408.5933 [hep-lat] (2014).

[106] Hansen, M. T. & Sharpe, S. R. "Expressing the three-particle finite-volume spectrum in terms of the three-to-three scattering amplitude". *Phys. Rev. D* **92**, 114509. arXiv: 1504.04248 [hep-lat] (2015).

[107] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three-particle quantization condition in a finite volume: 1. The role of the three-particle force". *JHEP* **09**, 109. arXiv: 1706.07700 [hep-lat] (2017).

[108] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three particle quantization condition in a finite volume: 2. general formalism and the analysis of data". *JHEP* **10**, 115. arXiv: 1707.02176 [hep-lat] (2017).

[109] Mai, M. & Döring, M. "Three-body Unitarity in the Finite Volume". *Eur. Phys. J. A* **53,** 240. arXiv: 1709.08222 [hep-lat] (2017).

[110] Mai, M. & Döring, M. "Finite-Volume Spectrum of $\pi^+\pi^+$ and $\pi^+\pi^+\pi^+$ Systems". *Phys. Rev. Lett.* **122,** 062503. arXiv: 1807.04746 [hep-lat] (2019).

[111] Roper, L. D. "Evidence for a P-11 Pion-Nucleon Resonance at 556 MeV". *Phys. Rev. Lett.* **12,** 340–342 (1964).

[112] Albert, J. & Rastelli, L. "Bootstrapping pions at large N". *JHEP* **08,** 151. arXiv: 2203.11950 [hep-th] (2022).

[113] Chen, H., Fitzpatrick, A. L. & Karateev, D. "Nonperturbative bounds on scattering of massive scalar particles in $d \geq 2$". *JHEP* **12,** 092. arXiv: 2207.12448 [hep-th] (2022).

[114] Fernandez, C., Pomarol, A., Riva, F. & Sciotti, F. "Cornering large-N_c QCD with positivity bounds". *JHEP* **06,** 094. arXiv: 2211.12488 [hep-th] (2023).

Overview of publications

> *Look on my works, ye mighty, and despair!*
>
> — *Percy Bysshe Shelley, from "Ozymandias" (1818)*

In the field of theoretical particle physics, all authors are listed alphabetically, rather than in order of the level of contribution. Below follows a short description of each of the publications included in this thesis and my contribution to each one.

Paper I

Higher-order tree-level amplitudes in the nonlinear sigma model

Johan Bijnens, Karol Kampf, **Mattias Sjö**
JHEP **11** 074 (2019) [Erratum: *JHEP* **03**, 066 (2021)]

In this paper, we derive the method now called DFO (see section 4.4 in the introduction), generalizing earlier leading-order work by Kampf, Novotný and Trnka (ref. [36] in the paper). My innovations included the handling of higher-order vertices, split flavor structures and singlets; the associated graphical notation; the automatic generation of flavor-ordered diagrams; and an early form of closed Mandelstam bases and deorbiting (see section 4.5 of the introduction), which was not fully developed until paper III. We apply it to the nonlinear sigma model, i.e., ChPT in the chiral limit.

The paper is based on my master thesis; in comparison, the paper contains a more mature understanding of the method, a greatly expanded number of amplitudes calculated with it, and a completely reworked method for diagram generation. With the supervision and feedback of Hans and Karol, I developed the methods, wrote the code, performed all analyses, and wrote the entire manuscript except for the introduction, which was mainly written by them.

Paper II

NNLO positivity bounds on chiral perturbation theory for a general number of flavours

Benjamin Alvarez, Johan Bijnens, **Mattias Sjö**
JHEP **03** 159 (2022)

In this paper, we derive positivity bounds on the LECs based on the principles of unitarity, analyticity, crossing symmetry and perturbativity, extending the existing methods and creating framework of *linear constraints* to formalize the management of large sets of bounds.

It is a substantial reworking and expansion of Benjamin's master thesis, which in turn extends the method of Manohar & Mateu (ref. [13] in the paper). I revised the methods to remove further limitations, extended the scope to NNLO, and introduced above-threshold integration, based on similar work by Wang, Zhou and collaborators (refs. [19-20] in the

paper). With Benjamin, I worked on what eventually became appendix B, which in turn led to new insights and better ways to obtain, view and interpret the results. The work went on throughout the brunt of the COVID-19 pandemic, which contributed to its extreme length and level of detail.

The project was suggested by Hans. I wrote all code, performed all analyses and produced all figures, with input and feedback from Hans. The definitions and proofs in section 4 and appendix B were worked out in collaboration with Benjamin. I wrote the entire manuscript, with thorough feedback from Hans and Benjamin. Paper VII provides a summary of the work.

Paper III

Six-meson amplitude in QCD-like theories

Johan Bijnens, Tomáš Husek, **Mattias Sjö**
Phys. Rev. D **106** (2022) 054021

In this paper, we compute the six-meson scattering amplitude at one-loop order in ChPT. We use n_f flavors and a wider set of symmetry breaking patterns beyond that associated with low-energy QCD, following earlier four-meson work by Hans and Jie Lu (ref. [30] in the paper). It is a continuation of Hans and Tomáš' previous work (ref. [19] in the paper), which covers other ChPT variants including that corresponding to two-flavor low-energy QCD.

Hans and Tomáš independently calculated the amplitude. I further developed techniques used in paper I (see sections 4.4 and 4.5 in the introduction), and while their utility for direct calculations was lesser than expected, they proved useful for simplifying and elucidating the results, bringing it to the compact form seen in the paper. I also assisted Tomáš with doing the numerics. The manuscript was mainly written in close collaboration between me and Tomáš, with further contributions and thorough feedback from Hans.

Paper IV

The isospin-3 three-particle K-matrix at NLO in ChPT

Jorge Baeza-Ballesteros, Johan Bijnens, Tomáš Husek, Fernando Romero-López, Stephen Sharpe, **Mattias Sjö**
JHEP **05** 187 (2023)

In this paper, we use the amplitude from paper III to derive the three-particle K-matrix (see section 5.3 in the introduction). We do this for three pions at maximum isospin [compare eq. (4.24) in the introduction], which is the simplest such case, and the only one for which good lattice data was available.

The project was suggested by Hans. Steve, Fernando and Jorge, experts on the K-matrix rather than ChPT, independently begun work on it; we later formed a collaboration.

My main concern was the so-called bull's head subtraction, for which I developed the

methods of section 4.3.4 and appendix E; I was involved to a lesser degree in the complementary methods of sections 4.3.1-2 and 4.3.3, which were mainly Hans' and Tomáš' inventions, respectively. I also performed the cutoff-dependence analysis behind appendix A and produced many of the figures. The manuscript was written in collaboration, with me and Tomáš taking on somewhat editorial roles. Papers VIII and IX contain summaries of the work, including some preliminary parts of paper V.

Paper V

The three-particle K-matrix at NLO in ChPT

Jorge Baeza-Ballesteros, Johan Bijnens, Tomáš Husek, Fernando Romero-López, Stephen Sharpe, **Mattias Sjö**
In preparation for submission to *JHEP*

In this paper, we extend the results of paper IV to all isospin channels, in anticipation of future lattice results. The overall method was derived earlier by Steve, Fernando and their collaborator Maxwell Hansen (ref. [20] in the paper); Steve and I determined the remaining terms in the threshold expansion that had not been derived there. I again mainly focused on the bull's head subtraction and the cutoff-dependence analysis of appendix A.

Like paper IV, the manuscript was written in collaboration. I took on a driving role in structuring it and establishing consistent conventions, and in the presentation of the results. As of the publication of this thesis, the manuscript is still in preparation.

Paper I

I

Higher-order tree-level amplitudes
in the nonlinear sigma model

Johan Bijnens,[1] Karol Kampf[2] and Mattias Sjö[1]

[1]Dept. of Astronomy and Theoretical Physics, Lund University,
 Sölvegatan 14A, Lund, Sweden
[2]Institute of Particle and Nuclear Physics, Charles University,
 V Holešovičkách 2, Prague, Czech Republic

Abstract: We present a generalisation of the flavour-ordering method applied to the chiral nonlinear sigma model with any number of flavours. We use an extended Lagrangian with terms containing any number of derivatives, organised in a power-counting hierarchy. The method allows diagrammatic computations at tree-level with any number of legs at any order in the power-counting. Using an automated implementation of the method, we calculate amplitudes ranging from 12 legs at leading order, $\mathcal{O}(p^2)$, to 6 legs at next-to-next-to-next-to-leading order, $\mathcal{O}(p^8)$. In addition to this, we generalise several properties of amplitudes in the nonlinear sigma model to higher orders. These include the double soft limit and the uniqueness of stripped amplitudes.

Note: Due to a publishing error, the supplementary material referenced in the text is not included with the original publication, but rather with the erratum.

1 Introduction

In 1960, Gell-Mann and Lévy [1] proposed a number of models for mesons and nucleons. Two of these, the linear and nonlinear sigma models, were extended to highly general quantum field theories with many different applications. One of the most important application is interaction of mesons described by the nonlinear sigma model (NLSM) extended by Weinberg [2] and Gasser and Leutwyler [3, 4] into chiral perturbation theory (ChPT). A recent introductory review is [5] and more introductory literature can be found at [6]. This effective field theory (EFT) of low-energy QCD is not only widely used today in many phenomenological applications, but also motivated further theoretical avenues for the beyond-standard-model physics such as technicolour and little Higgs models. Examples of recent work in ChPT is the calculation of meson-meson scattering for a general number of flavours at two loops [7], and masses and decays up to next-to-next-to-leading order [8]. In this paper, we will push the study of this type of models in a different direction.

Even at tree-level, diagrammatic many-particle calculations in EFTs become very complicated due to the rapidly increasing number of terms in the effective Lagrangian, but can be facilitated with tools similar to those used for gluon scattering in perturbative QCD. In recent years, the renewal of interest in the S-matrix program for the gauge theory and gravity has in fact led to progress in both simplification of complicated technical calculations as well as discoveries of new properties [9]. The possibility to apply similar amplitude methods to EFTs started recently and is mainly connected with studies of the NLSM. First, it was demonstrated that it is indeed possible to employ recursive methods in [10], further studied and developed in [11]. The crucial ingredient in developing the recursive formula is the existence of the so-called Adler zero [12], the vanishing of scattering amplitudes for soft momenta of Goldstone bosons (pions for NLSM), as a consequence of a spontaneous symmetry breaking in EFT. The argument can be also inverted and used for classification of the allowed space of EFT theories based on their soft properties. It turned out that the leading order of NLSM is one important representative of exceptional EFTs. The exceptional status of those theories is connected with the fact that all their interaction vertices are uniquely fixed by a single coupling constant, most conveniently the lowest four-point vertex. This can be labelled as a soft-bootstrap program, studied and developed in recent years by several groups [13–15]. It represents a rebirth of similar attempts at the end of the 1960s [16–18].

The exceptional theories have also appeared in completely different context, the so-called CHY scattering equation [19], studied more recently also in [20]. This indeed suggests their uniqueness, and though of completely different nature, it hints to deeper connections with gauge theory and gravity. It is probably one of the main motivation behind the recent increase of activities in studying theoretical properties of NLSM: [21–33]. This effort demonstrates the importance of NLSM; however, these studies mainly concentrated only on the leading, two-derivative ($\mathcal{O}(p^2)$) order. As pointed out in [15], it is important to expand the on-shell soft bootstrap program to higher orders. Our work aims in this direction. An early attempt is [34] and one that appeared during the writing up of this paper is [35].

We will mainly focus on the problem of calculating scattering amplitudes at tree-level with increasing number of legs and orders, with possible flavour splitting, i.e. beyond single-trace amplitudes. Using recursion relations, tree-level amplitudes based on the leading-order term in the Lagrangian have been computed with up to 10 external particles [36]. Using more general recursion relations based on soft limits [11], 6-particle tree-level interactions have been computed using the next-to-leading-order Lagrangian [15]. These methods suffer limitations when higher-order Lagrangian terms are used, and can not handle loops.

In this paper, we generalise an enhanced diagrammatic method called flavour-ordering, which was introduced in [36]. We apply it to a generalised version of the $SU(N)$ or $U(N)$ chiral NLSM, which includes terms with arbitrarily high power-counting order in the effective Lagrangian. This generalisation corresponds to removing all external fields from the general ChPT Lagrangian. The method allows computation of tree-level amplitudes with any number of external particles using Lagrangian terms of any order, and is valid also beyond tree-level. It is significantly more efficient than a brute-force Feynman diagram approach, and the caveats that appear beyond the leading order can be handled with simple rules. Preliminary results can be found in the Lund university master thesis [37].

In section 2, we describe the NLSM and introduce our notation. Our main new results on the method side are described in sections 3 and 4. Section 3 discusses our generalization of flavour-ordering, while section 4 discusses how this can be used to calculate more complex amplitudes as well as the kinematic methods needed. Section 5 discusses the amplitudes we have calculated using our methods; the longer expressions are relegated to appendix E and the supplementary material. Our main conclusions are reviewed in section 6. The Lagrangians are given in appendix A, together with some results regarding renormalisation of the amplitudes. Section B contains the proof of the orthogonality of flavour structures. The double soft limit with multiple traces is derived in appendix C, and appendix D derives the minimal bases of kinematic variables used in the amplitude calculations.

2 The nonlinear sigma model

The nonlinear sigma model describes the Nambu-Goldstone bosons that arise when a global symmetry group G is broken to a subgroup H. Each configuration of the Nambu-Goldstone fields can be uniquely mapped to an element of the coset space G/H, and from each such coset, a representative $\xi(\phi)$ may be chosen to represent the field configuration ϕ.

In the context of low-energy QCD, the group G is the chiral group $SU(N_f)_L \times SU(N_f)_R$, which is a global symmetry of the massless QCD Lagrangian with N_f quark flavours. It is broken to the diagonal subgroup $H = SU(N_f)_V$, so the coset space G/H is isomorphic to $SU(N_f)$. With a chiral decomposition of the coset representatives, $\xi = (\xi_L, \xi_R)$, we may represent the Nambu-Goldstone fields with the unitary matrix $u(\phi) = \xi_R(\phi) = \xi_L^\dagger(\phi)$ parametrised as

$$u(\phi) = \exp\left(\frac{i\Phi(\phi)}{F\sqrt{2}}\right), \qquad \Phi(\phi) = t^a \phi^a \tag{1.2.1}$$

with the flavour index with a running from 1 to $N_f^2 - 1$. Here, t^a are the generators of $SU(N_f)$, and F is a constant.[1] We use Einstein's summation notation without distinction between upper and lower flavour indices, and use the following normalisation for the generators:

$$\langle t^a t^b \rangle = \delta^{ab}, \qquad [t^a, t^b] = i f^{abc} t^c, \tag{1.2.2}$$

where $\langle \cdots \rangle$ denotes a trace over internal indices. Here, f^{abc} are the totally antisymmetric structure constants of $SU(N_f)$. With this convention, the $SU(2)$ generators can be chosen such that they relate to the Pauli matrices as $t^a = \sigma^a/\sqrt{2}$, $f^{abc} = \epsilon^{abc}\sqrt{2}$. Likewise, the $SU(3)$ generators can be chosen in terms of the Gell-Mann matrices like $t^a = \lambda^a/\sqrt{2}$.

Under a chiral transformation $g = (g_L, g_R)$, $u(\phi)$ transforms as

$$u \xrightarrow{g} g_R\, u\, h^\dagger(u, g) = h(u, g)\, u\, g_L, \tag{1.2.3}$$

where the compensating transformation $h(u, g) \in H$ is defined by the above relation.

When constructing the most general symmetry-consistent Lagrangian, it is more convenient to replace u by

$$u_\mu = i\left(u^\dagger \partial_\mu u - u \partial_\mu u^\dagger\right), \qquad u_\mu \xrightarrow{g} h(u, g) u_\mu h(u, g)^\dagger, \tag{1.2.4}$$

which was introduced in this context by [38]. Higher derivatives are applied through the covariant derivative

$$\nabla_\mu X = \partial_\mu X + [\Gamma_\mu, X], \qquad \Gamma_\mu = \frac{1}{2}\left(u^\dagger \partial_\mu u + u \partial_\mu u^\dagger\right), \tag{1.2.5}$$

which has the convenient properties

$$X \xrightarrow{g} h X h^\dagger \quad \Rightarrow \quad \nabla_\mu X \xrightarrow{g} h \nabla_\mu X h^\dagger, \qquad \nabla_\mu u_\nu - \nabla_\nu u_\mu = 0. \tag{1.2.6}$$

Note that we do not include the external fields that are used in chiral perturbation theory.

The Lagrangian is often written using derivatives of $U(\phi) = u(\phi)^2$ and its conjugate. It is possible to convert directly between $\partial_\mu U^{(\dagger)}$ and u_μ by using unitarity:

$$\partial_\mu U^\dagger \partial_\nu U = -(U^\dagger \partial_\mu U)(U^\dagger \partial_\nu U), \qquad U^\dagger \partial_\mu U = u^\dagger u^\dagger \partial_\mu u u - \partial_\mu u^\dagger u = -i u^\dagger u_\mu u. \tag{1.2.7}$$

This makes $\partial_\mu U^\dagger \partial_\nu U$ wholly interchangeable with $u_\mu u_\nu$ inside a trace.

With the above definitions, the simplest valid term in the NLSM Lagrangian is

$$\mathcal{L}_2 = \frac{F^2}{4}\langle u_\mu u^\mu \rangle, \tag{1.2.8}$$

where the constant in front is fixed by the canonical normalisation of the kinetic term.

---

[1] The above expression for $u(\phi)$ is only one of many possible parametrisations, but is the most common.

Beyond this, there is an infinite sequence of increasingly complex terms permitted by the chiral symmetries[2]. We also impose parity (P), charge-conjugation (C) and Lorentz invariance. We restrict to the sector that involves an even number of Levi-Civita tensors ($\epsilon_{\mu\nu\alpha\beta}$), which can always be rewritten in terms of the Minkowski metric only. The terms can be organised into a hierarchy based on power counting in the momentum scale p. Since each derivative in the Lagrangian brings down one factor of p into an amplitude, both u_μ and ∇_μ are $\mathcal{O}(p)$ and the power-counting at the Lagrangian level is simply counting derivatives. Thus, we may split the Lagrangian as

$$\mathcal{L} = \sum_{n=1}^{\infty} \mathcal{L}_{2n}, \tag{1.2.9}$$

where $\mathcal{L}_{2n}$ is $\mathcal{O}(p^{2n})$ and contains $2n$ derivatives carrying n pairs of Lorentz indices. Assuming a low momentum scale, we may then ignore all terms above a certain n.

The four-derivative $\mathcal{O}(p^4)$ Lagrangian is, for general N_f [3, 4, 39],

$$\mathcal{L}_4 = L_0 \langle u_\mu u_\nu u^\mu u^\nu \rangle + L_1 \langle u_\mu u^\mu \rangle \langle u_\nu u^\nu \rangle + L_2 \langle u_\mu u_\nu \rangle \langle u^\mu u^\nu \rangle + L_3 \langle u_\mu u^\mu u_\nu u^\nu \rangle. \tag{1.2.10}$$

The L_i are independent coupling constants, so-called low-energy constants (LECs). It is in principle possible to derive the LECs from any underlying theory (e.g. QCD), but in practice, they are unknown parameters that must be measured by experiments or lattice simulations.

The Lagrangian is known also at $\mathcal{O}(p^6)$ and $\mathcal{O}(p^8)$. The latter is the first 135 terms in the ChPT Lagrangian of [40]; the former has only been published with different notation and formulated in a way that gives redundant terms when naïvely reduced to the NLSM [39]. A more compatible version, constructed in conjunction with [40], is given in appendix A. The Lagrangian at $\mathcal{O}(p^{10})$ and above has not been studied.

2.1 Restrictions due to fixed N_f and dimensionality

The Lagrangians discussed above are the most general ones. They are valid in any dimension and for a generic number of flavours.

When N_f is small, the Cayley-Hamilton theorem gives additional linear relations that reduce the number of independent terms. The theorem states that for any $N_f \times N_f$ matrix M, the characteristic polynomial

$$p(\lambda) = \det(\lambda\mathbb{1} - M), \tag{1.2.11}$$

which is zero whenever λ is an eigenvalue of M, is also satisfied by M, i.e. $p(M) = 0$ when viewed as a matrix polynomial. When $N_f = 2$, this implies the identity

$$\{A, B\} = \langle AB \rangle, \tag{1.2.12}$$

---

[2]Many authors refer to $\mathcal{L}_2$ as the full Lagrangian of the NLSM. We instead use "the NLSM" to refer to the more general version, which includes also terms with more derivatives.

for traceless 2×2 matrices A, B. When $N_f = 3$, the identity is

$$\sum_{\substack{\text{permutations of} \\ \{ABC\}}} \langle ABCD \rangle \quad = \quad \sum_{\substack{\text{cyclic permutations} \\ \text{of } \{ABC\}}} \langle AB \rangle \langle CD \rangle, \qquad (1.2.13)$$

for traceless matrices. The relations when $\langle A \rangle \neq 0$ used in [39, 40] contain many more terms.

As an example, we may choose $A = C = u_\mu$ and $B = D = u_\nu$; these are traceless as a consequence of the identity $\partial_\mu \det(A) = \det(A) \langle A^{-1} \partial_\mu A \rangle$, which holds for any invertible A, and which reduces to $\langle A^\dagger \partial_\mu A \rangle = 0$ when $A \in SU(N_f)$.

The $N_f = 2$ identity allows for the elimination of all Lagrangian terms containing a product of two or more traces from any $\mathcal{L}_{2n}$; for instance, L_1 and L_2 may be eliminated from $\mathcal{L}_4$. The $N_f = 3$ identity allows for the removal of a single term from $\mathcal{L}_4$, 7 terms from $\mathcal{L}_6$, and so on. The standard choice is to remove the L_0-term of eq. (1.2.10) for $N_f = 3$ [4], and the L_0- and L_3-terms for $N_f = 2$ [3].

When the spacetime dimension D is finite, the Schouten identity implies

$$(f^{\mu_1 \mu_2 \cdots \mu_k} u_{\mu_1} u_{\mu_2} \cdots u_{\mu_k})^2 = 0 \quad \text{if } k > D, \qquad (1.2.14)$$

where $f^{\mu_1 \mu_2 \cdots \mu_k}$ is antisymmetric in all its indices. This results in additional linear relations among the terms in $\mathcal{L}_{2k}$ for $k > D$. For $D = 4$, this does not affect any of the currently known orders. In the sector involving a single Levi-Civita tensor it already removes a large number of terms at $\mathcal{O}(p^6)$.

3 Flavour-ordering

With the structure of the NLSM established, we are ready to use it for perturbative calculations of scattering amplitudes. However, the infinite number of interaction terms requires the use of some scheme for restricting it to a manageable subset. Even then, the resulting vertex factors are very intricate, both in their dependence on the particle momenta, and in their group-algebraic structure. This leaves only the simplest Feynman diagrams tractable by hand, and even computer algebra becomes highly time-consuming when tackling more complicated cases directly.

In this section, we will direct much effort towards the development of simpler ways to perform these calculations. As we will see, the group-algebraic structure of the flavour indices carried by the particles can be used to condense an amplitude into a much more easily manageable expression, for which simpler calculation rules exist. We will mostly follow the derivation of $\mathcal{O}(p^2)$ flavour-ordering as presented in [36], but insert the notation to support our own generalisations to higher-order vertices.

3.1 Some notation

In this section, we will need a compact notation for writing the flavour structures of scattering amplitudes. A flavour structure is a product of one or more traces containing group generators carrying the flavour indices of the external particles in some order. We will represent this as

$$\mathcal{F}_\sigma(r_1, r_2, \ldots) = \langle t^{a_{\sigma(1)}} \cdots t^{a_{\sigma(r_1)}} \rangle \langle t^{a_{\sigma(r_1+1)}} \cdots t^{a_{\sigma(r_1+r_2)}} \rangle \cdots . \tag{1.3.1}$$

The ith trace contains r_i generators, ordered by a permutation $\sigma \in \mathcal{S}_n$. For example, $\langle a_1 a_3 \rangle \langle a_2 a_4 \rangle = \mathcal{F}_{1324}(2, 2)$, and $\langle a_1 a_2 \cdots a_n \rangle = \mathcal{F}_{\mathrm{id}}(n)$, where $\mathrm{id}(i) = i$ is the identity permutation.

We encapsulate the r_i in $R = \{r_1, r_2, \ldots, r_{|R|}\}$. We call R a *flavour splitting*. $|R|$ is the number of traces in the flavour structure, and we write $\mathcal{F}_\sigma(R)$ rather than $\mathcal{F}_\sigma(r_1, \ldots)$. For a structure with n indices, we impose the restrictions

$$\sum_{i=1}^{|R|} r_i = n, \qquad r_1 \leq r_2 \leq \ldots \leq r_{|R|}. \tag{1.3.2}$$

The latter limits the number of equivalent ways to write a flavour structure.

Since traces are cyclic, $\mathcal{F}_\sigma(R)$ will be invariant under cyclic permutations of the indices inside each trace. If $r_i = r_j$, it will also be invariant under swapping the contents of the ith and jth trace. As a generalisation of the cyclic group $\mathbb{Z}_n$, we define $\mathbb{Z}_R$ to be the group of all permutations under which $\mathcal{F}_\sigma(R)$ is invariant. For instance,

$$\begin{aligned} \mathbb{Z}_{\{2,2\}} &= \{12\,34, 21\,34, 12\,43, 21\,43, 34\,12, 43\,12, 34\,21, 43\,21\}, \\ \mathbb{Z}_{\{2,4\}} &= \{12\,3456, 21\,3456, 12\,4563, 21\,4563, 12\,5634, 21\,5634, 12\,6345, 21\,6345\}, \end{aligned} \tag{1.3.3}$$

where we label a permutation by how $12 \ldots n$ ends up. We have inserted spaces between blocks of indices corresponding to different traces to make it more legible.

In this notation, we generalise the notion of two permutations being equivalent modulo a cyclic permutation: we write $\sigma \equiv \rho \pmod{\mathbb{Z}_R}$ if $\mathcal{F}_\sigma(R) = \mathcal{F}_\rho(R)$. For instance, $1234 \equiv 2341 \pmod{\mathbb{Z}_{\{4\}}}$ and $1234 \equiv 2134 \pmod{\mathbb{Z}_{\{2,2\}}}$.

$\mathbb{Z}_{\{2,2\}}$ is isomorphic to the dihedral group D_4. Other $\mathbb{Z}_R$ are not isomorphic to such well-known groups, but $\mathbb{Z}_{\{2,4\}} \simeq \mathbb{Z}_2 \times \mathbb{Z}_4$, and in general, $\mathbb{Z}_R \simeq \mathbb{Z}_{r_1} \times \mathbb{Z}_{r_2} \times \cdots$ whenever all r_i are different. When some r_i are equal (say, m in a row), the group will be non-abelian and isomorphic to a semidirect product, e.g. $\mathbb{Z}_{\{2,2,2\}} \simeq (\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes \mathcal{S}_3$. In general, $\mathbb{Z}_R \simeq (\mathbb{Z}_{r_1} \times \mathbb{Z}_{r_2} \times \cdots) \rtimes (\mathcal{S}_{m_1} \times \mathcal{S}_{m_2} \times \cdots)$, where each m_j is the length of a stretch of equal r_i.[3]

---

[3]The proof follows from the following definition of the semidirect product: if a group G has a subgroup H and a normal subgroup N, then $G = N \rtimes H$ if $G = \{nh \,|\, n \in N, h \in H\}$ and $N \cap H = e$, the identity element. The groups $N \simeq (\mathbb{Z}_{r_1} \times \cdots)$ of cyclings within traces and $H \simeq (\mathcal{S}_{m_1} \times \cdots)$ of swaps of equal-size traces are clearly subgroups of $G = \mathbb{Z}_R$, and N is normal since $gng^{-1} \in N$ for any $n \in N, g \in G$ — any trace swaps in g are cancelled by g^{-1}, leaving only cyclings. Any permutation in $\mathbb{Z}_R$ is the composition of a cycling and a trace swap, and the only element shared by N and H is id, which completes the proof.

3.2 Stripped vertex factors

Each term in the Lagrangian will produce an infinite tower of interaction vertices with increasingly many legs. Due to parity and the absence of Levi-Civita tensors, only terms with an even number of legs are produced. If the Lagrangian term contains a product of several traces, the flavour indices of the corresponding vertices will be distributed between the same number of traces in multiple ways. If a trace contains an even number of u_μ's in the Lagrangian, the corresponding trace in the vertices will only contain an even number of indices, again from parity.

We will organise the vertices by their power-counting order and flavour splitting. For instance, in the expansion of $\mathcal{L}_2$ eq. (1.2.8),

$$\mathcal{L}_2 = \frac{1}{2}\langle t^a t^b \rangle \partial_\mu \phi^a \partial^\mu \phi^b + \frac{1}{F^2}\langle t^a t^b t^c t^d \rangle \left(\frac{1}{6}\phi^a \partial_\mu \phi^b \phi^c \partial^\mu \phi^d - \frac{1}{12}\phi^a \phi^b \partial_\mu \phi^c \partial^\mu \phi^d \right) + \dots,$$
(1.3.4)

both terms attached to the 4-index trace will be part of the $\mathcal{O}(p^2)$ vertex with splitting $R = \{4\}$, which we label $\mathcal{V}_{2,\{4\}}^{abcd}$, a vertex (factor). At order p^m, a specific flavour splitting R for a vertex with n legs, and thus n flavour indices a_i, will have a vertex factor $\mathcal{V}_{m,R}^{a_1 \cdots a_n}$. It will in general contain contributions from many different Lagrangian terms, but we treat it as a single factor for the purposes of Feynman diagrams.

We can further organise the contents of an n-point $\mathcal{O}(p^m)$ vertex by flavour structure, i.e. all possible distributions of the n-flavour indices over the flavour splitting R:

$$\mathcal{V}_{m,R}^{a_1 a_2 \cdots a_n}(p_1, p_2 \dots, p_n) = \sum_{\sigma \in \mathcal{S}_n / \mathbb{Z}_R} \mathcal{F}_\sigma(R)\, \mathcal{V}_{m,R,\sigma}(p_1, p_2, \dots, p_n),$$
(1.3.5)

where $\mathcal{V}_{m,R,\sigma}$ contains whatever kinematic factors come attached to $\mathcal{F}_\sigma(R)$. Due to the derivatives, the kinematic factors $\mathcal{V}_{m,R,\sigma}$ are functions of the momenta p_i of the interacting particles. Here and in all other places, we treat all momenta as ingoing. Since $\mathcal{F}_\sigma(R)$ is invariant under $\mathbb{Z}_R$, the kinematic factors must also have this symmetry, i.e.

$$\mathcal{V}_{m,R,\sigma}(p_1, p_2, \dots, p_n) = \mathcal{V}_{m,R,\sigma}(p_{\rho(1)}, p_{\rho(2)}, \dots, p_{\rho(n)})$$
(1.3.6)

for any $\rho \in \mathbb{Z}_R$. Also, Bose symmetry implies that the act of rearranging the legs of the vertex by any permutation $\rho \in \mathcal{S}_n$ must have the effect

$$\mathcal{V}_{m,R,\sigma \circ \rho}(p_1, p_2, \dots, p_n) = \mathcal{V}_{m,R,\sigma}(p_{\rho(1)}, p_{\rho(2)}, \dots, p_{\rho(n)}),$$
(1.3.7)

where $\circ$ denotes composition of permutations. Specifically,

$$\mathcal{V}_{m,R,\sigma}(p_1, p_2, \dots, p_n) = \mathcal{V}_{m,R}(p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(n)}),$$
(1.3.8)

where $\mathcal{V}_{m,R} = \mathcal{V}_{m,R,\mathrm{id}}$ is called a *stripped* vertex factor.[4] It contains all the necessary information of the full vertex factor, but is only a kinematic factor with no flavour structure.

---

[4]The word "stripped" is typical in the context of EFTs. For the analogous concept in perturbative QCD (where "flavour" is replaced by "colour"), the word "primitive" is used instead; see e.g. [41, 42]. In older literature, the word "dual" is common.

It can be "dressed" into a full vertex factor by the simple act of multiplying by $\mathcal{F}_{\mathrm{id}}(R)$ and then summing over all $\sigma \in \mathcal{S}_n/\mathbb{Z}_R$.

A stripped vertex factor has the property of being *flavour-ordered*, since it is the kinematic factor attached to $\mathcal{F}_{\mathrm{id}}(R)$, where all flavour indices are sorted in ascending order. Thanks to this, its explicit form can be derived by expanding the relevant Lagrangian terms and discarding all terms where any flavour index appears out of order. This saves a significant amount of work for the more complicated vertices.

Stripped vertices serve as the first ingredient in our method. In the following sections, we treat diagrams and amplitudes along the same lines.

3.3 Stripped amplitudes

Like the vertices, we may organise the diagrams by their power-counting order and flavour structure. The order can be determined by using Weinberg's power-counting formula,

$$D = 2 + 2L + \sum_d (d-2)N_d, \tag{1.3.9}$$

which states that a diagram containing L loops and N_d $\mathcal{O}(p^d)$ vertices is $\mathcal{O}(p^D)$ overall. Due to the form $(d-2)$, a diagram may contain any number of $\mathcal{O}(p^2)$ vertices without changing its order. Each loop

As for the vertex factors, we may decompose the $\mathcal{O}(p^m)$ n-point amplitude as

$$\mathcal{M}_{m,n}^{a_1 a_2 \cdots a_n}(p_1, p_2, \ldots, p_n) = \sum_{R \in \mathcal{R}(m,n)} \sum_{\sigma \in \mathcal{S}_n/\mathbb{Z}_R} \mathcal{F}_\sigma(R) \mathcal{M}_{m,R,\sigma}(p_1, p_2, \ldots, p_n), \tag{1.3.10}$$

where $\mathcal{M}_{m,R,\sigma}$ carries all kinematic factors, and $\mathcal{R}(m,n)$ contains all flavour splittings that contribute to the amplitude. Its contents will become apparent when drawing diagrams.

The direct analogues of eqs. (1.3.6) to (1.3.8) hold also for $\mathcal{M}_{m,R,\sigma}$, and we may similarly define the *stripped amplitude* $\mathcal{M}_{m,R}$ with the property

$$\mathcal{M}_{N,R,\sigma}(p_1, p_2, \ldots, p_n) = \mathcal{M}_{m,R}(p_{\sigma(1)}, p_{\sigma(2)}, \ldots, p_{\sigma(n)}). \tag{1.3.11}$$

It is sufficient to compute the stripped amplitude, since summing over flavour splittings and permutations,

$$\mathcal{M}_{m,n}^{a_1 a_2 \cdots a_n}(p_1, p_2, \ldots, p_n) = \sum_{R \in \mathcal{R}(N,n)} \sum_{\sigma \in \mathcal{S}_n/\mathbb{Z}_R} \mathcal{F}_\sigma(R) \mathcal{M}_{m,R}(p_{\sigma(1)}, p_{\sigma(2)}, \ldots, p_{\sigma(n)}), \tag{1.3.12}$$

gives the full amplitude.

3.4 Flavour-ordered diagrams

Due to its relative simplicity, the stripped amplitude serves as the target of our methods. Like the stripped vertex factors, it is flavour-ordered, so when calculating it, we may discard all terms where two flavour indices appear out of order. We can derive further simplifications by studying how the flavour structures behave when two sub-diagrams are joined by propagators. The NLSM Feynman rule for a propagator with momentum q is

$$\underset{a}{\bullet} \xrightarrow{\quad q \quad} \underset{b}{\bullet} = \frac{i\delta^{ab}}{q^2 + i\epsilon}, \tag{1.3.13}$$

so the flavour structures are simply contracted by the delta. For $SU(N_f)$, the contraction can be performed through the Fierz identity,

$$t_{ij}^a t_{k\ell}^a = \delta_{i\ell}\delta_{jk} - \frac{1}{N_f}\delta_{ij}\delta_{k\ell}, \tag{1.3.14}$$

where $ijk\ell$ are the internal indices of the generators. Inside traces, the identity implies

$$\langle Xt^a\rangle\langle t^aY\rangle = \langle XY\rangle - \frac{1}{N_f}\langle X\rangle\langle Y\rangle, \tag{1.3.15}$$

$$\langle Xt^aYt^a\rangle = \langle X\rangle\langle Y\rangle - \frac{1}{N_f}\langle XY\rangle \tag{1.3.16}$$

for arbitrary X and Y. For future reference, we will name the first term on the right-hand side the *multiplet term* and the second term (containing N_f^{-1}) the *singlet term*. In $U(N_f)$, the corresponding identities contain only the multiplet term.

For tree-level diagrams, eq. (1.3.15) is the relevant identity. Its multiplet term preserves the ordering of X and Y; the singlet term does not, but we will ignore it for now and deal with it in section 3.5. We then see that the stripped amplitude only gets contributions from stripped vertex factors (if X or Y is not flavour-ordered, neither is XY) that are combined in ways that maintain their flavour-ordering. In a diagrammatic view, this is rather intuitive to achieve; for instance, the following constitutes all the distinct ways to assemble two 4-point vertices into an $\mathcal{O}(p^2)$ 6-point diagram:

$$\tag{1.3.17}$$

The labels on the legs refer to external momenta and flavour indices. Flavour-ordering corresponds to having all indices in cyclic order around the diagram labelled counterclockwise; we will keep this convention in the remainder. These three labellings give distinct kinematic factors, e.g. they have distinct propagator momenta $(p_1 + p_2 + p_3)^2$, $(p_2 + p_3 + p_4)^2$, and $(p_3 + p_4 + p_5)^2$, respectively. Due to the symmetry of the diagram, the remaining three cyclic permutations of the labels are not distinct from these three. All other labellings are not flavour-ordered, and can be ignored.

For compactness, we will draw flavour-ordered diagrams with unlabelled legs. These are defined as the sum over all distinct flavour-ordered ways to label them. Equivalently, they can be defined as any flavour-ordered labelling, summed over $\mathbb{Z}_R$, and divided by the factor needed to account for symmetry. For 4, 6 and 8 particles at $\mathcal{O}(p^2)$, the flavour-ordered diagrams are

$$\text{(1.3.18)}$$

respectively. The second 6-point diagram is the sum of the three in eq. (1.3.17). Stripped vertex factors are completely symmetric under their respective $\mathbb{Z}_R$ by virtue of eq. (1.3.6), so single-vertex diagrams always have only one distinct labelling. Therefore, the 4-point diagram and the first 6-point diagram in eq. (1.3.18) should not be summed over other labellings. The 8-point diagrams have 1, 8, 4, and 8 distinct labellings, respectively, as can be seen from their symmetry. Note that since the order of the legs of a stripped vertex factor matters, the last two diagrams are distinct.

Above $\mathcal{O}(p^2)$, we begin to encounter flavour-split vertices, but they can be integrated into the flavour-ordering routine. We still label the legs according to the identity permutation, but instead of summing over cyclic permutations, we sum over $\mathbb{Z}_R$, and once again only consider distinct labellings.

At higher orders, we also need to distinguish vertices of different order, which is done by attaching a number to all vertices above $\mathcal{O}(p^2)$. In order to distinguish vertices with split flavour structures, we leave a gap in the vertex, so that each contiguous piece of a diagram resides in a single trace. For instance, the 4-point $\mathcal{O}(p^4)$ diagrams are

$$\text{(1.3.19)}$$

for $R = \{4\}$ and $R = \{2, 2\}$, respectively. Neither diagram has more than one distinct labelling, since they contain only a single vertex each. The four lines in the right diagram are still kinematically connected, but are separated flavour-wise. Since there is a direct correspondence between traces in a flavour structure and contiguous pieces of a diagram, we will simply refer to the pieces as *traces*.

Some adjustment is needed when handling split diagrams. Since $\langle X \rangle \langle Y \rangle = \langle Y \rangle \langle X \rangle$, the traces may "float" to different positions around the same vertex. For instance,

$$\text{(1.3.20)}$$

are the same. By our conventions, the distinct labellings of this diagram are

$$(1.3.21)$$

Labels 1 and 2 are applied to the smaller trace (as per eq. (1.3.2)), and no cycling is needed due to the symmetry of the vertex. Labels 3456 must be summed over all four cyclings, since each cycling gives a different propagator. No other labelling is flavour-ordered; in particular,

$$(1.3.22)$$

which would be valid on a single-trace diagram, should not be counted, since it has flavour structure $\mathcal{F}_{\mathrm{id}}(4, 2)$ in disagreement with eq. (1.3.2). Including it would be double-counting when summing over all σ in eq. (1.3.12), since it is obtained from eq. (1.3.21) via a permutation in $\mathcal{S}_6/\mathbb{Z}_{\{2,4\}}$.

Extra caveats sometimes show up. For instance, the two $\mathcal{O}(p^4)$ diagrams

$$(1.3.23)$$

emerge from different orientations of the same three vertices, but have completely different flavour structure and properties. In the first diagram, the smaller trace should not be cycled at all, and the larger trace only halfway, since it is symmetric (compare to the $\mathcal{O}(p^2)$ 6-point diagram). In the second diagram, all $4 \cdot 4$ combined cyclings of the two traces are distict, but due to the symmetry of the diagram, swapping them, e.g.

$$(1.3.24)$$

does not produce a distinct kinematic structure and should not be counted.

In the $\mathcal{O}(p^6)$ diagrams

$$(1.3.25)$$

the component of $\mathbb{Z}_R$ that swaps equal-size traces does play a role. In the first diagram, we may place either 12 or 34 in the trace straddling the propagator, and we must sum over both placements. In addition to that, we must sum over cyclings of the trace that straddles the propagator. In the second diagram, the two smaller traces are equivalent under the $\mathbb{Z}_{\{2,2,2\}}$ symmetry of the vertex, and we should not sum over both ways of placing the labels 12 and 34.

3.5 The singlet problem and its solution

The construction of flavour-ordered diagrams hinges heavily on the use of eq. (1.3.15), or specifically the muliplet term, $\langle XY \rangle$. The singlet term, $\langle X \rangle \langle Y \rangle / N_f$, threatens the notion that the stripped amplitude is given by exactly the flavour-ordered diagrams. Consider the diagrams

$$. \qquad (1.3.26)$$

The first diagram is flavour-ordered according to both the multiplet and singlet terms. The second diagram is also flavour-ordered according to our definitions, but gives the non-flavour-ordered structure $\langle 234 \rangle \langle 561 \rangle$ under the singlet. The third diagram is not flavour-ordered, but the singlet gives the flavour-ordered structure $\langle 123 \rangle \langle 456 \rangle$. Since the only permutation contained in both $\mathbb{Z}_{\{6\}}$ and $\mathbb{Z}_{\{3,3\}}$ is id, the behaviour of the singlet and multiplet terms is clearly very different and must be treated carefully.

There is, however, an elegant solution. As stated previously, the singlet term in eq. (1.3.15) is not present in $U(N_f)$. Therefore, in the $U(N_f)$ NLSM we may always do flavour-ordering without singlet issues. We can extend this to $SU(N_f)$ by using its similarity to $U(N_f)$.

The $U(N_f)$ algebra differs from the $SU(N_f)$ algebra by a non-traceless generator t^0 that commutes with all other generators. Due to the latter property, its associated field ϕ^0 forms a $U(1)$ singlet separate from the $SU(N_f)$ mutliplet ϕ^a. With this in mind, a more elucidating form of eq. (1.3.15) is

$$\sum_{a=1}^{N_f^2-1} \langle Xt^a \rangle \langle t^a Y \rangle = \sum_{a=0}^{N_f^2-1} \langle Xt^a \rangle \langle t^a Y \rangle - \langle Xt^0 \rangle \langle t^0 Y \rangle, \qquad (1.3.27)$$

where we temporarily suppress Einstein summation. This expression suggests that a $SU(N_f)$ propagator (left) represents a $U(N_f)$ propagator (right) minus the singlet propagator, and explains our naming of the terms in eq. (1.3.15). The N_f^{-1} is absorbed into t^0 since $t^0 = \mathbb{1}/\sqrt{N_f}$.

Now, if we extend our Lagrangian-building field like

$$\hat{\Phi}(\phi^0, \phi) = t^0 \phi^0 + \Phi(\phi), \qquad \hat{u}(\phi^0, \phi) = \exp\left(\frac{i\hat{\Phi}}{F\sqrt{2}}\right) = \exp\left(\frac{i\phi^0 t^0}{F\sqrt{2}}\right) u(\phi), \quad (1.3.28)$$

where $u(\phi) \in SU(N_f)$ and $\hat{u}(\phi^0, \phi) \in U(N_f)$, we see that

$$\hat{U} = \hat{u}^2 \quad \Rightarrow \quad \hat{U}^\dagger \partial_\mu \hat{U} = \left(\frac{i\sqrt{2}}{F\sqrt{N_f}}\right) \partial_\mu \phi^0 + U^\dagger \partial_\mu U \qquad (1.3.29)$$

(remembering that $U^\dagger \partial_\mu U$ is equivalent to u_μ), and therefore

$$\hat{\mathcal{L}}_2 = \frac{1}{2} \partial_\mu \phi^0 \partial^\mu \phi^0 + \frac{F^2}{4} \langle u_\mu u^\mu \rangle. \qquad (1.3.30)$$

At this order, the singlet decouples from the other fields and forms a free theory. Therefore, no $\mathcal{O}(p^2)$ vertex involves the singlet, so there is no distinction between $U(N_f)$ and $SU(N_f)$ amplitudes at this order, and we may ignore the singlet term in eq. (1.3.15).

This observation was sufficient in [36], but we must handle the singlet problem beyond $\mathcal{O}(p^2)$. $\mathcal{L}_4$ and all higher-order Lagrangians introduce vertices that couple the singlet to the other particles. However, a singlet propagator can only exist if both vertices at its ends couple to it. Since this requires at least two vertices of at least $\mathcal{O}(p^4)$, the diagram as a whole must be at least $\mathcal{O}(p^6)$ to include such complications.[5,6] Therefore, flavour-ordering at $\mathcal{O}(p^4)$ works with no other complications than the introduction of split vertices.

At $\mathcal{O}(p^6)$ and above, the singlet term in eq. (1.3.15) can not be avoided in $SU(N_f)$, but the interpretation of eq. (1.3.27) still holds. In order to build a $SU(N_f)$ amplitude, we first work in $U(N_f)$ to build flavour-ordered diagrams using only the multiplet term. Then, we construct all diagrams with singlet propagators in a similar fashion, maintaining flavour-ordering independently. For instance, the full suite of $\mathcal{O}(p^6)$ 6-point diagrams is

$$\tag{1.3.31}$$

including one singlet propagator, indicated by a dashed line. It implicitly includes a factor of $-N_f^{-1}$, and its flavour structure is split $\{3,3\}$ over the propagator. All cyclings of the two traces should be counted as distinct, since the vertices are invariant under $\mathbb{Z}_4$, not $\mathbb{Z}_3$. By adding the singlet diagrams to the others, we get the stripped $SU(N_f)$ amplitude.

The singlet diagram contains all contractions that are flavour-ordered under the singlet term, like the first and last diagram in eq. (1.3.26). The decoupling of the singlet at $\mathcal{O}(p^2)$ means that these contributions must cancel in the amplitude at this order, which is not at all obvious from the individual diagrams. Still, recasting the singlet terms as flavour-ordered singlet diagrams is valid, as follows from the uniqueness of the stripped amplitude.

---

[5] If the singlet forms a loop, only one $\mathcal{O}(p^4)$ vertex is necessary, but the loop itself increases the power counting, so $\mathcal{O}(p^6)$ is needed in this case as well.

[6] An interesting parallel can be seen in [41], where $U(1)$ gluons similar to our singlets must be introduced in perturbative QCD. While our singlets only emerge with at least two higher-order vertices, their $U(1)$ gluons cancel unless the diagram contains at least two quark lines. In general, there are several intriguing analogies between the inclusion of quark lines in gluon scattering (where there are no higher-order vertices) and the inclusion of higher-order vertices in the NLSM (where there are no quark lines).

3.6 Uniqueness of stripped amplitudes

Above, we have blindly trusted the definition of the stripped amplitude as everything that comes attached to the flavour-ordered structure $\mathcal{F}_{\mathrm{id}}(R)$. If this definition is not unique, flavour-ordering would not necessarily be valid, and we could not rely on our use of singlet diagrams. However, we can show that the stripped amplitude is indeed unique, using a generalisation of a method presented by [43] and adapted to flavour-ordering by [36].

The uniqueness hinges on the orthogonality relation

$$\mathcal{F}_\sigma(Q) \cdot \left[\mathcal{F}_\rho(R)\right]^* = N_f^n \begin{cases} 1 + \mathcal{O}\left(N_f^{-2}\right) & \text{if } Q = R \text{ and } \sigma \equiv \rho \pmod{\mathbb{Z}_R}, \\ \mathcal{O}\left(N_f^{-\gamma}\right) & \text{otherwise } (\gamma \geq 1; \text{ see below}) \end{cases} \tag{1.3.32}$$

using the notation defined in section 3.1. The dot in the left-hand side indicates contraction over all flavour indices. If $Q \neq R$, $\gamma \geq 1$, and if $\sigma \not\equiv \rho \pmod{\mathbb{Z}_R}$, $\gamma \geq 2$; therefore, the single-trace version (i.e. that given in [36]) has $\mathcal{O}(N_f^{-2})$ as its second case. The more different the flavour structures are, the larger γ is. The relation eq. (1.3.32) is proven in appendix B and states that any given flavour structure $\mathcal{F}_\sigma(Q)$ is orthogonal at leading order in N_f to all other flavour structures whose permutations are not equivalent to σ, or whose flavour splittings are not equal to Q.

In the context of stripped amplitude uniqueness, eq. (1.3.32) can be applied as follows. In analogy with eqs. (1.3.5) and (1.3.12), we write some arbitrary quantity $\mathcal{X}$ in the form

$$\mathcal{X}^{a_1 \cdots a_n} = \sum_{R \in \mathcal{R}} \sum_{\sigma \in \mathcal{S}_n / \mathbb{Z}_R} \mathcal{F}_\sigma(R) \mathcal{X}_{\sigma, R}, \tag{1.3.33}$$

where $\mathcal{R}$ is some appropriate selection of flavour splittings. Then, we use the orthogonality relation eq. (1.3.32) to perform the projection

$$\mathcal{X}^{a_1 \cdots a_n} \left[\mathcal{F}_{\mathrm{id}}(R)\right]^* = N_f^n \left[\mathcal{X}_{\mathrm{id}, R} + \mathcal{O}\left(\frac{1}{N_f}\right)\right]. \tag{1.3.34}$$

This means that we can always project out the stripped $\mathcal{X}$, and that any overlap with other terms must come suppressed by at least N_f^{-1}. In a stripped amplitude of $\mathcal{O}(p^4)$ or lower, the stripped amplitude can not contain any powers of N_f^{-1} due to the decoupling of the singlet, so there can be no overlap for arbitrary N_f. This proves that stripped amplitudes are unique at $\mathcal{O}(p^4)$ or below.

At higher orders, things are not as simple, since there are possibly many factors of N_f^{-1}. This would allow mixing between different stripped $\mathcal{X}$'s, threatening to break uniqueness. However, it can be resolved by expressing $\mathcal{X}^{a_1 \cdots a_n}$ as a polynomial in N_f^{-1},

$$\mathcal{X}^{a_1 \cdots a_n} = \mathcal{X}_0^{a_1 \cdots a_n} + \frac{1}{N_f} \mathcal{X}_1^{a_1 \cdots a_n} + \frac{1}{N_f^2} \mathcal{X}_2^{a_1 \cdots a_n} + \dots \tag{1.3.35}$$

such that each $\mathcal{X}_i^{a_1 \cdots a_n}$, and therefore also its stripped counterpart, is independent of N_f^{-1}. We then apply the projection to each $\mathcal{X}_i^{a_1 \cdots a_n}$ independently, and ignore the $\mathcal{O}(N_f^{-1})$

completely. Thus, stripped amplitudes, vertex factors, and other analogous quantities are unique to all orders.[7]

The proof holds for general N_f, but for any specific N_f, there may be additional relations between the generators that break the uniqueness. The Cayley-Hamilton relations provide such relations for small N_f. However, we always assume that the relations have been "exhausted" by removing terms from the Lagrangian, so that they do not affect the uniqueness of stripped amplitudes.

This proof in this section has significant consequences. Most importantly, it guarantees the correctenss of our method of flavour ordering with split traces and singlets: gathering all flavour-ordered pieces of the full amplitude is guaranteed to equal the unique stripped amplitude. Also, uniqueness allows many properties of the full amplitude to carry over to the stripped amplitude, as is discussed below.

A second consequence is worthy of note. The full amplitude of some $\mathcal{O}(p^N)$ n-particle process is constructed from $|\mathcal{R}(N, n)|$ different stripped amplitudes. When summed over permutations according to eq. (1.3.12), the total number of flavour structures grows to

$$\mathcal{N}_{N,n} \sim \sum_{R \in \mathcal{R}(N,n)} \frac{|\mathcal{S}_n|}{|\mathbb{Z}_R|}, \tag{1.3.36}$$

which is a very rapidly growing number — even at $\mathcal{O}(p^2)$, $\mathcal{N}(2, n) \sim (n-1)!$. Since the flavour structures are not truly orthogonal, the expression for the cross section of the process, proportional to $\mathcal{M}_{N,n}^{a_1 \cdots a_n} [\mathcal{M}_{N,n}^{a_1 \cdots a_n}]^\dagger$, grows in length as $(\mathcal{N}_{N,n})^2$. However, the expression for the cross section contracts the flavour structures as in eq. (1.3.32), which suppresses products of non-equivalent flavour structures by a factor of N_f^{-1} for each difference (or N_f^{-2} in the single-trace case). Therefore, in the limit $N_f \to \infty$, flavour structures *are* orthogonal, and the cross section only grows as $\mathcal{N}_{N,n}$. Even with finite N_f, most cross-terms will be heavily suppressed, and can most likely be ignored.

An alternative approach would be to construct other bases for flavour space that are more orthogonal than the trace bases used here, as is done in perturbative QCD by [44]. Such methods have so far not been applied in the present context.

4 NLSM amplitudes

In this section, we introduce and generalise several concepts related to NLSM amplitudes and flavour-ordering.

---

[7]This uniqueness is of course only up to a permutation in $\mathbb{Z}_R$, but since we sum over those in the definition of the stripped quantity, they are unique for our purposes.

4.1 Adler zeroes and soft limits

In any effective field theory emerging from the spontanous beaking of a global symmetry, the amplitude possesses the so-called Adler zero,

$$\lim_{\varepsilon \to 0} \mathcal{M}^{a_1 \cdots a_n}(p_1, \ldots, \varepsilon p_i, \ldots, p_n) = 0, \tag{1.4.1}$$

for any i [12, 45]. The approach to zero will generally go as ε^σ, where the *soft degree* $\sigma \geq 1$ can be used to classify and construct EFTs [13, 46]. The NLSM has $\sigma = 1$. Due to the orthogonality of flavour structures and the uniqueness of stripped amplitudes, Adler zeroes may only exist in the full amplitude if they also exist, with the same soft degree, in the stripped amplitudes. Therefore, eq. (1.4.1) and any statement relying on it can equally well be applied to the stripped amplitudes.

The Adler zeroes may be used as a starting point to construct amplitudes through recursion relations [11, 15]. For our purposes, however, their main use is in validating the correctness of complicated stripped amplitudes. Since far from every term in the amplitude is proportional to p_i, the Adler zero must manifest itself through intricate cancellations. Therefore, any error in the amplitude is extremely likely to give a finite right-hand side in eq. (1.4.1).

Beside the Adler zeroes, there also exists the *double soft limit*, where two momenta are sent to zero at the same rate. It turns out that the double soft limit of any $(n+2)$-particle amplitude can be expressed in terms of n-particle amplitudes with the soft particles removed; for the NLSM, the specific form is

$$\lim_{\varepsilon \to 0} \mathcal{M}^{aba_1 \cdots a_n}_{m,n+2}(\varepsilon p, \varepsilon q, p_1, \ldots, p_n) =$$
$$-\frac{1}{4F^2} \sum_{i=1}^{n} f^{abc} f^{a_i dc} \frac{p_i \cdot (p-q)}{p_i \cdot (p+q)} \mathcal{M}^{a_1 \cdots a_{(i-1)} d a_{(i+1)} \cdots a_n}_{m,n}(p_1, \cdots, p_n). \tag{1.4.2}$$

This was conjectured in [47] and proven in [36]. Like the Adler zero, it can be projected to a relation for stripped amplitudes, although the projection is not entirely trivial. The result for single-trace flavour structures is given in [36]. We derive the counterpart for general flavour structures in appendix C, with the result being as follows. At any order m in the power counting and for any flavour split $R \in \mathcal{R}(m, n+2)$, the double soft limit

$$\lim_{\varepsilon \to 0} \mathcal{M}_{m,R}(p_1, \ldots, p_{i-1}, \varepsilon p_i, \ldots, \varepsilon p_j, p_{j+1}, \ldots, p_{n+2}) \tag{1.4.3}$$

is nonzero if the indices $i-1$, i, j and $j+1$ are consecutive and lie within the same trace; we will call this condition $\mathcal{C}$. It is also nonzero if the indices can be made to satisfy $\mathcal{C}$ by applying a permutation in $\mathbb{Z}_R$ and possibly swapping i and j. In all other cases, the double soft limit is zero.

Since $\mathcal{M}_{m,R}$ is invariant under $\mathbb{Z}_R$, we can without loss of generality assume that $\mathcal{C}$ holds whenever the double soft limit is nonzero. Assuming this, the double soft limit is

$$\lim_{\varepsilon \to 0} \mathcal{M}_{m,R}(p_1, \ldots, \varepsilon p_i, \varepsilon p_j, \ldots, p_{(n+2)})$$

$$= \frac{1}{4F^2} \left(\frac{p_{(j+1)} \cdot (p_i - p_j)}{p_{(j+1)} \cdot (p_i + p_j)} - \frac{p_{(i-1)} \cdot (p_i - p_j)}{p_{(i-1)} \cdot (p_i + p_j)} \right) \mathcal{M}_{m,R'}(p_1, \ldots, p_{(i-1)}, p_{(j+1)}, \ldots, p_{(n+2)}),$$

$$\text{(1.4.4)}$$

where $R' \in \mathcal{R}(m, n)$ is R with the location of the soft particles removed and $j = i + 1$. The result, which generalises that given in [36], is quite remarkable: for properly chosen i, j, the double soft limit amounts to removing the soft particles from the amplitude and multiplying by a simple kinetic factor. The factor is similar to those that arise in IR divergences, which is understandable — both arise from propagators going on-shell in the soft (IR) limit.

4.2 Generalised Mandelstam invariants

In order to express stripped amplitudes in a way that naturally includes on-shellness and conservation of momentum, we will employ bases of generalised Mandelstam invariants in the form

$$s_{ijk\cdots} = (p_i + p_j + p_k + \ldots)^2, \tag{1.4.5}$$

In this notation, the standard 4-particle Mandelstam invariants are

$$s = s_{12} = s_{34}, \qquad t = s_{13} = s_{24}, \qquad u = s_{23} = s_{41}. \tag{1.4.6}$$

Since $s + t + u = 0$, this basis is overcomplete, and one element can be removed. We will generally use bases where $ijk\ldots$ are consecutive, so we choose to keep $\{s, u\}$ as the 4-particle basis.

For n particles, the products of momenta are related to the invariants with consecutive indices through

$$2p_i \cdot p_{i+1} = s_{i(i+1)},$$
$$2p_i \cdot p_{i+2} = s_{i(i+1)(i+2)} - s_{i(i+1)} - s_{(i+1)(i+2)},$$
$$j > i + 2: \quad 2p_i \cdot p_j = s_{i\cdots j} - s_{i\cdots(j-1)} - s_{(i+1)\cdots j} + s_{(i+1)\cdots(j-1)}. \tag{1.4.7}$$

Based on this, a complete basis of invariants for $n = 6$ is

$$\mathcal{B}_{\{6\}} = \left\{ s_{12}, s_{23}, s_{34}, s_{45}, s_{56}, s_{61}, \; s_{123}, s_{234}, s_{345} \right\}, \tag{1.4.8}$$

where s_{456} etc. are not needed due to conservation of momentum in the form

$$s_{i\cdots(i+k-1)} = s_{(i+k)\cdots(i-1)}, \tag{1.4.9}$$

with indices cycling around from n to 1. The form of $\mathcal{B}_{\{6\}}$ can be carried on to any even n, giving

$$\mathcal{B}_{\{n\}} = \left\{ s_{12}, s_{23}, \ldots, s_{n(n-1)}, s_{n1}, \quad s_{123}, s_{234}, \ldots, s_{n12}, \quad \ldots, \right.$$
$$\left. s_{12\cdots(n/2)}, \ldots, s_{(n/2-1)\cdots(n-1)} \right\}. \tag{1.4.10}$$

This contains $n(n-3)/2$ invariants, which is also the number of independent products that can be formed from $\{p_1, \ldots, p_n\}$ with all $p_i^2 = 0$.[8] Note that all invariants have consecutive indices.

These bases are only linearly independent in sufficiently high spacetime dimensions D. If $D < n + 1$, the Gram determinant gives relations among the basis elements. In practice, these relations are so algebraically messy that we have found it simpler to always work in arbitrary D.

Mandelstam invariants have further benefits beyond taking care of on-shellness and conservation of momentum. In an n-point $\mathcal{O}(p^2)$ single trace flavour-ordered tree diagram, all propagators carry a momentum q such that $q^2 \in \mathcal{B}_{\{n\}}$. Therefore, $\mathcal{O}(p^2)$ stripped amplitudes will never contain a denominator with a sum of several invariants, making their algebraic handling simpler. This is not true for diagrams with multi-trace flavour structures. It is also not possible to find a different basis that contains all squared propagator momenta in the general case; for instance, the set of all possible q^2 under $R = \{2, 6\}$ is not linearly independent.

Another use of Mandelstam invariants is the shortening of stripped amplitudes. As a consequence of invariance under $\mathbb{Z}_R$, any stripped amplitude can be written in the form

$$\mathcal{M}_{m,R}(p_1, \ldots, p_n) = \text{(simpler expression)} + [\mathbb{Z}_R], \tag{1.4.11}$$

where we use a shorthand for the sum over $\mathbb{Z}_R$, generalising the familiar idiom "+cycl.". The simpler expression is rather obvious for simple amplitudes, but for more complicated cases, it is an enormous aid to readability.

Any stripped amplitude can be simplified as above by separating it into simple terms, separating the terms into cosets under $\mathbb{Z}_R$, and picking a single representative from each coset. The "simpler expression" in eq. (1.4.11) will then be the sum of the representatives. For $R = \{n\}$, this works because for any $s_{ij\ldots} \in \mathcal{B}_{\{n\}}$ and $\sigma \in \mathbb{Z}_{\{n\}}$, applying σ to the indices of $s_{ij\ldots}$ yields another element in $\mathcal{B}_{\{n\}}$. Thus, $\mathcal{B}_{\{n\}}$ can be said to be *closed* under $\mathbb{Z}_{\{n\}}$. However, the basis given in eq. (1.4.10) is not closed under any $\mathbb{Z}_R$ with $R \neq \{n\}$ (with the sole exception of $R = \{2, 2\}$), so the separation into cosets fails. Simplifying general amplitudes therefore requires either painstaking manual work, or a Mandelstam basis that is closed under $\mathbb{Z}_R$. We have no general method of finding such bases. In appendix D, we present closed bases for $\mathbb{Z}_{\{2,4\}}$, $\mathbb{Z}_{\{3,3\}}$ and $\mathbb{Z}_{\{2,2,2\}}$. These cover all flavour structures that appear for $n \leq 6$.

4.3 Diagram generation

For most amplitudes presented here, the number of diagrams is small enough that they are easily found by hand, but above a dozen or so diagrams, this becomes a slow and error-prone process. We therefore automated the process by designing a program called FODGE

---

[8]There are $n(n + 1)/2$ ways to form products of pairs of p_i, $i \in \{1, \ldots, n\}$. Of these, n vanish due to $p_i^2 = 0$. Conservation of momentum implies that $p_n = \sum_{i=1}^{n-1} p_i$, which gives n linear combinations among the remaining products, reducing the number of independent ones to $n(n + 1)/2 - 2n = n(n - 3)/2$.

(ꜰʟᴀᴠᴏᴜʀ-ᴏʀᴅᴇʀᴇᴅ ᴅɪᴀɢʀᴀᴍ ɢᴇɴerator) written in C++.[9] It produces TikZ code for drawing the diagrams, and generates the input to a set of ꜰᴏʀᴍ procedures that compute the amplitudes.[10] The same procedures were used with manual input for computing simpler amplitudes. Inspiration was taken from the diagram generator used in [50, 51].

The diagram generation works recursively. A list of all $\mathcal{O}(p^M)$ N-point diagrams can be generated by generating all $\mathcal{O}(p^m)$ n-point diagrams for $m \leq M$ and $n \leq N-2$, and then listing all ways to attach an $\mathcal{O}(p^{2+M-m})$ $(2+N-n)$-point vertex to their external legs. Adding a list of $\mathcal{O}(p^M)$ N-point single-vertex diagram and removing duplicates completes the list. The number of duplicates can be reduced by restricting m and n.

The number of independent labellings on each diagram must then be determined. Representing diagrams in a way that shows their symmetries turns out to be very difficult when complicated flavour structures are involved. This was not entirely successfully tried in the original ꜰᴏᴅɢᴇ used in [37]. Here, we take a different approach: each diagram is associated with all flavour-ordered labellings of its external legs that give unique kinematic structures. This removes the need to explicitly consider its symmetries; internally, the diagrams can be represented in whatever way is convenient.

As is pointed out below eq. (1.3.17), a kinematic structure is uniquely determined by the propagator momenta it contains. It is easy to see that this holds for any $\mathcal{O}(p^2)$ diagram. At higher orders, it is sufficient to add the order of the vertices at the ends of each propagator. The flavour splits of the vertices are not needed if the overall flavour split of the diagram is provided. For singlet diagrams, we must also specify how the vertex is cycled relative to the singlet propagator, by writing down the momentum carried by a vertex leg adjacent to the propagator. In general, any kinematic factor is uniquely determined by listing all vertices and the momenta carried by their legs, but this can be shown to reduce to these simpler rules when diagrams are flavour-ordered.

Thus, ꜰᴏᴅɢᴇ generates all diagrams of a given order and size, equips each diagram with an arbitrary flavour-ordered labelling, and determines the kinematic factor as described above. $\mathbb{Z}_R$ is then applied to generate all other labellings, but only a subset that gives distinct kinematic factors is kept. If the choice of subset is consistent, equivalent diagrams will always give an identical list of kinematic factors, so duplicates are easily removed.

Knowing the labellings also makes the diagram generation more efficient. There is no need to attach a vertex to several legs that are equivalent to each other under the symmetries. By dividing the set of labels into cosets under $\mathbb{Z}_R$, it is sufficient to attach vertices to legs that, in one of the distinct labellings, carries a coset representative as its label. This reduces the number of generated duplicates.

---

[9] The source code of ꜰᴏᴅɢᴇ can be found at `https://github.com/mssjo/fodge`.

[10] The ꜰᴏʀᴍ procedures can be found at `https://github.com/mssjo/flavour-order`. For ꜰᴏʀᴍ itself, see [48, 49].

5 Explicit amplitudes

Using the methods developed in the previous sections, we have computed several stripped NLSM amplitudes, several of which have not previously been determined. These we discuss in this section.

5.1 4-point amplitudes

These amplitudes are by far the simplest, since their tree-level diagrams contain no propagators and only carry two flavour structures ($\{4\}$ and $\{2,2\}$), or only one in the $\mathcal{O}(p^2)$ case. At $\mathcal{O}(p^6)$ and above, they only receive contributions from the Lagrangian terms with no more than four u_μ's, which is a tiny subset of the total Lagrangian.

The $\mathcal{O}(p^2)$ 4-point amplitude is given by a single diagram and a simple stripped amplitude,

$$-iF^2 \mathcal{M}_{2,\{4\}} = \frac{t}{2}, \tag{1.5.1}$$

where t is the Mandelstam invariant $(p_1 + p_3)^2$. We have pulled factors of i and F over to the left-hand side for clarity. The only independent $\mathcal{O}(p^2)$ kinematic structure that is invariant under $\mathbb{Z}_{\{4\}}$ is t, so the form of the right-hand side could have been guessed based on symmetry.

If we plug eq. (1.5.1) into eq. (1.3.12) and apply some $SU(2)$ group algebra, we recover the familiar $N_f = 2$ amplitude

$$\mathcal{M}_{2,4}^{abcd}(s,t,u) = \frac{-4i}{F^2}\left[s\delta^{ab}\delta^{cd} + t\delta^{ac}\delta^{bd} + u\delta^{ad}\delta^{bc} \right] \tag{1.5.2}$$

with the Mandelstam invariants defined as in section 4.2.

The $\mathcal{O}(p^4)$ 4-point amplitude consists of the two diagrams

$$-iF^4 \mathcal{M}_{4,\{4\}} = 2L_3(u^2 + s^2) + 4L_0 t^2, \tag{1.5.3}$$

$$-iF^4 \mathcal{M}_{4,\{2,2\}} = 8L_1 s^2 + 4L_2(t^2 + u^2), \tag{1.5.4}$$

which includes the simplest example of a flavour split. There are now two independent $\mathbb{Z}_4$-invariant kinematic structures, t^2 and $s^2 + u^2$, and likewise two independent $\mathbb{Z}_{\{2,2\}}$-invariant ones, s^2 and $t^2 + u^2$. All four appear equipped with one LEC each. The full amplitude is analogous to eq. (1.5.2), but with various linear combinations of the LECs and Mandelstam variables in place of s, t and u. The full amplitude agrees with the known results, see [7] and references therein.

The $\mathcal{O}(p^6)$ 4-point amplitude, like its $\mathcal{O}(p^4)$ analogue, has two diagrams,

$$-iF^4\mathcal{M}_{6,\{4\}} = -L_{6,3}t(s^2+u^2) - 2L_{6,4}t^3, \tag{1.5.5}$$

$$-iF^4\mathcal{M}_{6,\{2,2\}} = -2L_{6,1}(t^3+u^3) + \frac{2}{3}L_{6,2}(s^3+t^3+u^3). \tag{1.5.6}$$

As for $\mathcal{O}(p^4)$, there are two independent $\mathbb{Z}_4$-invariant kinematic structures, t^3 and $t(s^2 + u^2)$, and two independent $\mathbb{Z}_{\{2,2\}}$-invariant ones, s^3 and $s(t^2 + u^2)$. These four correspond to the four LECs — $s^3 + t^3 + u^3$ is a linear combination of s^3 and $s(t^2 + u^2)$. The full amplitude agrees with the result in [7].

The $\mathcal{O}(p^8)$ amplitude, like its lower-order analogues, has two diagrams,

$$-iF^4\mathcal{M}_{8,\{4\}} = L_{8,4}s^2u^2 + \frac{1}{2}L_{8,5}t^2(s^2+u^2) + L_{8,6}t^4, \tag{1.5.7}$$

$$-iF^4\mathcal{M}_{8,\{2,2\}} = L_{8,1}s^2(t^2+u^2) + L_{8,2}(t^4+u^4) + 2L_{8,3}t^2u^2, \tag{1.5.8}$$

There are now three independent $\mathbb{Z}_4$-invariant kinematic structures, t^4, $t^2(s^2 + u^2)$ and s^2u^2, and correspondingly three for $\mathbb{Z}_{\{2,2\}}$. This is reflected in the six LECs.

Similarly, $\mathcal{M}_{10,\{4\}}$ will be a linear combination of s^5, s^3tu and st^2u^2, and $\mathcal{M}_{10,\{2,2\}}$ will be a linear combination of t^5, t^3us and tu^2s^2, since these are the only independent $\mathcal{O}(p^{10})$ kinematic structures that are invariant under $\mathbb{Z}_{\{4\}}$ and $\mathbb{Z}_{\{2,2\}}$, respectively. The coefficients will be linear combinations of the LECs of the terms in $\mathcal{L}_{10}$ that only contain four u_μ's. These terms, along with the rest of $\mathcal{L}_{10}$, have not yet been studied. The same pattern can be applied to $\mathcal{O}(p^{12})$ and beyond.

Note that the above discussion is fully compatible with section 5 in [40] where we have two functions with properties $B(s,t,u) = B(u,t,s)$ and $C(s,t,u) = C(s,u,t)$. Independent combinations in B at order p^{2n} are made from $t^{n-2i}(s-u)^{2i}$ and in C from $s^{n-2i}(t-u)^{2i}$.

5.2 The $\mathcal{O}(p^2)$ 6- and 8-point amplitudes

The leading order in the power counting offers a relatively simple playground for flavour-ordering, free from splittings and singlets. It is relatively well explored, and the amplitudes presented here were also calculated in [36] using different methods.

The 6-point amplitude is given by the diagrams

$$\tag{1.5.9}$$

Each diagram represents the sum of all distinct labellings of its legs, as described in section 3. The amplitude is
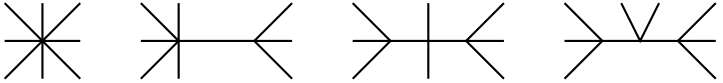
$$
\begin{aligned}
-4iF^4\mathcal{M}_{2,6} = {}& s_{12} + s_{23} + s_{34} + s_{45} + s_{56} + s_{61} \\
& - \frac{(s_{12} + s_{23})(s_{45} + s_{56})}{s_{123}} - \frac{(s_{23} + s_{34})(s_{56} + s_{61})}{s_{234}} - \frac{(s_{34} + s_{45})(s_{61} + s_{12})}{s_{345}},
\end{aligned}
$$

(1.5.10)

which suggests the simplified form as defined in eq. (1.4.11)

$$
-4iF^4\mathcal{M}_{2,6} = \left\{ s_{12} - \frac{1}{2}\frac{(s_{12} + s_{23})(s_{45} + s_{56})}{s_{123}} \right\} + [\mathbb{Z}_6],
$$

(1.5.11)

where $[\mathbb{Z}_6]$ indicates summation over all cyclic permutations. Note the factor of $1/2$, which expresses that the second term has twofold symmetry under rotation, a trait that is shared by the second diagram above.

The 8-point amplitude is given by the diagrams

(1.5.12)
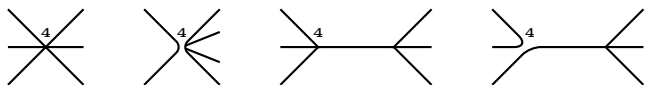
and its stripped amplitude is, in a similarly simplified form,

$$
\begin{aligned}
-8iF^6\mathcal{M}_{2,8} = \bigg\{ & \frac{4s_{12} + s_{1234}}{2} - \frac{(s_{12} + s_{23})(s_{45} + s_{56} + s_{67} + s_{78} + s_{4567} + s_{5678})}{s_{123}} \\
& + \frac{1}{2}\frac{(s_{12} + s_{23})(s_{1234} + s_{4567})(s_{56} + s_{67})}{s_{123}s_{567}} \\
& + \frac{(s_{12} + s_{23})(s_{1234} + s_{45})(s_{67} + s_{78})}{s_{123}s_{678}} \bigg\} + [\mathbb{Z}_8].
\end{aligned}
$$

(1.5.13)

The analogous 10- and 12-point amplitudes, the second of which has not been determined before, are presented in appendix E.2 and E.3.

5.3   The $\mathcal{O}(p^4)$ 6-point amplitude

The calculation of this amplitude hinges decisively on the use of split-trace flavour ordering. It was arrived at independently in a different form by [15] using recursion relations. Our result agrees with theirs. The amplitude is given by the four diagrams

(1.5.14)

Note that unlike its $\mathcal{O}(p^2)$ counterpart, the third diagram is not symmetric due to the asymmetric placement of vertices. The amplitude has a single-trace and a two-trace part. The single-trace stripped amplitude is

$$-iF^6\mathcal{M}_{4,\{6\}} = L_3 \left\{ s_{12} \left(s_{12} + s_{34} + s_{45}\right) - \frac{(s_{12} + s_{23})\left(s_{45}^2 + s_{56}^2\right)}{s_{123}} \right\} + [\mathbb{Z}_6]$$

$$+ 2L_0 \left\{ s_{12} \left(s_{12} + s_{34} + 2s_{45}\right) + s_{123} \left(s_{612} - s_{61}\right) - \frac{(s_{12} + s_{23})\left(s_{45} + s_{56}\right)^2}{s_{123}} \right\} + [\mathbb{Z}_6]$$

$$(1.5.15)$$

In order to find the simplified form of the two-trace part, it is extremely helpful to have a closed Mandelstam basis. In terms of the closed basis $\mathcal{B}_{\{2,4\}} = \{t_1, \ldots, t_9\}$ of eq. (1.4.1), it is

$$-iF^6\mathcal{M}_{4,\{2,4\}} = \frac{L_1}{2} \left\{ t_1\left[t_1 + 2t_2 + t_3 - 3t_5\right] + \frac{(t_2 + t_3 + t_4)^2\left[t_3 - 2t_5\right]}{2t_1} \right\} + [\mathbb{Z}_{\{2,4\}}]$$

$$+ \frac{L_2}{8} \left\{ t_1\left[t_1 + 2t_2 + \frac{t_3}{2} - 3t_5\right] + 4t_7^2 - 2t_9^2 \right.$$

$$\left. + \frac{\left[(t_2 + t_3 + t_4)^2 + 4(t_7 + t_8 + t_9)^2\right]\left[t_3 - 2t_5\right]}{2t_1} \right\} + [\mathbb{Z}_{\{2,4\}}]. \quad (1.5.16)$$

Note that the summation over cyclic permutations is replaced by summation over $\mathbb{Z}_{\{2,4\}}$.

5.4 Further amplitudes

We have computed the $\mathcal{O}(p^6)$ 6-point amplitude, and using the closed Mandelstam bases presented in appendix D, it is possible to present its reduced form given in appendix E.1. The $\mathcal{O}(p^6)$ divergent part is given explicitly in the supplementary material as well as the $\mathcal{O}(p^8)$ expression. The $\mathcal{O}(p^2)$ 10-point is given in appendix E.2. Finally the $\mathcal{O}(p^2)$ 12-point amplitude is given in appendix E.3.

We have also computed several amplitudes whose expressions are too large to overview. They have been verified by checking their Adler zeroes, and in some cases by running brute-force Feynman diagram calculations. Beyond these amplitudes, we have generated the flavour-ordered diagrams of many more amplitudes using our program FODGE, described in section 4.3. Here, we only summarise the number and general properties of the diagrams to give an idea of how the complexity scales. The summary is given in table 1.1.

For all entries labelled "Yes" in table 1.1 that are not included in the main text, the flavour-ordered diagrams are given in the supplementary material. Most of the amplitudes themselves are too long to be practically written down, but they can be generated by using the freely available programs described in section 4.3.

In the table, we note that the number of diagrams grows more rapidly with n (the number of particles) than with N (the power-counting order). Especially when $N > n$, the number of new diagrams is very small. This is also reflected in the computational effort needed: the $\mathcal{O}(p^2)$ 10-point, $\mathcal{O}(p^6)$ 8-point and $\mathcal{O}(p^8)$ 6-point amplitudes took approximately 10 minutes each to calculate with FORM [48, 49], while the $\mathcal{O}(p^4)$ 10-point amplitude took almost

Table 1.1: Summary of the number of $\mathcal{O}(p^N)$ n-point diagrams. The $SU(N_f)$ column shows the general count, and the $U(N_f)$ column shows the count without singlet diagrams. The $SU(3)$ and $SU(2)$ columns show the number of distinct diagrams left when some Lagrangian terms have been eliminated using the Cayley-Hamilton relation as discussed in section 2.1. Note that the distinction for $N_f = 2$ assumes we remove the L_1 and L_2 term and emerges first at $\mathcal{O}(p^4)$. For $N_f = 3$ it emerges first at $\mathcal{O}(p^6)$. The distinction between SU and U also emerges first at $\mathcal{O}(p^6)$. The rightmost column states whether an amplitude has been computed by us, and provides references to the explicit amplitudes when possible. Amplitudes marked with an asterisk have to our knowledge not been calculated before; the $\mathcal{O}(p^4)$ 6-point amplitude was recently independently reproduced by [15]. Amplitudes marked with a dagger have been verified with a brute-force Feynman diagram calculation; the remainder rely only on Adler zeroes for verification.

$\mathcal{O}(p^N)$	n	Number of diagrams						Computed?
		$SU(N_f)$	$U(N_f)$	$SU(3)$	$U(3)$	$SU(2)$	$U(2)$	
$\mathcal{O}(p^2)$	4	1	*[same as $SU(N_f)$]*					Yes (1.5.1)
	6	2						Yes (1.5.11)
	8	4						Yes (1.5.13)
	10	16						Yes (1.5.10)
	12	73						Yes* (1.5.11)
	14	414						No
$\mathcal{O}(p^4)$	4	2	*[same as $SU(N_f)$]*			1	1	Yes (1.5.4)
	6	4				2	2	Yes†* (1.5.15-1.5.16)
	8	18				8	8	Yes†* (1.5.2-1.5.9)
	10	90				43	43	Yes*
	12	577				283	283	No
$\mathcal{O}(p^6)$	4	2	2	2	2	1	1	Yes (1.5.6)
	6	10	9	9	8	4	3	Yes†*
	8	50	45	48	43	18	14	Yes*
	10	360	318	348	316	129	98	No
$\mathcal{O}(p^8)$	4	2	2	2	2	1	1	Yes* (1.5.8)
	6	11	10	10	9	4	3	Yes*
	8	105	85	97	77	34	21	No

and hour and the $\mathcal{O}(p^2)$ 12-point amplitude took over 2 days. At high N, the calculation of vertex factors takes significant time, while at high n, the conversion to Mandelstam variables is very time-consuming due to the large dimension of the kinematic space.

As the table shows, we have calculated all amplitudes with less than 100 diagrams, excluding $N \geq 10$, where the Lagrangian is not yet known. If we decide to push the frontier of large n further in the future, we expect the required computational effort to be severe.

6 Conclusions

In this work we have extended flavour ordering methods to include multiple traces and higher orders in derivatives. The uniqueness of the method relies on the extended orthogonality relation eq. (1.3.32). We implemented the constraints in a diagram generator and used this then to calulate a number of amplitudes in the NLSM with more legs and derivatives than obtained previously.

Our methods are fairly constrained in which models they can be applied to, since they hinge on the existence of flavour structures and the contraction identity eq. (1.3.15). On the other hand, they are readily extended to extremely high-order and many-particle amplitudes. They may also have some applicability to loop diagrams and massive particles under ChPT. A tentative discussion of these possibilities can be found in [37].

Flavour-ordering serves as an enhancement of the standard diagrammatic approach, and as such is rather brute-force in nature. This contrasts with the recursive approach developed in [11], in which subtler properties such as soft limits play a much clearer role. These methods can also be applied to a wider range of models. The downside is that practical calculations require algebraic manipulations that are not entirely obvious. Flavour-ordering calculations can be very extensive, but are mathematically trivial and easily automated. Further developments of recursion relations in [15] have offset the algebraic difficulties, but soft recursion retains the fundamental limitation that recursive calculation of an $\mathcal{O}(p^m)$ n-point amplitude requires $n > m$. Therefore, the $\mathcal{O}(p^6)$ 6-point can not be reached by such means, and must be supplied as a seed amplitude if $\mathcal{O}(p^6)$ amplitudes are to be calculated for more than 6 particles. For this, our methods seem to be the only viable option other than brute-force Feynman diagrams.

Acknowledgements

A The NNLO NLSM Lagrangian

The NNLO ChPT Lagrangian $\mathcal{L}_6^{\text{ChPT}}$ was first determined in [39]. It has 21 terms that do not vanish when external fields are removed, but this turns out to be an overcomplete basis for the NLSM. In tandem with the NNNLO ChPT Lagrangian $\mathcal{L}_8$ in [40], the authors of that paper produced a version of $\mathcal{L}_6^{\text{ChPT}}$ where removing external fields yields a *minimal*

NLSM Lagrangian with 19 terms. It was not published there, but we present it in table 1.2. The first 135 terms of the Lagrangian in [40] constitute a minimal NLSM Lagrangian $\mathcal{L}_8$.

The Lagrangian of [39] is expressed as

$$\mathcal{L}_6^{\text{ChPT}} = \sum_{i=1}^{112} K_i Y_i, \tag{1.1.1}$$

where K_i are LECs and Y_i are monomials in the fields. Terms 1–6 and 49–63 remain when external fields are removed. All monomials except Y_1, Y_2 and Y_6 correspond directly to monomials in the minimal NLSM Lagrangian

$$\mathcal{L}_6 = \sum_{i=1}^{19} L_{6,i} \mathcal{O}_{6,i}, \tag{1.1.2}$$

where $L_{6,i}$ are LECs and $\mathcal{O}_{6,i}$ are monomials. The remaining Y_i can be decomposed in terms of $\mathcal{O}_{6,i}$ using the relations described in [40], which for the NLSM simplify to

$$\nabla_\mu u_\nu = \nabla_\nu u_\mu, \qquad \nabla_\mu u^\mu = 0,$$
$$[\nabla_\mu, \nabla_\nu] u_\rho = \frac{1}{4}[[u_\mu, u_\nu], u_\rho]. \tag{1.1.3}$$

This yields the relations

$$Y_1 = -3\mathcal{O}_{6,3} + \mathcal{O}_{6,4} + \mathcal{O}_{6,15} - 2\mathcal{O}_{6,16} + \frac{1}{2}\mathcal{O}_{6,17} + \mathcal{O}_{6,18} - \frac{1}{2}\mathcal{O}_{6,19}, \tag{1.1.4}$$

$$Y_2 = -8\mathcal{O}_{6,2} + 2\mathcal{O}_{6,8} - 2\mathcal{O}_{6,9}, \tag{1.1.5}$$

$$Y_6 = 4\frac{\mathcal{O}_{6,2} - \mathcal{O}_{6,1}}{3} - 2\frac{\mathcal{O}_{6,9} - \mathcal{O}_{6,8}}{3} - 2\frac{\mathcal{O}_{6,12} - \mathcal{O}_{6,11}}{3}. \tag{1.1.6}$$

Furthermore, some factors of 2 appear since [39] includes higher derivatives in terms of $h_{\mu\nu} \equiv \nabla_\mu u_\nu + \nabla_\nu u_\mu$, which is just $2\nabla_\mu u_\nu$ in the NLSM.

A.1 Renormalisation

NNLO ChPT was renormalised in [52], based on [39]. For renormalisation in the NLSM, we transfer those results to the minimal Lagrangian given in table 1.2. For details on the renormalisation, see [52] and sources therein. At NLO, it is performed by splitting the LECs as

$$L_i = (c\mu)^{d-4}\left[L_i^r(\mu, d) + \Gamma_i \Lambda\right], \qquad \Lambda = \frac{1}{16\pi^2(d-4)}. \tag{1.1.7}$$

The measurable LECs are given by $L_i^r(\mu, d)$ as $d \to 4$, with

$$\Gamma_0 = \frac{N_f}{48}, \qquad \Gamma_1 = \frac{1}{16}, \qquad \Gamma_2 = \frac{1}{8}, \qquad \Gamma_3 = \frac{N_f}{24}. \tag{1.1.8}$$

Table 1.2: The terms of the NNLO NLSM Lagrangian $\mathcal{L}_6$. The numbering of the terms is taken from material produced in tandem with [40]; the choice of which terms to keep at small N_f (see section 2.1) is carried over from [39]. The rightmost column shows how K_i combine to give $L_{6,i}$ when the overcomplete Lagrangian is decomposed.

Monomial	Number in $\mathcal{L}_6$			Relation to [39]
	$SU(N_f)$	$SU(3)$	$SU(2)$	
$\langle u_\mu \nabla_\nu u_\rho\rangle\langle u^\mu \nabla^\nu u^\rho\rangle$	1	1	1	$K_4 - \frac{4}{3}K_6$
$\langle u_\mu \nabla_\nu u_\rho\rangle\langle u^\rho \nabla^\mu u^\nu\rangle$	2	2		$-8K_2 + \frac{4}{3}K_6$
$\langle u_\mu \nabla_\nu u^\mu u_\rho \nabla^\nu u^\rho\rangle$	3	3	2	$4K_5 - 3K_1$
$\langle u_\mu \nabla_\nu u_\rho u^\mu \nabla^\nu u^\rho\rangle$	4	4	3	$4K_3 + K_1$
$\langle u_\mu u^\mu\rangle\langle u_\nu u^\nu\rangle\langle u_\rho u^\rho\rangle$	5			K_{51}
$\langle u_\mu u^\mu\rangle\langle u_\nu u_\rho\rangle\langle u^\nu u^\rho\rangle$	6			K_{56}
$\langle u_\mu u^\nu\rangle\langle u_\nu u^\rho\rangle\langle u_\rho u^\mu\rangle$	7			K_{63}
$\langle u_\mu u^\mu\rangle\langle u_\nu u^\nu u_\rho u^\rho\rangle$	8	5		$K_{50} + 2K_2 + \frac{2}{3}K_6$
$\langle u_\mu u^\mu\rangle\langle u_\nu u_\rho u^\nu u^\rho\rangle$	9	6		$K_{57} - 2K_2 - \frac{2}{3}K_6$
$\langle u_\mu u^\mu u_\nu\rangle\langle u^\nu u_\rho u^\rho\rangle$	10	7		K_{53}
$\langle u_\mu u_\nu\rangle\langle u^\mu u^\nu u_\rho u^\rho\rangle$	11			$K_{55} + \frac{2}{3}K_6$
$\langle u_\mu u_\nu\rangle\langle u^\mu u_\rho u^\nu u^\rho\rangle$	12			$K_{62} - \frac{2}{3}K_6$
$\langle u_\mu u_\nu u_\rho\rangle\langle u^\mu u^\nu u^\rho\rangle$	13			K_{59}
$\langle u_\mu u_\nu u_\rho\rangle\langle u^\mu u^\rho u^\nu\rangle$	14			K_{61}
$\langle u_\mu u^\mu u_\nu u^\nu u_\rho u^\rho\rangle$	15	8	4	$K_{49} + K_1$
$\langle u_\mu u^\mu u_\nu u_\rho u^\nu u^\rho\rangle$	16	9	5	$K_{54} - 2K_1$
$\langle u_\mu u^\mu u_\nu u_\rho u^\rho u^\nu\rangle$	17	10		$K_{52} + \frac{1}{2}K_1$
$\langle u_\mu u_\nu u^\mu u_\rho u^\nu u^\rho\rangle$	18	11		$K_{60} + K_1$
$\langle u_\mu u_\nu u_\rho u^\mu u^\nu u^\rho\rangle$	19	12	6	$K_{58} - \frac{1}{2}K_1$

Likewise, at NNLO the LECs are split as

$$L_{6,i} = \frac{(c\mu)^{2(d-4)}}{F^2}\left[L_{6,i}^r(\mu,d) - \Gamma_i^{(2)}\Lambda^2 - \left(\Gamma_i^{(1)} + \Gamma_i^{(L)}(\mu,d)\right)\Lambda\right]. \tag{1.1.9}$$

The Γ's for the corresponding renormalisation of the K_i are given in [52]. Using the rightmost column of table 1.2, the renormalisation of the minimal $\mathcal{L}_6$ is given in table 1.3.

A.2 Explicit divergences

In analogy with eq. (1.1.7), we define

$$\mathcal{M}_{4,R} = (c\mu)^{d-4}\left[\mathcal{M}_{4,R}^r(\mu,d) + \mathcal{M}_{4,R}^{(1)}\Lambda\right], \tag{1.1.10}$$

where $\mathcal{M}_{4,R}$ is some $\mathcal{O}(p^4)$ stripped amplitude, $\mathcal{M}_{4,R}^r(\mu,d)$ is the corresponding measurable amplitude expressed in terms of $L_i^r(\mu,d)$, and $\mathcal{M}_{6,R}^{(1)}$ is its divergence.

Table 1.3: The coefficients used to renormalise $\mathcal{L}_6$ as per eq. (1.1.9). L_i^r are the renormalised LECs of $\mathcal{L}_4$ as per eqs. (1.1.7) and (1.1.8)). Note how the highest power of N_f in $\Gamma_i^{(1,2)}$ is 3 minus the number of traces in $\mathcal{O}_{6,i}$.

i	$\Gamma_i^{(2)}$	$16\pi^2\Gamma_i^{(1)}$	$\Gamma_i^{(L)}$
1	$-\frac{5}{72}N_f$	$-\frac{19}{864}N_f$	$-\frac{4}{3}L_3^r - 4L_0^r$
2	$-\frac{5}{9}N_f$	$\frac{1}{864}N_f$	$-\frac{16}{3}L_3^r - \frac{8}{3}L_0^r - \frac{8}{3}N_f L_2^r - 8N_f L_1^r$
3	$-\frac{5}{16} - \frac{5}{96}N_f^2$	$-\frac{1}{96} + \frac{35}{6912}N_f^2$	$-\frac{25}{6}L_2^r - \frac{5}{3}L_1^r - \frac{23}{12}N_f L_3^r - \frac{7}{6}N_f L_0^r$
4	$\frac{5}{16} + \frac{5}{288}N_f^2$	$\frac{1}{96} - \frac{17}{6912}N_f^2$	$\frac{25}{6}L_2^r + \frac{5}{3}L_1^r + \frac{3}{4}N_f L_3^r + \frac{1}{6}N_f L_0^r$
5	$\frac{1}{64}$	$\frac{5}{256}$	$\frac{1}{4}L_2^r$
6	$-\frac{1}{32}$	$\frac{3}{128}$	$-\frac{1}{2}L_2^r$
7	$-\frac{1}{8}$	$-\frac{1}{32}$	$-2L_2^r$
8	$\frac{1}{24}N_f$	$\frac{25}{576}N_f$	$\frac{17}{12}L_3^r + \frac{13}{6}L_0^r + \frac{2}{3}N_f L_2^r - \frac{1}{2}N_f L_1^r$
9	$-\frac{1}{96}N_f$	$-\frac{5}{1152}N_f$	$-\frac{13}{24}L_3^r - \frac{29}{12}L_0^r + \frac{1}{12}N_f L_2^r$
10	$-\frac{1}{64}N_f$	$-\frac{5}{256}N_f$	$-\frac{5}{4}L_3^r + L_0^r$
11	$-\frac{5}{144}N_f$	$-\frac{1}{1728}N_f$	$\frac{2}{3}L_3^r + \frac{4}{3}L_0^r - \frac{2}{3}N_f L_2^r$
12	$-\frac{13}{144}N_f$	$-\frac{53}{1728}N_f$	$-\frac{7}{6}L_3^r - \frac{19}{3}L_0^r - \frac{1}{3}N_f L_2^r$
13	$-\frac{1}{192}N_f$	$\frac{65}{2304}N_f$	$-\frac{3}{4}L_3^r + L_0^r$
14	$\frac{7}{192}N_f$	$-\frac{23}{2304}N_f$	$\frac{5}{4}L_3^r + L_0^r$
15	$\frac{5}{48} + \frac{1}{144}N_f^2$	$-\frac{7}{576} - \frac{25}{6912}N_f^2$	$\frac{5}{6}L_2^r + \frac{5}{3}L_1^r + \frac{1}{6}N_f L_3^r + \frac{1}{3}N_f L_0^r$
16	$-\frac{5}{24} - \frac{1}{96}N_f^2$	$-\frac{19}{576} + \frac{5}{1152}N_f^2$	$-\frac{2}{3}L_2^r - \frac{16}{3}L_1^r - \frac{1}{4}N_f L_3^r - \frac{1}{6}N_f L_0^r$
17	$\frac{5}{96} - \frac{1}{48}N_f^2$	$\frac{43}{576} + \frac{49}{13824}N_f^2$	$\frac{1}{4}L_2^r + \frac{7}{6}L_1^r - \frac{2}{3}N_f L_3^r - \frac{2}{3}N_f L_0^r$
18	$\frac{5}{48} + \frac{1}{64}N_f^2$	$-\frac{67}{576} - \frac{7}{1728}N_f^2$	$\frac{1}{6}L_2^r + 3L_1^r + \frac{17}{24}N_f L_3^r + \frac{1}{12}N_f L_0^r$
19	$-\frac{5}{96} - \frac{1}{144}N_f^2$	$\frac{25}{288} + \frac{5}{4608}N_f^2$	$-\frac{7}{12}L_2^r - \frac{1}{2}L_1^r - \frac{1}{3}N_f L_3^r$

Using this notation and eq. (1.1.8), the divergence of the $\mathcal{O}(p^4)$ 4-point amplitude eq. (1.5.4) is

$$-iF^4\mathcal{M}_{4,\{4\}}^{(1)} = N_f\frac{s^2 + t^2 + u^2}{12}, \qquad -iF^4\mathcal{M}_{4,\{2,2\}}^{(1)} = \frac{s^2 + t^2 + u^2}{2}. \tag{1.1.11}$$

These kinematic terms are highly symmetric, more so than the amplitude itself. The divergences of the 6-point amplitude eqs. (1.5.15) and (1.5.16)) are

$$-iF^6\mathcal{M}_{4,\{6\}}^{(1)} = \frac{N_f}{12}\left\{ s_{12}\left(s_{12} + s_{34} + \frac{3s_{45}}{2} + s_{234}\right) - \frac{s_{123}s_{234}}{2}\right.$$
$$\left. - \frac{(s_{12} + s_{23})(s_{45}^2 + s_{56}^2) + s_{12}s_{23}(s_{45} + s_{56})}{s_{123}}\right\} + [\mathbb{Z}_{\{6\}}]$$

$$-iF^6\mathcal{M}_{4,\{2,4\}}^{(1)} = \frac{1}{64}\left\{ 3\left(t_1^2 + 2t_1t_2 + \frac{t_1t_3}{2} - 3t_1t_5\right) + 4t_7^2 - 2t_9^2\right.$$
$$+ \frac{1}{t_1}\left[3(t_2t_3t_4 - 2t_2t_3t_5 - 2t_2t_4t_5 - 2t_3t_4t_5) + 3t_3^2(t_2 + t_4)\right.$$
$$\left.\left. + \frac{3(t_3 - 2t_5)}{2}\left(T_{234}^2 + 2T_{789}^2\right)\right]\right\} + [\mathbb{Z}_{\{2,4\}}]. \tag{1.1.12}$$

We use the closed basis eq. (1.4.1), and $T_{ijk} = t_i + t_j + t_k$.

An $\mathcal{O}(p^6)$ analogue of eq. (1.1.11) can be formed based on eq. (1.1.9):

$$\mathcal{M}_{6,R} = \frac{(c\mu)^{2(d-4)}}{F^2} \left[\mathcal{M}^r_{6,R}(\mu,d) - \mathcal{M}^{(2)}_{6,R}\Lambda^2 - \left(\mathcal{M}^{(1)}_{6,R} + \mathcal{M}^{(L)}_{6,R}(\mu,d) \right) \Lambda \right], \qquad (1.1.13)$$

where $\mathcal{M}^{(2,L)}_{6,R}$ will gain contributions from both eqs. (1.1.7) and (1.1.9).

Using this notation and the above renormalisation, the divergences of the $\mathcal{O}(p^6)$ 4-point amplitude are eq. (1.5.6) are

$$-iF^4 \mathcal{M}^{(1)}_{6,\{4\}} = \frac{1}{72} \left(2t^3 - s^3 - u^3 \right) + \frac{N_f^2}{5184} \left[8t^3 + 35(s^3 + u^3) \right],$$

$$-iF^4 \mathcal{M}^{(2)}_{6,\{4\}} = \frac{5}{12} \left(2t^3 - s^3 - u^3 \right) - \frac{5N_f^2}{72} \left(s^3 + u^3 \right),$$

$$-iF^4 \mathcal{M}^{(L)}_{6,\{4\}} = 10\frac{2L_1^r + 5L_2^r}{9} \left(2t^3 - s^3 - u^3 \right)$$
$$+ \frac{2N_f L_0^r}{9} \left[2t^3 - 7(s^3 + u^3) \right] - \frac{N_f L_3^r}{9} \left[2t^3 + 23(s^3 + u^3) \right], \qquad (1.1.14)$$

$$-iF^4 \mathcal{M}^{(1)}_{6,\{2,2\}} = \frac{N_f}{1296} \left[s^3 + 58(u^3 + t^3) \right],$$

$$-iF^4 \mathcal{M}^{(2)}_{6,\{2,2\}} = -\frac{5N_f}{108} \left[8s^3 + 5(u^3 + t^3) \right],$$

$$-iF^4 \mathcal{M}^{(L)}_{6,\{2,2\}} = -16N_f \frac{3L_1^r + L_2^r}{9} \left(s^3 + t^3 + u^3 \right)$$
$$+ \frac{2L_0^r}{9} \left[23(u^3 + t^3) - 8s^3 \right] - \frac{8L_3^r}{9} \left(u^3 + t^3 + 4s^3 \right), \qquad (1.1.15)$$

with the dependence on (μ, d) left out for compactness. These expressions do not share the simplicity and symmetry of their $\mathcal{O}(p^4)$ counterparts. The analogous divergences of the $\mathcal{O}(p^6)$ 6-point amplitude (appendix E.1) are given in the supplementary material.

B   The orthogonality of flavour structures

Here, we prove the orthogonality relation eq. (1.3.32) used in section 3.6 to prove the uniqueness of stripped amplitudes. It relies on notation defined in that and previous sections.

Let $\sigma, \rho \in \mathcal{S}_n$ be two permutations, and Q, R be two n-index flavour splittings. We use these to build two flavour structures, and begin by focusing on the trace in $\mathcal{F}_\sigma(Q)$ that contains $a_{\sigma(n)}$ and the trace in $\mathcal{F}_\rho(R)$ that containis $a_{\rho(m)}$, where we have picked m such that $\rho(m) = \sigma(n)$. If there are more traces present, we leave them as passive "spectators"

for the time being. Then, we use eq. (1.3.15) to contract $a_{\sigma(n)}$ in

$$\mathcal{F}_\sigma(Q) \cdot \left[\mathcal{F}_\rho(R)\right]^* = \left[\langle X a_{\sigma(n-1)} a_{\rho(m-1)} Y\rangle - \frac{1}{N_f}\langle X a_{\sigma(n-1)}\rangle\langle a_{\rho(m-1)} Y\rangle\right] \cdot \text{(spectators)},$$

(1.2.1)

where the product is defined as in eq. (1.3.32).

From here on, we work only to leading order in N_f, so we can omit the second term above. (Note that we do not do this because N_f is necessarily large, but because we wish to use power counting of N_f to separate orthogonal flavour structures.) We then move on to contracting $\sigma(n-1)$, followed by $\sigma(n-2)$, and so on. Each time we contract $\sigma(n-i)$, the situation may be one of the following cases:

1. $\rho(m-i) = \sigma(n-i)$. We carry on through a special case of the contraction identity eq. (1.3.16), and find

$$\langle X a_{\sigma(n-i)} a_{\sigma(n-i)} Y\rangle = \frac{N_f^2 - 1}{N_f}\langle XY\rangle.$$

(1.2.2)

   This may be repeated as long as there are indices left, and we gain a factor of N_f (plus $\mathcal{O}(N_f^{-1})$, which we ignore) each time.

2. $\rho(m-i) \neq \sigma(n-i)$, but $\rho(m') = \sigma(n-i)$ is in the same trace as $\sigma(n-i)$. Here, eq. (1.3.16) (after some reshuffling of X and Y) gives

$$\langle X a_{\sigma(n-i)} Y a_{\sigma(n-i)}\rangle = \left[\langle X\rangle\langle Y\rangle - \frac{1}{N_f}\langle XY\rangle\right].$$

(1.2.3)

   the second term is suppressed by a factor of N_f^{-1}, and the first must eventually take a detour through eq. (1.2.1) before continuing; in any case, this case falls behind case 1 by at least two factors of N_f.

3. $\rho(m') = \sigma(n-i)$ is in a different trace than $\sigma(n-i)$. This forces us to bring in the spectator trace containing $\rho(m')$ and go back to eq. (1.2.1), so this case falls behind case 1 by at least one factor of N_f.

4. The trace is empty. We gain a factor of $\langle \mathbb{1}\rangle = N_f$, and if there are no spectator traces left, we are done. Otherwise, we bring in the next pair of spectators and continue from eq. (1.2.1).

If $Q = R = \{n\}$ and $\sigma \equiv \rho \pmod{\mathbb{Z}_R}$, we will only encounter case 1 until we finish with a case 4, and will gain a total factor of $N_f^n[1 + \mathcal{O}(N_f^{-2})]$. If $Q = R \neq \{n\}$ on the other hand, we will encounter case 4 at each split, but the leading power of N_f stays the same.

If $\sigma \not\equiv \rho \pmod{\mathbb{Z}_R}$, we must eventually encounter case 2, so this falls behind the $\sigma \equiv \rho \pmod{\mathbb{Z}_R}$ case by at least two powers of N_f. If $Q \neq R$, we will encounter case 3 (without a corresponding case 4) whenever there is a mismatch in the flavour splits, so we will fall

behind the $Q = R$ case by at least one power of N_f. This is the reason for the values of γ stated below eq. (1.3.32).

Thus, we have proven

$$\mathcal{F}_\sigma(Q) \cdot \left[\mathcal{F}_\rho(R)\right]^* = N_f^n \begin{cases} 1 + \mathcal{O}\left(N_f^{-2}\right) & \text{if } Q = R \text{ and } \sigma \equiv \rho \ (\text{mod } \mathbb{Z}_R), \\ \mathcal{O}\left(N_f^{-\gamma}\right) & \text{otherwise } (\gamma \geq 1) \end{cases} \tag{1.2.4}$$

which is eq. (1.3.32).

C   The double soft limit

This appendix provides a derivation of eq. (1.4.4), which is used to calculate the double soft limit of stripped amplitudes. We start by quoting eq. (1.4.2), which is proven in [36] and gives the double soft limit of the full amplitude:

$$\lim_{\varepsilon \to 0} \mathcal{M}_{m,n+2}^{aba_1 \cdots a_n}(\varepsilon p, \varepsilon q, p_1, \ldots, p_n) =$$

$$-\frac{1}{F^2} \sum_{i=1}^n f^{abc} f^{a_i dc} \frac{p_i \cdot (p - q)}{p_i \cdot (p + q)} \mathcal{M}_{m,n}^{a_1 \cdots a_{(i-1)} da_{(i+1)} \cdots a_n}(p_1, \cdots, p_n). \tag{1.3.1}$$

In order to find the corresponding expression for a stripped amplitude, we project it out by contracting both sides with $[\mathcal{F}_{\mathrm{id}}(R)]^*$ over all flavour indices (see eq. (1.3.12) and section 3.6). On the left-hand side of eq. (1.3.1), this will project out $\lim_{\epsilon \to 0} \mathcal{M}_{m,R}(\varepsilon p, \varepsilon q, p_1, \ldots)$. For simplicity, we start with $R = \{n + 2\}$ before moving on to the general multi-trace case. According to eq. (1.3.12), the right-hand side of eq. (1.3.1) has the form (schematically, with kinematic terms omitted)

$$\sum_{\sigma \in \mathcal{S}_n / \mathbb{Z}_n} f^{abc} f^{a_i dc} \langle a_{\sigma(1)} \cdots a_{\sigma(i-1)} da_{\sigma(i+1)} \cdots a_{\sigma(n)} \rangle \tag{1.3.2}$$

plus flavour-split structures, but those can be ignored due to eq. (1.3.32). We have omitted the algebra generators for readability; a_i means t^{a_i}. The structure constants can be contracted in using eq. (1.3.15) and $f^{abc} = -i \langle t^a[t^b, t^c] \rangle$, leaving

$$-\langle a_{\sigma(1)} \cdots a_{\sigma(i-1)} [[a, b], a_{\sigma(i)}] a_{\sigma(i+1)} \cdots a_{\sigma(n)} \rangle. \tag{1.3.3}$$

With appendix B in mind, we immediately see that this is orthogonal to $\mathcal{F}_{\mathrm{id}}(n + 2)$ unless $\sigma = \mathrm{id}$. The nested commutators expand to

$$[[a, b], a_i] = aba_i - baa_i - a_i ab + a_i ba. \tag{1.3.4}$$

Since a comes before b in $\mathcal{F}_{\mathrm{id}}(n+2)$, the second and fourth terms vanish under the projection. Also, ab occurs at the beginning (or, equivalently, the end) of the flavour structure,

so the first term only contributes when $i = 1$, and the third term only contributes when $i = n$. This collapses the sum in eq. (1.3.1) to those two cases, leaving

$$\lim_{\varepsilon \to 0} \mathcal{M}_{m,\{n+2\}}(\varepsilon p, \varepsilon q, p_1, \ldots, p_n) =$$

$$\frac{1}{F^2} \left\{ \frac{p_1 \cdot (p - q)}{p_1 \cdot (p + q)} - \frac{p_n \cdot (p - q)}{p_n \cdot (p + q)} \right\} \mathcal{M}_{m,\{n\}}(p_1, \cdots, p_n). \quad (1.3.5)$$

If we now move on to general R, we see that a and b must reside in the same trace, since the nested commutator on the right-hand side is inside a single trace. This is essentially the condition stated for the validity of eq. (1.4.4), with (p_n, p, q, p_1) mapping to $(p_{i-1}, p_i, p_j, p_{j+1})$. The trace they reside in can be treated exactly like the single-trace flavour structure of eq. (1.3.5), and all other traces in the flavour structure follow along as "spectators", as in a normal application of eq. (1.3.32). The reduction $\{n+2\} \to \{n\}$ in eq. (1.3.5) then generalises to $R \to R'$ as described below eq. (1.4.4). This generalisation therefore results in eq. (1.4.4), which is thereby proven.

D   Closed Mandelstam bases

Here, we show the derivation of closed Mandelstam bases for all 6-particle flavour structures as described in section 4.2. Note that neither basis is unique, and that better basis choices may exist.

D.1   The basis for $R = \{2, 4\}$

This is the only basis other than $\mathcal{B}_{\{6\}}$ that is needed at $\mathcal{O}(p^4)$. This flavour split permits four different propagator momenta (corresponding to the labellings in eq. (1.3.21)). Since $\mathbb{Z}_{\{2,4\}}$ is Abelian and rather small, it is simple to handle, and some inspired trial-and-error gives the closed basis $\mathcal{B}_{\{2,4\}} = \{t_1, \ldots, t_9\}$ with elements[11]

$$t_1 = s_{123}, \quad t_2 = s_{124}, \quad t_3 = s_{125}, \quad t_4 = s_{126},$$

$$t_5 = s_{45} + s_{56} + \frac{s_{125} - s_{123}}{2}, \quad t_6 = s_{45} - s_{56} + \frac{2s_{124} - (s_{124} + s_{125})}{2},$$

$$t_7 = s_{14} + s_{15} + \frac{s_{123} + s_{126}}{2}, \quad t_8 = s_{15} + s_{16} + \frac{s_{123} + s_{124}}{2}, \quad (1.4.1)$$

$$t_9 = s_{14} + s_{16} + \frac{s_{123} + s_{125}}{2}.$$

Under the action of $\mathbb{Z}_{\{2,4\}}$, they transform as

$$21\,3456: \quad \{t_1, \ldots t_6, t_7, t_8, t_9\} \to \{t_1, \ldots t_6, -t_7, -t_8, -t_9\},$$
$$12\,4563: \quad \{t_1, \ldots t_6, t_7, t_8, t_9\} \to \{t_2, t_3, t_4, t_1, t_5, -t_6, +t_8, -t_7, -t_9\}, \quad (1.4.2)$$

[11]This basis is a slight improvement over the one used in [37]. It modifies t_5 and t_6 so that they map to themselves under $\mathbb{Z}_{\{2,4\}}$.

where the first permutation cycles the 2-trace, and the second cycles the 4-trace; together, they generate all of $\mathbb{Z}_{\{2,4\}}$. Note that $\mathbb{Z}_{\{2,4\}}$ does not act as a true permutation on the basis, since some elements change sign. This appears to be unavoidable, but is not a problem — in fact, any complex phase can be applied without hindering simplification.

D.2 The basis for $R = \{3, 3\}$

The group $\mathbb{Z}_{\{3,3\}}$ is generated by the permutations $g_1 = 231\,456$ and $g_2 = 456\,123$. The group is not abelian, which makes its effects less predictable. Among all kinematic invariants, only s_{123} maps to itself under both generators, and is also the only squared propagator momentum permitted by this flavour structure. The other 24 invariants decompose into a sextuplet and two nonets under the group, and can be mapped out in a variant of a Cayley graph:
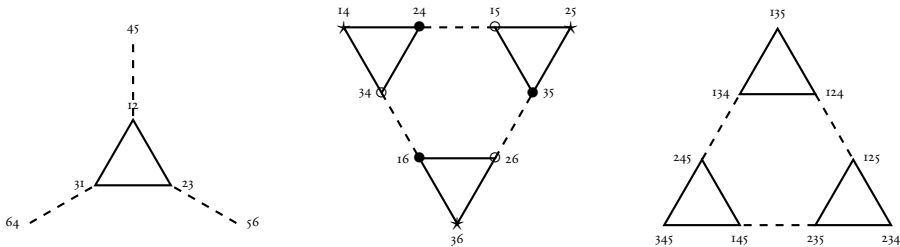
$$(1.4.3)$$

Each node in the graph represents $s_{ij\ldots}$ and is marked with $ij\cdots$. The action of g_1 is represented by following the solid-drawn triangles clockwise, and g_2 is represented by following the dashed lines.

We must now extract 9 basis elements $t_1, \ldots, t_9$ that are closed under $\mathbb{Z}_{\{3,3\}}$. In the first nonet, we have marked three sets of invariants with $\star$, $\bullet$ and $\circ$. They map to each other as $(\star, \bullet, \circ) \to (\bullet, \circ, \star)$ under g_1 and as $(\star, \bullet, \circ) \to (\star, \circ, \bullet)$ under g_2, so suitable linear combinations of the elements in each set will be closed under $\mathbb{Z}_{\{3,3\}}$. Similar constructions taken from the sextet and the other nonet turn out not to be linearly independent from these.

Unfortunately, it appears impossible to form a basis that contains the propagator momentum s_{123} as an element, but since there is only one propagator, this is not as much of a problem as it would be under a group that supports more operators. Also, it appears impossible to form real linear combinations without sacrificing either linear independence or closedness. Guided by the fact that g_1 has period 3, we instead insert the third root of

unity, $\omega = e^{2\pi i/3}$, and find the closed and complete basis $\mathcal{B}_{\{3,3\}}$ with elements[12]

$$t_1 = -\frac{s_{36} + s_{14} + s_{25}}{3}, \quad t_2 = -\frac{s_{24} + s_{35} + s_{16}}{3}, \quad t_3 = -\frac{s_{15} + s_{26} + s_{34}}{3},$$

$$t_6 = \frac{s_{36} + \omega s_{14} + \omega^2 s_{25}}{3}, \quad t_4 = \frac{s_{24} + \omega s_{35} + \omega^2 s_{16}}{3}, \quad t_5 = \frac{s_{15} + \omega s_{26} + \omega^2 s_{34}}{3},$$

$$t_9 = \frac{\omega^2 s_{36} + \omega s_{14} + s_{25}}{3}, \quad t_7 = \frac{\omega^2 s_{24} + \omega s_{35} + s_{16}}{3}, \quad t_8 = \frac{\omega^2 s_{15} + \omega s_{26} + s_{34}}{3},$$

$$\text{(1.4.4)}$$

In each row above, the first basis element comes from the $\star$ set, the second from the $\bullet$ set, and the third from the $\circ$ set. The propagator momentum is $s_{123} = \frac{3}{2}(t_1 + t_2 + t_3)$. The basis transforms as

$$
\begin{aligned}
g_1: & \quad \{t_1, t_2, t_3, \, t_4, t_5, t_6, \, t_7, t_8, t_9\} \to \{t_2, t_3, t_1, \, \omega t_5, \omega t_6, \omega t_4, \, \omega^2 t_8, \omega^2 t_9, \omega^2 t_7\}, \\
g_2: & \quad \{t_1, t_2, t_3, \, t_4, t_5, t_6, \, t_7, t_8, t_9\} \to \{t_1, t_3, t_2, \, t_4, t_6, t_5, \, t_7, t_9, t_8\}.
\end{aligned}
$$

$$\text{(1.4.5)}$$

Since stripped amplitudes are real, the complex basis must be compensated for by complex coefficients. Still, $\mathcal{B}_{\{3,3\}}$ is just as valid as a real basis, and is useable for simplification.

D.3 The basis for $R = \{2, 2, 2\}$

The group $\mathbb{Z}_{\{2,2,2\}}$ is also non-abelian, and can be tackled similarly to $\mathbb{Z}_{\{3,3\}}$. We choose the generators $g_1 = 34\,56\,12$, $g_2 = 21\,34\,56$ and $g_3 = 65\,43\,21$ with the hopes that they be well-behaved, since $\mathcal{B}_{\{6\}}$ is closed under two of them. This flavour structure permits six propagators that form a sextet under the group. The Cayley graph is

$$\text{(1.4.6)}$$
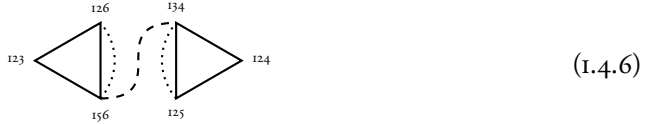
where g_1 and g_2 are represented as in eq. (1.4.3), and the dotted lines represent the action of g_3. The remaining invariants decompose into a triplet, a quadruplet, and a 12-plet:

$$\text{(1.4.7)}$$
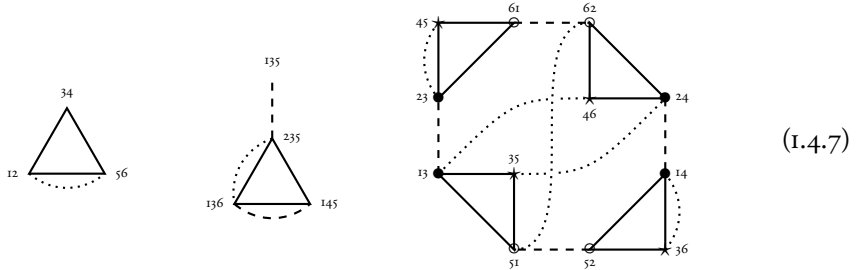
[12]The basis presented in [37] was not complete. This mistake was not discovered until after its publication, and is corrected here at the cost of losing the propagator.

Like in eq. (1.4.3), we have marked three closed sets of s_{ij}'s. From these, it is possible to construct three linearly independent elements that close the basis without need for the less structured triplet and quadruplet. Thus, $\mathcal{B}_{\{2,2,2\}}$ has elements

$$t_1 = s_{123}, \quad t_2 = s_{126}, \quad t_3 = s_{156}, \quad t_4 = s_{124}, \quad t_5 = s_{125}, \quad t_6 = s_{134},$$

$$t_7 = \frac{s_{61} - s_{62} + s_{52} - s_{51}}{2}, \quad t_8 = \frac{s_{23} - s_{24} + s_{14} - s_{13}}{2}, \quad t_9 = \frac{s_{45} - s_{46} + s_{36} - s_{35}}{2},$$

(1.4.8)

where the factors of $1/2$ remove some large powers of 2 that show up when writing amplitudes in this basis. Unlike in $\mathcal{B}_{\{3,3\}}$, there was no need to resort to complex numbers. The basis transforms as

$$\begin{aligned}
g_1 : &\quad \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\} \to \{t_1, t_2, t_6, t_4, t_5, t_3, -t_7, -t_8, t_9\}, \\
g_2 : &\quad \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\} \to \{t_2, t_3, t_1, t_5, t_6, t_4, t_8, t_9, t_7\}, \\
g_3 : &\quad \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\} \to \{t_1, t_3, t_2, t_4, t_6, t_5, t_7, t_9, t_8\}.
\end{aligned}$$

(1.4.9)

No element is a fixed point, which makes the basis harder to work in.

E Explicit amplitudes

E.1 The $\mathcal{O}(p^6)$ 6-point amplitude

This amplitude has been simplified using the closed bases of appendix D. The terms were reduced to coset representatives in FODGE followed by manual post-processing. Greater simplification might be possible for some terms. The amplitude consists of four stripped amplitudes with flavour split $\{6\}$, $\{2,4\}$, $\{3,3\}$, and $\{2,2,2\}$.

There are three diagrams with a single-trace flavour structure:

(1.5.1)

The corresponding stripped amplitude is

$$
\begin{aligned}
-iF^8 &\mathcal{M}_{6,\{6\}} \\
&= 2(L_3 + 2L_0)^2 \left\{ 2(s_{12} + s_{23})^2(s_{45} + s_{56}) - s_{123}(s_{12} + s_{23})(s_{45} + s_{56}) \right\} \\
&\quad + 16 L_3 L_0 s_{12} s_{23}(s_{45} + s_{56}) \\
&\quad - 8(L_0^2 + L_3 L_0) \frac{(s_{12} + s_{23})^2 (s_{45} + s_{56})^2}{s_{123}} - 2L_3^2 \frac{(s_{12}^2 + s_{23}^2)(s_{45}^2 + s_{56}^2)}{s_{123}} \\
&\quad + L_{6,3} \left\{ \left[s_{123} s_{34} - \frac{s_{123} s_{234}}{2} \right] (s_{12} + [\mathbb{Z}_6]) - s_{12} s_{23} s_{34} \right. \\
&\qquad\qquad \left. + \frac{2(s_{12} + s_{45})^3 + (s_{12}^3 + s_{45}^3)}{3} - \frac{\left[(s_{12} + s_{23})^3 + 2(s_{12}^3 + s_{23}^3) \right](s_{45} + s_{56})}{3 s_{123}} \right\}
\end{aligned}
$$

$$+ L_{6,4} \left\{ s_{123}s_{234}s_{345} + \frac{s_{123}^2}{2}(s_{234} + s_{345} - 2s_{34}) \right.$$

$$- 2s_{123}s_{234}(2s_{12} + s_{23} + 2s_{34}) - \frac{s_{12}s_{45}}{2}(s_{123} + s_{345})$$

$$+ s_{123}\left[s_{12} + s_{56} + 2s_{12}s_{34} + 2s_{34}s_{56} + 4s_{34}(s_{23} + s_{34} + s_{45})\right]$$

$$\left. + \frac{(s_{12} + s_{45})^3}{2} - \frac{(s_{12} + s_{23})^3(s_{45} + s_{56})}{s_{123}} \right\}$$

$$- L_{6,15} \left\{ s_{12}s_{45}s_{56} \right\}$$

$$+ L_{6,16} \left\{ s_{12}(s_{34} + s_{45})(s_{123} + s_{345}) - s_{12}s_{123}s_{345} - s_{12}s_{34}s_{56} - \frac{s_{12}s_{45}}{2}(s_{12} + [\mathbb{Z}_6]) \right\}$$

$$+ L_{6,17} \left\{ s_{12}s_{45}(s_{123} + s_{345} - s_{12} - s_{45}) \right\}$$

$$+ L_{6,18} \left\{ (s_{12} + s_{45} + s_{234}^2)(s_{123} + s_{345}) - s_{12}s_{234}(s_{123} + [\mathbb{Z}_6]) - 4s_{12}s_{123}s_{345} \right.$$

$$\left. + 2s_{12}s_{234}(s_{23} + s_{34} + s_{45} + s_{56} + s_{61}) - 2s_{12}s_{34}(s_{23} + s_{45} + s_{56} + s_{61}) \right\}$$

$$+ L_{6,19} \left\{ 2s_{123}s_{234}s_{345} + 3s_{234}^2(s_{123} + s_{345}) - 3s_{12}s_{234}(s_{123} + [\mathbb{Z}_6]) - 6s_{12}s_{123}s_{345} \right.$$

$$\left. + 6s_{12}s_{234}(s_{23} + s_{34} + s_{56} + s_{61}) - 6s_{12}s_{23}s_{34} - 2s_{12}s_{34}s_{56} \right\}$$

$$+ [\mathbb{Z}_6]. \tag{1.5.2}$$

The "$+ [\mathbb{Z}_6]$" acts on all terms in the amplitude.

There are also three diagrams with a $\{2, 4\}$-split flavour structure:

$$\tag{1.5.3}$$

Using the closed basis eq. (1.4.1), the stripped amplitude is

$$-iF^8 \mathcal{M}_{6,\{2,4\}}$$

$$= L_0 L_1 \left\{ t_1 \left[-t_1^2 + t_2^2 + 2t_3(3t_2 + t_3 - 2t_5) + t_4^2 + 4t_5^2 \right] - \frac{T_{234}^2(t_3 - 2t_5)^2}{t_1} \right\}$$

$$+ L_0 L_2 \left\{ t_1 \left[t_1\left(2t_5 - \frac{5t_1}{4}\right) + t_2\left(\frac{t_2}{4} + \frac{3t_2}{2} + 4t_5\right) \right. \right.$$

$$\left. + 3t_3\left(\frac{t_3}{2} - t_5\right) + \frac{t_4^2}{4} - 3t_5^2 + T_{789}^2 \right] - \frac{(t_3 - 2t_5)^2(T_{234}^2 + 4T_{789}^2)}{4t_1} \right\}$$

$$+ L_3 L_1 \left\{ -\frac{t_1^3}{2} + 2t_1^2(t_5 - t_6) \right.$$

$$+ t_1 \left[2t_2(t_3 + 2t_5 + t_6) + t_3(5t_3 + t_5 - 3t_6) - \frac{3t_4^2}{2} + t_5^2 + t_6^2 \right]$$

$$+ \frac{1}{t_1} \left[-t_2^3(t_2 + t_3 + 2t_4 + 2t_6) + t_2^2\left(\frac{t_3^2}{2} + t_3(t_5 - 3t_6) - t_4^2 - 4t_4 t_6\right) \right.$$

$$+ t_2(t_3^2 t_4 + t_3 t_4^2) + 2t_2 t_3 t_4(t_5 - t_6) - 2t_2 t_4^2 t_6 - \frac{t_3^4}{2}$$

$$\left. \left. + t_3^3(-t_4 + t_5 + t_6) - \frac{t_3^2 t_4^2}{2} + (t_5 + t_6)(t_3 t_4^2 + 2t_3^2 t_4) - T_{234}^2(t_5^2 + t_6^2) \right] \right\}$$

$$+ L_3 L_2 \left\{ -\frac{9t_1^3}{8} + t_1^2 \frac{t_5 + t_6}{2} \right.$$

$$+ t_1 \left[t_2 \left(\frac{11t_2}{2} + \frac{t_3}{2} + 3t_5 + \frac{5t_6}{2} \right) + \frac{t_3}{4} (5t_3 + t_5 + t_6) - \frac{3(t_5^2 + t_6^2)}{4} \right.$$

$$\left. - \frac{7t_4^2}{8} + t_7(t_9 - 3t_8) + 5t_8 t_9 + \frac{t_8^2 + t_9}{2} \right] + t_5(2t_7^2 + t_9^2) + 2t_6 t_7 t_8$$

$$+ \frac{1}{t_1} \left[\frac{T_{789}^2}{4} \left([2t_2 - t_3]^2 + [2t_5 + 2t_6 - t_3]^2 + 8t_6[t_2 - t_5] \right) \right.$$

$$- \frac{T_{234}^2}{4}(t_5^2 + t_6^2) + \frac{t_5 + t_6}{4}(2t_3^2 t_4 + t_3 t_4^2) - \frac{t_2^4}{4} - \frac{t_2^3}{4}(t_3 + 2t_4 + 2t_6)$$

$$+ \frac{t_2^2}{8}(t_3^2 + 2t_3 t_5 - 6t_3 t_6 - 2t_4^2 - 8t_4 t_6)$$

$$\left. \left. + \frac{t_2}{2} \left(\frac{t_3^2}{2}[t_4 + 2t_5] + \frac{t_3}{2} \left[t_4^2 + 2t_4 t_5 - 2t_4 t_6 \right] - t_4^2 t_6 \right) \right] \right\}$$

$$+ L_{6,1} \mathcal{H}(3) + L_{6,2} \mathcal{H}(-1)$$

$$+ L_{6,8} \left\{ \frac{t_1}{4} \left[t_2^2 - 2t_3^2 + t_4^2 - 2(t_5^2 + 2t_6^2) + 2t_6(t_1 + t_3) \right] \right\}$$

$$+ L_{6,9} \left\{ -t_1 \left[4t_5^2 + \frac{t_2^2 + t_4^2}{2} + t_2 t_3 - t_5(t_1 + 2t_2 + t_3) \right] \right\}$$

$$+ L_{6,11} \left\{ \frac{t_1^3}{16} - \frac{t_1}{16} \left[t_2^2 - t_3^2 + t_4^2 + 4(t_8^2 - t_7^2) + 2t_2(t_5 + t_6) + 4t_9(t_8 + t_7) \right] \right.$$

$$\left. - \frac{t_5 t_9^2}{8} - \frac{t_6 t_7^2}{4} \right\}$$

$$+ L_{6,12} \left\{ \frac{t_1^3}{16} - \frac{t_1}{16} \left[(t_2 + t_3)^2 + t_4^2 + 4(t_7 + t_8)^2 + 4t_9^2 - 4t_3 t_5 \right] + t_5 \frac{2t_7^2 - t_9^2}{4} \right\}$$

$$+ [\mathbb{Z}_{\{2,4\}}], \tag{1.5.4}$$

where $T_{ijk} = t_i + t_j + t_k$, and

$$\mathcal{H}(\eta) = \frac{3t_1^3}{128} - \frac{3t_1^2 t_5}{32}$$

$$+ t_1 \left[\frac{9t_2^2}{128} + t_2 \frac{3t_3 - 5t_5}{32} + \frac{t_3^2}{32} - \frac{5t_3 t_5}{64} + \frac{9t_4^2}{128} + \frac{\eta}{32}(t_7 + t_8 - t_9)^2 + \frac{\eta}{8}(t_7^2 + t_8^2 - t_9^2) \right]$$

$$- \eta t_5 \frac{2t_7^2 + t_9^2}{16} + \frac{3}{16t_1} \left[t_2^3 \frac{t_3 - 2t_5}{24} + \frac{t_2^2 t_3}{8}(t_3 + t_4 - 2t_5) - \frac{t_2^2 t_4 t_5}{4} \right.$$

$$+ \frac{t_2 t_3^2}{8}(t_3 + 2t_4 - 2t_5) + \frac{t_2 t_3}{8}(t_4^2 - 4t_4 t_5) + \frac{t_2 t_4^2 t_5}{4}$$

$$+ \frac{t_3^3}{24}(t_3 + 3t_4 + 2t_5) + \frac{t_3^2 t_4}{8}(t_4 - 2t_5) + \frac{t_4^4}{24} - \frac{t_4^2 t_5}{12}(3t_3 + t_4)$$

$$\left. + \frac{\eta T_{789}^2}{6} \left(t_3^2 + t_2 t_3 - t_2 t_5 + t_3 t_4 - 2t_3 t_5 - 2t_4 t_5 \right) \right] \tag{1.5.5}$$

is used for compactness.

There are two diagrams with a $\{3, 3\}$ flavour split:

$$(1.5.6)$$

Using the closed basis eq. (1.4.4), the stripped amplitude is

$$-iF^8 \mathcal{M}_{6,\{3,3\}}$$

$$= 64(L_0 + L_3)^2 \left\{ (s_{12} + s_{23})\left(s_{45}^2 + s_{45}s_{56} + s_{56}^2\right) + \left(s_{12}^2 + s_{12}s_{23} + s_{23}^2\right)(s_{45} + s_{56}) \right.$$

$$\left. - \frac{\left(s_{12}^2 + s_{12}s_{23} + s_{23}^2\right)\left(s_{45}^2 + s_{45}s_{56} + s_{56}^2\right)}{s_{123}} \right\}$$

$$+ L_{6,10}\left\{ \frac{t_1^3}{4} - (t_4^4 + t_7^3) + \frac{t_1 t_2 t_3}{2} + t_4 t_5 t_6 + t_7 t_8 t_9 + \frac{3t_1 t_2^2}{2} + 6\omega t_1 t_5(t_9 - t_8) \right\}$$

$$+ L_{6,13}\left\{ t_1^3 - (t_4^3 + t_7^3) - 3\omega t_1 t_4 t_7 \right\}$$

$$+ L_{6,14}\left\{ t_1 t_2 t_3 - (t_4 t_5 t_6 + t_7 t_8 t_9) + 6\omega t_1 t_5 t_9 \right\}$$

$$+ [\mathbb{Z}_{\{3,3\}}],$$

$$(1.5.7)$$

where $\omega = e^{2\pi i/3}$ is a third root of unity. The contribution from the singlet diagram turns out to be simpler to express in the standard basis $\mathcal{B}_{\{6\}}$ than in the closed basis.

Lastly, there are two diagrams for the $\{2, 2, 2\}$ flavour split:

$$(1.5.8)$$

Using the closed basis eq. (1.4.8), the stripped amplitude is

$$-iF^8 \mathcal{M}_{6,\{2,2,2\}}$$

$$= L_1^2 \left\{ -\frac{t_1^3}{2} + t_1(2t_2^2 + 6t_2 t_4 + t_4^2) \right.$$

$$\left. - \frac{1}{t_1}\left[\frac{t_4^4}{2} + (t_2 + t_5)^2(t_4^2 + 4t_3 t_4 t_3^2) + 2t_2 t_3 t_5 t_6 + 2t_2 t_3^2 t_5 + 8t_4^2 t_2 t_3 + 4t_4^3 t_2 \right] \right\}$$

$$+ L_1 L_2 \left\{ t_1 t_8^2 + \frac{3t_1 t_4^2}{2} + \frac{t_1 t_2}{2}(4t_9 + 5t_4 - 8t_3 - 5t_2) - \frac{5t_1^3}{4} \right.$$

$$- \frac{1}{t_1}\left[\frac{t_2^2}{2}\left(t_2^2 + 2t_3^3 + 3t_4^2 - t_5^2 + 4t_9^2 + 2t_2 t_4 + 4t_2 t_9 + 8t_4 t_9 + 8t_3 t_4 + 4t_3 t_6 \right) \right.$$

$$+ t_2 t_3(4t_4^2 + 4t_4 t_5 + t_5 t_6) + 2t_2 t_9(t_4^2 - t_5^2) + 2t_2 t_9^2(t_4 + t_5)$$

$$\left. \left. + 2t_2 t_4^3 + \frac{t_2 t_4^2 t_5}{2} - t_2 t_4 t_5^2 + t_4^2 t_8^2 + \frac{t_4^4}{4} \right] \right\}$$

$$- L_2^2 \left\{ \frac{21t_1^3}{32} + \frac{3t_1 t_8^2}{4} - \frac{11 t_1 t_4^2}{16} + \frac{t_1 t_2}{4}(t_4 + 6t_3 + t_2 - 10 t_9) \right.$$

$$+ \frac{1}{4t_1} \left[\frac{t_2^2}{2} \left(t_2^2 + 2[t_3 + t_4]^2 - t_5^2 + 4[t_8^2 + t_9^2] + 2t_2[t_4 + 2t_9] + 8t_4 t_9 + 8t_3 t_8 \right) \right.$$

$$+ t_2 t_3 \left(2t_4^2 + 2t_4 t_5 + t_5 t_6 - 4t_5 t_8 + 8t_8 t_9 \right)$$

$$+ t_2 t_4 \left(t_4^2 + 2t_4 t_9 - t_5^2 + 4t_9^2 \right) - 2t_2 t_5^2 t_9 + 2t_2 t_5 \left(t_9^2 - t_8^2 \right)$$

$$\left. \left. + t_8^2 \left(t_4^2 + 2t_9^2 + 8t_2 t_9 \right) + \frac{t_4^4}{8} \right] \right\}$$

$$+ \, [\mathbb{Z}_{\{2,2,2\}}] \tag{1.5.9}$$

This completes the amplitude.

E.2 The $\mathcal{O}(p^2)$ 10-point amplitude

Due to the absence of flavour splits, $\mathcal{O}(p^2)$ amplitudes are relatively easy to extend to many legs. The 10-point amplitude, which is also computed in [36], is given by the 16 diagrams[13]
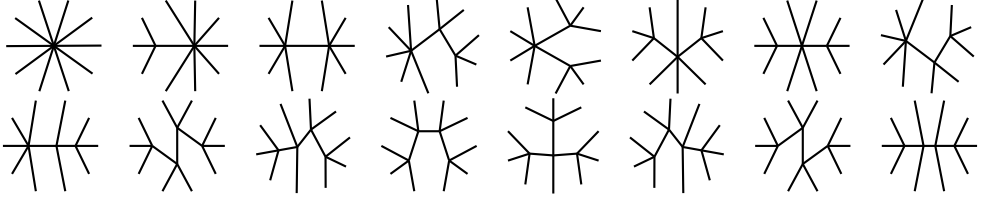
and has the stripped amplitude

$$-16 i F^8 \mathcal{M}_{2,\{10\}} = 5 s_{12} + 2 s_{1234}$$

$$- \frac{(s_{12} + s_{23} + s_{34} + s_{45} + s_{\overline{14}} + s_{\overline{25}})(s_{67} + s_{78} + s_{89} + s_{9A} + s_{\overline{69}} s_{\overline{7A}})}{2 s_{\overline{15}}}$$

$$- \frac{s_{12} + s_{23}}{s_{123}} \left\{ 2 \left(s_{45} + s_{56} + s_{67} + s_{78} + s_{89} + s_{9A} \right) + 2 \left(s_{\overline{A3}} + s_{\overline{14}} \right) \right.$$

$$- \frac{(s_{67} + s_{78})(s_{45} + s_{9A} + s_{\overline{A3}} + s_{\overline{14}} + s_{\overline{58}} + s_{\overline{69}})}{2 s_{678}}$$

$$- \frac{(s_{78} + s_{89})(s_{45} + s_{56} + s_{\overline{14}} + s_{\overline{47}} + s_{\overline{7A}} + s_{\overline{69}})}{s_{789}}$$

$$- \frac{(s_{89} + s_{9A})(s_{45} + s_{56} + s_{67} + s_{\overline{14}} + s_{\overline{47}} + s_{\overline{7A}})}{s_{89A}}$$

$$+ \frac{s_{9A} + s_{\overline{A3}}}{s_{\overline{48}}} \left[\frac{(s_{67} + s_{78})(s_{45} + s_{\overline{58}})}{2 s_{678}} - \left(s_{45} + s_{56} + s_{67} + s_{78} + s_{\overline{47}} + s_{\overline{58}} \right) \right]$$

---

[13]The circular shape is a result of the automatic diagram drawing in FODGE. The external legs are evenly distributed around a circle, and the location of each vertex is generated from the mean locations of all legs and vertices connected to it.

$$
+ \frac{s_{\overline{\text{A3}}} + s_{\overline{14}}}{s_{\overline{59}}} \left[\frac{(s_{67} + s_{78})(s_{\overline{58}} + s_{\overline{69}})}{2s_{678}} - (s_{56} + s_{67} + s_{78} + s_{89} + s_{\overline{58}} + s_{\overline{69}}) \right]
$$

$$
+ \frac{s_{\overline{14}} + s_{45}}{s_{\overline{15}}} \left[- (s_{67} + s_{78} + s_{89} + s_{9\text{A}} + s_{\overline{69}} + s_{\overline{7\text{A}}}) \right.
$$

$$
\left. + \frac{(s_{67} + s_{78})(s_{\overline{69}} + s_{9\text{A}})}{2s_{678}} + \frac{(s_{78} + s_{89})(s_{\overline{69}} + s_{\overline{7\text{A}}})}{s_{789}} + \frac{(s_{89} + s_{9\text{A}})(s_{67} + s_{\overline{7\text{A}}})}{s_{89\text{A}}} \right]
$$

$$
\left. + s_{\overline{47}} + s_{\overline{58}} + s_{\overline{69}} + s_{\overline{7\text{A}}} + \frac{(s_{\overline{7\text{A}}} + s_{\overline{\text{A3}}})(s_{45} + s_{56})(s_{78} + s_{89})}{s_{456}s_{789}} \right\} + [\mathbb{Z}_{10}]. \qquad (1.5.10)
$$

To avoid problems with multi-digit indices, we switch to hexadecimal and write A instead of 10. To abbreviate long index lists, we write $\overline{ij}$ for $i(i+1)\cdots(j-1)j$. Indices wrap around cyclically; $\overline{\text{A3}}$ means A123.

E.3 The $\mathcal{O}(p^2)$ 12-point amplitude

This is a novel amplitude, and takes the most time to compute of all amplitudes presented in this work. It consists of 73 diagrams:

135
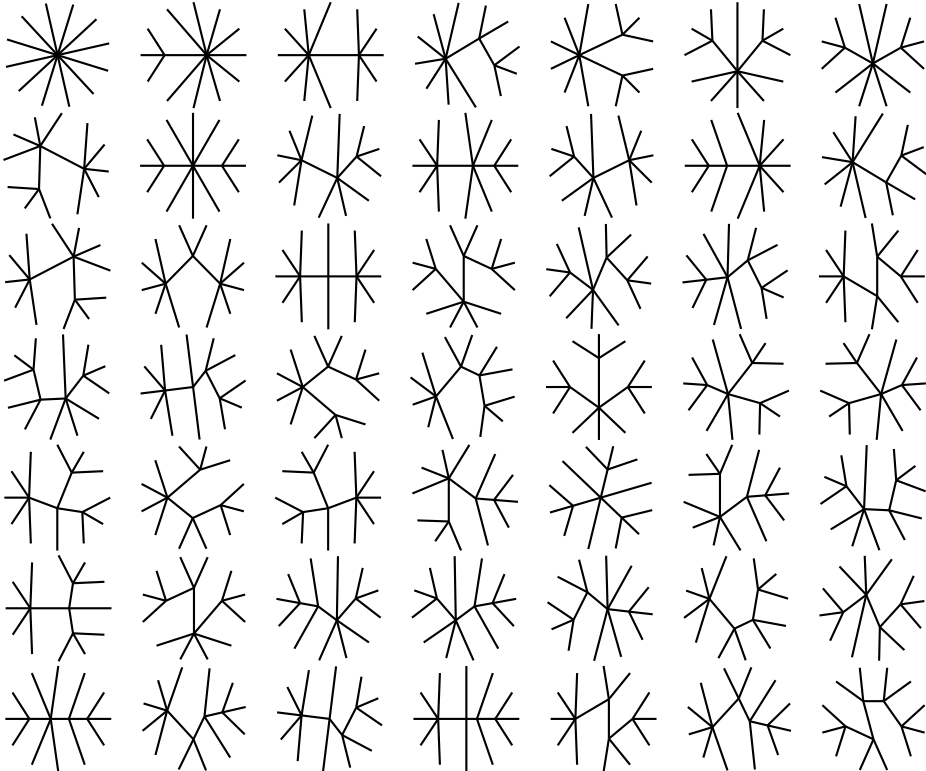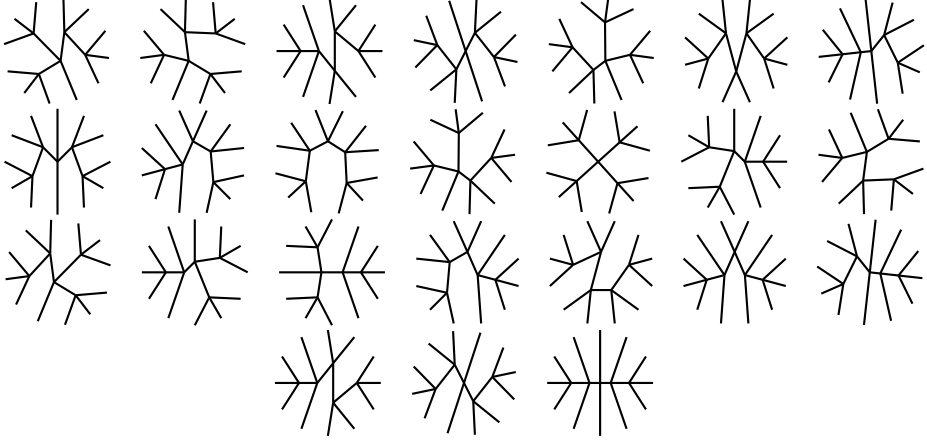
and has the stripped amplitude

$$
-32iF^{10}\mathcal{M}_{2,\{12\}} = 14s_{12} + 5s_{1234} + 2s_{123456}
$$

$$
+ \frac{s_{12} + \ldots + s_{45} + s_{\overline{14}} + s_{\overline{25}}}{s_{\overline{15}}} \left\{ -\frac{(s_{67} + \ldots + s_{9\mathrm{A}} + s_{\overline{69}} + s_{\overline{7\mathrm{A}}})(s_{\overline{69}} + s_{\overline{7\mathrm{A}}})}{s_{\overline{6\mathrm{A}}}} \right.
$$

$$
\times [2(s_{67} + \ldots + s_{\mathrm{BC}}) + (s_{\overline{69}} + \ldots + s_{\overline{9\mathrm{C}}}) + 2(s_{\overline{16}} + s_{\overline{6\mathrm{B}}})]
$$

$$
\left. + \frac{(s_{78} + \ldots + s_{\mathrm{AB}} + s_{\overline{7\mathrm{A}}} + s_{\overline{8\mathrm{B}}})(s_{\overline{16}} + s_{\overline{6\mathrm{B}}})}{2s_{\overline{7\mathrm{B}}}} \right\}
$$

$$
+ \frac{s_{12} + s_{23}}{s_{123}} \left\{ -[5(s_{45} + \ldots + s_{\mathrm{BC}} + s_{\overline{\mathrm{C}3}} + s_{\overline{14}}) + 2(s_{\overline{47}} + \ldots + s_{\overline{9\mathrm{C}}} + s_{\overline{49}} + \ldots + s_{\overline{7\mathrm{C}}})] \right.
$$

$$
+ \frac{s_{45} + s_{56}}{s_{456}} \left[2(s_{78} + \ldots + s_{\mathrm{BC}} + s_{\overline{\mathrm{C}3}} + s_{\overline{47}} + s_{\overline{16}}) + s_{\overline{7\mathrm{A}}} + s_{\overline{8\mathrm{B}}} + s_{\overline{9\mathrm{C}}} + s_{\overline{49}} \right.
$$

$$
+ \frac{s_{78} + s_{89}}{s_{789}} \left(-[s_{\mathrm{AB}} + s_{\mathrm{BC}} + s_{\overline{7\mathrm{A}}} + s_{\overline{\mathrm{C}3}} + s_{\overline{16}} + s_{\overline{49}}] + \frac{s_{\mathrm{AB}} + s_{\mathrm{BC}}}{2s_{\mathrm{ABC}}} \right.
$$

$$
\left. + \frac{(s_{\overline{7\mathrm{A}}} + s_{\mathrm{AB}})(s_{\overline{16}} + s_{\overline{\mathrm{C}3}})}{s_{\overline{7\mathrm{B}}}} + \frac{(s_{\overline{7\mathrm{A}}} + s_{\overline{49}})(s_{\mathrm{BC}} + s_{\mathrm{C}3})}{s_{\overline{\mathrm{B}3}}} \right)
$$

$$
+ \frac{s_{89} + s_{9\mathrm{A}}}{s_{89\mathrm{A}}} \left(-[s_{\mathrm{BC}} + s_{\overline{47}} + s_{\overline{7\mathrm{A}}} + s_{\overline{8\mathrm{B}}} + s_{\overline{\mathrm{C}3}} + s_{\overline{16}}] + \frac{(s_{\overline{7\mathrm{A}}} + s_{\overline{8\mathrm{B}}})(s_{\overline{\mathrm{C}3}} + s_{\overline{16}})}{s_{\overline{7\mathrm{B}}}} \right.
$$

$$
\left. + \frac{(s_{\overline{8\mathrm{B}}} + s_{\overline{\mathrm{BC}}})(s_{\overline{47}} + s_{\overline{16}})}{s_{\overline{8\mathrm{C}}}} + \frac{(s_{\overline{\mathrm{BC}}} + s_{\overline{\mathrm{C}3}})(s_{\overline{47}} + s_{\overline{7\mathrm{A}}})}{s_{\overline{\mathrm{B}3}}} \right)
$$

$$
+ \frac{s_{9\mathrm{A}} + s_{\mathrm{AB}}}{s_{9\mathrm{AB}}} \left(-[s_{\overline{47}} + s_{78} + s_{\overline{9\mathrm{C}}} + s_{\overline{\mathrm{C}3}} + s_{\overline{16}}] + \frac{(s_{\overline{47}} + s_{78})(s_{\overline{9\mathrm{C}}} + s_{\overline{\mathrm{C}3}})}{s_{\overline{48}}} \right.
$$

$$
\left. + \frac{(s_{78} + s_{\overline{8\mathrm{B}}})(s_{\overline{\mathrm{C}3}} + s_{\overline{16}})}{s_{\overline{7\mathrm{B}}}} + \frac{(s_{\overline{8\mathrm{B}}} + s_{\overline{9\mathrm{C}}})(s_{\overline{47}} + s_{\overline{16}})}{s_{\overline{8\mathrm{C}}}} \right)
$$

$$
+ \frac{s_{\overline{47}} + s_{78}}{s_{\overline{48}}} \left(\frac{(s_{\overline{49}} + s_{9\mathrm{A}})(s_{\mathrm{BC}} + s_{\overline{\mathrm{C}3}})}{s_{\overline{\mathrm{B}3}}} - [s_{9\mathrm{A}} + s_{\mathrm{AB}} + s_{\mathrm{BC}} + s_{\overline{9\mathrm{C}}} + s_{\overline{\mathrm{C}3}} + s_{\overline{49}}] \right)
$$

$$- \frac{(s_{\overline{C3}} + s_{\overline{16}})(s_{78} + \ldots + s_{AB} + s_{\overline{7A}} + s_{\overline{8B}})}{s_{\overline{7B}}}$$

$$- \frac{(s_{\overline{47}} + s_{\overline{16}})(s_{89} + \ldots + s_{BC} + s_{8B} + s_{\overline{9C}})}{s_{8C}}$$

$$- \left. \frac{(s_{BC} + s_{\overline{C3}})(s_{78} + s_{89} + s_{9A} + s_{\overline{47}} + s_{\overline{7A}} + s_{\overline{49}})}{s_{\overline{B3}}} \right]$$

$$+ \frac{s_{56} + s_{67}}{s_{567}} \left[2\left(s_{89} + \ldots + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{47}} + s_{\overline{58}} \right) + s_{\overline{8B}} + s_{\overline{9C}} + s_{\overline{49}} + s_{\overline{5A}} \right.$$

$$+ \frac{s_{9A} + s_{AB}}{s_{9AB}} \left(\frac{(s_{\overline{47}} + s_{\overline{58}})(s_{\overline{9C}} + s_{\overline{C3}})}{s_{\overline{48}}} - 1 \right)$$

$$+ \frac{s_{\overline{47}} + s_{\overline{58}}}{s_{\overline{48}}} \left(\frac{(s_{\overline{49}} + s_{9A})(s_{BC} + s_{\overline{C3}})}{s_{\overline{B3}}} - (s_{9A} + s_{AB} + s_{BC} + s_{\overline{9C}} + s_{\overline{C3}} + s_{\overline{49}}) \right)$$

$$+ \frac{s_{\overline{58}} + s_{89}}{s_{59}} \left(\frac{(s_{BC} + s_{\overline{C3}})(s_{\overline{49}} + s_{\overline{5A}})}{s_{\overline{B3}}} + \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{\overline{5A}} + s_{AB})}{s_{\overline{C4}}} \right.$$

$$\left. - (s_{AB} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{49}} + s_{\overline{5A}}) \right)$$

$$- \frac{(s_{\overline{14}} + s_{\overline{47}})(s_{89} + \ldots + s_{BC} + s_{\overline{8B}} + s_{\overline{9C}})}{s_{\overline{8C}}}$$

$$- \frac{(s_{BC} + s_{\overline{C3}})(s_{89} + s_{9A} + s_{\overline{47}} + s_{\overline{58}} + s_{\overline{49}} + s_{\overline{5A}})}{s_{\overline{B3}}}$$

$$- \left. \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{89} + s_{9A} + s_{AB} + s_{\overline{58}} + s_{\overline{8B}} + s_{\overline{5A}})}{s_{\overline{C4}}} \right]$$

$$+ \frac{s_{67} + s_{78}}{s_{678}} \left[2\left(s_{45} + s_{9A} + s_{AB} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{58}} + s_{\overline{69}} \right) + s_{\overline{9C}} + s_{\overline{49}} + s_{\overline{5A}} + s_{\overline{6B}} \right.$$

$$+ \frac{s_{45} + s_{\overline{14}}}{s_{\overline{15}}} \left(\frac{(s_{\overline{69}} + s_{9A})(s_{\overline{6B}} + s_{BC})}{s_{\overline{6A}}} - (s_{9A} + s_{AB} + s_{BC} + s_{\overline{69}} + s_{\overline{9C}} + s_{\overline{6B}}) \right)$$

$$+ \frac{s_{45} + s_{\overline{58}}}{s_{\overline{48}}} \left(\frac{(s_{\overline{49}} + s_{9A})(s_{BC} + s_{\overline{C3}})}{s_{\overline{B3}}} - (s_{9A} + s_{AB} + s_{BC} + s_{\overline{9C}} + s_{\overline{C3}} + s_{\overline{49}}) \right)$$

$$+ \frac{s_{\overline{58}} + s_{\overline{69}}}{s_{59}} \left(\frac{(s_{BC} + s_{\overline{C3}})(s_{\overline{49}} + s_{\overline{5A}})}{s_{\overline{B3}}} + \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{\overline{5A}} + s_{AB})}{s_{\overline{C4}}} \right.$$

$$\left. - (s_{AB} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{49}} + s_{\overline{5A}}) \right)$$

$$+ \frac{s_{\overline{69}} + s_{9A}}{s_{\overline{6A}}} \left(\frac{(s_{45} + s_{\overline{5A}})(s_{BC} + s_{\overline{C3}})}{s_{\overline{B3}}} + \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{\overline{5A}} + s_{\overline{6B}})}{s_{\overline{C4}}} \right.$$

$$\left. - (s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{5A}} + s_{\overline{6B}}) \right)$$

$$- \frac{(s_{BC} + s_{\overline{C3}})(s_{45} + s_{BC} + s_{\overline{58}} + s_{\overline{69}} + s_{\overline{49}} + s_{\overline{5A}})}{s_{\overline{B3}}}$$

$$- \left. \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{9A} + s_{AB} + s_{\overline{58}} + s_{\overline{69}} + s_{\overline{5A}} + s_{\overline{6B}})}{s_{\overline{C4}}} \right]$$

$$+ \frac{s_{78} + s_{89}}{s_{789}} \left[\frac{1}{2}(s_{45} + s_{56} + s_{\overline{14}} + s_{\overline{69}}) + 2(s_{\overline{16}} + s_{\overline{49}} + s_{\overline{5A}} + s_{\overline{6B}}) \right.$$

$$+ \frac{s_{\overline{14}} + s_{45}}{s_{\overline{15}}} \left(\frac{(s_{\overline{6B}} + s_{BC})(s_{\overline{69}} + s_{\overline{7A}})}{s_{\overline{6A}}} + \frac{(s_{\overline{7A}} + s_{AB})(s_{\overline{16}} + s_{\overline{6B}})}{2s_{\overline{7B}}} \right.$$

$$\left. - (s_{AB} + s_{BC} + s_{\overline{69}} + s_{\overline{7A}} + s_{\overline{16}} + s_{\overline{6B}}) \right)$$

$$+ \frac{s_{56} + s_{\overline{69}}}{s_{\overline{59}}} \left(\frac{(s_{BC} + s_{\overline{C3}})(s_{\overline{49}} + s_{\overline{5A}})}{2s_{\overline{B3}}} + \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{\overline{5A}} + s_{AB})}{s_{\overline{C4}}} \right.$$

$$\left. - (s_{AB} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + + s_{\overline{49}} + s_{\overline{5A}}) \right)$$

$$\left. + \frac{s_{\overline{69}} + s_{\overline{7A}}}{s_{\overline{6A}}} \left(\frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{\overline{5A}} + s_{\overline{6B}})}{2s_{\overline{C4}}} - (s_{45} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{5A}} + s_{\overline{6B}}) \right) \right]$$

$$+ \frac{s_{\overline{14}} + s_{45}}{s_{\overline{15}}} \left[2 \left(s_{67} + \ldots + s_{BC} + s_{\overline{16}} + s_{\overline{6B}} \right) + s_{\overline{69}} + \ldots + s_{\overline{9C}} \right.$$

$$- \frac{(s_{\overline{6B}} + s_{BC})(s_{67} + \ldots + s_{9A} + s_{\overline{69}} + s_{\overline{7A}})}{s_{\overline{6A}}}$$

$$- \frac{(s_{\overline{16}} + s_{\overline{6B}})(s_{78} + \ldots + s_{AB} + s_{\overline{7A}} + s_{\overline{8B}})}{s_{\overline{7B}}}$$

$$\left. - \frac{(s_{\overline{16}} + s_{67})(s_{89} + \ldots + s_{BC} + s_{\overline{8B}} + s_{\overline{9C}})}{s_{\overline{8C}}} \right]$$

$$+ \frac{s_{45} + \ldots + s_{78} + s_{\overline{47}} + s_{\overline{58}}}{s_{\overline{48}}} \left[\left(s_{9A} + s_{AB} + s_{BC} + s_{\overline{9C}} + s_{\overline{C3}} + s_{\overline{49}} \right) \right.$$

$$\left. - \frac{(s_{\overline{49}} + s_{9A})(s_{BC} + s_{\overline{C3}})}{s_{\overline{B3}}} \right]$$

$$+ \frac{s_{56} + \ldots + s_{89} + s_{\overline{58}} + s_{\overline{69}}}{s_{\overline{59}}} \left[\left(s_{AB} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{49}} + s_{\overline{5A}} \right) \right.$$

$$\left. - \frac{(s_{\overline{49}} + s_{\overline{5A}})(s_{BC} + s_{\overline{C3}}}{s_{\overline{B3}}} - \frac{(s_{\overline{5A}} + s_{\overline{AB}})(s_{\overline{C3}} + s_{\overline{14}})}{s_{\overline{C4}}} \right]$$

$$+ \frac{s_{67} + \ldots + s_{9A} + s_{\overline{69}} + s_{\overline{7A}}}{s_{\overline{6A}}} \left[\left(s_{45} + s_{BC} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{5A}} + s_{\overline{6B}} \right) \right.$$

$$\left. - \frac{(s_{\overline{45}} + s_{\overline{5A}})(s_{BC} + s_{\overline{C3}})}{s_{\overline{B3}}} - \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{\overline{5A}} + s_{\overline{6B}})}{s_{\overline{C4}}} \right]$$

$$+ \frac{s_{78} + \ldots + s_{AB} + s_{\overline{7A}} + s_{\overline{8B}}}{s_{\overline{7B}}} \left[\left(s_{45} + s_{56} + s_{\overline{C3}} + s_{\overline{14}} + s_{\overline{16}} + s_{\overline{6B}} \right) - \frac{(s_{\overline{C3}} + s_{\overline{14}})(s_{56} + s_{\overline{6B}})}{s_{\overline{C4}}} \right]$$

$$+ \frac{(s_{89} + \ldots + s_{BC} + s_{\overline{8B}} + s_{\overline{9C}})(s_{56} + s_{67} + s_{\overline{14}} + s_{\overline{47}} + s_{\overline{16}})}{s_{\overline{8C}}}$$

$$+ \frac{(s_{BC} + s_{\overline{C3}}) \left[2(s_{45} + \ldots + s_{9A} + s_{\overline{49}} + s_{\overline{9A}}) + s_{\overline{47}} + \ldots + s_{\overline{7A}} \right]}{s_{\overline{B3}}}$$

$$+ \left. \frac{(s_{\overline{C3}} + s_{\overline{14}}) \left[2(s_{56} + \ldots + s_{AB} + s_{\overline{5A}} + s_{\overline{6B}}) + s_{\overline{58}} + \ldots + s_{\overline{8B}} \right]}{s_{\overline{C4}}} \right\} + [\mathbb{Z}_{12}]. \qquad \text{(I.5.II)}$$

We use the same abbreviations as above, with $A, B, C = 10, 11, 12$. Furthermore, we contract sums like $s_{12} + s_{23} + s_{34} + s_{45}$ to $s_{12} + \ldots + s_{45}$.

References

[1] Gell-Mann, M. & Lévy, M. "The axial vector current in beta decay". *Nuovo Cim.* **16**, 705 (1960).

[2] Weinberg, S. "Phenomenological Lagrangians". *Physica A* **96**, 327–340 (1979).

[3] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory to One Loop". *Annals Phys.* **158**, 142 (1984).

[4] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark". *Nucl. Phys. B* **250**, 465–516 (1985).

[5] Pich, A. *Effective Field Theory with Nambu-Goldstone Modes* in *Les Houches summer school: EFT in Particle Physics and Cosmology* (2018). arXiv: 1804.05664 [hep-ph].

[6] http://home.thep.lu.se/~bijnens/chpt/.

[7] Bijnens, J. & Lu, J. "Meson-meson Scattering in QCD-like Theories". *JHEP* **03**, 028. arXiv: 1102.0172 [hep-ph] (2011).

[8] Bijnens, J. & Hermansson Truedsson, N. "The Pion Mass and Decay Constant at Three Loops in Two-Flavour Chiral Perturbation Theory". *JHEP* **11**, 181. arXiv: 1710.01901 [hep-ph] (2017).

[9] Elvang, H. & Huang, Y.-T. "Scattering Amplitudes". arXiv: 1308.1697 [hep-th] (2013).

[10] Kampf, K., Novotny, J. & Trnka, J. "Recursion relations for tree-level amplitudes in the SU(N) nonlinear sigma model". *Phys. Rev.* **D87**, 081701. arXiv: 1212.5224 [hep-th] (2013).

[11] Cheung, C., Kampf, K., Novotný, J., Shen, C.-H. & Trnka, J. "On-Shell Recursion Relations for Effective Field Theories". *Phys. Rev. Lett.* **116**, 041601. arXiv: 1509.03309 [hep-th] (2016).

[12] Adler, S. L. "Consistency conditions on the strong interactions implied by a partially conserved axial-vector current. II". *Phys. Rev.* **139**, B1638–B1643 (1965).

[13] Cheung, C., Kampf, K., Novotný, J., Shen, C.-H. & Trnka, J. "A Periodic Table of Effective Field Theories". *JHEP* **02**, 020. arXiv: 1611.03137 [hep-th] (2017).

[14] Elvang, H., Hadjiantonis, M., Jones, C. R. T. & Paranjape, S. "Soft Bootstrap and Supersymmetry". *JHEP* **01**, 195. arXiv: 1806.06079 [hep-th] (2019).

[15] Low, I. & Yin, Z. "Soft Bootstrap and Effective Field Theories". *JHEP* **11**, 078. arXiv: 1904.12859 [hep-th] (2019).

[16] Osborn, H. "Implications of Adler zeros for multipion processes". *Lett. Nuovo Cim.* **2**, 717–723 (1969).

[17] Susskind, L. & Frye, G. "Algebraic aspects of pionic duality diagrams". *Phys. Rev. D* **1**, 1682–1686 (1970).

[18] Ellis, J. R. & Renner, B. "On the relationship between chiral and dual models". *Nucl. Phys. B* **21**, 205–216 (1970).

[19] Cachazo, F., He, S. & Yuan, E. Y. "Scattering Equations and Matrices: From Einstein To Yang-Mills, DBI and NLSM". *JHEP* **07**, 149. arXiv: 1412.3479 [hep-th] (2015).

[20] Gomez, H. & Helset, A. "Scattering equations and a new factorization for amplitudes. Part II. Effective field theories". *JHEP* **05**, 129. arXiv: 1902.02633 [hep-th] (2019).

[21] Chen, G. & Du, Y.-J. "Amplitude Relations in Non-linear Sigma Model". *JHEP* **01**, 061. arXiv: 1311.1133 [hep-th] (2014).

[22] Chen, G., Du, Y.-J., Li, S. & Liu, H. "Note on off-shell relations in nonlinear sigma model". *JHEP* **03**, 156. arXiv: 1412.3722 [hep-th] (2015).

[23] Du, Y.-J. & Luo, H. "On single and double soft behaviors in NLSM". *JHEP* **08**, 058. arXiv: 1505.04411 [hep-th] (2015).

[24] Low, I. "Double Soft Theorems and Shift Symmetry in Nonlinear Sigma Models". *Phys. Rev.* **D93**, 045032. arXiv: 1512.01232 [hep-th] (2016).

[25] Du, Y.-J. & Fu, C.-H. "Explicit BCJ numerators of nonlinear simga model". *JHEP* **09**, 174. arXiv: 1606.05846 [hep-th] (2016).

[26] Carrasco, J. J. M., Mafra, C. R. & Schlotterer, O. "Abelian Z-theory: NLSM amplitudes and α'-corrections from the open string". *JHEP* **06**, 093. arXiv: 1608.02569 [hep-th] (2017).

[27] Du, Y.-J. & Luo, H. "Leading order multi-soft behaviors of tree amplitudes in NLSM". *JHEP* **03**, 062. arXiv: 1611.07479 [hep-th] (2017).

[28] Cheung, C., Remmen, G. N., Shen, C.-H. & Wen, C. "Pions as Gluons in Higher Dimensions". *JHEP* **04**, 129. arXiv: 1709.04932 [hep-th] (2018).

[29] Low, I. & Yin, Z. "Ward Identity and Scattering Amplitudes for Nonlinear Sigma Models". *Phys. Rev. Lett.* **120**, 061601. arXiv: 1709.08639 [hep-th] (2018).

[30] Low, I. & Yin, Z. "The Infrared Structure of Nambu-Goldstone Bosons". *JHEP* **10**, 078. arXiv: 1804.08629 [hep-th] (2018).

[31] Rodina, L. "Scattering Amplitudes from Soft Theorems and Infrared Behavior". *Phys. Rev. Lett.* **122**, 071601. arXiv: 1807.09738 [hep-th] (2019).

[32] Mizera, S. & Skrzypek, B. "Perturbiner Methods for Effective Field Theories and the Double Copy". *JHEP* **10**, 018. arXiv: 1809.02096 [hep-th] (2018).

[33] Bjerrum-Bohr, N. E. J., Gomez, H. & Helset, A. "New factorization relations for nonlinear sigma model amplitudes". *Phys. Rev.* **D99**, 045009. arXiv: 1811.06024 [hep-th] (2019).

[34] Cornwell, D. T. "The six pion amplitude to fourth order in momenta". *Nucl. Phys.* **B34**, 125–135 (1971).

[35] Carrillo-Gonzalez, M., Penco, R. & Trodden, M. "Shift symmetries, soft limits, and the double copy beyond leading order". arXiv: 1908.07531 [hep-th] (2019).

[36] Kampf, K., Novotný, J. & Trnka, J. "Tree-level Amplitudes in the Nonlinear Sigma Model". *JHEP* **05,** 032. arXiv: 1304.3048 [hep-th] (2013).

[37] Sjö, M. *Flavour-ordering in the nonlinear sigma model with more derivatives and legs* M.Sc. thesis (Lund University, 2019). http://lup.lub.lu.se/student-papers/record/8984750.

[38] Ecker, G., Gasser, J., Pich, A. & de Rafael, E. "The Role of Resonances in Chiral Perturbation Theory". *Nucl. Phys. B* **321,** 311–342 (1989).

[39] Bijnens, J., Colangelo, G. & Ecker, G. "The Mesonic chiral Lagrangian of order p^6". *JHEP* **02,** 020. arXiv: hep-ph/9902437 (1999).

[40] Bijnens, J., Hermansson-Truedsson, N. & Wang, S. "The order p^8 mesonic chiral Lagrangian". *JHEP* **01,** 102. arXiv: 1810.06834 [hep-ph] (2019).

[41] Reuschle, C. & Weinzierl, S. "Decomposition of one-loop QCD amplitudes into primitive amplitudes based on shuffle relations". *Phys. Rev. D* **88,** 105020. arXiv: 1310.0413 [hep-ph] (2013).

[42] Schuster, T. "Color ordering in QCD". *Phys. Rev. D* **89,** 105022. arXiv: 1311.6296 [hep-ph] (2014).

[43] Mangano, M. L. & Parke, S. J. "Multiparton amplitudes in gauge theories". *Phys. Rept.* **200,** 301–367. arXiv: hep-th/0509223 (1991).

[44] Sjodahl, M. & Thorén, J. "Decomposing color structure into multiplet bases". *JHEP* **09,** 055. arXiv: 1507.03814 [hep-ph] (2015).

[45] Weinberg, S. *The quantum theory of fields. Vol. 2: Modern applications* ISBN: 978-1-139-63247-8, 978-0-521-67054-8, 978-0-521-55002-4 (Cambridge University Press, 2013).

[46] Cheung, C., Kampf, K., Novotny, J. & Trnka, J. "Effective Field Theories from Soft Limits of Scattering Amplitudes". *Phys. Rev. Lett.* **114,** 221602. arXiv: 1412.4095 [hep-th] (2015).

[47] Arkani-Hamed, N., Cachazo, F. & Kaplan, J. "What is the Simplest Quantum Field Theory?" *JHEP* **09,** 016. arXiv: 0808.1446 [hep-th] (2010).

[48] Vermaseren, J. A. M. "New features of FORM". arXiv: math-ph/0010025 (2000).

[49] Kuipers, J., Ueda, T., Vermaseren, J. A. M. & Vollinga, J. "FORM version 4.0". *Comput. Phys. Commun.* **184,** 1453–1467. arXiv: 1203.6543 [cs.SC] (2013).

[50] Bijnens, J. & Carloni, L. "The Massive O(N) Non-linear Sigma Model at High Orders". *Nucl. Phys. B* **843,** 55–83. arXiv: 1008.3499 [hep-ph] (2011).

[51] Bijnens, J., Kampf, K. & Lanz, S. "Leading logarithms in N-flavour mesonic Chiral Perturbation Theory". *Nucl. Phys. B* **873,** 137–164. arXiv: 1303.3125 [hep-ph] (2013).

[52] Bijnens, J., Colangelo, G. & Ecker, G. "Renormalization of chiral perturbation theory to order p^6". *Annals Phys.* **280,** 100–139. arXiv: hep-ph/9907333 (2000).

Paper II

II

NNLO positivity bounds on chiral perturbation theory for a general number of flavours

Benjamin Alvarez,[1] Johan Bijnens [2] and Mattias Sjö[2]

[1]Aix Marseille Univ, Univ Toulon,
   CNRS, CPT, Marseille, France
[2]Department of Astronomy and Theoretical Physics, Lund University,
   Box 43, SE 22100 Lund, Sweden

Abstract: We present positivity bounds, derived from the principles of analyticity, unitarity and crossing symmetry, that constrain the low-energy constants of chiral perturbation theory. Bounds are produced for 2, 3 or more flavours in meson-meson scattering with equal meson masses, up to and including next-to-next-to-leading order (NNLO), using the second and higher derivatives of the amplitude. We enhance the bounds by using the most general isospin combinations posible (or higher-flavour counterparts thereof) and by analytically integrating the low-energy range of the discontinuities. In addition, we present a powerful and general mathematical framework for efficiently managing large numbers of positivity bounds.

1   Introduction

Chiral perturbation theory (ChPT) is the most widespread theory for low-energy quantum chromodynamics (QCD). It is an effective field theory (EFT) which reformulates the non-perturbative behaviour of low-energy QCD as a perturbative theory of new degrees of freedom, physically interpreted as bound states of quarks. When constructed using n light quark flavours, the degrees of freedom are the $n^2 - 1$ light pseudoscalar mesons: the pions for $n = 2$, with the kaons and eta added for $n = 3$. ChPT was developed by Gasser & Leutwyler [1, 2] based on earlier work by Weinberg [3]; see [4, 5] for modern introductions with further references.

At leading order in the low-energy expansion, the only parameters of ChPT are the meson mass and decay constant, but higher orders introduce a rapidly increasing number of Wilson coefficients or low-energy constants (LECs) which, while in principle derivable from the underlying QCD dynamics, must in practice be seen as unknowns. At next-to-leading order (NLO), the LECs can be measured reasonably well with experimental or lattice methods, although the precision is typically only one or two significant digits. At next-to-next-to-leading order (NNLO), only tentative results are presently available. For a review of LEC measurements, see [6].

All quantum field theories must obey the axioms of unitarity, analyticity and crossing symmetry, and normally do so by construction. However, it turns out that these axioms are not automatically satisfied by EFTs such as ChPT when perturbativity is assumed at a fixed order in the expansion. Therefore, imposing the axioms actually adds new information, typically by placing bounds on the scattering amplitudes. Pioneering work was done by Martin [7] before the development of ChPT as such. Bounds on NLO two-flavour ChPT amplitudes, which in turn translate to bounds on the LECs, were first obtained in [8–10] and extended in [11, 12]. Further improvements were made in [13] and extended to three-flavour ChPT in [14]. There is ongoing research in extending these methods, both specific to ChPT and with broader scope; recent examples include [15–18].

The method of [13, 14], which serves as the basis of our method, is to apply dispersion relations (a consequence of analyticity) to a meson-meson scattering amplitude decomposed into isospin components (for higher flavours, the Clebsch-Gordan decomposition is used). Then, crossing symmetry and the optical theorem (a consequence of unitarity) are applied to give a positivity condition on the decomposed amplitude. With the amplitude calculated in terms of the LECs to some order, this results in bounds on linear combinations of LECs. More recently, stronger bounds have been obtained in [19, 20] by improving this method; put extremely simply, this was done with more sophisticated use of dispersion relations and crossing symmetry, respectively. Put similarly simply, our work instead improves the handling of the isospin decompositions and the LEC bounds themselves, although some improvements similar to [19] are also made. Perhaps more importantly, we perform the first extension to NNLO ChPT with any number of flavours (two flavours was treated in [19]), albeit with the simplification that all mesons have the same mass. The LECs are independent of the chosen masses, although the bounds do depend on the mass. At NLO they depend only on the ratio of the meson mass and the subtraction scale μ, at

NNLO also on the ratio of the meson mass and decay constant.

Preliminary results of this work are presented in the Lund University master thesis [21]. Our work is structured as follows: section 2 introduces ChPT and its LECs; section 3 (backed by appendix A) presents the $2 \to 2$ meson scattering amplitude used to obtain the bounds; section 4 (backed by appendix B) introduces the mathematical framework used to manage them; section 5 (backed by appendix C) presents the method of [13, 14] and the improvements made to it; and section 6 displays the most interesting bounds we obtain, with final remarks given in section 7.

2 Chiral perturbation theory

n-flavour ChPT is based around a non-linear sigma model (NLSM), whose degrees of freedom are the $n^2 - 1$ Nambu-Goldstone bosons that arise when the chiral symmetry $G = \mathrm{SU}(n)_L \times \mathrm{SU}(n)_R$ of n-flavour massless QCD is spontaneously broken into its diagonal subgroup $H = \mathrm{SU}(n)_V$. The Goldstone bosons live in the coset space G/H, which is isomorphic to $\mathrm{SU}(n)$.

The presence of quark masses, electroweak interactions, etc. can be accounted for by including four external $n \times n$ flavour-space matrix fields — s (scalar), p (pseudoscalar), v_μ (vector) and a_μ (axial vector)[1] — into the massless QCD Lagrangian. These additions were introduced in [1, 2], and endow the Nambu-Goldstone bosons with masses and interactions that allow them to accurately model the light pseudoscalar mesons, turning the $\mathrm{SU}(n)$ NLSM into ChPT proper.

The Nambu-Goldstone boson fields can be organised into a $n \times n$ flavour-space matrix field $u(\phi)$ [24, 25]. Under the chiral transformation $(g_L, g_R) \in G$, $u(\phi)$ transforms as

$$u(\phi) \longrightarrow g_R\, u(\phi)\, h\big[g_L, g_R, u(\phi)\big] = h\big[g_L, g_R, u(\phi)\big]\, u(\phi)\, g_L^\dagger, \qquad \text{(II.2.1)}$$

where $h \in H$ is defined by this transformation. By requiring that G can be made local while leaving the extended QCD Lagrangian invariant, it can be shown that

$$\begin{aligned}
\chi &\equiv 2B(s + ip) \longrightarrow g_R \chi g_L^\dagger, \\
\ell_\mu &\equiv v_\mu - a_\mu \longrightarrow g_L \ell_\mu g_L^\dagger - i\partial_\mu g_L g_L^\dagger, \\
r_\mu &\equiv v_\mu + a_\mu \longrightarrow g_R r_\mu g_R^\dagger - i\partial_\mu g_R g_R^\dagger,
\end{aligned} \qquad \text{(II.2.2)}$$

where B is a constant related to the leading-order (LO) meson decay constant and the $\langle \bar{q}q \rangle$ condensate.

It is possible to rewrite $u(\phi), \chi, \ell_\mu, r_\mu$ in a basis of fields that transform entirely in terms of g_L and g_R, as is done in [2] to derive the NLO ChPT Lagrangian. We instead choose to

---

[1]One can add more types of externals fields to ChPT. Examples are symmetric or antisymmetric tensors [22, 23]. These extensions are not relevant for this work.

follow [26–28] and rewrite them in a basis of fields that all transform as $X \to hXh^\dagger$:

$$u_\mu \equiv i\left[u^\dagger(\partial_\mu - ir_\mu)u - u(\partial_\mu - i\ell_\mu)u^\dagger\right],$$
$$\chi_\pm \equiv u^\dagger\chi u^\dagger \pm u\chi^\dagger u,$$
$$f_\pm^{\mu\nu} \equiv uF_L^{\mu\nu}u^\dagger \pm u^\dagger F_R^{\mu\nu}u, \tag{II.2.3}$$

where $F_L^{\mu\nu} \equiv \partial^\mu\ell^\nu - \partial^\nu\ell^\mu - i[\ell^\mu, \ell^\nu]$ and similarly for $F_R^{\mu\nu}$ and r^μ. These transformation properties are conserved under the covariant derivative ∇_μ defined as

$$\nabla_\mu X = \partial_\mu X + [\Gamma_\mu, X], \qquad \Gamma_\mu \equiv \frac{1}{2}\left[u^\dagger(\partial_\mu - ir_\mu)u + u(\partial_\mu - i\ell_\mu)u^\dagger\right]. \tag{II.2.4}$$

2.1 The ChPT Lagrangian

There exists an infinite number of possible Lagrangian terms consistent with the symmetries of ChPT. They can be organised into a power-counting hierarchy in the small energy-momentum scale p, where $u_\mu, \nabla_\mu = \mathcal{O}(p)$ and $\chi_\pm, f_\pm^{\mu\nu} = \mathcal{O}(p^2)$. Thus,

$$\mathcal{L}_{\text{ChPT}} = \mathcal{L}_2 + \mathcal{L}_4 + \mathcal{L}_6 + \dots, \tag{II.2.5}$$

where $\mathcal{L}_{2n}$ is $\mathcal{O}(p^{2n})$; odd powers are forbidden by parity. The coefficient of each term in $\mathcal{L}_{2n}$ is a separate LEC.[2]

The LO Lagrangian is

$$\mathcal{L}_2 = \frac{F^2}{4}\langle u_\mu u^\mu + \chi_+ \rangle, \tag{II.2.6}$$

where F is a LEC related to the LO meson decay constant, and $\langle \dots \rangle$ indicates a trace over flavour-space indices. The LEC of the χ_+ term is $\frac{BF^2}{4}$ as defined in eq. (II.2.2). By requiring that the kinetic term is canonically normalised, one can fix $u(\phi) = 1 + \frac{it^a\phi^a}{F\sqrt{2}} + \dots$, where t^a are the generators of $SU(n)$ and Einstein's summation convention is used. The higher-order terms depend on the choice of parametrisation, which influences the computation of amplitudes but not the amplitudes themselves.

The next-to-leading-order (NLO) Lagrangian, which was first determined in [2], is in terms of our basis[3]

$$\begin{aligned}
\mathcal{L}_4 = \quad & \hat{L}_0\langle u_\mu u_\nu u^\mu u^\nu \rangle + \hat{L}_1\langle u_\mu u^\mu \rangle^2 + \hat{L}_2\langle u_\mu u_\nu \rangle\langle u^\mu u^\nu \rangle + \hat{L}_3\langle (u_\mu u^\mu)^2 \rangle \\
& + \hat{L}_4\langle u_\mu u^\mu \rangle\langle \chi_+ \rangle + \hat{L}_5\langle u_\mu u^\mu \chi_+ \rangle + \hat{L}_6\langle \chi_+ \rangle^2 + \hat{L}_7\langle \chi_- \rangle^2 + \hat{L}_8\langle \chi_+^2 + \chi_-^2 \rangle \\
& - i\hat{L}_9\langle f_+^{\mu\nu} u_\mu u_\nu \rangle + \hat{L}_{10}\langle f_+^{\mu\nu} f_{\mu\nu}^+ - f_-^{\mu\nu} f_{\mu\nu}^- \rangle,
\end{aligned} \tag{II.2.7}$$

[2]Some "terms", like the one associated with $\hat{L}_{10}$ in eq. (II.2.7) below, actually consist of several terms. These transform into each other under the discrete symmetries of the Lagrangian, and must therefore appear with the same LEC.

[3]There are two additional *contact terms* proportional to $\langle \chi_+^2 - \chi_-^2 \rangle$ and $\langle f_+^{\mu\nu} f_{\mu\nu}^- + f_-^{\mu\nu} f_{\mu\nu}^+ \rangle$. They are needed for renormalisation but make no physical contributions to the amplitudes considered here.

where the LECs are $\hat{L}_i$. The analogous NNLO Lagrangian with 112 LECs K_i was determined in [27]. The 1862-LEC NNNLO Lagrangian, which we do not use here, was determined in [28].

For small n, the Cayley-Hamilton identity reduces the number of independent terms, and consequently the number of LECs. At $n = 3$, it is standard to eliminate $\hat{L}_0$; the remaining LECs are conventionally labelled L_i with i preserved. At $n = 2$, it is customary to also redefine the LECs slightly, resulting in the l_i of the original Gasser-Leutwyler convention [1]. At NNLO, the 112+3 K_i (ordinary+contact terms) are reduced to 90+4 C_i at $n = 3$ and 52+4 c_i at $n = 2$ as detailed in [27]. For more details on the Lagrangians for different n, see [6, 29].

The NLO renormalisation was first carried out in [1, 2], and the extension to NNLO in [29]; for more information on ChPT renormalisation, see [30]. A slightly altered $\overline{\text{MS}}$ scheme is conventionally used, with renormalisation scale $\mu = 0.77$ GeV. The renormalised LECs are denoted X_i^r where $X = \ell, L, \hat{L}$, etc. At $n = 2$ flavours it is conventional to use $\bar{\ell}_i$ instead, related to ℓ_i^r through

$$\ell_i^r = \frac{\gamma_i}{32\pi^2} \left[\bar{\ell}_i + \ln \left(\frac{M_{\text{phys}}^2}{\mu^2} \right) \right] \tag{II.2.8}$$

where M_{phys} is the chosen meson mass and γ_i are coefficients found in [1]. Effectively, eq. (II.2.8) sets the renormalisation scale to M_{phys} for $\bar{\ell}_i$.

3 Scattering amplitudes

In this section, and in the remainder of the paper, we will restrict ourselves to a simplified version of ChPT. Firstly, we will not include the external (axial) vector fields a_μ, v_μ in the Lagrangian, which essentially amounts to ignoring electroweak corrections to the amplitude. Secondly, we will assume that all mesons have the same mass M_{phys}, as mentioned in the introduction. While this limits the phenomenological applicability of three-flavour ChPT, it is a reasonable approximation that simplifies the procedure for obtaining bounds (see section 5). More importantly, the full NNLO amplitude is currently not available in the general-mass case; available results only cover $\pi\pi$ scattering in two- [30, 31] and three-flavour [32] ChPT, as well as πK scattering [33], and are not expressed in terms of elementary functions. With equal masses, we normalise all Mandelstam variables so that $s + t + u = 4$.

For the general equal-mass n-flavour scattering process $a + b \rightarrow c + d$, there are nine independent flavour structures possible: the six distinct index permutations on $\langle t^a t^b t^c t^d \rangle$ and the three on $\langle t^a t^b \rangle \langle t^c t^d \rangle$. Due to charge conjugation symmetry, a permutation is not

independent of its reverse. Thus, the scattering amplitude M may be decomposed as

$$
\begin{aligned}
M(s,t,u) = \ & \big[\langle t^a t^b t^c t^d\rangle + \langle t^d t^c t^b t^a\rangle\big] B(s,t,u) \\
& + \big[\langle t^a t^c t^d t^b\rangle + \langle t^b t^d t^c t^a\rangle\big] B(t,u,s) \\
& + \big[\langle t^a t^d t^b t^c\rangle + \langle t^c t^b t^d t^a\rangle\big] B(u,s,t) \\
& + \delta^{ab}\delta^{cd} C(s,t,u) + \delta^{ac}\delta^{bd} C(t,u,s) + \delta^{ad}\delta^{bc} C(u,s,t),
\end{aligned}
\tag{II.3.1}
$$

where s, t, u are the normalised Mandelstam variables, and crossing symmetry imposes that only two distinct functions B, C are used.[4] This is the form used in [34], where the functions B, C are given to NNLO for $\mathrm{SU}(n)$ equal-mass ChPT. The NLO results were first obtained in [1, 35].

3.1 Other forms of the amplitude

With two flavours, the traces can be evaluated in terms of Kronecker δ's, giving[5]

$$
M(s,t,u) = \delta^{ab}\delta^{cd} A(s,t,u) + \delta^{ac}\delta^{bd} A(t,u,s) + \delta^{ad}\delta^{bc} A(u,s,t),
\tag{II.3.2}
$$

which is the form used in [13] (up to reordering the arguments as permitted by the symmetries of A). In terms of the functions above,

$$
A(s,t,u) = C(s,t,u) + B(s,t,u) + B(t,u,s) - B(u,s,t).
\tag{II.3.3}
$$

the function A was first determined to NLO in [1].

With n flavours, the traces can be evaluated using the anticommutation relation $\{t^a, t^b\} = \frac{2}{n}\delta^{ab} + d^{abc}t^c$ to give[6]

$$
\begin{aligned}
M(s,t,u) = \ & d^{abe}d^{cde} B'(s,t,u) + d^{ace}d^{bde} B'(t,u,s) + d^{ade}d^{bce} B'(u,s,t) \\
& + \delta^{ab}\delta^{cd} C'(s,t,u) + \delta^{ac}\delta^{bd} C'(t,u,s) + \delta^{ad}\delta^{bc} C'(u,s,t),
\end{aligned}
\tag{II.3.4}
$$

where

$$
\begin{aligned}
B'(s,t,u) &= \tfrac{1}{2}\big[B(s,t,u) + B(t,u,s) - B(u,s,t)\big], \\
C'(s,t,u) &= C(s,t,u) + \tfrac{4}{n}B'(s,t,u).
\end{aligned}
\tag{II.3.5}
$$

---

[4] These functions have the symmetries $B(s,t,u) = B(u,t,s)$ and $C(s,t,u) = C(s,u,t)$, which is consistent with the symmetries of the respective flavour structures. Likewise, $A(s,t,u) = A(s,u,t)$ holds in eq. (II.3.2).

[5] This form can be traced back to the original current-algebra calculation [36] of the $\pi\pi$ amplitude.

[6] The relevant identity is

$$
\langle t^a t^b t^c t^d\rangle + \langle t^a t^d t^c t^b\rangle = \tfrac{1}{2}\left(d^{abe}d^{cde} + d^{ade}d^{cbe} - d^{ace}d^{bde}\right) + \tfrac{2}{n}\left(\delta^{ab}\delta^{cd} + \delta^{ad}\delta^{cb} - \delta^{ac}\delta^{bd}\right).
$$

It is most easily derived by first using $t^a t^b = \tfrac{1}{n}\delta^{ab} + \tfrac{1}{2}(d^{abc} + if^{abc})t^c$ repeatedly, and then removing all occurrences of f with the Jacobi-like identity

$$
f^{abe}f^{cde} = d^{ace}d^{bde} - d^{bce}d^{ade} + \tfrac{4}{n}\left(\delta^{ac}\delta^{bd} - \delta^{ad}\delta^{bc}\right),
$$

which is derived from the observation that $[[t^a, t^b], t^c] = \{\{t^b, t^c\}, t^a\} - \{\{t^c, t^a\}, t^b\}$.

With three flavours, the Cayley-Hamilton theorem[7] allows for the removal of one term at the expense of symmetry, leaving

$$
\begin{aligned}
M(s,t,u) = \quad & \delta^{ab}\delta^{cd}A_1(s,t,u) + \delta^{ac}\delta^{bd}A_2(s,t,u) + \delta^{ad}\delta^{bc}A_3(s,t,u) \\
& + d^{abe}d^{cde}B_1(s,t,u) + d^{ace}d^{bde}B_2(s,t,u)
\end{aligned}
\tag{II.3.6}
$$

where

$$
\begin{aligned}
B_1(s,t,u) &= B(t,u,s) - B(u,s,t), \qquad B_2(s,t,u) = B(t,u,s) - B(s,t,u), \\
A_1(s,t,u) &= C(s,t,u) + B(s,t,u) + \tfrac{1}{3}B_1(s,t,u), \\
A_2(s,t,u) &= C(t,u,s) + B(u,s,t) + \tfrac{1}{3}B_2(s,t,u), \\
A_3(s,t,u) &= C(u,s,t) + B(s,t,u) + B(u,s,t) - B(t,u,s).
\end{aligned}
\tag{II.3.7}
$$

This is the form used in [14].

3.2 Structure of the amplitude

The functions $B(s,t,u)$ and $C(s,t,u)$ consist of one part that is polynomial in the Mandelstam variables and contains the LECs, plus the so-called *unitarity correction* that is non-polynomial in the Mandelstam variables.[8] The polynomial parts are quadratic at NLO and cubic at NNLO. At NLO, the unitarity correction does not contain any LECs; at NNLO, the unitarity correction depends on the NLO LECs.

The unitarity correction at NLO depends on the function $\bar{J}$, which originates in the loop integral as shown in [1]. The NNLO unitarity correction introduces four analogous functions $k_i, i = 1, \dots 4$ [31, 32, 34, 37]. More details about these functions can be found in appendix C.

The LEC content of the amplitude considered here is more limited than that of the full ChPT Lagrangian. About half of the Lagrangian terms are dropped by not including the external (axial) vector fields, and a significant part of the NNLO Lagrangian cannot appear in a 4-particle process below NNNLO. Also, the number of LECs is reduced by the Cayley-Hamilton theorem in the 2- and 3-flavour case as described in section 2.1. Lastly, $\hat{L}_7$, K_{12}, K_{24}, K_{30}, K_{34}, K_{36}, K_{41} and K_{42}, i.e. those whose Lagrangian terms contain $\langle\chi_-\rangle$, disappear in the equal-mass limit.[9] Even with these reductions, there are still 35 (27 at $n = 3$, 18 at $n = 2$) NNLO LECs that are involved in the amplitude at hand, in addition to 8 (7, 4) NLO LECs.

---

[7]More specifically the $n = 3$ Cayley-Hamilton theorem, recast as the SU(3)-specific identity
$$3\left(d^{abe}d^{cde} + d^{bce}d^{ade} + d^{cae}d^{bde}\right) = 2\left(\delta^{ab}\delta^{cd} + \delta^{bc}\delta^{ad} + \delta^{ca}\delta^{bd}\right).$$
[8]This split is not uniquely defined, but we adhere to the conventions of [34].
[9]This can be understood by noting that χ_- has odd parity, so all terms in its expansion contain an odd number of pseudoscalar fields. If the even-parity Lagrangian term contains two traces of odd-parity objects, it can therefore only result in six-point vertices or larger, since the trace of a single field vanishes. Therefore, K_{12}, K_{24} etc. do not appear in the NNLO four-point amplitude, whereas $\hat{L}_7$ only appears in s, t, u-independent tadpole diagrams. As will be shown in section 5, we only consider s-derivatives of the amplitude, so also $\hat{L}_7$ disappears for our purposes.

3.3 Irreducible amplitudes

The scattered particles are in the adjoint representation of $SU(n)$. The Clebsch-Gordan decomposition of the initial and final states is therefore[10]

$$\text{Adj} \otimes \text{Adj} = R_I + R_S + R_A + R_S^A + R_A^S + R_S^S + R_A^A, \qquad (\text{II.3.8})$$

where R_I is the singlet representation, and the sub(super)scripts on the other representations indicate lower (upper) index pairs that are symmetric (S) or antisymmetric (A). Details on the representations and their dimensions can be found in [34, 38]. From this, it follows that the scattering amplitude can be decomposed in terms of seven corresponding irreducible amplitudes T_J. In terms of eq. (II.3.1), these are

$$
\begin{aligned}
R_I: \quad T_I &= 2\frac{n^2-1}{n}\big[B(s,t,u) + B(t,u,s)\big] - \frac{2}{n}B(u,s,t) \\
&\quad + (n^2-1)C(s,t,u) + C(t,u,s) + C(u,s,t), \\
R_S: \quad T_S &= \frac{n^2-4}{n}\big[B(s,t,u) + B(t,u,s)\big] - \frac{4}{n}B(u,s,t) \\
&\quad + C(t,u,s) + C(u,s,t), \\
R_A: \quad T_A &= n\big[B(t,u,s) - B(s,t,u)\big] + C(t,u,s) - C(u,s,t), \\
R_S^A, R_A^S: \quad T_{AS} = T_{SA} &= C(t,u,s) - C(u,s,t), \\
R_S^S: \quad T_{SS} &= 2B(u,s,t) + C(t,u,s) + C(u,s,t), \\
R_A^A: \quad T_{AA} &= -2B(u,s,t) + C(t,u,s) + C(u,s,t).
\end{aligned}
\qquad (\text{II.3.9})
$$

Only six amplitudes are needed, since T_{SA} and T_{AS} are identical due to crossing symmetry etc., as mentioned in [14].

In SU(3), the R_A^A representation vanishes, so only five amplitudes are needed. In [14], the representations are labelled by their dimensions, which are $1, 8, 8, 10$ and 27 in the order they appear in eqs. (II.3.8) and (II.3.9).

In SU(2), only R_I, R_A and R_S^S remain and have dimension $1, 3$ and 5, respectively. The corresponding amplitudes can be identified with the isospin components T^0, T^1 and T^2, respectively. In terms of eq. (II.3.2), they are

$$
\begin{aligned}
T^0 &= 3A(s,t,u) + A(t,u,s) + A(u,s,t), \\
T^1 &= A(t,u,s) - A(u,s,t), \\
T^2 &= A(t,u,s) + A(u,s,t).
\end{aligned}
\qquad (\text{II.3.10})
$$

This well-known relation can be derived from eqs. (II.3.3) and (II.3.9).

---

[10][35] contains an intuitive description of how the decomposition is performed.

3.4 Eigenstate amplitudes

A general amplitude can be expressed as $a_J T^J$, where the index J runs over the representations in the order they appear in eq. (II.3.8). For a physically applicable scattering process, however, the initial and final states should typically be taken as a product of mass eigenstates such as π, K and η. This corresponds to fixing a_J to a small selection of values so that $T(ab \to cd) = a_J(ab \to cd)T^J$. Here, as in [13, 14], we consider only elastic scattering of eigenstates, with $a_J(ab \to ab) \equiv a_J(ab)$.

With two flavours, where J runs over I, A, SS (alternatively, isospin $0, 1, 2$), the eigenstates are[11]

$$a_J(\pi^0\pi^0) = \begin{pmatrix} \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}, \quad a_J(\pi^0\pi^\pm) = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad a_J(\pi^\pm\pi^\pm) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix},$$
(II.3.11)

and with three flavours, where J runs over I, S, A, AS, SS, they are[12]

$$
\begin{aligned}
a_J(\pi^0\pi^0) &= \begin{pmatrix} \frac{1}{8} & \frac{1}{5} & 0 & 0 & \frac{27}{40} \end{pmatrix}, & a_J(\pi^\pm\pi^\pm, K^\pm K^\pm, K^0K^0) &= \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \\
a_J(\pi^0\pi^\pm) &= \begin{pmatrix} 0 & 0 & \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{pmatrix}, & a_J(K^\pm\pi^\pm, K^\pm\pi^\mp, K^0K^\pm) &= \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \\
a_J(K^0\pi^\pm) &= \begin{pmatrix} 0 & \frac{3}{10} & \frac{1}{6} & \frac{1}{3} & \frac{1}{5} \end{pmatrix}, & a_J(K\pi^0) &= \begin{pmatrix} 0 & \frac{3}{20} & \frac{1}{12} & \frac{5}{12} & \frac{7}{20} \end{pmatrix}, \\
a_J(\pi\eta) &= \begin{pmatrix} 0 & \frac{1}{5} & 0 & \frac{1}{2} & \frac{3}{10} \end{pmatrix}, & a_J(K\eta) &= \begin{pmatrix} 0 & \frac{1}{20} & \frac{1}{4} & \frac{1}{4} & \frac{9}{20} \end{pmatrix};
\end{aligned}
$$
(II.3.12)

see e.g. [13, 14], respectively.[13] With four or more flavours, ChPT loses its applicability as low-energy QCD since there are only three light quarks in the Standard Model. Therefore, there is little sense in considering eigenstates for n flavours, although we can note that $T(\pi^\pm\pi^\pm \to \pi^\pm\pi^\pm) = T_{SS}$ regardless of n.

One of our extensions over previous work is that we use all possible values for the a_J, rather than restricting them to eigenstates (see section 5 for what constitutes "possible"). This can be done without complications, since the mass eigenstates are completely degenerate in the equal-mass limit. However, it is still useful to view those states that remain mass eigenstates in the unequal-mass case as special. Below, by "eigenstate" we will specifically mean scattering between these states. Note that by treating general a_J, we effectively include inelastic scattering such as $a_J(\pi^0\pi^0 \to \pi^+\pi^-) = \begin{pmatrix} \frac{1}{3} & 0 & -\frac{1}{3} \end{pmatrix}$. However, it turns out that inelastic scattering is useless for our purposes by invariably failing to satisfy eq. (II.5.6b). This (in addition to [13, 14]) is why this section has focused mainly on elastic scattering.

3.5 Crossing symmetry

Since all amplitudes can be expressed as $a_J T^J$, crossing symmetry implies that channel crossing must take the form of a linear transformation of a_J. For $s \leftrightarrow u$ crossing, the

---

[11]a_J is invariant under particle/antiparticle exchange, so $a_J(\pi^+\pi^+) = a_J(\pi^-\pi^-) \equiv a_J(\pi^\pm\pi^\pm)$. Note, however, that $a_J(\pi^\pm\pi^\mp) \neq a_J(\pi^\pm\pi^\pm)$ — they are instead related by crossing; see eq. (II.3.14).

[12]Here, π without superscript stands for any of $\pi^\pm$ or π^0 (and similarly for K) whenever a_J is agnostic about the particular choice. We use $a_J(ab, cd, \ldots)$ for $a_J(ab) = a_J(cd) = \ldots$.

[13]Equation (II.3.12) differs from the values given in [14]: there was an error or misprint in $a_J(\pi^0\pi^\pm)$, and all eigenstates were not included, with $a_J(K^\pm\pi^\pm)$ given as $a_J(K\pi)$.

transformation $T^I(u,t,s) = C_u^{IJ} T^J(s,t,u)$ is given by [21, 38]

$$C_u^{IJ} = \begin{pmatrix} \frac{1}{n^2-1} & 1 & -1 & \frac{4-n^2}{2} & \frac{n^2(n+3)}{4(n+1)} & \frac{n^2(n-3)}{4(n-1)} \\ \frac{1}{n^2-1} & \frac{n^2-12}{2(n^2-4)} & -\frac{1}{2} & 1 & \frac{n^2(n+3)}{4(n+1)(n+2)} & \frac{n^2(3-n)}{4(n-1)(n-2)} \\ \frac{1}{1-n^2} & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{n(n+3)}{4(n+1)} & \frac{n(3-n)}{4(n-1)} \\ \frac{1}{1-n^2} & \frac{2}{n^2-1} & 0 & \frac{1}{2} & \frac{n(n+3)}{4(n+1)(n+2)} & \frac{n(n-3)}{4(n-1)(n-2)} \\ \frac{1}{n^2-1} & \frac{1}{2+n} & \frac{1}{n} & \frac{n-2}{2n} & \frac{n^2+n+2}{4(n+1)(n+2)} & \frac{n-3}{4(n-1)} \\ \frac{1}{n^2-1} & \frac{1}{2-n} & -\frac{1}{n} & \frac{n+2}{2n} & \frac{n+3}{4(n+1)} & \frac{n^2-n+2}{4(n-1)(n-2)} \end{pmatrix} \tag{II.3.13}$$

which also works at $n = 2, 3$ by removing appropriate rows and columns:

$$C_u^{IJ}\big|_{\mathrm{SU}(2)} = \frac{1}{6}\begin{pmatrix} 2 & -6 & 10 \\ -2 & 3 & 5 \\ 2 & 3 & 1 \end{pmatrix} \qquad C_u^{IJ}\big|_{\mathrm{SU}(3)} = \begin{pmatrix} \frac{1}{8} & 1 & -1 & -\frac{5}{2} & \frac{27}{8} \\ \frac{1}{8} & -\frac{3}{10} & -\frac{1}{2} & 1 & \frac{9}{8} \\ -\frac{1}{8} & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{27}{40} \\ -\frac{1}{8} & \frac{2}{5} & 0 & \frac{1}{2} & \frac{9}{40} \\ \frac{1}{8} & \frac{1}{5} & \frac{1}{3} & \frac{1}{6} & \frac{7}{40} \end{pmatrix}. \tag{II.3.14}$$

These versions can be found in [13, 14] respectively.

4 Linear constraints

In this section, we will introduce a mathematical language of *linear constraints*. This formalism is introduced before positivity bounds (see section 5) so that they can be established in full generality. In order to make the handling of the bounds as general and powerful as possible, we dedicate this section to developing some useful mathematical definitions and results.[14]

4.1 Definition and combination of constraints

For a set of parameters b_i (e.g. the LECs), a linear constraint takes the general form

$$\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n - c \geq 0, \tag{II.4.1}$$

where c, α_i are known coefficients. By treating α_i, b_i as components of vectors, this is equivalent to

$$\boldsymbol{\alpha} \cdot \boldsymbol{b} \geq c. \tag{II.4.2}$$

---

[14]In this section employ mathematical notation that, depending on the background of the reader, may not be entirely familiar. We also define new notation for our own purposes. A glossary covering all potentially unfamiliar notation is provided in appendix B.7.

We say that b lives in the *parameter space*, whereas α lives in the *constraint space*.[15] Since α and c can be rescaled by any positive scalar without changing the inequality, any linear constraint can be described by the pair $\langle \alpha, c \rangle$ with $c \in \{1, 0, -1\}$.

We say that a point b *satisfies* a constraint $\langle \alpha, c \rangle$ if $\alpha \cdot b \geq c$. We denote by $\mathcal{B}(\langle \alpha, c \rangle)$ the subset of parameter space that satisfies $\langle \alpha, c \rangle$. For any α, it is clear that the origin $b = 0$ is contained in $\mathcal{B}(\langle \alpha, -1 \rangle)$ but not in $\mathcal{B}(\langle \alpha, 1 \rangle)$, and lies on the boundary of $\mathcal{B}(\langle \alpha, 0 \rangle)$ (except when $\alpha = 0$).

The LECs will typically be subject to many linear constraints simultaneously. We will normally use the letter Ω to denote a constraint, either a single one like $\langle \alpha, c \rangle$ or a combination of several such constraints. Given two constraints Ω, Ω', we write the constraint that imposes both of them simultaneously as $\Omega + \Omega'$. A point b satisfies $\Omega + \Omega'$ if and only if it satisfies both Ω and Ω'; thus, the $\mathcal{B}$ notation naturally generalises through $\mathcal{B}(\Omega + \Omega') \equiv \mathcal{B}(\Omega) \cap \mathcal{B}(\Omega')$. For combinations of many constraints, we will generalise $+$ into e.g. $\Omega = \sum_i \langle \alpha_i, c_i \rangle$.

4.2   Stronger and weaker constraints

A hierarchy can be established among the constraints based on how strong (restrictive) they are. For instance, $b_1 \geq 1$ is stronger than $b_1 \geq 0$. We will write the stronger-than relation as $\Omega \geq \Omega'$, which holds if all points that satisfy Ω also satisfy Ω'. Thus, $\Omega \geq \Omega'$ is equivalent to $\mathcal{B}(\Omega) \subseteq \mathcal{B}(\Omega')$. Naturally, we say $\Omega = \Omega'$ if $\mathcal{B}(\Omega) = \mathcal{B}(\Omega')$, and say $\Omega > \Omega'$ if $\Omega \geq \Omega'$ but $\Omega \neq \Omega'$. Just like subset relations, our stronger-than relation is not a total ordering, as there exist many pairs of constraints Ω, Ω' where neither is stronger than the other. From our definitions, it trivially follows that

$$\begin{aligned}
(\Omega + \Omega') \geq \Omega, \qquad (\Omega + \Omega') &\geq \Omega', \\
(\Omega + \Omega') = \Omega \quad \Leftrightarrow \quad \Omega &\geq \Omega',
\end{aligned} \tag{II.4.3}$$

so that if $\Omega \ngeq \Omega'$ and $\Omega' \ngeq \Omega$, their combination $\Omega + \Omega'$ is indeed a new, strictly stronger constraint. Furthermore, we see that, for all $\lambda > 0$, $\kappa > 1$, and $\alpha \neq 0$,

$$\langle \lambda \alpha, 1 \rangle > \langle \lambda \alpha, 0 \rangle > \langle \lambda \alpha, -1 \rangle, \tag{II.4.4}$$

$$\langle \kappa \alpha, 1 \rangle < \langle \alpha, 1 \rangle, \qquad \langle \lambda \alpha, 0 \rangle = \langle \alpha, 0 \rangle, \qquad \langle \kappa \alpha, -1 \rangle > \langle \alpha, -1 \rangle. \tag{II.4.5}$$

From the $c = 0$ version of eq. (II.4.5), we see that $\langle \alpha, 0 \rangle$ is not a unique representation of the constraint, since we can freely rescale α without changing it. We may remove this ambiguity by constraining α to be a unit vector.

There exists a constraint Ω_∞, equivalent to $\langle 0, 1 \rangle$ or e.g. $\langle \alpha, 1 \rangle + \langle -\alpha, 1 \rangle$, that is not satisfied by any point. It follows that $\Omega_\infty \geq \Omega$ and $\Omega_\infty + \Omega = \Omega_\infty$ for any Ω. A constraint that is satisfied by all points, i.e. $\langle 0, -1 \rangle$ or $\langle 0, 0 \rangle$, will be called a *trivial constraint*.

---

[15]We consistently use Roman letters for vectors in parameter space and Greek letters for vectors in constraint space. In general, parameter space may be any finite-dimensional real vector space, with constraint space considered as its dual.

4.3 Determining the relationship between constraints
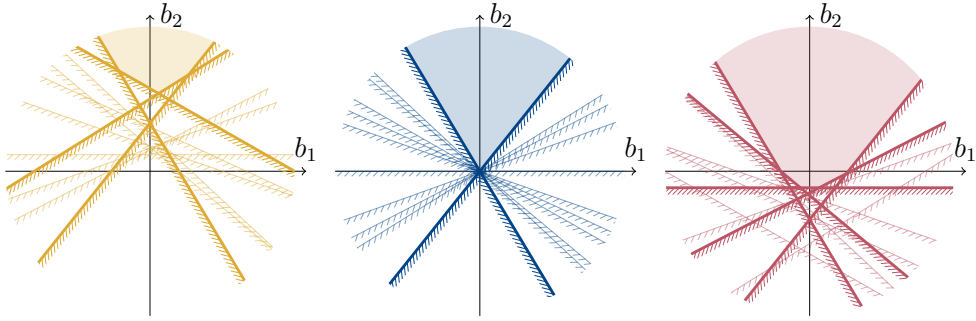
Figure 11.1: A cropped depiction of twelve random two-dimensional constraints $\langle \boldsymbol{\alpha}_i, c \rangle$ for $c = +1$ (left, **yellow**), $c = 0$ (middle, **blue**) and $c = -1$ (right, **red**) illustrated as the parameter-space lines $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} = c$. The side of the line that is *excluded* by the constraint is hatched. The region $\mathcal{B}(\omega_c)$ for $\omega_c = \sum_i \langle \boldsymbol{\alpha}_i, c \rangle$, i.e. the set of points that satisfy all the constraints, is shaded. The lines corresponding to *relevant* constraints (i.e. those that actually delimit $\mathcal{B}(\omega_c)$; this is more closely defined in section 4.4) are drawn more strongly than the rest.

We will now present a general result, which determines if a given linear constraint $\langle \boldsymbol{\beta}, c \rangle$ is weaker than an arbitrarily complicated constraint Ω. This will serve as the basis for all our uses of linear constraints.[16] For complete proofs, more details, and practical applications, see appendix B.

Consider a set of linear constraints $\langle \boldsymbol{\alpha}_i, c \rangle$ for i in some finite set $I \subset \mathbb{N}$.[17] Note that c is the same for all constraints. Then let $\omega_c \equiv \sum_{i \in I} \langle \boldsymbol{\alpha}_i, c \rangle$;[18] an example of such a constraint is given in fig. 11.1. Then define $\mathcal{A}(\omega_c)$ as the set of all points that can be expressed as

$$\sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i \geq 0, \qquad \sum_{i \in I} \lambda_i \in \Lambda_c, \tag{11.4.6}$$

where

$$\Lambda_1 = [1, \infty), \qquad \Lambda_0 = [0, \infty), \qquad \Lambda_{-1} = [0, 1]. \tag{11.4.7}$$

The shape of $\mathcal{A}(\omega_c)$ is illustrated in fig. 11.2. With these definitions, the following holds:

Proposition 4.1 (determining if constraint is weaker, special case). *Let $\langle \boldsymbol{\beta}, c \rangle$ be a single linear constraint, and let $\omega_c \neq \Omega_\infty$ be defined as above. Then $\langle \boldsymbol{\beta}, c \rangle \leq \omega_c$ if and only if $\boldsymbol{\beta} \in \mathcal{A}(\omega_c)$.*

---

[16]Proposition 4.1, along with a version of the notation we use here, was defined in [21], although the proof was completely different. An incorrect version of proposition 4.2 was also presented without proof. To the best of our knowledge, these results are novel, although the relevant literature is vast and lies outside our area of expertise. The closest we have found is [39], although their algorithm requires knowing a point that satisfies Ω, relies on more complicated mathematical machinery, and does not include all the extensions presented further below in sections B and 4.4.

[17]It is crucial that only finite combinations of constraints are considered, and it will normally be tacitly assumed that all sets like I are finite. A limited extension to infinite sets is covered in appendix B.6.

[18]We use lowercase ω here to emphasise that it is not a general constraint. A similar treatment of general Ω is given below.
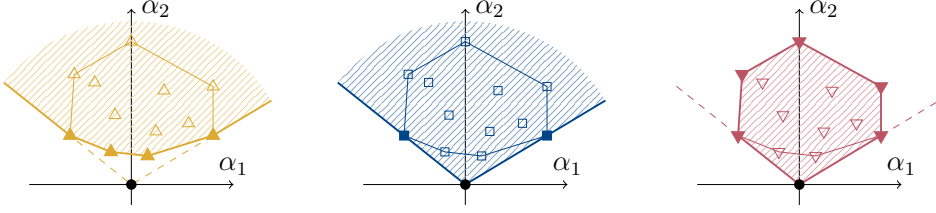
Figure II.2: Examples of the regions $\mathcal{A}(\omega_1)$ (left, yellow), $\mathcal{A}(\omega_0)$ (middle, blue) and $\mathcal{A}(\omega_{-1})$ (right, red) in constraint space, using the same $\boldsymbol{\alpha}_i$ as in fig. II.1. The $\boldsymbol{\alpha}_i$ are represented as points ($\triangle$, $\square$, $\triangledown$, respectively) in the space, and the relevant ones are filled (again, relevancy is defined in section 4.4). The convex hulls (as defined in eq. (II.4.8)) of the $\boldsymbol{\alpha}_i$ are outlined. For comparison to fig. II.1, it is helpful to remember that $\boldsymbol{\alpha}_i$ are normal vectors to the lines shown there, and that larger $|\boldsymbol{\alpha}_i|$ correspond to lines passing *closer* to the origin. Note how, given identical $\boldsymbol{\alpha}_i$, $\mathcal{A}(\omega_0)$ is the union of $\mathcal{A}(\omega_1)$ and $\mathcal{A}(\omega_{-1})$ (this is easy to see from eq. (II.4.7)) whereas the set of relevant constraints is the intersection of the respective sets.

This is proven in appendix B.1. If ω_c is a single linear constraint, this result reduces down to eq. (II.4.5). The condition $\omega_c \neq \Omega_\infty$ is necessary, since there exist corner cases where $\omega_c = \Omega_\infty$ but $\mathcal{A}(\omega_c)$ fails to cover the entire constraint space.[19] However, if $\mathcal{A}(\omega_c)$ does cover the entire space, then it is certain that $\omega_c = \Omega_\infty$.

In fig. II.2, we may note that $\mathcal{A}(\omega_c)$ is closely related to the *convex hull* of the $\boldsymbol{\alpha}_i$. In essence, $\mathcal{A}(\omega_c)$ is obtained by forming the hull, and then also including all points that give weaker constraints under eq. (II.4.5). We may also note that the convex hull can be defined as

$$\text{Hull}\left(\{\boldsymbol{\alpha}_i\}_{i\in I}\right) = \left\{ \sum_{i\in I} \lambda_i \boldsymbol{\alpha}_i \;\middle|\; \lambda_i \geq 0, \; \sum_{i\in I} \lambda_i = 1 \right\}, \tag{II.4.8}$$

which is very similar to eq. (II.4.6).

Now, let us handle the general case. The most general combination of a finite number of linear constraints can be expressed as

$$\Omega = \sum_{i\in I_1} \langle \boldsymbol{\alpha}_i, 1 \rangle + \sum_{i\in I_0} \langle \boldsymbol{\alpha}_i, 0 \rangle + \sum_{i\in I_{-1}} \langle \boldsymbol{\alpha}_i, -1 \rangle, \tag{II.4.9}$$

where I_c are some disjoint, finite, and possibly empty sets. We may compactly write this as $\Omega = \sum_{i\in I} \langle \boldsymbol{\alpha}_i, c_i \rangle$ where $I \equiv I_{+1} \cup I_0 \cup I_{-1}$ and $c_i = c$ if $i \in I_c$.

Similarly to $\mathcal{A}(\omega_c)$, let $\mathcal{A}_c(\Omega)$ be the set of all points that can be expressed as (recall that $c \in \{1, 0, -1\}$)

$$\sum_{i\in I_{-1}} \lambda_i \boldsymbol{\alpha}_i + \sum_{i\in I_0} \lambda_i \boldsymbol{\alpha}_i + \sum_{i\in I_1} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i \geq 0, \tag{II.4.10}$$

---

[19]A trivial example of this is $\omega_1 = \langle \mathbf{0}, 1 \rangle$, where $\mathcal{A}(\omega_1) = \{\mathbf{0}\}$.

with λ_i constrained by the condition

$$\sum_{i \in I_1} \lambda_i - \sum_{i \in I_{-1}} \lambda_i \geq c, \tag{II.4.11}$$

An illustration of $\mathcal{A}_c(\Omega)$ can be found in appendix B.4.5. With these definitions, the following holds:

Proposition 4.2 (determining if constraint is weaker, general case). *Let $\langle \boldsymbol{\beta}, c \rangle$ be a linear constraint, and let $\Omega \neq \Omega_\infty$ be defined as above. Then $\langle \boldsymbol{\beta}, c \rangle \leq \Omega$ if and only if $\boldsymbol{\beta} \in \mathcal{A}_c(\Omega)$.*

This is proven in appendix B.1.4. If only one of the I_c is nonempty, this reduces down to proposition 4.1.

While it is not as useful for the purposes of proposition 4.2, one may note that eqs. (II.4.10) and (II.4.11) can be more succinctly stated as

$$\mathcal{A}_c(\Omega) = \left\{ \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i \;\middle|\; \lambda_i \geq 0, \sum_{i \in I} \lambda_i c_i \geq c \right\}. \tag{II.4.12}$$

This definition of $\mathcal{A}_c(\Omega)$ works also if c, c_i are not constrained to $\{-1, 0, 1\}$.

4.4 Representations and degeneracy

Checking if $\boldsymbol{b}$ satisfies Ω becomes computationally expensive if Ω is the combination of many different linear constraints. However, Ω is usually not uniquely determined by how it is expressed as a sum of linear constraints, and it is possible to vastly reduce that redundancy. To that end, we define a *representation* of a constraint Ω as any finite set $\mathcal{S}$ of linear constraints with the property[20]

$$\Omega = \sum_{\langle \boldsymbol{\alpha}, c \rangle \in \mathcal{S}} \langle \boldsymbol{\alpha}, c \rangle. \tag{II.4.13}$$

If it is implicit which representation is used for Ω, we may call the $\langle \boldsymbol{\alpha}, c \rangle \in \mathcal{S}$ the *elements of* Ω.

It is clear that there exist *minimal representations*, i.e. representations with the smallest number of elements. As we will see below, there is often a unique minimal representation, which we will label $\mathcal{R}(\Omega)$. However, there is an important exception to this: when $\mathcal{B}(\Omega)$ is contained in a hyperplane. This happens when there are some $\boldsymbol{\delta}, d$ such that $\boldsymbol{\delta} \cdot \boldsymbol{b} = d$ for all $\boldsymbol{b} \in \mathcal{B}(\Omega)$, or equivalently $\langle \boldsymbol{\delta}, d \rangle + \langle -\boldsymbol{\delta}, -d \rangle \leq \Omega$. We will call Ω *degenerate* if so is the case, and *non-degenerate* otherwise.[21] With this in mind, we can state the following result:

---

[20]Clearly, all Ω also admit representation as a sum of an *infinite* number of constraints. However, we will not consider such representations, and proposition 4.3 below generally only holds if Ω can be expressed as a finite sum. See appendix B.6 for a discussion about infinite sums of constraints.

[21]As defined here, Ω_∞ would be considered a special case of a degenerate constraint. In the closer study of degenerate constraints given in appendix B.2.1, it turns out to be more useful to consider Ω_∞ seperately, viewing it as neither degenerate nor non-degenerate.

Proposition 4.3 (finding relevant constraints, non-degenerate case). *If Ω is a non-degenerate constraint, there exists a minimal representation $\mathcal{R}(\Omega)$ that is unique up to the normalisation of its elements. Furthermore, for any representation $\mathcal{S}$ of Ω, the relation $\mathcal{R}(\Omega) \subseteq \mathcal{S}$ is true up to normalisation.*

The elements of $\mathcal{R}(\Omega)$ are exactly those $\langle \boldsymbol{\alpha}, c \rangle \leq \Omega$ for which there is some $\boldsymbol{b} \in \mathcal{B}(\Omega)$ such that $\boldsymbol{\alpha} \cdot \boldsymbol{b} = c$ and $\boldsymbol{\beta} \cdot \boldsymbol{b} > d$ for all $\langle \boldsymbol{\beta}, d \rangle \leq \Omega$ with $\langle \boldsymbol{\beta}, d \rangle \neq \langle \boldsymbol{\alpha}, c \rangle$.

This is proven in appendix B.3. Due to this uniqueness, and the fact that $\mathcal{R}(\Omega)$ is a subset of any representation, we will call the elements of $\mathcal{R}(\Omega)$ the *relevant elements of Ω*, and call all other elements of any representation *irrelevant*, since they can be discarded without altering Ω. A more practical way of finding $\mathcal{R}(\Omega)$, based on proposition 4.2, is given in appendix B.4.4.

When Ω is degenerate, there is typically no unique minimal representation, although there is still a straightforward way to find *some* minimal representation, which we will also label $\mathcal{R}(\Omega)$. This generalisation of proposition 4.3 is discussed in appendix B.2.1, along with a more general method of replacing any degenerate constraint with a non-degenerate analogue in a lower-dimensional space. Note, however, that degenerate constraints are only a corner case with little practical relevance: a small perturbation, e.g. by numerical error, to the elements of a degenerate constraint will either render it non-degenerate, or render it equal to Ω_∞.

5 Positivity bounds

Equipped with the notion of linear constraints, we are ready to move on to the main topic of this paper: positivity bounds. (For a more detailed version of this derivation, see [13]; various generalisations can be found in e.g. [19, 20].) We start by writing down the fixed-t dispersion relation for the amplitude $a_J T^J$:

$$a_J T^J(s,t) = \frac{1}{2\pi i} \oint_\gamma \mathrm{d}z \frac{a_J T^J(z,t)}{z-s}. \tag{II.5.1}$$

The amplitude has two branch cuts along the real axis: a right-hand cut starting at $z = 4$ corresponding to the s-channel, and a left-hand cut starting at $z = -t$ corresponding to the u-channel. The discontinuity across these cuts is $T(z+i\varepsilon) - T(z-i\varepsilon) = 2i \operatorname{Im} T(z+i\varepsilon)$. For real s in the span $-t < s < 4$, we deform the contour γ as shown in fig. II.3. We can then reexpress the integral in terms of the discontinuities, which may require derivatives (subtractions) to make the contour at infinity vanish. Using the crossing relation derived in section 3.5 to rewrite the u-channel cut in terms of s, the result is

$$a_J \frac{\mathrm{d}^k}{\mathrm{d}s^k} T^J(s,t) = \frac{k!}{\pi} \int_4^\infty \mathrm{d}z \left[\frac{a_J}{(z-s)^{k+1}} + \frac{(-1)^k a_I C_u^{IJ}}{(z-u)^{k+1}} \right] \operatorname{Im} T^J(z+i\varepsilon, t). \tag{II.5.2}$$

The Froissart bound [40] shows that the integral converges whenever $k \geq 2$, since $T^J(z + i\varepsilon, t) = \mathcal{O}(s \ln^2 s)$.[22] We will discuss specific values for k in section 5.2; here, we keep it general.
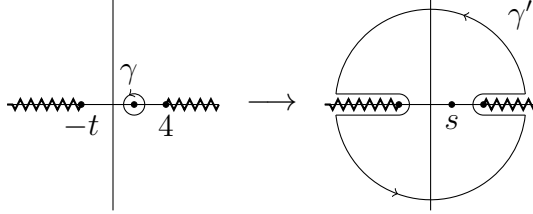
Figure II.3: The contour integral in the z-plane around $z = s$ used in the dispersion relation.

Above threshold, the partial-wave expansion of the amplitude takes the form[23]

$$T^J(s,t) = \sum_{\ell=0}^{\infty}(2\ell + 1)f_\ell^J(s)P_\ell\left(1 + \frac{2t}{s-4}\right), \tag{II.5.3}$$

where f_ℓ^J are partial wave amplitudes, P_ℓ are Legendre polynomials, and the expression in parentheses is the cosine of the scattering angle. The optical theorem then gives

$$\operatorname{Im} f_\ell^J(s) = s\beta(s)\sigma_\ell^J(s), \qquad \beta(s) \equiv \sqrt{1 - \frac{4}{s}}, \tag{II.5.4}$$

which is positive above threshold since the partial-wave cross-sections σ_ℓ^J are always positive. Therefore,

$$\operatorname{Im} T^J(s,t) = \sum_{\ell=0}^{\infty}(2\ell + 1)s\beta(s)\sigma_\ell^I P_\ell\left(1 + \frac{2t}{s-4}\right) \tag{II.5.5}$$

is positive above threshold as long as P_ℓ is. Since $P_\ell(z) \geq 0$ when $z \geq 1$, eq. (II.5.2) therefore imposes the constraint that, for any $t \in [0,4]$, $s \in [-t, 4]$ and any representation index J,

$$a_J \frac{\mathrm{d}^k}{\mathrm{d}s^k}T^J(s,t) \geq 0 \tag{II.5.6a}$$

$$\text{if} \quad a_I\left\{\delta^{IJ}\left[\frac{z-u}{z-s}\right]^{k+1} + (-1)^k C_u^{IJ}\right\} \geq 0 \quad \text{for all } z \geq 4. \tag{II.5.6b}$$

---

[22]Note that $T^J(z + i\varepsilon, t)$ on the right-hand side of eq. (II.5.2) is the exact, non-perturbative amplitude — see e.g. [13]. Indeed, the perturbative ChPT amplitude at any fixed order grows polynomially with s, so it violates the Froissart bound. We can insert the fixed-order perturbative amplitude at the left-hand side thanks to the smallness of s (and t), which guarantees good agreement with the exact one.

[23]There is a limited domain of validity for this expansion, but it does not affect the range of s, t used by us. Again, see [13] for details.

The region in the s, t plane where this holds is shown in fig. 11.4. Note that $u \in [-4, 4]$, so the expression in square brackets above is always positive.

Up to and including NNLO, the second derivative of T^J is linear in all LECs, so we obtain from eq. (11.5.6a) an expression of the form

$$\sum_i \alpha_i \hat{L}_i^r + \sum_j \beta_j K_j^r + \gamma \geq 0, \tag{11.5.7}$$

where the coefficients $\alpha_i, \beta_i, \gamma$ are functions of s, t and a_J, but not of the LECs. This constitutes a linear constraint, and each valid choice of s, t and a_J potentially yields a different constraint. The result of combining these constraints will be that only a limited region in parameter space ($\mathcal{B}(\Omega)$ in the notation of section 4) satisfies the positivity bounds. With some luck, the boundary of this region is close enough to the experimentally measured value to improve on its uncertainty (carefully considering also the uncertainty of the bounds).
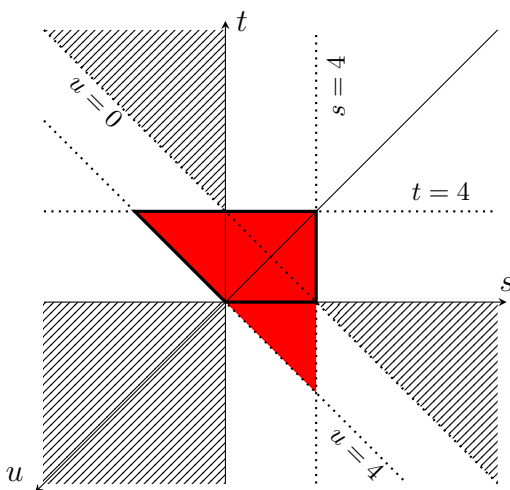
Figure 11.4: The plane of normalised Mandelstam variables. The red triangle is the region where the amplitude is real and free from singularities or branch cuts. The positivity conditions eq. (11.5.6) are valid inside the outlined part. The hatched regions with s, t or u positive are the physical regions for the respective channels.

5.1 Conditions on a_J

If we demand that eq. (11.5.6b) holds in the entire allowed s, t region, we see that the factor in square brackets can be made arbitrarily large or small by varying s, u, z. Therefore, we obtain the independent conditions $a_J \geq 0$ and $(-1)^k a_I C_u^{IJ} \geq 0$. However, we may apply the dispersion relation independently to each fixed s, t. Then, eq. (11.5.6b) can be made

less restrictive, and a wider range of constraints on the LECs can be generated. This also includes permitting odd k for some s, t.

While we may fix s and t (which in turn fixes u), we must still allow z to cover its entire range. Therefore, finding all valid a_J for given s, t presents some practical issues. We solve this by using the technology of section 4, since eq. (11.5.6b) is a set of linear constraints on the vector a_I; we may write it compactly as $\boldsymbol{a} \cdot \boldsymbol{\beta}^J(z) \geq 0$. Noting that $\left(\frac{z-u}{z-s}\right)^{k+1}$ is monotonic as a function of $z \in [4, \infty)$, we see that it is always possible to write $\boldsymbol{\beta}^J(z) = \mu \boldsymbol{\beta}^J(4) + (1-\mu)\boldsymbol{\beta}^J(\infty)$ for $\mu \in [0, 1]$. By propositions 4.1 and 4.3 (see also proposition B.3), it follows that only $\boldsymbol{\beta}^J(4)$ and $\boldsymbol{\beta}^J(\infty)$ are relevant constraints on $\boldsymbol{a}$. Thus, it is sufficient to evaluate eq. (11.5.6b) at $z = 4$ and $z = \infty$, rather than letting z cover its entire range.

Another practical problem is that the set of allowed a_J is typically unbounded. However, eq. (11.5.6a) is independent of the magnitude of a_J. The obvious solution is to fix the normalisation of the vector $\boldsymbol{a}$, but this is problematic since a linear constraint on $\boldsymbol{a}$ is not necessarily a linear constraint on $\frac{1}{|\boldsymbol{a}|}\boldsymbol{a}$. Instead, we may simply rescale $\boldsymbol{a}$ so that $\sum_J a_J = 1$. This does not cover all possible $\boldsymbol{a}$ (for that, we must also look at $\sum_J a_J = 0$ and $\sum_J a_J = -1$), but it turns out that eq. (11.5.6b) is only satisfied by $\boldsymbol{a}$ for which this works. Using this, constraints on a_I are shown in fig. 11.5.

5.2 The number of derivatives

As mentioned before, eq. (11.5.6) requires $k \geq 2$ to be valid, and $k = 2$ is sufficient; indeed, [41] claims that this value produces the best bounds. However, nothing prevents us from taking more derivatives, and with our generalised methods, we do find new relevant bounds from larger k; see e.g. figs. 11.6 and 11.12 below. Also [19] makes use of higher derivatives.

At NLO, the LECs only enter through the second-order polynomial part of the amplitude, so the third and higher derivatives are parameter-independent and do not generate any bounds. This is not the case at NNLO, where the polynomial part is third-order, and where the non-polynomial unitarity correction contains NLO LECs. Therefore, $k = 3$ should yield another set of bounds on the NNLO LECs, and $k \geq 4$ should add bounds on the NLO LECs not obtainable from the NLO-only amplitude.

It also turns out that odd k cannot be used at any order in the 2-flavour case. To see this, look explicitly at eq. (11.5.6b) at $z = \infty$:

$$\delta^{IJ} - C_u^{IJ} = \tfrac{1}{6}c_1^I c_2^J \qquad \text{where} \quad c_1 = \begin{pmatrix} 2 & 1 & -1 \end{pmatrix}, c_2 = \begin{pmatrix} 2 & 3 & -5 \end{pmatrix} \qquad (11.5.8)$$

Due to what seems to be a coincidence, the matrix factorises into a direct product, and since c_2 has different-sign elements, no nonzero a_I satisfies $(a_I c_1^I)c_2^J \geq 0$ for all J. No such coincidences hinder odd k at 3 or more flavours, and we have explicitly found a_J that satisfy eq. (11.5.6b) with odd k at 3 and 4 flavours (these turn out to produce very weak bounds, though). Even $k > 2$ remain permitted also at 2 flavours.
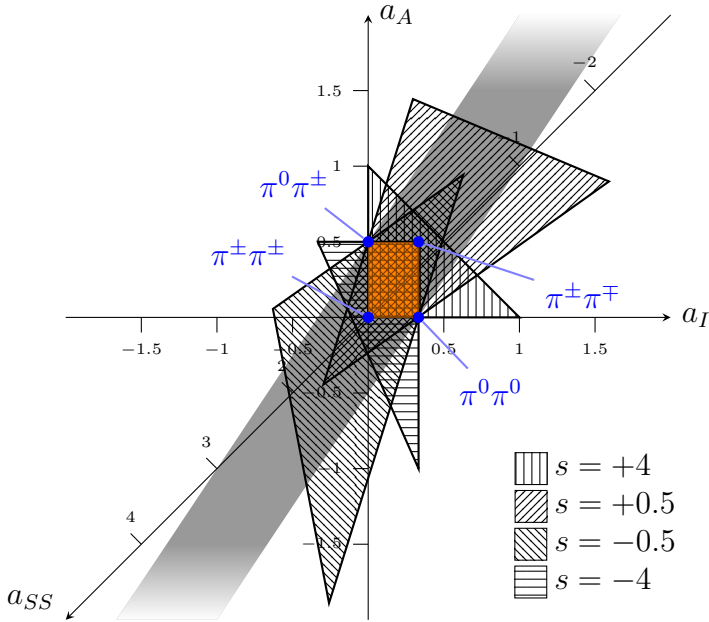
Figure II.5: Illustration of which a_I, a_A, a_{SS}, normalised so that $a_I + a_A + a_{SS} = 1$, are permitted by eq. (II.5.6b) for $n = 2$, $k = 2$ at $t = 4$ and various fixed s. The shaded band is the region permitted by the $z = \infty$ bounds (see the discussion in section 5.1). It is independent of s and extends to infinity. The hatched triangles are the regions permitted by the $z = 4$ bounds for various s as indicated. Thus, the a_J permitted at fixed s, t is the overlap between the triangle and the band. The **orange** rectangle is the region permitted at all s, t. The **blue** points represent the eigenstates, including $a_J(\pi^\pm \pi^\mp) = C_u^{IJ} a_I(\pi^\pm \pi^\pm)$ in addition to those given in eq. (II.3.11). Not shown is the inelastic scattering $a_J(\pi^0 \pi^0 \to \pi^+ \pi^-) = \left(\frac{1}{3} \quad 0 \quad -\frac{1}{3} \right)$, which never satisfies eq. (II.5.6b). For $n \geq 2$, the permitted region has an analogous shape, albeit in 4 ($n = 3$) and 5 ($n \geq 4$) dimensions, respectively. Like for $n = 2$, the eigenstate scattering amplitudes are mainly located in the corners of the always-permitted region.

5.3 The value of t

There is no immediately obvious reason to favour any specific part of the allowed s, t region when producing bounds. However, one may note that at NLO with $k \geq 2$, the only part of $\frac{\mathrm{d}^k}{\mathrm{d}s^k} T^J(s, t)$ that depends on s, t is the LEC-independent unitarity correction, which manifests itself as the term labelled γ in eq. (II.5.7). Therefore, at fixed a_J the most restrictive bound is obtained by minimising γ. It turns out that given s, the magnitude of $\frac{\mathrm{d}^k}{\mathrm{d}s^k} T^J(s, t)$ tends to increase with t, with minima and maxima always falling along the $t = 4$ line. Therefore, it is expected that all relevant constraints should be found with $t = 4$. While we see no clear *a priori* reason for it to be so, we have verified it by scanning

the entire s, t range for bounds; all relevant ones were found at $t = 4$, within numerical uncertainty.

At NNLO, also the α_i in eq. (II.5.7) may depend on s, t, so the simple argument above does not hold. However, the NNLO corrections are far too small to affect the overall shape of the amplitude, so for NLO LECs, fixing $t = 4$ should remain sufficient. The situation is yet more complicated for the NNLO LECs, since it turns out that certain combinations of them only feature in the amplitude when $t \neq 4$ (see the next section). Therefore, complete NNLO bounds require using the full s, t range.

5.4 Independently bounded parameters

While the parameter space affected by our bounds is technically the full space of (N)NLO LECs, it is of course impractical to work in such a large and redundant space. Many LECs do not receive any bounds at all by not appearing in the Lagrangian relevant for meson-meson scattering (see section 3.2), and others only appear in fixed linear combinations. Specifically, all NLO LECs that appear in the amplitude do so independently, but the NNLO LECs only appear in combinations; consequently, it is not possible to obtain bounds on the individual K_i^r. For instance, we shall see below that the combination $K_4^r - 2K_2^r$ appears in the amplitude and therefore receives bounds,[24] but its complement $2K_4^r + K_2^r$, which does not appear, is free to assume any value. Therefore, nothing can be said about the values of K_2^r and K_4^r themselves.

We will therefore reexpress our parameter space in terms of NLO LECs in addition to a new set of independently bounded parameters built from the NNLO LECs. Their form can be deduced from the polynomial parts of the functions B, C described in section 3, which, following [34], are

$$B_P(s, t, u) = \gamma_1 + \gamma_2 t + \gamma_3 t^2 + \gamma_4(s - u)^2 + \gamma_5 t^3 + \gamma_6 t(s - u)^2 + \text{(NLO)}, \quad \text{(II.5.9a)}$$

$$C_P(s, t, u) = \delta_1 + \delta_2 s + \delta_3 s^2 + \delta_4(t - u)^2 + \delta_5 s^3 + \delta_6 s(t - u)^2 + \text{(NLO)}. \quad \text{(II.5.9b)}$$

Here, γ_i, δ_i are linear combinations of the NNLO LECs, and "(NLO)" contains all NLO LECs, constant terms, etc.[25,26] Therefore, bounds on the NNLO LECs only come in the form of bounds on γ_i, δ_i. Of course, not all γ_i, δ_i are bounded either — those with $i < 3$ vanish in the second derivative of the amplitude and therefore receive no bounds, and those with $i < 6$ vanish in the third. The remaining combinations are also not necessarily independent, so we will proceed to remould them into a better set of parameters.

---

[24]This combination is, up to a scale factor, Δ_3 as defined below and explicitly given in appendix A.

[25]This differs from the convention in [34], where the NLO terms are included in γ_i, δ_i. The "NNLO parts" that we extract here are easily read off from the appendices to that paper.

[26][19] use a similar approach in their 2-flavour NNLO bounds, but do not separate NLO and NNLO parts. This results in a smaller parameter space (6 dimensions compared to our 8; see section 5.4.4), but our approach has the benefit of separating the relatively well-determined $\bar{l}_i$ from the much more uncertain NNLO values, allowing for figures such as fig. II.8. The fact that our parametrisation remains partly redundant is not a major issue, since we always fix some parameters rather than working in the full space.

5.4.1 General number of flavours

As is discussed in section 5.3, all relevant NLO bounds appear at $t = 4$, so we may expect that this particular t-value is special also at NNLO. Therefore, we express the polynomial parts in terms of s and $\bar{t} \equiv 4 - t$, using $u = \bar{t} - s$:

$$
\begin{aligned}
B_P(s,t,u) &= s^2\Gamma_1 + \bar{t}s^2(\Gamma_4 - 3\Gamma_3) + \dots, \\
B_P(t,u,s) &= s^2\Gamma_2 - s^3\Gamma_3 + \bar{t}s^2\Gamma_4 + \dots, \\
B_P(u,s,t) &= s^2\Gamma_2 + s^3\Gamma_3 + \bar{t}s^2(\Gamma_4 - 3\Gamma_3) + \dots,
\end{aligned}
\tag{II.5.10}
$$

where "..." consists of terms that vanish in the second derivative of the amplitude. Here, we have defined the parameters

$$
\begin{aligned}
\Gamma_1 &= 4\gamma_4 + 16\gamma_6, & \Gamma_2 &= \gamma_3 + \gamma_4 + 8\gamma_6, \\
\Gamma_3 &= \gamma_5 + \gamma_6, & \Gamma_4 &= 3\gamma_5 - \gamma_6,
\end{aligned}
\tag{II.5.11}
$$

of which Γ_4 only receives bounds when $t \neq 4$ due to the presence of $\bar{t}$. Similarly,

$$
\begin{aligned}
C_P(s,t,u) &= s^2\Delta_2 + s^3\Delta_3 + \bar{t}s^2(\Delta_4 - 3\Delta_3) + \dots, \\
C_P(t,u,s) &= s^2\Delta_1 + \bar{t}s^2(\Delta_4 - 3\Delta_3) + \dots, \\
C_P(u,s,t) &= s^2\Delta_2 - s^3\Delta_3 + \bar{t}s^2\Delta_4 + \dots,
\end{aligned}
\tag{II.5.12}
$$

where Δ_i are defined in terms of δ_i identically to eq. (II.5.11).

These 8 parameters Δ_i, Γ_i constitute a minimal set of parameters for NNLO bounds with a general number of flavours; explicit expressions are given in appendix A.[27] At 2 and 3 flavours, the Cayley-Hamilton identity allows for further reduction of the number of parameters.

5.4.2 Two flavours

Decomposing the polynomial part of $A(s,t,u)$ using eqs. (II.3.3), (II.5.10) and (II.5.12) reveals

$$
\begin{aligned}
A_P(s,t,u) &= s^2\Theta_2 - s^3\Theta_3 + \bar{t}s^2\Theta_4 + \dots + \dots, \\
A_P(t,u,s) &= s^2\Theta_1 + \bar{t}s^2(\Theta_4 - 3\Theta_3), \\
A_P(u,s,t) &= s^2\Theta_2 + s^3\Theta_3 + \bar{t}s^2(\Theta_4 - 3\Theta_3) + \dots,
\end{aligned}
\tag{II.5.13}
$$

where

$$
\begin{aligned}
\Theta_1 &= 2\gamma_3 - 2\gamma_4 + 4\delta_4 + 16\delta_6, & \Theta_2 &= 4\gamma_4 + 16\gamma_6 + \delta_3 + \delta_4 + 8\delta_6, \\
\Theta_3 &= 2\gamma_5 + 2\gamma_6 - \delta_5 - \delta_6, & \Theta_4 &= 3\gamma_5 - \gamma_6 - 4\delta_6.
\end{aligned}
\tag{II.5.14}
$$

Θ_4, like Γ_4 and Δ_4, is only bounded when $t \neq 4$. Explicit expressions and experiment-based reference values for Θ_i are given in appendix A.

---

[27] While they can technically be considered LECs, we will avoid confusion by referring to the Γ_i, Δ_i as *NNLO parameters*, reserving "LEC" for the coefficients appearing in the standard form of the Lagrangian.

5.4.3 Three flavours

A similar but less elegant simplification is possible in the 3-flavour case, using eq. (II.3.6):[28]

$$\begin{aligned}
B_{1P}(s,t,u) &= -2s^3\Gamma_3 + 3\bar{t}s^2\Gamma_3, \\
B_{2P}(s,t,u) &= s^2\Xi_1 - s^3\Gamma_3 + 3\bar{t}s^2\Xi_4, \\
A_{1P}(s,t,u) &= s^2\Xi_2 + s^3(\Delta_3 - \tfrac{2}{3}\Gamma_3) + \bar{t}s^2\Xi_4, \\
A_{2P}(s,t,u) &= s^2\Xi_3 + \tfrac{2}{3}s^3\Gamma_3 + \bar{t}s^2\Xi_4, \\
A_{3P}(s,t,u) &= s^2\Xi_2 + s^3(2\Gamma_3 - \Delta_3) + \bar{t}s^2(\Xi_4 + 3\Delta_3 - 4\Gamma_3),
\end{aligned} \qquad \text{(II.5.15)}$$

where

$$\begin{aligned}
\Xi_1 &= \gamma_3 - 3\gamma_4 - 8\gamma_6, & \Xi_2 &= 5\gamma_4 + 16\gamma_6 + \delta_3 + \delta_4 + 8\delta_6, \\
\Xi_3 &= \tfrac{4}{3}\gamma_3 + \tfrac{16}{3}\gamma_6 + 4\delta_4 + 16\delta_6, & \Xi_4 &= \gamma_5 - 3\gamma_6 - 4\delta_6.
\end{aligned} \qquad \text{(II.5.16)}$$

Again, expressions and values are given in appendix A.

5.4.4 The full parameter space

Table II.1 summarises the parameters affected by our bounds at different orders and number of flavours. Note how the dimension of the space ranges from 2 (NLO 2-flavour) to 16 (NNLO n-flavour, $n \geq 4$). If t is fixed to 4 at NNLO, this is reduced by 1 if $n \leq 3$ and by 2 otherwise; as discussed in section 5.3, $t = 4$ is the only relevant value at NLO. The difficulties associated with large parameter spaces are discussed in section 6.

Table II.1: Summary of the NLO LECs and NNLO parameters appearing in the second derivative of the amplitude. The parameters $\Gamma_i', \Delta_i', \Theta_i', \Xi_i'$ as defined in eqs. (II.5.11), (II.5.14) and (II.5.16). The LECs and parameters that remain in the third derivative are underlined, and those that remain also in the fourth and above are doubly underlined. Parameters that only feature in the amplitude when $t \neq 4$ are placed in parentheses.

Flavours	NLO	NNLO	
2	$\bar{l}_1, \bar{l}_2$	$\underline{\bar{l}_1, \dots \bar{l}_4}$;	$\Theta_1, \Theta_2, \underline{\Theta_3}, (\Theta_4)$
3	L_1^r, L_2^r, L_3^r	$\underline{L_1^r, \dots, L_6^r, L_8^r}$;	$\Xi_1, \Xi_2, \Xi_3, (\Xi_4), \underline{\Gamma_3, \Delta_3}$
≥ 4	$\hat{L}_0^r, \dots, \hat{L}_3^r$	$\underline{\underline{\hat{L}_0^r, \dots, \hat{L}_6^r, \hat{L}_8^r}}$;	$\Gamma_1, \Gamma_2, \underline{\Gamma_3}, (\Gamma_4), \Delta_1, \Delta_2, \underline{\Delta_3}, (\Delta_4)$

---

[28]A more symmetric result would have been obtained by eliminating the $d^{ace}d^{bde}$ term in eq. (II.3.4) instead, but we choose to follow [14].

5.5 The absence of catastrophic divergences

At NLO and above, the coefficient γ in eq. (II.5.7) diverges in the limit $s \to 0$ or $u \to 0$. If the divergence is towards positive infinity, this is not a problem — it simply means that the positivity bound becomes trivial in these limits. However, divergence towards negative infinity would be catastrophic, since no finite LECs could satisfy the positivity condition. If there were some value of a_J for which the divergence is in this direction, the theory would be inconsistent.

The situation becomes more complicated at NNLO, where also α_i diverge. If α_i diverge at the same rate or faster than γ, the positivity conditions remain sensible also in these limits, but if γ were to diverge towards negative infinity faster than α_i, we would again have inconsistencies.

As s approaches 4 from below, the kth derivative of the amplitude diverges as odd powers (up to $2k - 1$) of $1/\delta$, where $s = 4(1 - \delta^2)$; see appendix C.2 for details. Let q^J be the coefficient of the leading divergence $\mathrm{d}^k T^J / \mathrm{d} s^k$ that contributes to γ. Then consistency requires $a_J q^J \geq 0$ for all valid a_J. Since eq. (II.5.6a) requires $a_J \geq 0$ in the limit $s \to 4$, this is satisfied if $q^J \geq 0$. At both NLO and NNLO for any number of flavours n, this turns out to be true for the $s \to 4$ divergence (this was already noted in [13] for $n = 2$). Also, the divergences of α_i are of equal or lower powers than those of γ.

The same divergence structure appears in the $u \to 4$ limit, but here the coefficients of the leading divergences are not necessarily positive. However, we may use crossing symmetry to rewrite

$$\frac{\mathrm{d}^k}{\mathrm{d} s^k} T^I(s,t) = (-1)^k \frac{\mathrm{d}^k}{\mathrm{d} u^k} C_u^{IJ} T^J(u,t). \tag{II.5.17}$$

Here, we can simply relabel u as s. The coefficient of the leading divergence is here $(-1)^k C_u^{IJ} q^J$, and since eq. (II.5.6a) requires $(-1)^k a_I C_u^{IJ} \geq 0$, in the limit $u \to 4$, the fact that $q^J \geq 0$ in the $s \to 4$ limit guarantees that there are no catastrophic divergences in the $u \to 4$ limit either.

Since $\mathrm{d}\delta/\mathrm{d}s = -1/8\delta$, taking another derivative does not change the sign of q^J. Therefore, if no catastrophic divergences appear at the first k where γ diverges, they will not appear at larger k either.[29]

5.6 Integrals above threshold

The right-hand side of eq. (II.5.2) is, in its standard application, a non-perturbative quantity, about which the only knowledge we have is the fact that it is positive. However, ChPT is a low-energy theory, so its amplitude at any order should be an excellent approximation of the true amplitude for energies sufficiently close to threshold. Taking inspiration from the approach used in [19], we may therefore explicitly evaluate the lowest part of the integral on the right-hand side of eq. (II.5.2) and subtract it from both sides. Specifically, we

---

[29]The first divergence happens at $k = 1$ for $J \in \{I, S, SS, AA\}$ and at $k = 2$ for $J = A$, regardless of n. There is no divergence for $J = AS$, since $T^{AS}(s,t,u)$ does not contain $\bar{J}(s)$ or $k_i(s)$.

define

$$D_k^J(\lambda, v, t) = \frac{k!}{\pi} \int_4^\lambda \frac{\mathrm{d}z}{(z-v)^{k+1}} \operatorname{Im} T^J(z + i\varepsilon, t) \tag{II.5.18}$$

and modify eq. (II.5.2) to

$$a_J \left[\frac{\mathrm{d}^k}{\mathrm{d}s^k} T^J(s, t) - D_k^J(\lambda, s, t) - (-1)^k C^{JI} D_k^I(\lambda, u, t) \right] = \frac{k!}{\pi} \int_\lambda^\infty \mathrm{d}z \, [\dots] \operatorname{Im} T^J(z + i\varepsilon, t). \tag{II.5.19}$$

This has two benefits:

i) When s, t, u and a_J satisfy the conditions of eq. (II.5.6), both D_k^J and the right-hand side are positive, so we obtain a stronger positivity bound.

ii) The right-hand side of eq. (II.5.19) is positive under a wider range of conditions than that of eq. (II.5.2), so we obtain more positivity bounds. (This is because the constraint $\boldsymbol{a} \cdot \boldsymbol{\alpha}^J(4) \geq 0$ is replaced by the weaker $\boldsymbol{a} \cdot \boldsymbol{\alpha}^J(\lambda) \geq 0$, recalling the notation and discussion in section 5.1). Some of the new bounds are weakened by D_k^J being negative, but they may still contribute.

The size of λ presents a tradeoff: larger values amplify the benefits of using it, but also decrease the accuracy of the relation as the fixed-order ChPT amplitude strays from the exact result. The integral also requires some mathematical machinery; D_k^J is by no means a simple function, but we determine it up to NNLO in appendix C.3 (of course, it could also have been done numerically). By evaluating the NNLO corrections, we obtain a good idea of the accuracy of the NLO result.

An upper bound on λ is provided by [35], which determines the breakdown scale of n-flavour ChPT to be $s \sim (4\pi F)^2/n$. Using the value $F = 92.2(1)$ MeV adopted by [6] along with $M_{\mathrm{phys}} = M_\pi$, this places the breakdown at $\lambda \approx 35$ for $n = 2$ and at $\lambda \approx 25$ for $n = 3$. Thus, we cannot expect sensible results for λ anywhere close to this, and certainly not above it.[30]

6 Results

Here, we present the constraints obtained using the methods described in the preceding sections.

Following section 4, we will use the letter Ω to denote each collection of constraints, and $\mathcal{B}(\Omega)$ to denote the sets of parameter-space points that satisfy these. We will compare each $\mathcal{B}(\Omega)$ to a reference point, taken as the central value of the LEC estimates in [6]. These values can also be found in table II.2 in appendix A.

---

[30]Note that $(2M_K)^2 \approx 14$ and $(2M_\eta) \approx 16$ are already quite close to the breakdown scale. This offers some motivation as to why the equal-mass approximation is reasonable also at 3 flavours: with unequal, real-world masses, 3-flavour ChPT operates close to the limits of its range of validity even with nonrelativistic particles, which offsets the accuracy gained by increasing the realism of the model.

In all but the simplest cases, parameter space has too many dimensions to be visualised as a whole. Therefore, we will show lower-dimensional slices, with all omitted parameters set to their reference values. We will primarily show two-dimensional slices, since they are the easiest to understand, although some three-dimensional slices will be needed as well. It is not practical to show an exhaustive set of slices, so we will focus on grouping parameters that are, in some loose sense, related.

As a visual aid and a rough indicator of constraint strength, we define the quantity $\rho(\langle \boldsymbol{\alpha}, c \rangle)$ to be the shortest distance between the reference point and the hyperplane $\boldsymbol{\alpha} \cdot \boldsymbol{b} = c$.[31] Alternatively, we may use $\hat{\rho}(\langle \boldsymbol{\alpha}, c \rangle)$, which is the analogous distance if the space where all parameters are rescaled so that their reference values are 1.

6.1 Two flavours

Two-flavour ChPT constraints are in many regards quite simple: there are only 2 parameters at NLO and 7 at NNLO; there is only one reasonable choice for M_{phys}, namely M_π; and as shown in section 5.2, we do not have to consider odd numbers of derivatives. The NLO constraints have been extensively studied in e.g. [13, 19, 20], whereas the NNLO constraints are novel to this work.

Figure 11.6 shows constraints obtained using 2 derivatives and various amounts of above-threshold integration. The non-integrated ($\lambda = 4$) constraints are slightly stronger than those in [13], which only considered eigenstate a_J rather than the full space, but the constraints do not come close to the experimental uncertainty of the reference point without using λ that are far too large for the results to be trusted (recall that perturbative breakdown is expected at $\lambda \approx 35$). In [19, 20], comparable bounds are obtained with slightly less extreme λ, but in both cases, λ needs to be rather large to start cutting into the experimental uncertainty. The large discrepancy between the NLO and NNLO versions of the constraints indicate that the bounds are highly unreliable for all but the smallest λ used. Even with $\lambda = 4$, the difference is quite significant.

Figure 11.7 shows similar NNLO constraints over a more conservative λ range; however, here we display the effects of higher even derivatives (recall again that odd derivatives need not be considered with 2 flavours). Unlike the ones shown in fig. 11.6, these constraints impose *upper* bounds on the LECs as well as lower bounds. Note that in the upper-left part of the plot, the two-derivative bounds are less restrictive than their NLO counterpart. This can partly be seen as an artefact of introducing multiple new parameters and fixing them to imprecise experimental values (for instance, fig. 11.9 shows that smaller values of $\bar{l}_4$ strengthen the bounds on $\bar{l}_1, \bar{l}_2$), but one must keep in mind that switching to a more refined theory can both strengthen and relax the predictions.

Figure 11.8 shows similar bounds on $\bar{l}_3$ and $\bar{l}_4$, which are the only NLO LECs other than $\bar{l}_1, \bar{l}_2$ that appear in the NNLO amplitude. The bounds on $\bar{l}_3$ are extremely weak, since $\bar{l}_3$ figures in the amplitude with much smaller prefactors than the other LECs. It can be

---

[31]Note that this refers to distance in the full parameter space, which does not directly correspond to distance in the subspaces shown in the figures.
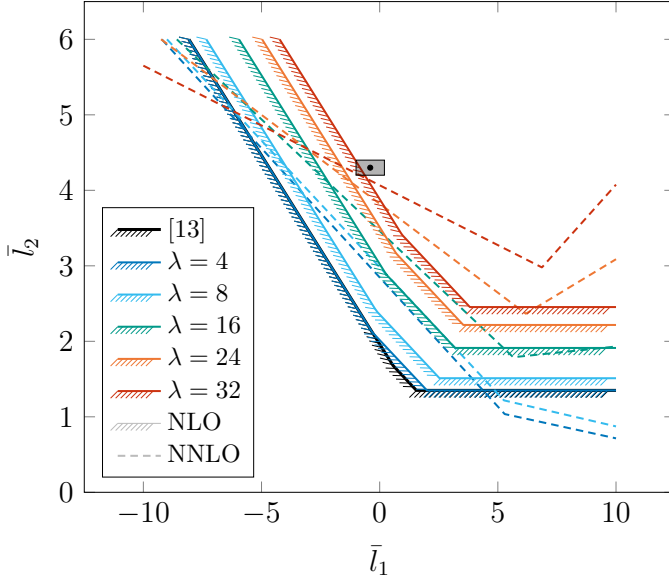
Figure 11.6: Two-flavour two-derivative constraints on $\bar{l}_1, \bar{l}_2$ for various λ, as indicated in the legend. The presentation is similar to fig. 11.1: each version of $\mathcal{B}(\Omega)$ is outlined, with the hatched side indicating the points excluded by the constraints. The bounds from [13] are also drawn. The reference point $\bar{l}_1 = -0.4(6)$, $\bar{l}_2 = 4.3(1)$ is drawn as a dot with an uncertainty region around it. For each set of constraints, the direct NNLO counterpart (i.e. using the same λ, s, t and a_J) is drawn as a dashed outline.

partly understood by noting that the $\bar{l}_3$ term in the Lagrangian, unlike the other terms, does not contain the field u_μ. Interestingly, the upper bounds on $\bar{l}_4$ become *weaker* as λ is increased. This does not necessarily contradict the arguments made in section 5.6, due to the complicated NNLO situation where both α and c in a constraint may depend on λ. Nevertheless, it is surprising to see, and does not seem to appear in other bounds, such as those on $\bar{l}_{1,2}$.

Figure 11.9 combines the $\bar{l}_4$ bounds with the $\bar{l}_1, \bar{l}_2$ bounds to form a summary of the effective bounds on the two-flavour NLO LECs obtained in this paper. Unfortunately, only the $k = 2, \lambda \approx 4$ constraints have a shape that is sensible to show in three dimensions.

Lastly, fig. 11.10 shows bounds on the four NNLO parameters Θ_i, using only two derivatives. Here, the bounds are not entirely consistent with the reference point, although not too much meaning should be read into this, as the reference values for the NNLO parameters are little more than educated guesses. It also showcases the phenomenon where some constraints become weaker at larger λ, at least at the particular values at which we have fixed the NLO LECs.

The NNLO parameter bounds are not particularly strong compared to the magnitude of
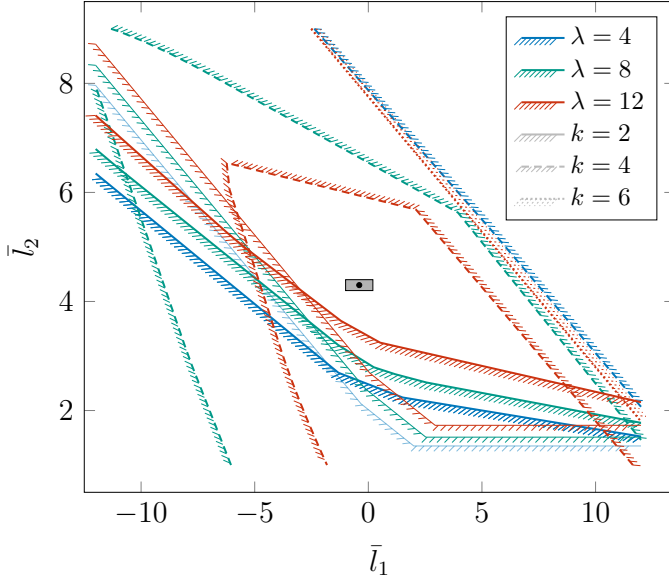
Figure II.7: NNLO bounds on $\bar{l}_1, \bar{l}_2$ with $k = 2, 4, 6$ derivatives, displayed similarly to fig. II.6 but over a slightly wider part of parameter space. For comparison to fig. II.6, the corresponding NLO bounds are drawn with weak solid lines. The six-derivative bounds are nearly independent of λ, so all versions are not drawn.
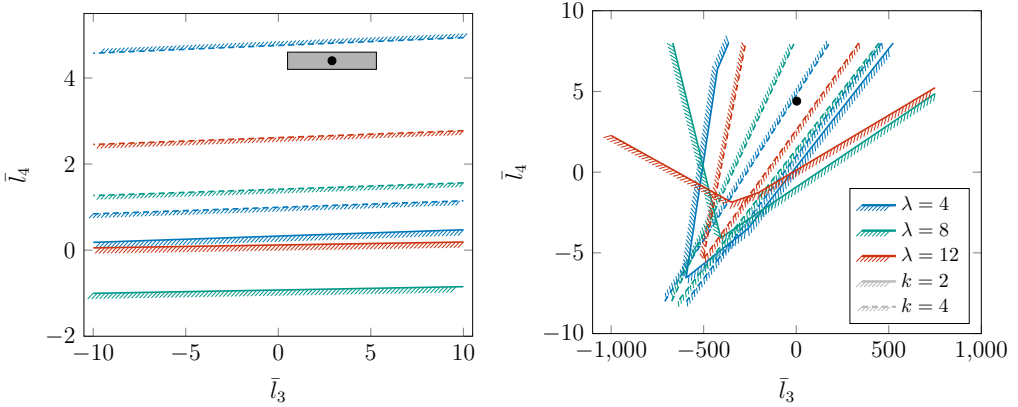
Figure II.8: NNLO k-derivative constraints on $\bar{l}_3, \bar{l}_4$, displayed like in fig. II.7, using similarly scaled axes (left) or an extreme scale on the $\bar{l}_3$ axis (right) to show the full constraint structure. In all cases, the bounds are satisfied by the reference point. The six-derivative bounds (which like in fig. II.6 are nearly λ-independent) exactly overlap with the $\lambda = 4$ two-derivative bounds and are not drawn.
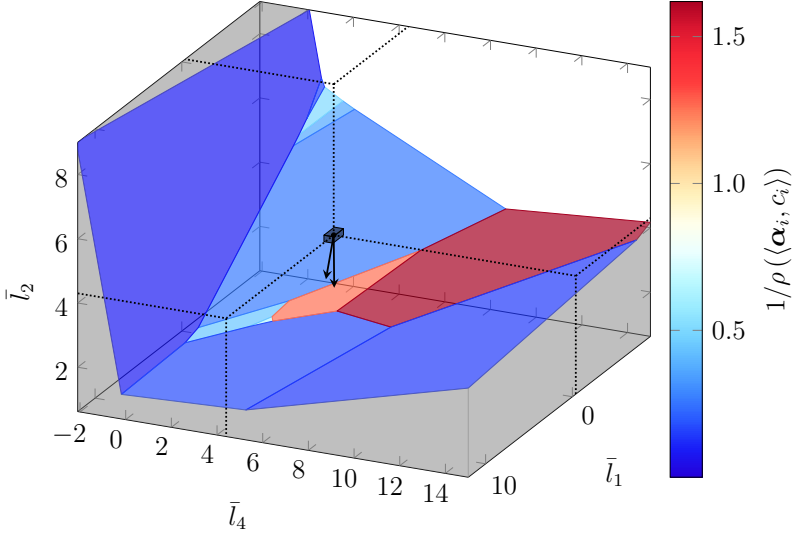
Figure II.9: NNLO two-derivative constraints on $\bar{l}_1, \bar{l}_2$ and $\bar{l}_4$ with $\lambda = 4$; a cross-section at $\bar{l}_4 = 4.4$ would yield part of fig. II.6. The constraint surfaces are coloured according to their proximity $1/\rho\left(\langle \boldsymbol{\alpha}_i, c_i \rangle\right)$ to the reference point, which is drawn similar to the 2D plots. The "open space" bounded by the constraint surfaces is part of $\mathcal{B}(\Omega)$; the grayed-out region is excluded by the constraints. To clarify its spatial position, the reference point is connected to the boundaries of the plotted region with black dotted lines parallel to the coordinate axes, whenever doing so is possible without intersecting a constraint surface. It is similarly connected to neighbouring surfaces with black arrows. These are orthogonal to the respective surfaces, even though the different scales of the axes makes it not appear so.

the reference values, but some are still notable. Θ_4 has a strict lower bound, with no values of $\Theta_{1,2,3}$ being permitted if $\Theta_4 \lesssim -2.5$. Θ_3 is bounded from both above and below, and the bounds are fairly independent of the values of the other parameters in a large part of parameter space. Thus, we may write down the tentative single-parameter bounds

$$-1.5 \lesssim 10^3 \Theta_3 \lesssim 1, \qquad -2.5 \lesssim 10^3 \Theta_4, \qquad (\text{II.6.1})$$

both of which are satisfied, with a margin of about an order of magnitude, by the reference values in appendix A.
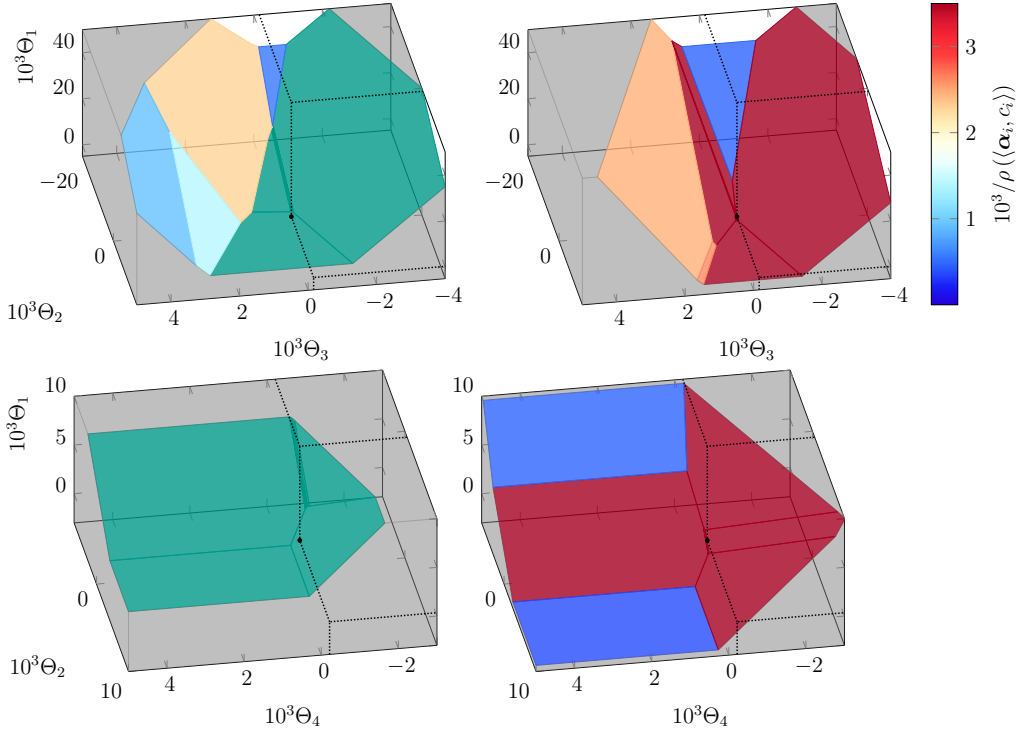
Figure 11.10: NNLO two-derivative bounds on Θ_i at $\lambda = 4$ (**left**) and $\lambda = 8$ (**right**) for $i = 1, 2, 3$ (**top**) and $1, 2, 4$ (**bottom**, note different $\Theta_{1,2}$-axes). These specific i choices are used, for if either Θ_1 or Θ_2 is fixed to its reference value, the remaining constraint is not satisfied by any value of the other Θ_i. Note the rather different scales on the axes; several facets in the top plots (especially for $\lambda = 8$) are in fact nearly perpendicular to the Θ_3 axis, placing upper and lower bounds on Θ_3 that do not depend very strongly on the other parameters. The green facets exclude the reference point. The exclusion is by a rather small amount (no orthogonal arrow can be sensibly drawn), and varying the L_i^r within their uncertainties is sufficient to remedy this. Note how increasing λ to 8 slightly weakens these constraints so that they no longer exclude the reference point.
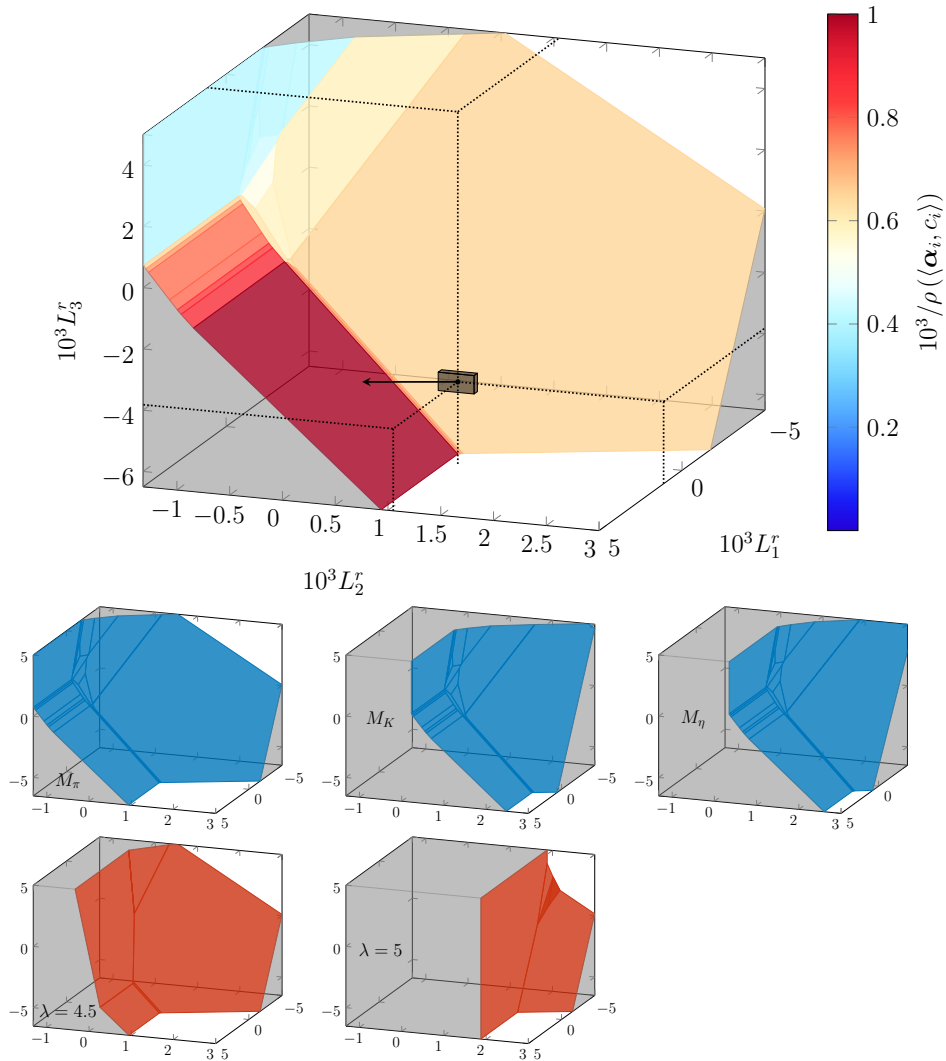
Figure 11.11: Two-derivative constraints on the three-flavour NLO LECs L_i^r, visualised similarly to previous figures. **Top:** detailed constraints with $M_{\text{phys}} = M_\pi$, $\lambda = 4$. The visualisation is similar to fig. 11.9; note the different scales on the axes. **Middle:** Constraints with (from left to right) $M_{\text{phys}} = M_\pi$, M_K, and M_η, all with $\lambda = 4$. Note that the leftmost figure is just a less detailed version of the top figure, and that the M_K and M_η surfaces only differ from each other by a very small amount. Both these higher-mass constraints exclude the reference point (not drawn), although not by much. **Bottom:** the effect of setting $\lambda = 4.5$ (left) or 5 (right) with $M_{\text{phys}} = M_\pi$, for comparison with the $\lambda = 4$ figure above. The $\lambda = 5$ constraint excludes the reference point (not drawn) rather severely.

174

6.2 Three flavours

Three-flavour ChPT bounds cover 3 parameters at NLO and 12 at NNLO, and three choices for M_{phys} (namely M_π, M_K and M_η) present themselves, with no *a priori* indication of which to choose. This would of course be resolved by working with inequal-mass mesons, but this NLO amplitude is far more complicated (see [42] and references therein) and its NNLO counterpart is so far undetermined in a simple analytic fashion; furthermore, inequal masses have implications for the construction of bounds that we do not address here (see [14]).

Figure 11.11 shows the bounds on L_1^r, L_2^r and L_3^r obtainable at NLO. With $M_{\text{phys}} = M_\pi$, the bounds are consistent with the reference point and qualitatively similar to the two-flavour bounds on $\bar{l}_1, \bar{l}_2$. Interestingly, the reference point does not satisfy the bounds at the other choices of M_{phys}, although only barely — the smallest distance between the reference point and $\mathcal{B}(\Omega)$ is $0.5 \cdot 10^{-3}$ for $M_{\text{phys}} = M_K$ and $0.6 \cdot 10^{-3}$ for $M_{\text{phys}} = M_\eta$, which is smaller than the uncertainty in the experimental values (approximately $(1\text{-}3)\cdot 10^{-3}$). Therefore, using these M_{phys} does not imply any significant inconsistency.

It is worth noting that the three-flavour constraints are much more sensitive to integration than their two-flavour counterparts. This can be partly understood by noting that integrals scale as roughly the square of the number of flavours (see fig. 11.22).

Figure 11.12 shows NNLO bounds on the same three LECs. Interestingly, the four-derivative constraints alone confine L_1^r, L_2^r, L_3^r to a bounded region, although it is very large in most directions. However, the bounds are reasonably strict between the two near-parallel faces shown in the figure, so we may write down another double-ended bound, similar to eq. (11.6.1):

$$-27 \lesssim 4.9L_1^r + 2.8L_2^r + 2.4L_3^r \lesssim 10 \qquad \text{for} \quad M_{\text{phys}} = M_\pi, \qquad (11.6.2)$$

where the linear combination of the LECs is chosen to be roughly orthogonal to the bounding faces.

Figure 11.13 shows NNLO bounds on the remaining L_i^r. These are rather weak, even though L_6^r obtains a double-ended bound. The weakness is understandable for similar reasons as the weakness of the $\bar{l}_3$ bounds. With $M_{\text{phys}} = M_K, M_\eta$, the bounds are inconsistent, not only with the reference values for $L_{4,5,6,8}^r$, but with *all* values of these parameters, unless $L_{1,2,3}^r$ are removed from their experimental values by an amount roughly one order of magnitude larger than their stated uncertainty (the NNLO parameters have little effect). This practically renders this version of the theory self-inconsistent, except at experimentally unreasonable points in parameter space. Although the $M_{\text{phys}} = M_\pi$ version remains consistent, only a small amount of integration excludes the reference point, so its validity is quite dubious.

Lastly, figs. 11.14 and 11.15 show bounds on the NNLO parameters; for the same reason as above, we keep $M_{\text{phys}} = M_\pi$. The former shows some features reminiscent of its SU(2) analogue, fig. 11.10, although there is not a clear double-ended bound on either parameter. The latter is more interesting, since all three parameters are confined to a small bounded
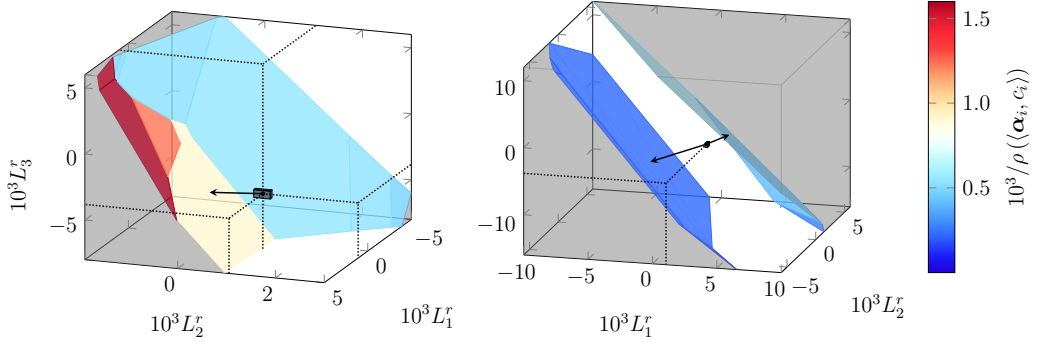
Figure 11.12: NNLO bounds on L_1^r, L_2^r and L_3^r with $M_{\mathrm{phys}} = M_\pi$ and $\lambda = 4$, using two (left) and four (right) derivatives; three derivatives gives very weak bounds. Thus, the left figure is essentially the NNLO version of fig. 11.11. Its response to larger λ or M_{phys} (not shown) is qualitatively similar to that exhibited at NLO. In the four-derivative case, $\mathcal{B}(\Omega)$ is actually a bounded region. It is a lentil-shaped body whose largest dimension is about two orders of magnitude larger than the region shown in the figure. Note that the axes have been rotated relative to the left figure in order to make the inside of $\mathcal{B}(\Omega)$ reasonably visible.
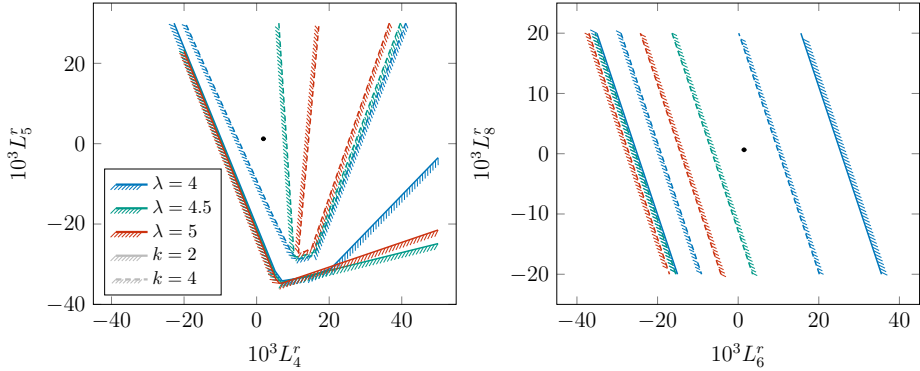
Figure 11.13: NNLO bounds on L_4^r, L_5^r (left) and L_6^r, L_8^r (right), visualised similarly to fig. 11.6. The uncertainty region around the reference point is hardly visible at this scale. Note how the four-derivative bounds exclude the reference point also for rather small $\lambda > 4$. However, the integrated two-derivative bounds are weaker and are not visible at all in the right figure. The corresponding bounds for $M_{\mathrm{phys}} = M_K, M_\eta$ are entirely inconsistent with the experimental values.

region and quite significantly excludes the reference point. The bounds on Γ_3, Δ_3 in particular are fairly independent of each other and Ξ_4, leading to the single-parameter bounds

$$0.08 \lesssim 10^3 \Gamma_3 \lesssim 0.34, \qquad -0.8 \lesssim 10^3 \Delta_3 \lesssim 0.25 \qquad \text{for} \quad M_{\mathrm{phys}} = M_\pi. \qquad (11.6.3)$$
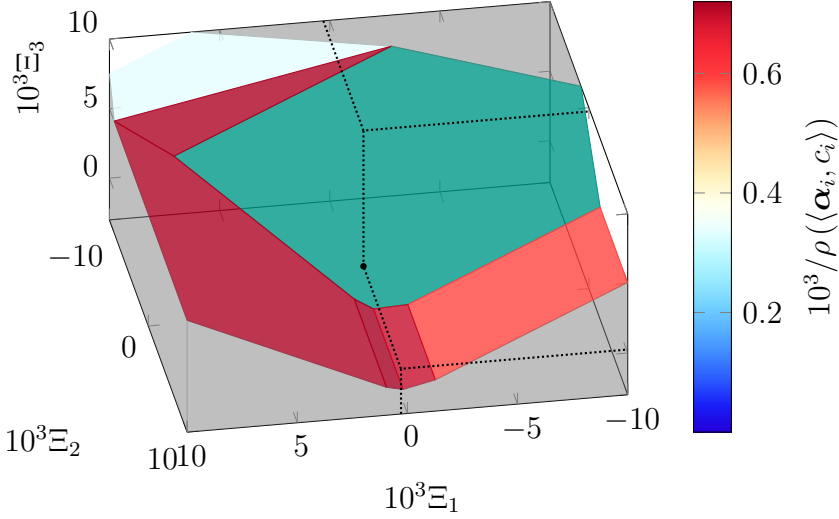
Figure 11.14: NNLO two-derivative bounds on Ξ_1, Ξ_2, Ξ_3 (those NNLO parameters whose contribution is independent of t) with $\lambda = 4$ and $M_{\text{phys}} = M_\pi$. Higher-derivative bounds are extremely weak and are not shown. The facets are coloured based on their distance to the reference point in the space of NNLO parameters rather than the full parameter space. Like in fig. 11.10, the **green** facet excludes the reference point, although by an extremely small amount, well within the drawn size of the reference point.
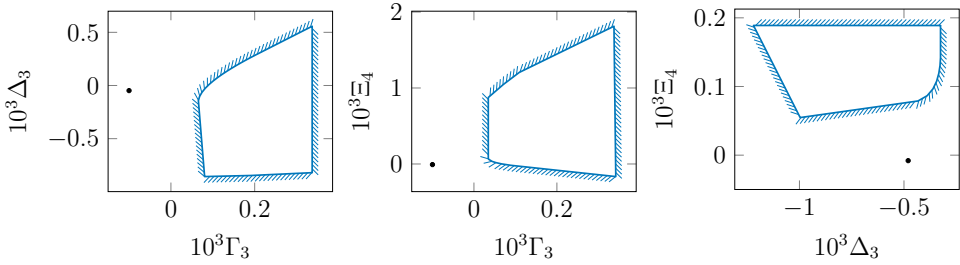
Figure 11.15: NNLO two-derivative bounds on Γ_3, Δ_3 and Ξ_4 (those NNLO parameters that obtain additional bounds at $t \neq 4$) with $\lambda = 4$ and $M_{\text{phys}} = M_\pi$. Note how they are bounded from all directions and confined to a rather small volume; this makes three-dimensional representation difficult, so we use two-dimensional slices instead. Note how the bounds exclude the reference point. A major source of uncertainty in these bounds is the fixing of the other parameters: as the NLO LECs are varied within their uncertainties, the boundary lines shift in various directions, with the typical amount displacement being roughly 0.2. However, at no values are the bounds consistent with the reference values of Γ_3, Δ_3 and Ξ_4.

177
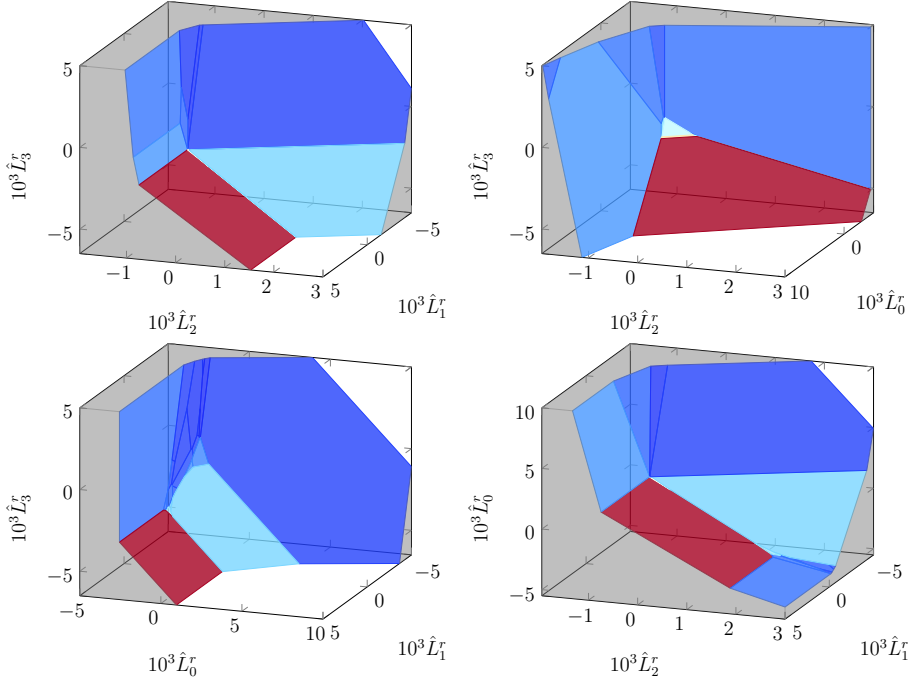
Figure II.16: Two-derivative constraints on the 4-flavour NLO LECs $L_0^r, \ldots, L_3^r$ with $M_{\text{phys}} = M_\pi$ and $\lambda = 4$. The four possible selections of three LECs are shown, with the top left figure (L_1^r, L_2^r, L_3^r) being almost exactly the 4-flavour equivalent of fig. II.11 (the LEC and colour ranges differ slightly). No reference point is shown, and the colour information should not be considered very meaningful, as discussed in the text; therefore, the colour bar has been omitted to save space.

6.3 Higher number of flavours

ChPT with more than three flavours are not of direct interest as low-energy QCD, since the large mass of the charm quark makes it entirely invalid as a model of mesons. An arbitrary number of flavours is useful when developing the methods, though, and is interesting in its own right in the context of EFT studies. Furthermore, ChPT has many uses other than QCD (for a review of some of these, see [43]). Besides various numbers of flavours, these commonly use different symmetry breaking patterns than $SU(n) \times SU(n) \to SU(n)$; some of these were treated in [34], so their amplitudes would be a drop-in replacement into our methods. However, that is beyond the scope of this paper, in which we are content to show some basic high-flavour results in "QCD-style" ChPT.
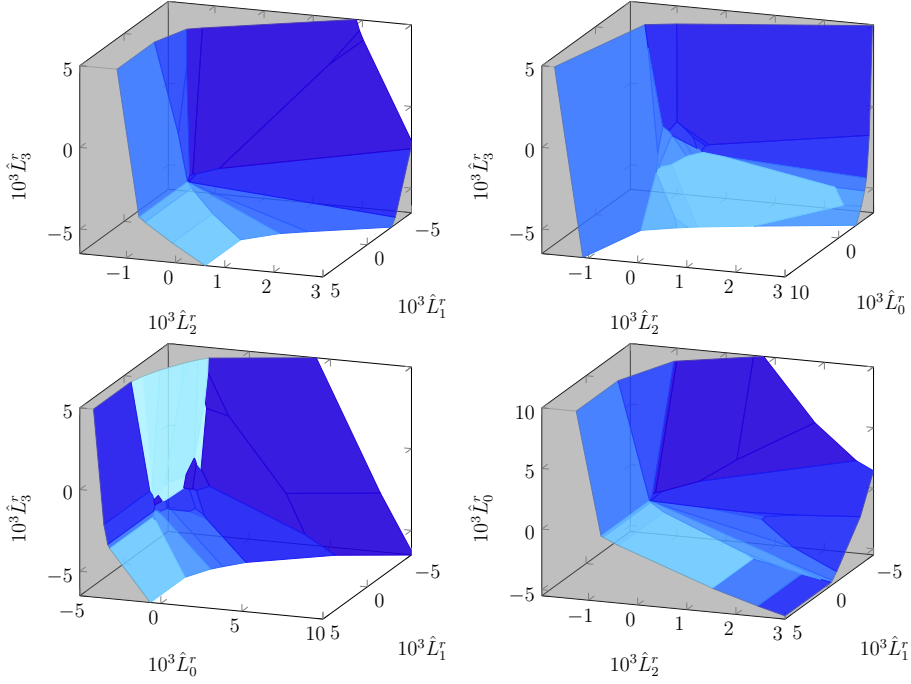
Figure 11.17: The exact 8-flavour equivalent of fig. 11.16.
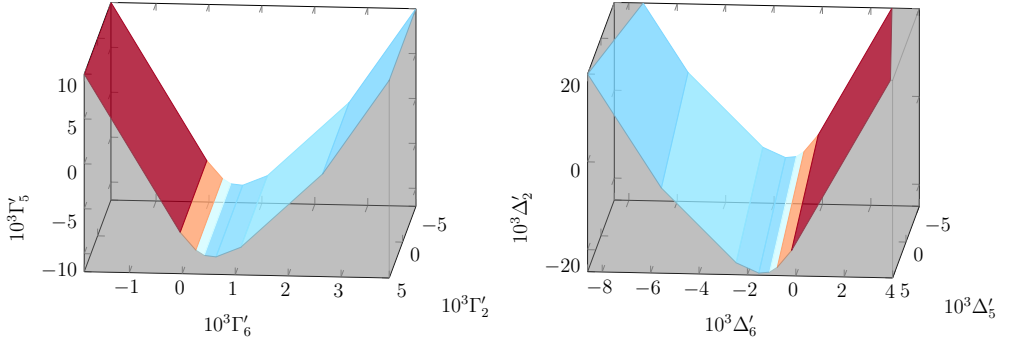
Figure 11.18: NNLO 4-flavour bounds on Γ_i' (left) and Δ_i' (right), analogous to fig. 11.10. Note the different scales on the axes; Γ_6' and Δ_6' are bounded by almost parallel faces.

It turns out that the bounds change quite gradually between different $n \geq 4$, to the extent that the difference between e.g. $n = 4$ and $n = 5$ is hardly visible upon first inspection of figures like those used here. Therefore, we have chosen to only display $n = 4$ (the "leading" high-flavour example) and $n = 8$ (a reasonable "very high flavour" example, also of historical technicolour relevance [44]). Understandably, there are no experimental reference values for the high-flavour LECs, so to perform two- or three-dimensional slices of parame-

179
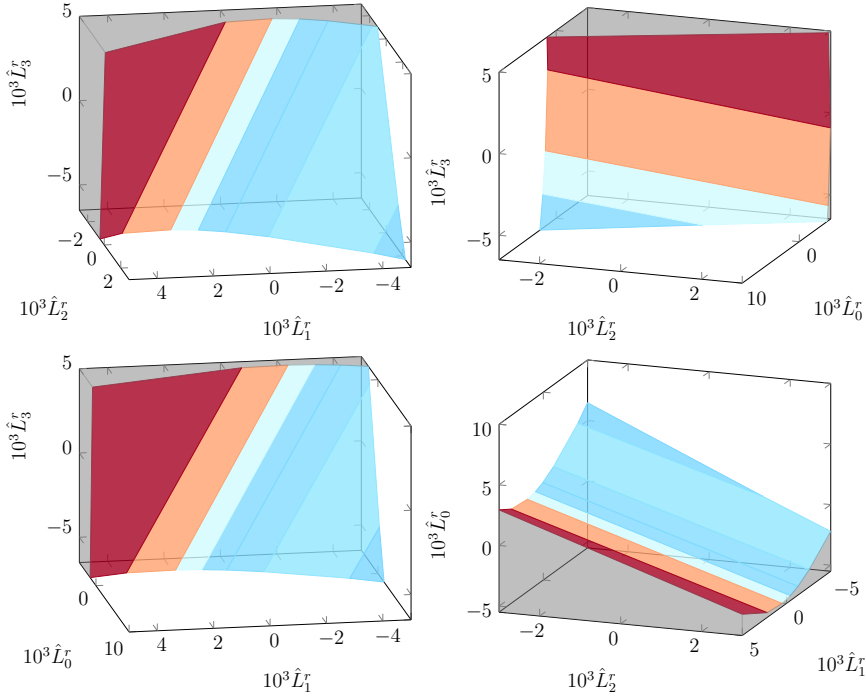
Figure 11.19: NNLO 4-flavour bounds on $L_0^r, \ldots, L_3^r$ presented in the same way as fig. 11.16, but with some axes rotated to give a better view. See fig. 11.12 for the 3-flavour counterpart (no analogous 4-derivative version has been produced due to issues with high-dimensional parameter space).

ter space, we provisionally set $\hat{L}_i^r = L_i^r$ and use the 3-flavour reference values when possible, and use 0 as the reference value for e.g. $\hat{L}_0^r$. We also retain the distance-to-reference-point colouring of the constraint surfaces to make them more visually distinguishable. Of course, not too much meaning should be read into these distances.

Figures 11.16 ($n = 4$) and 11.17 ($n = 8$) show the basic NLO bounds, similarly to fig. 11.11, to which the bounds are qualitatively similar. There is a trend towards weaker constraints as n increases, as can be seen from the amplitude: many of the most important terms go as powers of $1/n$ (cf. eq. (11.4.5); all constraints have $c = -1$). Note how we wholly abandon the M_{phys} debate and use M_π throughout.

Figures 11.18, 11.19 ($n = 4$) and 11.20, 11.21 ($n = 8$) show the most interesting NNLO bounds, which are again similar to their lower-flavour analogues. The bounds on $\hat{L}_4^r, \hat{L}_5^r, \hat{L}_6^r$ and $\hat{L}_8^r$, which are even weaker than the corresponding ones for $n = 3$, are not shown. Likewise, the bounds on Γ_4, Δ_4 are very weak and have been omitted. Note how the prominent lower bound on $\hat{L}_1^r$ in fig. 11.21 breaks the trend of weaker bounds at larger n.

Finally, one may ask what happens in the limit $n \to \infty$. Looking at the amplitudes in
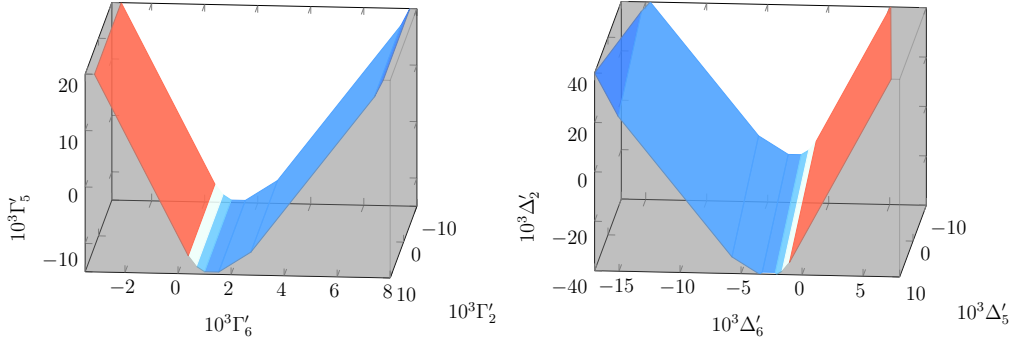
Figure 11.20: The 8-flavour equivalent of fig. 11.18. The axes have been zoomed out to adequately view the weaker constraints.
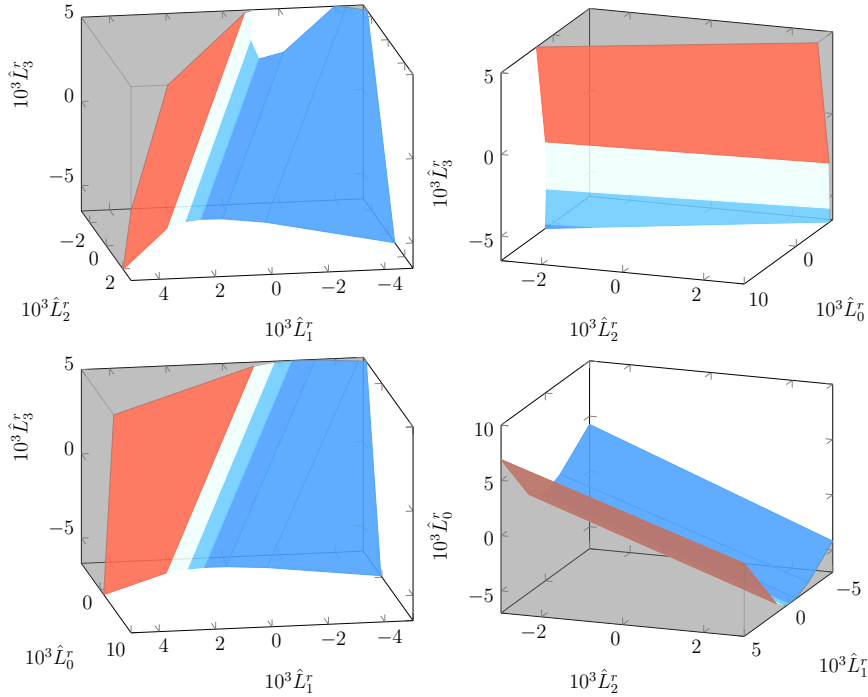
Figure 11.21: The exact 8-flavour equivalent of fig. 11.19.

[34], we see that they are independent of the LECs at leading order in n: at NLO, the amplitude is $\mathcal{O}(n)$ whereas the LEC parts are $\mathcal{O}(1)$, while at NNLO they are $\mathcal{O}(n^2)$ and $\mathcal{O}(n)$, respectively (this of course indicates convergence problems at high n, in agreement with [35]). Thus, the bounds, expressed in the schematic form of eq. (11.5.7), will eventually be dominated by γ, and will therefore asymptotically tend towards either $\langle \mathbf{0}, 1 \rangle$ or the trivial $\langle \mathbf{0}, -1 \rangle$; the gradual weakening in figs. 11.16 to 11.21 suggests the latter.
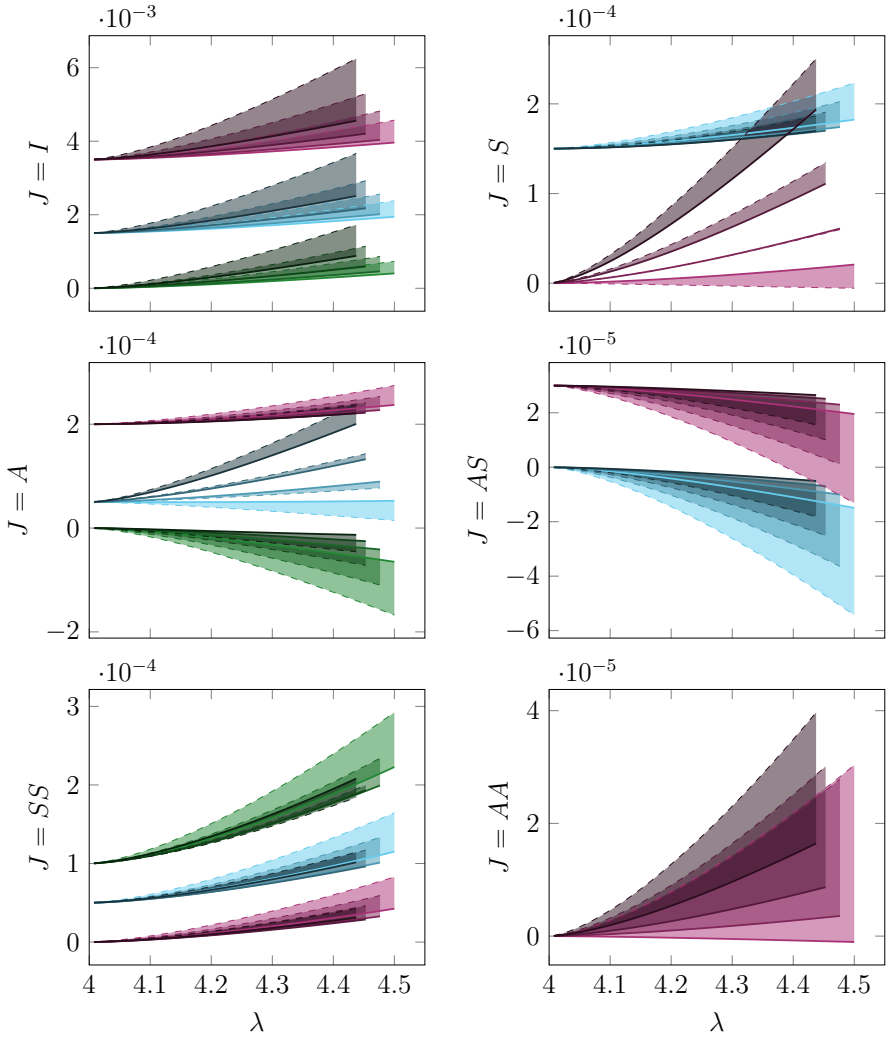
Figure II.22: The integrals $\frac{1}{n^2}\left[D_k^J(s,t;\lambda) + (-1)^k C^{JI} D_k^I(u,t;\lambda)\right]$, defined in eq. (II.5.19), as functions of λ between 4 and 4.5 at $k=2$, $t=4$, for the six possible J. The factor $1/n^2$ roughly cancels the n-dependence, making it possible to show all n at the same vertical scale. Each function is represented for $n=2$ (**green**), $n=3$ (**cyan**) and $n=4$ (**purple**) ChPT, with the latter representing the general n-flavour case reasonably well. Different values of s between 0 and 1.5 (the range of most relevant constraints) in increments of 0.5 are shown with the colour saturation (brighter colour = smaller s). The NLO result is drawn with solid lines, and the NNLO result (with the NLO LECs fixed to their reference values) is drawn with dashed lines. The difference between them is shaded, and the λ ranges are staggered slightly to improve readability. The graphs have been vertically shifted to enhance readability; all integrals are equal to 0 at $\lambda=4$. Note that the vertical scale differs greatly between different J.
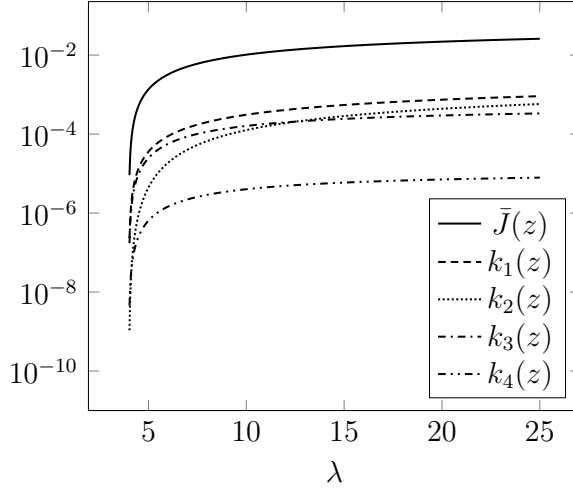
Figure 11.23: Logarithmic plot of the relative size of the different loop integral functions, covering a rather wide range of λ. Each line represents $\left|\int_4^\lambda X(z)\mathrm{d}z\right|$ for the given function X. Note how $\bar{J} \gg k_1 \approx k_3 \gg k_2 \approx k_4$ for small λ, which changes to $\bar{J} \gg k_1 \approx k_2 \approx k_3 \gg k_4$ for large λ. In either case, "$\gg$" is by about two orders of magnitude, and "$\approx$" is within one order of magnitude.

6.4 Considerations about the integrals

Figure 11.22 demonstrates the integrals of the components of the amplitude over relevant s, λ ranges. We may note that despite the great complexity seen in appendix C.3, the graphs are typically quite simple and qualitatively similar. There is, however, a very wide range of magnitudes; typically, "lower" J components have larger integrals.

It is interesting to note that the ratio between NLO and NNLO integrals is approximately constant in λ, varying only with s and J. This is perhaps unexpected since the NNLO integrand contains terms like $z^3 X(z)$, where X is one of $\bar{J}$ or k_i, whereas the NLO amplitude only contains $z^2 X(z)$. Therefore, we would expect the ratio to grow approximately linearly with λ. However, the $z^3 \bar{J}(z)$ terms are typically suppressed by small numerical coefficients or NLO LECs, and $\bar{J}(z)$ dominates $k_i(z)$, as shown in fig. 11.23. Therefore, this effect does not manifest until λ is much larger than the values relevant to this application, which in practice limits the importance of higher-order corrections to the bounds.

6.5 Considerations about a_J

We have made two innovations in the treatment of a_J:

(i) not restricting a_J to the physical eigenstates (this was done already in [21]),

(ii) employing the fixed-s (as opposed to all-s) constraints eq. (11.5.6b) on a_J.

This section investigates whether these changes actually give any improvements at all — it would be conceivable that the physical eigenstates, which are allowed for all s, were special in a way that guarantees that they generate the strongest bounds.

To measure how significantly (i) is used, we consider a_J as a point on the unit sphere and find the angle $\theta(a_J)$ between it and the closest eigenstate point. Thus, larger θ indicates, in a sense, more use of (i). Similarly, we may measure (ii) via the fraction of points on $s \in [-4, 4]$ for which eq. (11.5.6b) permits a_J.

In terms of the *ad hoc* measures together with $\hat{\rho}$ defined above, fig. 11.24 shows the distribution of NLO bounds for 2, 3 and 4 flavours, and fig. 11.25 shows a geometrically more intuitive version in the 2-flavour case.[32] We see that relevant constraints tend to have low $\hat{\rho}$, albeit with many exceptions — the orientation of the constraint is another important factor. The relevant constraints are rather evenly distributed over the θ range permitted by eq. (11.5.6b), indicating that there is, in this regard, nothing special about the eigenstates, validating (i). Most relevant constraints also occur at a_J that are permitted for very few s (i.e. coloured very pale in the plots), validating (ii).

There are, of course, severe limitations to the analysis in this section. Apart from the roughness of the chosen measures discussed above, it is difficult to assess just how great the benefits of (i) and (ii) are. Stronger constraints are obtained, but not necessarily much stronger: the improvement in fig. 11.6 over [13] is very slight, although it seems that the use of (i) is limited at $n = 2$ but more extensive at e.g. $n = 3$ by comparing the subfigures of fig. 11.24.

7   Conclusions and outlook

To recapitulate, our method has been as follows: We scan over the s, t range depicted in fig. 11.4 (or in many cases just the s range, with t fixed to 4, as discussed in section 5.3). At each (s, t)-value, we scan over those a_J that are permitted by eq. (11.5.6b), and compute the kth derivative of the n-flavour amplitude to either NLO or NNLO at that point, possibly with above-threshold integration up to λ, as described in section 5.6. Through eq. (11.5.6a), this yields positivity bounds on the LECs (or at NNLO, the parameters derived in section 5.4), which can be handled as linear constraints using the language and methods of section 4 (implemented as discussed in appendix B). In the end, this yields a manageable set of relevant constraints, which can be visualised and interpreted.

Our results at NLO consist of stronger bounds than in [13, 14], whereas the comparison to the more recent works [19, 20] is less clear-cut. As in previous works, most bounds consist of irregularly shaped and usually infinite regions, although some of the nicer cases allow for more clear-cut bounds such as eqs. (11.6.1) to (11.6.3). Although the bounds themselves are highly uncertain, basic compatibility suggests that one may assign uncertainties of roughly

---

[32]The points were sampled uniformly over the unit octahedron (i.e. the unit sphere under the 1-norm $\sum_J |a_J| = 1$) and its higher-dimensional analogues in a_J space. This shape was used rather than the unit sphere to preserve the linearity of bounds on a_J; compare the discussion in section 5.1.
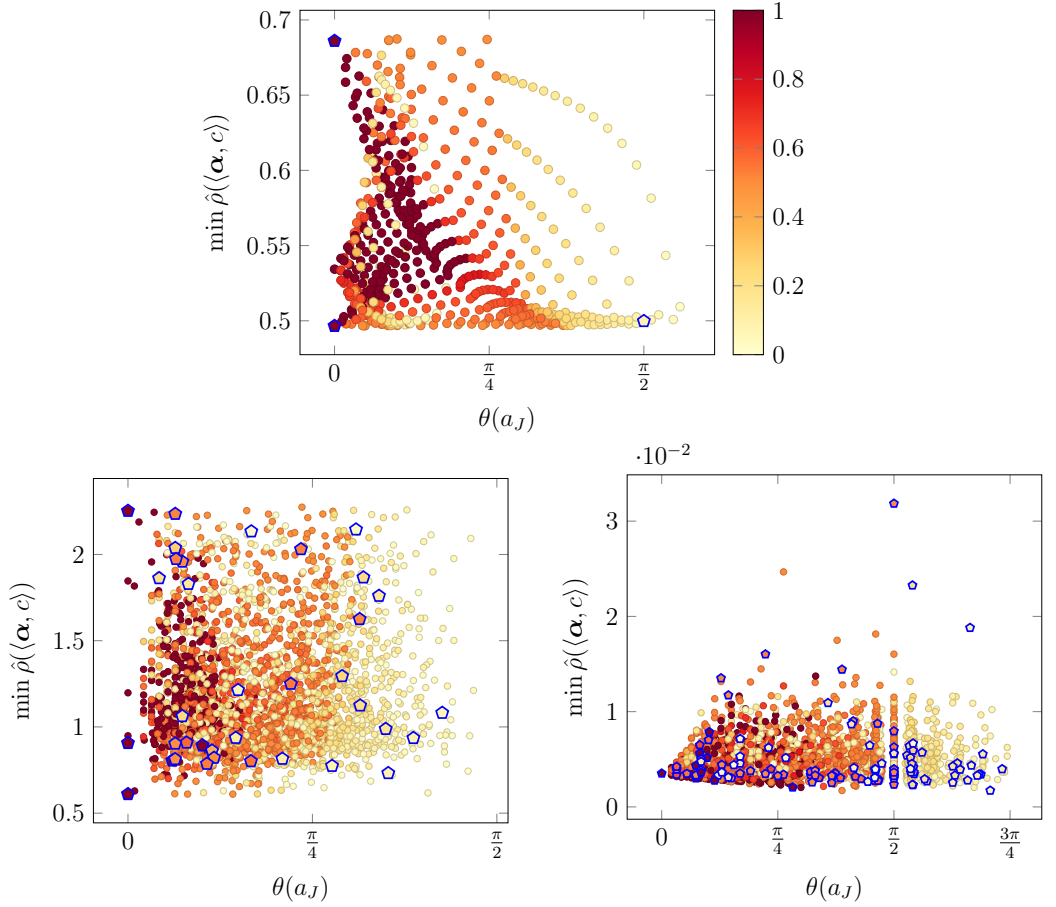
Figure 11.24: A discrete sample of the full set of a_J permitted by eq. (11.5.6b), plotted over θ as defined in section 6.5. For each a_J, the closest distance $\hat{\rho}$ between the reference point and any constraint generated at that a_J is indicated with a dot. Relevant constraints are outlined with a blue pentagon. The colour of the dots indicate the fraction of the full range $s \in [-4, 4]$ in which eq. (11.5.6b) permits that a_J. **Top:** Two flavours. The eigenstates, i.e. the points at $\theta = 0$, are, from top to bottom, $\pi^0\pi^+$, $\pi^0\pi^0$ and $\pi^+\pi^+$. **Lower left:** Three flavours. The eigenstates are, from top to bottom, $\pi^0\pi^+$, $K\eta$, $\pi^+\pi^+$, $K\pi^0$, $K^\pm\pi^\pm/K^0\pi^\pm/\pi^0\pi^0$ (overlapping) and $\eta\pi$. **Lower right:** Four flavours (representing $n > 3$ flavours). Since high-flavour ChPT is not directly applicable to meson physics, we do not use the full set of "physical" eigenstates, but for θ to be defined, we retain a single eigenstate: $\pi^+\pi^+$, whose decomposition (consisting of the SS component only) is uniquely flavour-independent.
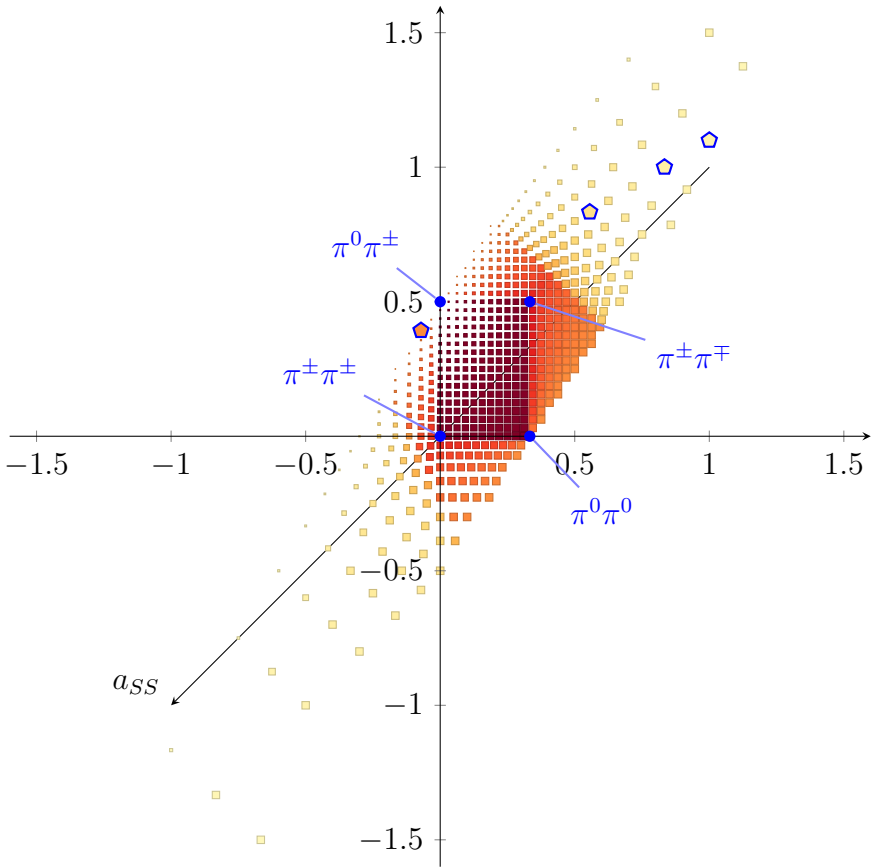
Figure II.25: The same 2-flavour data as in fig. II.24, but plotted over a_I, a_A plane with $a_I + a_A + a_{SS} = 1$, like in fig. II.5, rather than as a function of θ. The colours convey the same meaning as in fig. II.24, whereas the $\hat{\rho}$ values are indicated by the size of the points: smaller $\hat{\rho}$ (higher proximity to the reference point) corresponds to larger points. Note how the a_J that are permitted at all s can be clearly seen as a uniformly coloured patch, which is also featured in fig. II.5. It is interesting to note that most relevant constraints come from the $a_{SS} < 0$ region (upper right), which is permitted for significantly fewer s than those points where all a_J are positive.

200-1000% to $\Theta_i, \Gamma_i, \Delta_i$ and therefore also to the NNLO LECs, which are not given any error estimates in [6]. Obtaining better error estimates is a possible direction for future work.

The employment of above-threshold integration allows for very strong bounds, but carries the risk of going too far beyond the low-energy limit; the difference between our NLO and NNLO bounds, such as in fig. 11.6, does not inspire much confidence in integrated bounds for λ significantly larger than 4. Our improved handling of a_J, which is evaluated in section 6.5, improves bounds without additional assumptions ([19, 20] also uses assumption-less improvements). Of course, the choice of fixed-order ChPT is itself something that relies on the low-energy limit, although it is easier to motivate than a particular choice of $\lambda > 4$ is.

At NNLO, our methods suffer some practical problems due to the very high dimension of the parameter space, so we have only performed rather coarse scans of the available s, t and a_J ranges. This is remedied by fixing some parameters and focusing on the lower-dimensional space that remains. However, it is important to keep in mind that, even though we may fail to obtain many constraints due to limited scans and technical issues with determining convex hulls (see appendix B.4), it is guaranteed that the constraints we do find are true — the method automatically errs on the side of caution, so to speak. In particular, this means that if the bounds are inconsistent with the experimental values, then the error must lie either with the values, or with the theory itself. In general, the main problem with our method is not its ability to produce bounds, but our ability to rely on the assumptions behind them.

The most prominent example of this is the problems encountered at three flavours, as discussed in section 6.2. The obvious remedy is to replace equal-mass ChPT with the more realistic unequal-mass version, towards which the main hurdle is some so-far-unsolved two-loop integrals. The emergence of distinct mass eigenstates, as well as some other features discussed in [14], could possibly interfere with some of our method innovations, although we are confident that they can be remedied.

Beyond practical improvements and the use of unequal-mass ChPT, a possible step forward is to either go beyond $2 \rightarrow 2$ scattering, or to study NNNLO. The former would allow for bounds on LECs that do not appear in the $2 \rightarrow 2$ amplitude, as well as possibly new bounds on those covered here. While it would be infeasible to manage all the NNNLO LECs in the $2 \rightarrow 2$ amplitude, or even all the NNLO LECs in higher-multiplicity amplitudes, it is not unthinkable that useful results could be obtained by fixing most LECs and studying the rest. In any case, further explorations in these directions are mainly hindered by the lack of available amplitudes; the NLO 2-flavour $2 \rightarrow 4$ amplitude was recently calculated [45], but no higher-order or higher-multiplicity amplitudes are currently known. Furthermore, section 5 would need to be generalised to handle the 9-dimensional kinematic space of 6-particle amplitudes. Lastly, some parts of the NNNLO amplitude would be nonlinear in the NLO LECs, necessitating the development of proposition 4.2-like technology for *nonlinear* constraints, i.e. $\alpha_i b_i + \alpha_{ij} b_i b_j + \ldots \geq c$. We believe it possible that at least some of the tools in appendix B can be generalised to handle this, but have not investigated it much.

Lastly, these methods could lend themselves to application on EFTs other than ChPT, e.g. for beyond-the-Standard-Model (BSM) applications. This could be particularly promising if there are no experimentally measured values for the LECs, or if experiments have only yielded bounds. An experimental upper bound coupled with an analytic lower bound could confine the coupling of an unobserved process to a range, or exclude a BSM EFT altogether.

7.1 Acknowledgements

A   LEC details

This section contains further details about the LECs that are introduced in section 2.1, and the NNLO parameters $\Gamma_i, \Delta_i, \Theta_i, \Xi_i$ that are defined in section 5.4. Table II.2 contains the values determined in [6] that are used as references in section 6. Naturally, these do not include estimates of parameters that only appear above 3 flavours; similarly, the Γ_i, Δ_i that only appear in the $n \geq 3$ amplitude are just given provisional values based on the 3-flavour data.

Equation (II.1.1) shows how the independent NNLO parameters depend on the LECs K_i^r, for a general number of flavours n.[33]

$$\tfrac{1}{8}\Gamma_1 = 6K_1^r - 4K_5^r + K_7^r + K_{11}^r + 2K_{31}^r + nK_8^r + 2nK_{18}^r \tag{II.1.1a}$$

$$\tfrac{1}{2}\Gamma_2 = 12K_1^r - 48K_3^r - 12K_5^r + K_7^r + K_{11}^r - 16K_{13}^r + 32K_{17}^r - 8K_{19}^r - 8K_{23}^r - 16K_{28}^r$$
$$+ 2K_{31}^r - 8K_{33}^r + 16K_{37}^r + nK_8^r - 16nK_{14}^r + 34nK_{18}^r - 4nK_{20}^r \tag{II.1.1b}$$

$$\tfrac{1}{4}\Gamma_3 = K_1^r - 2K_3^r - 2K_5^r \tag{II.1.1c}$$

$$\tfrac{1}{8}\Gamma_4 = -3K_3^r - K_5^r \tag{II.1.1d}$$

$$\tfrac{1}{16}\Delta_1 = 6K_4^r - 2K_6^r + K_{15}^r + 2K_{29}^r + nK_{16}^r \tag{II.1.1e}$$

$$\tfrac{1}{4}\Delta_2 = -48K_2^r + 12K_4^r - 4K_6^r - 16K_9^r + K_{15}^r + 32K_{18}^r + 8K_{20}^r + 16K_{21}^r$$
$$+ 2K_{29}^r - 8K_{32}^r + 16K_{35}^r - 8K_{38}^r - 16nK_{10}^r + nK_{16}^r + 16nK_{22}^r \tag{II.1.1f}$$

$$\tfrac{1}{8}\Delta_3 = -2K_2^r + K_4^r \tag{II.1.1g}$$

$$\tfrac{1}{8}\Delta_4 = -6K_2^r + K_6^r \tag{II.1.1h}$$

---

[33]These are generated by FORM with minimal post-processing. Common factors have been extracted to make the expressions shorter.

Table II.2: All LECs and parameters covered by the bounds derived in this paper, along with their experimental reference values, taken from the most general fits in [6]. The values in parentheses indicate the uncertainties in the last decimal places. No uncertainties are provided for the NNLO parameters, since none are given for the NNLO LEC values in [6]; these values are little more than educated guesses. The values of Γ_i, Δ_i also depend directly on n for $i < 3$ (see eq. (II.1.1)); the listed values use $n = 4$.

NLO LEC		NNLO parameter $\cdot 10^3$	
$\bar{l}_1$	$-0.4(6)$	Θ_1	0.34
$\bar{l}_2$	$4.3(1)$	Θ_2	0.68
$\bar{l}_3$	$2.9(24)$	Θ_3	-0.16
$\bar{l}_4$	$4.4(2)$	Θ_4	-0.22
		Ξ_1	0.29
$\cdot 10^3$		Ξ_2	0.34
L_1^r	$1.11(10)$	Ξ_3	0.25
L_2^r	$1.05(17)$	Ξ_4	-0.008
L_3^r	$-3.82(30)$	Γ_1	0.008
L_4^r	$1.87(53)$	Γ_2	-0.71
L_5^r	$1.22(06)$	Γ_3	-0.10
L_6^r	$1.46(46)$	Γ_4	-0.22
L_8^r	$0.65(07)$	Δ_1	-0.032
		Δ_2	-0.83
		Δ_3	-0.048
		Δ_4	-0.14

Equation (II.1.2) shows the same for 3 flavours. The application of the $n = 3$ Cayley-Hamilton identity and the numbering of the C_i^r follows [27].

$$\frac{1}{4}\Xi_1 = -6C_1^r + 12C_3^r + 4C_4^r - C_5^r - 3C_6^r - C_8^r + 2C_{10}^r + 6C_{11}^r - 4C_{12}^r$$
$$- 18C_{13}^r + 4C_{22}^r - 4C_{25}^r \tag{II.1.2a}$$

$$\frac{3}{8}\Xi_2 = 50C_1^r + 27C_2^r + 22C_3^r - 18C_4^r - 48C_5^r + 54C_6^r + 45C_7^r - 21C_8^r + 24C_9^r$$
$$+ 3C_{10}^r + 9C_{11}^r + 6C_{13}^r + 6C_{16}^r - 4C_{24}^r + 12C_{25}^r - 4C_{26}^r + 4C_{29}^r \tag{II.1.2b}$$

$$\frac{3}{8}\Xi_3 = 70C_1^r - 18C_2^r + 68C_3^r - 16C_4^r + 43C_5^r + 21C_6^r - 30C_7^r + 25C_8^r$$
$$- 24C_9^r + 10C_{10}^r + 30C_{11}^r - 8C_{12}^r - 18C_{13}^r + 12C_{16}^r - 4C_{22}^r$$
$$+ 18C_{23}^r + 4C_{24}^r - 2C_{25}^r - 8C_{26}^r + 8C_{29}^r \tag{II.1.2c}$$

$$\frac{3}{4}\Xi_4 = -38C_1^r + 9C_2^r - 28C_3^r + 12C_4^r - 6C_{16}^r + 6C_{22}^r - 3C_{23}^r - 2C_{24}^r$$
$$+ 4C_{26}^r - 4C_{29}^r \tag{II.1.2d}$$

$$\frac{1}{4}\Gamma_3 = C_1^r - 2C_3^r - 2C_4^r \tag{II.1.2e}$$

$$\frac{1}{4}\Delta_3 = 4C_1^r - 5C_2^r + 2C_3^r \tag{II.1.2f}$$

Lastly, eq. (II.1.3) shows the same for 2 flavours. Again, we have followed [27].

$$\tfrac{3}{16}\Theta_1 = 14c_1^r + 40c_2^r - 6c_3^r + 15c_4^r + 30c_5^r - 30c_6^r + 12c_7^r + 9c_{12}^r - 3c_{13}^r$$
$$\qquad - 4c_{14}^r + 4c_{16}^r \tag{II.1.3a}$$
$$\tfrac{3}{8}\Theta_2 = 104c_1^r + 22c_2^r - 18c_3^r + 93c_4^r + 15c_5^r + 12c_7^r + 12c_{13}^r - 4c_{14}^r + 4c_{16}^r \tag{II.1.3b}$$
$$\tfrac{3}{16}\Theta_3 = 6c_1^r - 9/2c_2^r - 3c_3^r \tag{II.1.3c}$$
$$\tfrac{3}{4}\Theta_4 = -14c_1^r - 40c_2^r - 12c_7^r + 3c_{12}^r + 4c_{14}^r - 4c_{16}^r \tag{II.1.3d}$$

B Details and proofs regarding linear constraints

In this appendix, we prove the propositions stated in section 4 and provide some more details on how they may be applied.[34]

B.1 Proof of propositions 4.1 and 4.2

As has already been mentioned, proposition 4.1 is a direct consequence of proposition 4.2, obtained by leaving all but one I_c empty. However, directly proving proposition 4.2 is much less straightforward than the following chain of implications,

$$\text{Proposition 4.1 } (c = 0 \text{ only}) \quad \Rightarrow \quad \text{Proposition 4.2} \quad \Rightarrow \quad \text{Proposition 4.1}, \tag{II.2.1}$$

which we will demonstrate in this section. First, however, we will show an easily accessible partial result (section B.1.1), and then prove some properties that are necessary for the main proof (section B.1.2).

B.1.1 The trivial half of the proof

One side of proposition 4.2 is easy to prove, namely that $\langle \boldsymbol{\beta}, c \rangle \le \Omega$ if $\boldsymbol{\beta} \in \mathcal{A}_c(\Omega)$.

Assuming that $\Omega \ne \Omega_\infty$, take any point $\boldsymbol{b}$ that satisfies Ω. Then by eq. (II.4.10),

$$\boldsymbol{\beta} \cdot \boldsymbol{b} = \sum_{i \in I_1} \lambda_i \boldsymbol{\alpha}_i \cdot \boldsymbol{b} + \sum_{i \in I_0} \lambda_i \boldsymbol{\alpha}_i \cdot \boldsymbol{b} + \sum_{i \in I_{-1}} \lambda_i \boldsymbol{\alpha}_i \cdot \boldsymbol{b}, \tag{II.2.2}$$

and since $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} \ge c$ for $i \in I_c$, we have

$$\boldsymbol{\beta} \cdot \boldsymbol{b} \ge \sum_{i \in I_1} \lambda_i - \sum_{i \in I_{-1}} \lambda_i \ge c, \tag{II.2.3}$$

which uses (and motivates) eq. (II.4.11). The corresponding result for proposition 4.1 follows immediately.

---

[34]Like in section 4, we make use of potentially unfamiliar mathematical notation in this section, so the glossary (section B.7) may be helpful.

B.1.2 Proof that $\mathcal{A}(\omega_0)$ is closed and convex

As section B.1.3 will show, these properties of $\mathcal{A}(\omega_0)$ are crucial for the main proof. Convexity is easy to show for $\mathcal{A}(\omega_0)$: given any points $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ satisfying eq. (II.4.6), their convex combination

$$\mu\boldsymbol{\beta}_1 + (1-\mu)\boldsymbol{\beta}_2 = \sum_{i \in I} \left(\mu\lambda_{1i} + (1-\mu)\lambda_{2i}\right)\boldsymbol{\alpha}_i, \qquad \lambda_{1i} \geq 0, \quad \lambda_{2i} \geq 0 \qquad \text{(II.2.4)}$$

satisfies eq. (II.4.6) as well.

For the proof of closedness, we only need three basic facts: (i) the intersection of closed sets is closed, (ii) the union of a finite number of closed sets is closed, and (iii) for any $\boldsymbol{\alpha}$ and c the set $\mathcal{B}\left(\langle\boldsymbol{\alpha}, c\rangle\right)$ is closed.[35]

Now, we employ *Carathéodory's theorem for convex cones*,[36] which for our purposes can be formulated as follows:

> Let $I \subset \mathbb{N}$ be finite, and let $\{\boldsymbol{\alpha}_i\}_{i \in I}$ be vectors in $\mathbb{R}^D$. For any point $\boldsymbol{\beta} \in \mathbb{R}^D$ fulfilling
>
> $$\boldsymbol{\beta} = \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i \geq 0, \qquad \text{(II.2.6)}$$
>
> there exists a set $J \subseteq I$ with at most D elements such that
>
> $$\boldsymbol{\beta} = \sum_{j \in J} \lambda_j \boldsymbol{\alpha}_j, \qquad \lambda_j \geq 0, \qquad \text{(II.2.7)}$$
>
> where and all $\boldsymbol{\alpha}_j$ are linearly independent.

Equation (II.2.6) is clearly equivalent to eq. (II.4.6) for $c = 0$.

For any such J, there exists a set of vectors $\{\boldsymbol{\gamma}_k\}$ such that $\{\boldsymbol{\alpha}_j\}_{j \in J} \cup \{\boldsymbol{\gamma}_k\}$ is a basis of $\mathbb{R}^D$. Let $\mathbf{A}_J$ be the invertible matrix whose columns are these basis vectors, and let $\boldsymbol{\lambda}$ be the vector whose components are λ_j, where $\lambda_j = 0$ if $j \notin J$. Then eq. (II.2.7) can be rewritten as $\boldsymbol{\beta} = \mathbf{A}_J \boldsymbol{\lambda}$, or equivalently $\mathbf{A}_J^{-1}\boldsymbol{\beta} = \boldsymbol{\lambda}$. Since $\lambda_j \geq 0$, we therefore obtain the inequalities $\boldsymbol{a}_j \cdot \boldsymbol{\beta} \geq 0$, where $\boldsymbol{a}_j$ are the column vectors of $\mathbf{A}_J^{-1}$, and $\boldsymbol{a}_j \cdot \boldsymbol{\beta} = 0$ if $j \notin J$. In other words,[37]

$$\boldsymbol{\beta} \in \mathcal{B}\left(\sum_j \langle\boldsymbol{a}_j, 0\rangle + \sum_{j \notin J} \langle-\boldsymbol{a}_j, 0\rangle\right). \qquad \text{(II.2.8)}$$

---

[35]This last fact is easy to prove: take a point $\boldsymbol{b} \notin \mathcal{B}(\langle\boldsymbol{\alpha}, c\rangle)$, i.e. $\boldsymbol{\alpha} \cdot \boldsymbol{b} < c$. Then for any vector $\boldsymbol{d}$ such that $|\boldsymbol{d}| < \varepsilon$,

$$\boldsymbol{\alpha} \cdot (\boldsymbol{b} + \boldsymbol{d}) \leq \boldsymbol{\alpha} \cdot \boldsymbol{b} + |\boldsymbol{\alpha} \cdot \boldsymbol{d}| < \boldsymbol{\alpha} \cdot \boldsymbol{b} + \varepsilon|\boldsymbol{\alpha}|, \qquad \text{(II.2.5)}$$

where we used the Cauchy-Schwarz inequality in the last step. For $\varepsilon > 0$ sufficiently small, this is still less than c, so $\boldsymbol{b} + \boldsymbol{d} \notin \mathcal{B}(\langle\boldsymbol{\alpha}, c\rangle)$, proving that it is closed. (Note that this also works for $\mathcal{B}(\langle\mathbf{0}, 1\rangle) = \emptyset$ and $\mathcal{B}(\langle\mathbf{0}, -1\rangle) = \mathbb{R}^D$.)

[36]In common mathematical nomenclature, $\mathcal{A}(\omega_0)$ is a convex cone, $\mathcal{A}(\omega_1)$ is an affine cone, and $\mathcal{A}(\omega_{-1})$ is a convex hull. We have chosen not to use these classifications, partly because neither applies to the general $\mathcal{A}_c(\Omega)$.

[37]Note that $\boldsymbol{\beta}$, which normally is part of a linear constraint, is itself constrained here. This is not a problem; in fact, the expression $\boldsymbol{\alpha} \cdot \boldsymbol{b} \geq c$ and be interpreted both as $\boldsymbol{b} \in \mathcal{B}(\langle\boldsymbol{\alpha}, c\rangle)$ and as $\boldsymbol{\alpha} \in \mathcal{B}(\langle\boldsymbol{b}, c\rangle)$. We will return to this symmetric interpretation many times below.

By facts (i) and (iii), this set is always closed. Therefore, the set of all $\boldsymbol{\beta} \in \mathcal{B}(\omega_0)$ associated with the same J is closed. $\mathcal{A}(\omega_0)$ must then be the union of all such sets, but since I is finite, there are finitely many different subsets J, so $\mathcal{A}(\omega_0)$ is the union of a finite number of closed sets. By fact (ii), it is therefore closed. $\qquad\square$

Let us remark that this proof extends to the other cases, so that $\mathcal{A}_c(\Omega)$ is closed and convex for any Ω and c.

B.1.3 Proof of proposition 4.1 in the $c = 0$ case

We will now turn our attention to the statement that $\langle \boldsymbol{\beta}, 0 \rangle \not\leq \omega_0$ if $\boldsymbol{\beta} \notin \mathcal{A}(\omega_0)$, which will complete the proof of proposition 4.1 for $c = 0$.[38] We employ the *separating hyperplane theorem*, which can be formulated as follows:

> Let $\mathcal{X}$ and $\mathcal{Y}$ be disjoint convex sets, with $\mathcal{X}$ closed and $\mathcal{Y}$ compact. Then there exists a nonzero vector $\boldsymbol{h}$ and a real number d such that
>
> $$\boldsymbol{\chi} \cdot \boldsymbol{h} > d \quad \text{and} \quad \boldsymbol{\psi} \cdot \boldsymbol{h} < d \qquad (\text{II.2.9})$$
>
> for all $\boldsymbol{\chi} \in \mathcal{X}, \boldsymbol{\psi} \in \mathcal{Y}$. (The set $\{\boldsymbol{\pi} \mid \boldsymbol{\pi} \cdot \boldsymbol{h} = d\}$ is a hyperplane that separates $\mathcal{X}$ from $\mathcal{Y}$, hence the name.)

Since $\mathcal{A}(\omega_0)$ is closed and convex, and because the set consisting of the single point $\boldsymbol{\beta} \notin \mathcal{A}(\omega_0)$ is compact and convex, the separating hyperplane theorem implies that there exists $\langle \boldsymbol{h}, d \rangle$ such that

$$\forall \boldsymbol{\alpha} \in \mathcal{A}(\omega_0), \quad \boldsymbol{h} \cdot \boldsymbol{\alpha} > d \quad \text{and} \quad \boldsymbol{h} \cdot \boldsymbol{\beta} < d. \qquad (\text{II.2.10})$$

Since $\boldsymbol{0} \in \mathcal{A}(\omega_0)$, we know that $d < 0$. However, we claim that for any $\boldsymbol{\alpha} \in \mathcal{A}(\omega_0)$, we in fact have $\boldsymbol{h} \cdot \boldsymbol{\alpha} \geq 0$. Indeed, if we assume that there exists $\boldsymbol{\alpha} \in \mathcal{A}(\omega_0)$ such that $\boldsymbol{h} \cdot \boldsymbol{\alpha} < 0$, then for any $\lambda \geq \frac{|d|}{|\boldsymbol{\alpha} \cdot \boldsymbol{h}|}$, eq. (II.4.6) implies that $\lambda \boldsymbol{\alpha} \in \mathcal{A}(\omega_0)$. But $\lambda \boldsymbol{\alpha} \cdot \boldsymbol{h} \leq -|d|$, which contradicts the fact that $\boldsymbol{\alpha} \cdot \boldsymbol{h} > d$. Therefore, we have $\boldsymbol{\alpha} \cdot \boldsymbol{h} \geq 0$ and $\boldsymbol{\beta} \cdot \boldsymbol{h} < 0$, implying that $\langle \boldsymbol{\beta}, 0 \rangle \not\leq \omega_0$. This proves proposition 4.1. $\qquad\blacksquare$

B.1.4 Proof of proposition 4.2

In order to reduce the general Ω defined in eq. (II.4.9) to one that can be handled by proposition 4.1 for $c = 0$, we define the "lifted" vector

$$\ell_x(\boldsymbol{v}) = (v_1, v_2, \ldots, v_D, x), \qquad (\text{II.2.11})$$

where D is the dimension of the original vector $\boldsymbol{v}$. Then we note that

$$\boldsymbol{\alpha} \cdot \boldsymbol{b} \geq c \quad \Leftrightarrow \quad \ell_{-c}(\boldsymbol{\alpha}) \cdot \ell_1(\boldsymbol{b}) = \boldsymbol{\alpha} \cdot \boldsymbol{b} - c \geq 0. \qquad (\text{II.2.12})$$

---

[38]The same method is easy to apply to the $c = -1$ case and, with some slight complications, the $c = 1$ case. With considerable effort, it can also be extended to proposition 4.2. However, we will follow the outline eq. (II.2.1) and only prove what is necessary.

Thus, any D-dimensional linear constraint $\langle \boldsymbol{\alpha}, c \rangle$ can be lifted into a $(D+1)$-dimensional linear constraint $\langle \ell_{-c}(\boldsymbol{\alpha}), 0 \rangle$. We can now show that proposition 4.1 in the lifted space, where we only ever have $c = 0$, is equivalent to proposition 4.2 in the original space. Define

$$\omega_0^\ell = \sum_{i \in I_1} \langle \ell_{-1}(\boldsymbol{\alpha}_i), 0 \rangle + \sum_{i \in I_0} \langle \ell_0(\boldsymbol{\alpha}_i), 0 \rangle + \sum_{i \in I_{-1}} \langle \ell_1(\boldsymbol{\alpha}_i), 0 \rangle + \langle \ell_1(\boldsymbol{0}), 0 \rangle, \qquad \text{(II.2.13)}$$

where the extra constraint $\langle \ell_1(\boldsymbol{0}), 0 \rangle$ imposes that $\ell_x(\boldsymbol{b})$ only satisfies ω_0^ℓ if $x \geq 0$. ω_0^ℓ can be thought of as a lifted version of Ω, and as indicated by the notation, it fulfils the definition of ω_0 so that proposition 4.1 for $c = 0$ applies to it.

Now, assume that $\ell_{-c}(\boldsymbol{\beta}) \in \mathcal{A}(\omega_0^\ell)$. Looking at eq. (II.4.6), we find that

$$\begin{aligned} \boldsymbol{\beta} &= \sum_{i \in I_1} \lambda_i \boldsymbol{\alpha}_i + \sum_{i \in I_0} \lambda_i \boldsymbol{\alpha}_i + \sum_{i \in I_{-1}} \lambda_i \boldsymbol{\alpha}_i, + \lambda' \boldsymbol{0}, \\ -c &= \sum_{i \in I_{-1}} \lambda_i - \sum_{i \in I_1} \lambda_i + \lambda', \qquad \lambda_i \geq 0, \lambda' \geq 0. \end{aligned} \qquad \text{(II.2.14)}$$

These exactly reproduce eqs. (II.4.10) and (II.4.11), so we have shown that $\ell_{-c}(\boldsymbol{\beta}) \in \mathcal{A}(\omega_0^\ell)$ implies $\boldsymbol{\beta} \in \mathcal{A}_c(\Omega)$.

Conversely, let us now assume, $\ell_{-c}(\boldsymbol{\beta}) \notin \mathcal{A}(\omega_0^\ell)$. Proposition 4.1 then implies the existence of some $\ell_x(\boldsymbol{b}) \in \mathcal{B}(\omega_0^\ell)$ with $x \geq 0$ such that $\ell_{-c}(\boldsymbol{\beta}) \cdot \ell_x(\boldsymbol{b}) < 0$, and for any $i \in I$, $\ell_{-c_i}(\boldsymbol{\alpha}_i) \cdot \ell_x(\boldsymbol{b}) \geq 0$. We may moreover assume that $x > 0$, for if $x = 0$, we may choose any $\boldsymbol{a} \in \mathcal{B}(\Omega)$ and consider $\ell_\varepsilon(\boldsymbol{b}') = (1 - \varepsilon)\ell_0(\boldsymbol{b}) + \varepsilon \ell_1(\boldsymbol{a})$ for $0 < \varepsilon < 1$. By convexity, $\ell_\varepsilon(\boldsymbol{b}') \in \mathcal{B}(\omega_0^\ell)$, and

$$\ell_{-c}(\boldsymbol{\beta}) \cdot \ell_\varepsilon(\boldsymbol{b}') = \boldsymbol{\beta} \cdot \boldsymbol{b} + \varepsilon \left[\boldsymbol{\beta} \cdot (\boldsymbol{a} - \boldsymbol{b}) - c \right], \qquad \text{(II.2.15)}$$

so for ε small enough, $\ell_{-c}(\boldsymbol{\beta}) \cdot \ell_\varepsilon(\boldsymbol{b}') < 0$. Consequently,

$$\begin{cases} \boldsymbol{\alpha}_i \cdot \boldsymbol{b} \geq x, & i \in I_1, \\ \boldsymbol{\alpha}_i \cdot \boldsymbol{b} \geq 0, & i \in I_0, \\ \boldsymbol{\alpha}_i \cdot \boldsymbol{b} \geq -x, & i \in I_{-1}, \end{cases} \qquad \boldsymbol{\beta} \cdot \boldsymbol{b} < xc. \qquad \text{(II.2.16)}$$

Thus, $\frac{1}{x}\boldsymbol{b}$ is a point that satisfies Ω but not $\langle \boldsymbol{\beta}, c \rangle$, which means that $\ell_{-c}(\boldsymbol{\beta}) \notin \mathcal{A}(\omega_0^\ell)$ implies $\boldsymbol{\beta} \notin \mathcal{A}_c(\Omega)$.

We have now shown that $\ell_{-c}(\boldsymbol{\beta}) \in \mathcal{A}(\omega_0^\ell)$ is equivalent to $\boldsymbol{\beta} \in \mathcal{A}_c(\Omega)$, and since $\langle \ell_{-c}(\boldsymbol{\beta}), 0 \rangle \leq \omega_0^\ell$ is equivalent to $\langle \boldsymbol{\beta}, c \rangle \leq \Omega$, we have therefore proven proposition 4.2 as a consequence of the $c = 0$ version of proposition 4.1. As mentioned before, proposition 4.1 for $c = \pm 1$ follows easily. ∎

B.1.5 An important corollary

The following interesting result is a consequence of proposition 4.2:

Corollary B.1 (boundedness of $\mathcal{B}(\Omega)$). *For $\Omega \neq \Omega_\infty$, the region $\mathcal{B}(\Omega)$ is bounded if and only if the origin, $\mathbf{0}$, is in the interior of $\mathcal{A}_{-1}(\Omega)$. This happens if and only if $\{\boldsymbol{\alpha}_i\}_{i \in I}$ spans the full D-dimensional space and there are λ_i such that*

$$\mathbf{0} = \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i > 0 \tag{II.2.17}$$

PROOF. We will show the converse, namely that $\mathcal{B}(\Omega)$ being unbounded is equivalent to $\mathbf{0}$ not being in the interior of $\mathcal{A}_{-1}(\Omega)$.

First assume that $\mathbf{0}$ is not in the interior of $\mathcal{A}_{-1}(\Omega)$. Then for all $\varepsilon > 0$, there exists some $\boldsymbol{\beta}_\varepsilon$ such that $|\boldsymbol{\beta}_\varepsilon| < \varepsilon$ and $\boldsymbol{\beta}_\varepsilon \notin \mathcal{A}_{-1}(\Omega)$. Since $\langle \boldsymbol{\beta}_\varepsilon, -1 \rangle \not\preceq \Omega$, there consequently exists a point $\boldsymbol{b}_\varepsilon \in \mathcal{B}(\Omega)$ such that $\boldsymbol{\beta}_\varepsilon \cdot \boldsymbol{b}_\varepsilon < -1$. Now, the Cauchy-Schwarz inequality gives

$$1 < |\boldsymbol{\beta}_e \cdot \boldsymbol{b}_e| \leq |\boldsymbol{\beta}_e| \, |\boldsymbol{b}_e| < \varepsilon |\boldsymbol{b}_\varepsilon|. \tag{II.2.18}$$

Since this holds for arbitrarily small ε, there can be no upper bound on $|\boldsymbol{b}_\varepsilon|$; therefore, $\mathcal{B}(\Omega)$ is unbounded.

Conversely, assume that $\mathcal{B}(\Omega)$ is unbounded. Then for all $M > 0$, there must exist some $\boldsymbol{b}_M \in \mathcal{B}(\Omega)$ such that $|\boldsymbol{b}_M| > M$. Now define

$$\boldsymbol{\beta}_M \equiv -\frac{\boldsymbol{b}_M}{|\boldsymbol{b}_M|^{3/2}} \quad \Rightarrow \quad |\boldsymbol{\beta}_M| = \frac{1}{|\boldsymbol{b}_M|^{1/2}} < \frac{1}{\sqrt{M}}, \quad \boldsymbol{\beta}_M \cdot \boldsymbol{b}_M = -|\boldsymbol{b}_M|^{1/2} < -\sqrt{M}. \tag{II.2.19}$$

For sufficiently large M, the last inequality implies that $\langle \boldsymbol{\beta}_M, -1 \rangle \not\preceq \Omega$, so $\boldsymbol{\beta}_M \notin \mathcal{A}_{-1}(\Omega)$. However, the inequality before that tells us that $\boldsymbol{\beta}_M$ may lie arbitrarily close to the origin. Therefore, $\mathbf{0} \in \mathcal{A}_{-1}(\Omega)$ must lie on the boundary, not the interior, of $\mathcal{A}_{-1}(\Omega)$.

That completes the main proof, but we must also prove condition about the span of $\{\boldsymbol{\alpha}_i\}_{i \in I}$. If the span was lower-dimensional, then there would exist some vector $\boldsymbol{\beta}$ linearly independent of all $\boldsymbol{\alpha}_i$, and then clearly $\varepsilon \boldsymbol{\beta} \notin \mathcal{A}_{-1}(\Omega)$ for all $\varepsilon > 0$, implying that $\mathbf{0}$ is not in the interior of $\mathcal{A}_{-1}(\Omega)$. Lastly, we must prove eq. (II.2.17), which is essentially eq. (II.4.12) with $\lambda_i > 0$ rather than $\lambda_i \geq 0$. Since $\{\boldsymbol{\alpha}_i\}_{i \in I}$ spans the full space, any vector $\boldsymbol{v}$ of sufficiently small magnitude satisfies

$$\boldsymbol{v} \in \text{Hull}\left(\{\boldsymbol{\alpha}_i\}_{i \in I} \cup \left\{-\sum_{i \in I} \boldsymbol{\alpha}_i\right\}\right). \tag{II.2.20}$$

Thus, $\mathbf{0}$ is in the interior of $\mathcal{A}_{-1}(\Omega)$ if and only if for sufficiently small $\varepsilon > 0$[39]

$$-\varepsilon \sum_{i \in I} \boldsymbol{\alpha}_i \in \mathcal{A}_{-1}(\Omega) \quad \Leftrightarrow \quad -\varepsilon \sum_{i \in I} \boldsymbol{\alpha}_i = \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i \quad \Leftrightarrow \quad \mathbf{0} = \sum_{i \in I} (\lambda_i + \varepsilon) \boldsymbol{\alpha}_i \tag{II.2.21}$$

for $\lambda_i \geq 0$, which implies $\lambda_i + \varepsilon > 0$, thereby producing eq. (II.2.17). ∎

---

[39]Here, we neglect the condition $\sum_{i \in I} \lambda_i c_i \geq -1$ in the second equality, since it is always possible to multiply both sides by a positive factor to rescale the λ_i appropriately.

B.2   Some mathematical tools

Before moving on with proving proposition 4.3 and deriving further results, we need to establish some tools and terminology that range from useful to crucial in subsequent sections.

B.2.1   The degenerate constraint framework

In this section, we properly define what it means for a constraint to be degenerate, and derive notations and results that are not only useful for the proof of proposition 4.3 and its generalisation, but also for many other things later in this appendix.

In a D-dimensional space, consider an affine subspace E of dimension d. There exists two sets of vectors $\{\boldsymbol{g}_j\}_{j=1}^{d}, \{\boldsymbol{\delta}_k\}_{k=1}^{D-d}$ whose union forms an orthonormal basis for $\mathbb{R}^D$, that, given an arbitrary point $\boldsymbol{e} \in E$, allow E to be expressed in two complementary ways:[40]

$$E = \left\{ \boldsymbol{e} + \sum_{j=1}^{d} x_j \boldsymbol{g}_j \;\middle|\; x_j \in \mathbb{R} \right\}, \quad E = \mathcal{B}\left[\sum_{k=1}^{D-d} \left(\langle \boldsymbol{\delta}_k, \boldsymbol{\delta}_k \cdot \boldsymbol{e} \rangle + \langle -\boldsymbol{\delta}_k, -\boldsymbol{\delta}_k \cdot \boldsymbol{e} \rangle \right) \right]. \quad (\text{II.2.22})$$

Up to the choice of $\boldsymbol{e}$ and the basis vectors, any vector $\boldsymbol{b}$ in parameter space can be uniquely decomposed as

$$\boldsymbol{b} = \boldsymbol{e} + \sum_{j=1}^{d} x_j \boldsymbol{g}_j + \sum_{k=1}^{D-d} z_k \boldsymbol{\delta}_k. \quad (\text{II.2.23})$$

We then define $\lfloor \boldsymbol{b} \rfloor_E \equiv (x_1, x_2, \ldots x_d)$ and $\lceil \boldsymbol{b} \rceil^E \equiv (z_1, z_2, \ldots, z_{D-d})$. These are d- and $(D - d)$-dimensional vectors, respectively, and live in spaces separate from the D-dimensional space in which $E, \boldsymbol{e}$, etc. live. Note that if $\boldsymbol{b} \in E$, then $\lceil \boldsymbol{b} \rceil^E = \boldsymbol{0}$ and $\boldsymbol{b}$ is uniquely determined by $\lfloor \boldsymbol{b} \rfloor_E$. For constraint-space vectors, we instead make the decomposition

$$\boldsymbol{\alpha} = \sum_{j=1}^{d} \xi_j \boldsymbol{g}_j + \sum_{k=1}^{D-d} \zeta_k \boldsymbol{\delta}_k \quad (\text{II.2.24})$$

and analogously define $\lfloor \boldsymbol{\alpha} \rfloor_E \equiv (\xi_1, \xi_2, \ldots \xi_d)$ and $\lceil \boldsymbol{\alpha} \rceil^E \equiv (\zeta_1, \zeta_2, \ldots \zeta_{D-d})$.[41] These can form constraints acting on $\lfloor \boldsymbol{b} \rfloor_E$ and $\lceil \boldsymbol{b} \rceil^E$, respectively. As an extension, for any set $\mathcal{X}$ we define

$$\lfloor \mathcal{X} \rfloor_E \equiv \left\{ \lfloor \boldsymbol{x} \rfloor_E \;\middle|\; \boldsymbol{x} \in \mathcal{X} \right\}, \qquad \lceil \mathcal{X} \rceil^E \equiv \left\{ \lceil \boldsymbol{x} \rceil^E \;\middle|\; \boldsymbol{x} \in \mathcal{X} \right\}. \quad (\text{II.2.25})$$

---

[40]Note that we have written $\boldsymbol{g}_j$ as parameter-space vectors, and $\boldsymbol{\delta}_k$ as constraint-space vectors. This is consistent with their use in eq. (II.2.22), but eqs. (II.2.23) and (II.2.24) are in a sense breaking our conventions by adding vectors of different types. This is of course no problem when both parameter and constraint space are just $\mathbb{R}^D$, but if we considered constraints in more general spaces, we would have to make appropriate adjustments to our formulae.

[41]Our notation does not make the choice of reference point $\boldsymbol{e}$ explicit, and it is arbitrary for all purposes. Replacing $\boldsymbol{e} \to \boldsymbol{e}'$ simply entails translating all $\lfloor \boldsymbol{b} \rfloor_E \to \lfloor \boldsymbol{b} \rfloor_E + \lfloor \boldsymbol{e}' - \boldsymbol{e} \rfloor_E$ while leaving $\lceil \boldsymbol{b} \rceil^E$, $\lfloor \boldsymbol{\alpha} \rfloor_E$ and $\lceil \boldsymbol{\alpha} \rceil^E$ unchanged. Likewise, altering $\{\boldsymbol{g}_j\}_{j=1}^{d}, \{\boldsymbol{\delta}_k\}_{k=1}^{D-d}$ just corresponds to a change of basis in the spaces.

We reiterate how important it is to view $\mathcal{X}$, $\lfloor \mathcal{X} \rfloor_E$ and $\lceil \mathcal{X} \rceil^E$ as living in three different spaces. There is of course a straightforward mapping between $\lfloor \mathcal{X} \rfloor_E$ and $\mathcal{X} \cap E$ — indeed, the d-dimensional space can be seen as the vector space underlying the affine subspace E — but the notion of separate spaces makes the proofs below clearer.

Before moving on to constraints, let us make the folloing definition:

> The *dimension* of any nonempty set $\mathcal{X}$, written $\dim(\mathcal{X})$, is the affine dimension of the smallest (i.e. lowest-dimensional) affine subspace that contains $\mathcal{X}$. Equivalently, $\dim(\mathcal{X})$ is the dimension of the affine span of the points in $\mathcal{X}$.[42]

This smallest affine subspace is clearly unique, for if it is not, the intersection of all such subspaces is even smaller. For any nonempty convex set $\mathcal{C} \in \mathbb{R}^D, D > 0$, the following basic fact holds:

$$\dim(\mathcal{C}) < D \quad \Leftrightarrow \quad \mathrm{int}(\mathcal{C}) = \emptyset, \tag{II.2.26}$$

where $\mathrm{int}(\mathcal{C})$ is the interior of $\mathcal{C}$.

Consider now a constraint $\Omega \neq \Omega_\infty$ and define $d_\Omega \equiv \dim\big[\mathcal{B}(\Omega)\big]$. We formalise the definition of degeneracy stated in section 4.4 as follows:

> A constraint $\Omega \neq \Omega_\infty$ in D-dimensional space is *degenerate* if $d_\Omega < D$ and *non-degenerate* otherwise. Ω_∞, for which d_{Ω_∞} is undefined, does not fall into either category.

Let E_Ω be the unique d_Ω-dimensional affine subspace that contains $\mathcal{B}(\Omega)$. Given linear constraint $\langle \boldsymbol{\alpha}, c \rangle$, we define

$$\Big\lfloor \langle \boldsymbol{\alpha}, c \rangle \Big\rfloor_\Omega \equiv \langle\, \lfloor \boldsymbol{\alpha} \rfloor_{E_\Omega}, \quad c - \boldsymbol{\alpha} \cdot \boldsymbol{e} \rangle, \qquad \Big\lceil \langle \boldsymbol{\alpha}, c \rangle \Big\rceil^\Omega \equiv \langle\, \lceil \boldsymbol{\alpha} \rceil^{E_\Omega}, \quad c - \boldsymbol{\alpha} \cdot \boldsymbol{e} \rangle. \tag{II.2.27}$$

For these, the following holds:

Lemma B.1. *For any point $\boldsymbol{b}$,*

(a) If $\lceil \boldsymbol{b} \rceil^{E_\Omega} = \boldsymbol{0}$, then $\lfloor \boldsymbol{b} \rfloor_{E_\Omega}$ satisfies $\lfloor \langle \boldsymbol{\alpha}, c \rangle \rfloor_\Omega$ if and only if $\boldsymbol{b}$ satisfies $\langle \boldsymbol{\alpha}, c \rangle$.

(b) If $\lfloor \boldsymbol{\alpha} \rfloor_{E_\Omega} = \boldsymbol{0}$, then $\lceil \boldsymbol{b} \rceil^{E_\Omega}$ satisfies $\lceil \langle \boldsymbol{\alpha}, c \rangle \rceil^\Omega$ if and only if $\boldsymbol{b}$ satisfies $\langle \boldsymbol{\alpha}, c \rangle$.

PROOF. This follows directly from eqs. (II.2.23), (II.2.24) and (II.2.27) and the orthonormality of $\{\boldsymbol{g}_j\}_{j=1}^d \cup \{\boldsymbol{\delta}_k\}_{k=1}^{D-d}$. $\qquad\square$

Using this, we define for any $\Omega \neq \Omega_\infty$ with representation $\mathcal{S}$

$$\lfloor \Omega \rfloor \equiv \sum_{\langle \boldsymbol{\alpha}, c \rangle \in \mathcal{S}} \lfloor \langle \boldsymbol{\alpha}, c \rangle \rfloor_\Omega . \tag{II.2.28}$$

This constraint, which acts on the d_Ω-dimensional space of vectors $\lfloor \boldsymbol{b} \rfloor_{E_\Omega}$, has three important properties. Firstly, $\lfloor \Omega \rfloor$ is satisfied by $\lfloor \boldsymbol{b} \rfloor_{E_\Omega}$ if Ω is satisfied by $\boldsymbol{b}$, and the converse holds when $\lceil \boldsymbol{b} \rceil^{E_\Omega} = \boldsymbol{0}$ (this follows from lemma B.1), so

$$\mathcal{B}(\lfloor \Omega \rfloor) = \lfloor \mathcal{B}(\Omega) \rfloor_{E_\Omega} . \tag{II.2.29}$$

---

[42]Note that this definition agrees with the usual affine/linear dimension when $\mathcal{X}$ is itself an affine/linear subspace.

Secondly, $\lfloor \Omega \rfloor$ is independent of $\mathcal{S}$ as a consequence of this. Lastly, $\lfloor \Omega \rfloor$ is, by construction, non-degenerate.[43] Thanks to these properties, $\lfloor \Omega \rfloor$ is key to all further treatment of degenerate constraints.

B.2.2 K-faces

In this section, we introduce K-*faces*, which will be highly useful in subsequent sections; especially corollary B.2 and propositions B.3 and B.4 rely heavily on them. Like several other things introduced here, it is partially based on standard concepts and nomenclature, but has been adapted and extended to fit the context of linear constraints. K-faces are defined as follows:

> Let $\mathcal{F}$ be a non-empty convex subset of a closed convex set $\mathcal{C}$, with $K = \dim(\mathcal{F})$. Then $\mathcal{F}$ is called a K-*face of* $\mathcal{C}$ if the following holds: For every $\boldsymbol{\varphi} \in \mathcal{F}$, if there exists $\boldsymbol{\eta}_{1,2} \in \mathcal{C}$ such that $\boldsymbol{\varphi} = \mu\boldsymbol{\eta}_1 + (1-\mu)\boldsymbol{\eta}_2$ with $\mu \in (0,1)$, then $\boldsymbol{\eta}_{1,2} \in \mathcal{F}$.
>
> The single point in a 0-face of $\mathcal{C}$ is called an *extreme point* or *vertex* of $\mathcal{C}$; it is a point that cannot be expressed as a convex combination of any two points in $\mathcal{C}$ distinct from itself.
>
> A 1-face of $\mathcal{C}$ is called an *edge*.
>
> A $(D-1)$-face of $\mathcal{C}$ is called a *facet*.

For convex polygons, polyhedra, etc., these definitions agree with the usual concepts of vertices, edges and facets. A number of useful properties of K-faces easily follow from the definitions:

(i) No K-face of $\mathcal{C}$ is a strict subset of another K-face of $\mathcal{C}$ with the same K.

(ii) There is a unique $\dim(\mathcal{C})$-face of $\mathcal{C}$, namely $\mathcal{C}$ itself.

(iii) All K-faces of $\mathcal{C}$ (except possibly $\mathcal{C}$ itself) are contained in the boundary of $\mathcal{C}$.

(iv) If $\mathcal{F}$ is a K-face of $\mathcal{C}$, then for all $K' \leq K$ the K'-faces of $\mathcal{F}$ are also K'-faces of $\mathcal{C}$. Specifically, an edge may have up to two vertices, which are its endpoints.

(v) If $\mathcal{F}$ is a K-face of $\mathcal{C}$ and $\mathcal{C}' \subseteq \mathcal{C}$, then if $\mathcal{F} \cap \mathcal{C}'$ is non-empty, it is a K'-face $(K' \leq K)$ of $\mathcal{C}'$.

We will now prove some less obvious properties. In the remainder of this section, let $\mathcal{C}$ be any convex set such that there exists a finite set of constraints $\{\langle \boldsymbol{g}_\ell, c_\ell \rangle\}_{\ell \in L}$ fulfilling[44]

$$\mathcal{C} = \mathcal{B}\left(\sum_{\ell \in L} \langle \boldsymbol{g}_\ell, c_\ell \rangle\right). \tag{II.2.30}$$

We then begin with the following technical lemma:

---

[43]If $d_\Omega = 0$, so that $\mathcal{B}(\Omega)$ is a single point, then $\lfloor \Omega \rfloor$ is a zero-dimensional constraint. This is not conceptually a problem for proposition B.1 below: $\lfloor \Omega \rfloor$ is satisfied by $\boldsymbol{0}$, which is the only point in zero-dimensional space, and $\mathcal{R}(\lfloor \Omega \rfloor) = \emptyset$.

[44]This class of sets includes all convex hulls of finite sets, all linear and affine subspaces, as well as most other sets we work with, including $\mathcal{A}_c(\Omega)$ as we will prove in section B.4. For convenience, we express $\mathcal{C}$ as a subset of constraint space, but all results hold equally well if $\mathcal{C}$ is a subset of parameter space.

Lemma B.2. *Let $\mathcal{C}$ be a convex set defined as in eq. (II.2.30), and let $\mathcal{F} \subseteq \mathcal{C}$ be non-empty with dimension $K < D$. Then $\mathcal{F}$ is a K-face of $\mathcal{C}$ if and only if there exists $J \subset L$ with $|J| = D - K$ such that $\{\boldsymbol{g}_j\}_{j \in J}$ are linearly independent and*

$$\forall \boldsymbol{\varphi} \in \mathcal{F}, \forall j \in J, \quad \boldsymbol{g}_j \cdot \boldsymbol{\varphi} = c_j. \tag{II.2.31}$$

Note that there may be $\ell \in L \setminus J$ such that $\boldsymbol{g}_\ell \cdot \boldsymbol{\varphi} = c_\ell$. All such $\boldsymbol{g}_\ell$ are contained in the span of $\{\boldsymbol{g}_j\}_{j \in J}$, though.

PROOF. If such a J exists, then $\mathcal{F}' \equiv \{\boldsymbol{\varphi} \in \mathbb{R}^D \mid \forall j \in J, \boldsymbol{g}_j \cdot \boldsymbol{\varphi} = c_j\}$ is clearly an affine subspace of dimension K. Let $\boldsymbol{\varphi} \in \mathcal{F}'$ and $\boldsymbol{\eta}_{1,2} \in \mathcal{C}$, and assume $\boldsymbol{\varphi} = \mu \boldsymbol{\eta}_1 + (1 - \mu) \boldsymbol{\eta}_2$ with $\mu \in (0, 1)$. Then

$$\boldsymbol{g}_j \cdot \boldsymbol{\varphi} = d = \boldsymbol{g}_j \cdot \left[\mu \boldsymbol{\eta}_1 + (1 - \mu) \boldsymbol{\eta}_2 \right] \tag{II.2.32}$$

and since $\boldsymbol{g}_j \cdot \boldsymbol{\eta}_{1,2} \geq c_j$, this implies $\boldsymbol{g}_j \cdot \boldsymbol{\eta}_{1,2} = c_j$. Therefore $\boldsymbol{\eta}_{1,2} \in \mathcal{F}'$, so $\mathcal{F} = \mathcal{F}' \cap \mathcal{C}$ is a K-face.

For the less straightforward converse, let $\omega = \sum_{\ell \in L} \langle \boldsymbol{g}_\ell, c_\ell \rangle$, with $\mathcal{C} = \mathcal{B}(\omega)$. If ω is degenerate, then $\mathcal{C}$ clearly has no K-faces for $K > d_\omega$, and the unique d_ω-face is $\mathcal{C}$ itself by property (ii). In that case, the proof follows trivially from eq. (II.2.22). For the remaining cases, we may substitute $\omega \to \lfloor \omega \rfloor$, $\mathcal{C} \to \lfloor \mathcal{C} \rfloor_\omega = \mathcal{B}(\lfloor \omega \rfloor)$ and thus assume without loss of generality that ω is non-degenerate.

Let us then note that J cannot be empty. Indeed, if it were the case, then $\boldsymbol{\varphi} \in \mathcal{F}$ would be in the interior of $\mathcal{C}$ which via property (iii) contradicts the fact that $\mathcal{F}$ is a K-face with $K < D$. Also, $|J| \leq D - K$, since otherwise the set $\{\boldsymbol{\chi} \in \mathcal{C} \mid \forall j \in J, \boldsymbol{\chi} \cdot \boldsymbol{g}_j = c_j\}$, being the intersection of more than $D - K$ independent hyperplanes, would have dimension less than K. We will then proceed by induction on $D > K$. The result is trivial in $K + 1$ dimensions, since $|J| = 1$ is guaranteed by J being nonempty.

Assume then that the lemma holds in n dimensions, and consider $\mathcal{C}, \mathcal{F}$ in $(n+1)$-dimensional space. Since $J \neq \emptyset$, there is some $i \in L$ such that $\forall \boldsymbol{\varphi} \in \mathcal{F}, \boldsymbol{g}_i \cdot \boldsymbol{\varphi} = c_i$. Then consider

$$E_i \equiv \mathcal{B}\big(\langle \boldsymbol{g}_i, c_i \rangle + \langle -\boldsymbol{g}_i, -c_i \rangle\big), \qquad \lfloor \mathcal{C} \rfloor_{E_i} = \mathcal{B}\left(\sum_{\ell \in L} \lfloor \langle \boldsymbol{g}_\ell, c_\ell \rangle \rfloor_{E_i}\right). \tag{II.2.33}$$

By construction, $\lfloor \mathcal{F} \rfloor_{E_i}$ is a K-face of $\lfloor \mathcal{C} \rfloor_{E_i}$. Since these are sets in a n-dimensional space, we know by the induction hypothesis that we have J' with $|J'| = n - K$ such that $\{\lfloor \boldsymbol{g}_j \rfloor_{E_i}\}_{j \in J'}$ are linearly independent and eq. (II.2.31) is satisfied. Furthermore, $\lfloor \boldsymbol{g}_i \rfloor_{E_i} = \boldsymbol{0}$, which cannot be expressed as a linear combination of $\{\lfloor \boldsymbol{g}_j \rfloor_{E_i}\}_{j \in J'}$ with nonzero coefficients. Therefore, $\boldsymbol{g}_i$ is linearly independent of $\{\boldsymbol{g}_j\}_{j \in J'}$. Thus, $J = J' \cup \{i\}$ has $|J| = n + 1$, $\{\boldsymbol{g}_j\}_{j \in J}$ linearly indpendent and satisfies eq. (II.2.31). This proves that the lemma holds in $n + 1$ dimensions, and completes the induction. □

Based on this, we can prove two more interesting lemmata:

Lemma B.3. *Let $\mathcal{C}$ be a convex set satisfying eq. (II.2.30). Let $\mathcal{F}$ be a nonempty convex subset of $\mathcal{C}$, and let $K = \dim(\mathcal{F})$. Then $\mathcal{F}$ is a K-face of $\mathcal{C}$ if and only if there is some constraint $\langle \boldsymbol{h}, d \rangle$ such that $\mathcal{C} \subseteq \mathcal{B}(\langle \boldsymbol{h}, d \rangle)$ and $\{ \boldsymbol{\chi} \in \mathcal{C} \mid \boldsymbol{h} \cdot \boldsymbol{\chi} = d \} = \mathcal{F}.$*[45]

PROOF. First note that this trivially holds when $\{ \boldsymbol{\chi} \in \mathcal{C} \mid \boldsymbol{h} \cdot \boldsymbol{\chi} = d \} = \mathcal{C}$ (compare property (ii)), which may happen even when $\mathcal{C}$ is not contained in any hyperplane if $\langle \boldsymbol{h}, d \rangle = \langle \boldsymbol{0}, 0 \rangle$.

Setting aside the trivial cases, assume that such a $\langle \boldsymbol{h}, d \rangle$ exists. Then $\mathcal{F}$ is a K-face by essentially the same argument that was made around eq. (II.2.32).

Conversely, if $\mathcal{F}$ is a K-face then lemma B.2 holds. In the notation of that lemma, let

$$\boldsymbol{h} = \sum_{\ell \in J} \boldsymbol{g}_\ell, \qquad d = \sum_{\ell \in J} c_\ell, \qquad (\text{II}.2.34)$$

so that $\forall \boldsymbol{\varphi} \in \mathcal{F}, \boldsymbol{h} \cdot \boldsymbol{\varphi} = d$. Let us now consider any $\boldsymbol{\eta} \in \mathcal{C}$ such that $\boldsymbol{h} \cdot \boldsymbol{\eta} = d$. Let E_J be defined as in eq. (II.2.33). Then by construction, $\lfloor \mathcal{F} \rfloor_{E_J}$ has the same dimension as its native space, i.e. $D - K$, and has nonempty interior by eq. (II.2.26) (recall that $D - K > 0$). Thus, there exists $\boldsymbol{\varphi} \in \mathcal{F}$ such that $\lfloor \boldsymbol{\varphi} \rfloor_{E_J}$ is in the interior of $\lfloor \mathcal{F} \rfloor_{E_J}$. For $\varepsilon > 0$ small enough, we therefore have

$$\lfloor \boldsymbol{\eta}_\varepsilon \rfloor_{E_J} \equiv \varepsilon \lfloor \boldsymbol{\eta} \rfloor_{E_J} + (1 - \varepsilon) \lfloor \boldsymbol{\varphi} \rfloor_{E_J} \in \lfloor \mathcal{F} \rfloor_{E_J}. \qquad (\text{II}.2.35)$$

Since by construction $\lceil \boldsymbol{\chi} \rceil^{E_J} = \boldsymbol{0}$ for all $\boldsymbol{\chi}$ such that $\boldsymbol{h} \cdot \boldsymbol{\chi} = d$, lemma B.1 implies that

$$\boldsymbol{\eta}_\varepsilon = \varepsilon \boldsymbol{\eta} + (1 - \varepsilon) \boldsymbol{\varphi}. \qquad (\text{II}.2.36)$$

By the definition of a K-face, this implies that $\boldsymbol{\eta} \in \mathcal{F}$, which concludes the proof.[46] $\qquad \square$

Lemma B.4. *Let $\mathcal{C} \subseteq \mathcal{C}'$ be convex sets satisfying eq. (II.2.30), and let $\mathcal{F}$ be a K-face of $\mathcal{C}$. Then there exists a K-face $\mathcal{F}'$ of $\mathcal{C}'$ with $\mathcal{F} \subseteq \mathcal{F}'$ if there is some $\boldsymbol{\varphi} \in \mathcal{F}$ and $\varepsilon > 0$ such that $\{ \boldsymbol{\chi} \in \mathcal{C}' \mid \varepsilon \geq |\boldsymbol{\chi} - \boldsymbol{\varphi}| \} \subseteq \mathcal{C}$.*

Note how this complements property (v). Note also that $\mathcal{F}' \supset \mathcal{F}$ may exist even if $\mathcal{F}$ fails to satisfy the given conditions.

PROOF. By lemma B.3, there exists $\langle \boldsymbol{h}, d \rangle$ such that and $\{ \boldsymbol{\chi} \in \mathcal{C} \mid \boldsymbol{h} \cdot \boldsymbol{\chi} = d \} = \mathcal{F}$ and $\mathcal{C} \subseteq \mathcal{B}(\langle \boldsymbol{h}, d \rangle)$. We moreover claim that $\mathcal{C}' \subseteq \mathcal{B}(\langle \boldsymbol{h}, d \rangle)$. Indeed, let $\boldsymbol{\varphi} \in \mathcal{F}$ and $\varepsilon > 0$ such that $\{ \boldsymbol{\chi} \in \mathcal{C}' \mid \varepsilon \geq |\boldsymbol{\chi} - \boldsymbol{\varphi}| \} \subseteq \mathcal{C}$. For any $\boldsymbol{\eta} \in \mathcal{C}'$ it follows from the triangle inequality that

$$\left| \boldsymbol{\varphi} - \left[\left(1 - \tfrac{\varepsilon}{|\boldsymbol{\varphi}| + |\boldsymbol{\eta}|} \right) \boldsymbol{\varphi} + \tfrac{\varepsilon}{|\boldsymbol{\varphi}| + |\boldsymbol{\eta}|} \boldsymbol{\eta} \right] \right| = \varepsilon \frac{|\boldsymbol{\varphi} - \boldsymbol{\eta}|}{|\boldsymbol{\varphi}| + |\boldsymbol{\eta}|} \leq \varepsilon, \qquad (\text{II}.2.37)$$

---

[45]Although we do not use it, this holds for $K = D - 1$ for *any* convex set $\mathcal{C}$, not just those satisfying eq. (II.2.30) (this is proven in a later footnote). This is not the case for smaller K: take e.g. the $D = 2$ example $\mathcal{C} = \{ (x, y) \in \mathbb{R}^2 \mid y \geq \max(0, x^3) \}$, for which $(0, 0)$ is an extreme point but $\langle \boldsymbol{h}, d \rangle$ does not exist.

[46]When $K = D - 1$, the following proof, which does not use lemma B.2 and therefore holds for all convex sets $\mathcal{C}$, works for the converse: Let $\mathcal{F}'$ be the affine span of $\mathcal{F}$, which is a hyperplane. By a variant of the separating hyperplane theorem (for any disjoint convex sets $\mathcal{X}$ and $\mathcal{Y}$ (no closedness/compactness needed), there exists $\langle \boldsymbol{h}, d \rangle$ with $\boldsymbol{h} \cdot \boldsymbol{\chi} \leq d'$ for all $\boldsymbol{\chi} \in \mathcal{X}$ and $\boldsymbol{h} \cdot \boldsymbol{\psi} \geq d'$ for all $\boldsymbol{\psi} \in \mathcal{Y}$) applied to the disjoint convex sets $\mathcal{F}'$ and $\mathcal{C} \setminus \mathcal{F}$, we immediately find our desired $\langle \boldsymbol{h}, d \rangle$. This hinges on the properties of hyperplanes: the separating hyperplane must be parallel to $\mathcal{F}'$, or else they would intersect. Therefore, this does not work if $K < D - 1$.

so $\left(1 - \frac{\varepsilon}{|\varphi|+|\eta|}\right)\varphi + \frac{\varepsilon}{|\varphi|+|\eta|}\eta \in \mathcal{C}$. In particular,

$$\boldsymbol{h} \cdot \left[\left(1 - \tfrac{\varepsilon}{|\varphi|+|\eta|}\right)\varphi + \tfrac{\varepsilon}{|\varphi|+|\eta|}\eta\right] \geq d \quad \Rightarrow \quad \boldsymbol{h} \cdot \eta \geq d. \tag{II.2.38}$$

Therefore, lemma B.3 guarantees that $\mathcal{F}' = \{\eta \in \mathcal{C}' \mid \boldsymbol{h} \cdot \eta = d\}$ is a K'-face for some K'. Since $\mathcal{F} \subseteq \mathcal{F}'$ we have $K \leq K'$, so it remains to prove that $K \geq K'$. For any $\varphi' \in \mathcal{F}'$, we have $\varphi + \varepsilon\frac{\varphi'}{|\varphi'|} \in \mathcal{F}$ since $\left|\varphi - \left(\varphi + \varepsilon\frac{\varphi'}{|\varphi'|}\right)\right| \leq \varepsilon$. However, since $-\varphi$ belongs to the span of $\mathcal{F}$, we can deduce that $\varepsilon\frac{\varphi'}{|\varphi'|}$, and therefore also φ', belongs to the span of $\mathcal{F}$. Thus, the span of $\mathcal{F}'$ is included in the span of $\mathcal{F}$, implying that $K \geq K'$, which completes the proof. $\square$

B.3   Proof and generalisation of proposition 4.3

To simplify this proof, we will introduce the following terminology:

> A linear constraint $\langle \boldsymbol{\alpha}, c \rangle$ with $\boldsymbol{\alpha} \neq \boldsymbol{0}$ *supports* a point $\boldsymbol{b}$ if $\boldsymbol{\alpha} \cdot \boldsymbol{b} = c$.[47]

> A linear constraint $\langle \boldsymbol{0}, c \rangle$ supports no point (this is natural for $c = \pm 1$, but we define it to be so also for $c = 0$; this simplifies most statements expressed in terms of support).

> Given Ω, a linear constraint $\langle \boldsymbol{\alpha}, c \rangle$ *uniquely supports* a point $\boldsymbol{b} \in \mathcal{B}(\Omega)$ if it supports $\boldsymbol{b}$, and there is no other $\langle \boldsymbol{\beta}, d \rangle \leq \Omega$ with $\langle \boldsymbol{\beta}, d \rangle \neq \langle \boldsymbol{\alpha}, c \rangle$ that supports $\boldsymbol{b}$.

> A constraint representation $\mathcal{S}$ is *non-redundant* if it contains no trivial constraints, and there are no two elements $\langle \boldsymbol{\alpha}, c \rangle \in \mathcal{S}$ and $\langle \boldsymbol{\beta}, d \rangle \in \mathcal{S}$ such that $\langle \boldsymbol{\alpha}, c \rangle = \langle \boldsymbol{\beta}, d \rangle$.

Reducing a representation to a non-redundant one is of course trivial. Noting that the second paragraph of proposition 4.3 can be reduced to "The elements of $\mathcal{R}(\Omega)$ are exactly those $\langle \boldsymbol{\alpha}, c \rangle \leq \Omega$ that uniquely support a point $\boldsymbol{b} \in \mathcal{B}(\Omega)$", we will then begin with the following lemma:

Lemma B.5. *Let $\Omega = \sum_{i \in I} \langle \boldsymbol{\alpha}_i, c_i \rangle$ be a non-degenerate constraint, as defined in section 4.4, and let the representation $\{\langle \boldsymbol{\alpha}_i, c_i \rangle\}_{i \in I}$ be non-redundant. Then for any $j \in I$ and $\boldsymbol{b} \in \mathcal{B}(\Omega)$, $\langle \boldsymbol{\alpha}_j, c_j \rangle$ uniquely supports $\boldsymbol{b}$ if and only if $\boldsymbol{\alpha}_j \cdot \boldsymbol{b} = c_j$ and $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} > c_i$ for all $i \neq j$.*

Note that compared to the definition of unique support, this only concerns the elements of a non-redundant representation rather than all $\langle \boldsymbol{\beta}, c \rangle \leq \Omega$.

Proof. Assume that $\boldsymbol{\alpha}_j \cdot \boldsymbol{b} = c_j$ and $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} > c_i$ for all $i \neq j$, and assume there is some $\langle \boldsymbol{\beta}, c \rangle < \Omega$ that supports $\boldsymbol{b}$, i.e. that $\langle \boldsymbol{\alpha}_j, c_j \rangle$ does not support it uniquely. By proposition 4.2, there exist some positive numbers $\{\lambda_i\}_{i \in I}$ such that

$$c = \boldsymbol{\beta} \cdot \boldsymbol{b} = \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i \cdot \boldsymbol{b} \quad \Rightarrow \quad \begin{cases} c = \lambda_j c_j & \text{if } \lambda_i = 0 \text{ for all } i \neq j, \\ c > \sum_{i \in I} \lambda_i c_i \geq c & \text{otherwise.} \end{cases} \tag{II.2.39}$$

---

[47]This is inspired by the standard concept of a *supporting hyperplane*.

The second case is a contradiction, so $\boldsymbol{\beta} = \lambda_j \boldsymbol{\alpha}_j$ and $c = \lambda_j c_j$ with $\lambda_j \geq 0$. If $c_j = \pm 1$, then either $\boldsymbol{\beta} = -\boldsymbol{\alpha}_j$ and $c = -c_j$ (contradicting non-degeneracy), or $\boldsymbol{\beta} = \mathbf{0}$ (impossible since $\langle \mathbf{0}, \pm 1 \rangle$ supports no point), or $\boldsymbol{\beta} = \boldsymbol{\alpha}_j$ and $c = c_j$ (contradicting non-redundancy). If $c_j = 0$, then $c = 0$ and $\boldsymbol{\beta}$ is proportional to $\boldsymbol{\alpha}_j$ (again contradicts non-rendundancy, via eq. (II.4.5)). With no non-contradictory cases left, we have proven that $\langle \boldsymbol{\alpha}_j, c_j \rangle$ supports $\boldsymbol{b}$ uniquely if $\boldsymbol{\alpha}_j \cdot \boldsymbol{b} = c_j$ and $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} > c_i$ for all $i \neq j$. The converse is trivial. □

Let us then prove the following more significant lemma:

Lemma B.6. *Let $\langle \boldsymbol{\alpha}, c \rangle$ be a linear constraint, and let Ω be a constraint such that $\Omega + \langle \boldsymbol{\alpha}, c \rangle$ is non-degenerate. Assume that Ω uses a non-redundant representation $\mathcal{S}$ and that $\mathcal{S} \cup \{\langle \boldsymbol{\alpha}, c \rangle\}$ is a non-redundant representation for $\Omega + \langle \boldsymbol{\alpha}, c \rangle$. Then $\Omega + \langle \boldsymbol{\alpha}, c \rangle > \Omega$ if and only if $\langle \boldsymbol{\alpha}, c \rangle$ uniquely supports some point $\boldsymbol{b} \in \mathcal{B}[\Omega + \langle \boldsymbol{\alpha}, c \rangle]$.*

PROOF. Assume that $\langle \boldsymbol{\alpha}, c \rangle$ uniquely supports some point $\boldsymbol{b} \in \mathcal{B}(\Omega + \langle \boldsymbol{\alpha}, c \rangle)$, so that $\boldsymbol{\alpha} \cdot \boldsymbol{b} = c$ and $\boldsymbol{\beta} \cdot \boldsymbol{b} > d$ for all $\langle \boldsymbol{\beta}, d \rangle \in \mathcal{S}$. Then define

$$\boldsymbol{b}_\varepsilon \equiv \boldsymbol{b} - \varepsilon \frac{\boldsymbol{\alpha}}{|\boldsymbol{\alpha}|^2}, \quad \varepsilon > 0 \tag{II.2.40}$$

so that

$$\boldsymbol{\alpha} \cdot \boldsymbol{b}_\varepsilon = c - \varepsilon < c, \qquad \boldsymbol{\beta} \cdot \boldsymbol{b}_\varepsilon > d - \varepsilon \frac{\boldsymbol{\beta} \cdot \boldsymbol{\alpha}}{|\boldsymbol{\alpha}|^2}. \tag{II.2.41}$$

Thus, $\boldsymbol{b}_\varepsilon \notin \mathcal{B}(\Omega + \langle \boldsymbol{\alpha}, c \rangle)$, but $\boldsymbol{b}_\varepsilon \in \mathcal{B}(\Omega)$ for sufficiently small ε. Then by definition, $\Omega + \langle \boldsymbol{\alpha}, c \rangle > \Omega$.

Conversely, assume that $\Omega + \langle \boldsymbol{\alpha}, c \rangle > \Omega$. Then there exists some $\boldsymbol{b} \in \mathcal{B}(\Omega)$ such that $\boldsymbol{b} \notin \mathcal{B}(\Omega + \langle \boldsymbol{\alpha}, c \rangle)$. Furthermore, since $\Omega + \langle \boldsymbol{\alpha}, c \rangle$ is non-degenerate, the interior of $\mathcal{B}(\Omega + \langle \boldsymbol{\alpha}, c \rangle)$ is non-empty;[48] therefore, it contains some point $\boldsymbol{n}$. These points have the properties

$$\begin{aligned} \boldsymbol{\alpha} \cdot \boldsymbol{b} < c, \qquad \boldsymbol{\beta} \cdot \boldsymbol{b} \geq d, \\ \boldsymbol{\alpha} \cdot \boldsymbol{n} > c, \qquad \boldsymbol{\beta} \cdot \boldsymbol{n} > d. \end{aligned} \tag{II.2.42}$$

where again $\langle \boldsymbol{\beta}, d \rangle \in \mathcal{S}$. By the intermediate value theorem, there must therefore exist some $\mu \in (0, 1)$ such that[49]

$$\boldsymbol{\alpha} \cdot \big[\mu \boldsymbol{b} + (1 - \mu)\boldsymbol{n}\big] = c, \qquad \boldsymbol{\beta} \cdot \big[\mu \boldsymbol{b} + (1 - \mu)\boldsymbol{n}\big] > d, \tag{II.2.43}$$

which, through lemma B.5, proves that $\langle \boldsymbol{\alpha}, c \rangle$ supports $\mu \boldsymbol{b} + (1 - \mu)\boldsymbol{n}$ uniquely. □

We now move on to proving proposition 4.3. Let a non-degenerate constraint Ω be expressed as

$$\Omega = \sum_{i \in I} \langle \boldsymbol{\alpha}_i, c_i \rangle = \sum_{j \in J} \langle \boldsymbol{\gamma}_j, d_j \rangle \tag{II.2.44}$$

---

[48]Recall that Ω being degenerate is equivalent to $\mathcal{B}(\Omega)$ being contained in a hyperplane, and a convex set has empty interior if and only if it is contained in a hyperplane; see also eq. (II.2.26). In general, the existence of a point in the interior of $\mathcal{B}(\Omega)$ is the *only* property of non-degenerate constraints used in this proof. Circumventing this requirement is key to proposition B.1 below.

[49]Specifically, $\mu = \dfrac{\boldsymbol{\alpha} \cdot \boldsymbol{n} - c}{\boldsymbol{\alpha} \cdot \boldsymbol{n} - \boldsymbol{\alpha} \cdot \boldsymbol{b}}$.

where $\{\langle \boldsymbol{\alpha}_i, c_i \rangle\}_{i \in I}$ is any representation, whereas $\{\langle \boldsymbol{\gamma}_j, d_j \rangle\}_{j \in J}$ is minimal. A minimal representation must exist, since Ω can be written as a sum of a finite number of constraints. For any $k \in J$, minimality implies that

$$\sum_{j \in J} \langle \boldsymbol{\gamma}_j, d_j \rangle > \sum_{\substack{j \in J \\ j \neq k}} \langle \boldsymbol{\gamma}_j, d_j \rangle, \tag{II.2.45}$$

but by lemma B.6, $\langle \boldsymbol{\gamma}_k, d_k \rangle$ must then uniquely support some $\boldsymbol{b} \in \mathcal{B}(\Omega)$. Then, there must also be some $i_k \in I$ such that $\langle \boldsymbol{\alpha}_{i_k}, c_{i_k} \rangle$ also supports $\boldsymbol{b}$. To see this, assume that no $\langle \boldsymbol{\alpha}_i, c_i \rangle, i \in I$ supports $\boldsymbol{b}$, and consider

$$\boldsymbol{b}_\varepsilon \equiv \boldsymbol{b} - \varepsilon \frac{\boldsymbol{\gamma}_k}{|\boldsymbol{\gamma}_k|^2} \quad \Rightarrow \quad \boldsymbol{\alpha}_i \cdot \boldsymbol{b}_\varepsilon > c - \varepsilon \frac{\boldsymbol{\alpha} \cdot \boldsymbol{\gamma}_j}{|\boldsymbol{\gamma}_j|^2}, \tag{II.2.46}$$

so that for sufficiently small ε, $\boldsymbol{\alpha}_i \cdot \boldsymbol{b}_\varepsilon > c$ for all $i \in I$, and thus $\boldsymbol{b}_\varepsilon \in \mathcal{B}(\Omega)$. On the other hand,

$$\boldsymbol{\gamma}_k \cdot \boldsymbol{b}_\varepsilon = d_k - \varepsilon \tag{II.2.47}$$

implying $\boldsymbol{b} \notin \mathcal{B}(\Omega)$, a contradiction. Therefore, $i_k \in I$ does exist. By the definition of unique support, we must then have $\langle \boldsymbol{\gamma}_k, d_k \rangle = \langle \boldsymbol{\alpha}_{i_k}, c_{i_k} \rangle$; that is, they are identical up to normalisation. By repeating this argument, we see that a distinct i_k exists for each $k \in J$, so $\{\langle \boldsymbol{\gamma}_j, d_j \rangle\}_{j \in J} \subseteq \{\langle \boldsymbol{\alpha}_i, c_i \rangle\}_{i \in I}$ up to normalisation. Having shown this, we may without loss of generality normalise and re-index $\langle \boldsymbol{\gamma}_j, d_j \rangle$ so that $J \subseteq I$. Then

$$\begin{aligned} \Omega &= \sum_{i \in I} \langle \boldsymbol{\alpha}_i, c_i \rangle \\ &= \sum_{i \in J} \langle \boldsymbol{\alpha}_i, c_i \rangle + \sum_{i \in I \setminus J} \langle \boldsymbol{\alpha}_i, c_i \rangle \\ &= \Omega + \sum_{i \in I \setminus J} \langle \boldsymbol{\alpha}_i, c_i \rangle. \end{aligned} \tag{II.2.48}$$

By lemma B.6, the sum in the last line can only contain constraints that do not uniquely support any point. Thus, the minimal representation $\mathcal{R}(\Omega) = \{\langle \boldsymbol{\alpha}_i, c_i \rangle\}_{i \in J}$ consists (up to normalisation) of *exactly* those elements of any representation that uniquely support a point. From this it follows that $\mathcal{R}(\Omega)$ consists of exactly all those $\langle \boldsymbol{\gamma}, d \rangle \leq \Omega$ that uniquely support a point, since $\{\langle \boldsymbol{\alpha}_i, c_i \rangle\}_{i \in I}$ could be made to include all $\langle \boldsymbol{\alpha}, c \rangle \leq \Omega$, and from that it follows that $\mathcal{R}(\Omega)$ is unique up to normalisation. ∎

B.3.1 The treatment of degenerate constraints

The following result generalises proposition 4.3 to all Ω:

Proposition B.1 (finding relevant constraints, general case). *Let Ω be any constraint in D-dimensional space. Then a minimal representation $\mathcal{R}(\Omega)$ can be determined as follows:*

(i) *If Ω is non-degenerate, then proposition 4.3 applies. $\mathcal{R}(\Omega)$ is therefore unique up to normalisation, and is a subset of any representation of Ω.*

(ii) If $\Omega = \Omega_\infty$, then trivially $\mathcal{R}(\Omega) = \{\langle \mathbf{0}, 1 \rangle\}$. This is unique, but not necessarily a subset of other representations.

(iii) If Ω is degenerate, let E_Ω be the unique d_Ω-dimensional affine subspace that contains $\mathcal{B}(\Omega)$. Let $\left\{ \langle \boldsymbol{\alpha}_i, c_i \rangle \right\}_{i \in I}$ be any set such that $\left\{ \lfloor \langle \boldsymbol{\alpha}_i, c_i \rangle \rfloor_\Omega \right\}_{i \in I} = \mathcal{R}(\lfloor \Omega \rfloor)$, with $\lfloor \langle \boldsymbol{\alpha}_i, c_i \rangle \rfloor_\Omega \neq \lfloor \langle \boldsymbol{\alpha}_j, c_j \rangle \rfloor_\Omega$ for all $i \neq j$. Let $\{ \boldsymbol{\sigma}_k \}_{k=0}^{d_\Omega}$ be any set of $(D - d_\Omega + 1)$ vectors with the following properties:

- *$\lfloor \boldsymbol{\sigma}_k \rfloor_{E_\Omega} = \mathbf{0}$;*
- *The dimension of $\mathrm{span} \left(\{ \boldsymbol{\sigma}_k \}_{k=0}^{D-d_\Omega} \right)$ is $D - d_\Omega$;*
- *There exists a solution to*

$$\mathbf{0} = \sum_{k=0}^{D-d_\Omega} \lambda_k \boldsymbol{\sigma}_k, \qquad \lambda_k > 0. \tag{II.2.49}$$

Then for arbitrary $\boldsymbol{e} \in E_\Omega$,

$$\mathcal{R}(\Omega) = \left\{ \langle \boldsymbol{\alpha}_i, c_i \rangle \right\}_{i \in I} \cup \left\{ \langle \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_k \cdot \boldsymbol{e} \rangle \right\}_{k=0}^{D-d_\Omega}. \tag{II.2.50}$$

This $\mathcal{R}(\Omega)$ is generally not unique, and is not neccesarily a subset of any given representation of Ω. However, there is no minimal representation of Ω that is not of this form.

The non-unqiueness in case (iii) comes about in two ways. Firstly, $\lceil \boldsymbol{\alpha} \rceil^{E_\Omega}$ is arbitrary for $\lfloor \langle \boldsymbol{\alpha}, c \rangle \rfloor_\Omega \in \mathcal{R}(\lfloor \Omega \rfloor)$. Secondly, there is clearly freedom in the choice of $\{ \boldsymbol{\sigma}_k \}_{k=0}^{D-d_\Omega}$. Given $\{ \boldsymbol{\delta}_k \}_{k=1}^{D-d_\Omega}$ as defined above eq. (II.2.22), a straightforward choice is

$$\{ \boldsymbol{\sigma}_k \}_{k=1}^{D-d_\Omega} = \{ \boldsymbol{\delta}_k \}_{k=1}^{D-d_\Omega}, \qquad -\boldsymbol{\sigma}_0 = \sum_{k=1}^{D-d_\Omega} \boldsymbol{\sigma}_k. \tag{II.2.51}$$

Proof. We only need to prove case (iii). For brevity, we will omit some sub/superscripts: $\lfloor \langle \boldsymbol{\alpha}, c \rangle \rfloor$ should be read as $\lfloor \langle \boldsymbol{\alpha}, c \rangle \rfloor_\Omega$, $\lceil \boldsymbol{b} \rceil$ as $\lceil \boldsymbol{b} \rceil^{E_\Omega}$, and so on.

Let $\left\{ \langle \boldsymbol{\alpha}_i, c_i \rangle \right\}_{i \in J}$ be a minimal representation of Ω, and subdivide it as

$$J_1 \equiv \{ i \in J \mid \lfloor \langle \boldsymbol{\alpha}_i, c_i \rangle \rfloor \neq \langle \mathbf{0}, 0 \rangle \}, \qquad J_2 \equiv \{ i \in J \mid \lfloor \langle \boldsymbol{\alpha}_i, c_i \rangle \rfloor = \langle \mathbf{0}, 0 \rangle \}, \tag{II.2.52}$$

for which the following holds:

Lemma B.7. $\{ \lfloor \langle \boldsymbol{\alpha}_i, c_i \rangle \rfloor \}_{i \in J_1} = \mathcal{R}(\lfloor \Omega \rfloor)$ *(up to normalisation).*

Proof. By proposition 4.3, it is clear that $\{ \lfloor \langle \boldsymbol{\alpha}_i, c_i \rangle \rfloor \}_{i \in J} \supseteq \mathcal{R}(\lfloor \Omega \rfloor)$, since it it a representation of $\lfloor \Omega \rfloor$. Then, for some $i \in J$, consider

$$\Omega' \equiv \sum_{\substack{j \in J \\ j \neq i}} \langle \boldsymbol{\alpha}_j, c_j \rangle. \tag{II.2.53}$$

Let us assume that $\lfloor\langle\boldsymbol{\alpha}_i, c_i\rangle\rfloor \notin \mathcal{R}(\lfloor\Omega\rfloor)$, so that $\lfloor\Omega'\rfloor = \lfloor\Omega\rfloor$. However, $\Omega' \neq \Omega$, since $\{\langle\boldsymbol{\alpha}_i, c_i\rangle\}_{i\in J}$ is minimal. Thus, by lemma B.1, Ω' must be satisfied by some point $\boldsymbol{a} \notin E_\Omega$. For each $\boldsymbol{b} \in \mathcal{B}(\Omega)$, consider then $x\boldsymbol{a} + (1 - x)\boldsymbol{b}$ for $x \in (0, 1]$. This point satisfies Ω' but not Ω, since it lies outside E_Ω. From this, we conclude that it does not satisfy $\langle\boldsymbol{\alpha}_i, c_i\rangle$. Thus, the continuous function

$$f_{\boldsymbol{b}}(x) = \boldsymbol{\alpha}_i \cdot [x\boldsymbol{a} + (1 - x)\boldsymbol{b}] \tag{II.2.54}$$

has $f_{\boldsymbol{b}}(x) < c_i$ for $x > 0$. However, $f_{\boldsymbol{b}}(0) \geq c_i$ since $\boldsymbol{b}$ satisfies $\langle\boldsymbol{\alpha}_i, c_i\rangle$, and this is only consistent with continuity if $f_{\boldsymbol{b}}(0) = c_i$, i.e. that $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} = c_i$, for all $\boldsymbol{b} \in \mathcal{B}(\Omega)$. Then $\mathcal{B}(\Omega) \subseteq \mathcal{B}(\langle\boldsymbol{\alpha}_i, c_i\rangle + \langle-\boldsymbol{\alpha}_i, -c_i\rangle)$, so it follows from eq. (II.2.22) that $\boldsymbol{\alpha}_i$ is a linear combination of $\{\boldsymbol{\delta}_k\}_{k=1}^{D-d_\Omega}$, i.e. that $\lfloor\boldsymbol{\alpha}_i\rfloor = \boldsymbol{0}$. Equation (II.2.27) then shows that $i \in J_2$, completing the proof. $\qquad\square$

Now, let

$$\Omega = \sum_{i\in J_1} \langle\boldsymbol{\alpha}_i, c_i\rangle + \sum_{i\in J_2} \langle\boldsymbol{\alpha}_i, c_i\rangle. \tag{II.2.55}$$

Lemma B.7 connects the J_1 part with $\mathcal{R}(\lfloor\Omega\rfloor)$, so it remains to study the J_2 part. We claim that eq. (II.2.55) holds true if and only if

$$\mathcal{B}\left(\sum_{i\in J_2} \lceil\langle\boldsymbol{\alpha}_i, c_i\rangle\rceil\right) = \{\boldsymbol{0}\}. \tag{II.2.56}$$

It follows immediately from lemma B.7 that

$$\boldsymbol{0} \in \mathcal{B}\left(\sum_{i\in J_2} \lceil\langle\boldsymbol{\alpha}_i, c_i\rangle\rceil\right). \tag{II.2.57}$$

Assume then that this set also contains some $\lceil\boldsymbol{v}\rceil \neq \boldsymbol{0}$. Since $\lfloor\Omega\rfloor$ is non-degenerate, there exists $\lfloor\boldsymbol{b}\rfloor \in \text{int}\left[\mathcal{B}(\lfloor\Omega\rfloor)\right]$; choosing $\lceil\boldsymbol{b}\rceil = \boldsymbol{0}$, it follows that $\forall i \in J_1$, $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} > c_i$. Therefore, for $\varepsilon > 0$ small enough we have $\forall i \in J_1$, $\boldsymbol{\alpha}_i \cdot (\boldsymbol{b} + \varepsilon\boldsymbol{v}) \geq c_i$. Since $\lceil\boldsymbol{b} + \varepsilon\boldsymbol{v}\rceil = \varepsilon\lceil\boldsymbol{v}\rceil$, we also have $\forall i \in J_2$, $\boldsymbol{\alpha}_i \cdot (\boldsymbol{b} + \varepsilon\boldsymbol{v}) \geq c_i$. Thus, $\boldsymbol{b} + \varepsilon\boldsymbol{v} \in \mathcal{B}(\Omega)$, but since $\lceil\boldsymbol{v}\rceil \neq \boldsymbol{0}$, we have $\boldsymbol{b} + \varepsilon\boldsymbol{v} \notin E_\Omega$, a contradiction. Along with its trivial converse, this proves the equivalence between eqs. (II.2.55) and (II.2.56).

What is then the minimal set $\{\langle\boldsymbol{\alpha}_i, c_i\rangle\}_{i\in J_2}$ that produces eq. (II.2.56)? Since $\{\boldsymbol{0}\}$ is a bounded set, corollary B.1 states that $\{\lceil\boldsymbol{\alpha}_i\rceil\}_{i\in J_2}$ spans the full $(D - d_\Omega)$-dimensional space, and that there are $\lambda_i > 0, i \in J_2$ such that

$$\boldsymbol{0} = \sum_{i\in J_2} \lambda_i\boldsymbol{\alpha}_i. \tag{II.2.58}$$

The span condition requires $|J_2| \geq D - d_\Omega$, but in order for there to be a nontrivial linear combination equal to zero, we must in fact have $|J_2| \geq D - d_\Omega + 1$. It is easy to see that this bound is sufficient (for details, see the proof of corollary B.1), so the minimal set must have $|J_2| = D - d_\Omega + 1$. Identifying $\{\boldsymbol{\alpha}_i\}_{i\in J_2}$ with $\{\boldsymbol{\sigma}_k\}_{k=0}^{D-d_\Omega}$, we see that we have just derived all conditions stated in the proposition, so the proof is complete. $\qquad\blacksquare$

B.3.2 An important corollary

The following interesting result, which is also our first use of the K-faces defined in section B.2.2, is a consequence of proposition 4.3:

Corollary B.2 (facet supported by relevant element). *Let $\Omega \neq \Omega_\infty$ be a non-degenerate constraint, and let $\langle \boldsymbol{\alpha}, c \rangle \leq \Omega$. Then $\langle \boldsymbol{\alpha}, c \rangle \in \mathcal{R}(\Omega)$ (up to normalisation) if and only if the set $\mathcal{F} \equiv \{ \boldsymbol{b} \in \mathcal{B}(\Omega) \mid \boldsymbol{\alpha} \cdot \boldsymbol{b} = c \}$ is a facet of $\mathcal{B}(\Omega)$.*[50]

PROOF. It is guaranteed via lemma B.3 that $\mathcal{F}$ is a K-face; we only need to show that $K = D - 1$ so that it is a facet.

Assume $\langle \boldsymbol{\alpha}, c \rangle \in \mathcal{R}(\Omega)$. Let $\langle \boldsymbol{\beta}, d \rangle$ be any constraint that supports all of $\mathcal{F}$, and construct the constraints

$$\Omega' \equiv \sum_{\langle \boldsymbol{\alpha}', c' \rangle \in \mathcal{R}(\Omega)} \langle \boldsymbol{\alpha}', c' \rangle + \langle \boldsymbol{\beta}, d \rangle, \qquad \Omega'' \equiv \sum_{\langle \boldsymbol{\alpha}', c' \rangle \in \mathcal{R}(\Omega)} \langle \boldsymbol{\alpha}', c' \rangle + \langle -\boldsymbol{\beta}, d \rangle. \quad \text{(II.2.59)}$$

Then it is clear that $\mathcal{B}(\Omega) = \mathcal{B}(\Omega') \cup \mathcal{B}(\Omega'')$, so at least one of Ω', Ω'' must be non-degenerate. Without loss of generality, assume Ω' is non-degenerate. Since $\langle \boldsymbol{\alpha}, c \rangle$ does not uniquely support any point in $\mathcal{B}(\Omega')$ (all points supported by it are also supported by $\langle \boldsymbol{\beta}, d \rangle$), proposition 4.3 (or rather lemma B.6) gives

$$\Omega' = \sum_{\substack{\langle \boldsymbol{\alpha}', c' \rangle \in \mathcal{R}(\Omega) \\ \langle \boldsymbol{\alpha}', c' \rangle \neq \langle \boldsymbol{\alpha}, c \rangle}} \langle \boldsymbol{\alpha}', c' \rangle + \langle \boldsymbol{\beta}, d \rangle. \quad \text{(II.2.60)}$$

Recall that $\langle \boldsymbol{\alpha}, c \rangle$, viewed as an element of Ω, uniquely supports some point $\boldsymbol{b} \in \mathcal{F}$. Therefore, $\boldsymbol{\alpha}' \cdot \boldsymbol{b} > c'$ for all $\langle \boldsymbol{\alpha}', c' \rangle \in \mathcal{R}(\Omega) \setminus \{\langle \boldsymbol{\alpha}, c \rangle\}$. By eq. (II.2.60), $\langle \boldsymbol{\beta}, d \rangle$ then uniquely supports $\boldsymbol{b} \in \mathcal{B}(\Omega')$, but since $\langle \boldsymbol{\alpha}, c \rangle \leq \Omega'$ and $\boldsymbol{\alpha} \cdot \boldsymbol{b} = c$, the definition of unique support gives $\langle \boldsymbol{\alpha}, c \rangle = \langle \boldsymbol{\beta}, d \rangle$. Thus, the only constraints that support all of $\mathcal{F}$ are, up to normalisation, $\langle \boldsymbol{\alpha}, c \rangle$ and $\langle -\boldsymbol{\alpha}, -c \rangle$, confirming via lemma B.2 that $\dim(\mathcal{F}) = D - 1$ (compare also eq. (II.2.22)). The converse is trivial. □

B.4 Practical construction of $\mathcal{A}_c(\Omega)$ and $\mathcal{R}(\Omega)$

This section describes how to leverage propositions 4.2 and 4.3 for the practical management of linear constraints. The results stated here, namely propositions B.2 and B.3 along with their corollaries, double as the algorithms which we used in practice to obtain the results presented in section 6.[51]

---

[50]It may seem obvious that this condition is equivalent to uniquely supporting a point, but is in fact rather subtle, and crucially depends on Ω having a finite representation. For instance, if $\mathcal{B}(\Omega)$ were a closed unit ball, then every point $\boldsymbol{b}$ on its surface would be uniquely supported by the constraint $\langle -\boldsymbol{b}, -1 \rangle$. For such a constraint, $\mathcal{F} = \{\boldsymbol{b}\}$, which is not a facet in $D > 1$ dimensions.

[51]The implementation code is available from Mattias Sjö upon request.

B.4.1 Construction of $\mathcal{A}_c(\Omega)$

When using proposition 4.2 to determine if $\langle \boldsymbol{\beta}, c \rangle \leq \Omega$ for some $\boldsymbol{\beta}$, c and Ω, the rather indirect definition in eq. (II.4.10) is of little practical use. Instead, we will take the approach of finding a constraint that is satisfied by $\boldsymbol{\beta}$ if and only if $\boldsymbol{\beta} \in \mathcal{A}_c(\Omega)$.

To understand why such a constraint exists and has a finite representation, note that eq. (II.4.11) along with $\lambda_i \geq 0$ are nothing more than an obfuscated set of linear constraints on the set $\mathcal{A}(\Omega)$. Provided a finite set of $\boldsymbol{\alpha}_i$, a considerable amount of linear algebra will determine a finite representation this way. Here, however, we present a simpler method in which the only complicated operation is the determination of the convex hull of a set of points. Highly efficient algorithms for determining convex hulls exist; we use the Quick-Hull algorithm [48] and the associated `qhull` implementation.[52]

The key to the construction is that, given a finite set of points $\{\boldsymbol{\beta}_j\}_{j \in J}$, a side-effect of the QuickHull algorithm is the creation of a set of constraints $\{\langle \boldsymbol{n}_\ell, r_\ell \rangle\}_{\ell \in L}$ such that

$$\mathcal{B}\left(\sum_{\ell \in L} \langle \boldsymbol{n}_\ell, r_\ell \rangle\right) = \mathrm{Hull}\left(\{\boldsymbol{\beta}_j\}_{j \in J}\right), \tag{II.2.61}$$

since $\boldsymbol{n}_\ell, r_\ell$ are the normals and offsets of the facets of the hull.[53] This is, by construction, a minimal representation. If we choose $\boldsymbol{\beta}_j$ such that $\mathrm{Hull}\left(\{\boldsymbol{\beta}_j\}_{j \in J}\right)$ is a suitable subset of $\mathcal{A}_c(\Omega)$, we will see that it is possible to write a simple rule that selects a subset $M \subseteq L$ such that

$$\mathcal{B}\left(\sum_{\ell \in M} \langle \boldsymbol{n}_\ell, r_\ell \rangle\right) = \mathcal{A}_c(\Omega). \tag{II.2.62}$$

In order to do this, let Ω be given as in eq. (II.4.9). Importantly, assume without loss of generality that $\langle \boldsymbol{\alpha}_i, -1 \rangle = \langle \mathbf{0}, -1 \rangle$ for some $i \in I_{-1}$, but that $\boldsymbol{\alpha}_i \neq \mathbf{0}$ for all $i \in I_0$.[54] Then, define the following sets of points:

$$
\begin{aligned}
\mathcal{P}_\pm(\Omega) &= \{\boldsymbol{\alpha}_i \mid i \in I_{\pm 1}\}, \\
\mathcal{Z}_\pm^{(p)}(\Omega) &= \{\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j \mid i \in I_{\pm 1}, j \in I_0\}, \\
\mathcal{N}_\pm^{(p)}(\Omega) &= \{p\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j \mid i \in I_{\pm 1}, j \in I_{\mp 1}\}
\end{aligned}
\tag{II.2.63}
$$

for integer $p > 1$. Also define

$$\mathcal{H}_\pm^{(p)}(\Omega) = \mathrm{Hull}\left[\mathcal{P}_\pm(\Omega) \cup \mathcal{Z}_\pm^{(p)}(\Omega) \cup \mathcal{N}_\pm^{(p)}(\Omega)\right], \tag{II.2.64}$$

where Hull denotes the convex hull; see eq. (II.4.8). From now on, we will often drop the "(Ω)" for brevity. In terms of these, we have the following result:

---

[52]For up-to-date information about `qhull`, see http://www.qhull.org/.

[53]`qhull` uses a different sign convention, but the conversion to the format given above is trivial.

[54]For maximum efficiency, the representation used for Ω should otherwise have as few elements as possible. The best easily accessible one is

$$\mathcal{R}\left(\sum_{i \in I_1} \langle \boldsymbol{\alpha}_i, c_i \rangle\right) \cup \mathcal{R}\left(\sum_{i \in I_0} \langle \boldsymbol{\alpha}_i, c_i \rangle\right) \cup \mathcal{R}\left(\sum_{i \in I_{-1}} \langle \boldsymbol{\alpha}_i, c_i \rangle\right) \cup \langle \mathbf{0}, -1 \rangle,$$

where the minimal representations are determined with proposition B.3 applied to corollary B.4.

Proposition B.2 (constructing $\mathcal{A}_c(\Omega)$). *Let Ω be a constraint, and arbitrarily select an integer $p > 1$. Construct a minimal representation $\{\langle \boldsymbol{n}_\ell, r_\ell \rangle\}_{\ell \in L_\pm^{(p)}}$ such that*[55]

$$\mathcal{B}\left[\sum_{\ell \in L_\pm^{(p)}} \langle \boldsymbol{n}_\ell, r_\ell \rangle \right] = \mathcal{H}_\pm^{(p)}(\Omega). \tag{II.2.65}$$

with $\mathcal{H}_\pm^{(p)}(\Omega)$ defined as in eq. (II.2.64). Let $M_\pm^{(p)} \subseteq L_\pm^{(p)}$ be the set of all ℓ for which $\langle \boldsymbol{n}_\ell, r_\ell \rangle$ supports at least one point $\boldsymbol{\pi} \in \mathcal{P}_\pm$. Then

$$\mathcal{B}\left[\sum_{\ell \in M_\pm^{(p)}} \langle \boldsymbol{n}_\ell, r_\ell \rangle \right] = \mathcal{A}_{\pm 1}(\Omega), \qquad \mathcal{B}\left[\sum_{\ell \in M_\pm^{(p)}} \langle \boldsymbol{n}_\ell, 0 \rangle \right] = \mathcal{A}_0(\Omega). \tag{II.2.66}$$

Note that $\mathcal{A}_0(\Omega)$ can be constructed from either $M_+^{(p)}$ or $M_-^{(p)}$. The exception is when $I_1 = \emptyset$, in which case the construction of $\mathcal{A}_1(\Omega)$ fails; $\mathcal{A}_1(\Omega) = \emptyset$ trivially, and $\mathcal{A}_0(\Omega)$ can only be constructed from $M_-^{(p)}$.

This result (along with corollaries B.3 and B.4 and proposition B.3 below) outlines the procedure we use in practice to obtain minimal representations. An example of this construction for $c = -1$ is given in fig. II.26.

PROOF. For brevity, we will write

$$\mho_\pm^{(p)} \equiv \sum_{\ell \in M_\pm^{(p)}} \langle \boldsymbol{n}_\ell, r_\ell \rangle. \tag{II.2.67}$$

The goal is then to show that $\mathcal{B}(\mho_\pm^{(p)}) = \mathcal{A}_{\pm 1}(\Omega)$, which we will do by showing that $\mathcal{B}(\mho_\pm^{(p)}) \supseteq \mathcal{A}_{\pm 1}(\Omega)$ followed by $\mathcal{B}(\mho_\pm^{(p)}) \subseteq \mathcal{A}_{\pm 1}(\Omega)$. We will then show the $c = 0$ case as a consequence of the others. First, however, we will establish some lemmata.

Lemma B.8. $r_\ell \geq 0$ *for all $\ell \in M_+^{(p)}$, and $r_\ell \leq 0$ for all $\ell \in M_-^{(p)}$.*

PROOF. The latter inequality is trivial, since $\boldsymbol{0} \in \mathcal{H}_-^{(p)}$. In the $M_+^{(p)}$ case, recall that by definition, $\boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_{i_\ell} = r_\ell$ for some $i_\ell \in I_1$. We must also have $r_\ell \leq \boldsymbol{n}_\ell \cdot (p\boldsymbol{\alpha}_{i_\ell} + \boldsymbol{0}) = pr_\ell$ since $(p\boldsymbol{\alpha}_{i_\ell} + \boldsymbol{0}) \in \mathcal{N}_+^{(p)}$, but $r_\ell \leq pr_\ell$ for $p > 1$ implies that $r_\ell \geq 0$. $\square$

Lemma B.9. $\boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_i \geq c_i |r_\ell|$ *for all $i \in I, \ell \in M_\pm^{(p)}$.*

PROOF. For $\ell \in M_\pm^{(p)}$, let $i_\ell \in I_{\pm 1}$ be such that $\boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_{i_\ell} = r_\ell$. Then consider two specific points in $\mathcal{Z}_\pm^{(p)}$ and $\mathcal{N}_\pm^{(p)}$:

$$\forall j \in I_0, \qquad \boldsymbol{n}_\ell \cdot (\boldsymbol{\alpha}_{i_\ell} + (p-1)\boldsymbol{\alpha}_j) \geq r_\ell \quad \Rightarrow \quad \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_j \geq 0 \tag{II.2.68}$$

---

[55]To avoid clutter, we do not indicate any p-dependence on $\langle \boldsymbol{n}_\ell, r_\ell \rangle$, but one should bear in mind that they may be entirely different constraints for different p (and different $\pm$). To remember this, it can be useful to think of $L_\pm^{(p)}$ as disjoint sets for different $p, \pm$.
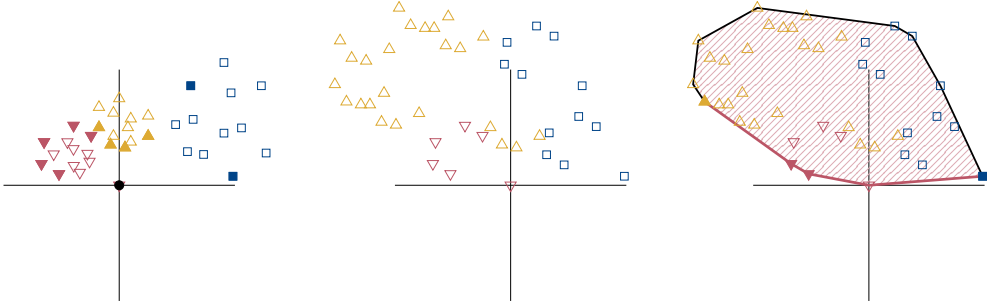
Figure II.26: Example of the construction of $\mathcal{A}_{-1}(\Omega)$ as described in proposition B.2. **Left:** Three sets of points $\boldsymbol{\alpha}_i$ for $i \in I_{-1}$ (▼), $i \in I_0$ (■) and $i \in I_1$ (▲). The sets are the same up to rescaling and rotation about the origin, and the addition of $\langle \mathbf{0}, -1 \rangle$. The relevant elements of the respective ω_c are marked as filled points, and the rest are left empty; compare to the similar sets in fig. II.2. **Middle:** The sets $\mathcal{P}_-$ (▽), $\mathcal{Z}_-^{(2)}$ (□) and $\mathcal{N}_-^{(2)}$ (△) constructed using only the relevant elements, as remarked above eq. (II.2.63). **Right:** The convex hull thereof (red), i.e. $\mathcal{H}_-^{(2)}$. The segments of its boundary correspond to $\langle \boldsymbol{n}_\ell, r_\ell \rangle, \ell \in L_-^{(2)}$. The ones for which $\ell \in M_-^{(2)}$ are highlighted in red. Removing the other segments leaves behind the unbounded region $\mathcal{A}_{-1}(\Omega)$, which can be seen in fig. II.27 below. The relevant elements of Ω have been marked as filled points. They follow from proposition B.3, although it can be intuitively seen that they alone influence the shape of $\mathcal{A}_{-1}(\Omega)$.

$$\forall j' \in I_{\mp 1}, \quad \boldsymbol{n}_\ell \cdot (p\boldsymbol{\alpha}_{i_\ell} + (p-1)\boldsymbol{\alpha}_{j'}) \geq r_\ell \quad \Rightarrow \quad \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_{j'} \geq \frac{1-p}{p-1} r_\ell = -r_\ell. \quad \text{(II.2.69)}$$

This, together with lemma B.8, implies lemma B.9. $\qquad\qquad\square$

Now for the main proof. Let $\boldsymbol{\beta} \in \mathcal{A}_{\pm 1}(\Omega)$ (note that this excludes the exceptional case $\mathcal{A}_1(\Omega) = \emptyset$). Using eq. (II.4.12), it can therefore be written

$$\boldsymbol{\beta} = \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i \geq 0, \quad \sum_{i \in I} \lambda_i c_i \geq c, \quad \text{(II.2.70)}$$

so for all $\ell \in M_\pm^{(p)}$, lemma B.9 gives

$$\boldsymbol{n}_\ell \cdot \boldsymbol{\beta} \geq |r_\ell| \sum_{i \in I} \lambda_i c_i \geq |r_\ell| c. \quad \text{(II.2.71)}$$

Since $|r_\ell| c = r_\ell$ for $c = \pm 1$ by lemma B.8, this means that all $\boldsymbol{\beta} \in \mathcal{A}_{\pm 1}(\Omega)$ satisfy $\langle \boldsymbol{n}_\ell, r_\ell \rangle$ for all $\ell \in M_\pm^{(p)}$, thereby proving that $\mathcal{A}_{\pm 1}(\Omega) \subseteq \mathcal{B}(\mho_\pm^{(p)})$.

For the converse, we first note a direct consequence of eqs. (II.2.63) and (II.4.10),

$$\mathcal{H}_\pm^{(p)}(\Omega) \subseteq \mathcal{A}_{\pm 1}(\Omega), \qquad \mathcal{H}_\pm^{(p)} \subseteq \mathcal{H}_\pm^{(q)} \quad \text{if } q \geq p. \quad \text{(II.2.72)}$$

Then, the proof hinges on the following deceptively simple result:

Lemma B.10. *For any $\beta \in \mathcal{B}(\mho_{\pm}^{(p)})$, there is some q such that $\beta \in \mathcal{H}_{\pm}^{(q)}$.*

We will postpone its lengthy proof until after the main proof is complete. Lemma B.10, along with eq. (II.2.72), shows that

$$\mathcal{B}(\mho_{\pm}^{(p)}) \subseteq \bigcup_{q=p}^{\infty} \mathcal{H}_{\pm}^{(q)} \subseteq \mathcal{A}_{\pm 1}(\Omega). \tag{II.2.73}$$

This proves that $\mathcal{B}(\mho_{\pm}^{(p)}) = \mathcal{A}_{\pm 1}(\Omega)$.

For $\mathcal{A}_0$, let us turn to eq. (II.4.12), which lets us straightforwardly generalise $\mathcal{A}_c(\Omega)$ to non-integer c. One easily finds the following generalisation of eq. (II.4.4):

$$\langle \boldsymbol{\alpha}, 0 \rangle \leq \Omega \quad \Leftrightarrow \quad \exists \varepsilon > 0, \ \langle \boldsymbol{\alpha}, +\varepsilon \rangle \leq \Omega \quad \Leftrightarrow \quad \forall \varepsilon > 0, \ \langle \boldsymbol{\alpha}, -\varepsilon \rangle \leq \Omega, \tag{II.2.74}$$

from which it follows that

$$\bigcup_{n=1}^{\infty} \mathcal{A}_{+\frac{1}{n}}(\Omega) = \mathcal{A}_0(\Omega) = \bigcap_{n=1}^{\infty} \mathcal{A}_{-\frac{1}{n}}(\Omega). \tag{II.2.75}$$

Using that we have proven proposition B.2 for $c = \pm 1$, which generalises to all $c \neq 0$ by rescaling, we have for all positive integers n

$$\mathcal{A}_{\pm\frac{1}{n}}(\Omega) = \mathcal{B}\left(\sum_{\ell \in M_{\pm}^{(p)}} \langle \boldsymbol{n}_l, \tfrac{1}{n} r_l \rangle \right), \tag{II.2.76}$$

so by extension,

$$\bigcup_{n=1}^{\infty} \mathcal{A}_{\pm\frac{1}{n}}(\Omega) = \bigcup_{n=1}^{\infty} \mathcal{B}\left(\sum_{\ell \in M_{\pm}^{(p)}} \langle \boldsymbol{n}_l, \tfrac{1}{n} r_l \rangle \right),$$
$$\bigcap_{n=1}^{\infty} \mathcal{A}_{\pm\frac{1}{n}}(\Omega) = \bigcap_{n=1}^{\infty} \mathcal{B}\left(\sum_{\ell \in M_{\pm}^{(p)}} \langle \boldsymbol{n}_l, \tfrac{1}{n} r_l \rangle \right). \tag{II.2.77}$$

With the sign of r_ℓ given by lemma B.8, we have

$$\bigcup_{n=1}^{\infty} \mathcal{B}\left(\sum_{\ell \in L'_+} \langle \boldsymbol{n}_l, \tfrac{1}{n} r_l \rangle \right) = \mathcal{B}\left(\sum_{\ell \in L'_+} \langle \boldsymbol{n}_l, 0 \rangle \right),$$
$$\bigcap_{n=1}^{\infty} \mathcal{B}\left(\sum_{\ell \in L'_-} \langle \boldsymbol{n}_l, \tfrac{1}{n} r_l \rangle \right) = \mathcal{B}\left(\sum_{\ell \in L'_-} \langle \boldsymbol{n}_l, 0 \rangle \right), \tag{II.2.78}$$

so by eqs. (II.2.75) and (II.2.77), both of these sets are equal to $\mathcal{A}_0(\Omega)$, which completes the proof. ∎

B.4.2 Proof of lemma B.10

This lemma is the key to proving proposition B.2, and relies on several other lemmata that we will now establish. They also serve to elucidate some aspects of the proposition and its proof; for instance, lemma B.14 explains why the choice of p is arbitrary.

Lemma B.11. *Let $\langle m, s \rangle$ be a constraint such that $\mathcal{H}_\pm^{(p)} \in \mathcal{B}(\langle m, s \rangle)$, and which supports a point $\alpha_i + (p-1)\alpha_j \in \mathcal{Z}_\pm^{(p)}$ or a point $p\alpha_i + (p-1)\alpha_{j'} \in \mathcal{N}_\pm^{(p)}$. Then $\langle m, s \rangle$ either supports $\alpha_i \in \mathcal{P}_\pm$ for that same $i \in I_{\pm 1}$, or supports no point in $\mathcal{P}_\pm$ at all.*

PROOF. Assume $\langle m, s \rangle$ supports some $\pi \in \mathcal{P}_\pm$, but that it does not support α_i. Then

$$m \cdot (p\alpha_i + (p-1)\alpha_j) = s \quad \Rightarrow \quad m \cdot (p\pi + (p-1)\alpha_j) < s, \tag{II.2.79}$$

since $m \cdot \alpha_i > m \cdot \pi = s$. This is a contradiction, since

$$p\pi + (p-1)\alpha_j \in \mathcal{N}_\pm^{(p)} \subseteq \mathcal{H}_\pm^{(p)} \subseteq \mathcal{B}(\langle m, s \rangle). \tag{II.2.80}$$

The argument for $\alpha_i + (p-1)\alpha_j \in \mathcal{Z}_\pm^{(p)}$ is the same. $\square$

Lemma B.12. *Let $\langle m, s \rangle$ be a constraint with $\mathcal{H}_\pm^{(p)} \subseteq \mathcal{B}(\langle m, s \rangle)$, and let it support at least one point in $\mathcal{P}_\pm$. Then if it supports $\alpha_i + (p-1)\alpha_j \in \mathcal{Z}_\pm^{(p)}$, it also supports $\alpha_i + (q-1)\alpha_j \in \mathcal{Z}_\pm^{(q)}$ for all $q > 1$. Likewise, if it supports $p\alpha_i + (p-1)\alpha_j \in \mathcal{N}_\pm^{(p)}$, it also supports $q\alpha_i + (q-1)\alpha_j \in \mathcal{N}_\pm^{(q)}$ for all $q > 1$.*[56]

PROOF. Assume $m \cdot [p\alpha_i + (p-1)\alpha_j] = s$. Then by lemma B.11, $\langle m, s \rangle$ also supports α_i, and thus $m \cdot (\alpha_i + \alpha_j) = 0$. Adding $(q-p)(\alpha_i + \alpha_j)$ therefore gives $m \cdot [q\alpha_i + (q-1)\alpha_j] = s$. The argument for $\alpha_i + (p-1)\alpha_j \in \mathcal{Z}_\pm^{(p)}$ is the same. $\square$

Lemma B.13. $\dim(\mathcal{H}_\pm^{(q)})$, *as defined in section B.2.1, is independent of q for $q > 1$. Furthermore, for any $J_c \subseteq I_c$, $\dim[\mathcal{J}_q(J_{\pm 1}, J_0, J_{\mp 1})]$ is independent of q for $q > 1$, where*

$$\mathcal{J}_q(J_{\pm 1}, J_0, J_{\mp 1}) \equiv \{\alpha_i\}_{i \in J_{\pm 1}} \cup \{\alpha_i + (q-1)\alpha_j\}_{\substack{i \in J_{\pm 1} \\ j \in J_0}} \cup \{q\alpha_i + (q-1)\alpha_j\}_{\substack{i \in J_{\pm 1} \\ j \in J_{\mp 1}}}. \tag{II.2.81}$$

PROOF. The main statement is actually a special case of the "furthermore" statement, since $\mathcal{P}_\pm \cup \mathcal{Z}_\pm^{(q)} \cup \mathcal{N}_\pm^{(q)} = \mathcal{J}_q(I_{\pm 1}, I_0, I_{\mp 1})$, and since for any set $\mathcal{X}$, $\dim[\text{Hull}(\mathcal{X})] = \dim(\mathcal{X})$ because convex combinations are a special case of affine combinations.

By definition, the affine span of $\mathcal{J}_q(J_{\pm 1}, J_0, J_{\mp 1})$ has the same dimension as the linear span of $\{\alpha - \beta \mid \alpha, \beta \in \mathcal{J}_q(J_{\pm 1}, J_0, J_{\mp 1})\}$; according to eq. (II.2.81), this set consists of

$$\begin{cases} \alpha_i - \alpha_j & \text{for } i, j \in J_{\pm 1}, \\ (q-1)\alpha_j & \text{for } j \in J_0, \\ (q-1)(\alpha_i - \alpha_j) & \text{for } i \in J_{\pm 1} \cup J_0 \cup J_{\mp 1} \text{ and } j \in J_0 \cup J_{\mp 1}, \\ \text{various linear combinations of the above.} \end{cases} \tag{II.2.82}$$

The linear span is unaffected by the inclusion of extra linear combinations or nonzero scale factors, so as long as $(q-1) \neq 0$ we can drop these and be left with the span of

$$\{\alpha_i - \alpha_j \mid i, j \in J_{\pm 1} \cup J_0 \cup J_{\mp 1}\} \cup \{\alpha_j \mid j \in J_0\}, \tag{II.2.83}$$

which is clearly q-independent. $\square$

---

[56] Lemmas B.12 and B.13 actually work for all $q \neq 1$. The only places where proposition B.2 actually requires $p > 1$ rather than $p < 1$ are in lemmas B.8 and B.9 and in eq. (II.2.72).

Lemma B.14. *If* $\dim(\mathcal{H}_{\pm}^{(q)}) = D$ *for some* $q > 1$*, then* $\mho_{\pm}^{(q)}$ *is non-degenerate for all* $q > 1$ *and* $\mho_{\pm}^{(p)} = \mho_{\pm}^{(q)}$ *for all* $p, q > 1$.

This makes it quite clear why p is arbitrary in proposition B.2.

PROOF. By lemma B.13, $\dim(\mathcal{H}_{\pm}^{(q)})$ equals D for all $q > 1$ if it does for some $q > 1$. Then $\mho_{\pm}^{(q)}$ is non-degenerate, since $\mathcal{H}_{\pm}^{(p)} \subseteq \mathcal{B}(\mho_{\pm}^{(p)})$ has dimension D.

Now for the converse. Given $\ell \in M_{\pm}^{(p)}, p > 1$, define

$$\mathcal{G}_{\ell}^{(p)} = \{\boldsymbol{\pi} \in \mathcal{P}_{\pm} \cup \mathcal{Z}_{\pm}^{(p)} \cup \mathcal{N}_{\pm}^{(p)} \mid \boldsymbol{n}_{\ell} \cdot \boldsymbol{\pi} = r_{\ell}\}. \tag{II.2.84}$$

By the definition of $M_{\pm}^{(p)}$ along with lemma B.11, there are some $J_c^{\ell} \subseteq I_c$ such that $\mathcal{G}_{\ell}^{(p)} = \mathcal{J}_p(J_{\pm 1}^{\ell}, J_0^{\ell}, J_{\mp 1}^{\ell})$ as defined in eq. (II.2.81). Then by lemma B.12, that same $\langle \boldsymbol{n}_{\ell}, r_{\ell} \rangle$ also supports $\mathcal{G}_{\ell}^{(q)} \equiv \mathcal{J}_q(J_{\pm 1}^{\ell}, J_0^{\ell}, J_{\mp 1}^{\ell})$ for all $q > 1$. From this and eq. (II.2.72), it follows that $\langle \boldsymbol{n}_{\ell}, r_{\ell} \rangle \leq \mho_{\pm}^{(q)}$.

Recall now that $\langle \boldsymbol{n}_{\ell}, r_{\ell} \rangle \in \mathcal{R}(\mho_{\pm}^{(p)})$, so corollary B.2 implies that $\dim\left[\mathcal{G}_{\ell}^{(p)}\right] = D - 1$.[57] Then by lemma B.13, $\dim[\mathcal{G}_{\ell}^{(q)}] = D - 1$ for all $q > 1$, so again by corollary B.2, $\langle \boldsymbol{n}_{\ell}, r_{\ell} \rangle \in \mathcal{R}(\mho_{\pm}^{(q)})$. By doing this for all $\ell \in M_{\ell}^{(p)}$ and repeating with p and q exchanged, we see that $\mathcal{R}(\mho_{\ell}^{(q)}) = \mathcal{R}(\mho_{\ell}^{(p)})$ up to normalisation. This implies $\mho_{\ell}^{(q)} = \mho_{\ell}^{(p)}$. $\qquad\square$

Thanks to lemma B.14, we will drop the "(p)" superscript on $\mho_{\pm}$ from now on. However, we face the problem that $\mho_{\pm}$ may be degenerate, which would make lemma B.14 inapplicable. It can be circumvented by using the notion of induced constraints developed in section B.2.1: just substitute

$$\mho_{\pm} \to \lfloor \mho_{\pm} \rfloor, \qquad \mathcal{A}_{\pm 1}(\Omega) \to \lfloor \mathcal{A}_{\pm 1}(\Omega) \rfloor_{E_{\mho_{\pm}}}, \tag{II.2.85}$$

since it follows from lemma B.1 (along with $\mathcal{A}_{\pm 1}(\Omega) \subseteq E_{\mho_{\pm}}$ which we proved earlier) that[58]

$$\lfloor \mathcal{A}_{\pm 1}(\Omega) \rfloor_{E_{\mho_{\pm}}} \supseteq \mathcal{B}(\lfloor \mho_{\pm} \rfloor) \quad \Leftrightarrow \quad \mathcal{A}_{\pm 1}(\Omega) \supseteq \mathcal{B}(\mho_{\pm}). \tag{II.2.86}$$

Thus, we may for the remainder assume that $\mho_{\pm}$ is non-degenerate.

We are now, at long last, ready to prove lemma B.10 itself. Consider $\mathcal{H}_{\pm}^{(q)}$ for arbitrary q, and presume $\boldsymbol{\beta} \notin \mathcal{H}_{\pm}^{(q)}$. Select some $\boldsymbol{\alpha} \in \mathcal{H}_{\pm}^{(q)}$, and draw the line segment joining $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. It must intersect the boundary of $\mathcal{H}_{\pm}^{(q)}$ in some point $\boldsymbol{\gamma}$, which is supported by one or more $\langle \boldsymbol{n}_{\ell}, r_{\ell} \rangle, \ell \in L_{\pm}^{(q)}$, at least one of which is not satisfied by $\boldsymbol{\beta}$. We therefore have some $\ell \in L_{\pm}^{(q)}$ such that

$$\boldsymbol{n}_{\ell} \cdot \boldsymbol{\alpha} \geq r_{\ell}, \qquad \boldsymbol{n}_{\ell} \cdot \boldsymbol{\beta} < r_{\ell}, \qquad \boldsymbol{n}_{\ell} \cdot \boldsymbol{\gamma} = r_{\ell}. \tag{II.2.87}$$

---

[57]What corollary B.2 calls "$\mathcal{F}$" is here $\text{Hull}\left(\mathcal{G}_{\ell}^{(p)}\right)$, but $\dim[\mathcal{X}] = \dim\left[\text{Hull}(\mathcal{X})\right]$ for any set $\mathcal{X}$ since convex combinations are a special case of affine combinations.

[58]For practical applications, there is the additional problem that qhull does not function properly when its output would be degenerate. This has not been a problem for us, and is of course no issue for the present proof, but if needed, one could identify the affine subspace E containing all points defined in eq. (II.2.63), apply $\lfloor \cdots \rfloor_E$ to them, and work entirely in the lower-dimensional space where there are no degeneracies.

If $\ell \in M_\pm^{(q)}$, then $\boldsymbol{\beta} \notin \mathcal{B}(\mho_\pm)$; this is where the q-independence of $\mho_\pm$ proven in lemma B.14 is crucial. Thus, we can assume that $\ell \notin M_\pm^{(q)}$.

The magnitude of $\boldsymbol{\gamma}$ is bounded from above by $|\boldsymbol{\gamma}| \leq \max(|\boldsymbol{\alpha}|, |\boldsymbol{\beta}|)$. We will now attempt to prove that $|\boldsymbol{\gamma}| > \max(|\boldsymbol{\alpha}|, |\boldsymbol{\beta}|)$, which leads to a contradiction, proving that $\boldsymbol{\gamma}$ does not exist and consequently that $\boldsymbol{\beta} \in \mathcal{H}_\pm^{(q)}$.

For any $\ell \in L_\pm^{(q)}$, let $\mathcal{G}_\ell^{(q)} \subseteq \left(\mathcal{P}_\pm \cup \mathcal{Z}_\pm^{(q)} \cup \mathcal{N}_\pm^{(q)}\right)$ be defined as in eq. (II.2.84), and let us think about the structure of $\mathcal{G}_\ell^{(q)}$. Each of its elements is of the form $\boldsymbol{\alpha}_i + (q-1)\boldsymbol{\alpha}'$, where $i \in I_{\pm 1}$ and $\boldsymbol{\alpha}'$ may take the form $\mathbf{0}, \boldsymbol{\alpha}_j$ or $\boldsymbol{\alpha}_i + \boldsymbol{\alpha}_j$ depending on whether the element is part of $\mathcal{P}_\pm, \mathcal{Z}_\pm^{(q)}$ or $\mathcal{N}_\pm^{(q)}$. This observation allows us to write the Minkowski sum

$$\text{Hull}\left(\mathcal{G}_\ell^{(q)}\right) = \mathcal{U}_\ell^{(q)} + (q-1)\mathcal{V}_\ell^{(q)}, \qquad (\text{II.2.88})$$

where $\mathcal{U}_\ell^{(q)} \subseteq \text{Hull}\left(\mathcal{P}_\pm\right)$. $\mathcal{V}_\ell^{(q)}$ is the convex hull of a subset of $\{\boldsymbol{\alpha}_i\}_{i \in I}$; the details are messy and unimportant, so we will not write it explicitly, but the important thing is that it only depends on q and ℓ through the specific choice of subset. This means that when counted over the infinitely many choices of q and ℓ, there is only a finite number of distinct $\mathcal{V}_\ell^{(q)}$ that appear: at most as many as there are subsets of I.

Now, focus on the case $\ell \in L_\pm^{(q)} \setminus M_\pm^{(q)}$, where we find that $\mathbf{0} \notin \mathcal{V}_\ell^{(q)}$: otherwise, $\text{Hull}\left(\mathcal{G}_\ell^{(q)}\right) \cap \text{Hull}\left(\mathcal{P}_\pm\right) \neq \emptyset$, which straightforwardly leads to a contradiction of the definition of $M_\pm^{(q)}$. We may also observe that $\mathcal{V}_\ell^{(q)}$ (and $\mathcal{U}_\ell^{(q)}$) are closed sets, being the convex hulls of finite sets of points.

Consider then the finite set

$$\{\mathcal{V} \mid \mathcal{V}_\ell^{(q)} = \mathcal{V} \quad \text{for some } q \geq 2, \ell \in L_\pm^{(q)} \setminus M_\pm^{(q)}\}, \qquad (\text{II.2.89})$$

and let $\mathcal{W}_\pm$ be the union of all elements of this set. By the observations we have made about $\mathcal{V}_\ell^{(q)}$, this is a closed set (being the finite union of closed sets) that does not contain $\mathbf{0}$. Consequently, there is some $m > 0$ such that $\mathcal{W}_\pm$ contains no vector of magnitude less than m.

Recall now that $\boldsymbol{\gamma} \in \mathcal{G}_\ell^{(q)}$ for some $\ell \in L_\pm^{(q)} \setminus M_\pm^{(q)}$. Thus, eq. (II.2.88) gives

$$\boldsymbol{\gamma} = \boldsymbol{\pi} + (q-1)\boldsymbol{\eta}, \qquad \boldsymbol{\pi} \in \text{Hull}\left(\mathcal{P}_\pm\right), \boldsymbol{\eta} \in \mathcal{W}_\pm, \qquad (\text{II.2.90})$$

and this is true no matter the value of q and no matter which $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are used to obtain $\boldsymbol{\gamma}$. As derived above, $|\boldsymbol{\eta}| \geq m$, and since $\text{Hull}\left(\mathcal{P}_\pm\right)$ is a bounded set, $|\boldsymbol{\pi}| \leq M$ for sufficiently large M. For q sufficiently large that $(q-1)m > M$, the triangle inequality then gives

$$|\boldsymbol{\gamma}| \geq (q-1)m - M, \qquad (\text{II.2.91})$$

which can be made arbitrarily large by further increasing q, thereby providing the desired contradiction and completing the proof. $\qquad \square$

B.4.3 Some important corollaries of proposition B.2

We can refine the treatment of $\mathcal{A}_0(\Omega)$ with the following:[59]

Corollary B.3 (construction of $\mathcal{A}_0(\Omega)$)**.** *Let* Ω*,* $M_\pm^{(p)}$*, etc. be as in proposition B.2. Let* $N_\pm^{(p)} \subseteq M_\pm^{(p)}$ *be the set of those* ℓ *for which* $\langle \boldsymbol{n}_\ell, r_\ell \rangle$ *also supports at least one point in* $\left(\mathcal{Z}_\pm^{(p)} \cup \mathcal{N}_\pm^{(p)} \right)$*. If* $\sum_{\ell \in M_\pm^{(p)}} \langle \boldsymbol{n}_\ell, 0 \rangle$ *is a non-degenerate constraint,[60] then*

$$\mathcal{A}_0(\Omega) = \mathcal{B}\left(\sum_{\ell \in N_\pm^{(p)}} \langle \boldsymbol{n}_\ell, 0 \rangle \right). \tag{II.2.92}$$

As in proposition B.2, the $M_+^{(p)}$ *construction does not work when* $I_1 = \emptyset$*.*

PROOF. For ℓ contained in $M_\pm^{(p)}$ but not in $N_\pm^{(p)}$, lemma B.9 more specifically gives

$$\forall i \in I_{\pm 1}, \ \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_i \geq r_\ell, \qquad \forall i \in I_0, \ \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_i > 0, \qquad \forall i \in I_{\mp 1}, \ \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_i > -r_\ell \ \ (\text{II.2.93})$$

with lemma B.8 dictating the sign of r_ℓ. Then if $\boldsymbol{n}_\ell \cdot \boldsymbol{\beta} = 0$ for $\boldsymbol{\beta} \in \mathcal{A}_0(\Omega)$, a look at eq. (II.4.10) tells us that all $\lambda_i, i \in I_0 \cup I_1$ must be zero due to the above inequalities, and eq. (II.4.11) then implies that also $\lambda_i, i \in I_{-1}$ must be zero. Thus, the only point $\boldsymbol{\beta} \in \mathcal{A}_0(\Omega)$ supported by $\langle \boldsymbol{n}_\ell, 0 \rangle$ is the trivial $\boldsymbol{\beta} = \boldsymbol{0}$, which is supported by $\langle \boldsymbol{\alpha}, 0 \rangle$ for all $\boldsymbol{\alpha} \neq \boldsymbol{0}$. Therefore, $\langle \boldsymbol{n}_\ell, 0 \rangle$ does not uniquely support any point, so by proposition 4.3 (which requires non-degeneracy), it can be omitted. ∎

One also easily finds the following simplification:

Corollary B.4 (construction of $\mathcal{A}_c(\omega_c)$)**.** *For those* ω_c *covered by proposition 4.1, proposition B.2 reduces down to the following:*

$c = +1 :$ $\mathcal{H}_+^{(p)} = \text{Hull}\left(\{ \boldsymbol{\alpha}_i, p\boldsymbol{\alpha}_i \}_{i \in I} \right)$*.* $M_+^{(p)}$ *consists of those* ℓ *for which* $r_\ell \geq 0$*.*

$c = 0 :$ $\mathcal{H}_-^{(p)} = \text{Hull}\left(\{ \boldsymbol{0} \} \cup \{ p\boldsymbol{\alpha}_i \}_{i \in I} \right)$*.* $N_-^{(p)} = M_-^{(p)}$ *consists of those* ℓ *where* $r_\ell = 0$*.*

$c = -1 :$ $\mathcal{H}_-^{(p)} = \text{Hull}\left(\{ \boldsymbol{0} \} \cup \{ \boldsymbol{\alpha}_i \}_{i \in I} \right)$*. All* $\langle \boldsymbol{n}_\ell, r_\ell \rangle$ *are kept, since* $M_-^{(p)} = L_-^{(p)}$*.* ∎

An alternative to using proposition B.2 is to apply corollary B.4 to ω_0^ℓ as defined in section B.1.4, and then "unlifting" the result. This requires the treatment of a much smaller

---

[59]An intuitive understanding of the construction of $\mathcal{A}_0(\Omega)$ can be gained by noting that in a sense, one can make $c \to 0$ by "zooming out" on constraint space. This is the principle that is formalised in the end of the proof of proposition B.2. Each facet of the body $\mathcal{A}_c(\Omega)$, consisting of the points supported by one $\langle \boldsymbol{n}_\ell, r_\ell \rangle$, is thereby shifted so that it passes through the origin (hence $r_\ell \to 0$), and if it was bounded, it shrinks down to a point. Corollary B.3 identifies those bounded facets and removes them. It is easy to see why it works: all facets of $\mathcal{H}_\pm^{(p)}$ are bounded, and only by being adjacent to a facet that is removed in the restriction $L_\pm^{(p)} \to M_\pm^{(p)}$ can a facet become unbounded. The removed facets are those that only support points in $\mathcal{Z}_\pm^{(p)} \cup \mathcal{N}_\pm^{(p)}$, so their neighbours are the ones that support at least one point in these sets.

[60]This is not equivalent to $\sum_{\ell \in M_\pm^{(p)}} \langle \boldsymbol{n}_\ell, r_\ell \rangle$ being non-degenerate. Consider as a counterexample $\Omega = \langle \binom{1}{1}, -1 \rangle + \langle \binom{1}{-1}, -1 \rangle + \langle \binom{1}{0}, 0 \rangle$, for which $\mathcal{A}_0(\Omega)$ is contained in a hyperplane (i.e. a line) whereas $\mathcal{A}_{-1}(\Omega)$ is not.

number of points, which makes the QuickHull algorithm run faster; on the other hand, lifting increases the dimension, which makes the QuickHull algorithm run slower and be less numerically stable. The time complexity of the qhull implementation suggests that asymptotically, lifting should be the faster method, but since qhull is vastly more efficient in 2 and 3 dimensions, not lifting should be preferable when the number of dimensions is small. In practice, we only used proposition B.2 directly without lifting.

B.4.4   Construction of $\mathcal{R}(\Omega)$

The marriage of propositions B.2 and 4.3 makes for a practical way of determining the minimal representation $\mathcal{R}(\Omega)$ of any non-degenerate constraint Ω. As remarked before, degenerate constraints are of little practical relevance, although if needed, the degenerate case can be covered by adapting proposition B.1.

When forming a convex hull, qhull produces a list of its vertices and readily checks if two vertices form the endpoints of an edge. Based on that, we devise the following:

Proposition B.3 (constructing $\mathcal{R}(\Omega)$). *Let Ω be a non-degenerate constraint, and let $\mathcal{H}_\pm^{(p)}$, $M_\pm^{(p)}$, etc. be defined as in proposition B.2. Then the unique minimal representation $\mathcal{R}(\Omega)$ is determined as follows:*

- *For $i \in I_{\pm 1}$, $\langle \boldsymbol{\alpha}_i, \pm 1 \rangle \in \mathcal{R}(\Omega)$ if and only if $\boldsymbol{\alpha}_i$ is an extreme point of $\mathcal{A}_{\pm 1}(\Omega)$. (The exception is $\langle \mathbf{0}, -1 \rangle$, which is of course not in $\mathcal{R}(\Omega)$.)*

  Equivalently, $\langle \boldsymbol{\alpha}_i, \pm 1 \rangle \in \mathcal{R}(\Omega)$ if and only if $\boldsymbol{\alpha}_i \in \mathcal{P}_\pm$ is an extreme point of $\mathcal{H}_\pm^{(p)}$ for some $p > 1$, with the same exception.

- *For $j \in I_0$, $\langle \boldsymbol{\alpha}_j, 0 \rangle \in \mathcal{R}(\Omega)$ if and only if there is some $i \in I_{\pm 1}$ such that the ray $\{\boldsymbol{\alpha}_i + \lambda \boldsymbol{\alpha}_j \mid \lambda \geq 0\}$ is an edge of $\mathcal{A}_{\pm 1}(\Omega)$ that contains no point in $\mathcal{N}_\pm^{(p)}(\Omega)$. (This is up to normalisation; several equivalent $\langle \boldsymbol{\alpha}_i, 0 \rangle$ may satisfy this condition.)*

  Equivalently, $\langle \boldsymbol{\alpha}_j, 0 \rangle \in \mathcal{R}(\Omega)$ if and only if there is some $i \in I_{\pm 1}$ such that the line segment between $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j$ is an edge of $\mathcal{H}_\pm^{(p)}(\Omega)$ that contains no point in $\mathcal{N}_\pm^{(p)}(\Omega)$. (This breaks the normalisation ambiguity: if several $\langle \boldsymbol{\alpha}_j, 0 \rangle$ are equivalent under eq. (II.4.5), then only the one with the largest $|\boldsymbol{\alpha}_j|$ will form the endpoint of their edge and be included in $\mathcal{R}(\Omega)$.)

Note how this can be applied to corollary B.4 without modification. Illustrations can be found in figs. II.26 to II.29.

PROOF. We will prove that the stated conditions are equivalent to $\langle \boldsymbol{\alpha}_i, c_i \rangle$ uniquely supporting a point; the rest follows from proposition 4.3. We will focus on proving the conditions based on $\mathcal{H}_\pm^{(p)}$ rather than $\mathcal{A}_{\pm 1}(\Omega)$; that they are equivalent follows easily from property (v) of K-faces along with lemma B.4 ($\boldsymbol{\alpha}_i \in \mathcal{P}_\pm$ serves as the point φ).[61]

---

[61]There is one subtlety for $\langle \boldsymbol{\alpha}_j, 0 \rangle \in \mathcal{R}(\Omega)$: the line segment between $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j$ may fail to be an edge of $\mathcal{H}_\pm^{(p)}$ even though $\{\boldsymbol{\alpha}_i + \lambda \boldsymbol{\alpha}_j \mid \lambda \geq 0\}$ is an edge of $\mathcal{A}_{\pm 1}(\Omega)$, if it is contained in the line segment

Consider first $i \in I_{\pm 1}$. If $\langle \boldsymbol{\alpha}_i, \pm 1 \rangle \in \mathcal{R}(\Omega)$, then there is some $\boldsymbol{h}$ uniquely supported by $\langle \boldsymbol{\alpha}_i, \pm 1 \rangle$; in particular, for any $\boldsymbol{\beta} \in \mathcal{A}_{\pm 1}(\Omega), \boldsymbol{\beta} \neq \boldsymbol{\alpha}$ we have $\boldsymbol{\beta} \cdot \boldsymbol{h} > \boldsymbol{\alpha}_i \cdot \boldsymbol{h} = \pm 1$. Since by eq. (II.2.73)

$$\mathcal{A}_{\pm 1}(\Omega) = \bigcup_{p=2}^{\infty} \mathcal{H}_{\pm}^{(p)} \tag{II.2.94}$$

we have that for any $p > 1$

$$\{\boldsymbol{\eta} \in \mathcal{H}_{\pm}^{(p)} \mid \boldsymbol{\eta} \cdot \boldsymbol{h} = \pm 1\} = \{\boldsymbol{\alpha}_i\}. \tag{II.2.95}$$

From lemma B.3, we conclude that $\boldsymbol{\alpha}_i$ is an extreme point of $\mathcal{H}_{\pm}^{(p)}$.

Conversely, if $\boldsymbol{\alpha}_i$ is an extreme point of $\mathcal{H}_{\pm}^{(p)}$, then eq. (II.2.95) holds. In particular, if $j \in I_{\mp 1}$ then $\boldsymbol{h} \cdot (p\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j) > \pm 1$ so $\boldsymbol{\alpha}_j \cdot \boldsymbol{h} > \mp 1$. Similarly, for any $j \in I_0$ we have $\boldsymbol{\alpha}_j \cdot \boldsymbol{h} > 0$. It then follows from lemma B.5 and proposition 4.3 that $\langle \boldsymbol{\alpha}_i, \pm 1 \rangle \in \mathcal{R}(\Omega)$.

Consider then $j \in I_0$. If the line segment between $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j$ is an edge of $\mathcal{H}_{\pm}^{(p)}$, then by lemma B.3 there exists $\langle \boldsymbol{h}, d \rangle$ such that

$$\{\boldsymbol{\eta} \in \mathcal{H}_{\pm}^{(p)} \mid \boldsymbol{\eta} \cdot \boldsymbol{h} = d\} = \{\boldsymbol{\alpha}_i + \lambda \boldsymbol{\alpha}_j \mid \lambda \in [0, (p-1)]\} \tag{II.2.96}$$

and $\mathcal{H}_{\pm}^{(p)} \subseteq \mathcal{B}(\langle \boldsymbol{h}, d \rangle)$. We then immediately find that $\boldsymbol{h} \cdot \boldsymbol{\alpha}_i = d, \boldsymbol{h} \cdot \boldsymbol{\alpha}_j = 0$.

What about the other $\boldsymbol{\alpha}_k, k \in I$? If $k \in I_{\mp}$ then $p\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_k \in \mathcal{N}_{\pm}^{(p)}$ which is by assumption not supported by $\langle \boldsymbol{h}, d \rangle$, so it follows that $\boldsymbol{h} \cdot \boldsymbol{\alpha}_k > -d$. In the same way, for any $k \in I_0$ we have $\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_k \in \mathcal{Z}_{\pm}^{(p)}$, implying that $\boldsymbol{h} \cdot \boldsymbol{\alpha}_k \geq 0$. If $\boldsymbol{h} \cdot \boldsymbol{\alpha}_k = 0$ then it follows that $\boldsymbol{\alpha}_k$ is collinear with $\boldsymbol{\alpha}_j$, so either $\langle \boldsymbol{\alpha}_i, 0 \rangle = \langle \boldsymbol{\alpha}_k, 0 \rangle$ or $\langle \boldsymbol{\alpha}_i, 0 \rangle = \langle -\boldsymbol{\alpha}_k, 0 \rangle$ under eq. (II.4.5); the latter is excluded by non-degeneracy. Therefore, we can without loss of generality assume $\boldsymbol{h} \cdot \boldsymbol{\alpha}_k > 0$ for all $k \in I_0 \setminus \{j\}$ by omitting equivalent constraints. Lastly, if $k \in I_{\pm}$ then $\boldsymbol{h} \cdot \boldsymbol{\alpha}_k \geq d$; here, the non-strict inequality is unavoidable but does not pose a problem.

If $i \in I_+$, then the fact that $p\boldsymbol{\alpha}_i \in \mathcal{N}_+^{(p)}$ implies that $d > 0$. If $i \in I_-$, we can use the fact that $\boldsymbol{0} \in \mathcal{P}_-$ to prove that $d \leq 0$, with equality only if $\boldsymbol{\alpha}_i = \boldsymbol{0}$. Combining this with the previous paragraph, we conclude that

$$\forall k \in I \setminus \{i, j\}, \ \boldsymbol{h} \cdot \boldsymbol{\alpha}_k \geq c_k |d| \tag{II.2.97}$$

(equality only possible if $k \in I_{\pm}$). For sufficiently small $\varepsilon > 0$, we may replace it by $|d| \to |d| \mp \varepsilon$ without invalidating the above inequality; this also guards against problems when $d = 0$. Then

$$\tfrac{1}{|d| \mp \varepsilon} \boldsymbol{h} \cdot \boldsymbol{\alpha}_i > \pm 1, \qquad \tfrac{1}{|d| \mp \varepsilon} \boldsymbol{h} \cdot \boldsymbol{\alpha}_j = 0, \qquad \forall k \in I \setminus \{i, j\}, \ \tfrac{1}{|d| \mp \varepsilon} \boldsymbol{h} \cdot \boldsymbol{\alpha}_k > c_k. \tag{II.2.98}$$

Thus, $\langle \boldsymbol{\alpha}_j, 0 \rangle$ uniquely supports $\frac{1}{|d| \mp \varepsilon} \boldsymbol{h}$ by lemma B.5.

---

between $\boldsymbol{\alpha}_i$ and some $\boldsymbol{\alpha}_{i'}, i' \in I_{\pm 1}$. This can be remedied by using sufficiently large p. For $\langle \boldsymbol{\alpha}_i, \pm 1 \rangle \in \mathcal{R}(\Omega)$, all p work equally well, as should be apparent from the proof.

Conversely, assume $\boldsymbol{\alpha}_j$ uniquely supports a point $\boldsymbol{b}$. Then

$$\boldsymbol{b} \cdot \boldsymbol{\alpha}_j = 0, \qquad \forall i \in I \setminus j, \ \boldsymbol{b} \cdot \boldsymbol{\alpha}_i > c_i. \tag{II.2.99}$$

In particular,

$$\mathcal{F} = \{\boldsymbol{\beta} \in \mathcal{A}_0(\Omega) \mid \boldsymbol{b} \cdot \boldsymbol{\beta} = 0\} = \{\lambda \boldsymbol{\alpha}_j \mid \lambda \geq 0\}. \tag{II.2.100}$$

(Non-degeneracy ensures $\lambda \geq 0$). By lemma B.3 and proposition B.2, $\mathcal{F}$ is an edge of $\mathcal{A}_0(\Omega) = \mathcal{B}\big(\sum_{\ell \in M_p^{(\pm)}} \langle \boldsymbol{n}_\ell, 0\rangle\big)$. Therefore, lemma B.2 guarantees that there exists $Q_\pm \subseteq M_\pm^{(p)}$ such that

$$\forall \ell \in Q_\pm, \ \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_j = 0, \qquad \forall \ell \in M_\pm^{(p)} \setminus Q_\pm, \ \boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_j > 0 \tag{II.2.101}$$

where $\{\boldsymbol{n}_\ell\}_{\ell \in Q_\pm}$ contains a subset of $D-1$ linearly independent vectors. Now, consider the set

$$\mathcal{C} \equiv \mathcal{B}\bigg(\sum_{\ell \in Q_\pm} \langle \boldsymbol{n}_\ell, r_\ell \rangle \bigg) \supseteq \mathcal{A}_{\pm 1}(\Omega). \tag{II.2.102}$$

This set must have at least one edge, since the intersection of $D-1$ independent hyperplanes is a line. By lemma B.2, there is then a subset $J_\pm \subseteq Q_\pm$ with $|J_\pm| = D-1$ and $\{\boldsymbol{n}_\ell\}_{\ell \in J_\pm}$ linearly independent, that describes that edge according to eq. (II.2.31) (if there are several, we choose one arbitrarily). By construction, $\mathcal{C}$ has no extreme points: $\{\boldsymbol{n}_\ell\}_{\ell \in Q_\pm}$ has no subset of D linearly independent vectors. Therefore, the chosen edge can be written like

$$\mathcal{F}_\pm \equiv \{\boldsymbol{\beta} \in \mathbb{R}^D \mid \forall \ell \in J_\pm, \ \boldsymbol{n}_\ell \cdot \boldsymbol{\beta} = r_\ell\}. \tag{II.2.103}$$

We now claim that

$$\mathcal{F}_\pm \cap \mathcal{A}_{\pm 1}(\Omega) = \mathcal{F}_\pm \cap \mathcal{B}\bigg(\sum_{\ell \in M_\pm^{(p)}} \langle \boldsymbol{n}_\ell, r_\ell \rangle \bigg) \neq \emptyset. \tag{II.2.104}$$

All points in $\mathcal{F}_\pm$ are satisfied by $\langle \boldsymbol{n}_\ell, r_\ell \rangle$ for $\ell \in Q_\pm$ by the arguments made above about $\mathcal{C}$. For $\ell \notin Q_\pm$, we note that for any $\boldsymbol{\beta} \in \mathcal{F}$, we also have $\boldsymbol{\beta} + \lambda \boldsymbol{\alpha}_j \in \mathcal{F}$ for all $\lambda \in \mathbb{R}$, so since $\boldsymbol{n}_\ell \cdot \boldsymbol{\alpha}_j > 0$, $\boldsymbol{n}_\ell \cdot (\boldsymbol{\beta} + \lambda \boldsymbol{\alpha}_j) \geq r_\ell$ for sufficiently large λ no matter what $\boldsymbol{n}_\ell \cdot \boldsymbol{\beta}$ is. Therefore, at least some subset of $\mathcal{F}_\pm$ is satisfied by all $\langle \boldsymbol{n}_\ell, r_\ell \rangle, \ell \in M_\pm^{(p)} \setminus Q_\pm$, which proves eq. (II.2.104). By then considering the constraint

$$\langle \boldsymbol{h}, d \rangle \equiv \langle \sum_{\ell \in J_\pm} \boldsymbol{n}_\ell, \sum_{\ell \in J_\pm} r_\ell \rangle, \tag{II.2.105}$$

it follows from lemma B.3 that $\mathcal{F}_\pm \cap \mathcal{A}_{\pm 1}(\Omega)$ is a K-facet of $\mathcal{A}_{\pm 1}(\Omega)$.

Let us now note that $\mathcal{F}_\pm \cap \mathcal{A}_{\pm 1}(\Omega)$ is an edge, not an extreme point, since for any $\boldsymbol{\beta} \in \mathcal{F}_\pm$ and $\lambda \geq 0$,[62]

$$\boldsymbol{\beta} + \lambda \boldsymbol{\alpha}_j \in \mathcal{F}_\pm \cap \mathcal{A}_{\pm 1}(\Omega). \tag{II.2.106}$$

---

[62]It also follows from this that $\mathcal{F}_\pm \cap \mathcal{A}_{\pm 1}(\Omega)$ is not a line segment. That it is a ray rather than a full line is not difficult to prove from non-degeneracy, and it is then possible to show that $\mathcal{F}_\pm \cap \mathcal{A}_{\pm 1}(\Omega) = \{\boldsymbol{\alpha}_i + \lambda \boldsymbol{\alpha}_j \mid \lambda \geq 0\}$ which completes the proof, but we choose an easier path.

From eqs. (II.2.72), (II.2.94) and (II.2.104), there is some $q > 1$ such that $\mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)} \neq \emptyset$ for all $p \geq q$, and by property (v) of K-facets, $\mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$ is an edge of $\mathcal{H}_\pm^{(p)}$. Unlike $\mathcal{A}_{\pm 1}(\Omega)$, $\mathcal{H}_\pm^{(p)}$ is always a compact set, so all its edges must be line segments with exactly two endpoints, which are extreme points of $\mathcal{H}_\pm^{(p)}$ by property (iv). Thus, $\mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$ is the line segment between two extreme points of $\mathcal{H}_\pm^{(p)}$, and these must clearly be elements of $\mathcal{P}_\pm \cup \mathcal{Z}_\pm^{(p)} \cup \mathcal{N}_\pm^{(p)}$.[63]

Then, let us prove that at least one of these endpoints of $\mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$ is in $\mathcal{P}_\pm$ by considering the alternatives. For $\mathcal{N}_\pm^{(p)}$, i.e. if $p\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_k$ is one endpoint for some $i \in I_{\pm 1}, k \in I_{\mp 1}$, then it follows from lemma B.11 that $\boldsymbol{\alpha}_i \in \mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$, but also (see the proof of lemma B.12) that $(p-1)(\boldsymbol{\alpha}_i + \boldsymbol{\alpha}_k) = \lambda \boldsymbol{\alpha}_j$ for some λ. This contradicts the fact that λ uniquely supports a point, since any point in $\mathcal{B}(\Omega)$ it supports is also supported by $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_k$. For $\mathcal{Z}_\pm^{(p)}$, i.e. if $\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_k$ is one endpoint for some $i \in I_{\pm 1}, k \in I_0$, then it similarly follows that $\boldsymbol{\alpha}_i \in \mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$, and that $\boldsymbol{\alpha}_k = \lambda \boldsymbol{\alpha}_j$, where $\lambda > 0$ by non-degeneracy.

Thus, $\mathcal{N}_\pm^{(p)}$ is excluded and $\mathcal{Z}_\pm^{(p)}$ can account for at most one endpoint of $\mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$, since the line segment between $\boldsymbol{\alpha}_i + \lambda(p-1)\boldsymbol{\alpha}_j$ and $\boldsymbol{\alpha}_{i'} + \lambda'(p-1)\boldsymbol{\alpha}_j$ cannot contain both $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_{i'}$, which are clearly contained in $\mathcal{F}_\pm \cap \mathcal{H}_\pm^{(p)}$. Hence, at least one of the endpoints is $\boldsymbol{\alpha}_i$ for some $i \in I_{\pm 1}$, and the other may either be $\boldsymbol{\alpha}_{i'}, i' \in I_{\pm 1}$ or $\boldsymbol{\alpha}_i + (p-1)\boldsymbol{\alpha}_j$ (possibly after exchanging $\langle \boldsymbol{\alpha}_j, 0 \rangle$ for an equivalent constraint). The former case was covered in the first paragraph of this proof, and is removed by considering sufficiently large p; the latter completes our proof. ∎

B.4.5 Visualisation of $\mathcal{A}_c(\Omega)$

We provided fig. II.2 for illustration along with the statement of proposition 4.1, since the shapes of $\mathcal{A}(\omega_c)$ are quite simple to interpret, and provide some insight into the result. The same cannot be said for the general $\mathcal{A}_c(\Omega)$, however, so we have put off a similar display until now. Figure II.26 showed a single example in great detail, and now figs. II.27 to II.29 illustrate $\mathcal{A}_c(\Omega)$ in a similar manner to how fig. II.2 illustrated $\mathcal{A}(\omega_c)$. For legibility, we have omitted the supporting sets $\mathcal{P}_\pm$, $\mathcal{Z}_\pm^{(p)}$ and $\mathcal{N}_\pm^{(p)}$, but if one wishes, it is not difficult to imagine them in the figures like in fig. II.26 to make sense of the shapes.

B.5 The duality between $\mathcal{A}_c(\Omega)$ and $\mathcal{B}(\Omega)$

As has been used many times above, parameter space and constraint space are dual in the sense that points in one correspond to hyperplanes in the other. This extends to $\mathcal{B}(\Omega)$ and $\mathcal{A}_c(\Omega)$, which are similarly related in ways we will explore in this section.

---

[63]This is made rigorous by comparing eq. (II.2.64) with the *Krein-Milman theorem*:

Let $\mathcal{C}$ be a compact convex set and $\mathcal{E}$ be the set of its extreme points. Then $\mathcal{C} = \text{Hull}(\mathcal{E})$, and if $\mathcal{C} = \text{Hull}(\mathcal{E}')$ then $\mathcal{E}' \supseteq \mathcal{E}$.
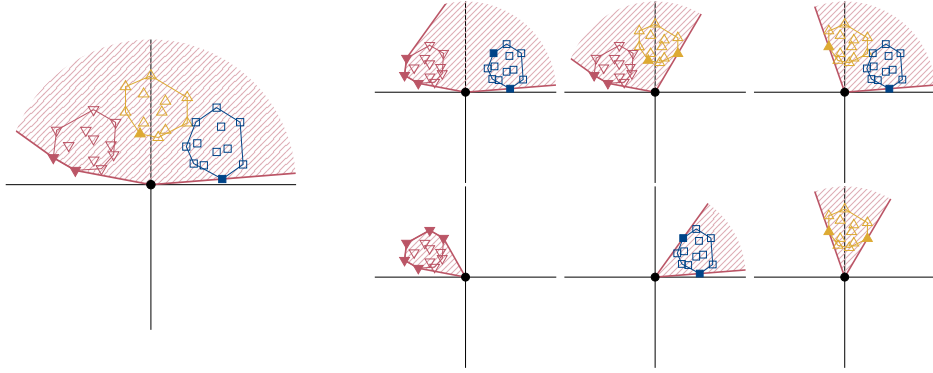
Figure II.27: Example of the region $\mathcal{A}_{-1}(\Omega)$ (red). The $\boldsymbol{\alpha}_i$, identical (up to $c_i = 0$ rescaling) to those in fig. II.26, are marked with $\triangledown$ for $i \in I_{-1}$, $\square$ for I_0 and $\triangle$ for I_1, and have their convex hulls outlined. The $\boldsymbol{\alpha}_i$ that are identified as relevant from the construction of $\mathcal{A}_{\pm 1}(\Omega)$, as per proposition B.3, are filled; the rest are left empty. Note how the $c_i = +1$ ones can strictly speaking only be deduced from $\mathcal{A}_1(\Omega)$ (fig. II.28 below), although their relevance can be easily seen in "nice" cases such as this. In the large figure, the full $\mathcal{A}_{-1}(\Omega)$ is drawn. The smaller figures demonstrate the effect of omitting $\boldsymbol{\alpha}_i$ from the definition of Ω, with i in various combinations of I_1, I_0 and I_{-1}. The only-I_{-1} figure (bottom left) is thus analogous to $\mathcal{A}(\omega_{-1})$ in fig. II.2. Note how in the only-I_1 figure (bottom right), proposition B.2 still works since $\mathcal{P}_{-1} = \{\mathbf{0}\}$ rather than being empty.

A taste of this duality can be found in the following result, which will also be useful further on:

Corollary B.5 (duality). *With $\Omega \neq \Omega_\infty$, let $\mho_\pm$ be defined as in eq. (II.2.67), so that $\mathcal{B}(\mho_\pm) = \mathcal{A}_{\pm 1}(\Omega)$. Then $\mathcal{B}(\Omega) = \mathcal{A}_1(\mho_+) \cap \mathcal{A}_{-1}(\mho_-)$ if $I_1 \neq \emptyset$, and $\mathcal{B}(\Omega) = \mathcal{A}_{-1}(\mho_-)$ otherwise.*

Loosely, one can think of this as "proposition B.2 is its own inverse": applying it to Ω gives $\mho_\pm$, and applying it to $\mho_\pm$ gives Ω.

PROOF. By definition and proposition B.2,

$$
\begin{aligned}
\boldsymbol{b} \in \mathcal{B}(\Omega) &\quad \Leftrightarrow \quad \forall \langle \boldsymbol{\alpha}, c \rangle \leq \Omega, \quad \boldsymbol{\alpha} \cdot \boldsymbol{b} \geq c, \\
\boldsymbol{b} \in \mathcal{A}_{\pm 1}(\mho_{\pm 1}) &\quad \Leftrightarrow \quad \mathcal{A}_{\pm 1}(\Omega) \subseteq \mathcal{B}(\langle \boldsymbol{b}, \pm 1 \rangle).
\end{aligned}
\tag{II.2.107}
$$

Looking at eq. (II.4.12), we see that the latter is equivalent to

$$
\forall \{\lambda_i\}_{i \in I} \text{ with } \lambda_i \geq 0 \text{ and } \sum_{i \in I} \lambda_i c_i \geq \pm 1, \qquad \boldsymbol{b} \cdot \sum_{i \in I} \lambda_i \boldsymbol{\alpha}_i \geq \pm 1.
\tag{II.2.108}
$$

This is true if $\forall i \in I$, $\boldsymbol{b} \cdot \boldsymbol{\alpha}_i \geq c_i$, but it only implies $\forall i \in I \setminus I_{\mp 1}$, $\boldsymbol{b} \cdot \boldsymbol{\alpha}_i \geq c_i$, as some straightforward algebra shows. By using $\mathcal{A}_1(\mho_+) \cap \mathcal{A}_{-1}(\mho_-)$, we avoid this shortcoming and complete the proof. When $I_1 = \emptyset$, we cannot use $\mho_+$ but also do not need it, since $I \setminus I_1 = I$. $\qquad \square$
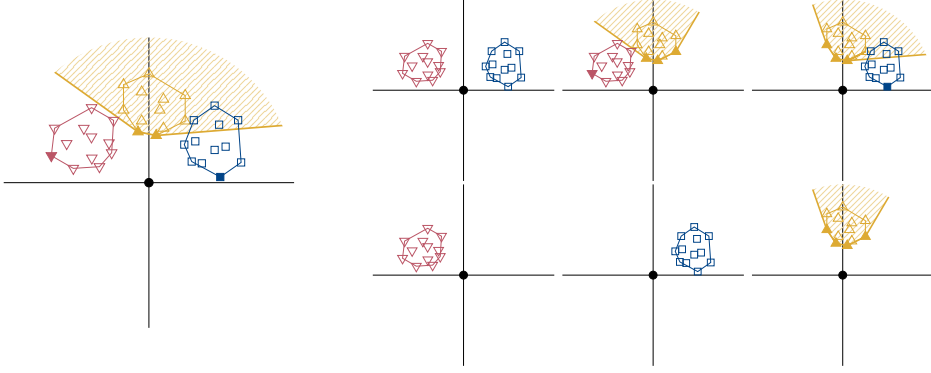
Figure II.28: Example of the region $\mathcal{A}_1(\Omega)$ (yellow), with the same α_i and analogous presentation as in fig. II.27. Note how some relevant $\alpha_i, i \notin I_1$ lie outside $\mathcal{A}_1(\Omega)$; this is also possible for $\mathcal{A}_0(\Omega)$, but not for $\mathcal{A}_{-1}(\Omega)$, since it contains all α_i. As in fig. II.27, the smaller figures illustrate the effect of omitting some α_i. The only-I_1 figure (bottom right) is thus analogous to $\mathcal{A}(\omega_1)$ in fig. II.2. Note how $\mathcal{A}_1(\Omega) = \emptyset$ when $i \in I_1$ are omitted, as can be seen from eq. (II.4.11) with $\sum_{i \in I_1} \lambda_i = 0$ (this is the exception to proposition B.2).

This relation can be extended into a geometric duality between $\mathcal{B}(\Omega)$ and $\mathcal{R}_c(\Omega)$: the vertices of one correspond directly to the facets of the other.[64]

Proposition B.4 (precise duality relations)**.** *Let $M_\pm^{(p)}$, etc. be defined as in proposition B.2, and let $\mho_\pm$ be defined as in eq. (II.2.67). Let $\pm 1 = 1$ if $I_1 \neq \emptyset$ and $\pm 1 = -1$ otherwise, as in corollary B.5. If Ω and $\mho_\pm$ are non-degenerate, then the following correspondences hold:*

(i) α_i *for $i \in I_{\pm 1}$ is a vertex of $\mathcal{A}_{\pm 1}(\Omega)$ if and only if $\langle \alpha_i, \pm 1 \rangle \in \mathcal{R}(\Omega)$.*

(ii) $\frac{1}{|r_\ell|} n_\ell$ *is a vertex of $\mathcal{B}(\Omega)$ if and only if $\ell \in M_\pm^{(p)}$ and $r_\ell \neq 0$.*

(iii) $\mathcal{F} \equiv \{ \beta \in \mathcal{A}_{\pm 1}(\Omega) \mid n_\ell \cdot \beta = r_\ell \}$ *is a facet of $\mathcal{A}_{\pm 1}(\Omega)$ if and only if $\ell \in M_\pm^{(p)}$.*

(iv) $\mathcal{F} \equiv \{ b \in \mathcal{B}(\Omega) \mid \alpha \cdot b = c \}$ *is a facet of $\mathcal{B}(\Omega)$ if and only if $\langle \alpha, c \rangle \in \mathcal{R}(\Omega)$.*

All relations are exhaustive: there is no vertex of $\mathcal{A}_{\pm 1}(\Omega)$ that is not covered by correspondence (i), etc. There is the single exception that $\mathbf{0}$ may be a vertex without corresponding to a facet.[65] Also, the exceptions to proposition B.2 apply.

This was used to obtain the visualisations of $\mathcal{B}(\Omega)$ in section 6. An illustrative example is given in fig. II.30.

---

[64]This is closely related to the concept of *dual polytopes*, where a polytope is a D-dimensional generalisation of a polygon or polyhedron; $\mathcal{B}(\Omega)$ and $\mathcal{A}_c(\Omega)$ are polytopes, if the definition is relaxed to permit unbounded polytopes.

[65]The origin is a vertex of $\mathcal{B}(\Omega)$ if $I_1 = \emptyset$ and $\{\alpha_i\}_{i \in I_0}$ is a basis of $\mathbb{R}^D$. It is never a vertex of (or even contained in) $\mathcal{A}_1(\Omega)$, but is always a vertex of $\mathcal{A}_{-1}(\Omega)$ if Ω is non-degenerate.
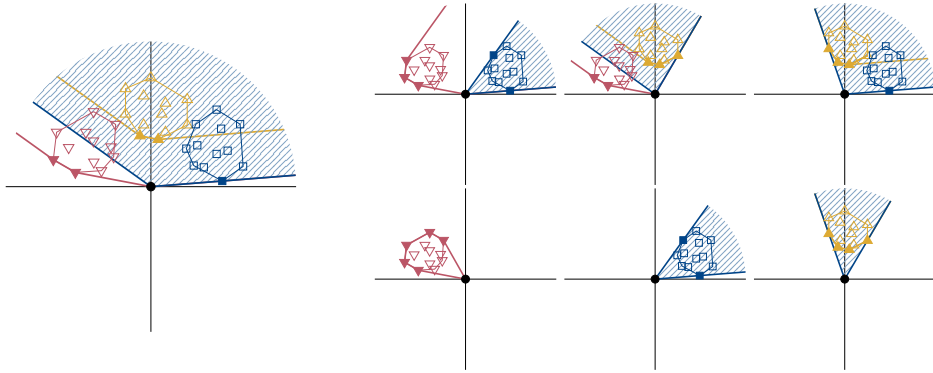
Figure II.29: Example of the region $\mathcal{A}_0(\Omega)$ (**blue**), with the same $\boldsymbol{\alpha}_i$ as in figs. II.27 and II.28. $\mathcal{A}_{\pm 1}(\Omega)$ are outlined to demonstrate how $\mathcal{A}_0(\Omega)$ can be constructed by taking their unbounded facets and shifting them so they pass through the origin, as discussed in section B.4.3. Note that unlike in figs. II.27 and II.28, not all relevant $\boldsymbol{\alpha}_i$ influence the shape of $\mathcal{A}_0(\Omega)$. As in figs. II.27 and II.28, the smaller figures illustrate the effect of omitting some $\boldsymbol{\alpha}_i$. The only-I_0 figure (bottom centre) is thus analogous to $\mathcal{A}(\omega_0)$ in fig. II.2. Note how in the only-I_{-1} figure (bottom left), $\mathcal{A}_0(\Omega) = \{\mathbf{0}\}$ since $\mathcal{A}_{-1}(\Omega)$ has no unbounded faces (this is one of the exceptions to corollary B.3).

Note how proposition B.4 is rather negligent of $c = 0$ constraints, partly because they are complicated to handle, and partly because an exact zero is an unlikely thing when constraints are generated with numerical inaccuracies. For similar reasons, we do not consider the degenerate case.

PROOF. Correspondence (i) is just proposition B.3.

Correspondence (ii), which is the most useful correspondence for visualisation, follows by applying proposition B.3 to $\mho_{\pm}$ instead. To do this, we first apply the normalisation

$$\langle \boldsymbol{n}_\ell, r_\ell \rangle \quad \rightarrow \quad \begin{cases} \langle \frac{1}{|r_\ell|} \boldsymbol{n}_\ell, \pm 1 \rangle & \text{if } r_\ell \neq 0, \\ \langle \boldsymbol{n}_\ell, 0 \rangle & \text{otherwise,} \end{cases} \tag{II.2.109}$$

recalling lemma B.8. Granted that $\mho_{\pm}$ is non-redundant, proposition B.3 therefore states that the extreme points (vertices) of $\mathcal{A}_{\pm 1}(\mho_{\pm})$, except $\mathbf{0}$, are exactly $\frac{1}{|r_\ell|} \boldsymbol{n}_\ell$, since $\langle \frac{1}{|r_\ell|} \boldsymbol{n}_\ell, \pm 1 \rangle$ are exactly the relevant elements of $\mho_{\pm}$ with $r_\ell \neq 0$. The correspondence then follows from corollary B.5, as long as we can ensure that the intersection does not introduce any new vertices that are not vertices of $\mathcal{A}_{\pm 1}(\mho_{\pm})$. Such a vertex would, by lemma B.3, be uniquely supported by some constraint that is unaccounted for by proposition B.3, so it does not exist.

Correspondences (iii) and (iv) both follow from proposition 4.3, since each facet contains all points that are uniquely supported by some relevant constraint. It follows from corollary B.2 and lemma B.3 that a point $\boldsymbol{b}$ is uniquely supported by some constraint if and
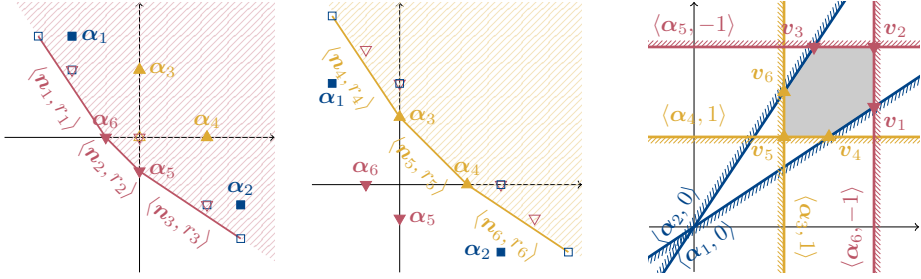
Figure 11.30: Example of how $\mathcal{A}_{\pm 1}\Omega$ are dual to $\mathcal{B}(\Omega)$. A different, simpler Ω than in previous figures is used. **Left:** Constraint space, with $\mathcal{A}_{-1}(\Omega)$ constructed in a similar way to fig. 11.26. The points $\boldsymbol{\alpha}_i$ are marked with filled dots and labelled; points in $\mathcal{P}_-(\Omega)$, $\mathcal{Z}_-^{(2)}(\Omega)$ and $\mathcal{N}_-^{(2)}(\Omega)$ are marked with empty dots. The facets corresponding to constraints $\langle \boldsymbol{n}_\ell, r_\ell \rangle$ are also labelled. **Middle:** Constraint space, with $\mathcal{A}_1(\Omega)$ constructed similarly. **Right:** Parameter space, drawn similarly to fig. 11.1. $\mathcal{B}(\Omega)$ is shaded; the six constraints $\langle \boldsymbol{\alpha}_i, c_i \rangle$ and vertices $\boldsymbol{v}_\ell = \boldsymbol{n}_\ell \frac{\pm 1}{r_\ell}$ are labelled. Vertices from $\mathcal{A}_{-1}(\Omega)$ are marked with ▼, and those from $\mathcal{A}_1(\Omega)$ with ▲. Note that two vertices sharing an edge of $\mathcal{B}(\Omega)$ corresponds to two constraints supporting a common point in $\mathcal{A}_{\pm 1}\Omega$. In the case of $c = 0$ edges, it corresponds to supporting $(\boldsymbol{\alpha}_i + p\boldsymbol{\alpha}_j) \in \mathcal{Z}_+^{(p)}(\Omega)$ and $(\boldsymbol{\alpha}_{i'} + p\boldsymbol{\alpha}_j) \in \mathcal{Z}_-^{(p)}(\Omega)$ for a common $j \in I_0$. Note also how corollary B.5 applies (compare fig. 11.2).

only if $\boldsymbol{b}$ is contained in some $(D-1)$-face, but not in any K-face for $K < D - 1$. It also follows from the definition of a K-face that all K-faces contain at least one point that is not contained in any K'-face, $K' < K$. Thus, every facet gives rise to a relevant constraint, and no constraint can be relevant if it does not give rise to a facet. ∎

B.5.1 Duality with a bounding box

It would be possible to extend proposition B.4 to relate the $\boldsymbol{n}_\ell$ with $r_\ell = 0$ to unbounded edges of $\mathcal{B}(\Omega)$ by equating them to the ray $\{\frac{1}{|r_k|}\boldsymbol{n}_k + \lambda \boldsymbol{n}_\ell \mid \lambda \geq 0\}$ for some k (see proposition B.3). However, it is not obvious how to find k, and it is not guaranteed that we exhaustively cover the unbounded edges this way. In either case, unbounded edges are far less pleasant to deal with than vertices, even though they do provide shape information that the vertices alone cannot provide.

Unbounded edges can be wholly avoided by artificially introducing $2D$ extra constraints that constrain $\mathcal{B}(\Omega)$ to a D-dimensional bounding box. All unbounded facets of $\mathcal{B}(\Omega)$ are cropped, and proposition B.4, correspondence (ii) seamlessly provides all the vertices that define the intersections between $\mathcal{B}(\Omega)$ and the walls of the box. This method, demonstrated in fig. 11.31, was used extensively in section 6.

Everything mentioned here can of course be applied equally well to $\mathcal{R}_{\pm 1}(\Omega)$.
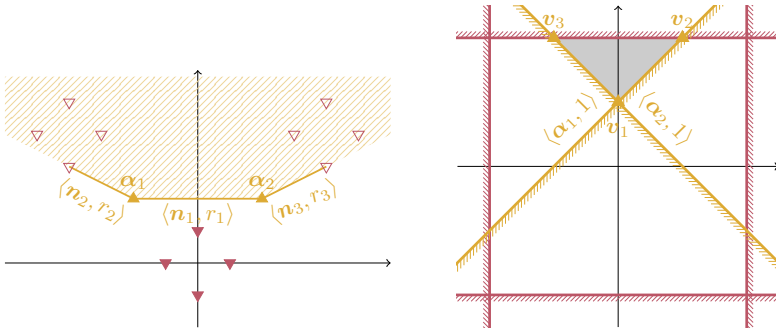
Figure 11.31: Example of how $\mathcal{B}(\Omega) = \langle\boldsymbol{\alpha}_1, 1\rangle + \langle\boldsymbol{\alpha}_2, 1\rangle$, which would normally have a single vertex $\boldsymbol{v}_1$ and two unbounded edges corresponding to $r_k = 0$ constraints, is made more manageable by imposing an artifical set of bounding-box constraints ($\blacktriangledown$), giving rise to some points in $\mathcal{N}_+^{(2)}(\Omega)$ ($\triangledown$). Like in fig. 11.30, $\mathcal{A}_1(\Omega)$ is shown on the left, and $\mathcal{B}(\Omega)$ is shown on the right.

B.6 A note on infinite sums of constraints

It is extremely important that we only ever consider sums of a finite number of constraints, not only for the validity of our proofs, but also for the validity of the propositions themselves. Consider as a counterexample the following countably infinite sum of one-dimensional constraints:

$$\omega_{-1} = \sum_{k=1}^{\infty} \langle 1 - \tfrac{1}{k}, -1\rangle \quad \Rightarrow \quad \mathcal{A}(\omega_{-1}) = [0, 1), \quad \mathcal{B}(\omega_{-1}) = [-1, \infty). \qquad (\text{11.2.110})$$

Here, $\mathcal{A}(\omega_{-1})$ is not a closed set. On the other hand, $\mathcal{B}(\omega_{-1})$ is closed, since for any $\varepsilon > 0$, the point $-(1+\varepsilon)$ fails to satisfy the element $\langle 1 - \tfrac{1}{k}, -1\rangle$ for sufficiently large k (specifically, larger than $\tfrac{1+\varepsilon}{\varepsilon}$). Thus, $\mathcal{B}(\omega_{-1}) = \mathcal{B}(\langle 1, -1\rangle)$, so $\langle 1, -1\rangle \leq \omega_{-1}$ even though $1 \notin \mathcal{A}(\omega_{-1})$. Thus, proposition 4.1, and therefore proposition 4.2, fails.

There is, however, a straightforward generalisation. We let eqs. (11.4.9) and (11.4.10) define Ω and $\mathcal{A}_c(\Omega)$ as before, but now with infinite sums permitted, so that $\{\boldsymbol{\alpha}_i\}_{i \in I_c}$ may be any subset of $\mathbb{R}^D$. Then we have the following:

Proposition B.5 (determining if constraint is weaker, generalised). *Let $\langle\boldsymbol{\beta}, c\rangle$ be a linear constraint, and let $\Omega \neq \Omega_\infty$ be a possibly infinite sum of linear constraints. Then $\langle\boldsymbol{\beta}, c\rangle \leq \Omega$ if and only if $\boldsymbol{\beta} \in \mathrm{cl}\,[\mathcal{A}_c(\Omega)]$.*

The simple introduction of the closure solves all issues with infinite combinations, such as the counterexample above. Note how the only part of sections B.1.2 to B.1.4 that relies on the finiteness of I is the proof that $\mathcal{A}_c(\Omega)$ is closed. The closure is by definition closed, so the remaining arguments in these sections remain valid when applied to $\mathrm{cl}\,[\mathcal{A}_c(\Omega)]$ instead.

The only part of the proof of proposition 4.2 that does not immediately carry over is section B.1.1. However, by the limit definition of closure, for any $\boldsymbol{\beta} \in \mathrm{cl}[\mathcal{A}_c(\Omega)]$ there exists

a sequence $\boldsymbol{\beta}_n$ in $\mathcal{A}_c(\Omega)$ such that $\lim_{n\to\infty}\boldsymbol{\beta}_n = \boldsymbol{\beta}$. Since the function $\boldsymbol{\beta} \to \boldsymbol{\beta} \cdot \boldsymbol{b}$ is continuous, it follows that for any $\boldsymbol{b}$, $\lim_{n\to\infty}\boldsymbol{\beta}_n \cdot \boldsymbol{b} = \boldsymbol{\beta} \cdot \boldsymbol{b}$. Thus, since the arguments of section B.1.1 hold for all $\boldsymbol{\beta}_n$, they also hold for $\boldsymbol{\beta} \in \mathrm{cl}[\mathcal{A}_c(\Omega)]$. ∎

Corollary B.1 does not require any adjustment, since the interior of any convex set in $\mathbb{R}^D$ is equal to the interior of its closure.

The generalisation of proposition 4.3, for which counterexamples abound, is less straightforward, partly because there does not necessarily exist a minimal representation if all representations are infinite. However, the notion of a minimal representation is mainly motivated as being the most practical format of a constraint, so it is not very useful in the infinite case, which is anyway only of theoretical interest. For the same reason, it is not relevant to our study to attempt to generalise any of the propositions presented in this appendix.

B.7   Mathematical glossary

The table below contains a list of notations and terms that may be unfamiliar to some readers, depending on their background. (We have chosen to employ such notation, since it makes some things much more brief and expressive, even though it necessitates this table.)

Concept	Description
$\square, \blacksquare$	**End of proof.** We use $\blacksquare$ for main proofs and $\square$ for lemmata.
$\forall$	**The universal quantifier.** Informally, short for "for all".
$\exists$	**The existence quantifier.** Informally, short for "there exists".
$x \in \mathcal{X}$	**The object x is contained in (is an element of) the set $\mathcal{X}$.** We typically denote sets using a calligraphic font, but use ordinary italics for sets of indices.
$\emptyset$	**The empty set.** The set that contains no elements.
$|\mathcal{X}|, |\boldsymbol{v}|, |s|$	**Cardinality or magnitude.** For a set, this indicates its cardinarlity (number of elements). For a vector or scalar, the same notation indicates its magnitude $\sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$ or $\sqrt{s^2}$.
$\mathcal{X} \subseteq \mathcal{Y}$	**$\mathcal{X}$ is a subset of $\mathcal{Y}$.** All elements of $\mathcal{X}$ are also contained in $\mathcal{Y}$. Just $\mathcal{X} \subset \mathcal{Y}$ means the same, except that $\mathcal{X} \neq \mathcal{Y}$; that is, $\mathcal{Y}$ has at least one element not contained in $\mathcal{X}$.
$\mathcal{X} \cup \mathcal{Y}$	**The union of two sets.** The set that contains all elements contained in either $\mathcal{X}$ or $\mathcal{Y}$, or both. Clearly, $\mathcal{X} \subseteq \mathcal{X} \cup \mathcal{Y}$ and $\mathcal{Y} \subseteq \mathcal{X} \cup \mathcal{Y}$.
$\mathcal{X} \cap \mathcal{Y}$	**The intersection of two sets.** The set that contains all elements contained in *both* $\mathcal{X}$ or $\mathcal{Y}$. Clearly, $\mathcal{X} \supseteq \mathcal{X} \cup \mathcal{Y}$ and $\mathcal{Y} \supseteq \mathcal{X} \cup \mathcal{Y}$.
Disjoint sets	Said of two sets $\mathcal{X}, \mathcal{Y}$ if $\mathcal{X} \cap \mathcal{Y} = \emptyset$.
$\mathcal{X} \setminus \mathcal{Y}$	**Relative complement.** Consists of all elements of $\mathcal{X}$ that are not also elements of $\mathcal{Y}$. Has the properties $(\mathcal{X} \setminus \mathcal{Y}) \cap \mathcal{Y} = \emptyset$, $(\mathcal{X} \setminus \mathcal{Y}) \cup \mathcal{Y} = \mathcal{X}$.
Complement	The complement of $\mathcal{X}$ consists of all elements not contained in $\mathcal{X}$; equal to $\mathcal{U} \setminus \mathcal{X}$ where $\mathcal{U}$ is the implicit "universal" set, e.g. $\mathbb{R}^D$ for sets of D-dimensional vectors.
$\bigcup, \bigcap$	**The union/intersection of many sets.** Used similarly to $\sum, \prod$.

Concept	Description
$\{x \mid A, B, \ldots\}$	**Set-builder notation.** Denotes the set of all objects x for which all conditions $A, B, \ldots$ are true. $\{x \mid x \in \mathcal{X}, A, \ldots\}$ is more compactly written $\{x \in \mathcal{X} \mid A, \ldots\}$.
$\{a, b, c\}$, $\{a_i\}_{i \in I}$, $\{a_n\}_{n=0}^{N}$	**Various shorthands used to define sets.** The latter two are equivalent to $\{a_i \mid i \in I\}$ and $\{a_n \mid 0 \le n \le N\}$, respectively.
Open set	A subset $\mathcal{X} \subseteq \mathbb{R}^D$ is open if for each $\boldsymbol{\chi} \in \mathcal{X}$, there is some $\varepsilon > 0$ such that $\boldsymbol{\xi} \in \mathcal{X}$ for all $|\boldsymbol{\xi} - \boldsymbol{\chi}| < \varepsilon$. This is for subsets of $\mathbb{R}^D$; definitions of openness exist for more general sets, but we do not use them.
Closed set	A set whose complement is open.
Finite set	A set that contains a finite (or zero) number of elements. Many intuitive properties, such as the existence of a smallest subset with a given property, are only guaranteed for finite sets.
Bounded set	A set $\mathcal{X} \subseteq \mathbb{R}^D$ is bounded if there is some M such that $|\boldsymbol{\chi}| < M$ for all $\boldsymbol{\chi} \in \mathcal{X}$.
Compact set	A set that is closed and bounded. This is for subsets of $\mathbb{R}^D$; we do not use the more general versions. Many properties of finite sets carry over to compact sets, such as having a (not necessarily unique) element that is the minimum or maximum of some property.
$\mathrm{cl}(\mathcal{X})$	**Closure.** The smallest closed set that has $\mathcal{X}$ as a subset. Equal to $\mathcal{X}$ itself if it is closed.
$\mathrm{int}(\mathcal{X})$	**Interior.** The largest open set that is a subset of $\mathcal{X}$. Equal to $\mathcal{X}$ itself if it is open.
Boundary	The boundary of $\mathcal{X}$ is $\mathrm{cl}(\mathcal{X}) \setminus \mathrm{int}(\mathcal{X})$. Note that a non-closed set does not necessarily contain its boundary, and that *boundary* and *boundedness* are unrelated concepts.
$(a, b), [a, b), [a, b]$	**Open, half-open and closed intervals.** Denotes the range between a and b. A square bracket indicates that the endpoint is included in the interval, a parenthesis that it is not. Note the imperfect agreement with the concept of open and closed sets: (a, b) is open and $[a, b]$ is closed for finite a, b, but $[a, b)$ is neither open nor closed. Also, intervals such as $[a, \infty)$ and $(-\infty, \infty)$ are closed.
Linear/affine/convex combination	A sum of vectors of the form $\sum_i a_i \boldsymbol{v}_i$. It is a linear combination for all a_i. It is an affine combination if $\sum_i a_i = 1$. It is a convex combination if $\sum_i a_i = 1$ and $a_i \ge 0$ for all i.
Convex set	A set that contains all convex combinations of its elements.
$\mathrm{Hull}(\mathcal{X})$	**Convex hull.** The set of all convex combinations of the elements of $\mathcal{X}$. Equivalently, the smallest convex set that contains $\mathcal{X}$.
Linear/affine span	Of a set of vectors: the set of all linear/affine combinations thereof. The linear span is often just called "span".
Linear/affine dimension	The linear/affine dimension of a set $\mathcal{X} \subseteq \mathbb{R}^D$ is the smallest number of vectors whose linear/affine span contains $\mathcal{X}$. The dimension is $\le D$. When it is clear from context, just "dimension" is often used.
Linear/affine subspace	A subset of $\mathbb{R}^D$ that contains all linear/affine combinations of its elements. Equivalently, the linear/affine span of some set of vectors.
Hyperplane	An affine subspace of affine dimension $D-1$. Generalises the notion of a plane in 3D space and a line in 2D space. Given a nonzero vector $\boldsymbol{v}$, the set $\{\boldsymbol{x} \mid \boldsymbol{v} \cdot \boldsymbol{x} = u\}$ is a distinct hyperplane for each u. The intersection of K hyperplanes in $\mathbb{R}^D$ with linearly independent $\boldsymbol{v}$ vectors is an affine subspace of dimension $D - K$.
$\mathcal{A} + \mathcal{B}$	**Minkowski sum.** The set $\{\boldsymbol{a} + \boldsymbol{b} \mid \boldsymbol{a} \in \mathcal{A}, \boldsymbol{b} \in \mathcal{B}\}$ (assuming the elements of $\mathcal{A}, \mathcal{B}$ support addition).
$x\mathcal{A}$	The set $\{x\boldsymbol{a} \mid \boldsymbol{a} \in \mathcal{A}\}$ (assuming the elements of $\mathcal{A}$ support scalar multiplication).

C   The loop integral functions

This appendix contains details on the functions appearing in the NLO and NNLO amplitudes, which originate in loop integrals. Sections C.1 and C.2 contain expansions important to their numerical evaluation, and section C.3 contains the derivation of their analytic integrals over the Mandelstam variables.

Using the conventions of [34],[66] the function $\bar{J}$ can be defined as

$$\bar{J} = \pi_{16}(\beta^2 h + 2), \qquad h = \frac{1}{\beta}\ln\frac{\beta - 1}{\beta + 1}, \qquad \beta = \sqrt{1 - \frac{4}{s}}, \tag{II.3.1}$$

where $\pi_{16} \equiv 1/16\pi^2$. To reduce clutter, we define $\hat{J} \equiv \bar{J}/\pi_{16}$. Similarly, with $\hat{k}_i \equiv k_i/\pi_{16}^2$ for brevity, the additional functions at NNLO are defined as

$$\hat{k}_1 = \beta^2 h^2, \qquad \hat{k}_3 = \frac{\beta^2 h^3}{s} + \frac{\pi^2 h}{s} - \frac{\pi^2}{2},$$

$$\hat{k}_2 = \beta^4 h^2 - 4, \qquad \hat{k}_4 = \frac{1}{s\beta^2}\left[\frac{\hat{k}_1}{2} + \frac{\hat{k}_3}{3} + \hat{J} + \frac{(\pi^2 - 6)s}{12}\right]. \tag{II.3.2}$$

$\bar{J}$ and k_i are real below threshold, and are finite as $s \to 4$ from below. However, β is not real when $0 < s < 4$, which poses a problem for numerical evaluation. This can be remedied by defining $\tilde{\beta} = -i\beta$ and rewriting h as

$$h = \frac{2\tan^{-1}(1/\tilde{\beta})}{\tilde{\beta}}, \tag{II.3.3}$$

which can be evaluated for $0 < s < 4$ using only real numbers.

The functions have further numerical problems. β diverges at $s = 0$, which leaves $\bar{J}$ and k_i with removable singularities there. These are rendered harmless with a series expansion, as shown below. As $s \to 4$ from below, h diverges while $\bar{J}$ and k_i stay finite. Reliable evaluation of this limit also requires series expansion. The derivatives of $\bar{J}$ and k_i diverge in this limit (starting at the second derivative for k_2 and the first derivative for the others), which also necessitates series expansion for reliable handling.

C.1   Expanding around $s = 0$

Near $s = 0$, we make an expansion in $s = 4\varepsilon^2$, where ε may be complex. To make the NNLO unitarity corrections numerically well-behaved in all cases used by us, an $\mathcal{O}(\varepsilon^8) = \mathcal{O}(s^4)$ expansion is needed. For β and $1/\beta$, it is

$$\beta = \frac{\sqrt{\varepsilon^2 - 1}}{\varepsilon} = \frac{i}{\varepsilon}\left(1 - \frac{\varepsilon^2}{2} - \frac{\varepsilon^4}{8} - \frac{\varepsilon^6}{16} - \frac{5\varepsilon^8}{128} + \mathcal{O}(\varepsilon^{10})\right), \tag{II.3.4}$$

[66]We use β rather than σ for consistency with eq. (II.5.4).

$$\frac{1}{\beta} = \frac{\varepsilon}{\sqrt{\varepsilon^2 - 1}} = -i\varepsilon\left(1 + \frac{\varepsilon^2}{2} + \frac{3\varepsilon^4}{8} + \frac{5\varepsilon^6}{16} + \frac{35\varepsilon^8}{128} + \mathcal{O}(\varepsilon^{10})\right). \tag{II.3.5}$$

By using the expansion

$$h = -2\sum_{n=1}^{\infty} \frac{1}{(2n-1)\beta^{2n}}, \tag{II.3.6}$$

it follows that

$$h = \;\; 2\varepsilon^2 \quad\quad + \tfrac{4}{3}\varepsilon^4 \quad\quad\quad + \tfrac{16}{15}\varepsilon^6 \quad\quad\quad\quad + \tfrac{32}{35}\varepsilon^8, \tag{II.3.7a}$$

$$\hat{J} = \;\; \tfrac{2}{3}\varepsilon^2 \quad\quad + \tfrac{4}{15}\varepsilon^4 \quad\quad\quad + \tfrac{16}{105}\varepsilon^6 \quad\quad\quad\quad + \tfrac{32}{35}\varepsilon^8 \tag{II.3.7b}$$

$$\hat{k}_1 = -4\varepsilon^2 \quad\quad - \tfrac{4}{3}\varepsilon^4 \quad\quad\quad - \tfrac{32}{45}\varepsilon^6 \quad\quad\quad\quad - \tfrac{16}{35}\varepsilon^8 \tag{II.3.7c}$$

$$\hat{k}_2 = -\tfrac{8}{3}\varepsilon^2 \quad\quad - \tfrac{28}{15}\varepsilon^4 \quad\quad\quad - \tfrac{64}{63}\varepsilon^6 \quad\quad\quad\quad - \tfrac{592}{175}\varepsilon^8 \tag{II.3.7d}$$

$$\hat{k}_3 = \left(\tfrac{\pi^2}{3} - 2\right)\varepsilon^2 + \left(\tfrac{4\pi^2}{15} - 2\right)\varepsilon^4 + \left(\tfrac{8\pi^2}{45} - \tfrac{28}{15}\right)\varepsilon^6 \;\; + \tfrac{328}{189}\varepsilon^8 \tag{II.3.7e}$$

$$\hat{k}_4 = \left(1 - \tfrac{\pi^2}{9}\right)\varepsilon^2 + \left(\tfrac{19}{15} - \tfrac{\pi^2}{15}\right)\varepsilon^4 + \left(\tfrac{464}{315} - \tfrac{16\pi^2}{105}\right)\varepsilon^6 + \left(\tfrac{820}{567} - \tfrac{16\pi^2}{105}\right)\varepsilon^8, \tag{II.3.7f}$$

where we have omitted "$+\mathcal{O}(\varepsilon^{10})$" for brevity. These functions are real, so only even powers of ε appear. Note also how all numerical coefficients stay roughly order 1.

C.2   Expanding around $s = 4$

As $s \to 4$ from below, we expand $s = 4(1 - \delta^2)$ with $\delta > 0$, and again need an eighth-order expansion at NNLO. The expansion of β and $1/\beta$ is

$$\beta = \frac{\delta}{\sqrt{\delta^2 - 1}} = -i\delta\left(1 + \tfrac{1}{2}\delta^2 + \tfrac{3}{8}\delta^4 + \tfrac{5}{16}\delta^6 + \tfrac{35}{128}\delta^8 + \mathcal{O}(\delta^{10})\right), \tag{II.3.8}$$

$$\frac{1}{\beta} = \frac{\sqrt{\delta^2 - 1}}{\delta} = \;\; \frac{i}{\delta}\left(1 - \tfrac{1}{2}\delta^2 - \tfrac{1}{8}\delta^4 - \tfrac{1}{16}\delta^6 - \tfrac{5}{128}\delta^8 + \mathcal{O}(\delta^{10})\right). \tag{II.3.9}$$

By using the expansion

$$h = -\frac{i\pi}{\beta} + 2i\sum_{n=0}^{\infty} \frac{\beta^{2n}}{2n+1}, \tag{II.3.10}$$

it follows that (suppressing "$+\mathcal{O}(\delta^{10})$")

$$h = \frac{\pi}{\delta} - 2\left[1 \;\; + \tfrac{1}{3}\delta^2 \;\; + \tfrac{2}{15}\delta^4 + \tfrac{8}{105}\delta^6 + \tfrac{16}{315}\delta^8\right] - \frac{\pi\delta}{2}\left[1 + \tfrac{1}{4}\delta^2 \;\; + \tfrac{1}{8}\delta^4 \;\; + \tfrac{5}{64}\delta^6 \;\;\right], \tag{II.3.11a}$$

$$\hat{J} = \quad 2\left[1 \;\; + \delta^2 \;\; + \tfrac{2}{3}\delta^4 \;\; + \tfrac{8}{15}\delta^6 \;\; + \tfrac{16}{35}\delta^8 \;\;\right] - \pi\delta\left[1 + \tfrac{1}{2}\delta^2 \;\; + \tfrac{3}{8}\delta^4 \;\; + \tfrac{5}{16}\delta^6 \;\;\right], \tag{II.3.11b}$$

$$\hat{k}_1 = \quad -4\left[\tfrac{\pi^2}{4} + \delta^2 \;\; + \tfrac{1}{3}\delta^4 \;\; + \tfrac{8}{45}\delta^6 \;\; + \tfrac{4}{35}\delta^8 \;\;\right] + \pi\delta\left[4 + \tfrac{2}{3}\delta^2 \;\; + \tfrac{3}{10}\delta^4 \;\; + \tfrac{5}{28}\delta^6 \;\;\right], \tag{II.3.11c}$$

$$\hat{k}_2 = \quad -4 \left[1 \qquad\qquad - \delta^4 \quad - \tfrac{4}{3}\delta^6 \quad - \tfrac{68}{45}\delta^8 \right] - \pi\delta \left[\qquad 4\delta^2 \quad + \tfrac{14}{3}\delta^4 \quad + \tfrac{149}{30}\delta^6 \right]$$
$$+ \pi^2 \left[\qquad \delta^2 \quad + \delta^4 \quad + \delta^6 \quad + \delta^8 \qquad \right], \tag{II.3.11d}$$

$$\hat{k}_3 = \quad 2 \left[\qquad \delta^2 \quad + \delta^4 \quad + \tfrac{14}{15}\delta^6 \quad + \tfrac{164}{189}\delta^8 \right] - \pi\delta \left[3 + \tfrac{5}{2}\delta^2 \quad + \tfrac{259}{120}\delta^4 + \tfrac{3229}{1680}\delta^6 \right]$$
$$+ \pi^2 \left[\tfrac{1}{2} \quad + \tfrac{2}{3}\delta^2 \quad + \tfrac{8}{15}\delta^4 + \tfrac{16}{35}\delta^6 \quad + \tfrac{128}{315}\delta^8 \right], \tag{II.3.11e}$$

$$\hat{k}_4 = \quad -\frac{1}{3} \left[2 \quad + \delta^2 \quad + \delta^4 \quad + \tfrac{128}{105}\delta^6 + \tfrac{844}{945}\delta^8 \right] + \frac{\pi\delta}{4} \left[1 + \tfrac{17}{18}\delta^2 + \tfrac{311}{360}\delta^4 + \tfrac{2227}{2800}\delta^6 \right]$$
$$+ \frac{\pi^2}{3} \left[\tfrac{1}{12} \quad - \tfrac{2}{15}\delta^2 - \tfrac{4}{35}\delta^4 - \tfrac{32}{315}\delta^6 + \tfrac{187}{1260}\delta^8 \right] - \frac{\pi^3\delta}{4} \left[\qquad\qquad\qquad \tfrac{121}{1536}\delta^6 \right]. \tag{II.3.11f}$$

Since δ is real, odd powers of δ are permitted. Note again how the coefficients stay roughly order 1. Due to $\mathrm{d}\delta/\mathrm{d}s = -1/8\delta$, the derivative of the linear terms diverges as $\delta \to 0$. Since $\hat{k}_2$ lacks a linear term, its derivative remains finite. Since the derivative changes the powers in steps of two, only odd negative powers appear in derivatives of any order.

C.3 Integrals above threshold

Here, we seek to analytically determine the function $D_k^J(\lambda, v, t)$ defined in eq. (II.5.19). In the relevant z range, $\beta(z)$ remains real while $h(z)$ obtains an imaginary part.[67] The most convenient form of h is

$$h(z) = \frac{H(\beta) + i\pi}{\beta}, \qquad H(\beta) \equiv \ln\frac{1-\beta}{1+\beta} = -\tfrac{1}{2}\tanh^{-1}(\beta), \tag{II.3.12}$$

where we take the branch with positive imaginary part. We make the following easily verifiable observations about the real function $H(\beta)$:

$$H' = \frac{-2}{1-\beta^2} = -\tfrac{1}{2}z, \tag{II.3.13a} \qquad\qquad \beta^2 = 1 + \frac{2}{H'}, \tag{II.3.13c}$$

$$H'' = -\beta H'^2, \tag{II.3.13b} \qquad \mathrm{d}z = -2H''\mathrm{d}\beta = 2\beta H'^2 \mathrm{d}\beta. \tag{II.3.13d}$$

$D_k^J(\lambda, v, t)$ will be a linear combination of integrals of the form $\int_4^\lambda \frac{z^n\,\mathrm{d}z}{(z-v)^{k+1}} \operatorname{Im} X(z)$, where X is one of $\bar{J}$ and k_i. We will ignore the denominator for now, and show later how to reduce all integrals to the form $\int z^n \operatorname{Im} X(z)\,\mathrm{d}z$. This will involve a wide range of values for n, so it is easiest to treat general n and then read off the special cases. Reading from eqs. (II.3.1) and (II.3.2), we find that above threshold,

$$\operatorname{Im}\hat{J} = \pi\beta, \qquad \operatorname{Im}\hat{k}_1 = 2\pi H, \qquad \operatorname{Im}\hat{k}_2 = 2\pi\beta^2 H,$$
$$\operatorname{Im}\hat{k}_3 = \frac{3\pi H^2}{z\beta}, \qquad \operatorname{Im}\hat{k}_4 = \frac{1}{z\beta^2}\left[H + \frac{H^2}{z\beta} + \beta \right]. \tag{II.3.14}$$

---

[67] We will drop the dependence on z from now on; everything in this section implicitly depends on it unless otherwise specified.

Using eq. (II.3.13), we see that everything can be expressed in terms of the functions

$$Z_p^{m,n}(\beta) \equiv \int \zeta_p^{m,n}(\beta)\, \mathrm{d}\beta, \qquad \zeta_p^{m,n}(\beta) \equiv \frac{H^m H'^n}{\beta^p}. \tag{II.3.15}$$

Thanks to eq. (II.3.13b), the family of functions $\zeta_p^{m,n}$ is closed under derivatives:

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\zeta_p^{m,n} = m\zeta_p^{m-1,n+1} - n\zeta_{p-1}^{m,n+1} - p\zeta_{p+1}^{m,n}. \tag{II.3.16}$$

It is therefore our hope that the highly nontrivial integral $Z_p^{m,n}$ can mostly be expressed as a sum of $\zeta_{p'}^{m',n'}$, plus some special cases (for instance, $\zeta_p^{m,0}$ is not the derivative of another $\zeta_{p'}^{m',n'}$). We will therefore attempt to find recurrence relations on m, n, p that allow Z to be reduced to ζ's and a few special cases.

C.3.1 The integral $Z_p^{m,n}$ for $p = 0$

First, we note that n can be assumed non-negative, since

$$Z_p^{m,n} = \tfrac{1}{2}\big[Z_{p-2}^{m,n+1} - Z_p^{m,n+1}\big], \tag{II.3.17}$$

according to eq. (II.3.13c).[68] Then, before treating general p, we consider $Z^{m,n} \equiv Z_0^{m,n}$ for $n, m \geq 0$. The integral of eq. (II.3.16) gives

$$\begin{aligned}
\zeta^{m-1,n-1} &= (m-1)Z^{m-2,n} - (n-1)Z_{-1}^{m-1,n}, \\
\zeta_{-1}^{m,n-1} &= mZ_{-1}^{m-1,n} - (n-1)Z_{-2}^{m,n} + Z^{m,n-1},
\end{aligned} \tag{II.3.18}$$

where the second line allows for the removal of $Z_{-1}^{m-1,n}$ in the first. Equation (II.3.17) can be invoked to turn $Z_{-2}^{m,n}$ into $2Z^{m,n-1} + Z^{m,n}$, and after extracting $Z^{m,n}$, we get a recurrence relation where both m and n decrease:

$$Z^{m,n} = \frac{3-2n}{n-1}Z^{m,n-1} - \frac{\zeta_{-1}^{m,n-1}}{n-1} + \frac{m}{(n-1)^2}\big[(m-1)Z^{m-2,n} - \zeta^{m-1,n-1}\big]. \tag{II.3.19}$$

This is valid for all $m \geq 0, n > 1$. The $n = 1$ case is covered by the trivial identity[69]

$$Z^{m,1} = \frac{\zeta^{m+1,0}}{m+1}. \tag{II.3.20}$$

$Z^{m,0}$ requires some more thought and will be treated later (see eq. (II.3.26)).

---

[68]We will not need to treat negative m when integrating the loop integral functions. As we will see below, $Z_p^{m,n}$ is tractable for all integer n and p, but m has to stay non-negative.

[69]We suppress the constant of integration here and everywhere else.

C.3.2 Reduction of $Z_p^{m,n}$ to $Z_0^{m,n}$ and $Z_1^{m,n}$

We can integrate and restructure eq. (II.3.16) into the recurrence relation

$$Z_p^{m,n} = \frac{m}{p-1} Z_{p-1}^{m-1,n+1} - \frac{n}{p-1} Z_{p-2}^{m,n+1} - \frac{\zeta_{p-1}^{m,n}}{p-1}, \qquad p \neq 1. \tag{II.3.21}$$

This allows any $p > 0$ to be reduced to the cases $p = 0$ and $p = 1$. When p is negative, (II.3.17) furnishes the simpler relation

$$Z_p^{m,n} = Z_{p+2}^{m,n} + 2Z_{p+2}^{m,n-1}. \tag{II.3.22}$$

This reduces any $p < 0$ to $p = 0$ and $p = 1$, as long as n stays positive. In the $n = 0$ case, we again turn to eq. (II.3.16):

$$Z_p^{m,0} = (p+1)Z_{p+2}^{m,-1} - mZ_{p+1}^{m-1,0} - \zeta_{p+1}^{m,-1}. \tag{II.3.23}$$

Using eq. (II.3.17) to treat the $n = -1$ terms, we arrive at

$$Z_p^{m,0} = \frac{p+1}{p-1} Z_{p+2}^{m,0} + \frac{2m}{p-1} Z_{p+1}^{m-1,0} + \frac{\zeta_{p+1}^{m,0} - \zeta_{p-1}^{m,0}}{p-1}, \tag{II.3.24}$$

which works for $n = 0, p \neq 1$.[70]

In the case $p = 1$, $n > 0$ we integrate by parts:

$$Z_1^{m,n} = \frac{Z^{m,n}}{\beta} + \int \frac{Z^{m,n}}{\beta^2} \, d\beta. \tag{II.3.25}$$

Through eqs. (II.3.19) and (II.3.21), the integral will reduce to $p = 0$ terms that are easy to handle, plus various $Z_1^{m',n'}$ where $m' \leq m, n' \leq n$. At least one of the inequalities is strict, so we will eventually arrive at $Z_1^{0,n''}$ and $Z_1^{m'',0}$. The former can be run through the recurrence again, but the latter requires separate consideration.

C.3.3 Special cases for $n = 0$

The relations above are capable of reducing almost all Z's to ζ's, but they are unable to get rid of $Z_p^{m,0}$ for $p = 0, 1, 2$. If we could define a function $\Phi_\ell^{(p)}$ with derivative $-H'\Phi_{\ell-1}^{(p)}$ and base case $\frac{d}{d\beta}\Phi_0^{(p)} = \beta^{-p}$, then an elegant solution to this would be

$$Z_p^{m,0} = \sum_{\ell=0}^{m} \frac{m!}{(m-\ell)!} \Phi_\ell^{(p)} H^{m-\ell}, \qquad m, p \geq 0. \tag{II.3.26}$$

Such a function can be constructed using polylogarithms, since

$$\frac{d}{d\beta} \text{Li}_\ell[f(\beta)] = \text{Li}_{\ell-1}[f(\beta)] \frac{f'(\beta)}{f(\beta)}. \tag{II.3.27}$$

---

[70]It could also be adapted to other $n \neq \frac{1-p}{2}$, but in those cases eq. (II.3.22) is simpler.

The correct recurrence relation is obtained by solving a simple differential relation for $f(\beta)$, which gives

$$f(\beta) = K\frac{1+\beta}{1-\beta}, \qquad \mathrm{Li}_0[f(\beta)] = \frac{K(1+\beta)}{(1-K)-\beta(1+K)}. \tag{II.3.28}$$

We immediately see that $K = -1$ is suitable for $p = 0$, and $K = 1$ for $p = 2$:[71,72]

$$\Phi_\ell^{(0)} = -2\,\mathrm{Li}_\ell\left(\frac{\beta+1}{\beta-1}\right), \qquad \Phi_\ell^{(2)} = 2\,\mathrm{Li}_\ell\left(\frac{1+\beta}{1-\beta}\right), \tag{II.3.29}$$

We cannot handle $p = 1$ directly this way, but using the close connection between $\mathrm{Li}_1(x)$ and the natural logarithm, we find that

$$\Phi_\ell^{(1)} = -\frac{\Phi_{\ell+1}^{(0)} + \Phi_{\ell+1}^{(2)}}{2} \tag{II.3.30}$$

gives the correct result.[73]

C.3.4 The treatment of $(z - v)$

The factors that arise due to the dispersion relations are highly problematic when $v \neq 0$ (which corresponds to $s \neq 0$ or $u \neq 0$ in eq. (II.5.2)). We only need to consider $k = 0$, since

$$\int \frac{\zeta_p^{m,n}}{(z-v)^{k+1}}\mathrm{d}z = \frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}v^k}\int\frac{\zeta_p^{m,n}}{z-v}\mathrm{d}z; \tag{II.3.31}$$

the derivative can be taken after evaluating the integral. Trying to get rid of the last power of $(z-v)$ is futile, so we instead change variables to β and find

$$\int\frac{\zeta_p^{m,n}}{z-v}\mathrm{d}z = 2\int\frac{4}{zv}\frac{\zeta_{p-1}^{m,n+2}}{\frac{4}{v}-\frac{4}{z}}\mathrm{d}\beta = \frac{-4}{v}\int\frac{\zeta_{p-1}^{m,n+1}}{\beta^2-\beta_v^2}\mathrm{d}\beta, \tag{II.3.32}$$

where $\beta_v \equiv \beta(v) = \sqrt{1-4/v}$.[74] This relation is singular when $v = 0$, but then $\frac{z^n}{(z-v)^{k+1}} = z^{n-(k+1)}$ so it is not needed.[75]

---

[71]This process can in principle be continued to treat all $p \geq 2$, but it is more practical to rely on the recurrence relations to get rid of larger p.

[72]Note that $\Phi_\ell^{(2)}$ is complex-valued for $\ell > 0$, since $\frac{1+\beta}{1-\beta} > 1$ for z above threshold.

[73]Although $\frac{\mathrm{d}}{\mathrm{d}\beta}\Phi_0^{(1)} = 1/\beta$, we have $\Phi_0^{(1)} = \ln(-\beta)$, not $\ln\beta$, so one must be careful to place β on the correct side of the branch cut.

[74]The same relation is useful if one wishes to explicitly evaluate a partial-wave expansion like eq. (II.5.3), since the Legendre polynomials consist of powers of $1/(z-4)$. It is further simplified by $\beta(4) = 0$.

[75]Note that β_v is imaginary for $v \in (0,4)$, which forces a more involved detour through the complex plane than above. For instance, the functions $\Psi_\ell^\pm$ below will be complex-valued, although in the end the integral will of course remain real.

It turns out that $\int H^m/(\beta^2 - \beta_v^2)\,\mathrm{d}\beta$ is tractable, so the strategy is to separate that from the rest of the integral by repeatedly applying partial fractions:

$$\frac{\zeta_p^{m,n}}{\beta^2 - \beta_v^2} = \frac{v}{4}\left[\zeta_p^{m,n} - \frac{2\zeta_p^{m,n-1}}{\beta^2 - \beta_v^2}\right]. \tag{II.3.33}$$

If n is negative, we instead use eq. (II.3.17). Once n has been reduced to zero in the $(\beta^2 - \beta_v^2)$-containing term this way, we remove p similarly:

$$\frac{\zeta_p^{m,0}}{\beta^2 - \beta_v^2} = \zeta_{p+2}^{m,0}\left[\frac{\beta_v^2}{\beta^2 - \beta_v^2} + 1\right] = \frac{1}{\beta_v^2}\left[\frac{\zeta_{p-2}^{m,0}}{\beta^2 - \beta_v^2} - \zeta_p^{m,0}\right]. \tag{II.3.34}$$

This reduces p to 0 or -1; the latter can be handled with the expansion

$$\frac{\beta}{\beta^2 - \beta_v^2} = \frac{1 - \beta_v}{(1 - \beta)(\beta - \beta_v)} - \frac{\beta_v}{\beta^2 - \beta_0^2} + \tfrac{1}{2}(1 + \beta)H'. \tag{II.3.35}$$

The second term on the right-hand side corresponds to $p = 0$, and the third does not involve β_v at all. In the spirit of eq. (II.3.26), the first is equal to $\frac{\mathrm{d}}{\mathrm{d}\beta}\Psi_0^-$ if we identify

$$\Psi_\ell^\pm = -\mathrm{Li}_{\ell+1}\left[\frac{(1 + \beta)(1 \pm \beta_v)}{(1 - \beta)(1 \mp \beta_v)}\right] \tag{II.3.36}$$

guided by eq. (II.3.28). Since $\frac{\mathrm{d}}{\mathrm{d}\beta}\Psi_\ell^\pm = -H'\Psi_{\ell-1}^\pm$,

$$\int\frac{\zeta_{-1}^{m,0}}{\beta^2 - \beta_v^2}\mathrm{d}\beta = \frac{Z^{m,1} + Z_{-1}^{m,1}}{2} - \int\frac{\beta_v\zeta^{m,0}}{\beta^2 - \beta_v^2}\mathrm{d}\beta + \sum_{\ell=0}^m\frac{m!}{(m-\ell)!}\Psi_\ell^- H^{m-\ell}. \tag{II.3.37}$$

This leaves $p = 0$, for which a similar solution is

$$\int\frac{\beta_v\zeta^{m,0}}{\beta^2 - \beta_v^2}\mathrm{d}\beta = \sum_{\ell=0}^m\frac{m!}{(m-\ell)!}\frac{\Psi_\ell^- - \Psi_\ell^+}{2}H^{m-\ell}. \tag{II.3.38}$$

The very last piece in the puzzle of evaluating $D_k^J(\lambda, v, t)$ is the conceptually simple derivative in eq. (II.3.31):

$$\frac{\mathrm{d}\beta_v}{\mathrm{d}v} = \frac{2}{\beta_v v^2}, \qquad \frac{\mathrm{d}}{\mathrm{d}v}\Psi_\ell^\pm = \pm\frac{v}{\beta_v}\Psi_{\ell-1}^\pm, \qquad \frac{\mathrm{d}\beta}{\mathrm{d}v} = \frac{\mathrm{d}}{\mathrm{d}v}Z_p^{m,n} = 0. \tag{II.3.39}$$

C.3.5 The completed integral

The above recurrence relations allow for the integration of all terms that appear in $D_k^J(\lambda, v, t)$. $Z_p^{m,n}$ diverges at $z \to 4$ for some values of p, m, n (in particular those that contain $\Phi_1^{(2)}$ or negative powers of β), so obtaining a finite lower limit of the overall integral requires careful (albeit straightforward) extraction and cancellation of those divergences. The resulting expressions are very lengthy in most cases, so we do not reproduce them here.[76]

---

[76]The FORM implementation of the relations, and the expressions produced by it, are available from Mattias Sjö upon request.

References

[1]   Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory to One Loop". *Annals Phys.* **158,** 142 (1984).

[2]   Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark". *Nucl. Phys. B* **250,** 465–516 (1985).

[3]   Weinberg, S. "Phenomenological Lagrangians". *Physica A* **96,** 327–340 (1979).

[4]   Scherer, S. & Schindler, M. R. *A Primer for Chiral Perturbation Theory* 1st ed. ISBN: 978-3-642-19253-1 (Springer-Verlag, 2012).

[5]   Pich, A. *Effective Field Theory with Nambu-Goldstone Modes* in *Les Houches summer school: EFT in Particle Physics and Cosmology* (2018). arXiv: 1804.05664 [hep-ph].

[6]   Bijnens, J. & Ecker, G. "Mesonic low-energy constants". *Ann. Rev. Nucl. Part. Sci.* **64,** 149–174. arXiv: 1405.6488 [hep-ph] (2014).

[7]   Martin, A. *Scattering Theory: Unitarity, Analyticity and Crossing* 1st ed. (Springer-Verlag, 1969).

[8]   Pham, T. N. & Truong, T. N. "Evaluation of the Derivative Quartic Terms of the Meson Chiral Lagrangian From Forward Dispersion Relation". *Phys. Rev. D* **31,** 3027 (1985).

[9]   Ananthanarayan, B., Toublan, D. & Wanders, G. "Consistency of the chiral pion pion scattering amplitudes with axiomatic constraints". *Phys. Rev. D* **51,** 1093–1100. arXiv: hep-ph/9410302 (1995).

[10] Pennington, M. R. & Portoles, J. "The Chiral Lagrangian parameters, l_1, l_2, are determined by the rho resonance". *Phys. Lett. B* **344,** 399–406. arXiv: hep-ph/9409426 (1995).

[11] Diţă, P. "Positivity constraints on chiral perturbation theory pion pion scattering amplitudes". *Phys. Rev. D* **59,** 094007. arXiv: hep-ph/9809568 (1999).

[12] Distler, J., Grinstein, B., Porto, R. A. & Rothstein, I. Z. "Falsifying Models of New Physics via WW Scattering". *Phys. Rev. Lett.* **98,** 041601. arXiv: hep-ph/0604255 (2007).

[13] Manohar, A. V. & Mateu, V. "Dispersion Relation Bounds for $\pi\pi$ Scattering". *Phys. Rev. D* **77,** 094019. arXiv: 0801.3222 [hep-ph] (2008).

[14] Mateu, V. "Universal Bounds for SU(3) Low Energy Constants". *Phys. Rev. D* **77,** 094020. arXiv: 0801.3627 [hep-ph] (2008).

[15] Bellazzini, B., Elias Miró, J., Rattazzi, R., Riembau, M. & Riva, F. "Positive moments for scattering amplitudes". *Phys. Rev. D* **104,** 036006. arXiv: 2011.00037 [hep-th] (2021).

[16] Caron-Huot, S. & Van Duong, V. "Extremal Effective Field Theories". *JHEP* **05,** 280. arXiv: 2011.02957 [hep-th] (2021).

[17] Sinha, A. & Zahed, A. "Crossing Symmetric Dispersion Relations in Quantum Field Theories". *Phys. Rev. Lett.* **126,** 181601. arXiv: 2012.04877 [hep-th] (2021).

[18] Zahed, A. "Positivity and geometric function theory constraints on pion scattering". *JHEP* **12,** 036. arXiv: 2108.10355 [hep-th] (2021).

[19] Wang, Y.-J., Guo, F.-K., Zhang, C. & Zhou, S.-Y. "Generalized positivity bounds on chiral perturbation theory". *JHEP* **07,** 214. arXiv: 2004.03992 [hep-ph] (2020).

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

[21] Alvarez, B. *Updating bounds on the low-energy constants of Chiral Perturbation Theory from exact bounds on amplitudes* M.Sc. thesis (Lund University, 2016).

[22] Catà, O. & Mateu, V. "Chiral perturbation theory with tensor sources". *JHEP* **09,** 078. arXiv: 0705.2948 [hep-ph] (2007).

[23] Donoghue, J. F. & Leutwyler, H. "Energy and momentum in chiral theories". *Z. Phys. C* **52,** 343–351 (1991).

[24] Coleman, S. R., Wess, J. & Zumino, B. "Structure of phenomenological Lagrangians. 1." *Phys. Rev.* **177,** 2239–2247 (1969).

[25] Callan Curtis G., J., Coleman, S. R., Wess, J. & Zumino, B. "Structure of phenomenological Lagrangians. 2." *Phys. Rev.* **177,** 2247–2250 (1969).

[26] Ecker, G., Gasser, J., Pich, A. & de Rafael, E. "The Role of Resonances in Chiral Perturbation Theory". *Nucl. Phys. B* **321,** 311–342 (1989).

[27] Bijnens, J., Colangelo, G. & Ecker, G. "The Mesonic chiral Lagrangian of order p^6". *JHEP* **02,** 020. arXiv: hep-ph/9902437 (1999).

[28] Bijnens, J., Hermansson-Truedsson, N. & Wang, S. "The order p^8 mesonic chiral Lagrangian". *JHEP* **01,** 102. arXiv: 1810.06834 [hep-ph] (2019).

[29] Bijnens, J., Colangelo, G. & Ecker, G. "Renormalization of chiral perturbation theory to order p^6". *Annals Phys.* **280,** 100–139. arXiv: hep-ph/9907333 (2000).

[30] Bijnens, J., Colangelo, G., Ecker, G., Gasser, J. & Sainio, M. E. "Pion-pion scattering at low energy". *Nucl. Phys. B* **508.** [Erratum: Nucl.Phys.B **517,** 639–639 (1998)], 263–310. arXiv: hep-ph/9707291 (1997).

[31] Bijnens, J., Colangelo, G., Ecker, G., Gasser, J. & Sainio, M. E. "Elastic $\pi\pi$ scattering to two loops". *Phys. Lett. B* **374,** 210–216. arXiv: hep-ph/9511397 (1996).

[32] Bijnens, J., Dhonte, P. & Talavera, P. "$\pi\pi$ scattering in three flavor ChPT". *JHEP* **01,** 050. arXiv: hep-ph/0401039 (2004).

[33] Bijnens, J., Dhonte, P. & Talavera, P. "πK scattering in three flavor ChPT". *JHEP* **05,** 036. arXiv: hep-ph/0404150 (2004).

[34] Bijnens, J. & Lu, J. "Meson-meson Scattering in QCD-like Theories". *JHEP* **03,** 028. arXiv: 1102.0172 [hep-ph] (2011).

[35] Chivukula, R. S., Dugan, M. J. & Golden, M. "Analyticity, crossing symmetry and the limits of chiral perturbation theory". *Phys. Rev. D* **47,** 2930–2939. arXiv: hep-ph/9206222 (1993).

[36] Weinberg, S. "Pion scattering lengths". *Phys. Rev. Lett.* **17,** 616–621 (1966).

[37] Gasser, J. & Sainio, M. E. "Two loop integrals in chiral perturbation theory". *Eur. Phys. J. C* **6,** 297–306. arXiv: hep-ph/9803251 (1999).

[38] Neville, D. E. "Elastic Scattering of Pseudoscalar Mesons and $SU(N)$ Symmetry". *Phys. Rev.* **132**, 844–851 (1963).

[39] Preparata, F. P. & Muller, D. E. "Finding the intersection of n half-spaces in time $\mathcal{O}(n \log n)$". *Theoretical Computer Science* **8**, 45–55 (1979).

[40] Froissart, M. "Asymptotic behavior and subtractions in the Mandelstam representation". *Phys. Rev.* **123**, 1053–1057 (1961).

[41] Manohar, A. V. "Effective field theories". *Lect. Notes Phys.* **479** (eds Latal, H. & Schweiger, W.) 311–362. arXiv: hep-ph/9606222 (1997).

[42] Gomez Nicola, A. & Pelaez, J. R. "Meson meson scattering within one loop chiral perturbation theory and its unitarization". *Phys. Rev. D* **65**, 054009. arXiv: hep-ph/0109056 (2002).

[43] Cacciapaglia, G., Pica, C. & Sannino, F. "Fundamental Composite Dynamics: A Review". *Phys. Rept.* **877**, 1–70. arXiv: 2002.04914 [hep-ph] (2020).

[44] Farhi, E. & Susskind, L. "A Technicolored G.U.T." *Phys. Rev. D* **20**, 3404–3411 (1979).

[45] Bijnens, J. & Husek, T. "Six-pion amplitude". *Phys. Rev. D* **104**, 054046. arXiv: 2107.06291 [hep-ph] (2021).

[46] Tol, P. *Colour Schemes* 2018. https://personal.sron.nl/~pault/.

[47] Brewer, C. A. *ColorBrewer, a web tool for selecting colors for maps* 2009. http://colorbrewer2.org.

[48] Barber, C. B., Dobkin, D. P. & Huhdanpaa, H. "The quickhull algorithm for convex hulls". *ACM Transactions on Mathematical Software (TOMS)* **22**, 469–483 (1996).

Paper III

III

Six-meson amplitude in QCD-like theories

Johan Bijnens,[1] Tomáš Husek [1] and Mattias Sjö[1]

[1]Department of Astronomy and Theoretical Physics, Lund University,
Box 43, SE 22100 Lund, Sweden

Abstract: We calculate the relativistic six-meson scattering amplitude at low energy within the framework of QCD-like theories with n degenerate quark flavors at next-to-leading order in the chiral counting. We discuss the cases of complex, real and pseudo-real representations, i.e. with global symmetry and breaking patterns $\mathrm{SU}(n) \times \mathrm{SU}(n)/\mathrm{SU}(n)$ (extending the QCD case), $\mathrm{SU}(2n)/\mathrm{SO}(2n)$, and $\mathrm{SU}(2n)/\mathrm{Sp}(2n)$. In case of the one-particle-irreducible part, we obtain analytical expressions in terms of 10 six-meson sub-amplitudes based on the flavor and group structures. We extend on our previous results obtained within the framework of $\mathrm{O}(N + 1)/\mathrm{O}(N)$ non-linear sigma model, with N being the number of meson flavors. This work allows for studying a number of properties of six-particle amplitudes at one-loop level.

1 Introduction

Quantum chromodynamics (QCD), the fundamental theory of the strong interaction, becomes non-perturbative at low energy and it is therefore impractical for phenomenology in that regime. From the large-distance perspective, the fundamental quark and gluon degrees of freedom are effectively replaced by composite colorless states, the lightest of which are the mesons. These can be approximately interpreted as the Nambu–Goldstone bosons of the associated spontaneous breaking of the chiral symmetry of massless QCD. With appropriate explicit symmetry breaking added to account for quark masses and non-strong interactions, the resulting effective field theory (EFT) is known as chiral perturbation theory (ChPT) [1–3] and is commonly used with great success for low-energy hadron phenomenology. See Refs. [4, 5] for modern introductions to ChPT.

There has been recent interest in the $3 \to 3$ meson scattering amplitude driven by advances in lattice QCD [6–18]. While many ChPT observables are known to high loop level, the six-meson amplitude was only recently calculated to one-loop level [19], and then only for two quark flavors, i.e. a meson spectrum of only pions. The case of three or more flavors is largely unexplored; the tree-level part is known up to next-to-next-to-next-to-leading order (N³LO) in the massless case [20]. The leading-order (LO) massless pion case was initially done with current algebra methods and predates ChPT [21, 22].

While QCD is the canonical example, strongly coupled gauge theories can have different patterns of spontaneous symmetry breaking. These were first discussed in the context of technicolor theories [23–25]. When the gauge group is vector-like and all fermions have the same mass, only three patterns show up as discussed in ref. [26]; earlier work can be traced from there. If all n fermions are in a complex representation, the global symmetry group is $\mathrm{SU}(n) \times \mathrm{SU}(n)$ and is broken spontaneously to the diagonal (vector) $\mathrm{SU}(n)$, which corresponds to the n-quark QCD case. If the fermions are in a real or pseudo-real representation, the global symmetry group is $\mathrm{SU}(2n)$ and is spontaneously broken to $\mathrm{SO}(2n)$ and $\mathrm{Sp}(2n)$, respectively. We will refer to these cases as SU, SO and Sp, respectively, and collectively dub them 'QCD-like theories'. ChPT has been extended to these, and results can be found e.g. in refs. [27–29]. The similarity between all cases, and a number of calculations to two-loop order (vacuum expectation value, mass and decay constant, and four-meson amplitudes), were worked out in refs. [30–32].

In the context of studying general properties of amplitudes, much attention has been paid to the structure of (massive) nonlinear sigma models (corresponding to ChPT without additional fields) at tree level including higher orders using various techniques [20, 33–36]. However, not all of these properties generalize to loop level. Some loop-level progress can be found in ref. [37–39]. In this work, we calculate the six-meson amplitude to next-to-leading order (NLO) for the three symmetry-breaking patterns. This generalizes the earlier work of ref. [19] for the symmetry-breaking pattern $\mathrm{O}(N + 1)/\mathrm{O}(N)$.

In section 2, Chiral Perturbation Theory for QCD-like theories is shortly discussed; a more extensive discussion can be found in ref. [30]. We introduce here also a notation that explicitly brings out the similarities for the three cases. We do not describe the calculation in great detail; it follows the standard Feynman diagram method and does sums over flavor

indices using eqs. (III.2.7) and (III.2.8). The flavor structure of the general four- and six-meson amplitudes is discussed in Sec. III–A. The expressions can be very much simplified by using all symmetry properties of the amplitudes, as expected from general considerations. This is described later in section 3. We have checked that our results are UV finite and independent of the parametrization of the Nambu–Goldstone boson manifold, and that they reduce to the results of refs. [19, 31] in the appropriate cases. At the end of section 3, we also present a limit of the six-meson amplitude in which the three-momenta of the particles of modulus p go symmetrically to zero. In this particular kinematic setting, we plot the flavor-stripped amplitudes with respect to p and show the results in section 4. Our conclusions are shortly discussed in section 5, followed by several technical appendices that fix the notation and explain further subtleties and broader context. Explicit expressions for our main result — the NLO six-meson amplitude — in terms of deorbited group-universal subamplitudes can be found in appendix D.

The analytical work in this manuscript was done both using Wolfram *Mathematica* with the FeynCalc package [40–42] and a FORM [43] implementation. The numerical results use LoopTools [44, 45].

2 Theoretical setting

2.1 Lagrangian

We consider a theory of n fermions with some symmetry group G, which is spontaneously broken to a subgroup H. This gives rise to an EFT whose degrees of freedom are pseudo-Nambu–Goldstone bosons transforming under the quotient group G/H. In analogy with the QCD case, we will refer to these as 'mesons'.

We choose G/H from the patterns of symmetry breaking present in the QCD-like theories described in the introduction. The mesons are parametrized through a flavor-space matrix field u, also called the Nambu–Goldstone boson matrix. In addition, the Lagrangian can be extended in terms of vector, axial-vector, scalar and pseudoscalar external fields [2, 3]. These correspond to vector and scalar sources for conserved and broken generators in general. The symmetry may be explicitly broken by introducing quark masses in the scalar external field. Except for the definition of the decay constant and the introduction of quark masses, we do not need external fields in this work.

The Lagrangian for the meson–meson scattering at NLO relevant for all the discussed theories can be written as

$$\mathcal{L} = \mathcal{L}^{(2)} + \mathcal{L}^{(4)} \,, \tag{III.2.1}$$

separating the LO and NLO terms in the chiral counting.[1,2] These take the form

$$\mathcal{L}^{(2)} = \frac{F^2}{4} \langle u_\mu u^\mu + \chi_+ \rangle, \tag{III.2.2}$$

$$\mathcal{L}^{(4)} = L_0 \langle u_\mu u_\nu u^\mu u^\nu \rangle + L_1 \langle u_\mu u^\mu \rangle \langle u_\nu u^\nu \rangle + L_2 \langle u_\mu u_\nu \rangle \langle u^\mu u^\nu \rangle + L_3 \langle u_\mu u^\mu u_\nu u^\nu \rangle$$
$$+ L_4 \langle u_\mu u^\mu \rangle \langle \chi_+ \rangle + L_5 \langle u_\mu u^\mu \chi_+ \rangle + L_6 \langle \chi_+ \rangle^2 + L_7 \langle \chi_- \rangle^2 + \tfrac{1}{2} L_8 \langle \chi_+^2 + \chi_-^2 \rangle. \tag{III.2.3}$$

Above, $\langle \cdots \rangle$ denotes a flavor-space trace over $n \times n$ matrices for the SU and $2n \times 2n$ matrices for the SO and Sp cases. Moreover,

$$u_\mu \equiv i\big(u^\dagger \partial_\mu u - u \partial_\mu u^\dagger\big), \tag{III.2.4}$$

$$\chi_\pm \equiv u^\dagger \chi u^\dagger \pm u \chi^\dagger u. \tag{III.2.5}$$

Under G, both u_μ and $\chi_\pm$ transform as $X \to hXh^\dagger$, where $h \in H$. Above, as usual, $\chi \equiv 2B_0 \mathcal{M}$, with $\mathcal{M} = s - ip$, where $s(p)$ are the (pseudo)scalar external fields and B_0 is a parameter related to the scalar singlet quark condensate $\langle 0|\bar{q}q|0\rangle$ (not to be confused with the integral B_0 in appendix A). For our application and in the case with all the mesons having the same (lowest-order) mass M, we can simply put $\chi = M^2 \mathbb{1}$.

The Nambu–Goldstone boson matrix u can be parametrized as

$$u = \exp\left(\frac{i}{\sqrt{2}F} \phi^a t^a\right), \tag{III.2.6}$$

where ϕ^a denote the pseudoscalar meson fields and t^a are Hermitian generators of G/H normalized to $\langle t^a t^b \rangle = \delta^{ab}$. Besides the 'exponential' parametrization (III.2.6), there are other options available in the literature. For practical calculations, it is useful to employ several different parametrizations in parallel. This serves as a neat cross-check since, as anticipated, the final amplitude should be parametrization-independent. We discuss the most general reparametrization in appendix B. In the case of the six-meson amplitude at NLO, 18 free parameters appear in the expansion of u in terms of $\phi^a t^a$. We have checked that all our physical results are independent of these parameters.

2.2 Flavor structures

Each meson ϕ^a carries a *flavor index* a, which appears in the amplitude carried by a G/H generator residing in a flavor-space trace. When a pair of fields is Wick contracted, the corresponding flavor indices are summed over; under SU, the resulting expressions are evaluated using the Fierz identities

$$\langle t^a A \rangle \langle t^a B \rangle = \langle AB \rangle - \tfrac{1}{n} \langle A \rangle \langle B \rangle, \tag{III.2.7a}$$

---

[1]The next-to-next-to-leading-order (NNLO) terms $\mathcal{L}^{(6)}$ [46] and N^{3}LO terms $\mathcal{L}^{(8)}$ [47] are also known but not used here.

[2]Recall that the chiral counting order of an ℓ-loop diagram with n_k vertices from $\mathcal{L}^{(k)}$ is $m = 2 + 2\ell + \sum_k n_k(k-2)$. Thus, NLO ($m = 4$) diagrams have either one loop or one vertex from $\mathcal{L}^{(4)}$.

$$\langle t^a A t^a B \rangle = \langle A \rangle \langle B \rangle - \tfrac{1}{n} \langle AB \rangle \,, \tag{III.2.7b}$$

where A and B are arbitrary flavor-space matrices. The analogous identities for SO and Sp are quite similar, so we will use the abbreviation $S_{\mathrm{P}}^{\mathrm{O}}$ in correlation with $\pm$: SO is paired with $+$, and Sp with $-$. Thus, for $S_{\mathrm{P}}^{\mathrm{O}}$ one uses

$$\langle t^a A \rangle \langle t^a B \rangle = \tfrac{1}{2} \big[\, \langle AB \rangle \; + \langle AB^\dagger \rangle \big] - \tfrac{1}{2n} \langle A \rangle \langle B \rangle, \tag{III.2.8a}$$

$$\langle t^a A t^a B \rangle = \tfrac{1}{2} \big[\langle A \rangle \langle B \rangle \pm \langle AB^\dagger \rangle \big] - \tfrac{1}{2n} \langle AB \rangle \,. \tag{III.2.8b}$$

Here, B must be a string of generators t^a or the unit matrix, so $\dagger$ effectively denotes reversal.[3]

Note that the implicitly summed index a has different dimensions in eqs. (III.2.7) and (III.2.8), corresponding to the number of mesons: $n^2 - 1$ under SU and $2n^2 \pm n - 1$ under $S_{\mathrm{P}}^{\mathrm{O}}$. Note also, that due to the formally identical Lagrangians, eqs. (III.2.7) and (III.2.8) are the only source of formal dissimilarity between the amplitudes for the different cases.

2.3   Low-energy constants and renormalization

At LO, we have two low-energy parameters: the mass M (related to the aforementioned B_0) and decay constant F. At NLO, 9 more constants (LECs) L_i accompanying additional allowed chirally symmetric structures (operators) relevant for our application appear, as shown in eq. (III.2.3). These constants contain UV-divergent parts represented by coefficients Γ_i, which are uniquely fixed from the requirement that physical NLO amplitudes should be finite, and UV-finite parts $L_i^{\mathrm{r}} \equiv L_i^{\mathrm{r}}(\mu)$, renormalized at a scale μ, that are free parameters in the theory:

$$L_i = (c\mu)^{d-4} \left(\frac{1}{16\pi^2} \frac{1}{d-4} \Gamma_i + L_i^{\mathrm{r}}(\mu) \right). \tag{III.2.9}$$

Above, d is the space-time dimension in the vicinity of 4 and c is such that

$$\log c = -\frac{1}{2} \left(1 - \gamma_{\mathrm{E}} + \log 4\pi \right). \tag{III.2.10}$$

Consequently, in terms of $\epsilon = 2 - d/2$ and

$$\frac{1}{\bar{\epsilon}} \equiv \frac{1}{\epsilon} - \gamma_{\mathrm{E}} + \log 4\pi - \log \mu^2 + 1 \,, \tag{III.2.11}$$

one writes (to NLO)

$$L_i = -\kappa \frac{\Gamma_i}{2} \frac{1}{\bar{\epsilon}} + L_i^{\mathrm{r}} \,, \tag{III.2.12}$$

with $\kappa \equiv \frac{1}{16\pi^2}$. The extra '+1' term in eq. (III.2.11) with respect to the standard $\overline{\mathrm{MS}}$ renormalization scheme is customary in ChPT.

---

[3]The general version is given in ref. [31]. The version here follows since the generators are Hermitian: $(t^a t^b \cdots t^c)^\dagger = t^c \cdots t^b t^a$.

Studying and renormalizing the four-meson amplitude at NLO (i.e. considering one-loop diagrams with vertices from $\mathcal{L}^{(2)}$ and tree-level counterterms from $\mathcal{L}^{(4)}$) determines all the Γ_i except for one: The divergent part of L_7 remains unset. It can, however, be fixed from the six-meson amplitude. Using the heat-kernel technique, all NLO divergences were derived in ref. [30]. For the reader's convenience, we list the Γ_i here in a group-universal form:

$$
\begin{aligned}
\Gamma_0 &= \tfrac{1}{48}(n + 4\xi), & \Gamma_5 &= \tfrac{n}{8}, \\
\Gamma_1 &= \tfrac{1}{16\zeta}, & \Gamma_6 &= \tfrac{1}{16\zeta} + \tfrac{1}{8\zeta^2 n^2}, \\
\Gamma_2 = \Gamma_4 &= \tfrac{1}{8\zeta}, & \Gamma_7 &= 0, \\
\Gamma_3 &= \tfrac{1}{24}(n - 2\xi), & \Gamma_8 &= \tfrac{1}{16}\left(n + \xi - \tfrac{4}{\zeta n}\right).
\end{aligned}
\tag{III.2.13}
$$

Above, ξ and $\zeta \equiv (1 + \xi^2)$ parametrize the groups as follows:

$$
\xi \equiv \begin{cases} 0 & [\,\mathrm{SU}\,], \\ \pm 1 & [\,S_{\mathrm{p}}^{\mathrm{o}}\,], \end{cases} \qquad \zeta = \begin{cases} 1 & [\,\mathrm{SU}\,], \\ 2 & [\,S_{\mathrm{p}}^{\mathrm{o}}\,]. \end{cases}
\tag{III.2.14}
$$

Another check on our calculation is that, with the expressions in eq. (III.2.13), all our results are finite.

2.4 Mass and decay constant

The Z factor used for the wave-function renormalization is related to the meson self-energy Σ as

$$
\frac{1}{Z} = 1 - \left.\frac{\partial \Sigma(p^2)}{\partial p^2}\right|_{p^2 = M_\pi^2},
\tag{III.2.15}
$$

with $-i\Sigma$ being represented by a tadpole graph with two external legs plus counterterms stemming from the Lagrangian (III.2.3). Note that in our application the physical mass of all mesons is equal and denoted as M_π. At NLO, the LO vertex and propagator are extended in terms of the replacements

$$
\begin{aligned}
M^k &\to M_\pi^k \left(1 - \frac{k}{2}\frac{\overline{\Sigma}}{M_\pi^2}\right), \\
\frac{1}{F^k} &\to \frac{1}{F_\pi^k}(1 + k\delta F),
\end{aligned}
\tag{III.2.16}
$$

at the given order equivalent to the standard $M_\pi^2 = M^2 + \overline{\Sigma}$, $F_\pi = F(1 + \delta F)$, with

$$
\begin{aligned}
\overline{\Sigma} &= \frac{M_\pi^4}{F_\pi^2}\left\{-8\left[L_5^{\mathrm{r}} - 2L_8^{\mathrm{r}} + n\zeta(L_4^{\mathrm{r}} - 2L_6^{\mathrm{r}})\right] + \left(\frac{1}{\zeta n} - \frac{\xi}{2}\right)L\right\} + \mathcal{O}\left(\frac{1}{F_\pi^4}\right), \\
\delta F &= \frac{M_\pi^2}{F_\pi^2}\left[4(L_5^{\mathrm{r}} + n\zeta L_4^{\mathrm{r}}) - \frac{n}{2}L\right] + \mathcal{O}\left(\frac{1}{F_\pi^4}\right).
\end{aligned}
\tag{III.2.17}
$$

Here, we again present the group-universal form. Above and later on, we use $L \equiv \kappa \log \frac{M^2}{\mu^2}$. Needless to say, in the final result one only retains the terms relevant at order $\mathcal{O}(p^4)$. Thus, in the rest of the NLO expressions one simply takes $M \to M_\pi$ and $F \to F_\pi$. Note that we recalculated the results of eq. (III.2.17) and that they agree with refs. [27, 29, 30].

3   The amplitudes

In terms of Feynman diagrams, loop integrals, etc., the present calculation proceeds along the same lines as the one performed in ref. [19]. However, the result is considerably more cumbersome, largely because the flavor indices are carried by more structures beyond Kronecker δ's. There is also the matter of treating the SU, SO and Sp variants in parallel without tripling the amount of material to present. We will therefore devote much of this section to simplifying the amplitude expressions.

3.1   Flavor structure of the four- and six-meson amplitudes

The general flavor structure of the SU case for meson–meson scattering is well-known; see e.g. refs. [28, 30]. We consider four incoming mesons with flavor indices $b_1, \ldots, b_4$ and momenta $p_1, \ldots, p_4$. The usual Mandelstam variables are defined as[4]

$$s = (p_1 + p_2)^2 , \ t = (p_1 + p_3)^2 , \ u = (p_2 + p_3)^2 . \tag{III.3.1}$$

The amplitude is then conventionally decomposed as

$$\begin{aligned}
A_{4\pi}(s,t,u) = {}& \left(\langle t^{b_1} t^{b_2} t^{b_3} t^{b_4} \rangle + \langle t^{b_4} t^{b_3} t^{b_2} t^{b_1} \rangle \right) B(s,t,u) \\
& + \left(\langle t^{b_1} t^{b_3} t^{b_4} t^{b_2} \rangle + \langle t^{b_2} t^{b_4} t^{b_3} t^{b_1} \rangle \right) B(t,u,s) \\
& + \left(\langle t^{b_1} t^{b_4} t^{b_2} t^{b_3} \rangle + \langle t^{b_3} t^{b_2} t^{b_4} t^{b_1} \rangle \right) B(u,s,t) \\
& + \delta^{b_1 b_2} \delta^{b_3 b_4} C(s,t,u) + \delta^{b_1 b_3} \delta^{b_2 b_4} C(t,u,s) \\
& + \delta^{b_1 b_4} \delta^{b_2 b_3} C(u,s,t) .
\end{aligned} \tag{III.3.2}$$

The functions satisfy $B(s,t,u) = B(u,t,s)$ and $C(s,t,u) = C(s,u,t)$. This structure follows from requiring invariance under the unbroken group, Bose symmetry and charge conjugation for SU. Under $S_{\mathrm{P}}^{\mathrm{O}}$, $\langle t^a t^b t^c t^d \rangle = \langle t^d t^c t^b t^a \rangle$ without relying on charge conjugation.

A similar decomposition for the six-meson amplitude in terms of six flavor labels and momenta is

$$\begin{aligned}
A_{6\pi}(p_1, \ldots, p_6) = \sum_{\mathcal{S}_6} \Big\{ & \tfrac{1}{12} \left[\langle t^{b_1} \cdots t^{b_6} \rangle + \langle t^{b_6} \cdots t^{b_1} \rangle \right] D(p_1, \ldots, p_6) \\
& + \tfrac{1}{16} \delta^{b_1 b_2} \left[\langle t^{b_3} \cdots t^{b_6} \rangle + \langle t^{b_6} \cdots t^{b_3} \rangle \right] E(p_1, \ldots, p_6) \\
& + \tfrac{1}{36} \left[\langle t^{b_1} t^{b_2} t^{b_3} \rangle \langle t^{b_4} t^{b_5} t^{b_6} \rangle + \langle t^{b_3} t^{b_2} t^{b_1} \rangle \langle t^{b_6} t^{b_5} t^{b_4} \rangle \right] F(p_1, \ldots, p_6) \\
& + \tfrac{1}{48} \delta^{b_1 b_2} \delta^{b_3 b_4} \delta^{b_5 b_6} G(p_1, \ldots, p_6) \Big\} ,
\end{aligned} \tag{III.3.3}$$

where $\mathcal{S}_6$ represents the $6! = 720$ permutations of $\{1, \ldots, 6\}$ and the symmetry factors correspond to how many permutations leave the traces and δ's (the 'flavor structure') unchanged; thus, D, E, F and G are summed over 60, 45, 20 and 15 distinct permutations,

---

[4]We have chosen the specific definition here to later define an off-shell extension.

respectively, just like B and C are summed over 3. Charge conjugation and group structure imply the following properties:

$$
\begin{aligned}
D(p_1, \ldots, p_6) &= D(p_2, p_3, p_4, p_5, p_6, p_1) \\
&= D(p_6, p_5, p_4, p_3, p_2, p_1) \,, & \text{(III.3.4a)} \\
E(p_1, \ldots, p_6) &= E(p_2, p_1, p_3, p_4, p_5, p_6) \\
&= E(p_1, p_2, p_4, p_5, p_6, p_3) \\
&= E(p_1, p_2, p_6, p_5, p_4, p_3) \,, & \text{(III.3.4b)} \\
F(p_1, \ldots, p_6) &= F(p_4, p_5, p_6, p_1, p_2, p_3) \\
&= F(p_2, p_3, p_1, p_4, p_5, p_6) \\
&= F(p_1, p_3, p_2, p_4, p_6, p_5) \,, & \text{(III.3.4c)} \\
G(p_1, \ldots, p_6) &= G(p_2, p_1, p_3, p_4, p_5, p_6) \\
&= G(p_3, p_4, p_5, p_6, p_1, p_2) \\
&= G(p_3, p_4, p_1, p_2, p_5, p_6) \,. & \text{(III.3.4d)}
\end{aligned}
$$

These are discussed in a more formal and general way in the next subsection.

3.2 General flavor-based simplification

In order to formalize the structure seen in eqs. (III.3.2) and (III.3.3), we follow the notation of ref. [20] and define a k-particle *flavor structure* as

$$
\mathcal{F}_R(b_1, \ldots, b_k) = \langle t^{b_1} t^{b_2} \cdots t^{b_{r_1}} \rangle \langle t^{b_{r_1+1}} \cdots t^{b_{r_1+r_2}} \rangle \cdots \langle t^{b_{k-r_{|R|}}+1} \cdots t^{b_k} \rangle \,, \qquad \text{(III.3.5)}
$$

where $R = \{r_1, \ldots, r_{|R|}\}$ with $\sum r_i = k$ is a *flavor split*: The flavors are split across $|R|$ traces, each containing r_i indices. Without loss of generality, we may impose $r_1 \leq r_2 \leq \cdots \leq r_{|R|}$. For a permutation σ that maps $i \to \sigma_i$, we define $\mathcal{F}_R^\sigma(b_1, \ldots, b_k) \equiv \mathcal{F}_R(b_{\sigma_1}, \ldots, b_{\sigma_k})$ and denote by $\mathbb{Z}_R$ the group of permutations that preserve $\mathcal{F}_R$: For all $\sigma \in \mathbb{Z}_R$, $\mathcal{F}_R^\sigma(b_1, \ldots, b_k) = \mathcal{F}_R(b_1, \ldots, b_k)$.[5] The group $\mathbb{Z}_R$ is, of course, related to the symmetries in eq. (III.3.4).

In general, an amplitude can be decomposed as

$$
A_{k\pi}(p_1, b_1; p_2, b_2; \ldots; p_k, b_k) = \sum_R \sum_\sigma A_R^\sigma(p_1, \ldots, p_k) \mathcal{F}_R^\sigma(b_1, \ldots, b_k) \,, \qquad \text{(III.3.6)}
$$

where σ is summed over all permutations that do not preserve $\mathcal{F}_R$, i.e. $\mathcal{S}_k / \mathbb{Z}_R$. It follows from Bose symmetry that $A_R^\sigma(p_1, \ldots, p_k) = A_R^{\text{id}}(p_{\sigma_1}, \ldots, p_{\sigma_k})$ where id is the identity permutation. It is therefore sufficient to work with $A_R \equiv A_R^{\text{id}}$, the *stripped amplitude*, for all R; the full amplitude follows from eq. (III.3.6).

The stripped amplitude is easily obtained from the full amplitude by taking the coefficient of $\mathcal{F}_R$. In SU, it is guaranteed to be unique, as was proven in ref. [20]. This carries

---

[5] $\mathbb{Z}_R$ is the cyclic group $\mathbb{Z}_k$ when $R = \{k\}$, hence the notation. In general, it combines cyclic symmetry of each trace with exchanging the contents of same-size traces. It is Abelian as long as all r_i are different.

over to SO and Sp; the ambiguity created by $\langle X \rangle = \langle X^\dagger \rangle$ is easily resolved by applying $\langle X \rangle \to \frac{1}{2}[\langle X \rangle + \langle X^\dagger \rangle]$ before extracting A_R.

In a four-meson amplitude, the stripped amplitudes $A_{\{4\}}$ and $A_{\{2,2\}}$ are the functions called $B(s,t,u)$ and $C(s,t,u)$, respectively, in eq. (III.3.2). For six mesons, one has $A_{\{6\}}$, $A_{\{2,4\}}$, $A_{\{3,3\}}$ and $A_{\{2,2,2\}}$, which correspond to D, E, F and G in eq. (III.3.3), respectively. In the SU($n = 2$) case (equivalent to the O(4)/O(3) case treated in ref. [19]), the Cayley–Hamilton theorem allows all the trace structures to be reduced to $R = \{2,2,2\}$. When $n = 2,3,4,5$ for SU and $n = 1,2$ for $S_{\mathrm{P}}^{\mathrm{O}}$, respectively, the $\mathcal{F}_R$ satisfy a number of linear relations (see ref. [48] for explicit expressions), which in turn relate the A_R to each other. Otherwise, $\mathcal{F}_R$ are linearly independent for different R.[6]

As follows from its definition, A_R inherits $\mathbb{Z}_R$ symmetry (acting on $\{p_1, \ldots, p_k\}$) from $\mathcal{F}_R$. We must also consider another permutation of the external particles, which we dub *trace-reversal* (TR). It is the permutation which reverses the product of generators in each trace: $\langle t^a t^b \cdots t^c \rangle \to \langle t^c \cdots t^b t^a \rangle$. Under SU, this is *not* a symmetry of $\mathcal{F}_R$, but CP invariance nevertheless requires it to be a symmetry of A_R: Charge conjugation maps $t^a \to (t^a)^\mathsf{T}$, and thus $\langle t^a t^b \cdots t^c \rangle \to \langle t^{a\mathsf{T}} t^{b\mathsf{T}} \cdots t^{c\mathsf{T}} \rangle = \langle t^c \cdots t^b t^a \rangle$. This is why eqs. (III.3.2) and (III.3.3) pair each trace with its reverse (except for the reversal-symmetric $\langle t^a t^b \rangle$). We will denote the general symmetry group of A_R, i.e. $\mathbb{Z}_R$ plus TR, by $\mathbb{Z}_R^{+\mathrm{TR}}$.

Under $S_{\mathrm{P}}^{\mathrm{O}}$, $\mathbb{Z}_R^{+\mathrm{TR}}$ is a symmetry also of $\mathcal{F}_R$; in fact, $\langle t^a t^b \cdots t^c \rangle = \langle t^c \cdots t^b t^a \rangle$ makes $\mathcal{F}_R$ symmetric under the reversal of any single trace (CP only requires symmetry under the simultaneous reversal of all traces). This enhanced symmetry is inherited by A_R, and is very important for the relation between the amplitudes of the different QCD-like theories (see appendix E).

The size of the amplitude expressions can be further reduced by writing them in terms of a quantity $\tilde{A}_R$ such that

$$A_R(p_1, \ldots, p_k) = \sum_{\sigma \in \mathbb{Z}_R^{+\mathrm{TR}}} \tilde{A}_R(p_{\sigma_1}, \ldots, p_{\sigma_k}). \tag{III.3.7}$$

This clearly exists (consider e.g. $\tilde{A}_R = A_R/|\mathbb{Z}_R^{+\mathrm{TR}}|$) but is not unique. A method for obtaining a minimal-length $\tilde{A}_R$, the *deorbited* stripped amplitude, is described in appendix C.

3.3 Group-universal formulation

One can expect the amplitudes of the SU, SO and Sp theories to have many similarities, since the only differences relevant to the amplitude are the variations of the Fierz identity, eqs. (III.2.7) and (III.2.8), and the substitution $n \to \zeta n$. In fact, comparison of the amplitudes suggests that one might introduce four subamplitudes $\mathcal{A}^{(i)}$ such that

$$A = \left\{ \mathcal{A}^{(1)} + \xi \mathcal{A}^{(\xi)} + \xi^2 \mathcal{A}^{(\xi^2)} + \frac{\mathcal{A}^{(\zeta)}}{\zeta} \right\}_{n \to \zeta n}, \tag{III.3.8}$$

---

[6]They are in fact orthogonal in a certain sense, as shown in ref. [20].

Figure III.1: The single LO four-meson diagram, with the vertex stemming from $\mathcal{L}^{(2)}$. In formulae we refer to it as $i\mathcal{M}_{\text{LO}}^{(2)}$ or, after NLO mass and decay-constant redefinitions (III.2.17) are applied, $i\mathcal{M}_{\text{LO}}^{(4)}$.

where $(\xi, \zeta) = (0, 1)$ for SU and $(\pm 1, 2)$ for S_{p}^0, as defined in eq. (III.2.14). This decomposition is clearly redundant: Three amplitudes are expressed as a combination of four subamplitudes. However, we find it natural and choose it for its simplicity and clarity; very few terms appear in more than one subamplitude, and $\mathcal{A}^{(\xi^2)}$ is a relatively short expression. The decomposition (III.3.8) can be combined with stripping and deorbiting, allowing the amplitude to be formulated using the very concise quantities $\tilde{\mathcal{A}}_R^{(i)}$. Furthermore, many of these are actually zero. The patterns for which (R, i) combinations are allowed, what LECs, loop integrals and powers of n may appear where, etc., are studied in appendix D and explained in appendix E.

3.4 The four-meson amplitude

The notation of the previous sections allows the four-meson amplitude to be written very compactly. We will use the ordinary Mandelstam variables (III.3.1). At LO, there is a single nonzero subamplitude, stemming from the diagram in fig. III.1,[7]

$$\mathcal{A}_{\{4\}}^{(\text{LO},1)} = 8\tilde{\mathcal{A}}_{\{4\}}^{(\text{LO},1)} = -\frac{t - 2M_\pi^2}{2F_\pi^2} \, . \tag{III.3.9}$$

At NLO, one has one-loop diagrams (two topologies of four one-loop diagrams in total) combined with counterterms, as shown in fig. III.2. Moreover, one needs to take into account NLO wave-function renormalization $(Z^{1/2} - 1)$ applied for every external leg, and mass and decay-constant redefinitions [at the given order based on eq. (III.2.17)] applied to the LO graph $\mathcal{M}_{\text{LO}}^{(2)}$, giving $\mathcal{M}_{\text{LO}}^{(4)}$. Schematically, this can be summed up as

$$A_{4\pi}^{(\text{NLO})} = \mathcal{M}_{\text{1-loop}} + \mathcal{M}_{\text{CT}} + 4(Z^{1/2} - 1)\mathcal{M}_{\text{LO}}^{(2)} + \mathcal{M}_{\text{LO}}^{(4)} \, . \tag{III.3.10}$$

Note that while the above combination is parametrization-independent and UV finite, the separate terms are not. Altogether, the nonzero stripped and deorbited group-universal NLO subamplitudes read

$$F_\pi^4 \tilde{\mathcal{A}}_{\{4\}}^{(\text{NLO},1)} = \frac{M_\pi^4}{4n}\Big\{L + \kappa - \bar{J}(s)\Big\} + 2M_\pi^4(L_0^{\text{r}} + L_8^{\text{r}})$$
$$- \frac{M_\pi^2 t}{2}(4L_0^{\text{r}} + L_5^{\text{r}}) + \frac{t^2}{8}(4L_0^{\text{r}} + L_3^{\text{r}}) + \frac{s(s-u)}{4}L_3^{\text{r}} \, , \tag{III.3.11a}$$

---

[7]We remind the reader that the $R = \{\dots\}$ subscript indicates stripping according to eq. (III.3.6), the tilde indicates deorbiting according to eq. (III.3.7), and the calligraphic $\mathcal{A}$ indicates group-universal formulation according to eq. (III.3.8), with the subamplitude label '1' sharing the superscript with the 'LO' label. All three simplifications can be applied independently and commute with each other.

| (a) 1× | (b) 1× | (c) 3× |
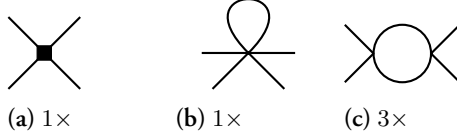
Figure III.2: NLO topologies for the four-meson amplitude. The unmarked vertices stem from $\mathcal{L}^{(2)}$, while the square vertices stem from $\mathcal{L}^{(4)}$ and contain L_i, $i = 1, \ldots, 8$. The numbers (diagram multiplicities) indicate the number of distinct diagrams with the same topology but different permutations of the external legs. Diagram (a) corresponds to $i\mathcal{M}_{\mathrm{CT}}$ and diagrams (b,c) sum up to $i\mathcal{M}_{\mathrm{1\text{-}loop}}$.

$$
F_\pi^4 \tilde{\mathcal{A}}_{\{4\}}^{(\mathrm{NLO},\xi)} = \frac{\bar{J}(t)}{64}\left\{(t - 2M_\pi^2)^2\right\} + \frac{\bar{J}(s)}{192}\left\{(4M_\pi^2 - s)(3s + t - u)\right\}
$$
$$
- \frac{6L + 5\kappa}{1152}\left\{28M_\pi^4 - 16M_\pi^2 t + 3t^2 - 2s(s - u)\right\} - \frac{M_\pi^4 \kappa}{96}, \tag{III.3.11b}
$$
$$
F_\pi^4 \tilde{\mathcal{A}}_{\{4\}}^{(\mathrm{NLO},\zeta)} = \frac{n\bar{J}(s)}{192}\left\{4M_\pi^2(t - u) + s(3s - t + u)\right\}
$$
$$
- \frac{n(3L + 2\kappa)}{1152}\left\{32M_\pi^4 - 20M_\pi^2 t + 3t^2 + 2s(s - u)\right\}
$$
$$
- \frac{n\kappa}{288}\left\{4M_\pi^4 - M_\pi^2 t + s(s - u)\right\}, \tag{III.3.11c}
$$
$$
F_\pi^4 \tilde{\mathcal{A}}_{\{2,2\}}^{(\mathrm{NLO},1)} = \frac{M_\pi^4}{4n^2}\left\{\bar{J}(s) - (L + \kappa)\right\} + \frac{u(u - t)}{2}L_2^{\mathrm{r}} + 4M_\pi^4(L_1^{\mathrm{r}} - L_4^{\mathrm{r}} + L_6^{\mathrm{r}})
$$
$$
+ \frac{s^2}{4}(4L_1^{\mathrm{r}} + L_2^{\mathrm{r}}) - 2M_\pi^2 s(2L_1^{\mathrm{r}} - L_4^{\mathrm{r}}), \tag{III.3.11d}
$$
$$
F_\pi^4 \tilde{\mathcal{A}}_{\{2,2\}}^{(\mathrm{NLO},\zeta)} = \frac{s^2 \bar{J}(s)}{32} + \frac{\bar{J}(u)}{16}\left\{(u - 2M_\pi^2)^2\right\} - \frac{L + \kappa}{64}\left\{3s^2 - 2u(t - u)\right\}. \tag{III.3.11e}
$$

(Recall κ and L from section 2; $\bar{J}$ is defined in appendix A.) This is identical to the results given in ref. [31].

3.5 Poles and factorization

The six-meson amplitude has a simple pole whenever an internal propagator goes on-shell, i.e. $p_{ijk}^2 = M_\pi^2$ with $p_{ijk} = p_i + p_j + p_k$ for any indices i, j and k. As in ref. [19], the amplitude can therefore be separated into a part containing the pole and a nonpole part,

$$
A_{6\pi} = A_{6\pi}^{(\mathrm{pole})} + A_{6\pi}^{(\mathrm{nonpole})}, \tag{III.3.12}
$$

where the pole part can be factorized in terms of four-meson amplitudes:

$$
A_{6\pi}^{(\mathrm{pole})} = \sum_{P_{10}, b_\mathrm{o}} A_{4\pi}(p_i, b_i; p_j, b_j; p_k, b_k; -p_{ijk}, b_\mathrm{o})
$$
$$
\times \frac{-1}{p_{ijk}^2 - M_\pi^2} A_{4\pi}(p_\ell, b_\ell; p_m, b_m; p_n, b_n; p_{ijk}, b_\mathrm{o}). \tag{III.3.13}
$$

(a) 1×   **(b)** 10×
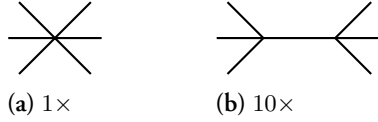
Figure III.3: LO topologies for the six-meson amplitude. As in fig. III.2, multiplicities are indicated.

Above, P_{10} represents the 10 distinct ways of distributing the indices $1, \dots, 6$ into two triples i, j, k and ℓ, m, n, and b_{o} is the flavor of the off-shell leg, i.e. the propagator.

This factorization can also be done at the stripped-amplitude level. With eq. (III.3.13) schematically summarized as $A_{6\pi}^{(\mathrm{pole})} \sim A_{4\pi} \times A_{4\pi}$, we correspondingly have

$$
\begin{aligned}
2\left(A_{\{6\}}^{(\mathrm{pole})} - \tfrac{1}{n} A_{\{3,3\}}^{(\mathrm{pole})}\right) &\sim A_{\{4\}} \times A_{\{4\}}\,, \\
A_{\{2,4\}}^{(\mathrm{pole})} &\sim A_{\{2,2\}} \times A_{\{4\}}\,, \\
2 A_{\{2,2,2\}}^{(\mathrm{pole})} &\sim A_{\{2,2\}} \times A_{\{2,2\}}\,,
\end{aligned}
\tag{III.3.14}
$$

with each $A_R^{(\mathrm{pole})}$ summed over $\mathbb{Z}_R$ instead of P_{10}, causing some symmetry factors.[8]

In eq. (III.3.13), the four-pion subamplitude is defined as usual, although $s + t + u = 3M_\pi^2 + p_{ijk}^2$ since one leg is off-shell. The residue at the pole is unique (since the on-shell four-meson amplitude is), but the extrapolation away from $p_{ijk}^2 = M_\pi^2$ is not. Correspondingly, the distribution of terms between the parts in eq. (III.3.12) is not unique. We choose to express $A_{4\pi}$ in *exactly* the form (III.3.11), which in turn fixes $A_{6\pi}^{(\mathrm{nonpole})}$. This choice by definition leaves both $A_{6\pi}^{(\mathrm{pole})}$ and $A_{6\pi}^{(\mathrm{nonpole})}$ parametrization-independent. However, the distribution of contributions from individual one-particle-reducible (1PR) diagrams remains parametrization-dependent, while one-particle-irreducible (1PI) diagrams only contribute to $A_{6\pi}^{(\mathrm{nonpole})}$.

By suitably deforming $A_{4\pi}$, it is in fact possible to make the tree-level $A_{6\pi}^{(\mathrm{nonpole})}$ vanish. This is the principle underlying Britto–Cachazo–Feng–Witten recursion [49, 50] and similar techniques, wherein many-particle amplitudes are recursively built up from smaller ones. This technique was used for the first published calculation of the NLO tree-level six-meson amplitude [33], but at least its standard configuration suffers from convergence problems at NNLO [20]. Significant work has been done on the topic of loop-level recursion techniques but is typically limited to loop integrands rather than complete amplitudes; see ref. [38, 39] and references therein. We make no use of such techniques here.

(a) $1\times$ **(b)** $20\times$ **(c)** $10\times$

(d) $1\times$ **(e)** $20\times$ **(f)** $10\times$

(g) $15\times$ **(h)** $60\times$ **(i)** $15\times$
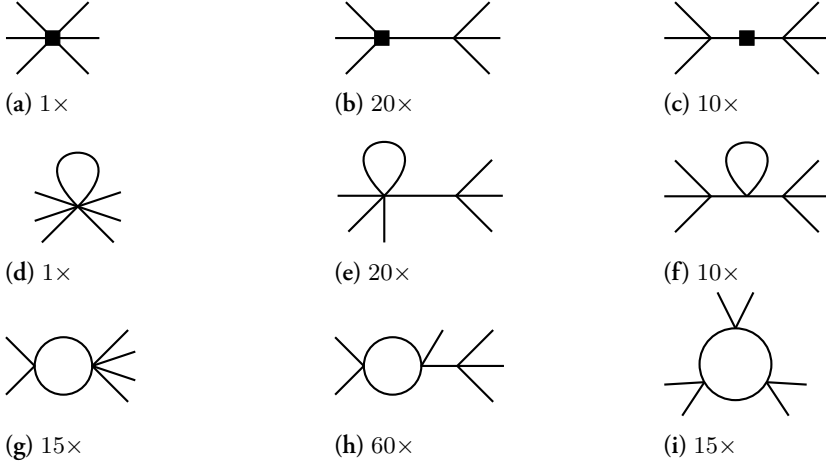
Figure III.4: NLO topologies for the six-meson amplitude. As in fig. III.2, multiplicities are indicated.

3.6   The six-meson amplitude

The nonpole LO six-meson amplitude contains only two nonzero subamplitudes, stemming from the 1 + 10 tree diagrams in fig. III.3:

$$\mathcal{A}_{\{6\}}^{(\text{LO, nonpole, 1})} = \frac{p_1 \cdot p_3 + p_3 \cdot p_5 + p_5 \cdot p_1}{2F_\pi^4}, \tag{III.3.15a}$$

$$\mathcal{A}_{\{3,3\}}^{(\text{LO, nonpole, 1})} = \frac{M_\pi^2 - p_1 \cdot p_2 - p_1 \cdot p_3 - p_2 \cdot p_3}{2nF_\pi^4}. \tag{III.3.15b}$$

The pole part is given by eq. (III.3.13). Note that eq. (III.3.15) is group-invariant, i.e. equal for SU and S_P^0 up to $n \to \zeta n$. This is true for all analogous LO k-meson amplitudes, and indeed all tree-level contributions, as is proven in appendix E.

The NLO amplitude, which is the main result of this work, stems from the diagrams in fig. III.4. Even when maximally simplified, it is rather lengthy, so we leave its explicit expressions to appendix D.

Let us now describe briefly the renormalization procedure for the NLO six-meson amplitude analogously to eq. (III.3.10). Regarding the 1PI diagrams [figs. III.4a, III.4d, III.4g and III.4i] which only contribute to the nonpole part, we can again write, schematically,

$$A_{6\pi}^{(\text{NLO})}\big|_{1\text{PI}} = \mathcal{M}_{1\text{-loop},1\text{PI}}^{(6\pi)} + \mathcal{M}_{\text{CT},1\text{PI}}^{(6\pi)} + 6(Z^{1/2} - 1)\mathcal{M}_{\text{LO}}^{(a)(2)} + \mathcal{M}_{\text{LO}}^{(a)(4)}. \tag{III.3.16}$$

The discussion of the 1PR part is a bit more involved. The double-pole part $\mathcal{M}^{2\text{-pole}}$ stemming from the contributions represented by diagrams depicted in figs. III.4c and III.4f

---

[8]Note that $A_{\{2,4\}}^{(\text{pole})} = 0$ at LO and $A_{\{2,2,2\}}^{(\text{pole})} = 0$ at NLO, since $A_{\{2,2\}} = 0$ at LO.

cancels with the piece due to the NLO propagator mass renormalization in the LO pole contribution $\mathcal{M}_{\text{LO}}^{(b)(2)}$ from fig. III.3b, denoted as $\mathcal{M}_{\text{LO}}^{(b)(4)}\big|_{\text{prop}}$, and consequently

$$\mathcal{M}^{\text{2-pole}} + 6(Z^{1/2} - 1)\mathcal{M}_{\text{LO}}^{(b)(2)} + \mathcal{M}_{\text{LO}}^{(b)(4)}\big|_{\text{prop}} = 8(Z^{1/2} - 1)\mathcal{M}_{\text{LO}}^{(b)(2)}. \qquad \text{(III.3.17)}$$

This, together with the LO contribution itself, $\mathcal{M}_{\text{LO}}^{(b)(2)}$, with vertices promoted to NLO in the same contribution, $\mathcal{M}_{\text{LO}}^{(b)(4)}\big|_{\text{vert}}$, and contributions stemming from the 1PR topologies shown in figs. III.4b, III.4e and III.4h gives, schematically,

$$A_{6\pi}^{(\text{LO+NLO})}\big|_{\text{1PR}} = \mathcal{M}_{\text{1-loop,1PR}}^{(6\pi)} + \mathcal{M}_{\text{CT,1PR}}^{(6\pi)} + 8(Z^{1/2} - 1)\mathcal{M}_{\text{LO}}^{(b)(2)} + \mathcal{M}_{\text{LO}}^{(b)(4)}\big|_{\text{vert}} + \mathcal{M}_{\text{LO}}^{(b)(2)},$$
$$\text{(III.3.18)}$$

which is the equivalent of two up-to-NLO $\pi\pi$ scatterings [analogous to $A_{4\pi}^{(\text{LO})} + A_{4\pi}^{(\text{NLO})}$ as from eqs. (III.3.9) and (III.3.10)] connected with the propagator, i.e. precisely the structure of eq. (III.3.13). Choosing the particular form of $A_{6\pi}^{(\text{pole})}$ as discussed earlier, the remainder with respect to $A_{6\pi}^{(\text{LO+NLO})}\big|_{\text{1PR}}$ is deferred to $A_{6\pi}^{(\text{nonpole})}$. What we call the nonpole part of the six-meson amplitude is thus the combination of such a remainder and the contributions of the 1PI diagrams from eq. (III.3.16).

3.7 Zero-momentum limit

In what follows, we choose a symmetric $3 \to 3$ scattering configuration given by the four-momenta

$$
\begin{aligned}
p_1 &= (E_p, p, 0, 0), & p_{2,3} &= \left(E_p, -\tfrac{1}{2}p, \pm\tfrac{\sqrt{3}}{2}p, 0\right), \\
p_4 &= (-E_p, 0, 0, p), & p_{5,6} &= \left(-E_p, \pm\tfrac{\sqrt{3}}{2}p, 0, -\tfrac{1}{2}p\right),
\end{aligned}
\qquad \text{(III.3.19)}
$$

with $E_p = \sqrt{p^2 + M_\pi^2}$. These only depend on a single parameter p, the modulus of three-momenta of all the mesons. In this kinematic setting, the zero-momentum limit of the stripped nonpole amplitudes up-to-and-including NLO take a simple group-universal form

$$
\begin{aligned}
F_\pi^2 \lim_{p\to 0} A_{\{6\}} = \;&-\frac{1}{2}\frac{M_\pi^2}{F_\pi^2} + \frac{M_\pi^4}{F_\pi^4}\left\{\frac{\xi}{4}(8L - \kappa) + (3L - 5\kappa)\left(\frac{10n + \xi}{36} - \frac{1}{\zeta n}\right)\right. \\
&\left. + \frac{\kappa}{2\zeta n} - 4(8L_0^{\text{r}} - L_5^{\text{r}} + 6L_8^{\text{r}})\right\}, \\
F_\pi^2 \lim_{p\to 0} A_{\{2,4\}} = \;&\frac{M_\pi^4}{F_\pi^4}\left\{\frac{1}{\zeta}(3L + \kappa) + 2(L - 2\kappa)\frac{1}{\zeta^2 n^2} - 16(L_2^{\text{r}} + L_4^{\text{r}} + 2L_6^{\text{r}})\right\}, \\
F_\pi^2 \lim_{p\to 0} A_{\{3,3\}} = \;&-\frac{1}{\zeta n}\frac{M_\pi^2}{F_\pi^2} + \frac{M_\pi^4}{F_\pi^4}\left\{\frac{3}{2\zeta}L - \frac{3}{2}(L - \kappa)\left(\frac{\xi}{\zeta n} - \frac{4}{\zeta^2 n^2}\right) - 16L_7^{\text{r}}\right. \\
&\left. - \frac{48}{\zeta n}(L_5^{\text{r}} - L_8^{\text{r}})\right\}, \\
F_\pi^2 \lim_{p\to 0} A_{\{2,2,2\}} = \;&\frac{M_\pi^4}{F_\pi^4}\left\{\frac{4\kappa}{\zeta^3 n^3}\right\}. \qquad \text{(III.3.20)}
\end{aligned}
$$

Due to the Adler zero $\lim_{\epsilon \to 0} A(p_1, \ldots, \epsilon p_i, \ldots) = 0$, which holds for any i in the massless case [51, 52], the zero-momentum limit is proportional to M_π^2 also in the general case.

It seems that eq. (III.3.20) is valid also for general momentum configurations rather than just eq. (III.3.19); this is explained in the next section. However, it is specifically the zero-momentum limit of $A_R(p_1, \ldots, p_6)$ where particles $1, 2, 3$ are in the initial state and $3, 4, 5$ in the final state. Different assignments of initial- and final-state particles will yield different zero-momentum limits. After accounting for $\mathbb{Z}_R^{+TR}$ symmetry, time-reversal symmetry, and the freedom to exchange particles within the initial and final states (which changes the stripped amplitude but not its zero-momentum limit[9]), there are 10 distinct limits, produced by the following:

$$
\left.
\begin{aligned}
&\lim_{p \to 0} A_R(p_1, p_2, p_3, p_4, p_5, p_6)\,, \\
&\lim_{p \to 0} A_R(p_1, p_4, p_2, p_5, p_3, p_6)\,,
\end{aligned}
\right\} \quad \text{for all } R\,,
$$
$$
\lim_{p \to 0} A_{\{6\}}(p_1, p_2, p_4, p_3, p_5, p_6)\,,
$$
$$
\lim_{p \to 0} A_{\{2,4\}}(p_1, p_4, p_2, p_3, p_5, p_6)\,.
$$
$$\tag{III.3.21}$$

Note that the first line reproduces eq. (III.3.20); in the interest of space, we do not reproduce the other cases. Also note that this is for $3 \to 3$ scattering, and that different limits will be obtained for $2 \to 4$ scattering.

4   Numerical results

We only present a few numerical results here since the full analysis of the finite volume and the subtraction of the two-body rescatterings is very nontrivial; see refs. [53, 54] and references therein. The numerical inputs we use are

$$
\begin{aligned}
M_\pi &= 0.139570\,\text{GeV}\,, &\quad \mu &= 0.77\,\text{GeV}\,, \\
F_\pi &= 0.0927\,\text{GeV}\,, &\quad n &= 3\,.
\end{aligned}
\tag{III.4.1}
$$

For LECs, we use the p^4 fit from Table 1 of ref. [55]:

$$
\begin{aligned}
L_1^{\mathrm{r}} &= 1.0 \times 10^{-3}\,, &\quad L_5^{\mathrm{r}} &= 1.2 \times 10^{-3}\,, \\
L_2^{\mathrm{r}} &= 1.6 \times 10^{-3}\,, &\quad L_6^{\mathrm{r}} &= 0\,, \\
L_3^{\mathrm{r}} &= -3.8 \times 10^{-3}\,, &\quad L_7^{\mathrm{r}} &= -0.3 \times 10^{-3}\,, \\
L_4^{\mathrm{r}} &= 0\,, &\quad L_8^{\mathrm{r}} &= 0.5 \times 10^{-3}\,.
\end{aligned}
\tag{III.4.2}
$$

For $n = 3$, we use $L_0^{\mathrm{r}} = 0$. Throughout this section, we use the kinematic setting of eq. (III.3.19).

---

[9] This is not necessarily true if the limit depends nontrivially on how it is approached. This seems not to be the case for this amplitude, although the analytic structure of C could hide some subtleties.
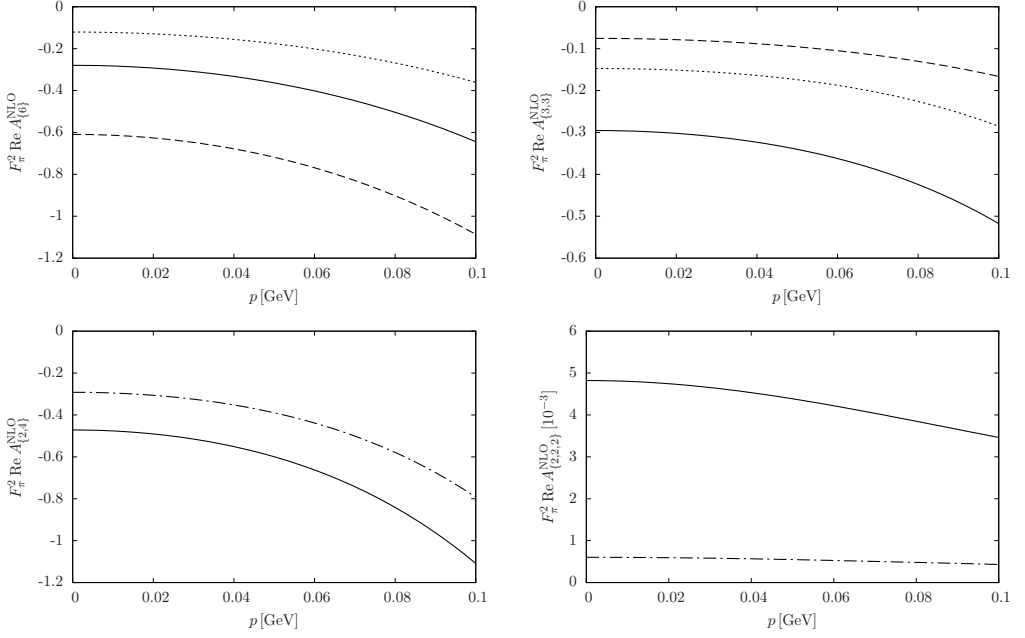
Figure III.5: Stripped amplitudes at NLO in the kinematic setting of eq. (III.3.19). The solid line stands for SU, the dashed line for SO and the dotted line for Sp. The dash-dotted line used in the two bottom panels represents the cases in which SO and Sp coincide. Note the extra factor of 10^{-3} for the $R = \{2,2,2\}$ stripped amplitude.

The resulting plots are shown in fig. III.5. Interestingly, $\mathcal{A}^{(\mathrm{NLO},\xi^2)}_{\{3,3\}}$ vanishes in this kinematic setting and hence does not contribute to the top right panel of fig. III.5.

We use the following shorthand notation for the value at $p = 0.1\,\mathrm{GeV}$:

$$A(p) \;\longrightarrow\; \hat{A} \equiv F_\pi^2 \times \mathrm{Re}\,A(p = 0.1\,\mathrm{GeV})\,. \tag{III.4.3}$$

Note that $\hat{A}$ is dimensionless. Using this notation, values for general n are given in table III.1. As fig. III.5 shows, the relative sizes of these values are representative across a broader energy range.

The pole part is clearly the dominant contribution, but in a sense the nonpole part is the interesting one, since it is not directly related to the previously known four-meson amplitude. The NLO part is smaller than the LO part, but not by much; perturbative convergence is understandably poor, with breakdown expected at a scale of $4\pi F_\pi/\sqrt{n}$ [28], i.e. $\approx 5 M_\pi$ at $n = 3$.

Among the NLO nonpole subamplitudes, neither is clearly dominant: $A_{\{6\}}$, $A_{\{2,4\}}$ and $A_{\{3,3\}}$ are comparable in magnitude, as are $\mathcal{A}^{(1)}$, $\mathcal{A}^{(\xi)}$ and $\mathcal{A}^{(\zeta)}$. Among the NLO pole subamplitudes, $\mathcal{A}^{(1)}_{\{6\}}$, $\mathcal{A}^{(\zeta)}_{\{6\}}$ and $\mathcal{A}^{(1)}_{\{2,4\}}$ dominate. Thus, the results of the three QCD-like

Table III.1: The different contributions to the six-meson amplitude, evaluated at $p = 0.1\,\text{GeV}$ as described in eq. (III.4.3), using the momentum configuration (III.3.19). Note that we only quote the real part, and that we multiply by a suitable power of F_π to make the result dimensionless. We omit subamplitudes that are identically zero.

	Pole part	Non-pole part
$\hat{\mathcal{A}}^{(\text{LO},1)}_{\{6\}}$	86.4	-1.68
$\hat{\mathcal{A}}^{(\text{LO},1)}_{\{3,3\}}$	$5.06\frac{1}{n}$	$-4.89\frac{1}{n}$
$\hat{\mathcal{A}}^{(\text{NLO},1)}_{\{6\}}$	$-36.2 - 10.2\frac{1}{n}$	$-0.0119 + 0.477\frac{1}{n}$
$\hat{\mathcal{A}}^{(\text{NLO},\xi)}_{\{6\}}$	0.248	-0.363
$\hat{\mathcal{A}}^{(\text{NLO},\zeta)}_{\{6\}}$	$12.4\,n$	$-0.264\,n$
$\hat{\mathcal{A}}^{(\text{NLO},1)}_{\{2,4\}}$	$20.5 + 4.84\frac{1}{n^2}$	$-0.492 - 0.293\frac{1}{n^2}$
$\hat{\mathcal{A}}^{(\text{NLO},\zeta)}_{\{2,4\}}$	24.4	-0.584
$\hat{\mathcal{A}}^{(\text{NLO},1)}_{\{3,3\}}$	$1.75\frac{1}{n} + 0.624\frac{1}{n^2}$	$0.0247 - 0.606\frac{1}{n} - 0.741\frac{1}{n^2}$
$\hat{\mathcal{A}}^{(\text{NLO},\xi)}_{\{3,3\}}$	$-0.328\frac{1}{n}$	$0.357\frac{1}{n}$
$\hat{\mathcal{A}}^{(\text{NLO},\xi^2)}_{\{3,3\}}$	—	0
$\hat{\mathcal{A}}^{(\text{NLO},\zeta)}_{\{3,3\}}$	-0.501	-0.258
$\hat{\mathcal{A}}^{(\text{NLO},1)}_{\{2,2,2\}}$	—	$0.0935\frac{1}{n^3}$

theories are of similar magnitude. In the large-n limit, we expect $A_{\{6\}}$ to dominate due to the positive power of n in $\mathcal{A}^{(\zeta)}_{\{6\}}$; following eq. (III.3.8), this also implies that the stripped amplitudes of the three theories will become equal in this limit.

Besides the kinematic configuration (III.3.19), we have numerically evaluated the amplitude for a random sample of $3 \to 3$ scattering events generated with the RAMBO algorithm [56]. These samples confirm that $\mathcal{A}^{(\text{NLO},\xi^2)}_{\{3,3\}}$ is not generally zero. We obtained zero-momentum limits by uniformly scaling the random three-momenta by a factor $\epsilon \to 0$ while keeping the particles on-shell.[10] This consistently resulted in the same numerical values as eq. (III.3.20), allowing us to conclude that eq. (III.3.20) is the general uniform zero-momentum limit of $A_R(p_1, \ldots, p_6)$ in $3 \to 3$ scattering, rather than a special case for the configuration (III.3.19). The same is true for the other limits described in eq. (III.3.21).

---

[10] This introduces an energy-conservation-violating term of order ϵ^2. Violating either on-shellness or conservation of energy (or momentum) is inevitable when taking such a zero-momentum limit. It would be simpler to scale the four-momenta, but that would give the massless zero-momentum limit, which is zero and therefore not very interesting.

5 Conclusions

In this paper, we calculated the meson mass, decay constant, four-meson and six-meson amplitudes to NLO in the QCD-like theories, i.e. besides the QCD-related case SU, we also consider the SO and Sp symmetry-breaking patterns. In section 2, we introduced the relevant NLO Lagrangian in analogy with three(-quark)-flavor ChPT [2, 3]. The mass, decay constant and the four-meson amplitude agree for $n = 3$ with refs. [2, 3, 27, 29, 30, 57, 58] and with the general-n results of refs. [28, 31, 59–61].

Our main result is the six-meson amplitude, which can be written in terms of four independent flavor-stripped amplitudes [for detailed structure, see eq. (III.3.3)], as compared to a single amplitude in the $O(N+1)/O(N)$ case studied in ref. [19]. We split the whole amplitude into pole and nonpole parts; see eq. (III.3.12). The pole part is given in eq. (III.3.13), where we chose to employ the off-shell four-meson amplitude in the form of eqs. (III.3.9) and (III.3.11) generalizing (beyond $n = 3$) the amplitude given in refs. [57, 58] and exactly matching that in ref. [31].

The expression for the nonpole part is rather lengthy. We thus further divide the four flavor-stripped amplitudes into group-universal subamplitudes in order to account for all the three QCD-like theories in a concise way. By employing symmetries through the deorbiting procedure described in appendix C, we obtain the resulting 10 non-vanishing subamplitudes presented in appendix D. The nontrivial choice of a redundant but highly symmetric basis of tensor triangle loop integrals (for details, see appendix A) and of kinematic invariants (appendix C) allows for a fairly compact expression. While the result is still too lengthy and complicated to be grasped fully, the division into subamplitudes along with further analysis in appendix E allows many of its features to be understood.

In the kinematic setting of eq. (III.3.19), we present the analytical results for the zero-momentum limit in eq. (III.3.20). Some numerical results for this particular momentum configuration are presented in section 4.

In the process of our calculations, we devised a systematic procedure (deorbiting) for simplifying amplitudes beyond what is possible with stripping alone. Previous work, e.g. refs. [19, 31, 35], manually structure their results in similar ways, but this quickly becomes difficult with larger numbers of kinematic variables and more complicated symmetries. These issues are, at least partly, resolved by our simplification scheme, which should be applicable also beyond the present scope.

We see limited interest in computing the NNLO counterpart of this result. Several LECs (terms 49–63 in $\mathcal{L}^{(6)}$ [46]) that do not appear in lower-multiplicity amplitudes enter here and are so far undetermined. All relevant two-loop integrals are known (see e.g. ref. [31]) except for the five-propagator sunset topology, which we expect to be very difficult. There is also the matter of expressing the two-loop integrals in a symmetry-compliant way like in appendix A.

We believe that our techniques would make the NLO eight-meson amplitude accessible, but such a calculation is currently not motivated by lattice developments. Besides the

larger number of diagrams and longer expressions, the main technical hurdles would be extending appendix A to a similar treatment of box integrals, and extending appendix C to $\mathbb{Z}^{+\text{TR}}_{\{8\}}, \mathbb{Z}^{+\text{TR}}_{\{2,6\}}, \mathbb{Z}^{+\text{TR}}_{\{2,3,3\}}$, etc.

Work is in progress to combine our results with the methods for extracting three-body scattering from finite volume in lattice QCD. We expect that our results may also be of interest for the amplitude community.

Acknowledgements

A   Conventions for the loop integrals

Throughout the paper, we treat the momenta $(p_1, \ldots, p_6)$ as incoming, and we introduce the following independent combinations of momenta:

$$q_1 = p_1 + p_2 , \qquad q_2 = p_3 + p_4 , \qquad q_3 = p_5 + p_6 ,$$
$$r_1 = p_1 - p_2 , \qquad r_2 = p_3 - p_4 , \qquad r_3 = p_5 - p_6 . \tag{III.1.1}$$

Note that we use the same notation and conventions as in ref. [19]. In particular, the integrals are defined in Appendix A therein; here, we restate them along with some clarifications.

The functions we use to represent our results are very closely related to the standard Passarino–Veltman one-loop integrals A_0, B_0 and C_0. To fix our notation, let us present explicitly the simpler integrals with one and two propagators. In what follows, we use the compact notation for the Feynman denominators with loop momentum l,

$$D(q_i) \equiv (l - q_i)^2 - M^2 , \tag{III.1.2}$$

while setting $D_0 \equiv D(0) = l^2 - M^2$. Having in mind eq. (III.2.11),

$$\frac{1}{\tilde{\epsilon}} \equiv \frac{1}{\epsilon} - \gamma_E + \log 4\pi - \log \mu^2 + 1 , \tag{III.1.3}$$

and that we, like in section 2, set

$$\kappa = \frac{1}{16\pi^2} , \qquad L = \kappa \log \frac{M^2}{\mu^2} , \tag{III.1.4}$$

the integrals read

$$A \equiv \kappa A_0(M^2) = \frac{1}{i} \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{1}{D_0} = M^2 \kappa \frac{1}{\tilde{\epsilon}} - M^2 L , \tag{III.1.5}$$

$$B(q^2) \equiv \kappa B_0(q^2, M^2, M^2) = \frac{1}{i} \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{1}{D_0 D(q)} = \kappa \frac{1}{\bar{\epsilon}} - \kappa - L + \bar{J}(q^2) . \qquad \text{(III.1.6)}$$

We employ the standard definition for $\bar{J}(q^2)$:

$$\bar{J}(q^2) \equiv \kappa \left(2 + \beta \log \frac{\beta - 1}{\beta + 1} \right) , \qquad \text{(III.1.7)}$$

with $\beta \equiv \beta(q^2) = \sqrt{1 - \frac{4M^2}{q^2}}$. The terms L and $\bar{J}(q^2)$ we use to express our results thus absorb the factors of $\frac{1}{16\pi^2}$.

Let us emphasize that it is the tensor triangle one-loop integrals of higher ranks which generate lengthy expressions upon reduction to the scalar ones. It therefore turned out to be more convenient to use a specific basis for the tensor integrals with particular symmetry properties. Regarding the rank-3 integrals, the combination

$$
\begin{aligned}
C_3(p_1, p_2, \ldots, p_6) = {} & \frac{1}{3} \frac{1}{i} \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{l \cdot r_1 \, l \cdot r_2 \, l \cdot r_3}{D_0} \\
& \times \left[\frac{1}{D(q_1) D(-q_2)} + \frac{1}{D(q_2) D(-q_3)} + \frac{1}{D(q_3) D(-q_1)} \right] \quad \text{(III.1.8)}
\end{aligned}
$$

has more symmetries than the first term only and is UV finite. It is antisymmetric under the interchange of the momenta inside each pair [pairs being here (p_1, p_2), (p_3, p_4) and (p_5, p_6)] and antisymmetric under the interchange of two pairs. The rank-2 integral can be defined as

$$C_{21}(p_1, p_2, \ldots, p_6) = \frac{1}{i} \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{l \cdot r_1 \, l \cdot r_2}{D_0 D(q_1) D(-q_2)} . \qquad \text{(III.1.9)}$$

It is antisymmetric under the interchange $p_1 \leftrightarrow p_2$ and symmetric under $(p_1, p_2) \leftrightarrow (p_3, p_4)$ and $p_5 \leftrightarrow p_6$. The integral with one product $l \cdot r_i$ in the numerator can be defined as

$$C_{11}(p_1, p_2, \ldots, p_6) = \frac{1}{i} \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{l \cdot r_3}{D_0 D(q_1) D(-q_2)} . \qquad \text{(III.1.10)}$$

It is antisymmetric under the interchange $p_5 \leftrightarrow p_6$, $(p_1, p_2) \leftrightarrow (p_3, p_4)$ and is symmetric under $p_1 \leftrightarrow p_2$ and $p_3 \leftrightarrow p_4$. Owing to the symmetries, other integrals of ranks 2 and 1 can be expressed in terms of C_{21} and C_{11} and integrals with lower ranks, respectively, so we only need those to write out our final result. Finally, we define

$$C(p_1, p_2, \ldots, p_6) = \frac{1}{i} \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{1}{D_0 D(q_1) D(-q_2)} , \qquad \text{(III.1.11)}$$

which is symmetric under $p_1 \leftrightarrow p_2$ and under all pair interchanges. It is related to C_0 as

$$C(p_1, p_2, \ldots, p_6) = \kappa C_0(q_1^2, q_2^2, q_3^2, M^2, M^2, M^2) , \qquad \text{(III.1.12)}$$

in which case the mentioned symmetries are seen trivially due to equal masses.

We express the amplitude in terms of C_3, C_{21}, C_{11} and C. As already mentioned, the former three can be expressed in terms of C, but the expressions are cumbersome and lead to a very long expression for the amplitude. We have therefore kept all these four, among which only C_{21} contains a UV-infinite part:

$$C_{21}(p_1, p_2, \ldots, p_6) = \kappa \, \frac{r_1 \cdot r_2}{4} \frac{1}{\tilde{\epsilon}} + \overline{C}_{21}(p_1, p_2, \ldots, p_6) \,. \tag{III.1.13}$$

B General parametrization

In this appendix, we show how to parametrize a special unitary matrix u in full generality. Special cases of what follows give parametrizations such as the four used in ref. [62].

A special unitary matrix $\hat{U}$ can always be written as an exponential of a Hermitian traceless matrix ϕ:

$$\hat{U} = \exp(i\phi) \,, \quad \phi^\dagger = \phi \,, \quad \langle \phi \rangle = 0 \,. \tag{III.2.1}$$

The obvious way to write a general parametrization is

$$\hat{U} \longrightarrow \sum_{m=0}^{\infty} b_m \phi^m \,, \tag{III.2.2}$$

with unitarity conditions relating the b_m. As proven in ref. [63], the only generally valid solution where b_m are c-numbers is $\hat{U} = \exp(i\phi)$, although sufficiently low order, other such parametrizations are valid and useful; see e.g. refs. [35, 64]. Generally, however, b_m are functions of traces of powers of ϕ, as seen in ref. [62]; this complicates the unitarity conditions.

Here, we take the alternative approach of redefining ϕ to $\phi'(\phi)$ with $\phi'^\dagger = \phi'$ and $\langle \phi' \rangle = 0$, and keeping $\hat{U} = \exp(i\phi')$. Under the unbroken (vector) part of the chiral transformation, $\phi \to g_V \phi g_V^\dagger$, and we want $\phi' \to g_V \phi' g_V^\dagger$. The redefined ϕ' is thus a series in ϕ and traces of powers of ϕ:

$$\phi' = \phi + \sum_{\substack{m \geq 2, \, j \geq 0, \\ i_0 + \cdots + i_j = m, \\ i_0 \geq 0; \, i_1 \geq i_2 \geq \cdots i_j \geq 2}} a_{i_0 \ldots i_j} \phi^{i_0} \langle \phi^{i_1} \rangle \cdots \langle \phi^{i_j} \rangle \,. \tag{III.2.3}$$

Further restrictions follow from using intrinsic parity, i.e. employing $\phi' \to -\phi'$ if $\phi \to -\phi$, thus allowing only for odd values of m, and applying $\phi'^\dagger = \phi'$, which requires the $a_{i_0 \ldots i_j}$ to be real. The condition $\langle \phi' \rangle = 0$ determines all the $a_{0 i_1 \ldots i_j}$ (those with $i_0 = 0$) to be $a_{0 i_1 \ldots i_j} = -a_{i_1 \ldots i_j}/\zeta n$. Hence, all terms relevant for the six-meson amplitude discussed in this work, introducing 18 extra unconstrained parameters,[11] are in group-universal form

$$\phi' = \phi + a_3 \Big(\phi^3 - \tfrac{1}{\zeta n} \langle \phi^3 \rangle \Big) + a_{12} \, \phi \langle \phi^2 \rangle$$

---

[11]In general, the number of free parameters at order ϕ^m is equal to the number of ways to partition $m - 1$ into positive integers, i.e. Online Encyclopedia of Integer Sequences (OEIS) sequence A058696 starting with $2, 5, 11, 22, 43, 77, 135$. Let us briefly sketch the proof of this. Let $i_0 \geq 1$ since $a_{0 i_1 \ldots i_j} = -a_{i_1 \ldots i_j}/\zeta n$. Then,

$$
+ a_5\left(\phi^5 - \tfrac{1}{\zeta n}\langle\phi^5\rangle\right) + a_{32}\left(\phi^3\langle\phi^2\rangle - \tfrac{1}{\zeta n}\langle\phi^3\rangle\langle\phi^2\rangle\right) + a_{23}\left(\phi^2\langle\phi^3\rangle - \tfrac{1}{\zeta n}\langle\phi^3\rangle\langle\phi^2\rangle\right)
$$
$$
+ a_{14}\,\phi\langle\phi^4\rangle + a_{122}\,\phi\langle\phi^2\rangle^2
$$
$$
+ a_7\left(\phi^7 - \tfrac{1}{\zeta n}\langle\phi^7\rangle\right) + a_{52}\left(\phi^5\langle\phi^2\rangle - \tfrac{1}{\zeta n}\langle\phi^5\rangle\langle\phi^2\rangle\right) + a_{43}\left(\phi^4\langle\phi^3\rangle - \tfrac{1}{\zeta n}\langle\phi^4\rangle\langle\phi^3\rangle\right)
$$
$$
+ a_{34}\left(\phi^3\langle\phi^4\rangle - \tfrac{1}{\zeta n}\langle\phi^4\rangle\langle\phi^3\rangle\right) + a_{322}\left(\phi^3\langle\phi^2\rangle^2 - \tfrac{1}{\zeta n}\langle\phi^3\rangle\langle\phi^2\rangle^2\right)
$$
$$
+ a_{25}\left(\phi^2\langle\phi^5\rangle - \tfrac{1}{\zeta n}\langle\phi^5\rangle\langle\phi^2\rangle\right) + a_{232}\left(\phi^2\langle\phi^3\rangle\langle\phi^2\rangle - \tfrac{1}{\zeta n}\langle\phi^3\rangle\langle\phi^2\rangle^2\right)
$$
$$
+ a_{16}\,\phi\langle\phi^6\rangle + a_{142}\,\phi\langle\phi^4\rangle\langle\phi^2\rangle + a_{133}\,\phi\langle\phi^3\rangle^2 + a_{1222}\,\phi\langle\phi^2\rangle^3
$$
$$
+ \mathcal{O}\!\left(\phi^9\right). \tag{III.2.4}
$$

The presence of traces in all terms except the first confirms and generalizes the conclusions of ref. [63].

Taking this general form to define u via eq. (III.2.6) with $\phi \equiv \phi^a t^a/\sqrt{2}F$ and plugging it into the Lagrangian (III.2.1) adds an extra cross-check of one's calculations, since the physical amplitude cannot depend on $a_{i_0 \ldots i_j}$.

C Deorbiting and closed bases of kinematic invariants

In this appendix, we briefly describe the method used for the final simplification step of reducing $\mathcal{A}_R^{(i)}$ to $\tilde{\mathcal{A}}_R^{(i)}$ with the property (III.3.7). This is a development of an *ad hoc* technique used in ref. [20]. Recall that $\tilde{\mathcal{A}}_R^{(i)}$ is not unique and that our aim is to make its expression as short as possible.

Consider some class of objects x and a group $\mathcal{G}$ (in our case, x are products of one-loop integral functions and kinematic invariants, and $\mathcal{G} = \mathbb{Z}_R^{+\text{TR}}$). In standard nomenclature, the set of objects obtained by acting with $\mathcal{G}$ on x is called the *orbit* of x, denoted $\mathcal{G}\cdot x$; formally,

$$
\mathcal{G}\cdot x = \{g\cdot x \mid g \in \mathcal{G}\}. \tag{III.3.1}
$$

Consider then an expression X composed of a sum of objects x. Reducing it to $\tilde{X}$ such that $X = \sum_{g\in\mathcal{G}} g\cdot\tilde{X}$ (where $g\cdot\tilde{X}$ indicates acting with g on each term in $\tilde{X}$) is done using the following algorithm:

1. Start with $\tilde{X} = 0$.

---

rewrite each term in eq. (III.2.3) as

$$
a_{i_0 i_1,\ldots i_j}\,\phi\langle\phi^{i_1}\rangle\langle\phi^{i_2}\rangle\cdots\langle\phi^{i_j}\rangle\,\underbrace{\phi\phi\cdots\phi}_{i_0 - 1\ \phi\text{'s}}
$$

In total, there are m factors of ϕ, of which all but the first can be arbitrarily partitioned like $m - 1 \to i_1,\ldots,i_j,1,\ldots,1$. Since $\langle\phi\rangle = 0$ and $i_1 \geq i_2 \geq \cdots i_j \geq 2$, we can unambiguously associate the 1's with ϕ's outside traces and the other elements of the partition with $i_1,\ldots i_j$. This demonstrates the one-to-one correspondence between independent parameters $a_{i_0 i_1,\ldots i_j}$ and partitions of $m - 1$.

2. Select the first term x in X, under some arbitrary but consistent ordering of the terms.[12]

3. Compute the orbit $\mathcal{G} \cdot x$ and the symmetry factor $S = |\mathcal{G}|/|\mathcal{G} \cdot x|$ (this is always an integer).

4. Add x to $\tilde{X}$, and subtract $\frac{1}{S} \sum_{g \in \mathcal{G}} g \cdot x$ from X. (Now, no element of $\mathcal{G} \cdot x$ appears in X.)

5. Repeat from step 2 until $X = 0$.

The symmetry factor compensates for how each element of $\mathcal{G} \cdot x$ appears S times in the sum $\sum_{g \in \mathcal{G}} g \cdot x$.

Optimally, each object x that appears in any orbit should be a single term, not a sum of other objects. We will call this property being *closed* under $\mathcal{G}$. Without this property, the algorithm may yield poor results or not terminate at all. However, if the class of objects is closed under $\mathcal{G}$, it is easy to see that no orbits overlap and that the algorithm results in an $\tilde{X}$ that is the shortest possible subexpression of X, granted that there are no additional symmetries that are not taken into account.

In the context of our amplitudes, we therefore need to carefully choose our basis of kinematic invariants. We define them in terms of *generalized Mandelstam variables*

$$\hat{s}_{ij\ldots} \equiv (p_i + p_j + \ldots)^2 \,, \tag{III.3.2}$$

which relate to the usual ones as

$$s = \hat{s}_{12}, \quad t = \hat{s}_{13}, \quad u = \hat{s}_{23}. \tag{III.3.3}$$

Thus, the pairs s, u and t, u form closed bases under $\mathbb{Z}_R^{+\text{TR}}$ for $R = \{4\}$ and $\{2, 2\}$, respectively. The former generalizes straightforwardly to all $R = \{2k\}$, including the $R = \{6\}$ basis $s_1, \ldots, s_9$ with[13]

$$s_1 = \hat{s}_{12}, \ s_2 = \hat{s}_{23}, \ \ldots, \ s_5 = \hat{s}_{56}, \ s_6 = \hat{s}_{61}, \ s_7 = \hat{s}_{123}, \ s_8 = \hat{s}_{234}, \ s_9 = \hat{s}_{345}. \tag{III.3.4}$$

Finding closed bases is much harder in the other cases. For $R = \{2, 4\}$ we use the basis[14]

$$\begin{aligned}
t_1 &= \hat{s}_{123}, \quad t_2 = \hat{s}_{124}, \quad t_3 = \hat{s}_{125}, \quad t_4 = \hat{s}_{126}, \\
t_5 &= \hat{s}_{234}, \quad t_6 = \hat{s}_{245}, \quad t_7 = \hat{s}_{256}, \quad t_8 = \hat{s}_{263}, \\
t_9 &= \hat{s}_{135} - \hat{s}_{146},
\end{aligned} \tag{III.3.5}$$

---

[12]In practice, we use the internal ordering of FORM, with some modifications.

[13]This basis is valid for 5 or more space-time dimensions; with 4, the correct number of kinematic degrees of freedom is 8, not 9. However, the 9th variable is related to the other 8 through the nonlinear Gram determinant relation, so for the sake of simplicity we ignore this and use 9-element bases.

For a k-particle process in d dimensions, a similar basis of generalized Mandelstam variables will have $k(k-3)/2$ elements, as is easily found by counting products $p_i \cdot p_j$ and accounting for $p_i^2 = M_\pi^2$ and $\sum_i p_i^\mu = 0$. This is redundant when $k < d-1$; then, with only $d-1$ independent components in each p_i, the number of kinematic degrees of freedom after accounting for $\sum_i p_i^\mu = 0$ and subtracting the dimension of the Lorentz group gives $(d-1)k - d(d+1)/2$, i.e. $3k-10$ when $d = 4$. This naïve counting does not apply for large d, since the momentum vectors live in a $(k-1)$-dimensional subspace; thus, there are $k(k-3)/2$ degrees of freedom when $k \geq d-1$.

[14]This is a new basis; the one in ref. [20] is not closed under trace-reversal. It was obtained using similar methods to the $\{3, 3\}$ and $\{2, 2, 2\}$ bases derived in that paper (note that $\mathbb{Z}_{\{2,4\}}^{+\text{TR}}$, unlike $\mathbb{Z}_{\{2,4\}}$, is non-Abelian).

and for $R = \{3,3\}$[15]

$$
\begin{aligned}
&u_1 = \hat{s}_{14}, && u_4 = \hat{s}_{25}, && u_7 = \hat{s}_{36}, \\
&u_2 = \hat{s}_{24}, && u_5 = \hat{s}_{35}, && u_8 = \hat{s}_{16}, \\
&u_3 = \hat{s}_{34}, && u_6 = \hat{s}_{15}, && u_9 = \hat{s}_{26}.
\end{aligned}
\tag{III.3.6}
$$

No basis is needed for $R = \{2,2,2\}$ here due to the simplicity of $A_{\{2,2,2\}}$. References [19, 20] provide two different $R = \{2,2,2\}$ bases.

There is no need to apply similar considerations to the loop integrals $\bar{J}$ and C_X, since the inherent symmetries of these functions are much simpler than those imposed on $\hat{s}$ by the kinematics. C (C_3) is (anti)symmetric under $\mathbb{Z}_{\{2,2,2\}}$ acting on its arguments, while C_{11} and C_{21} are symmetric or antisymmetric under various subgroups thereof. The symmetries of C_X and those of the stripped amplitudes interplay nontrivially, giving rise to several orbits. Denoting by $i \cdots j(p)$ the orbit that has p distinct elements including $C_X(p_i, \ldots, p_j)$, the orbits of C and C_3 under various $\mathbb{Z}_R$ are

$$
\begin{aligned}
\mathbb{Z}_6 : \quad &\mathbf{123456(2)}, 142536(1), 152436(3), \mathbf{162435(6)}, \mathbf{162534(3)}, \\
\mathbb{Z}_{\{2,4\}} : \quad &\mathbf{123456(2)}, 123546(1), 162435(4), \mathbf{162345(8)}, \\
\mathbb{Z}_{\{3,3\}} : \quad &\mathbf{123456(9)}, \mathbf{142536(3)}, \mathbf{162534(3)}, \\
\mathbb{Z}_{\{2,2,2\}} : \quad &\mathbf{123456(1)}, 162435(8), 162534(6),
\end{aligned}
\tag{III.3.7}
$$

and those of C_{11} and C_{21} are

$$
\begin{aligned}
\mathbb{Z}_6 : \quad &\left\{ \begin{aligned} &\mathbf{123456(6)}, \mathbf{235614(3)}, 253614(3), 254613(6), 263415(6), 263514(3), \\ &264513(6), 354612(6), 364512(6), \end{aligned} \right. \\
\mathbb{Z}_{\{2,4\}} : \quad &\left\{ \begin{aligned} &123456(4), 124635(2), 162435(4), \mathbf{132456(8)}, 263415(8), 263514(8), \\ &264513(8), \mathbf{354612(1)}, \mathbf{345612(2)}, \end{aligned} \right. \\
\mathbb{Z}_{\{3,3\}} : \quad &123456(18), \mathbf{234516(9)}, 263415(9), 263514(9), \\
\mathbb{Z}_{\{2,2,2\}} : \quad &123456(3), 263415(12), 264513(24), 364512(6).
\end{aligned}
\tag{III.3.8}
$$

Only those marked in bold actually appear in the amplitude. This can be understood from the limited arrangements of legs around the diagram fig. III.4i that produce $\mathcal{F}_R(b_1, \ldots, b_6)$ as a flavor structure, as is clarified by the technology of appendix E.

D   The six-meson amplitude: expressions

In this appendix, we explicitly present the subamplitudes $\tilde{\mathcal{A}}_R^{(i)}$, whose relation to the full amplitude is given in section 3. Note that we only present the nonpole part here; the pole part is implicitly expressed in eq. (III.3.12). We use kinematic variables s_i, t_i, u_i ($i =$

---

[15]This is quite different from the one used in ref. [20] and is much simpler — it is just one of the nonets formed under $\mathbb{Z}_{\{3,3\}}$. The reason for it being overlooked can be traced back to ref. [37], where an effort was made to include the element $\hat{s}_{123}$ in the basis.

$1, \ldots, 9$) which are motivated and defined in appendix C. With these, the LO amplitude of eq. (III.3.15) can be deorbited to

$$F_\pi^4 \tilde{\mathcal{A}}_{\{6\}}^{(\mathrm{LO},1)} = \frac{M_\pi^2 - 2s_1 + s_9}{16}, \tag{III.4.1a}$$

$$F_\pi^4 \tilde{\mathcal{A}}_{\{3,3\}}^{(\mathrm{LO},1)} = \frac{9u_9 - 8M_\pi^2}{288n}, \tag{III.4.1b}$$

and the NLO part is[16,17,18]

$$
\begin{aligned}
F_\pi^6 \tilde{\mathcal{A}}_{\{6\}}^{(\mathrm{NLO},1)} = {} & \frac{M_\pi^4}{4n}\left\{ \bar{J}(p_1,p_2) - L - \kappa \right\} \\
& + \frac{M_\pi^4}{8n}\Big\{ 2C_{11}(p_1,\ldots,p_6) + 2C(p_1,\ldots,p_6)\big[s_7 - M_\pi^2\big] \\
& \qquad + C(p_1,p_6,p_2,p_5,p_3,p_4)\big[s_8 - 2s_6 + s_9\big] \Big\} \\
& - L_0^{\mathrm{r}}\Big\{ 2M_\pi^4 + 4M_\pi^2(s_7 - 2s_1) + s_1(s_1 + 2s_4 + 3s_5 + 2s_6 - 3s_7) \\
& \qquad - s_7(3s_2 + 2s_3 - s_7 - s_9) \Big\} \\
& - \tfrac{1}{4}L_3^{\mathrm{r}}\Big\{ 7M_\pi^4 - 2M_\pi^2(7s_1 - 4s_7) + s_1(2s_1 + 2s_4 + 2s_5 - 3s_7 - 3s_9) + s_7^2 \Big\} \\
& + M_\pi^2 L_5^{\mathrm{r}}\big\{ 2M_\pi^2 - 2s_1 + s_7 \big\} - 2M_\pi^4 L_8^{\mathrm{r}} \tag{III.4.2a}
\end{aligned}
$$

$$
\begin{aligned}
F_\pi^6 \tilde{\mathcal{A}}_{\{6\}}^{(\mathrm{NLO},\xi)} = {} & -\tfrac{1}{48}C_3(p_1,\ldots,p_6) \\
& - \tfrac{1}{96}\Big\{ \overline{C}_{21}(p_1,\ldots,p_6)\big[4M_\pi^2 - s_1 - s_3 + 4s_5 + 2s_9\big] \\
& \qquad - 3\overline{C}_{21}(p_2,p_3,p_5,p_6,p_1,p_4)\big[2s_5 - s_7 - s_8\big] \Big\} \\
& - \tfrac{1}{64}\Big\{ C_{11}(p_1,\ldots,p_6)\big[2M_\pi^2(s_1 + s_3) - s_1 s_3\big] \\
& \qquad + 8C_{11}(p_3,p_5,p_4,p_6,p_1,p_2)\big[(p_3 \cdot p_5)(p_4 \cdot p_6)\big] \Big\} \\
& + \tfrac{1}{384}C(p_1,\ldots,p_6)\Big\{ 24M_\pi^4 s_1 + 12M_\pi^2 s_1(s_1 - s_3 - s_9 - s_7) \\
& \qquad - 3(s_1^2 s_3 + s_1 s_3^2) + 2s_1 s_3(2s_5 + 3s_9) \Big\} \\
& + \tfrac{1}{128}C(p_1,p_6,p_2,p_5,p_3,p_4)\big\{ s_3 s_6(2s_6 - s_8 - s_9) \big\} \\
& + \tfrac{1}{16}C(p_1,p_6,p_2,p_4,p_3,p_5)\Big\{ (p_2 \cdot p_4)(p_3 \cdot p_5) \\
& \qquad \times (4M_\pi^2 - 2s_1 - s_2 - s_4 - 2s_5 - 2s_6 + 2s_7 + s_8 + s_9) \Big\} \\
& + \tfrac{1}{192}C(p_1,p_4,p_2,p_5,p_3,p_6)\Big\{ 6s_3 s_4 s_5 + 2s_2 s_4 s_6 + 6s_1(s_8^2 - s_6 s_7) \\
& \qquad - s_9\big[6s_2 s_4 + 6s_5(s_3 + s_4) + s_8\big(2s_7 - 6(s_3 + s_4 + s_5) + 3(s_8 + s_9)\big)\big] \Big\}
\end{aligned}
$$

$$
\begin{aligned}
&- \tfrac{1}{384}\bar{J}(p_1,p_2)\Big\{32M_\pi^4 - M_\pi^2\big[2s_1 + 8s_2 + 5s_3 + 5s_5 + 2(4s_6 + s_7 - 2s_8 + s_9)\big] \\
&\qquad\qquad - s_1(3s_1 - 2s_2 - 5s_3 - 12s_4 - 5s_5 - 2s_6 + 4s_7 + s_8 + 4s_9)\Big\} \\
&- \tfrac{1}{32}\bar{J}(p_1,p_3)\Big\{(p_1\cdot p_3)(\hat{s}_{46} + \hat{s}_{135} - 5M_\pi^2)\Big\} \\
&+ \tfrac{1}{256}\bar{J}(p_1,p_4)\Big\{2M_\pi^2(2s_2 - s_7 - s_8) + s_8(2s_5 + 4s_1 + 4s_3 - s_8 - 4s_9) \\
&\qquad\qquad - 4s_2(s_1 + s_3 + s_4 + s_6 - 2s_9) \\
&\qquad\qquad + s_7(4s_1 + 2s_2 + 4s_3 - s_7 - 2s_8 - 4s_9)\Big\} \\
&+ \tfrac{L+\kappa}{384}\Big\{20M_\pi^4 + M_\pi^2(26s_7 - 68s_1) + s_1(32s_4 + 54s_5 + 40s_6 + 9s_1 - 47s_7) \\
&\qquad\qquad + s_9(27s_8 - 47s_4 - 49s_5 + 15s_9)\Big\} \\
&- \tfrac{\kappa}{1152}\Big\{10M_\pi^4 - 2M_\pi^2(23s_1 - 8s_9) + s_1(16s_4 + 18s_5 + 8s_6 + 3s_1 - 16s_7) \\
&\qquad\qquad - 2s_7(8s_2 + s_3 - 3s_7)\Big\}
\end{aligned}
\tag{III.4.2b}
$$

$$
\begin{aligned}
F_\pi^6 \tilde{\mathcal{A}}_{\{6\}}^{(\mathrm{NLO},\zeta)} =\; &-\tfrac{n}{48}C_3(p_1,\ldots,p_6) + \tfrac{n}{96}\overline{C}_{21}(p_1,\ldots,p_6)\big\{2M_\pi^2 + s_1 + s_3 - 4s_5 - 2s_9\big\} \\
&- \tfrac{ns_1 s_3}{64}C_{11}(p_1,\ldots,p_6) \\
&+ \tfrac{n}{384}C(p_1,\ldots,p_6)\big\{3s_1^2 s_3 + 3s_3 s_5(2M_\pi^2 + s_5) - 2s_1 s_5(2s_3 + 3s_7)\big\} \\
&- \tfrac{n}{384}\bar{J}(p_1,p_2)\Big\{20M_\pi^4 - M_\pi^2\big[5s_1 + 8s_2 + 2s_3 + 2s_5 + 8s_6 - 4(s_7 + s_8 + s_9)\big] \\
&\qquad\qquad + s_1(9s_1 + 2s_2 - s_3 - s_5 + 2s_6 - 4s_7 - s_8 - 4s_9)\Big\} \\
&+ \tfrac{n(3L+\kappa)}{1152}\Big\{2M_\pi^4 + M_\pi^2(8s_9 - 23s_1) + s_1(9s_1 + 20s_4 + 18s_5 + 4s_6 - 20s_7) \\
&\qquad\qquad - s_7(20s_5 + s_6 - 6s_7)\Big\} \\
&- \tfrac{n\kappa}{384}\Big\{2M_\pi^4 - s_1(5s_1 + 8s_4 + 6s_5 - 8s_7 - 8s_9) - 2s_7^2\Big\}
\end{aligned}
\tag{III.4.2c}
$$

$$
\begin{aligned}
F_\pi^6 \tilde{\mathcal{A}}_{\{2,4\}}^{(\mathrm{NLO},1)} =\; &\tfrac{M_\pi^4}{8n^2}\big\{L + \kappa - 2M_\pi^2 C(p_1,\ldots,p_6) - \bar{J}(p_1,p_2)\big\} \\
&+ L_1^{\mathrm{r}}\big\{4M_\pi^4 + t_1 t_6 + t_4 t_8 - 2M_\pi^2(t_1 + 2t_8)\big\} + \tfrac{1}{2}L_4^{\mathrm{r}}\big\{2M_\pi^2(M_\pi^2 - t_1 + t_8)\big\} \\
&+ \tfrac{1}{16}L_2^{\mathrm{r}}\big\{16M_\pi^4 - 8M_\pi^2(2t_1 + t_5) + 4t_1 t_6 + 4t_4 t_8 - t_9^2\big\} - 2M_\pi^4 L_6^{\mathrm{r}}
\end{aligned}
\tag{III.4.2d}
$$

$$
\begin{aligned}
F_\pi^6 \tilde{\mathcal{A}}_{\{2,4\}}^{(\mathrm{NLO},\zeta)} =\; &\tfrac{1}{32}\overline{C}_{21}(p_3,\ldots,p_6,p_1,p_2)\big\{2M_\pi^2 - t_1 - t_4\big\} \\
&- \tfrac{1}{4}C_{11}(p_1,p_3,p_2,p_4,p_5,p_6)\big\{(p_1\cdot p_3)(p_2\cdot p_4)\big\} \\
&+ \tfrac{1}{512}C(p_1,\ldots,p_6)\Big\{32M_\pi^6 - 32M_\pi^4(t_1 + t_2 + t_5) \\
&\qquad\qquad - t_1^2(t_3 + t_4 + 2t_5) - t_2\big[t_2 t_4 + 2(t_2 + t_4)t_5\big] - 2t_4 t_5 t_7 - 2t_4 t_7^2 \\
&\qquad\qquad + 2M_\pi^2\big[2t_1^2 + 4t_1 t_2 + 2t_2^2 + 3t_2(t_4 + 4t_5) \\
&\qquad\qquad\qquad + 3t_1(t_3 + 2t_4 + 4t_5) + 2t_7(t_5 + t_7)\big] \\
&\qquad\qquad - t_1\big[t_4^2 + 4t_4 t_5 + 2t_2(t_3 + t_4 + 2t_5) + 2t_5(t_3 + t_5 + t_7)\big]\Big\} \\
&+ \tfrac{1}{8}C(p_1,p_6,p_2,p_3,p_4,p_5)\big\{(p_1\cdot p_6)(p_2\cdot p_3)(2M_\pi^2 + s_2 - 2s_8 - 2s_4 + s_6)\big\} \\
&+ \tfrac{1}{128}\bar{J}(p_1,p_2)\big\{M_\pi^2(6t_1 - 2t_5) + t_1(t_6 - 2t_2 - t_3) + t_4(t_8 - t_4) - 2M_\pi^4\big\}
\end{aligned}
$$

$$- \frac{1}{16}\bar{J}(p_1,p_3)\Big\{(p_1\cdot p_3)(\hat{s}_{46} + \hat{s}_{246} - 5M_\pi^2)\Big\}$$
$$+ \frac{L+\kappa}{256}\Big\{4M_\pi^2(t_1+3t_5) + 2t_1(t_1+2t_2+t_3-3t_6) - 6t_4t_8 + t_9^2 - 12M_\pi^4\Big\}$$

(III.4.2e)

$$F_\pi^6\tilde{\mathcal{A}}_{\{3,3\}}^{(\mathrm{NLO},1)} = \frac{M_\pi^4}{6n^2}\Big\{L+\kappa - \bar{J}(p_1,p_2)\Big\}$$
$$- \frac{1}{16n}(4L_0^{\mathrm{r}}+L_3^{\mathrm{r}})\Big\{4M_\pi^2 u_1 + 2u_1(u_4+u_5) - u_9(4u_8+3u_9)\Big\}$$
$$- \frac{1}{2n}L_3^{\mathrm{r}}\Big\{u_1(u_4+u_5) - u_9(u_8+u_9)\Big\} + \frac{1}{6n}L_5^{\mathrm{r}}\Big\{M_\pi^2(3u_9-8M_\pi^2)\Big\}$$
$$+ \frac{4M_\pi^4}{3}\Big\{\frac{1}{n}L_8^r - \frac{1}{3}L_7^r\Big\}$$

(III.4.2f)

$$F_\pi^6\tilde{\mathcal{A}}_{\{3,3\}}^{(\mathrm{NLO},\xi)} = \frac{M_\pi^4}{24n}\bar{J}(p_1,p_2) - \frac{M_\pi^4\kappa}{144n}$$
$$- \frac{6L+5\kappa}{2304n}\Big\{16M_\pi^4 - 12M_\pi^2 u_1 + 2u_1(u_4+u_5) + u_9(4u_8+u_9)\Big\}$$

(III.4.2g)

$$F_\pi^6\tilde{\mathcal{A}}_{\{3,3\}}^{(\mathrm{NLO},\xi^2)} = \frac{u_8 - 2M_\pi^2 u_8}{32}\overline{C}_{21}(p_2,\ldots,p_5,p_1,p_6)$$
$$+ \frac{1}{384}C(p_1,p_6,p_2,p_5,p_3,p_4)\Big\{8M_\pi^6 - 12M_\pi^4 u_8 + 6M_\pi^2 u_4 u_8 - u_3 u_4 u_8\Big\}$$
$$- \frac{1}{384}C(p_1,p_4,p_2,p_5,p_3,p_6)\Big\{8M_\pi^6 - 12M_\pi^4 u_7 + 6M_\pi^2 u_4 u_7 - u_1 u_4 u_7\Big\}$$
$$- \frac{1}{128}\bar{J}(p_1,p_4)\Big\{(p_1\cdot p_4)(u_4+u_7-2u_9)\Big\} - \frac{L+\kappa}{128}\Big\{u_1 u_5 - u_6 u_9\Big\}$$

(III.4.2h)

$$F_\pi^6\tilde{\mathcal{A}}_{\{3,3\}}^{(\mathrm{NLO},\zeta)} = -\frac{p_1\cdot p_6}{8}\overline{C}_{21}(p_2,\ldots,p_5,p_1,p_6)$$
$$+ \frac{1}{256}C(p_1,\ldots,p_6)\Big\{2(p_3\cdot p_4)\big[64M_\pi^4 - 2(u_1+u_2)^2 + u_3^2 + 2u_4^2$$
$$+ 2u_1 u_5 + 4u_4 u_6 + u_6^2$$
$$- 2u_3(2u_4+u_6) + 2u_7(2u_1+u_2)$$
$$+ 2u_4 u_8 + 16M_\pi^2(u_3 - 2u_4 - u_6)\big]$$
$$+ u_9\big[32M_\pi^4 - 4M_\pi^2(3u_3+2u_4+u_6)$$
$$+ 2u_3(u_6-u_3+2u_8) + u_9(u_3-2M_\pi^2)\big]\Big\}$$
$$- \frac{1}{96}C(p_1,p_6,p_2,p_5,p_3,p_4)\Big\{8M_\pi^6 - 12M_\pi^4 u_8 + 6M_\pi^2 u_4 u_8 - u_3 u_4 u_8\Big\}$$
$$+ \frac{1}{384}\bar{J}(p_1,p_2)\Big\{32M_\pi^4 - 36M_\pi^2(u_1+u_2-u_3)$$
$$+ 3\big[u_1^2 + u_3^2 + 2u_3 u_5 - 2u_2(u_3-u_4+u_5+u_7)$$
$$+ 2u_8 u_9 + u_9^2 + 2u_1(u_4-u_3-u_5+u_6-u_7+u_9)\big]\Big\}$$
$$+ \frac{1}{64}\bar{J}(p_1,p_4)\Big\{(p_1\cdot p_4)(u_4+u_7-2u_9)\Big\}$$
$$+ \frac{3L+2\kappa}{2304}\Big\{32M_\pi^4 - 12M_\pi^2 u_1 + 2u_1(u_4+13u_5) - u_9(20u_8+17u_9)\Big\}$$
$$+ \frac{\kappa}{576}\Big\{8M_\pi^4 - 6M_\pi^2 u_1 + u_1(u_4+7u_5) - 4u_9(u_8+u_9)\Big\}$$

(III.4.2i)

$$F_\pi^6\tilde{\mathcal{A}}_{\{2,2,2\}}^{(\mathrm{NLO},1)} = \frac{M_\pi^6}{6n^3}C(p_1,\ldots,p_6)$$

(III.4.2j)

Here, $\bar{J}(p_i,p_j) \equiv \bar{J}\big((p_i+p_j)^2\big)$. A few features can be observed:

- The NLO LECs only appear in $\mathcal{A}^{(\mathrm{NLO},1)}$.

- $\mathcal{A}_R^{(\mathrm{NLO},\xi^2)}$ only exists for $R = \{3,3\}$.

- $\mathcal{A}_R^{(\mathrm{NLO},\zeta)}$ and $\mathcal{A}_R^{(\mathrm{NLO},\xi^2)}$ are independent of n for $R \neq \{6\}$, while $\mathcal{A}_{\{6\}}^{(\mathrm{NLO},\zeta)}$ is proportional to n and is the only place where positive powers of n appear.

- $\mathcal{A}_{\{2,2,2\}}^{(\mathrm{NLO},1)}$ is the only $R = \{2,2,2\}$ subamplitude, is proportional to n^{-3}, and is the only place where this power appears.

These features and their generalizations are derived in appendix E.

E   Symmetries and group-dependent features of the amplitude

In this appendix, we derive the features described in the previous section using the technique we here dub *diagrammatic flavor-ordering*, wherein modified Feynman diagrams allow direct calculation of stripped amplitudes without going through the full amplitude. Simpler cases of the technique have been used for a long time [21, 22, 35], but the extension beyond LO and $R = \{k\}$ is more recent [20, 37]. A somewhat similar approach can be found in ref. [33]. In the preparation of this paper, we refined the technique and performed the first loop calculations using it, but it turned out that the proliferation of diagrams caused by the inclusion of loops and nonzero masses — nearly 200 distinct topologies compared to 9 without flavor-ordering — outweighed any efficiency advantages the technique had over standard Feynman diagrams, rendering it impractical for our purposes. Nevertheless, the manifest relation between kinematics and flavor structure in flavor-ordered diagrams can be used to illuminate some features that are obscured with the standard approach.

E–1   Diagrammatic flavor-ordering

Here, we give a brief summary of this technique; see ref. [20] for a detailed version, and ref. [37] for one including loops.

By 'flavor-ordered', we mean a quantity whose flavor structure is $\mathcal{F}_R$ for some R, i.e. whose flavor indices are in natural order (up to $\mathbb{Z}_R$). Such a quantity is invariant under $\mathbb{Z}_R$ acting simultaneously on its flavor indices and momenta. The stripped amplitude is obtained by keeping only the flavor-ordered parts of the amplitude and then dropping the flavor structure. Diagrammatic flavor-ordering is based on the observation that the Fierz identity, eqs. (III.2.7) and (III.2.8), generally preserves flavor-ordering: if two sub-diagrams contain $\langle t^a A \rangle$ and $\langle t^a B \rangle$, joining them will result in $\langle AB \rangle$, $\langle A \rangle \langle B \rangle$ or $\langle AB^\dagger \rangle$, all of which keep the (possibly reversed) order of flavor indices in A and B. Therefore, a diagram is flavor-ordered only if its sub-diagrams, all the way down to the vertices, are flavor-ordered.

We now desire a set of modified diagram-drawing rules that make diagrams inherently flavor-ordered, with the flavor structures manifest from the graphical shape of the diagram.
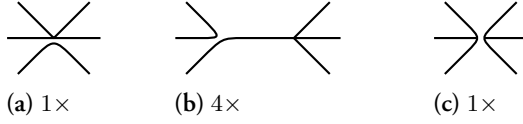
Figure III.6: Flavor-ordered variants of figs. III.4a and III.4b with split vertices. Note that the multiplicities given here refer to permutations in $\mathbb{Z}_R$. Diagrams (a) and (b) have flavor structure $\mathcal{F}_{\{2,4\}}$ and contain L_i, $i = 1, 2, 4, 6$. Diagram (c) has $\mathcal{F}_{\{3,3\}}$ and contains L_7. Flavor-ordered diagrams identical to figs. III.4a to III.4c have $\mathcal{F}_{\{6\}}$, contain L_i, $i = 0, 3, 5, 8$, and have multiplicity $1\times$, $6\times$ and $3\times$, respectively.

We think of each external leg as labeled by an index i, corresponding to momentum p_i and flavor index b_i. We will call two legs *flavor-connected* if their flavor indices reside in the same trace in the flavor structure.

For single-vertex diagrams, we indicate the flavor structure by adding gaps in the vertex between the groups of legs that are not flavor-connected, as is done in figs. III.6a and III.6c. We treat the vertices in multi-vertex diagrams like fig. III.6b similarly. Since the two terms on the right-hand side of eq. (III.2.7a) treat flavor structures differently, we represent them by different propagators as if there were two species of particles: ordinary (solid line) and singlets (dashed line), the latter of which carry a factor of $-\frac{1}{\zeta n}$.[19]

When combined with eq. (III.2.7b), these rules allow the flavor structure of any SU diagram to be read off, as illustrated in fig. III.7. Two legs are flavor-connected if and only if the following conditions hold:

- They are joined by an uninterrupted path through the diagram (vertex gaps and singlet propagators interrupt it).

- They can be joined by a line that does not intersect the diagram at any point (it can pass through vertex gaps and singlet propagators).

This is illustrated in fig. III.7b.

To read the indices, follow the outline of each flavor-connected set of legs, keeping the diagram to the right of the path (thus reading the indices of tree diagrams in clockwise order), as illustrated in fig. III.7c.[20] The starting point is arbitrary due to $\mathbb{Z}_R$ symmetry.

---

[19]The name "singlet" stems from how adding a singlet field ϕ^0, whose associated generator $t^0 = \frac{1}{\sqrt{\zeta n}}$ commutes with all t^a, results in the removal of the $1/n$ terms from the Fierz identity since e.g. $\langle t^0 A \rangle \langle t^0 B \rangle = \frac{1}{\zeta n} \langle A \rangle \langle B \rangle$. Thus, the $1/n$ terms can be interpreted as the subtraction of diagrams with internal singlet lines, allowing other contractions to be done using only the n-independent terms. The singlet decouples from the other fields in $\langle u_\mu u^\mu \rangle$, so LO singlet vertices stem from $\langle \chi_+ \rangle$ and therefore depend on the mass but not on the momenta [this is easiest to see in the exponential parametrization (III.2.6)]. This simplifies LO and NLO singlet diagrams and causes them to vanish in the massless limit.

[20]These rules become more complicated at two-loop level and above, where non-planar diagrams may appear. However, all diagrams can be drawn without self-intersections on a surface of sufficiently high topological genus (planar diagrams on a sphere, non-planar two-loop diagrams on a torus, etc.). One must then imagine the diagram drawn on such a surface (but not one of higher genus than necessary) when determining flavor-connectedness or assigning indices.
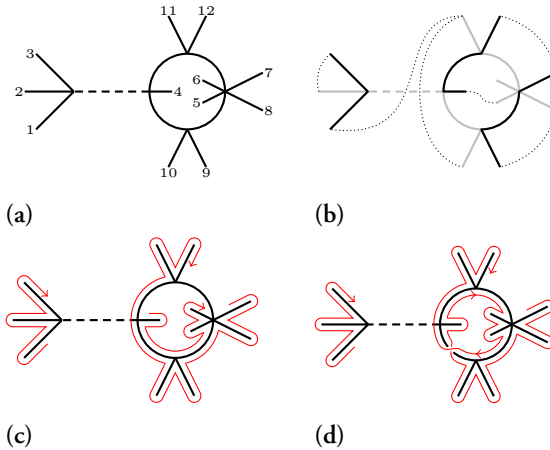
Figure III.7: An illustration of how indices are read from a flavor-ordered diagram. (a) The diagram with indices assigned (up to $\mathbb{Z}_{\{3,3,6\}}$). (b) Examples of paths connecting legs without intersecting the diagram (dotted) and paths within the diagram (solid, with the rest of the diagram grayed out). (c) Paths showing the order in which the indices are read to give the indexing in (a). (d) The alternative path giving the $\xi/2$ term under $S^{\mathrm{o}}_{\mathrm{p}}$. Note that this results in a different indexing than in (a) and flavor structure $\mathcal{F}_{\{3,9\}}$.

When a loop is 'empty', like those in fig. III.8, a factor of $\langle \mathbb{1} \rangle = \zeta n$ is added.

For the purposes of momentum flow, flavor-ordered diagrams are treated just like ordinary diagrams, and the two kinds of propagators are kinematically identical. However, flavor-ordered diagrams are typically sensitive to the order in which legs are arranged around a vertex (see e.g. fig. III.9). This, along with the combination of singlet and ordinary propagators, leads to the proliferation mentioned earlier. All diagrams must be summed over $\mathbb{Z}_R$ (with appropriate symmetry factors) and added up to obtain A_R.

The above rules hold for SU, and to a large extent also for $S^{\mathrm{o}}_{\mathrm{p}}$. In fact, the cases where they are equivalent (up to the substitution $n \to \zeta n$) exactly correspond to $\mathcal{A}^{(1)}$ of eq. (III.3.8).

E–2 Differences between the groups

We will now discuss all contexts in which differences between SU and $S^{\mathrm{o}}_{\mathrm{p}}$ may arise. The following fully accounts for the patterns seen in the six-meson amplitude:

Tree diagrams. View a $S^{\mathrm{o}}_{\mathrm{p}}$ diagram as being built by adding vertices one by one. With A belonging to the partially completed diagram and B to the vertex, $\langle t^a A \rangle \langle t^a B \rangle \to \frac{1}{2}[\langle AB \rangle + \langle AB^\dagger \rangle]$ gives one flavor-ordered term and one that is discarded. Adding a structurally identical diagram but with some indices permuted so that B is reversed gives $\langle t^a A \rangle \langle t^a B^\dagger \rangle \to \frac{1}{2}[\langle AB^\dagger \rangle + \langle AB \rangle]$: Again, one term is kept and one discarded. However, $\langle t^a B \rangle = \langle t^a B^\dagger \rangle$

(a) 18×   **(b)** 18×   **(c)** 9×
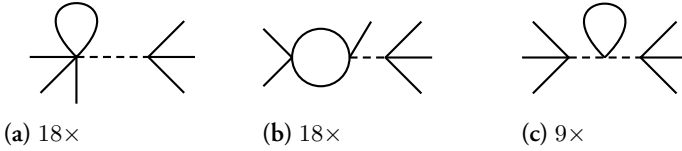
Figure III.8: A few of the flavor-ordered diagrams that contribute to $\mathcal{A}^{(\xi)}_{\{3,3\}}$. The multiplicities refer to permutations in $\mathbb{Z}_{\{3,3\}}$. All diagrams contain a single factor $1/n$ from singlets; in (c) the $1/n$ from the second singlet propagator is canceled by $\langle AB^\dagger \rangle = \langle \mathbb{1} \rangle = n$ (such a factor appears whenever a loop is not flavor-connected to any external leg). In their contributions to $\mathcal{A}^{(\zeta)}$, all n-dependence is canceled by $\langle A \rangle \langle B \rangle = \langle A \rangle \langle \mathbb{1} \rangle$ for (a,b) and $\langle A \rangle \langle B \rangle = \langle \mathbb{1} \rangle \langle \mathbb{1} \rangle$ for (c).

under $S^{\mathrm{o}}_{\mathrm{P}}$, so the kinematic structure of the vertex must be invariant under that index permutation. Thus, the two flavor-ordered terms are identical and add up to the same $\langle AB \rangle$ given by SU. This proves, to all orders in the chiral counting, that eqs. (III.2.7a) and (III.2.8a) fail to introduce any differences between SU and $S^{\mathrm{o}}_{\mathrm{P}}$. In other words, SU *and $S^{\mathrm{o}}_{\mathrm{P}}$ are equivalent at tree level (up to $n \to \zeta n$)*, so all tree diagrams go into $\mathcal{A}^{(1)}$. The only caveat is if any differences are introduced at the Lagrangian level, but this happens first at NNLO (see below).

Loops. Viewing the loop as being formed by joining two legs of a tree diagram, we see that eqs. (III.2.7b) and (III.2.8b) must be applied if those legs are part of the same trace. The term $\langle t^a A t^a B \rangle \to \langle A \rangle \langle B \rangle$ is the same in SU and $S^{\mathrm{o}}_{\mathrm{P}}$, up to a factor $\frac{1}{\zeta}$, giving rise to $\mathcal{A}^{(\zeta)}$. The term $\langle t^a A t^a B \rangle \to \langle AB^\dagger \rangle$ is unique to $S^{\mathrm{o}}_{\mathrm{P}}$ and contains $\pm$, giving rise to $\mathcal{A}^{(\xi)}$.[21] Since this term gives a single trace, it almost always results in $R = \{6\}$, which explains why $\mathcal{A}^{(\xi)}_{\{6\}}$ is the longest subamplitude. A few singlet diagrams, including those in fig. III.8, give $\mathcal{A}^{(\xi)}_{\{3,3\}}$ instead.

When $B = \mathbb{1}$ in the $\mathcal{A}^{(\zeta)}$ case, we get a factor of $n = \langle \mathbb{1} \rangle$; this corresponds graphically to an 'empty' loop as mentioned above. This is the only source of positive powers of n, and explains why they only appear in $\mathcal{A}^{(\zeta)}$. Those diagrams still contribute to $\mathcal{A}^{(\xi)}$ without a factor of n.

Singlets. The $\frac{1}{\zeta n}$ terms of the Fierz identity are the same for SU and $S^{\mathrm{o}}_{\mathrm{P}}$, so singlet propagators behave the same in both cases. When a singlet is part of a loop, one can let the singlet propagator 'close' the loop, thereby avoiding all differences stemming from eqs. (III.2.7b)

---

[21]This necessitates the extra rule illustrated in fig. III.7d. In detail, it is as follows: Starting from the regular index-assigning path [fig. III.7c], create arbitrarily located 'gaps' in loop propagators until all legs can be joined by lines that do not intersect the diagram. Following the path, split the diagram into two terms every time a gap is reached. In one term, continue on the original path and multiply by $1/\zeta$. In the other, go through the gap and multiply by $\xi/2$. Do this recursively if you come across another gap, but the second time a given gap is reached, go through it unconditionally. Every time you go through a gap, swap the orientation of your path (clockwise to counterclockwise, and vice versa).

(a) 9×   **(b)** 9×   **(c)** 9×
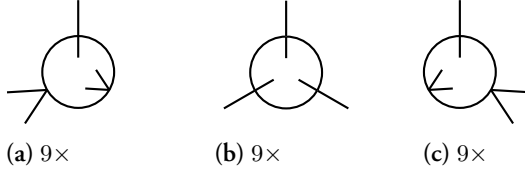
Figure III.9: A few of the flavor-ordered diagrams that contribute to $\mathcal{A}^{(\xi^2)}_{\{3,3\}}$. The multiplicities are for SU; S^0_P permits more permutations, which is exactly why they give $\mathcal{A}^{(\xi^2)}_{\{3,3\}}$.

and (III.2.8b). Therefore, such diagrams go into $\mathcal{A}^{(1)}$. This does not apply when a singlet is outside a loop, but for our amplitude this only happens with $R = \{3,3\}$ diagrams like in fig. III.8 and variations thereof. The 'empty' loop cancels the n-dependence in their contributions to $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(\zeta)}$. Therefore, negative powers of n, which only arise from singlets, only show up in $\mathcal{A}^{(1)}$ and (due to these diagrams) $\mathcal{A}^{(\xi)}_{\{3,3\}}$.

Trace-reversal. The greatest differences come from the fact that $\langle t^a t^b \cdots t^c \rangle = \langle t^c t^b \cdots t^a \rangle$ under S^0_P but not SU. *Therefore, A_R is invariant under reversal of individual traces under S^0_P, but only under simultaneous reversal of all traces under* SU. Among the cases considered here, these types of reversal are only inequivalent when $R = \{3,3\}$.[22] For instance, in the diagrams in fig. III.9, the "inside" of the loop can be read both clockwise and counterclockwise under S^0_P, but only one way under SU, and the momentum dependence will be correspondingly different. Such disorganized difference is what gives rise to $\mathcal{A}^{(\xi^2)}_{\{3,3\}}$. In all $R = \{3,3\}$ diagrams involving singlets, at least one trace can be reversed as a symmetry of the diagram (i.e. reversing a single-trace vertex), so they do not contribute to $\mathcal{A}^{(\xi^2)}_{\{3,3\}}$.

E–3 More particles, higher orders

The patterns discussed here are straightforward to generalize. At $N^\ell LO$ (i.e. ℓ loops), eq. (III.3.8) becomes

$$A = \left\{ \mathcal{A}^{(1)} + \xi \mathcal{A}^{(\xi)} + \xi^2 \mathcal{A}^{(\xi^2)} + \sum_{j=1}^{\ell} \frac{\mathcal{A}^{(\zeta^j)}}{\zeta^j} \right\}_{n \to \zeta n}, \qquad \text{(III.5.1)}$$

since each loop can give another factor of $\frac{1}{\zeta}$. Most of the features discussed above remain, although some are softened: Negative powers of n can appear in $\mathcal{A}^{(\zeta^j)}$ and $\mathcal{A}^{(\xi^2)}$, and positive powers of n in most subamplitudes except $\mathcal{A}^{(1)}$. $\mathcal{A}^{(\xi^2)}_R$ exists for $R = \{2,3,3\}$, $\{4,4\}$, etc. In an $N^\ell LO$ amplitude, A_R with $|R| > \ell + 1$ requires singlets breaking loops, which severely restricts the structure of that subamplitude; specifically, if $R = \{2, \ldots, 2\}$, it will be similarly simple to our $\mathcal{A}^{(1)}_{\{2,2,2\}}$.

As mentioned above, the equivalence of SU and S^0_P at tree level can only be broken by Lagrangian effects. The first such effect is in the NNLO Lagrangian $\mathcal{L}^{(6)}$ [46], where the

---

[22]In general, R must contain at least two elements greater than 2, since $\langle t^a t^b \rangle = \langle t^b t^a \rangle$ under all groups.

59th term $\langle u_\mu u_\nu u_\rho \rangle \langle u^\mu u^\nu u^\rho \rangle$ and the 61st term $\langle u_\mu u_\nu u_\rho \rangle \langle u^\rho u^\nu u^\mu \rangle$ are distinct under SU but equal under $S_\mathbb{P}^0$. (The N³LO Lagrangian $\mathcal{L}^{(8)}$ [47] contains several such cases.) This only results in additional relations between the LECs; the functional form of A is retained, and eq. (III.5.1) remains valid, albeit a bit more redundant.

References

[1] Weinberg, S. "Phenomenological Lagrangians". *Physica A* **96**, 327–340 (1979).

[2] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory to One Loop". *Annals Phys.* **158**, 142 (1984).

[3] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark". *Nucl. Phys. B* **250**, 465–516 (1985).

[4] Scherer, S. & Schindler, M. R. *A Primer for Chiral Perturbation Theory* 1st ed. ISBN: 978-3-642-19253-1 (Springer-Verlag, 2012).

[5] Pich, A. *Effective Field Theory with Nambu-Goldstone Modes* in *Les Houches summer school: EFT in Particle Physics and Cosmology* (2018). arXiv: 1804.05664 [hep-ph].

[6] Hansen, M. T. & Sharpe, S. R. "Relativistic, model-independent, three-particle quantization condition". *Phys. Rev. D* **90**, 116003. arXiv: 1408.5933 [hep-lat] (2014).

[7] Hansen, M. T. & Sharpe, S. R. "Expressing the three-particle finite-volume spectrum in terms of the three-to-three scattering amplitude". *Phys. Rev. D* **92**, 114509. arXiv: 1504.04248 [hep-lat] (2015).

[8] Hansen, M. T. & Sharpe, S. R. "Lattice QCD and Three-particle Decays of Resonances". *Ann. Rev. Nucl. Part. Sci.* **69**, 65–107. arXiv: 1901.00483 [hep-lat] (2019).

[9] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three-particle quantization condition in a finite volume: 1. The role of the three-particle force". *JHEP* **09**, 109. arXiv: 1706.07700 [hep-lat] (2017).

[10] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three particle quantization condition in a finite volume: 2. general formalism and the analysis of data". *JHEP* **10**, 115. arXiv: 1707.02176 [hep-lat] (2017).

[11] Mai, M. & Döring, M. "Three-body Unitarity in the Finite Volume". *Eur. Phys. J. A* **53**, 240. arXiv: 1709.08222 [hep-lat] (2017).

[12] Mai, M. & Döring, M. "Finite-Volume Spectrum of $\pi^+\pi^+$ and $\pi^+\pi^+\pi^+$ Systems". *Phys. Rev. Lett.* **122**, 062503. arXiv: 1807.04746 [hep-lat] (2019).

[13] Mai, M., Döring, M., Culver, C. & Alexandru, A. "Three-body unitarity versus finite-volume $\pi^+\pi^+\pi^+$ spectrum from lattice QCD". *Phys. Rev. D* **101**, 054510. arXiv: 1909.05749 [hep-lat] (2020).

[14] Culver, C., Mai, M., Brett, R., Alexandru, A. & Döring, M. "Three pion spectrum in the $I = 3$ channel from lattice QCD". *Phys. Rev. D* **101**, 114507. arXiv: 1911.09047 [hep-lat] (2020).

[15] Mai, M., Döring, M. & Rusetsky, A. "Multi-particle systems on the lattice and chiral extrapolations: a brief review". *Eur. Phys. J. ST* **230**, 1623–1643. arXiv: 2103.00577 [hep-lat] (2021).

[16] Brett, R. *et al.* "Three-body interactions from the finite-volume QCD spectrum". *Phys. Rev. D* **104**, 014501. arXiv: 2101.06144 [hep-lat] (2021).

[17] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "$I = 3$ Three-Pion Scattering Amplitude from Lattice QCD". *Phys. Rev. Lett.* **124**, 032001. arXiv: 1909.02973 [hep-lat] (2020).

[18] Blanton, T. D. *et al.* "Interactions of two and three mesons including higher partial waves from lattice QCD". *JHEP* **10**, 023. arXiv: 2106.05590 [hep-lat] (2021).

[19] Bijnens, J. & Husek, T. "Six-pion amplitude". *Phys. Rev. D* **104**, 054046. arXiv: 2107.06291 [hep-ph] (2021).

[20] Bijnens, J., Kampf, K. & Sjö, M. "Higher-order tree-level amplitudes in the nonlinear sigma model". *JHEP* **11**. [Erratum: *JHEP* **03** (2021) 066], 074. arXiv: 1909.13684 [hep-th] (2019).

[21] Osborn, H. "Implications of Adler zeros for multipion processes". *Lett. Nuovo Cim.* **2**, 717–723 (1969).

[22] Susskind, L. & Frye, G. "Algebraic aspects of pionic duality diagrams". *Phys. Rev. D* **1**, 1682–1686 (1970).

[23] Peskin, M. E. "The Alignment of the Vacuum in Theories of Technicolor". *Nucl. Phys. B* **175**, 197–233 (1980).

[24] Preskill, J. "Subgroup Alignment in Hypercolor Theories". *Nucl. Phys. B* **177**, 21–59 (1981).

[25] Dimopoulos, S. "Technicolored Signatures". *Nucl. Phys. B* **168** (ed Zichichi, A.) 69–92 (1980).

[26] Kogut, J. B., Stephanov, M. A., Toublan, D., Verbaarschot, J. J. M. & Zhitnitsky, A. "QCD-like theories at finite baryon density". *Nucl. Phys. B* **582**, 477–513. arXiv: hep-ph/0001171 (2000).

[27] Gasser, J. & Leutwyler, H. "Light Quarks at Low Temperatures". *Phys. Lett. B* **184**, 83–88 (1987).

[28] Chivukula, R. S., Dugan, M. J. & Golden, M. "Analyticity, crossing symmetry and the limits of chiral perturbation theory". *Phys. Rev. D* **47**, 2930–2939. arXiv: hep-ph/9206222 (1993).

[29] Splittorff, K., Toublan, D. & Verbaarschot, J. J. M. "Diquark condensate in QCD with two colors at next-to-leading order". *Nucl. Phys. B* **620**, 290–314. arXiv: hep-ph/0108040 (2002).

[30] Bijnens, J. & Lu, J. "Technicolor and other QCD-like theories at next-to-next-to-leading order". *JHEP* **11**, 116. arXiv: 0910.5424 [hep-ph] (2009).

[31] Bijnens, J. & Lu, J. "Meson-meson Scattering in QCD-like Theories". *JHEP* **03**, 028. arXiv: 1102.0172 [hep-ph] (2011).

[32] Bijnens, J. & Lu, J. "Two-Point Functions and S-Parameter in QCD-like Theories". *JHEP* **01,** 081. arXiv: 1111.1886 [hep-ph] (2012).

[33] Low, I. & Yin, Z. "Soft Bootstrap and Effective Field Theories". *JHEP* **11,** 078. arXiv: 1904.12859 [hep-th] (2019).

[34] Low, I., Rodina, L. & Yin, Z. "Double Copy in Higher Derivative Operators of Nambu-Goldstone Bosons". *Phys. Rev. D* **103,** 025004. arXiv: 2009.00008 [hep-th] (2021).

[35] Kampf, K., Novotný, J. & Trnka, J. "Tree-level Amplitudes in the Nonlinear Sigma Model". *JHEP* **05,** 032. arXiv: 1304.3048 [hep-th] (2013).

[36] Kampf, K., Novotný, J., Shifman, M. & Trnka, J. "New Soft Theorems for Goldstone Boson Amplitudes". *Phys. Rev. Lett.* **124,** 111601. arXiv: 1910.04766 [hep-th] (2020).

[37] Sjö, M. *Flavour-ordering in the nonlinear sigma model with more derivatives and legs* M.Sc. thesis (Lund University, 2019). http://lup.lub.lu.se/student-papers/record/8984750.

[38] Farrow, J. A., Geyer, Y., Lipstein, A. E., Monteiro, R. & Stark-Muchão, R. "Propagators, BCFW recursion and new scattering equations at one loop". *JHEP* **10,** 074. arXiv: 2007.00623 [hep-th] (2020).

[39] Bartsch, C., Kampf, K. & Trnka, J. "Recursion Relations for One-Loop Goldstone Boson Amplitudes". arXiv: 2206.04694 [hep-th] (2022).

[40] Mertig, R., Bohm, M. & Denner, A. "FEYN CALC: Computer algebraic calculation of Feynman amplitudes". *Comput. Phys. Commun.* **64,** 345–359 (1991).

[41] Shtabovenko, V., Mertig, R. & Orellana, F. "New Developments in FeynCalc 9.0". *Comput. Phys. Commun.* **207,** 432–444. arXiv: 1601.01167 [hep-ph] (2016).

[42] Shtabovenko, V., Mertig, R. & Orellana, F. "FeynCalc 9.3: New features and improvements". *Comput. Phys. Commun.* **256,** 107478. arXiv: 2001.04407 [hep-ph] (2020).

[43] Vermaseren, J. A. M. "New features of FORM". arXiv: math-ph/0010025 (2000).

[44] Van Oldenborgh, G. J. & Vermaseren, J. A. M. "New Algorithms for One Loop Integrals". *Z. Phys. C* **46,** 425–438 (1990).

[45] Hahn, T. & Pérez-Victoria, M. "Automatized one-loop calculations in four and D dimensions". *Comput. Phys. Commun.* **118,** 153–165. arXiv: hep-ph/9807565 (1999).

[46] Bijnens, J., Colangelo, G. & Ecker, G. "The Mesonic chiral Lagrangian of order p^6". *JHEP* **02,** 020. arXiv: hep-ph/9902437 (1999).

[47] Bijnens, J., Hermansson-Truedsson, N. & Wang, S. "The order p^8 mesonic chiral Lagrangian". *JHEP* **01,** 102. arXiv: 1810.06834 [hep-ph] (2019).

[48] Zhang, H.-H., Yan, W.-B. & Li, X.-S. "Trace Formulae of Characteristic Polynomial and Cayley–Hamilton's Theorem, and Applications to Chiral Perturbation Theory and General Relativity". *Commun. Theor. Phys.* **49,** 801. arXiv: hep-th/0701116 (2008).

[49] Britto, R., Cachazo, F. & Feng, B. "New recursion relations for tree amplitudes of gluons". *Nucl. Phys. B* **715**, 499–522. arXiv: `hep-th/0412308` (2005).

[50] Britto, R., Cachazo, F., Feng, B. & Witten, E. "Direct proof of tree-level recursion relation in Yang-Mills theory". *Phys. Rev. Lett.* **94**, 181602. arXiv: `hep-th/0501052` (2005).

[51] Adler, S. L. "Consistency conditions on the strong interactions implied by a partially conserved axial vector current". *Phys. Rev.* **137**, B1022–B1033 (1965).

[52] Adler, S. L. "Consistency conditions on the strong interactions implied by a partially conserved axial-vector current. II". *Phys. Rev.* **139**, B1638–B1643 (1965).

[53] Romero-López, F., Rusetsky, A., Schlage, N. & Urbach, C. "Relativistic N-particle energy shift in finite volume". *JHEP* **02**, 060. arXiv: `2010.11715 [hep-lat]` (2021).

[54] Blanton, T. D. & Sharpe, S. R. "Equivalence of relativistic three-particle quantization conditions". *Phys. Rev. D* **102**, 054515. arXiv: `2007.16190 [hep-lat]` (2020).

[55] Bijnens, J. & Ecker, G. "Mesonic low-energy constants". *Ann. Rev. Nucl. Part. Sci.* **64**, 149–174. arXiv: `1405.6488 [hep-ph]` (2014).

[56] Kleiss, R., Stirling, W. J. & Ellis, S. D. "A New Monte Carlo Treatment of Multiparticle Phase Space at High-energies". *Comput. Phys. Commun.* **40**, 359 (1986).

[57] Bijnens, J., Colangelo, G., Ecker, G., Gasser, J. & Sainio, M. E. "Elastic $\pi\pi$ scattering to two loops". *Phys. Lett. B* **374**, 210–216. arXiv: `hep-ph/9511397` (1996).

[58] Bijnens, J., Colangelo, G., Ecker, G., Gasser, J. & Sainio, M. E. "Pion-pion scattering at low energy". *Nucl. Phys. B* **508**. [Erratum: Nucl.Phys.B **517**, 639–639 (1998)], 263–310. arXiv: `hep-ph/9707291` (1997).

[59] Dobado, A. & Morales, J. "Pion mass effects in the large N limit of chiral perturbation theory". *Phys. Rev. D* **52**, 2878–2890. arXiv: `hep-ph/9407321` (1995).

[60] Bijnens, J. & Carloni, L. "Leading Logarithms in the Massive O(N) Nonlinear Sigma Model". *Nucl. Phys. B* **827**, 237–255. arXiv: `0909.5086 [hep-ph]` (2010).

[61] Bijnens, J. & Carloni, L. "The Massive O(N) Non-linear Sigma Model at High Orders". *Nucl. Phys. B* **843**, 55–83. arXiv: `1008.3499 [hep-ph]` (2011).

[62] Bijnens, J., Kampf, K. & Lanz, S. "Leading logarithms in N-flavour mesonic Chiral Perturbation Theory". *Nucl. Phys. B* **873**, 137–164. arXiv: `1303.3125 [hep-ph]` (2013).

[63] Cronin, J. A. "Phenomenological model of strong and weak interactions in chiral U(3) × U(3)". *Phys. Rev.* **161**, 1483–1494 (1967).

[64] Ellis, J. R. & Renner, B. "On the relationship between chiral and dual models". *Nucl. Phys. B* **21**, 205–216 (1970).

Paper IV

IV

The isospin-3 three-particle K-matrix at NLO in ChPT

Jorge Baeza-Ballesteros,[1] Johan Bijnens,[2] Tomáš Husek,[2,3]
Fernando Romero-López,[4] Stephen R. Sharpe[5] and Mattias Sjö[2]

[1]IFIC, CSIC-Universitat de València, 46980 Paterna, Spain
[2]Department of Physics, Lund University, Box 118, SE 22100 Lund, Sweden
[3]Institute of Particle and Nuclear Physics, Charles University,
   V Holešovičkách 2, 180 00 Prague, Czech Republic
[4]CTP, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[5]Physics Department, University of Washington, Seattle, WA 98195-1560, USA

ABSTRACT: The three-particle K-matrix, $\mathcal{K}_{\mathrm{df},3}$, is a scheme-dependent quantity that parametrizes short-range three-particle interactions in the relativistic-field-theory three-particle finite-volume formalism. In this work, we compute its value for systems of three pions at maximal isospin through next-to-leading order (NLO) in Chiral Perturbation Theory (ChPT). We compare the values to existing lattice QCD results and find that the agreement between lattice QCD data and ChPT in the first two coefficients of the threshold expansion of $\mathcal{K}_{\mathrm{df},3}$ is significantly improved with respect to leading order once NLO effects are incorporated.

1 Introduction

Lattice QCD provides a systematically improvable approach to calculate strongly-inter-acting processes, including several that are inaccessible in experiment. One example is the scattering of three particles, e.g., $3\pi^+ \to 3\pi^+$. Experimentally, such processes can be very difficult due to the challenge of creating and scattering three beams and because all hadrons (except the nucleons) are short-lived. By contrast, lattice QCD provides a method for obtaining multihadron scattering amplitudes and allows one to treat the strong force in isolation, thus rendering the lightest mesons (and some other hadrons) stable due to the absence of weak and electromagnetic interactions.

The extraction of scattering amplitudes from lattice QCD is a very active topic of research (see refs. [1–8] for recent reviews), and, in particular, three-particle processes have recently received a lot of attention. The formalism to extract these amplitudes from the three-particle finite-volume spectrum computed on the lattice has been developed over the last decade [9–39], using three main approaches, and has been applied to results from lattice simulations for a number of scattering amplitudes [40–53]. So far, the system that has been most extensively explored is that of three pions at maximal isospin, i.e., $3\pi^+ \to 3\pi^+$ scattering. Several of these works (refs. [44, 47, 48, 52]) use the relativistic-field-theory (RFT) three-particle finite-volume formalism, which parametrizes short-range three-body interactions via an intermediate cutoff-dependent quantity, the three-particle K-matrix, $\mathcal{K}_{\mathrm{df},3}$. As explained in ref. [14], $\mathcal{K}_{\mathrm{df},3}$ is related to the physical three-particle amplitude, $\mathcal{M}_3$, via integral equations.

An alternative approach to QCD is the use of effective field theories, with Chiral Pertur-bation Theory (ChPT) [54, 55] being the paradigm for meson dynamics at low energies. Besides its many phenomenological applications, the synergy between ChPT and lattice QCD is indeed frequently exploited in the literature. ChPT expressions allow one to address the quark-mass dependence, discretization effects, and finite-volume effects for certain quantities. To highlight one example, the $\pi\pi$ scattering lengths can be very well constrained by combining lattice QCD results at heavier-than-physical pion masses with ChPT extrapolations to the real-world value of the mass; see the dedicated chapter in the FLAG report [56] and refs. [57–59]. Given the recent progress in three-particle scattering amplitudes [60, 61], one hopes that a similar path can be followed for three-pion quanti-ties. So far, however, comparisons between ChPT and lattice QCD results for three-pion systems have only been qualitative.

The leading-order (LO) ChPT prediction for $\mathcal{K}_{\mathrm{df},3}$ for the $3\pi^+$ system was determined in ref. [44]. When compared to lattice QCD results from refs. [44, 47, 52], however, a significant disagreement was observed. This finding was surprising, given how well the two-particle counterpart, the maximal-isospin $\pi\pi$ K-matrix, is described by LO ChPT. It is thus important to understand the cause of this discrepancy. One source could be sys-tematic errors in the lattice QCD calculation, since extracting $\mathcal{K}_{\mathrm{df},3}$ from the finite-volume spectrum is numerically challenging as the shifts in the finite-volume energy levels from their free values are primarily determined by two-particle interactions. Another source could be the importance of higher-order ChPT corrections. In this work, we address the

latter possibility by determining the next-to-leading-order (NLO) predictions for $\mathcal{K}_{df,3}$.

The first step in this direction was carried out in refs. [60, 61], where the three-meson scattering amplitude, $\mathcal{M}_3$, was determined to NLO for a number of mesonic effective theories, including the one relevant for pions, i.e., ChPT with two quark flavors. It is, however, not obvious how to connect $\mathcal{M}_3$ to $\mathcal{K}_{df,3}$, since their relation, based on integral equations, needs to be inverted. The aim of this work is to combine the results of ref. [60] with the RFT approach to provide the NLO ChPT prediction for $\mathcal{K}_{df,3}$. We focus here on the case of three pions at maximal isospin, where most of the lattice QCD data is available.

The derivation of the RFT formalism in ref. [13] leads to $\mathcal{K}_{df,3}$ having the key properties of being real, smooth, and invariant under the same symmetries as $\mathcal{M}_3$ (i.e. Lorentz, parity, and time-reversal symmetries). In particular, all the branch cuts present in $\mathcal{M}_3$ due to unitarity (two- and three-particle cuts), as well as the single-particle pole due to one-particle exchange (OPE), are absent in $\mathcal{K}_{df,3}$ by construction.[1] An important aspect of an NLO ChPT calculation is that it can provide a check of these properties (and the RFT derivation) much more extensively than the LO result.

Since $\mathcal{K}_{df,3}$ is smooth, it can be expanded about threshold, constrained only by the above-mentioned symmetries. Such a "threshold expansion" is the three-particle analog of the effective-range expansion for the two-particle phase shift (or K-matrix). It has been worked out for the $3\pi^+$ scattering amplitude up to quadratic order as an expansion in relativistic invariants. As described in ref. [22], at this order there are only five unknown constants[2] in $\mathcal{K}_{df,3}$, i.e., $\mathcal{K}_0$, $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{K}_A$, and $\mathcal{K}_B$; see eq. (IV.2.2). The leading two orders, $\mathcal{K}_0$ and $\mathcal{K}_1$, give rise only to isotropic terms, i.e., those that are independent of the angles between particles. Angular dependence enters through the $\mathcal{K}_A$ and $\mathcal{K}_B$ terms. At LO in ChPT, ref. [44] finds that all terms but $\mathcal{K}_0$ and $\mathcal{K}_1$ vanish. At NLO, all five terms are expected to be nonzero, and our aim here is to determine their values.

Several technical complications need to be addressed to obtain the NLO result for $\mathcal{K}_{df,3}$. First, a relation between $\mathcal{K}_{df,3}$ and $\mathcal{M}_3$ that is valid at NLO in ChPT has to be established. As we will show, both at LO and NLO, the relation between $\mathcal{K}_{df,3}$ and $\mathcal{M}_3$ is algebraic and linear, which simplifies the subsequent calculation. Second, $\mathcal{K}_{df,3}$ depends, in general, on a cutoff function, and is thus unphysical. This cutoff function appears in the subtraction to cancel the aforementioned divergences in $\mathcal{M}_3$, and evaluating its contribution requires a tailored numerical approach. Third, a strategy to isolate the different terms in the threshold expansion of $\mathcal{K}_{df,3}$ is needed, both in the contributions that are analytical as well as in those that can only be evaluated numerically. Here, we succeeded in having an essentially fully analytical result checked by numerical calculations.

Once the NLO results for the different terms in the threshold expansion of $\mathcal{K}_{df,3}$ have been worked out, we can compare to lattice QCD data. In particular, we use the results of ref. [52], which provides values for different coefficients of the threshold expansion at three values of the pion mass. As we will see, the agreement in $\mathcal{K}_0$ and $\mathcal{K}_1$ between lattice

---

[1]The absence of the OPE pole is the reason why this is denoted a df = "divergence-free" quantity.

[2]In ref. [22], the constants are called $\mathcal{K}_{df,3}^{iso}$, $\mathcal{K}_{df,3}^{iso,1}$, $\mathcal{K}_{df,3}^{iso,2}$, $\mathcal{K}_{df,3}^{(2,A)}$ and $\mathcal{K}_{df,3}^{(2,B)}$, respectively, where "iso" marks the coefficients of the isotropic terms. We have chosen to use an abbreviated notation here.

QCD results and ChPT is significantly improved once NLO effects are incorporated. It is, however, interesting to note that NLO effects seem to be rather large, in particular in $\mathcal{K}_1$.

The remainder of this paper is structured as follows. In section 2, we provide the necessary background to compute $\mathcal{K}_{\mathrm{df},3}$ in ChPT at NLO. Section 3 then presents the central results of this paper, while we leave the technical part of the calculation to section 4. Finally, some conclusions are presented in section 5. This paper contains 5 appendices, detailing the cutoff dependence of $\mathcal{K}_{\mathrm{df},3}$ (appendix A) and the loop integrals in $\mathcal{M}_3$ (appendix B), verifying the cancellation of imaginary parts in $\mathcal{K}_{\mathrm{df},3}$ (appendix C), and supplementing section 4.2 (appendix D) and section 4.3 (appendix E).

2   The three-particle K-matrix from ChPT

In this section, we provide the necessary background to compute $\mathcal{K}_{\mathrm{df},3}$ at NLO in ChPT. Section 2.1 introduces $\mathcal{K}_{\mathrm{df},3}$ and describes its role in the three-particle formalism, and section 2.2 establishes its connection to $\mathcal{M}_3$ at NLO in ChPT. Then, an explicit calculation of $\mathcal{K}_{\mathrm{df},3}$ at LO is provided in section 2.4, and the strategy to follow at NLO is outlined in section 2.5. We defer all technical details of the computation to section 4.

2.1   The role of $\mathcal{K}_{\mathrm{df},3}$ in the three-particle formalism

The three-particle finite-volume formalism for identical scalar particles without two-to-three transitions was derived in ref. [13]. In (isospin-symmetric) QCD, it applies for three pions or three kaons at maximal isospin, $I = 3$. We will focus here on the former case. The central equation of the formalism is the quantization condition, whose solutions correspond to the energy levels E_n of a three-pion system with total three-momentum $\boldsymbol{P}$ in a box of side L,

$$\det\left[F_3^{-1}(E, \boldsymbol{P}, L) + \mathcal{K}_{\mathrm{df},3}(E^*) \right] = 0 \qquad \text{at} \qquad E = E_n \,. \qquad (\text{IV.2.1})$$

This is valid in the energy range where only three-pion intermediate states can go on shell, i.e., $M_\pi < E^* < 5M_\pi$,[3] with E^* the total energy in the center-of-mass frame (CMF). F_3 is a quantity that depends on the volume, kinematic functions and two-particle interactions, and $\mathcal{K}_{\mathrm{df},3}$ is a real, Lorentz-invariant and smooth function of E^* that parametrizes short-range three-particle interactions. Both quantities are matrices in a space that describes the kinematics of three on-shell particles, and the determinant is taken over those indices. In particular, the choice in the RFT formalism is to describe the three-particle system as composed by a pair of particles with angular momentum indices ℓ and m, usually called the *interacting pair* or *dimer*, plus a third particle with three-momentum $\boldsymbol{k}$, called the *spectator*.

---

[3]In the following, M_π will denote the renomalized mass of the pions, i.e., $p_i^2 = M_\pi^2$ for on-shell pions of momenta p_i. In amplitude calculations, this is called the "physical" mass to distinguish it from the non-renormalized mass appearing in the Lagrangian. In the lattice community, the "physical" mass typically refers to the real-world value $M_{\mathrm{phys}} \approx 139.570$ MeV. In this work we will use the latter convention, and so M_π will in general be different from M_{phys}.

Note that $\mathcal{K}_{\mathrm{df},3}$ is a scheme-dependent quantity, with the scheme being determined by the choice of a cutoff function applied to the momentum of the spectator particle. This function ensures that the matrices appearing in the quantization conditions have finite size. For more details on the implementation of the formalism, we refer the reader to refs. [20, 22, 26, 36].

The key feature of eq. (IV.2.1) needed for this work is that, given a set of $3\pi^+$ finite-volume energy levels, $\mathcal{K}_{\mathrm{df},3}$ can be extracted by fitting the predicted spectrum to the measured one. For this, one needs a parametrization of $\mathcal{K}_{\mathrm{df},3}$ in terms of few independent quantities. A systematic approach is to expand $\mathcal{K}_{\mathrm{df},3}$ in terms of relativistic invariants organized by the distance to the three-particle threshold. To reduce the number of independent parameters, one can use the fact that $\mathcal{K}_{\mathrm{df},3}$ has the same symmetries as the scattering amplitude, i.e., parity, time-reversal and particle-exchange symmetries. This leads to the threshold expansion worked out in ref. [22]. As explained in that work, only five independent terms contribute to the expansion up to quadratic order:

$$M_\pi^2 \mathcal{K}_{\mathrm{df},3} = \mathcal{K}_0 + \mathcal{K}_1 \Delta + \mathcal{K}_2 \Delta^2 + \mathcal{K}_A \Delta_A + \mathcal{K}_B \Delta_B + \mathcal{O}(\Delta^3)\,, \qquad \text{(IV.2.2)}$$

where the following kinematic quantities have been defined:

$$\Delta \equiv -\frac{1}{2}\sum_{i,j} \tilde{t}_{ij} = \frac{P^2 - 9M_\pi^2}{9M_\pi^2}\,,$$

$$\Delta_A \equiv \sum_i \left(\Delta_i^2 + \Delta_i'^2\right) - \Delta^2\,, \qquad \Delta_B \equiv \sum_{i,j} \tilde{t}_{ij}^2 - \Delta^2\,. \qquad \text{(IV.2.3)}$$

Here, $P = (E, \boldsymbol{P})$ is the total four-momentum of the system, and we define

$$\tilde{t}_{ij} \equiv \frac{(p_i - k_j)^2}{9M_\pi^2}\,,$$

$$\Delta_j \equiv \sum_i \tilde{t}_{ij} + \Delta = \frac{(P - k_j)^2 - 4M_\pi^2}{9M_\pi^2}\,, \qquad \text{(IV.2.4)}$$

$$\Delta_i' \equiv \sum_j \tilde{t}_{ij} + \Delta = \frac{(P - p_i)^2 - 4M_\pi^2}{9M_\pi^2}\,.$$

We choose k_1, k_2, k_3 to be the incoming momenta, and p_1, p_2, p_3 the outgoing ones, so that $P = k_1 + k_2 + k_3 = p_1 + p_2 + p_3$. We reiterate that $\mathcal{K}_X$ with $X = 0, 1, 2, A, B$ are unknown, dimensionless constants to be determined. As noted in the introduction, the only terms that lead to nontrivial dependence on the relative angles in the initial or final three-particle state are Δ_A and Δ_B. Moreover, only Δ_B leads to contributions with overall angular momentum differing from zero.

Once the coefficients of eq. (IV.2.2) are determined from lattice QCD simulation, one has to connect the scheme-dependent $\mathcal{K}_{\mathrm{df},3}$ to the physical scattering amplitude, $\mathcal{M}_3$. The relation between both quantities was derived in ref. [14] and involves integral equations. In this paper, we use that relation extensively, so as to obtain $\mathcal{K}_{\mathrm{df},3}$ at a consistent order in ChPT provided the corresponding prediction for $\mathcal{M}_3$. Therefore, we reproduce here the

key results from ref. [14]. For numerical solutions of the integral equations, see refs. [62, 63].

We begin by recalling the kinematic variables used in the RFT approach. As already anticipated, the configuration of three on-shell particles is described by singling out one as a spectator with three-momentum $\boldsymbol{k}$, and boosting the interacting pair to their CMF, in which the two particles of the pair have three-momentum $\boldsymbol{a}_k^*$ and $-\boldsymbol{a}_k^*$, respectively.[4] The magnitude of these momenta is denoted $q_{2,k}^* \equiv |\boldsymbol{a}_k^*|$ and is given by

$$q_{2,k}^{*2} = \frac{1}{4}\left(E_{2,k}^{*2} - 4M_\pi^2\right), \qquad \text{with} \quad E_{2,k}^{*2} = (P-k)^2, \tag{IV.2.5}$$

where $E_{2,k}^*$ is the energy of the pair in their rest frame. Expressing the initial-state kinematics this way, and expressing the final state analogously as a spectator and an interacting pair with three-momenta $\boldsymbol{p}$ and $\pm\boldsymbol{a}_p'^*$, respectively, the three-particle amplitude can be written as a function of the spectator momenta and the *directions* of the pair momenta,

$$\mathcal{M}_3(\boldsymbol{p}, \hat{\boldsymbol{a}}_p'^*; \boldsymbol{k}, \hat{\boldsymbol{a}}_k^*), \tag{IV.2.6}$$

since magnitudes are fixed by eq. (IV.2.5); the hats denote unit vectors. Here and in what follows, the dependence on P is left implicit. This description is somehow redundant, as it involves 10 (11 if we include P^2) variables, while there are only 8 independent kinematic quantities describing a general three-to-three process. In particular, the rotational invariance of the amplitude in, say, the overall CMF is not taken into account.[5] Nevertheless, these variables are the natural choice in the RFT approach.

A further step is to decompose the angular dependence in the pair CMFs into spherical harmonics:[6]

$$\mathcal{M}_3(\boldsymbol{p}, \hat{\boldsymbol{a}}_p'^*; \boldsymbol{k}, \hat{\boldsymbol{a}}_k^*) = \sum_{\ell'm'\ell m} 4\pi Y_{\ell'm'}^*(\hat{\boldsymbol{a}}_p'^*)\mathcal{M}_3(\boldsymbol{p}, \boldsymbol{k})_{\ell'm';\ell m}Y_{\ell m}(\hat{\boldsymbol{a}}_k^*). \tag{IV.2.7}$$

This is needed because, as will be seen shortly, some of the subtractions that appear in the definition of the divergence-free version of $\mathcal{M}_3$ depend on the pair angular momenta. Subsequent relations will be written for $\mathcal{M}_3(\boldsymbol{p}, \boldsymbol{k})_{\ell'm';\ell m}$, with the angular momentum indices treated as matrix indices and often left implicit. Other three-particle quantities entering the following equations, such as $\mathcal{K}_{\mathrm{df},3}$ itself, are also written in this hybrid notation. For two-particle quantities, in which the spectator is unchanged, we follow ref. [14] and label

---

[4]Here $*$ indicates that quantities are expressed in the CMF of the interacting pair, and the subscript is used to emphasize that the quantity is expressed in the CMF of the pair for which the spectator has that particular momentum.

[5]A convenient set of 8 parameters, from which the 11-parameter set $\{P^2, \boldsymbol{p}, \hat{\boldsymbol{a}}_p'^*, \boldsymbol{k}, \hat{\boldsymbol{a}}_k^*\}$ is easy to obtain, is $\{E^*, |\boldsymbol{p}|, |\boldsymbol{k}|, \psi, \hat{\boldsymbol{a}}_p'^*, \hat{\boldsymbol{a}}_k^*\}$, where E is the total CMF energy and ψ is the angle between $\boldsymbol{p}$ and $\boldsymbol{k}$ in the overall CMF. The remaining three degrees of freedom correspond to rotations of the full system.

[6]Later in this paper, we discuss the real and imaginary parts of quantities like $\mathcal{M}_3$. When doing so, the fact that, in the standard basis, the spherical harmonics are complex should be ignored. One can show that the spherical harmonics arise as overall factors in the unitarity-like relations, and it is the imaginary part of the remainder of the expression that matters. To avoid this issue in numerical calculations, one can use the real spherical harmonics, as is done in present implementations of the three-particle quantization condition (see, e.g., ref. [22]).

them with a single spectator momentum,[7] e.g.,

$$\mathcal{M}_2(\boldsymbol{p})_{\ell'm';\ell m} = \delta_{\ell'\ell}\delta_{m'm}\mathcal{M}_{2,\ell}(q^*_{2,p}),\tag{IV.2.8}$$

and similarly for the two-particle phase-space factor,

$$\rho(\boldsymbol{p})_{\ell'm';\ell m} = \delta_{\ell'\ell}\delta_{m'm}\,\bar{\rho}(q^*_{2,p}),\qquad \bar{\rho}(q^*_{2,p}) \equiv -i\,\frac{q^*_{2,p}}{16\pi E^*_{2,p}},\tag{IV.2.9}$$

which we only define above threshold, i.e., for $E^*_{2,p} \geq 2M_\pi$, as this is all we need in this work. A full definition of $\rho(\boldsymbol{p})$ is given in ref. [14], including its subthreshold behavior, where it is real and cutoff-dependent.

We are now ready to define the divergence-free three-particle amplitude, $\mathcal{M}_{\mathrm{df},3}$:

$$\mathcal{M}_{\mathrm{df},3}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{M}_3(\boldsymbol{p},\boldsymbol{k}) - \mathcal{S}\Big\{\mathcal{D}^{(u,u)}(\boldsymbol{p},\boldsymbol{k})\Big\}.\tag{IV.2.10}$$

Here, $\mathcal{S}$ indicates symmetrization over choices of initial and final spectators (9 terms in total), while $\mathcal{D}^{(u,u)}$ is the unsymmetrized subtraction term, with (u,u) indicating unsymmetrized for both final and initial states. This symmetrization procedure is explained in detail in ref. [14], and we give only specific examples below. It is important to keep in mind that $\mathcal{M}_3$ and $\mathcal{K}_{\mathrm{df},3}$ are, by definition, fully symmetrized quantities.

The subtraction term solves an integral equation that can be expanded in powers of $\mathcal{M}_2$ to the order we need as

$$\mathcal{D}^{(u,u)}(\boldsymbol{p},\boldsymbol{k}) = -\mathcal{M}_2(\boldsymbol{p})G^\infty(\boldsymbol{p},\boldsymbol{k})\mathcal{M}_2(\boldsymbol{k}) + \int_r \mathcal{M}_2(\boldsymbol{p})G^\infty(\boldsymbol{p},\boldsymbol{r})\mathcal{M}_2(\boldsymbol{r})G^\infty(\boldsymbol{r},\boldsymbol{k})\mathcal{M}_2(\boldsymbol{k}) + \dots,\tag{IV.2.11}$$

where $\int_r \equiv \int d^3r/[2\omega_r(2\pi)^3]$, with $\omega_r = \sqrt{\boldsymbol{r}^2 + M_\pi^2}$, and[8]

$$G^\infty(\boldsymbol{p},\boldsymbol{k})_{\ell'm';\ell m} = \left(\frac{k^*_p}{q^*_{2,p}}\right)^{\ell'}\frac{4\pi Y_{\ell'm'}(\hat{\boldsymbol{k}}^*_p)H(x_p)H(x_k)Y^*_{\ell m}(\hat{\boldsymbol{p}}^*_k)}{b^2_{pk} - M^2_\pi + i\epsilon}\left(\frac{p^*_k}{q^*_{2,k}}\right)^{\ell}.\tag{IV.2.12}$$

Here, $b_{pk} \equiv P - p - k$, and $H(x)$ is a smooth cutoff function that is 0 when $x \leq 0$ and 1 when $x \geq 1$, with $x_k \equiv (P-k)^2/(4M_\pi^2)$ and similarly for x_p.[9] Note that k^*_p refers to the magnitude of $\boldsymbol{k}$ taken in the CMF of the pair associated with p, and analogously for p^*_k. We stress that the cutoff functions do not violate Lorentz symmetry, because x_k and x_p are both Lorentz invariant.

---

[7]An alternative notation involving two variables and an adjustment of factors of the energy, 2ω, has been used in some subsequent works, e.g., ref. [32].

[8]Comparing to eq. (81) of ref. [14], we note that here we use the relativistic form of the denominator, which is needed to obtain a relativistically invariant $\mathcal{K}_{\mathrm{df},3}$.

[9]In the range $0 < x < 1$, we let

$$H(x) = \exp\Big[-\tfrac{1}{x}\,\exp\big(-\tfrac{1}{1-x}\big)\Big]$$

in accordance with, for example, eqs. (28) and (29) of ref. [13], but any smooth function that interpolates between the constant values may be used. Some other choices of $H(x)$ are studied in appendix A.

The factors of $k_p^*/q_{2,p}^*$ and $p_k^*/q_{2,k}^*$ in G^∞ result from the analysis of the power-law volume-dependent contributions to the *finite-volume* correlator. In particular, it is important in that analysis that the dependence on $\boldsymbol{k}_p^*$ and $\boldsymbol{p}_k^*$ is smooth near threshold, and this requires the presence of the harmonic polynomials, e.g., $(k_p^*)^\ell Y_{\ell m}(\hat{\boldsymbol{k}}_p^*)$, rather than the spherical harmonics alone. We stress that the factors of $k_p^*/q_{2,p}^*$ and $p_k^*/q_{2,k}^*$ imply that the sum over angular momentum indices cannot be performed analytically.

Some features of $\mathcal{D}^{(u,u)}$ will be important below and we comment on them here. First, it depends only on the *on-shell* two-particle amplitudes. Second, while in general one needs to keep the entire infinite series in $\mathcal{D}^{(u,u)}$ to determine $\mathcal{M}_{\mathrm{df},3}$, when working at some fixed order in a perturbative expansion such as ChPT, only a finite subset of the terms in eq. (IV.2.11) appear. Finally, we stress that $\mathcal{D}^{(u,u)}$ depends on the choice of cutoff function, and is thus not a physical quantity.

With $\mathcal{M}_{\mathrm{df},3}$ in hand, the full relation to $\mathcal{K}_{\mathrm{df},3}$ is given by[10]

$$\mathcal{M}_{\mathrm{df},3}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{S}\left\{ \int_s \int_r \mathcal{L}^{(u,u)}(\boldsymbol{p},\boldsymbol{s})\mathcal{T}(\boldsymbol{s},\boldsymbol{r})\mathcal{R}^{(u,u)}(\boldsymbol{r},\boldsymbol{k}) \right\}, \tag{IV.2.13}$$

where

$$\mathcal{L}^{(u,u)}(\boldsymbol{p},\boldsymbol{k}) \equiv \left[\frac{1}{3} - \mathcal{M}_2(\boldsymbol{p})\rho(\boldsymbol{p})\right]\bar{\delta}(\boldsymbol{p}-\boldsymbol{k}) - \mathcal{D}^{(u,u)}(\boldsymbol{p},\boldsymbol{k})\rho(\boldsymbol{k}), \tag{IV.2.14}$$

$$\mathcal{R}^{(u,u)}(\boldsymbol{p},\boldsymbol{k}) \equiv \bar{\delta}(\boldsymbol{p}-\boldsymbol{k})\left[\frac{1}{3} - \rho(\boldsymbol{p})\mathcal{M}_2(\boldsymbol{p})\right] - \rho(\boldsymbol{p})\mathcal{D}^{(u,u)}(\boldsymbol{p},\boldsymbol{k}), \tag{IV.2.15}$$

with $\bar{\delta}(\boldsymbol{p}-\boldsymbol{k}) \equiv 2\omega_k(2\pi)^3\delta^{(3)}(\boldsymbol{p}-\boldsymbol{k})$, and, finally,

$$\mathcal{T}(\boldsymbol{p},\boldsymbol{k}) \equiv \mathcal{K}_{\mathrm{df},3}(\boldsymbol{p},\boldsymbol{k}) - \int_s \int_r \mathcal{K}_{\mathrm{df},3}(\boldsymbol{p},\boldsymbol{s})\rho(\boldsymbol{s})\mathcal{L}^{(u,u)}(\boldsymbol{s},\boldsymbol{r})\mathcal{K}_{\mathrm{df},3}(\boldsymbol{r},\boldsymbol{k}) + \ldots \tag{IV.2.16}$$

The last equation shows the first two terms in the expansion of the integral equation for $\mathcal{T}$ in powers of $\mathcal{K}_{\mathrm{df},3}$.

2.2 Relation between $\mathcal{M}_3$ and $\mathcal{K}_{\mathrm{df},3}$ in ChPT

ChPT describes the low-energy regime of QCD in terms of mesonic degrees of freedom. It allows a perturbative determination of mesonic observables in terms of momenta and masses and as a function of some a priori unknown parameters—the so-called low-energy constants (LECs). These can be determined from experiment or first principles (the latter usually via matching to lattice QCD). We refer the unfamiliar reader to refs. [64, 65] for an introduction to ChPT and to refs. [60, 61] for brief summaries of what is needed for NLO three-meson scattering. In this work, we restrict ourselves to the case of two-flavor ChPT.

---

[10]Here and below, we have rearranged some factors of $2\omega_r$ relative to ref. [14] in order to simplify the notation.

Regarding the integral equations that are part of the RFT formalism, we can implement the usual ChPT power counting by expanding in powers of $1/F_\pi^2$, where F_π is the pion decay constant.[11] From here on, we will focus on the $3\pi^+$ system, and so $\mathcal{M}_2$ and $\mathcal{M}_3$ will refer to the two-particle $I = 2$ and three-particle $I = 3$ scattering amplitudes, respectively. We note that $\mathcal{M}_2 = \mathcal{O}(1/F_\pi^2)$ and $\mathcal{M}_3 = \mathcal{O}(1/F_\pi^4)$. Since $\mathcal{L}$ and $\mathcal{R}$ begin at $\mathcal{O}(1)$,

$$\mathcal{L}^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k}) = \frac{1}{3}\bar{\delta}(\boldsymbol{p}-\boldsymbol{k}) = \mathcal{R}^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k})\,, \qquad (\text{IV.2.17})$$

we have that $\mathcal{K}_{\text{df},3} = \mathcal{O}(1/F_\pi^4)$, and thus that $\mathcal{T}^{\text{LO}} = \mathcal{K}_{\text{df},3}^{\text{LO}}$. Putting this together and noting that the symmetrization of $\mathcal{K}_{\text{df},3}/9$ simply yields $\mathcal{K}_{\text{df},3}$, the LO version of eq. (IV.2.13) reduces to

$$\mathcal{K}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{M}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{M}_3^{\text{LO}}(\boldsymbol{p},\boldsymbol{k}) - \mathcal{S}\left\{\mathcal{D}^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k})\right\}\,, \qquad (\text{IV.2.18})$$

$$\mathcal{D}^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k}) = -\mathcal{M}_2^{\text{LO}}(\boldsymbol{p})G^\infty(\boldsymbol{p},\boldsymbol{k})\mathcal{M}_2^{\text{LO}}(\boldsymbol{k})\,. \qquad (\text{IV.2.19})$$

The second equation can be further simplified by noting that $\mathcal{M}_2^{\text{LO}}$ is purely s-wave, so we can make the replacement

$$G^\infty(\boldsymbol{p},\boldsymbol{k}) \longrightarrow G_{ss}^\infty(\boldsymbol{p},\boldsymbol{k})\,, \qquad G_{ss}^\infty(\boldsymbol{p},\boldsymbol{k})_{\ell'm';\ell m} \equiv \delta_{\ell'0}\delta_{m'0}\delta_{\ell 0}\delta_{m0}\,\frac{H(x_p)H(x_k)}{b_{pk}^2 - M_\pi^2 + i\epsilon}\,. \qquad (\text{IV.2.20})$$

We can also rewrite the expression for $\mathcal{K}_{\text{df},3}$ as

$$\mathcal{K}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{S}\left\{\mathcal{K}_{\text{df},3}^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k})\right\} = \mathcal{S}\left\{\mathcal{M}_3^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k}) - \mathcal{D}^{(u,u)\text{LO}}(\boldsymbol{p},\boldsymbol{k})\right\}\,, \qquad (\text{IV.2.21})$$

where $\mathcal{M}_3^{(u,u)}$ is the unsymmetrized amplitude. This form was used to perform the LO calculation of $\mathcal{K}_{\text{df},3}$ in ref. [44], which we reproduce below in section 2.4. We note that the subtraction produces a divergence-free quantity that is automatically real.

Moving to NLO, we need to keep terms up to $\mathcal{O}(1/F_\pi^6)$. We note that the second term in the right-hand side of eq. (IV.2.16) is $\mathcal{O}(1/F_\pi^8)$, so we still have $\mathcal{T} = \mathcal{K}_{\text{df},3}$ at NLO. Since $\mathcal{D}^{(u,u)} = \mathcal{O}(1/F_\pi^4)$, the NLO expressions for $\mathcal{L}^{(u,u)}$ and $\mathcal{R}^{(u,u)}$ are equal and given by

$$\mathcal{L}^{(u,u)\text{NLO}}(\boldsymbol{p},\boldsymbol{k}) = -\mathcal{M}_2^{\text{LO}}(\boldsymbol{p})\rho(\boldsymbol{p})\bar{\delta}(\boldsymbol{p}-\boldsymbol{k}) = \mathcal{R}^{(u,u)\text{NLO}}(\boldsymbol{p},\boldsymbol{k})\,. \qquad (\text{IV.2.22})$$

Here, we are adopting the notation, also used below, that NLO indicates the next-to-leading-order contribution alone rather than the sum of LO and NLO contributions. Applying these results to eq. (IV.2.13), we find

$$\mathcal{M}_{\text{df},3}^{\text{NLO}}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{K}_{\text{df},3}^{\text{NLO}}(\boldsymbol{p},\boldsymbol{k}) - \frac{1}{3}\mathcal{S}\left\{\mathcal{K}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k})\rho(\boldsymbol{k})\mathcal{M}_2^{\text{LO}}(\boldsymbol{k}) + \mathcal{M}_2^{\text{LO}}(\boldsymbol{p})\rho(\boldsymbol{p})\mathcal{K}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k})\right\}\,. \qquad (\text{IV.2.23})$$

Using the equality of $\mathcal{K}_{\text{df},3}$ and $\mathcal{M}_{\text{df},3}$ at LO, this can be reorganized into

$$\mathcal{K}_{\text{df},3}^{\text{NLO}}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{M}_{\text{df},3}^{\text{NLO}}(\boldsymbol{p},\boldsymbol{k}) + \frac{1}{3}\mathcal{S}\left\{\mathcal{M}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k})\rho(\boldsymbol{k})\mathcal{M}_2^{\text{LO}}(\boldsymbol{k}) + \mathcal{M}_2^{\text{LO}}(\boldsymbol{p})\rho(\boldsymbol{p})\mathcal{M}_{\text{df},3}^{\text{LO}}(\boldsymbol{p},\boldsymbol{k})\right\}\,. \qquad (\text{IV.2.24})$$

---

[11]Regarding what is considered "physical", F_π is treated the same way as M_π. Its real-world value is $F_{\text{phys}} \approx 92.2$ MeV.

The final quantity that we need is the NLO part of the subtraction term $\mathcal{D}^{(u,u)}$, which is given by

$$
\begin{aligned}
\mathcal{D}^{(u,u)\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) = {} & -\mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{p})G^\infty(\boldsymbol{p},\boldsymbol{k})\mathcal{M}_2^{\mathrm{NLO}}(\boldsymbol{k}) - \mathcal{M}_2^{\mathrm{NLO}}(\boldsymbol{p})G^\infty(\boldsymbol{p},\boldsymbol{k})\mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{k}) \\
& + \int_r \mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{p})G_{ss}^\infty(\boldsymbol{p},\boldsymbol{r})\mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{r})G_{ss}^\infty(\boldsymbol{r},\boldsymbol{k})\mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{k}) \, . \quad \text{(IV.2.25)}
\end{aligned}
$$

Note that we have made the replacement $G^\infty \to G_{ss}^\infty$ in the final term since there all two-particle interactions are LO and thus purely s-wave. Accordingly, this replacement cannot be made in the other two terms, since $\mathcal{M}_2^{\mathrm{NLO}}$ contains all (even) partial waves.

The final equation we need to completely specify the NLO contribution to $\mathcal{K}_{\mathrm{df},3}$ is

$$
\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) = \mathcal{M}_3^{\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) - \mathcal{S}\Big\{ \mathcal{D}^{(u,u)\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) \Big\} \, . \quad \text{(IV.2.26)}
$$

The procedure is thus, given $\mathcal{M}_3^{\mathrm{NLO}}$, to subtract $\mathcal{D}^{(u,u)\mathrm{NLO}}$, eq. (IV.2.25), after symmetrization, and then add in the "ρ terms" on the right-hand side of eq. (IV.2.24). In fact, the latter are purely imaginary since both LO amplitudes are real, and thus, given the fact that $\mathcal{K}_{\mathrm{df},3}$ is real, we obtain a simplified result

$$
\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) = \mathrm{Re}\,\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) \, . \quad \text{(IV.2.27)}
$$

In principle, we do not need to calculate the ρ terms. However, an important cross-check on the formalism and the calculations can be obtained by showing explicitly that the imaginary part of $\mathcal{K}_{\mathrm{df},3}$ vanishes based on the unitarity of off-shell amplitudes. This is presented in appendix C.

2.3 A note on off-shell conventions

The subtraction term $\mathcal{D}$ naturally separates into a part that cancels the OPE poles, namely the symmetrization of the first line of eq. (IV.2.25), and a remainder. Below, we will find it useful to similarly separate $\mathcal{M}_3$ into an OPE and a non-OPE part. The same separation of $\mathcal{M}_{\mathrm{df},3}$ and $\mathcal{K}_{\mathrm{df},3}$ follows from this. However, this separation (unlike that of $\mathcal{D}$) is not unique.

Feynman diagrams can be categorized as either OPE (e.g., figs. IV.1a and IV.10) or non-OPE (e.g., figs. IV.1b, IV.8 and IV.9), but the contribution of each diagram in ChPT depends on the parametrization of the Nambu–Goldstone manifold.[12] This dependence must cancel when all contributions are summed into a physical amplitude, but separating based on diagrams does introduce parametrization dependence into the OPE and non-OPE parts.

Alternatively, one may view the OPE part as two four-point amplitudes with one leg off-shell, with the OPE propagator joining the off-shell legs. The off-shell four-point amplitude can be modified by the addition of a smooth function of the kinematic variables that

---

[12]See appendix B of ref. [61] for an in-depth discussion on parametrizations.

vanishes on shell, and the remainder of the complete amplitude is deferred to the non-OPE part. Thus, we say that the separation is determined by an *off-shell convention*. Note that each parametrization will naturally give rise to a specific off-shell convention, as can be seen in appendix E.

Here, we follow the off-shell convention of ref. [60], where the off-shell amplitude is defined by directly replacing the on-shell Mandelstam variables by their off-shell counterparts in a particular form of the on-shell four-pion amplitude.[13] In this approach, the off-shell amplitude is unique up to the freedom to rewrite the on-shell amplitude using $s + t + u = 4M_\pi^2$, which does not hold off-shell. This prescription leads to the same OPE and non-OPE parts independently of the underlying initial parametrization, and both parts are separately scale-independent. Nevertheless, we stress that the separation into parts depends on a choice, and that only their sum is physical; also, contributions from individual diagrams remain parametrization dependent. This is in contrast to the subtraction terms, which are unique up to the choice of cutoff function, since they are built from on-shell quantities.

2.4 Explicit calculation of $\mathcal{K}_{\mathrm{df},3}^{\mathrm{LO}}$ for $3\pi^+$ scattering

To illustrate the subtractions needed to obtain a divergence-free quantity, we work through the calculation of $\mathcal{M}_{\mathrm{df},3}^{\mathrm{LO}} = \mathcal{K}_{\mathrm{df},3}^{\mathrm{LO}}$ for the $3\pi^+$ system, the results of which were first presented in ref. [44]; see also the calculation of the corresponding quantities for $\pi^+\pi^+K^+$ and $\pi^+K^+K^+$ systems in ref. [36]. This calculation also illustrates how intermediate results may depend on the convention used for off-shell amplitudes, as described in the previous section, although the final result should be (and is) independent of such conventions.

(a) OPE diagram   **(b)** Contact diagram
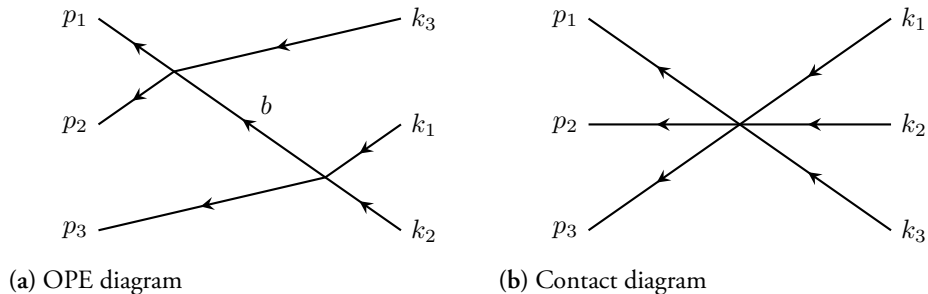
Figure IV.1: Feynman diagrams contributing to $\mathcal{M}_3$ at LO for maximal isospin. For diagram (a), there are an additional eight diagrams corresponding to the symmetrization of initial and final momenta.

To calculate $\mathcal{M}_{\mathrm{df},3}^{\mathrm{LO}}$, we use eq. (IV.2.21). Only two diagrams contribute to $\mathcal{M}_3^{\mathrm{LO}}$, which are shown in fig. IV.1, and only the first of them, the OPE diagram, requires a subtraction,

---

[13] This form is presented in eqs. (18) and (23) of ref. [60] and repeated (for maximum isospin) in eqs. (IV.2.31) and (IV.4.58) here. Ultimately, it comes from the form used in ref. [66], which is not entirely arbitrary but rather based on crossing symmetry considerations.

that is given by eq. (IV.2.19). We consider the OPE contributions first.

The results are simplified by the fact that $\mathcal{M}_2^{\mathrm{LO}}$ is purely s-wave, so we need only keep the $\ell' = \ell = 0$ contribution to $\mathcal{K}_{\mathrm{df},3}$. We obtain (with the s-wave limitation indicated by the subscripts s and using the momentum labeling of fig. IV.1)

$$
\begin{aligned}
\mathcal{K}_{\mathrm{df},3,s}^{(u,u)\mathrm{LO,OPE}}&(\boldsymbol{p}_3, \boldsymbol{k}_1) \\
&= -\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}_3) \frac{1}{b^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k}_3) + \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}_3) G_{ss}^\infty(\boldsymbol{p}_3, \boldsymbol{k}_3) \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}_3) . \quad \text{(IV.2.28)}
\end{aligned}
$$

Here, $\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}$ is the two-particle amplitude with a single leg off shell. Since both p_3 and k_3 are on shell, the H functions in G_{ss}^∞ both equal unity, and we can combine the terms to find the following result:

$$
\begin{aligned}
\mathcal{K}_{\mathrm{df},3,s}^{(u,u)\mathrm{LO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3) &= -\delta\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}_3) \frac{1}{b^2 - M_\pi^2 + i\epsilon} \delta\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k}_3) \\
&- \delta\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}_3) \frac{1}{b^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}_3) - \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}_3) \frac{1}{b^2 - M_\pi^2 + i\epsilon} \delta\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k}_3) , \quad \text{(IV.2.29)}
\end{aligned}
$$

where we define the difference between the off- and on-shell amplitudes,

$$
\delta\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}) = \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}) - \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) . \quad \text{(IV.2.30)}
$$

As we will see explicitly below, this quantity is proportional to $b^2 - M_\pi^2$ and thus cancels the poles appearing in eq. (IV.2.29).

Using the results and notation of ref. [60], the $I = 2\ \pi\pi$ off-shell amplitude is

$$
\begin{aligned}
\mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}_3) &= A^{(2)}(t_2, u_2, s_2) + A^{(2)}(u_2, s_2, t_2) \\
&= \frac{1}{F_\pi^2}(t_2 + u_2 - 2M_\pi^2) \quad \text{(IV.2.31)} \\
&= \frac{1}{F_\pi^2}\left[-s_2 + 2M_\pi^2 + (b^2 - M_\pi^2)\right] = \frac{1}{F_\pi^2}\left[-2p_1 \cdot p_2 + (b^2 - M_\pi^2)\right],
\end{aligned}
$$

where we use the subscript 2 on the Mandelstam variables to indicate that these are two-particle quantities, while $(b^2 - M_\pi^2)$ is the off-shellness of one of the legs. For example, using the labeling of momenta given in fig. IV.1a and focusing on the left vertex, we have $s_2 = (p_1 + p_2)^2$, $t_2 = (k_3 - p_1)^2$, and $u_2 = (k_3 - p_2)^2$, with $s_2 + t_2 + u_2 = 3M_\pi^2 + b^2$. The on-shell amplitude $\mathcal{M}_{2s}^{\mathrm{LO}}$ is then obtained by setting $b^2 = M_\pi^2$. Given these results, eq. (IV.2.29) yields

$$
F_\pi^4 \mathcal{K}_{\mathrm{df},3,s}^{(u,u)\mathrm{LO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3) = 2p_1 \cdot p_2 + 2k_1 \cdot k_2 - (b^2 - M_\pi^2) . \quad \text{(IV.2.32)}
$$

To symmetrize, we use the following relations:

$$
\mathcal{S}\{2p_1 \cdot p_2\} = \mathcal{S}\{2k_1 \cdot k_2\} = 3P^2 - 9M_\pi^2 , \qquad \mathcal{S}\{b^2 - M_\pi^2\} = 9M_\pi^2 - P^2 , \quad \text{(IV.2.33)}
$$

and we arrive at the final OPE contribution to $\mathcal{M}_{\mathrm{df},3}^{\mathrm{LO}}$,

$$F_\pi^4 \mathcal{M}_{\mathrm{df},3}^{\mathrm{LO,OPE}} = 6(P^2 - 3M_\pi^2) - (9M_\pi^2 - P^2) = 7P^2 - 27M_\pi^2 \,. \tag{iv.2.34}$$

The contribution from the contact term (six-point vertex) of fig. iv.1b can be read off from eqs. (30), (33), and (34) of ref. [60] which, however, describes the interaction of six pion fields $\pi^{f_1}(p_1)\pi^{f_2}(p_2)\pi^{f_3}(p_3)\pi^{f_4}(p_4)\pi^{f_5}(p_5)\pi^{f_6}(p_6)$, with flavor indices $f_i = 1,2,3$, in the "all-incoming" convention. In this language, our amplitude corresponds to the interaction of $\pi^+(k_1)\pi^+(k_2)\pi^+(k_3)\pi^-(-p_1)\pi^-(-p_2)\pi^-(-p_3)$. The connection between our conventions and those of ref. [60] is given by using $\pi^\pm = (\pi^1 \pm i\pi^2)/\sqrt{2}$ and replacing $\{p_1,\ldots,p_6\}$ by $\{k_1, k_2, k_3, -p_1, -p_2, -p_3\}$. Thus,

$$\begin{aligned}
\mathcal{M}_{\mathrm{df},3}^{\mathrm{LO,6pt}} = {} & A^{(2)}(k_1, -p_1, k_2, -p_2, k_3, -p_3) + A^{(2)}(k_1, -p_2, k_2, -p_1, k_3, -p_3) \\
& + A^{(2)}(k_1, -p_1, k_2, -p_3, k_3, -p_2) + A^{(2)}(k_1, -p_2, k_2, -p_3, k_3, -p_1) \\
& + A^{(2)}(k_1, -p_3, k_2, -p_1, k_3, -p_2) + A^{(2)}(k_1, -p_3, k_2, -p_2, k_3, -p_1) \,, \tag{iv.2.35}
\end{aligned}$$

with

$$F_\pi^4 A^{(2)}(k_1, -p_1, k_2, -p_2, k_3, -p_3) = -2k_1 \cdot p_1 - 2k_2 \cdot p_2 - 2k_3 \cdot p_3 + 3M_\pi^2 \,. \tag{iv.2.36}$$

Here, no subtraction or symmetrization is needed. Summing all six terms, we find

$$F_\pi^4 \mathcal{M}_{\mathrm{df},3}^{\mathrm{LO,6pt}} = -4P^2 + 18M_\pi^2 \,. \tag{iv.2.37}$$

Combining the OPE and contact terms, we finally obtain

$$F_\pi^4 \mathcal{M}_{\mathrm{df},3}^{\mathrm{LO}} = 3P^2 - 9M_\pi^2 = M_\pi^2(18 + 27\Delta) \,, \tag{iv.2.38}$$

which agrees with the result of ref. [44]. However, we note that the values for the separate OPE and contact contributions are not the same as in that reference, due to our different off-shell conventions.

2.5 Procedure to calculate $\mathcal{K}_{\mathrm{df},3}$ at NLO

Thanks to eq. (iv.2.27), the determination of our main result, $\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO}}$, is equivalent to calculating $\mathrm{Re}\,\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO}}$. We subdivide the calculation into multiple pieces according to the following schematic equation, which is also represented graphically in fig. iv.2:[14]

$$\mathrm{Re}\,\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO}} = \mathrm{Re}\,\mathcal{M}_3^{\mathrm{NLO,non\text{-}OPE}} - \mathrm{Re}\,\mathcal{D}^{\mathrm{BH}} + \mathrm{Re}\left\{\mathcal{M}_3^{\mathrm{NLO,OPE}} - \mathcal{D}^{\mathrm{NLO,OPE}}\right\} \,. \tag{iv.2.39}$$

Employing this separation, the one-particle-reducible diagrams contribute to the term $\mathcal{M}_3^{\mathrm{NLO,OPE}}$, while the remainder of the full $I = 3$ amplitude is denoted $\mathcal{M}_3^{\mathrm{NLO,non\text{-}OPE}}$.

---

[14]Colorblind- and monochrome-safe colors, courtesy of P. Tol (https://personal.sron.nl/~pault/), are used in figs. iv.2, iv.7 and iv.11.
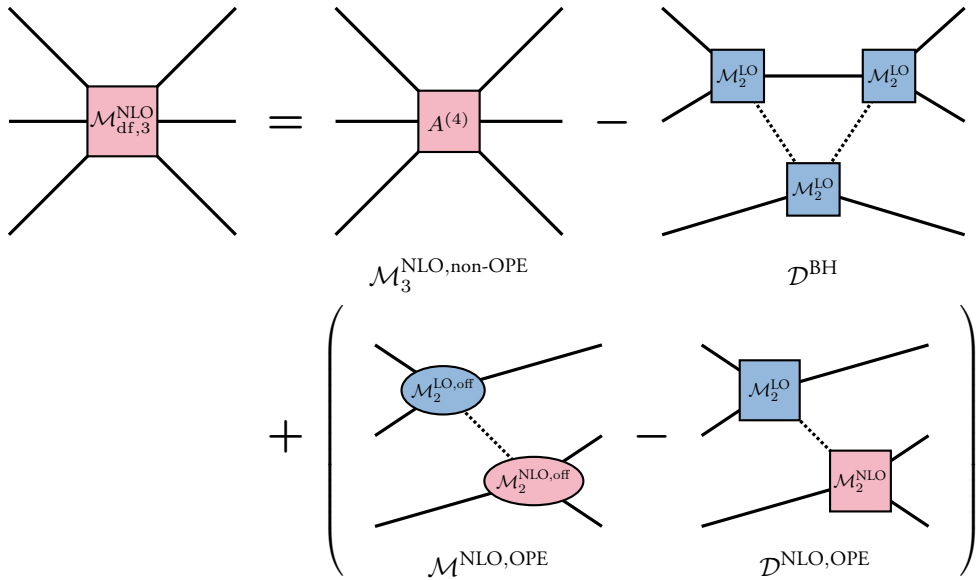
Figure ɪᴠ.2: Sketch of eq. (ɪᴠ.2.39). Solid lines represent on-shell pions, while dotted lines are off-shell propagators. Square boxes indicate fully on-shell amplitudes, while oval boxes have one leg off shell (factors of G^∞ ensure only on-shell amplitudes are needed in $\mathcal{D}$). Finally, blue and pink filling indicate, respectively, LO and NLO quantities. We leave implicit that we take only the real parts of all quantities.

Recall from section 2.3 that $\mathcal{M}_3^{\text{NLO,non-OPE}}$ and $\mathcal{M}_3^{\text{NLO,OPE}}$, while parametrization- and scale-independent, depend on our choice of off-shell convention.

We find that the real part of $\mathcal{M}_3^{\text{NLO,non-OPE}}$ is smooth at threshold, so it can be expanded directly without first subtracting any divergences, unlike for the imaginary part. Since the non-OPE contribution is smooth, the corresponding part of the subtraction must also be smooth, and can be calculated individually. This contribution, given by the third term in eq. (ɪᴠ.2.25), is the subtraction corresponding to the "bull's head" (BH) triangle diagram shown below in fig. ɪᴠ.9a. In eq. (ɪᴠ.2.39), $\mathcal{D}^{\text{BH}}$ is the symmetrization of this term. It is the only piece that depends on the cutoff function $H(x)$.

Calculating each of these pieces leads to distinctive difficulties and methods of solution. Since the full calculation is rather long and technical, we first present the main results in section 3, and then describe the full calculation in section 4: In section 4.1, we provide a general form of the threshold expansion of $\mathcal{K}_{\text{df},3}$; section 4.2 then covers the full threshold expansion of $\operatorname{Re}\mathcal{M}_3^{\text{NLO,non-OPE}}$, section 4.3 deals with $\operatorname{Re}\mathcal{D}^{\text{BH}}$, and section 4.4 covers the $\operatorname{Re}\{\mathcal{M}_3^{\text{NLO,OPE}} - \mathcal{D}^{\text{NLO,OPE}}\}$ term. All pieces can be computed analytically except $\operatorname{Re}\mathcal{D}^{\text{BH}}$. Since the real part is applied term-by-term in eq. (ɪᴠ.2.39), each of sections 4.2 to 4.4 gives

a real, finite result. We explicitly verify the cancellation of imaginary parts in appendix C.

3 Summary of results

We now present, in section 3.1, the threshold expansion of $\mathcal{K}_{\mathrm{df},3}$ at NLO in ChPT for the 3π system at maximal isospin, which constitutes the main result of this work. Some of the coefficients are then compared against lattice results from ref. [52] in section 3.2. Finally, we compare the threshold expansion to the full NLO result for a particular kinematic configuration in section 3.3.

3.1 Complete results

Including both LO and NLO contributions from ChPT, i.e., combining eqs. (IV.2.38), (IV.4.17), (IV.4.47) and (IV.4.86), the results are

$$\mathcal{K}_0 = \left(\frac{M_\pi}{F_\pi}\right)^4 18 + \left(\frac{M_\pi}{F_\pi}\right)^6 \left[-3\kappa(35 + 12\log 3) - \mathcal{D}_0 + 111L + \ell^{\mathrm{r}}_{(0)}\right], \qquad \text{(IV.3.1a)}$$

$$\mathcal{K}_1 = \left(\frac{M_\pi}{F_\pi}\right)^4 27 + \left(\frac{M_\pi}{F_\pi}\right)^6 \left[-\frac{\kappa}{20}(1999 + 1920\log 3) - \mathcal{D}_1 + 384L + \ell^{\mathrm{r}}_{(1)}\right], \quad \text{(IV.3.1b)}$$

$$\mathcal{K}_2 = \left(\frac{M_\pi}{F_\pi}\right)^6 \left[\frac{207\kappa}{1400}(2923 - 420\log 3) - \mathcal{D}_2 + 360L + \ell^{\mathrm{r}}_{(2)}\right], \quad \text{(IV.3.1c)}$$

$$\mathcal{K}_{\mathrm{A}} = \left(\frac{M_\pi}{F_\pi}\right)^6 \left[\frac{9\kappa}{560}(21809 - 1050\log 3) - \mathcal{D}_{\mathrm{A}} - 9L + \ell^{\mathrm{r}}_{(\mathrm{A})}\right], \quad \text{(IV.3.1d)}$$

$$\mathcal{K}_{\mathrm{B}} = \left(\frac{M_\pi}{F_\pi}\right)^6 \left[\frac{27\kappa}{1400}(6698 - 245\log 3) - \mathcal{D}_{\mathrm{B}} + 54L + \ell^{\mathrm{r}}_{(\mathrm{B})}\right]. \quad \text{(IV.3.1e)}$$

Here, $\kappa \equiv 1/(16\pi^2)$, $L \equiv \kappa \log(M_\pi^2/\mu^2)$, $\mathcal{D}_X$ are cutoff-dependent numerical constants related to the bull's head subtraction (see section 4.3),[15]

$$\mathcal{D}_0 \approx -0.0563476589, \qquad \mathcal{D}_1 \approx 0.129589681, \qquad \mathcal{D}_2 \approx 0.432202370,$$
$$\mathcal{D}_{\mathrm{A}} \approx 9.07273890 \cdot 10^{-4}, \qquad \mathcal{D}_{\mathrm{B}} \approx 1.62394747 \cdot 10^{-4}, \qquad \text{(IV.3.2)}$$

and we have defined the following linear combinations of LECs:

$$\ell^{\mathrm{r}}_{(0)} = -288\ell^{\mathrm{r}}_1 - 432\ell^{\mathrm{r}}_2 - 36\ell^{\mathrm{r}}_3 + 72\ell^{\mathrm{r}}_4, \qquad \ell^{\mathrm{r}}_{(1)} = -612\ell^{\mathrm{r}}_1 - 1170\ell^{\mathrm{r}}_2 + 108\ell^{\mathrm{r}}_4,$$
$$\ell^{\mathrm{r}}_{(2)} = -432\ell^{\mathrm{r}}_1 - 864\ell^{\mathrm{r}}_2, \qquad \ell^{\mathrm{r}}_{(\mathrm{A})} = 27\ell^{\mathrm{r}}_1 + \frac{27}{2}\ell^{\mathrm{r}}_2, \qquad \ell^{\mathrm{r}}_{(\mathrm{B})} = -162\ell^{\mathrm{r}}_1 - 81\ell^{\mathrm{r}}_2. \qquad \text{(IV.3.3)}$$

A numerical comparison of the different contributions to each coefficient at the physical point is given in table IV.1. Above, $\ell^{\mathrm{r}}_i \equiv \ell^{\mathrm{r}}_i(\mu)$ are scale-dependent LECs, with μ being the

---

[15]All digits shown are exact, and we have computed the values numerically to much higher precision, with the first twenty digits being verified by at least two independent methods. Higher-precision values and codes for calculating them are available upon request.

Table IV.1: Numerical comparison of the different contributions to $\mathcal{K}_{\mathrm{df},3}$ presented in eq. (IV.3.1), evaluated at the physical point, $M_{\mathrm{phys}} = 139.57$ MeV, $F_{\mathrm{phys}} = 92.2$ MeV. Errors inherited from the LECs, as listed in eq. (IV.3.7), are given where applicable; the BH (bull's head subtraction) numbers are exact up to rounding. The top part of the table covers the NLO contributions, with factors of M_π/F_π removed. In the bottom part, these factors are included so that LO and NLO can be compared; the uncertainty of M_π/F_π is not taken into account. The LO piece comes from eq. (IV.2.38). Of the NLO pieces, non-OPE comes from eq. (IV.4.17), OPE from eq. (IV.4.86), and BH from eq. (IV.4.47), with the numerical residues $\mathcal{D}_X$, also given in eq. (IV.3.2), separated out.

	$\left(\frac{F_\pi}{M_\pi}\right)^6\mathcal{K}_0$	$\left(\frac{F_\pi}{M_\pi}\right)^6\mathcal{K}_1$	$\left(\frac{F_\pi}{M_\pi}\right)^6\mathcal{K}_2$	$\left(\frac{F_\pi}{M_\pi}\right)^6\mathcal{K}_A$	$\left(\frac{F_\pi}{M_\pi}\right)^6\mathcal{K}_B$
non-OPE	$-2.04(28)$	$-3.75(61)$	$1.43(37)$	$3.00(14)$	$0.25(28)$
OPE	$0.50(53)$	$-1.8(1.0)$	$-5.11(58)$	$-2.76(15)$	$-0.22(37)$
BH, excl. $\mathcal{D}_X$	-1.16234	-3.35289	-1.67334	1.97425	0.08225
BH, only $\mathcal{D}_X$	0.05635	-0.12959	-0.43220	-0.00091	-0.00016
Total NLO	$-2.65(26)$	$-9.04(46)$	$-5.79(24)$	$2.212(16)$	$0.118(93)$
	$\mathcal{K}_0$	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_A$	$\mathcal{K}_B$
LO	94.5186	141.778	0	0	0
NLO	$-31.9(3.1)$	$-108.8(5.5)$	$-69.6(2.9)$	$26.62(19)$	$1.4(1.1)$
Total	$62.6(3.1)$	$34.0(5.5)$	$-69.6(2.9)$	$26.62(19)$	$1.4(1.1)$

renormalization scale. Different scales are related via

$$\ell_i^{\mathrm{r}}(\mu_2) = \ell_i^{\mathrm{r}}(\mu_1) + \frac{\gamma_i \kappa}{2} \log \frac{\mu_1^2}{\mu_2^2}\,, \qquad (\text{IV.3.4})$$

with

$$\gamma_1 = 1/3\,, \qquad \gamma_2 = 2/3\,, \qquad \gamma_3 = -1/2\,, \qquad \gamma_4 = 2\,. \qquad (\text{IV.3.5})$$

In combination with the L terms, this ensures the scale independence of the results in eq. (IV.3.1) and thus of $\mathcal{K}_{\mathrm{df},3}$. Often, scale-independent variants of the LECs, $\bar{\ell}_i$, are used. They are related to ℓ_i^{r} via

$$\ell_i^{\mathrm{r}}(\mu) = \kappa \frac{\gamma_i}{2}\left(\bar{\ell}_i + \log \frac{M_{\mathrm{phys}}^2}{\mu^2}\right)\,, \qquad (\text{IV.3.6})$$

where $M_{\mathrm{phys}} \approx 139.57$ MeV is the real-world pion mass. The $\bar{\ell}_i$ are fairly well known, either from phenomenology or from lattice QCD.

3.2 Comparison to lattice results

We are now in position to compare our results to lattice determinations of $\mathcal{K}_{\mathrm{df},3}$. Several works [44, 47, 48, 52] have applied the RFT formalism to the study of three pions at max-

imal isospin, and in all cases similar qualitative disagreement with LO ChPT predictions was found. We will see that this can be explained by NLO ChPT contributions. In particular, we compare to ref. [52], which studied the scattering process for pion masses of 200, 280 and 340 MeV. Note this is the only work in which $\mathcal{K}_{\mathrm{df},3}$ has been (partially) determined to quadratic order. The disagreement with LO ChPT in $\mathcal{K}_0$ and $\mathcal{K}_1$ is seen also in ref. [47] and is resolved similarly at NLO, but we do not include this in the plots due to its large uncertainties.

We take the following reference values for the scale-independent LECs,

$$\bar{\ell}_1 = -0.4(6)\,, \qquad \bar{\ell}_2 = 4.3(1)\,, \qquad \bar{\ell}_3 = 3.07(64)\,, \qquad \bar{\ell}_4 = 4.02(45)\,, \qquad \text{(iv.3.7)}$$

where $\bar{\ell}_1$ and $\bar{\ell}_2$ are determined by combining experiment, ChPT and dispersion relations [67], while $\bar{\ell}_3$ and $\bar{\ell}_4$ come from the averaged $N_{\mathrm{f}} = 2+1$ lattice QCD results [56], based on refs. [68–72]. We also take into account correlations between $\bar{\ell}_1$ and $\bar{\ell}_2$ using the covariance matrix from ref. [67]:

$$\mathrm{Cov}(\bar{\ell}_1, \bar{\ell}_2) = \begin{pmatrix} 0.35 & -0.033 \\ -0.033 & 0.012 \end{pmatrix}. \qquad \text{(iv.3.8)}$$

Although the results in eq. (iv.3.1) are independent of the choice of μ, it is customary in lattice computations to choose μ such that the results depend only on M_π/F_π. In our case this can be achieved by taking $\mu = 4\pi F_\pi$ in the L terms of eq. (iv.3.1), and approximating it as $\mu \approx 4\pi F_{\mathrm{phys}}$ in eq. (iv.3.6). This approximation only impacts the results for $\mathcal{K}_{\mathrm{df},3}$ at NNLO since F_π is independent of M_π at LO. With this choice, for example, eq. (iv.3.1a) can be rewritten as

$$\mathcal{K}_0 = \left(\frac{M_\pi}{F_\pi}\right)^4 18 \;+\; \left(\frac{M_\pi}{F_\pi}\right)^6 \left[-3\kappa(35 + 12\log 3) - \mathcal{D}_0 + 111\kappa \log\frac{\xi}{\xi_{\mathrm{phys}}} + \kappa\bar{\ell}_{(0)}\right], \quad \text{(iv.3.9)}$$

with $\bar{\ell}_{(0)} = -48\bar{\ell}_1 - 144\bar{\ell}_2 + 9\bar{\ell}_3 + 72\bar{\ell}_4$, $\xi \equiv M_\pi^2/(4\pi F_\pi)^2$ and $\xi_{\mathrm{phys}} \equiv M_{\mathrm{phys}}^2/(4\pi F_{\mathrm{phys}})^2$.

In fig. iv.3, we compare the ChPT predictions for $\mathcal{K}_0$ and $\mathcal{K}_1$ including NLO contributions to lattice results. We also show the LO predictions for comparison. We observe how the agreement with the lattice results is vastly improved by the inclusion of NLO terms. For $\mathcal{K}_0$ the addition of the NLO term leads to smaller values that are closer to the lattice results, while for $\mathcal{K}_1$ the correction produces a change of sign (for all except very small pion masses) that brings the sign and rough magnitude into agreement with that of the lattice results. A conservative interpretation of these results could be that, since the NLO corrections are so large, particularly for $\mathcal{K}_1$, the convergence of the chiral expansion is poor in the regime where lattice results have been obtained and the ChPT results cannot be trusted. A more optimistic interpretation is that the NLO results are, for some reason, larger than the LO contributions, but are nevertheless reliable. This could be because new classes of diagrams, such as the bull's-head diagram, appear at NLO. In either view, however, the discrepancy between lattice results and LO ChPT is resolved.

Due to the similarity between the ChPT prediction and the lattice results, as an exercise we perform a fit to the lattice data for $\mathcal{K}_0$ and $\mathcal{K}_1$ to estimate how well the values of the
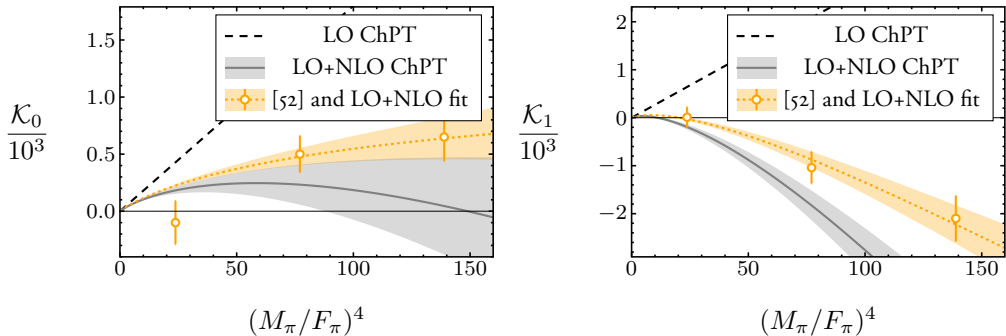
Figure IV.3: LO (dashed black line) and NLO (grey line and band) ChPT predictions for $\mathcal{K}_0$ (left) and $\mathcal{K}_1$ (right) as functions of $(M_\pi/F_\pi)^4$, using LECs from refs. [56, 67] [see eq. (IV.3.7)]. These are compared to lattice results from ref. [52] (orange points); for reference, the physical point is at $(M_{\text{phys}}/F_{\text{phys}})^4 \approx 5.25$. We also present the best fit to the lattice data (dotted orange line and orange band).

LECs in eq. (IV.3.3) can be constrained. We obtain the following results:

$$
\begin{aligned}
\ell^{\text{r}}_{(0)} &= 1.55(11), &\quad \chi^2/\text{dof} &= 2.93/2, \\
\ell^{\text{r}}_{(1)} &= 4.09(25), &\quad \chi^2/\text{dof} &= 0.36/2,
\end{aligned}
\tag{IV.3.10}
$$

which are not that dissimilar to those computed using eq. (IV.3.7),

$$
\ell^{\text{r}}_{(0)} = 1.19(25), \qquad \ell^{\text{r}}_{(1)} = 2.71(46). \tag{IV.3.11}
$$

These fits are also shown in fig. IV.3.

We show the predictions for $\mathcal{K}_2$, $\mathcal{K}_A$ and $\mathcal{K}_B$ in fig. IV.4. Here there are only NLO contributions, since these quantities vanish at LO in ChPT. In the case of $\mathcal{K}_B$, we also compare the expectations to results from ref. [52]. This time, however, we observe a much larger discrepancy, with the ChPT prediction taking the opposite sign to the lattice results, although the magnitude is roughly correct. This discrepancy is superficially similar to that between $\mathcal{K}_1$ and its leading nonzero prediction in ChPT, so it is possible that it is resolved by NNLO terms.

3.3 Range of validity of the threshold expansion

The determination of $\mathcal{K}_{\text{df},3}$ in ref. [52] assumes that the threshold expansion truncated at quadratic order provides an accurate representation of the full amplitude up to the inelastic threshold at $E^* = 5M_\pi$. Our results allow a test of this assumption, as we can compare the threshold expansion against the full result evaluated numerically for a particular kinematical configuration. Throughout this section, we make use of momentum family I (see
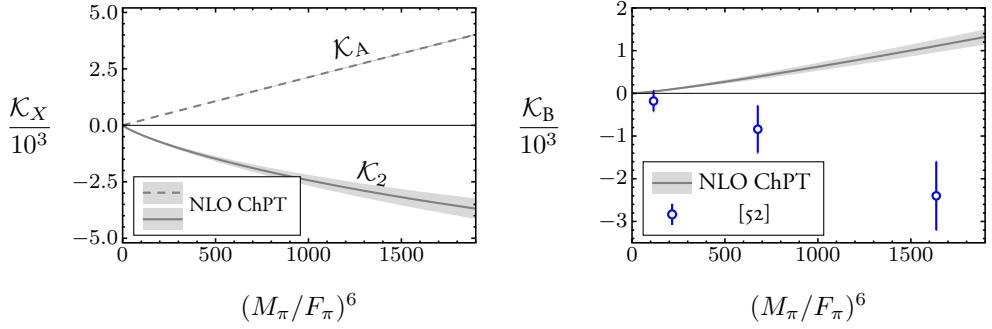
Figure IV.4: NLO ChPT predictions for $\mathcal{K}_2$ (left panel, solid line), $\mathcal{K}_A$ (left panel, dashed line) and $\mathcal{K}_B$ (right panel, solid line) as a function of $(M_\pi/F_\pi)^6$, using LECs from refs. [56, 67] [see eq. (IV.3.7)]. In the case of $\mathcal{K}_B$, we compare to lattice results from ref. [52] (blue points); for reference, the physical point is at $(M_{\text{phys}}/F_{\text{phys}})^6 \approx 12.0$. Note that all three coefficients vanish at LO in ChPT.
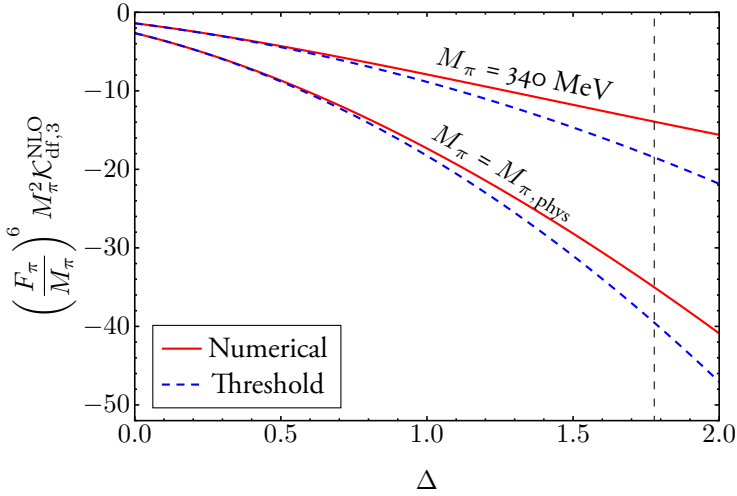
Figure IV.5: Comparison between numerical results and the threshold expansion for $\mathcal{K}_{\text{df},3}$, evaluated for momentum family I (see table IV.2). The comparison is presented for $M_\pi = M_{\text{phys}}$ and $M_\pi = 340$ MeV, the latter corresponding to the heaviest pion mass used in ref. [52]. The dashed vertical line indicates the inelastic threshold, which occurs at $E^* = 5M_\pi$.

table IV.2 in appendix D). We have checked that the results for other kinematic families are comparable.

In fig. IV.5 we perform this test for the total $\mathcal{K}_{\text{df},3}^{\text{NLO}}$, both for the physical pion mass and for $M_\pi = 340$ MeV, the latter being the heaviest pion mass used in ref. [52]. In both
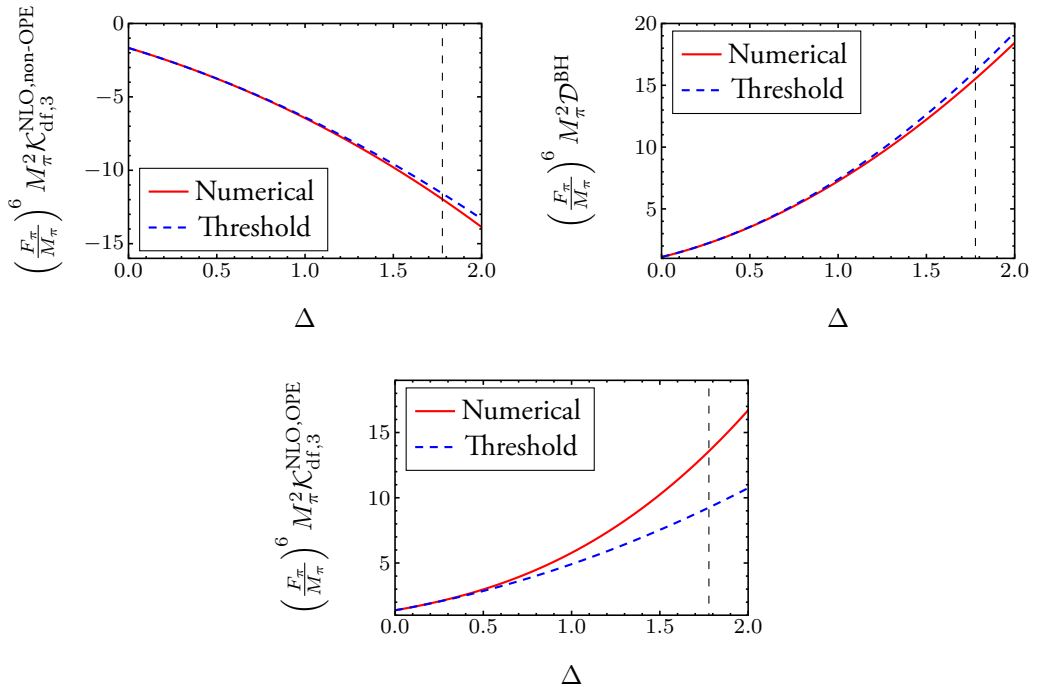
Figure IV.6: Comparison between numerical results and the threshold expansion for the different terms in the right-hand side of eq. (IV.2.39), evaluated for momentum family I (see table IV.2). The panels correspond to the terms in the right-hand side of eq. (IV.2.39): the non-OPE (top left), the BH subtraction (top right) and the OPE (bottom) contributions to $\mathcal{K}_{\mathrm{df},3}$. The black vertical line indicates the inelastic threshold, which occurs at $E^* = 5M_\pi$. The comparison is made for $M_\pi = 340$ MeV, corresponding to the heaviest pion mass used in ref. [52].

cases, we observe that the threshold expansion works reasonably well up to the inelastic threshold, where the discrepancy is $\sim 10\%$ for physical pions, and $\sim 20\%$ in the heavier case. As expected, the convergence is better for smaller pion masses.

We can perform the same comparison separately for each of the three components appearing in eq. (IV.2.39) and fig. IV.2. This is presented in fig. IV.6 for $M_\pi = 340$ MeV. We recall from section 2.3 that each of the three components is independent of the renormalization scale in ChPT, although they do depend on the off-shell convention. Of the three, only the BH subtraction depends on the cutoff function H. We observe that the difference is $\lesssim 5\%$ at the inelastic threshold in the case of the non-OPE and the BH subtraction contribution, while for the OPE part it is $\sim 30\%$.

For the OPE contribution we can also study the convergence as higher partial waves of the interacting pair are included. The threshold expansion contains only s- and d-waves, while
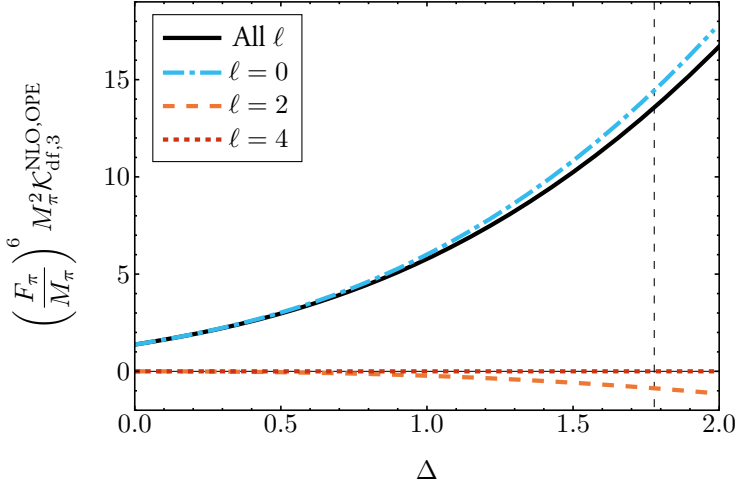
Figure IV.7: Comparison of contributions to $\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO,OPE}}$ from different interacting pair partial waves, numerically evaluated for momentum family 1 (see table IV.2). Results are shown for $M_\pi = 340$ MeV, corresponding to the heaviest pion mass in ref. [52]. The black solid line is the full result including all partial waves, and the dashed vertical line indicates the inelastic threshold. Contributions for $\ell \gtrsim 4$ are negligible.

the full result contains all even pair partial waves.[16] In fig. IV.7 we show the contributions of the first three nonzero partial waves to the numerical result for the OPE contribution for $M_\pi = 340$ MeV. We find a rapid convergence with ℓ. In particular, the contribution is negligible below the inelastic threshold for $\ell \geq 4$. A similar result holds for smaller pion masses.

4   Details of the calculation of $\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO}}$

In this section, we go in detail through the calculation outlined in section 2.5 to obtain the results of section 3. Given the length of the expressions, the algebra is done with either Wolfram Mathematica [**Mathematica**] or FORM [73], and is cross-checked independently by several of the authors. Likewise, all numerical calculations are performed using Mathematica or C+, using CHIRON [74], *LoopTools* [75] and GSL [76], with independent cross-checks and preferably using several different methods.

---

[16]The decomposition into partial waves is obtained by evaluating $\operatorname{Re}\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE}}$ using eq. (IV.4.49), and decomposing $\mathcal{M}_{2,\ell,\mathrm{off}}^{\mathrm{NLO}}$ into partial waves by numerical integration.

4.1 General form of the threshold expansion

Throughout the calculation of $\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO}}$, it will often be useful to work with $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$, the un-symmetrized version of $\mathcal{K}_{\mathrm{df},3}$. These quantities are related by $\mathcal{K}_{\mathrm{df},3} = \mathcal{S}\{\mathcal{K}_{\mathrm{df},3}^{(u,u)}\}$, where $\mathcal{S}$ indicates symmetrization of both initial and final states, as in eq. (IV.2.10). It is thus necessary to identify the form of the terms that can contribute to $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$ up to quadratic order.[17]

The symmetries that constrain $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$ are the same as those for a $2+1$ theory (i.e., with two identical particles and one that is different). Thus, we can use the threshold expansion for the latter theories derived in ref. [34]:[18]

$$
\begin{aligned}
\mathcal{K}_{\mathrm{df},3}^{(u,u)} = {}& c_0 + c_1\Delta + c_2\Delta_3^S + c_3\tilde{t}_{33} \\
& + c_4\Delta^2 + c_5\Delta\Delta_3^S + c_6\Delta\tilde{t}_{33} + c_7\Delta_3\Delta_3' + c_8(\Delta_3^S)^2 + c_9\Delta_3^S\tilde{t}_{33} + c_{10}\tilde{t}_{33}^2 \\
& + c_{11}\mathcal{Q}_{--} + c_{12}\mathcal{Q}_{+-} + c_{13}\mathcal{Q}_{3-} + c_{14}\mathcal{Q}_{tu} + \mathcal{O}(\Delta^3)\,. \quad\text{(IV.4.1)}
\end{aligned}
$$

Some quantities used above have already been defined in eq. (IV.2.4). Moreover, $\Delta_3^S \equiv \Delta_3 + \Delta_3'$ and

$$
\begin{aligned}
\mathcal{Q}_{--} &\equiv 4\left[\frac{p_-\cdot k_-}{9M_\pi^2}\right]^2\,, & \mathcal{Q}_{+-} &\equiv 2\left(\frac{p_+\cdot k_-}{9M_\pi^2}\right)^2 + 2\left(\frac{p_-\cdot k_+}{9M_\pi^2}\right)^2\,, \\
\mathcal{Q}_{3-} &\equiv 4\frac{p_-\cdot k_+}{9M_\pi^2}\frac{p_-\cdot k_3}{9M_\pi^2} + 4\frac{p_+\cdot k_-}{9M_\pi^2}\frac{p_3\cdot k_-}{9M_\pi^2}\,, & \mathcal{Q}_{tu} &\equiv \tilde{t}_{13}\tilde{t}_{23} + \tilde{t}_{31}\tilde{t}_{32}\,,
\end{aligned}
$$
(IV.4.2)

with $p_\pm = p_1 \pm p_2$ and $k_\pm = k_1 \pm k_2$. Here, the initial and final spectators are taken to have momenta k_3 and p_3, respectively. We note that, in eq. (IV.4.1), only terms on the final line contain values of ℓ other than zero.

Symmetrizing $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$ then leads to the following contributions to $\mathcal{K}_{\mathrm{df},3}$:

$$
\begin{aligned}
\mathcal{K}_0 &= 9c_0\,, & \text{(IV.4.3a)} \\
\mathcal{K}_1 &= 9c_1 + 6c_2 - 2c_3\,, & \text{(IV.4.3b)} \\
\mathcal{K}_2 &= 9c_4 + 6c_5 - 2c_6 + c_7 + 5c_8 - c_9 + c_{10} + 4c_{11} + 2c_{12} - 2c_{13} + \tfrac{1}{2}c_{14}\,, & \text{(IV.4.3c)} \\
\mathcal{K}_A &= 3c_8 + c_9 - 3c_{11} + c_{12} + 4c_{13} + \tfrac{1}{2}c_{14}\,, & \text{(IV.4.3d)} \\
\mathcal{K}_B &= c_{10} + 9c_{11} + 3c_{12} - 6c_{13} - c_{14}\,. & \text{(IV.4.3e)}
\end{aligned}
$$

As expected, several terms from $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$ are often combined in each term in $\mathcal{K}_{\mathrm{df},3}$ because of the larger symmetry of the latter. Note that some of the terms that are purely s-wave in eq. (IV.4.1) can lead to nontrivial angular dependence after symmetrization, and thus

---

[17]Note that in this subsection, all statements concerning $\mathcal{K}_{\mathrm{df},3}/\mathcal{K}_{\mathrm{df},3}^{(u,u)}$ apply equally well to $\mathcal{M}_{\mathrm{df},3}/\mathcal{M}_{\mathrm{df},3}^{(u,u)}$ or $\mathcal{D}/\mathcal{D}^{(u,u)}$, for those cases in which divergences are absent such that the threshold expansion is well-defined.

[18]Two operators are missed by the analysis of ref. [34], $\mathcal{Q}_{3-}$ and $\mathcal{Q}_{tu}$. Note that we have used a somewhat different basis than in ref. [34].

contribute to $\mathcal{K}_A$ and $\mathcal{K}_B$, because there can be nonzero angular momentum between the spectator and the pair.

We also make use of another basis of operators, in which all quantities are expressed in terms of $\tilde{t}_{ij}$ [defined in eq. (IV.2.4)], for expanding the symmetrized K-matrix directly. We can write

$$\mathcal{K}_{\mathrm{df},3} = c_0' + c_1' \mathcal{Q}_0' + c_2' \mathcal{Q}_1' + c_3' \mathcal{Q}_2' + c_4' \mathcal{Q}_3' + \mathcal{O}(\Delta^3) \,, \qquad (\text{IV.4.4})$$

where

$$\begin{aligned}
\mathcal{Q}_0' &\equiv \mathcal{S}[\tilde{t}_{11}] & &= -2\Delta \,, \\
\mathcal{Q}_1' &\equiv \mathcal{S}[\tilde{t}_{11}\tilde{t}_{11}] & &= \Delta^2 + \Delta_B \,, \\
\mathcal{Q}_2' &\equiv \mathcal{S}[\tilde{t}_{11}\tilde{t}_{12} + \tilde{t}_{11}\tilde{t}_{21}] & &= \tfrac{1}{2}(2\Delta^2 + \Delta_A - 2\Delta_B) \,, \\
\mathcal{Q}_3' &\equiv \mathcal{S}[\tilde{t}_{11}\tilde{t}_{22} + \tilde{t}_{21}\tilde{t}_{12}] & &= \tfrac{1}{2}(2\Delta^2 - \Delta_A + \Delta_B) \,.
\end{aligned} \qquad (\text{IV.4.5})$$

Here, we have indicated the relation to the basis used in eq. (IV.2.2). Employing these relations, we find that the coefficients in eq. (IV.2.2) are given by

$$\begin{aligned}
\mathcal{K}_0 &= c_0' \,, & \mathcal{K}_1 &= -2c_1' \,, & \mathcal{K}_2 &= c_2' + c_3' + 2c_4' \,, \\
& & \mathcal{K}_A &= c_3' - c_4' \,, & \mathcal{K}_B &= c_2' - 2c_3' + c_4' \,.
\end{aligned} \qquad (\text{IV.4.6})$$

The requirement for the two methods of expansion to agree serves as a cross-check. Specifically, eqs. (IV.4.3) and (IV.4.6) [or, equivalently, eqs. (IV.2.2) and (IV.4.4)] must match.

4.2 Threshold expansion of the non-OPE part of $\mathcal{M}_3$
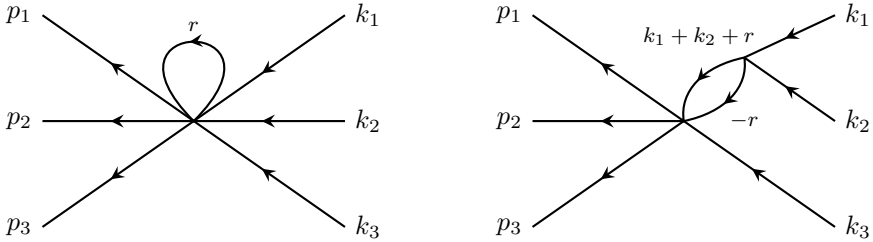
Figure IV.8: Examples of non-OPE NLO diagrams for which no subtraction is needed.

In this section, we focus on threshold-expanding the non-OPE piece of the amplitude in eq. (IV.2.39). This is the piece related to one-particle-irreducible diagrams, like those in figs. IV.8 and IV.9, for which the contribution to $\mathcal{K}_{\mathrm{df},3}$ can be computed independently of any subtraction. For the $I = 3$ system, this contribution is given by

$$\begin{aligned}
\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}} &= A^{(4)}(k_1, -p_1, k_2, -p_2, k_3, -p_3) + A^{(4)}(k_1, -p_2, k_2, -p_1, k_3, -p_3) \\
&\quad + A^{(4)}(k_1, -p_1, k_2, -p_3, k_3, -p_2) + A^{(4)}(k_1, -p_2, k_2, -p_3, k_3, -p_1) \\
&\quad + A^{(4)}(k_1, -p_3, k_2, -p_1, k_3, -p_2) + A^{(4)}(k_1, -p_3, k_2, -p_2, k_3, -p_1) \,,
\end{aligned} \qquad (\text{IV.4.7})$$

where $A^{(4)}$ is defined in eq. (35) of ref. [60], and we have converted to our conventions as described above eq. (IV.2.35).

This part of the amplitude can be divided as

$$A^{(4)} = A_C + A_J + A_\pi + A_L + A_l\,, \qquad\qquad (\text{IV.4.8})$$

where the division is similar to that in eq. (35) of ref. [60], although here we group some terms together. In particular, A_C is the part that contains the $\overline{C}$ functions, i.e., $\overline{C}$, $\overline{C}_{11}$ and $\overline{C}_{21}$ ($\overline{C}_3$ does not contribute to $I = 3$), A_J contains all terms with $\bar{J}$ functions, and A_π, A_L, and A_l are terms containing only factors of κ, L or LECs, respectively (see sections B and 3.1 for definitions). Some of these amplitudes have imaginary parts, but we only need to calculate the real parts since the imaginary parts do not contribute to $\mathcal{K}_{\text{df},3}$ [see eq. (IV.2.27)].

The computation of c_i' [following eq. (IV.4.4)] is straightforward for A_l, A_L, and A_π. The results for the latter two are given in eq. (IV.4.19). The remaining parts, which contain loop integrals, require more care. As is shown in the following subsections, they can be expanded in terms of squared momenta that are either small or close to threshold. These can then be straightforwardly related to $\tilde{t}_{ij}$ and $\Delta_i^{(\prime)}$ through eq. (IV.2.4). Alternatively, one can use (at least) three different particular kinematic configurations to numerically determine the corresponding coefficients in the Δ variables. This cross-check is described in appendix D.

4.2.1 Threshold expansion of A_J

Terms containing $\bar{J}(q^2)$ functions must be expanded about either $q^2 = 0$ or $q^2 = 4M_\pi^2$ (the two-particle threshold). In the notation of ref. [60], $A_J = A_J^{(1)} + A_J^{(2)}$, where all $\bar{J}$ functions in $A_J^{(1)}$ are expanded about $q^2 = 4M_\pi^2$, while in $A_J^{(2)}$ both cases appear. It is, however, easy to separate $A_J^{(2)}$ in two parts, each of which contains one single case. Using again the notation of ref. [60] [in particular eq. (B9) thereof],

$$\begin{aligned}
A_J^{(2)} = (&R_{132456} + R_{241356} + R_{152634} + R_{152634} + R_{261543} + R_{536412} + R_{645321} \\
&+ R_{142356} + R_{231456} + R_{251634} + R_{162543} + R_{635421} + R_{546312})A_{J,0}^{(2)}\,, \quad (\text{IV.4.9})
\end{aligned}$$

where R_{ijklmn} indicates an operator, acting on $A_{J,0}^{(2)}$, that permutes $p_1 \to p_i$, $p_2 \to p_j$, etc., with $\{p_1, \ldots, p_6\} = \{k_1, k_2, k_3, -p_1, -p_2, -p_3\}$ as described above eq. (IV.2.35). It is easy to check that, after projecting to $I = 3$, the first line above leads to terms that contain $\bar{J}$ functions that have to be expanded about the two-particle threshold, while the second has those that need to be expanded around $q^2 = 0$.

Using the definition (IV.2.3), the expansions in the two cases are

$$\frac{1}{\kappa}\,\text{Re}\,\bar{J}(4M_\pi^2 + \bar{s}) = 2 - \frac{1}{2}\frac{\bar{s}}{M_\pi^2} + \frac{1}{12}\frac{\bar{s}^2}{M_\pi^4} - \frac{1}{60}\frac{\bar{s}^3}{M_\pi^6} + \mathcal{O}(\bar{s}^4)\,, \qquad (\text{IV.4.10})$$

$$\frac{1}{\kappa}\bar{J}(t) = \frac{1}{6}\frac{t}{M_\pi^2} + \frac{1}{60}\frac{t^2}{M_\pi^4} + \frac{1}{420}\frac{t^3}{M_\pi^6} + \mathcal{O}(t^4)\,. \qquad \text{(IV.4.11)}$$

Note that only $\mathrm{Re}\,\bar{J}$, not $\mathrm{Im}\,\bar{J}$, is smooth at threshold. After conversion to threshold-expansion parameters, the results are given in eq. (IV.4.20).

4.2.2 Threshold expansion of A_C

All $\overline{C}$, which are functions of three pairs of momenta as defined in appendix B, can be expanded either for all three pairs being small or one small and the other two near threshold. Given that the expansions are analytic, for $\mathrm{Re}\,\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}}$ the Feynman integrals can be performed naïvely with a principal-value prescription. For C, defined in eq. (IV.2.4), the starting point is the Feynman-parameter representation

$$C \equiv C(p_1, p_2, \ldots, p_6) = -\kappa \int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\frac{\delta(1 - x - y - z)}{M_\pi^2 - xyq_1^2 - yzq_2^2 - zxq_3^2}\,, \qquad \text{(IV.4.12)}$$

with $q_1 = p_1 + p_2$, $q_2 = p_3 + p_4$ and $q_3 = p_5 + p_6$. For the three q_i^2 small, we straightforwardly expand the denominator and then perform the Feynman integrals, arriving at

$$\frac{C}{\kappa} = -\frac{1}{2M_\pi^2} - \frac{1}{24M_\pi^4}\left(q_1^2 + q_2^2 + q_3^2\right) - \frac{1}{180M_\pi^6}\left(q_1^4 + q_2^4 + q_3^4 + q_1^2 q_2^2 + q_2^2 q_3^2 + q_3^2 q_1^2\right) + \cdots$$
$$\text{(IV.4.13)}$$

For the expansion about threshold, take for example q_1^2, $\bar{s}_2$, and $\bar{s}_3$ small, where $\bar{s}_2 = q_2^2 - 4M_\pi^2$ and $\bar{s}_3 = q_3^2 - 4M_\pi^2$. We can write the Feynman parametrization of C as

$$C = -\kappa \int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\frac{\delta(1 - x - y - z)}{M_\pi^2 - 4M_\pi^2(yz + zx) - xyq_1^2 - yz\bar{s}_2 - zx\bar{s}_3}$$
$$= -\kappa \int_0^1 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\frac{\delta(1 - x - y - z)}{M_\pi^2(1 - 2z)^2 - xyq_1^2 - yz\bar{s}_2 - zx\bar{s}_3}\,. \qquad \text{(IV.4.14)}$$

Since we know the integral is analytic above threshold, we expand naïvely in q_1^2, $\bar{s}_2$, and $\bar{s}_3$ and perform the Feynman integrals, first doing the x, y integrals and then the z integral using the principal-value prescription, discarding the imaginary part. In particular, after expanding and performing the x, y integrals, one obtains integrals of the type

$$\mathcal{P}\int_0^1 \mathrm{d}z\,\frac{1}{(1 - 2z)^n} = \frac{1}{2}\mathcal{P}\int_{-1}^1 \mathrm{d}v\,\frac{1}{v^n} = \frac{1}{4}\int_\pi^0 e^{-in\theta}\,\mathrm{d}e^{i\theta} + \frac{1}{4}\int_\pi^0 e^{in\theta}\,\mathrm{d}e^{-i\theta}\,, \qquad \text{(IV.4.15)}$$

where we changed variables to $v \equiv 1 - 2z$ and implemented the principal-value ($\mathcal{P}$) prescription as the average of the contours above and below the singularity at $v = 0$ in the complex plane, letting $v = e^{\pm i\theta}$.[19] The integrals vanish for odd values of n and are equal to $-1/(n - 1)$ for even values. This way, the result for eq. (IV.4.14) becomes

---

[19]We refer readers unfamiliar with principal values for multiple poles to ref. [77].

$$\frac{C}{\kappa} = \frac{1}{2M_\pi^2} + \frac{1}{M_\pi^4}\left(\frac{5}{72}q_1^2 - \frac{1}{24}\bar{s}_2 - \frac{1}{24}\bar{s}_3\right)$$
$$+ \frac{1}{M_\pi^6}\left[\frac{2}{225}q_1^4 - \frac{1}{90}q_1^2(\bar{s}_2 + \bar{s}_3) + \frac{1}{180}\left(\bar{s}_2^2 + \bar{s}_3^2 + \bar{s}_2\bar{s}_3\right)\right] + \cdots \quad \text{(IV.4.16)}$$

The other triangle integrals are expanded using the same methods. We have checked the expansions numerically against *LoopTools* [75]. The contributions to $\mathcal{K}_{\mathrm{df},3}$ from A_C are given in eq. (IV.4.21).

4.2.3 The full non-OPE threshold expansion

Performing the expansion described above to second order, we find that the total contributions from $\operatorname{Re}\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}}$ are

$$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0 \supset 14\kappa + 33L + 36(8\ell_1^{\mathrm{r}} + \ell_3^{\mathrm{r}} - 2\ell_4^{\mathrm{r}})\,, \tag{IV.4.17a}$$

$$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1 \supset -\frac{35}{2}\kappa + 12L + 36(20\ell_1^{\mathrm{r}} + \ell_2^{\mathrm{r}} - 4\ell_4^{\mathrm{r}})\,, \tag{IV.4.17b}$$

$$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2 \supset -\frac{9747}{50}\kappa - 216L + 324(2\ell_1^{\mathrm{r}} + \ell_2^{\mathrm{r}})\,, \tag{IV.4.17c}$$

$$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A \supset \frac{576}{5}\kappa - 54L - 81(2\ell_1^{\mathrm{r}} - 3\ell_2^{\mathrm{r}})\,, \tag{IV.4.17d}$$

$$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B \supset -\frac{13797}{50}\kappa - 162L + 243(2\ell_1^{\mathrm{r}} + \ell_2^{\mathrm{r}})\,. \tag{IV.4.17e}$$

For comparison, the LO results from the non-OPE diagram in fig. IV.1b are

$$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_0 \supset -18\,, \qquad \frac{F_\pi^4}{M_\pi^4}\mathcal{K}_1 \supset -26\,. \tag{IV.4.18}$$

This is specific to our off-shell convention (see section 2.3).

The results for the separate parts of $A^{(4)}$ in eq. (IV.4.8) are as follows. The part stemming from A_l can be directly read from eq. (IV.4.17) as the terms containing ℓ_i^{r}. The remaining parts not containing $\bar{J}$ or $\overline{C}$ loop functions, i.e., A_π and A_L, give

$$\frac{F_\pi^6}{M_\pi^4}\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}} \supset 17\kappa + 33L + \Delta\big(10\kappa + 48L\big) + \Delta^2\left(-\frac{387}{2}\kappa - \frac{351}{2}L\right)$$
$$+ \Delta_A\left(-\frac{495}{8}\kappa - \frac{675}{8}L\right) + \Delta_B\left(-\frac{1323}{8}\kappa - \frac{1215}{8}L\right). \tag{IV.4.19}$$

The part A_J containing solely $\bar{J}$ functions, using the results from section 4.2.1, reads

$$\frac{F_\pi^6}{M_\pi^4}\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}} \supset -72\kappa + \Delta(-223\kappa) + \Delta^2\left(-\frac{243}{4}\kappa\right) + \Delta_A\left(\frac{1233}{4}\kappa\right) + \Delta_B\left(-\frac{117}{2}\kappa\right),$$
$$\tag{IV.4.20}$$

and the $\overline{C}$ functions of A_C, calculated employing the methods presented in section 4.2.2, all together give

$$\frac{F_\pi^6}{M_\pi^4}\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}} \supset 69\kappa + \Delta\left(-36L + \frac{391}{2}\kappa\right) + \Delta^2\left(-\frac{81}{2}L + \frac{5931}{100}\kappa\right)$$

$$+ \Delta_{\mathrm{A}}\left(\frac{243}{8}L - \frac{5247}{40}\kappa\right) + \Delta_{\mathrm{B}}\left(-\frac{81}{8}L - \frac{10413}{200}\kappa\right). \quad \text{(IV.4.21)}$$

The sum of eqs. (IV.4.19) to (IV.4.21), plus the ℓ_i^{r} terms, gives eq. (IV.4.17).

4.3 Bull's head subtraction contribution

(a) The "bull's head" diagram.

(b) The "crossed bull's head" diagram.
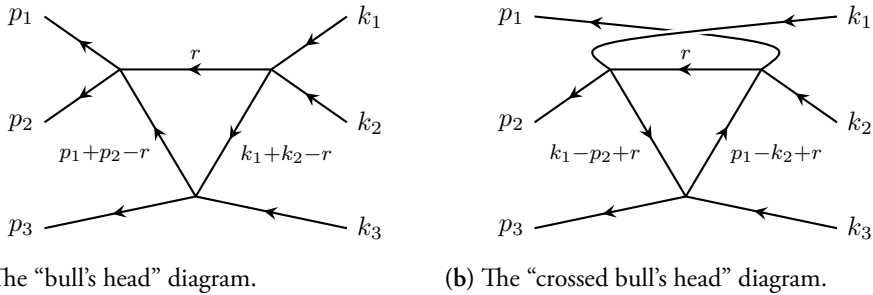
Figure IV.9: Two configurations of the triangle-loop diagram, of which only (a) requires subtraction [in the form of $\mathcal{D}^{(u,u)\mathrm{BH}}$, eq. (IV.4.23)]. There are a total of 15 diagrams with the triangle topology, of which 9 correspond to the configuration (a) [so their sum corresponds to the symmetrization of (a)] and 6 to the configuration (b).

The BH piece of the subtraction, $\mathcal{D}^{\mathrm{BH}}$ in eq. (IV.2.39), concerns the symmetrization of the final term in eq. (IV.2.11), which subtracts the "bull's head" diagram in fig. IV.9a. This is the only cutoff-dependent contribution to $\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO}}$, and due to the integral over $H(x)$, it requires numerical evaluation.

Since there are only s-wave contributions, the subtraction term reduces to

$$\mathcal{D}^{(u,u)\mathrm{BH}}(\boldsymbol{p}_3, \boldsymbol{k}_3) = -\frac{1}{F_\pi^6}(2p_1 \cdot p_2)\, I(\boldsymbol{p}_3, \boldsymbol{k}_3)(2k_1 \cdot k_2)\,, \quad \text{(IV.4.22)}$$

$$I(\boldsymbol{p}_3, \boldsymbol{k}_3) \equiv \int_r \frac{H(x_r)\big[(P-r)^2 - 2M_\pi^2\big]H(x_r)}{\big[(p_+-r)^2 - M_\pi^2 + i\epsilon\big]\big[(k_+-r)^2 - M_\pi^2 + i\epsilon\big]}\,, \quad \text{(IV.4.23)}$$

where $p_+ \equiv p_1 + p_2$, $k_+ \equiv k_1 + k_2$, and $x_r \equiv (P-r)^2/(4M_\pi^2)$. For brevity, we define $G(r; p_+, k_+)$ (not to be confused with G^∞ above) such that

$$\mathcal{D}^{(u,u)\mathrm{BH}} = -\frac{1}{F_\pi^6}(2p_1 \cdot p_2)(2k_1 \cdot k_2)\int_r H^2(x_r)G(r; p_+, k_+)\,. \quad \text{(IV.4.24)}$$

We stress that r is on-shell and that the integral is Lorentz-invariant.

4.3.1 Threshold expansion

The integrand of eq. (IV.4.23) can be expressed to any order in the threshold expansion using $\Delta, \Delta_3, \Delta_3'$, and $\tilde{t}_{33}$ only. Since the amplitude is free of singularities in the real part, it follows that the real part of the subtraction is also singularity-free. However, due to possible singularities in the integrand, some manipulations need to be done before the coefficients can be evaluated.

Starting with eq. (IV.4.23), we stay in the CMF of the whole system, where $\boldsymbol{P} = \boldsymbol{0}$ and $(P - r)^2 = E^2 - 2E\omega_r + M_\pi^2$, and set $\mathrm{d}^3 r = r^2 \mathrm{d}r \, \mathrm{d}\cos\theta \, \mathrm{d}\phi$. We rewrite the first denominator as

$$(p_+ - r)^2 - M_\pi^2 = p_+^2 - 2p_+ \cdot r = p_+^2 - 2E_{p_+}\omega_r + 2\vec{p}_+ \cdot \boldsymbol{r} \equiv 4M_\pi^2 - 4M_\pi\omega_r + \Delta_{p_+} \, , \quad \text{(IV.4.25)}$$

which defines Δ_{p_+}; Δ_{k_+} is defined similarly from the second denominator. Note that $p_+ = (E_{p_+}, \vec{p}_+)$ is the momentum of the two-pion system. We then expand naïvely in Δ_{p_+} and Δ_{k_+}, which are both $\mathcal{O}(\sqrt{\Delta})$, so

$$I(\boldsymbol{p}_3, \boldsymbol{k}_3) = \int_r H^2(x_r) \big[E^2 - 2E\omega_r - M_\pi^2 \big] \sum_{a,b=0}^{\infty} \frac{\Delta_{p_+}^a \Delta_{k_+}^b}{(4M_\pi^2 - 4M_\pi\omega_r)^{a+b+2}} \, , \quad \text{(IV.4.26)}$$

and cut off the sums at a suitable order. After that, we can perform all angular integrals using

$$\int \mathrm{d}\cos\theta \, \mathrm{d}\phi \, (\vec{p}_+ \cdot \boldsymbol{r})(\vec{k}_+ \cdot \boldsymbol{r}) = 4\pi \frac{r^2}{3} \vec{p}_+ \cdot \vec{k}_+ \quad \text{(IV.4.27)}$$

and its generalizations.[20] Finally, we expand the explicit and implicit [via $E = 3M_\pi\sqrt{1 + \Delta}$ and $x_r = (E^2 - 2E\omega_r + M_\pi^2)/4M_\pi^2$] dependence on Δ to the order needed. We will show in section 4.3.2 that naïvely expanding $H(x_r)$ this way is valid, despite the non-analyticity of H.

The remaining integral over r can be rewritten as an integral over ω_r. Introducing the variable z, defined via $\omega_r = M_\pi(1 + 2z^2)$ so that $r^2 \mathrm{d}r/[2\omega_r(2\pi)^3] = M_\pi^2 z^2 \sqrt{1 + z^2} \, \mathrm{d}z/(2\pi^3)$, leads to integrals of the type

$$H_{m,n} \equiv \frac{1}{\pi^2} \int_0^{1/\sqrt{3}} \mathrm{d}z \, \frac{\sqrt{1 + z^2}}{z^m} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \big[H^2(x) \big] \, . \quad \text{(IV.4.28)}$$

Since $x = 1 - 3z^2$, the integration limits are $x = 1$ at $z = 0$ and $x = 0$ at $z = 1/\sqrt{3}$. Due to H and all its derivatives vanishing at $x = 0$, the E-dependence in the upper limit of

---

[20]The general case can be compactly written as

$$\int \mathrm{d}\cos\theta \, \mathrm{d}\phi \prod_{a=1}^{2n} (\boldsymbol{p}_a \cdot \boldsymbol{r}) = 4\pi \frac{r^{2n}}{(2n+1)!!} \delta^{i_1 i_2 \cdots i_{2n}} p_1^{i_1} p_2^{i_2} \cdots p_n^{i_{2n}}$$

(the case with an odd number of $\boldsymbol{p}_a$ vanishes by symmetry). Repeated indices are summed, and $\delta^{i_1 \cdots i_n}$ generalizes the Kronecker δ to the totally symmetric tensor obtained by summing all distinct ways of distributing the indices $i_1, \ldots, i_{2n}$ among n Kronecker δ's, i.e., $\delta^{ijkl} = \delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}$, etc.

the integral is not captured by the expansion, meaning this is an expansion of asymptotic nature. However, as can be seen in the upper-right panel of fig. IV.6, the expansion to quadratic order approximates the full numerical result very well for the range of energies of interest.

We will return to the evaluation of $H_{m,n}$ in section 4.3.2. For now, we use it to state the result, which to second order in the threshold expansion is

$$
\frac{F_\pi^6}{M_\pi^4} \operatorname{Re} \mathcal{D}^{(u,u)\mathrm{BH}} = \left(\tfrac{3}{2}H_{0,0} - \tfrac{1}{4}H_{2,0}\right) + \left(\tfrac{21}{4}H_{0,1} + \tfrac{1}{8}H_{4,0}\right)\Delta + \left(\tfrac{81}{16}H_{0,0} - \tfrac{9}{32}H_{2,0} - \tfrac{3}{32}H_{4,0}\right)\Delta_3^S
$$
$$
+ \left(\tfrac{9}{16}H_{0,0} + \tfrac{15}{32}H_{2,0} - \tfrac{3}{32}H_{4,0}\right)\tilde{t}_{33} + \left(\tfrac{63}{8}H_{0,2} - \tfrac{33}{64}H_{0,1} - \tfrac{13}{256}H_{6,0} - \tfrac{19}{64}H_{4,0}\right)\Delta^2
$$
$$
+ \left(\tfrac{63}{4}H_{0,1} + \tfrac{3}{64}H_{6,0} + \tfrac{3}{32}H_{4,0}\right)\Delta\Delta_3^S + \left(\tfrac{9}{128}H_{6,0} - \tfrac{63}{128}H_{4,0} + \tfrac{27}{64}H_{2,0}\right)\Delta\tilde{t}_{33}
$$
$$
+ \left(-\tfrac{9}{32}H_{4,0} + \tfrac{81}{32}H_{2,0} - \tfrac{81}{16}H_{2,0}\right)\left(\Delta_3^S\right)^2 + \left(\tfrac{891}{32}H_{0,0} - \tfrac{243}{64}H_{2,0} - \tfrac{9}{64}H_{4,0}\right)\Delta_3\Delta_3'
$$
$$
+ \left(-\tfrac{9}{128}H_{6,0} + \tfrac{9}{64}H_{4,0} + \tfrac{189}{128}H_{2,0} + \tfrac{81}{64}H_{0,0}\right)\Delta_3^S\tilde{t}_{33}
$$
$$
+ \left(-\tfrac{27}{640}H_{6,0} + \tfrac{27}{160}H_{4,0} + \tfrac{297}{640}H_{2,0} + \tfrac{81}{320}H_{0,0}\right)\tilde{t}_{33}^2 . \tag{IV.4.29}
$$

This is the same expansion as in eq. (IV.4.1), but only a subset of the terms is needed. The integration-by-parts relation

$$
H_{m,n+1} + H_{m-2,n+1} = \frac{1}{6}\Big[(2-m)H_{m,n} - (m+1)H_{m+2,n}\Big]
$$
$$
- \frac{1}{6}\left[\left(f'_{m-1}(z) + f'_{m+1}(z)\right)\frac{\mathrm{d}^n}{\mathrm{d}x^n}H^2(x)\right]_0^{1/\sqrt{3}} , \tag{IV.4.30}
$$

where f_n is defined in eq. (IV.4.32), has been used extensively in simplifying eq. (IV.4.29); the "surface terms" [the second line of eq. (IV.4.30)] vanish identically, but are revisited in section 4.3.3. After symmetrization over all 9 possibilities, analogous to eq. (IV.4.3), we get

$$
\frac{F_\pi^6}{M_\pi^4} \operatorname{Re} \mathcal{D}^{\mathrm{BH}} = \left[\tfrac{27}{2}H_{0,0} - \tfrac{9}{4}H_{2,0}\right] + \Delta\left[\tfrac{117}{4}H_{0,0} - \tfrac{21}{8}H_{2,0} + \tfrac{3}{4}H_{4,0} + \tfrac{189}{4}H_{0,1}\right]
$$
$$
+ \Delta^2\left[\tfrac{243}{160}H_{0,0} + \tfrac{2241}{320}H_{2,0} - \tfrac{423}{160}H_{4,0} - \tfrac{369}{1280}H_{6,0} + \tfrac{5751}{64}H_{0,1} + \tfrac{567}{8}H_{0,2}\right]
$$
$$
+ \Delta_A\left[-\tfrac{891}{64}H_{0,0} + \tfrac{1161}{128}H_{2,0} - \tfrac{45}{64}H_{4,0} - \tfrac{9}{128}H_{6,0}\right]
$$
$$
+ \Delta_B\left[\tfrac{81}{320}H_{0,0} + \tfrac{297}{640}H_{2,0} + \tfrac{27}{160}H_{4,0} - \tfrac{27}{640}H_{6,0}\right] , \tag{IV.4.31}
$$

where $\mathcal{D}^{\mathrm{BH}} = \mathcal{S}\{\mathcal{D}^{(u,u)\mathrm{BH}}\}$ is the bull's head term in eq. (IV.2.39).

4.3.2 Hadamard finite-part integration

Equation (IV.4.28) for $n = 0$ and $m > 0$ has a troublesome singularity in the endpoint $z = 0$, so it is not possible to naïvely apply the Cauchy principal value to evaluate $H_{m,n}$. However, it is possible to use the *Hadamard finite-part* prescription. Let us present its

(rather simple) application before turning to the question of its validity. We define the functions $f_m(z)$ via

$$\frac{\mathrm{d}}{\mathrm{d}z} f_m(z) = \frac{1}{\pi^2} \frac{\sqrt{1+z^2}}{z^m}. \tag{IV.4.32}$$

Using partial integration on eq. (IV.4.28), we arrive at

$$H_{m,0} = f_m(z)H^2(x)\big|_{z=0}^{1/\sqrt{3}} - \int_0^{1/\sqrt{3}} \mathrm{d}z\, f_m(z)(-6z)\frac{\mathrm{d}}{\mathrm{d}x}H^2(x), \tag{IV.4.33}$$

an expression that is independent of the integration constant in $f_m(z)$. The $z = 0$ limit of the first term is singular, whereas the second term (the integral) is non-singular since derivatives of $H(x)$ vanish exponentially as $z \to 0$ or $x \to 1$. The $z = 1/\sqrt{3}$ or $x = 0$ limit vanishes by construction since H and its derivatives all vanish at $x = 0$. The Hadamard finite part of $H_{m,0}$ is obtained by dropping the singular $z = 0$ limit, and if the prescription is valid, $\mathcal{D}^{(u,u)\mathrm{BH}}$ is obtained by replacing all divergent $H_{m,n}$ with their finite parts.

It is at first not obvious that the Hadamard finite-part prescription is valid in our case, since the standard proofs involve complex integrals that break down due to the non-analyticity of $H(x)$. However, ref. [78] presents a proof using only smoothness criteria, which our integrands do satisfy. It also requires $m > 1$, but $H_{0,0}$ is non-singular, and we do not need $H_{1,0}$ (which only has an integrable singularity). Lastly, the prescription of course requires $\mathcal{D}^{(u,u)\mathrm{BH}}$ to be finite, but we know from section 4.2 that such is the case, at least for the real part: $\operatorname{Re}\mathcal{M}_3$ lacks the divergences that $\mathcal{D}^{(u,u)\mathrm{BH}}$ would subtract.

As a last remark, we note that the Hadamard finite-part integration validates the naïve threshold expansion used to obtain eq. (IV.4.28), which involved Taylor expanding $H(x)$. This Taylor series converges for $0 < x < 1$, but the convergence is extremely poor in the vicinity of the essential singularities at the endpoints. However, after applying eq. (IV.4.33), all integrands contain a derivative of $H(x)$, which is exponentially suppressed near those endpoints. Therefore, nothing remains that is sensitive to the Taylor expansion in the region where it converges poorly. Thus, the non-analyticity of $H(x)$ causes no problems for our method.

4.3.3 Analytic approximation

The result eq. (IV.4.33) is sufficient to obtain a numerical result for $\mathcal{D}^{\mathrm{BH}}$: One simply evaluates the Hadamard finite part of $H_{m,n}$ numerically. However, it is possible, at least for a wide class of functions $H(x)$, to find an analytic approximation to eq. (IV.4.28) that approximates $\mathcal{D}^{\mathrm{BH}}$ well, leaving a cutoff-dependent residue that must be evaluated numerically. Doing this, we are able to express $\mathcal{K}_{\mathrm{df},3}$ almost entirely as an analytic, cutoff-independent expression, with small numerical cutoff-dependent corrections.

Letting $H^2(x) \equiv 1 + \widetilde{H}^2(x)$, we get

$$H_{m,n} = \widetilde{H}_{m,n} + \delta_{n,0}\frac{1}{\pi^2}\int_0^{1/\sqrt{3}} \mathrm{d}z\, \frac{\sqrt{1+z^2}}{z^m}, \tag{IV.4.34}$$

where $\widetilde{H}_{m,n}$ is obtained by substituting $H(x) \to \widetilde{H}(x)$ in eq. (IV.4.28), and the remaining integral can be evaluated analytically, taking the Hadamard finite part when $m \neq 0$. Thus, making the choice $f_0(0) = 0$, we have

$$H_{m,n} = \widetilde{H}_{m,n} + f_m(1/\sqrt{3})\,\delta_{n,0}\,. \tag{IV.4.35}$$

The analytic approximation is then obtained by setting $\widetilde{H}_{m,n} = 0$, which allows everything to be expressed in terms of $f_m \equiv f_m(1/\sqrt{3})$. We may thus rewrite eq. (IV.4.31) as follows:

$$\begin{aligned}
\frac{F_\pi^6}{M_\pi^4}\,\mathrm{Re}\,\mathcal{D}^{\mathrm{BH}} =\ & \left[\tfrac{27}{2}f_0 - \tfrac{9}{4}f_2 + \mathcal{D}_0\right] + \Delta\left[36f_0 - 6f_2 - \tfrac{15}{16}f_4 + \mathcal{D}_1\right] \\
& + \Delta^2\left[\tfrac{2313}{160}f_0 + \tfrac{171}{320}f_2 - \tfrac{1677}{640}f_4 - \tfrac{519}{1280}f_6 + \mathcal{D}_2\right] \\
& + \Delta_{\mathrm{A}}\left[-\tfrac{891}{64}f_0 + \tfrac{1161}{128}f_2 - \tfrac{45}{64}f_4 - \tfrac{9}{128}f_6 + \mathcal{D}_{\mathrm{A}}\right] \\
& + \Delta_{\mathrm{B}}\left[\tfrac{81}{320}f_0 + \tfrac{297}{640}f_2 + \tfrac{27}{160}f_4 - \tfrac{27}{640}f_6 + \mathcal{D}_{\mathrm{B}}\right]. \tag{IV.4.36}
\end{aligned}$$

Here, $\mathcal{D}_X$ are H-dependent numerical corrections stemming from $\widetilde{H}_{m,n}$. They are defined by the requirement that eq. (IV.4.36) equals eq. (IV.4.31), and their values are given in eq. (IV.3.2). In appendix A, we investigate the dependence of the $\mathcal{D}_X$ upon the choice of cutoff function. Note that $\widetilde{H}_{m,n} \to 0$ is not a good approximation for individual $H_{m,n}$, but as the smallness of $\mathcal{D}_X$ shows, it clearly works at the level of $\mathcal{D}^{\mathrm{BH}}$ for our choice of H.

We stress that eq. (IV.4.36) is *not* obtained by substituting $H_{m,n} \to f_m \delta_{n,0}$ in eq. (IV.4.31), because the surface terms in eq. (IV.4.30) do not vanish for $\widetilde{H}_{m,n}$ and $f_m \delta_{n,0}$ separately. In other words, $H_{m,n} \to f_m \delta_{n,0}$ must be applied before eq. (IV.4.30).

However, looking at eq. (IV.4.30) for the $f_m \delta_{n,0}$ terms alone, we find that

$$(m+1)f_{m+2}(z) + (m-2)f_m(z) = -(1+z^2)f'_{m+1}(z)\,. \tag{IV.4.37}$$

This allows eq. (IV.4.30) to be restated in a more symmetrical form as

$$\begin{aligned}
H_{m,n+1} + H_{m-2,n+1} =\ & \frac{1}{6}\left[(2-m)H_{m,n} - (m+1)H_{m+2,n}\right] \\
& - \frac{1}{6}\left[\left[(2-m)f_m(z) - (m+1)f_{m+2}(z)\right]\frac{\mathrm{d}^n}{\mathrm{d}x^n}H^2(x)\right]_0^{1/\sqrt{3}}, \tag{IV.4.38}
\end{aligned}$$

but more importantly, since $f'_{m+1}(1/\sqrt{3})$ is just a rational number times κ, we can use eq. (IV.4.37) to reduce eq. (IV.4.36) entirely in terms of κ and $f_0 \equiv f_0(1/\sqrt{3}) = \tfrac{4}{3}\kappa(4 + 3\log 3)$:

$$\begin{aligned}
\frac{F_\pi^6}{M_\pi^4}\,\mathrm{Re}\,\mathcal{D}^{\mathrm{BH}} =\ & \left[96\kappa + 9f_0 + \mathcal{D}_0\right] + \Delta\left[296\kappa + 24f_0 + \mathcal{D}_1\right] + \Delta^2\left[\tfrac{5661}{50}\kappa + \tfrac{621}{40}f_0 + \mathcal{D}_2\right] \\
& + \Delta_{\mathrm{A}}\left[-\tfrac{1764}{5}\kappa + \tfrac{135}{32}f_0 + \mathcal{D}_{\mathrm{A}}\right] + \Delta_{\mathrm{B}}\left[-\tfrac{612}{25}\kappa + \tfrac{189}{160}f_0 + \mathcal{D}_{\mathrm{B}}\right]. \tag{IV.4.39}
\end{aligned}$$

4.3.4 Direct numerical evaluation

For a cross-check of the above results and for determining $\mathcal{D}^{(u,u)\mathrm{BH}}$ further away from threshold, it is necessary to evaluate eq. (IV.4.23) numerically without expansion. It is possible to evaluate it directly using numerical integration with a suitably chosen small $\epsilon > 0$, but the singularities in the integrand can lead to poor convergence or errors. Nevertheless, we have successfully performed this direct integration with 20 digits of accuracy, more than sufficient to reproduce eq. (IV.3.2). In this subsection, however, we present a less numerically demanding approach, which complements and validates the direct evaluation. Another approach, making no assumptions about the smoothness of $\mathcal{M}_3$, is presented in appendix E.

We note that the singularities in the integrand of eq. (IV.4.23) occur only where $H(x_r) = 1$,[21] so we separate

$$\int_r H^2(x_r) G(r; p_+, k_+) = \int_r \widetilde{H}^2(x_r)\, G(r; p_+, k_+)\, \theta(R - |\boldsymbol{r}|) + \int_r G(r; p_+, k_+)\, \theta(R - |\boldsymbol{r}|)\,, \tag{IV.4.40}$$

for some suitable cutoff R such that $H(x_r) = 0$ when $|\boldsymbol{r}| > R$. Here, θ is the Heaviside step function and $\widetilde{H}^2(x_r) \equiv H^2(x_r) - 1$ is the same as in section 4.3.1. The first term on the right-hand side is free from singularities and safe to evaluate numerically, while the second term, now free from $H(x_r)$, admits further simplification. This is easier in the CMF of $p_+ + k_+$, similar to the Breit frame in scattering, which we mark by $\diamond$. Setting up the $\boldsymbol{r}^\diamond$-integration in spherical coordinates with suitably aligned axes, i.e.,

$$\boldsymbol{r}^\diamond = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r\cos\theta)\,,$$
$$-\vec{p}_+^\diamond = \vec{k}_+^\diamond = (0, 0, q)\,, \qquad \boldsymbol{P}^\diamond = (0, Q\sin\gamma, Q\cos\gamma)\,, \tag{IV.4.41}$$

where we reuse the symbol r as $|\boldsymbol{r}^\diamond|$, we find

$$G^\diamond(r; p_+, k_+) = \frac{a_1 + a_2 \cos\theta + a_3 \sin\theta \sin\phi}{(b_1 - c\cos\theta)(b_2 + c\cos\theta)}\,, \tag{IV.4.42}$$

where

$$
\begin{aligned}
a_1 &= E^2 - M_\pi^2 - 2E^\diamond \omega_r\,, & a_2 &= 2rQ\cos\gamma\,, & a_3 &= 2rQ\sin\gamma\,, \\
b_1 &= p_+^2 - 2p_{+0}^\diamond \omega_r + i\epsilon\,, & b_2 &= k_+^2 - 2k_{+0}^\diamond \omega_r + i\epsilon\,, & c &= 2rq\,.
\end{aligned}
\tag{IV.4.43}
$$

Equation (IV.4.42) is a convenient parametrization for evaluating the first term in eq. (IV.4.40), which must be performed in the same frame as the second in order for the integration limits to match. More importantly it allows the angular integrals in the second term to be

---

[21]This follows from the construction of $\mathcal{K}_{\mathrm{df},3}$, but can also be proven directly as follows. At the pole where $(p_+ - r)^2 = M_\pi^2$,

$$4M_\pi^2 x_r = (P - r)^2 = (p_3 + p_+ - r)^2 = p_3^2 + (p_+ - r)^2 + 2p_3 \cdot (p_+ - r) = 2M_\pi^2 + 2\omega_{p_3}\omega_{p_+-r}\,,$$

where in the last equality we dropped the spatial part by going to the rest frame of either momentum. We have $\omega_{p_3} \geq m$ and $|\omega_{p_+-r}| \geq m$, so as long as $(p_+ - r)$ has positive energy, this proves that $x_r \geq 1$. At threshold, $\omega_{p_+-r} = 2m - m$ is positive, and since it is an analytic function of the kinematics, there is a path from threshold to any other configuration, and such a path clearly does not involve a switch to the negative-energy branch. The same holds for $(k_+ - r)^2 = M_\pi^2$.

performed analytically, leaving

$$\int_r G^\diamond(r; p_+, k_+)\, \theta(R - r) = \int_0^R \frac{2\pi r^2 \mathrm{d}r}{2\omega_r^\diamond (2\pi)^3}\, g(a_1, a_2; b_1, b_2; c)\,, \qquad \text{(IV.4.44)}$$

where

$$g(a_1, a_2; b_1, b_2; c) = \begin{cases} \dfrac{2a_1}{b_1 b_2}\,, & \text{if } c = 0\,, \\[2ex] \displaystyle\sum_{i=1,2} \dfrac{A_i}{c(b_1 + b_2)} \big[\log(b_i + c) - \log(b_i - c)\big] & \text{otherwise}\,, \end{cases} \qquad \text{(IV.4.45)}$$

with

$$A_1 = a_1 + \frac{a_2 b_1}{c}\,, \qquad A_2 = a_1 - \frac{a_2 b_2}{c}\,. \qquad \text{(IV.4.46)}$$

There are singularities at $b_1 + b_2 = 0$ and $b_i = \pm c$, both of which are regulated by the $+i\epsilon$ from the propagators. Even without regulation, the singularities are integrable, except at threshold where actual divergences develop at $b_1 + b_2 \to 0$ and $r \to 0$. These cancel to keep $\mathcal{D}^{(u,u)\mathrm{BH}}$ finite, but nevertheless present a numerical problem.

We have used three successful approaches to numerically evaluating eq. (IV.4.44), providing cross-checks against each other, against the brute-force evaluation of eq. (IV.4.23), and against the semi-analytic threshold expansion in the previous sections.

One method is to keep ϵ small but finite, giving an integrand with narrow spikes rather than singularities, which is numerically manageable with the right precautions. For sufficiently small ϵ, the ϵ-dependence of the result is very weak, and the $\epsilon \to 0$ limit is well approximated. It suffers stability issues near threshold, and we find that $\sqrt{\epsilon} \lesssim \Delta$ is required.

Another method is to deform the integration path into the complex plane, using, e.g., $r = z - i\alpha z(1 - z)$ for $z \in [0, R]$, where $\alpha > 0$ is arbitrary. With reasonably large α (but smaller than $4M_\pi$ to avoid the $\omega_r = 0$ pole at $r = iM_\pi$), the path avoids the singularities by a wide margin, allowing $\epsilon = 0$ and giving a smooth integrand that is easy to integrate. The deformation crosses some branch cuts if near thresholds, so it is only usable for the real part, and it cannot reach the threshold limit due to the divergence at $r \to 0$.

The third method is to express the $\epsilon \to 0$ limit in terms of Cauchy principal values, i.e., we replace the arguments of the logarithms in eq. (IV.4.45) with their absolute values. This approach is full of subtleties, requiring careful consideration of the details in ref. [77]. Basically, in addition to the principal-value part, we need to add the naïve double-pole contribution multiplied by a factor of 2.

4.3.5 The full bull's head subtraction

Evaluating f_0 in eq. (IV.4.39), the contributions of $\mathcal{D}^{\mathrm{BH}}$ to $\mathcal{K}_{\mathrm{df},3}$ are

$$\frac{F_\pi^6}{M_\pi^6}\, \mathcal{K}_0 \supset -36\kappa(4 + \log 3) - \mathcal{D}_0\,, \qquad \text{(IV.4.47a)}$$

$$\frac{F_\pi^6}{M_\pi^6}\,\mathcal{K}_1 \supset -8\kappa(53 + 12\log 3) - \mathcal{D}_1\,, \tag{IV.4.47b}$$

$$\frac{F_\pi^6}{M_\pi^6}\,\mathcal{K}_2 \supset -\frac{27\kappa}{50}(363 + 115\log 3) - \mathcal{D}_2\,, \tag{IV.4.47c}$$

$$\frac{F_\pi^6}{M_\pi^6}\,\mathcal{K}_A \supset \frac{9\kappa}{40}(1468 - 75\log 3) - \mathcal{D}_A\,, \tag{IV.4.47d}$$

$$\frac{F_\pi^6}{M_\pi^6}\,\mathcal{K}_B \supset \frac{9\kappa}{200}(404 - 105\log 3) - \mathcal{D}_B\,. \tag{IV.4.47e}$$

The numerical values of these contributions are given in table IV.1. From there, we find that $\mathcal{D}_X$ are indeed quite small compared to the full bull's head contributions: With the exception of $\mathcal{D}_2$, they differ by more than an order of magnitude.

It is perhaps unexpected that the analytic approximation, which effectively amounts to the invalid cutoff choice $H(x) = \theta(x-1)$, should be so accurate. However, as seen in section 4.3.4, the integral is dominated by the pole at small r, and $H(x_r)$ is very close to 1 in the vicinity of this pole. Conversely, the region where $H(x_r)$ differs meaningfully from 1 contributes very little. Section A looks more closely at this.

4.4 OPE diagrams
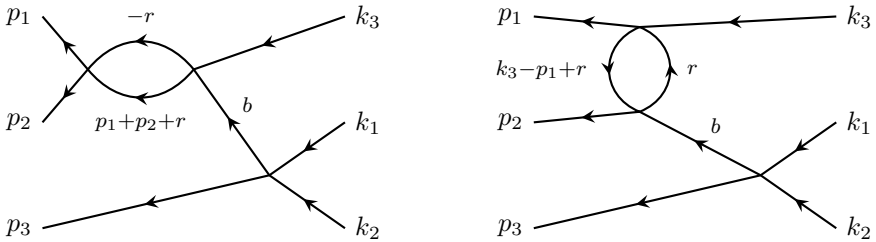
Figure IV.10: Examples of OPE NLO diagrams. There are also diagrams where the loop appears on the lower right.

The contributions of the OPE diagrams, such as those in fig. IV.10, can be computed analytically by directly performing the pertinent subtractions. Unlike in the LO case described in section 2.4, an additional challenge here is how to deal with the higher two-pion partial waves in the two-pion scattering subamplitude. While all partial waves appear in this subamplitude, it is possible to show that only two types of contributions are relevant at quadratic order: purely s-wave ones ($\ell = \ell' = 0$) and those with d-wave on one side ($\ell = 2$, $\ell' = 0$ or vice versa).

The starting point is the unsymmetrized version of the "master equation" (IV.2.27), which we now restrict to OPE contributions:

$$\mathcal{K}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3)_{\ell'm',\ell m} = \operatorname{Re}\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3)_{\ell'm',\ell m}\,. \tag{IV.4.48}$$

In turn, we need the definition of the divergence-free three-particle amplitude

$$\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3)_{\ell'm',\ell m}$$
$$= \mathcal{M}_3^{(u,u)\mathrm{NLO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3)_{\ell'm',\ell m} - \mathcal{D}^{(u,u)\mathrm{NLO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3)_{\ell'm',\ell m}, \quad \text{(IV.4.49)}$$

where $\mathcal{M}_3^{(u,u)\mathrm{NLO,OPE}}$ is given by one-particle-reducible diagrams like those shown in fig. IV.10. Before projection onto pair angular momenta, we have

$$\mathcal{M}_3^{(u,u)\mathrm{NLO,OPE}} = -\mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(\bar{s}_2', t_2', u_2') \frac{1}{b^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\bar{s}_2, t_2, u_2) \quad \text{(IV.4.50)}$$

$$- \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\bar{s}_2', t_2', u_2') \frac{1}{b^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(\bar{s}_2, t_2, u_2), \quad \text{(IV.4.51)}$$

where b is the momentum of the exchanged particle and

$$\begin{aligned} \bar{s}_2' &\equiv (p_1 + p_2)^2 - 4M_\pi^2, & t_2' &= (p_1 - k_3)^2, & u_2' &= (p_2 - k_3)^2, \\ \bar{s}_2 &= (k_1 + k_2)^2 - 4M_\pi^2, & t_2 &= (k_1 - p_3)^2, & u_2 &= (k_2 - p_3)^2. \end{aligned} \quad \text{(IV.4.52)}$$

The off-shell amplitudes have a single leg off shell (that of the exchanged particle), thus

$$\bar{s}_2 + t_2 + u_2 = b^2 - M_\pi^2 \equiv \bar{b}^2 = \bar{s}_2' + t_2' + u_2'. \quad \text{(IV.4.53)}$$

The quantities $\bar{s}_2$, $\bar{s}_2'$, and $\bar{b}^2$ are defined so that they vanish at threshold (as do t_2, u_2, t_2' and u_2'), and are thus convenient in a threshold expansion.

For the subtraction term, $\mathcal{D}^{(u,u)\mathrm{NLO,OPE}}$, we must use the $\{\ell, m\}$ basis, where (keeping the indices implicit) we have

$$\mathcal{D}^{(u,u)\mathrm{NLO,OPE}}(\boldsymbol{p}_3, \boldsymbol{k}_3)$$
$$= -\mathcal{M}_2^{\mathrm{NLO}}(\boldsymbol{p}_3) G^\infty(\boldsymbol{p}_3, \boldsymbol{k}_1) \mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{k}_3) - \mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{p}_3) G^\infty(\boldsymbol{p}_3, \boldsymbol{k}_1) \mathcal{M}_2^{\mathrm{NLO}}(\boldsymbol{k}_3), \quad \text{(IV.4.54)}$$

recalling eqs. (IV.2.8) and (IV.2.12) for relevant definitions. The absence of the subscript "off" on the factors of $\mathcal{M}_2$ in eq. (IV.4.54) indicates that these amplitudes are all evaluated on shell.

Next, we note that $\mathcal{M}_2^{\mathrm{LO}}$ contains only s-waves and is purely real,

$$F_\pi^2 \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(s_2, t_2, u_2) = -2M_\pi^2 - \bar{s}_2 + \bar{b}^2. \quad \text{(IV.4.55)}$$

Since the poles at $\bar{b}^2 = 0$ in $\mathcal{M}_3^{(u,u)}$ and $\mathcal{D}^{(u,u)}$ cancel, we are able to set $\epsilon \to 0$. Thus, to obtain the real part of $\mathcal{M}_{\mathrm{df},3}^{(u,u)}$, we can use the expressions above with $\mathcal{M}_2^{\mathrm{NLO}}$ replaced by its real part. Furthermore, since we are matching to the threshold expansion of $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$, keeping up to quadratic terms, we can expand $\mathrm{Re}\,\mathcal{M}_2^{\mathrm{NLO}}$ about threshold and drop terms of higher-than-cubic order in the quantities that vanish at threshold. Cubic order is required because of the $\bar{b}^2$ in the denominator coupled with the fact that $\mathcal{M}_2^{\mathrm{LO}}$ does not vanish at

threshold. It then turns out, as shown explicitly below, that the only term that contains other than s-waves is that proportional to $t_2 u_2$. Thus, we write

$$F_\pi^4 \operatorname{Re} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(\bar{s}_2, t_2, u_2) = F_\pi^4 \operatorname{Re} \mathcal{M}_{2s,\mathrm{off}}^{\mathrm{NLO}}(\bar{s}_2, t_2, u_2) + e_{tu} t_2 u_2 + e'_{tu} \frac{(\bar{b}^2 - \bar{s}_2)}{M_\pi^2} t_2 u_2 \,,$$

$$(\text{IV.4.56})$$

where the first term on the right-hand side contains the purely s-wave contributions. We treat the purely s-wave and the "$t_2 u_2$" terms separately. The coefficients e_{tu} and e'_{tu} are real by construction.

4.4.1 Expression for $\operatorname{Re} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}$

As explained in section 2.3, we use the choice of off-shell two-particle amplitude given in ref. [60]. For the $I = 2$ channel, the NLO amplitude is

$$F_\pi^4 \mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(s, t, u) = A^{(4)}(t_2, u_2, s_2) + A^{(4)}(u_2, s_2, t_2) \,, \qquad (\text{IV.4.57})$$

where

$$\begin{aligned} A^{(4)}(s_2, t_2, u_2) &= d_1(t_2 - u_2)^2 + d_2 M_\pi^2 s_2 + d_3 s_2^2 + d_4 M_\pi^4 \\ &\quad + f_1(s_2)\bar{J}(s_2) + f_2(s_2, t_2)\bar{J}(t_2) + f_2(s_2, u_2)\bar{J}(u_2) \,, \end{aligned} \qquad (\text{IV.4.58})$$

$$f_1(s) = d_5 s_2^2 + d_6 M_\pi^2 s_2 + d_7 M_\pi^4 \,, \qquad (\text{IV.4.59})$$

$$f_2(s, t) = d_8 t_2^2 + d_9 t_2 M_\pi^2 + d_{10} s_2 M_\pi^2 + d_{11} s_2 t_2 + d_{12} M_\pi^4 \,, \qquad (\text{IV.4.60})$$

with the function $\bar{J}$ defined in eq. (IV.2.3). We revert to the standard Mandelstam variable s_2 rather than $\bar{s}_2 = s_2 - 4M_\pi^2$ in order to simplify the comparison with ref. [60]. The constants in the above expressions are [60]

$$\begin{aligned} d_1 &= -\tfrac{5}{36}\kappa - \tfrac{1}{6}L + \tfrac{1}{2}\ell_2^{\mathrm{r}} \,, \\ d_2 &= (N - \tfrac{29}{9})\kappa + (N - \tfrac{11}{3})L - 8\ell_1^{\mathrm{r}} + 2\ell_4^{\mathrm{r}} \,, \\ d_3 &= (\tfrac{11}{12} - \tfrac{1}{2}N)\kappa + (1 - \tfrac{1}{2}N)L + 2\ell_1^{\mathrm{r}} + \tfrac{1}{2}\ell_2^{\mathrm{r}} \,, \\ d_4 &= (\tfrac{20}{9} - \tfrac{1}{2}N)\kappa + (\tfrac{8}{3} - \tfrac{1}{2}N)L + 8\ell_1^{\mathrm{r}} + 2\ell_3^{\mathrm{r}} - 2\ell_4^{\mathrm{r}} \,, \\ d_5 &= \tfrac{1}{2}N - 1 \,, \qquad d_6 = (3 - N) \,, \qquad d_7 = \tfrac{1}{2}N - 2 \,, \\ d_8 &= \tfrac{1}{3} \,, \qquad d_9 = -\tfrac{5}{3} \,, \qquad d_{10} = -\tfrac{2}{3} \,, \qquad d_{11} = \tfrac{1}{6} \,, \qquad d_{12} = \tfrac{7}{3} \,, \end{aligned} \qquad (\text{IV.4.61})$$

with $N = 3$, and remaining definitions are given below eq. (IV.4.17). Imaginary parts arise only from the $\bar{J}$ function. Its real part is analytic above threshold, and the expansions that we need are given in eq. (IV.4.11). Combining these results, we obtain

$$\begin{aligned} F_\pi^4 \operatorname{Re} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(\bar{s}_2, t_2, u_2) &= e_0 M_\pi^4 + e_1 M_\pi^2 \bar{s}_2 + e_2 \bar{s}_2^2 + e_3 M_\pi^2 \bar{b}^2 \\ &\quad + e_4 \bar{s}_2 \bar{b}^2 + e_5 (\bar{b}^2)^2 + e_{tu} t_2 u_2 \\ &\quad + e_6 \frac{\bar{s}_2^2 \bar{b}^2}{M_\pi^2} + e_7 \frac{\bar{s}_2 (\bar{b}^2)^2}{M_\pi^2} + e_8 \frac{(\bar{b}^2)^3}{M_\pi^6} + e'_{tu} \frac{(\bar{b}^2 - \bar{s}_2)}{M_\pi^2} t_2 u_2 \,, \end{aligned} \qquad (\text{IV.4.62})$$

where the constants e_i and e_{tu} are known in terms of the d_i. Cubic terms $\bar{s}_2^3$ are not needed for the quadratic threshold expansion, since they give rise only to higher-order terms. The on-shell amplitude is obtained by setting $\bar{b} \to 0$.

4.4.2 Decomposition of $t_2 u_2$

As noted above, only the $t_2 u_2$ term in eq. (IV.4.62) contains nonzero angular momenta in the pair CMF. To show this explicitly, we consider the final-state pair (with momenta p_1 and p_2) for which

$$t_2' u_2' = \tfrac{1}{4}(\bar{s}_2' - \bar{b}^2)^2 - 4(\boldsymbol{a}_p^* \cdot \boldsymbol{k}_p^*)^2 \,, \tag{IV.4.63}$$

where $\boldsymbol{a}_p^*$ and $\boldsymbol{k}_p^*$ are the three-momenta $\boldsymbol{p}_1$ and $\boldsymbol{k}_3$ boosted to the CMF of the final-state pair. Their magnitudes are given by

$$|\boldsymbol{a}_p^*|^2 = q_{2,p}^{*2} = \tfrac{1}{4}\bar{s}_2' \,, \qquad k_p^{*2} = \frac{(s_2' - \bar{b}^2)^2}{4 s_2'} - M_\pi^2 \,. \tag{IV.4.64}$$

To pull out the d-wave part, we use

$$(\boldsymbol{a}_p^* \cdot \boldsymbol{k}^*)^2 = q_{2,p}^{*2} k^{*2} \left[\frac{8\pi}{15} \sum_m Y_{2m}^*(\hat{\boldsymbol{a}}_p^*) Y_{2m}(\hat{\boldsymbol{k}}_p^*) + \frac{1}{3} \right] \,. \tag{IV.4.65}$$

Thus, we can separate s- and d-wave parts as $t_2' u_2' = [t_2' u_2']_s + [t_2' u_2']_d$, where

$$[t_2' u_2']_s = \tfrac{1}{4}(\bar{s}_2' - \bar{b}^2)^2 - \tfrac{4}{3} q_{2,p}^{*2} k_p^{*2} \,, \qquad [t_2' u_2']_d = q_{2,p}^{*2} k_p^{*2} \frac{8\pi}{15} \sum_m Y_{2m}^*(\hat{\boldsymbol{a}}_p^*) Y_{2m}(\hat{\boldsymbol{k}}_p^*) \,. \tag{IV.4.66}$$

An analogous expression holds for $t_2 u_2$.

In the following, we will need the expansion of k_p^{*2} about threshold, which is given by

$$k_p^{*2} = \tfrac{1}{4}\bar{s}_2' - \tfrac{1}{2}\bar{b}^2 + \tfrac{1}{16}(\bar{b}^2)^2 + \dots \,, \tag{IV.4.67}$$

where the ellipsis contains terms of higher order that do not contribute at the order we work. Note that the partial-wave decomposition of the term $(\bar{b}^2 - \bar{s}_2) t_2 u_2$ is analogous since $(\bar{b}^2 - \bar{s}_2)$ is purely s-wave.

4.4.3 s-wave contributions

Including the $[t_2 u_2]_s$ and $(\bar{b}^2 - \bar{s}_2)[t_2 u_2]_s$ terms, the s-wave part of the real part of the NLO amplitude becomes

$$F_\pi^4 \operatorname{Re} \mathcal{M}_{2s,\text{off}}^{\text{NLO}}(\bar{s}_2, \bar{b}^2) = e_0 M_\pi^4 + e_1 M_\pi^2 \bar{s}_2 + e_2' \bar{s}_2^2 + e_3 M_\pi^2 \bar{b}^2 + e_4' \bar{s}_2 \bar{b}^2 + e_5' (\bar{b}^2)^2$$

$$+ e_6' \frac{\bar{s}_2^2 \bar{b}^2}{M_\pi^2} + e_7' \frac{\bar{s}_2 (\bar{b}^2)^2}{M_\pi^2} + e_8' \frac{(\bar{b}^2)^3}{M_\pi^2} + \dots \,, \tag{IV.4.68}$$

where

$$e_2' = e_2 + \tfrac{1}{6}e_{tu}\,, \qquad e_4' = e_4 - \tfrac{1}{3}e_{tu}\,, \qquad\qquad e_5' = e_5 + \tfrac{1}{4}e_{tu}\,,$$
$$e_6' = e_6 + \tfrac{1}{2}e_{tu}'\,, \qquad e_7' = e_7 - \tfrac{1}{48}e_{tu} - \tfrac{7}{12}e_{tu}'\,, \qquad e_8' = e_8 + \tfrac{1}{4}e_{tu}'\,. \tag{IV.4.69}$$

We thus have

$$\mathcal{R}_s^{\mathrm{NLO}}(\bar{s}_2) \equiv F_\pi^4\,\mathrm{Re}\,\mathcal{M}_{2s,\mathrm{on}}^{\mathrm{NLO}}(\bar{s}_2) = e_0 M_\pi^4 + e_1 M_\pi^2 \bar{s}_2 + e_2' \bar{s}_2^2 + \ldots\,, \tag{IV.4.70}$$

$$\delta_s^{\mathrm{NLO}}(\bar{s}_2,\bar{b}^2) \equiv F_\pi^4\,\mathrm{Re}\,\mathcal{M}_{2s,\mathrm{off}}^{\mathrm{NLO}}(\bar{s}_2) - F_\pi^4\,\mathrm{Re}\,\mathcal{M}_{2s,\mathrm{on}}^{\mathrm{NLO}}(\bar{s}_2)$$
$$= \bar{b}^2\left(e_3 M_\pi^2 + e_4'\bar{s}_2 + e_5'\bar{b}^2 + e_6'\frac{\bar{s}_2^2}{M_\pi^2} + e_7'\frac{\bar{s}_2\bar{b}^2}{M_\pi^2} + e_8'\frac{(\bar{b}^2)^2}{M_\pi^2} + \ldots\right). \tag{IV.4.71}$$

The LO result from eq. (IV.4.55) gives

$$\mathcal{R}_s^{\mathrm{LO}}(\bar{s}_2) \equiv F_\pi^2\,\mathcal{M}_{2s,\mathrm{on}}^{\mathrm{LO}}(\bar{s}_2) = -2M_\pi^2 - \bar{s}_2\,, \tag{IV.4.72}$$

$$\delta_s^{\mathrm{LO}}(\bar{b}^2) \equiv F_\pi^2\,\mathcal{M}_{2s,\mathrm{off}}^{\mathrm{LO}}(\bar{s}_2,\bar{b}^2) - F_\pi^2\,\mathcal{M}_{2s,\mathrm{on}}^{\mathrm{LO}}(\bar{s}_2) = \bar{b}^2\,. \tag{IV.4.73}$$

We can now perform the required subtraction. Since we are considering purely s-wave terms, the projection onto pair angular momenta is trivial, and we can work with Mandelstam variables. Thus, the contribution of s-wave two-particle amplitudes is

$$-\bar{b}^2 F_\pi^6 \mathcal{K}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE},s}$$
$$= \mathcal{R}_s^{\mathrm{NLO}}(\bar{s}_2')\delta_s^{\mathrm{LO}}(\bar{b}^2) \quad + \delta_s^{\mathrm{NLO}}(\bar{s}_2',\bar{b}^2)\mathcal{R}_s^{\mathrm{LO}}(\bar{s}_2) + \delta_s^{\mathrm{NLO}}(\bar{s}_2',\bar{b}^2)\delta_s^{\mathrm{LO}}(\bar{b}^2)$$
$$+ \mathcal{R}_s^{\mathrm{LO}}(\bar{s}_2')\delta_s^{\mathrm{NLO}}(\bar{s}_2,\bar{b}^2) + \delta_s^{\mathrm{LO}}(\bar{b}^2)\mathcal{R}_s^{\mathrm{NLO}}(\bar{s}_2) \quad + \delta_s^{\mathrm{LO}}(\bar{b}^2)\delta_s^{\mathrm{NLO}}(\bar{s}_2,\bar{b}^2)\,. \tag{IV.4.74}$$

Substituting the results above, we obtain

$$F_\pi^6 \mathcal{K}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE},s} = g_0^s M_\pi^4 + g_1^s M_\pi^2(\bar{s}_2' + \bar{s}_2) + g_2^s \bar{s}_2'\bar{s}_2 + g_3^s(\bar{s}_2' + \bar{s}_2)^2$$
$$+ g_4^s M_\pi^2 \bar{b}^2 + g_5^s \bar{b}^2(\bar{s}_2' + \bar{s}_2) + g_6^s(\bar{b}^2)^2 + \ldots\,, \tag{IV.4.75}$$

where the coefficients g_i^s are known in terms of the e_i, and the ellipsis indicates higher-order terms that are not needed.

The final step is to convert the variables to those used in the threshold expansion for $\mathcal{K}_{\mathrm{df},3}^{(u,u)}$, eq. (IV.4.1), using the results

$$\bar{s}_2' + \bar{s}_2 = 9M_\pi^2\Delta_3^S\,, \qquad \bar{s}_2'\bar{s}_2 = 81M_\pi^4\Delta_3'\Delta_3\,, \qquad \bar{b}^2 = 9M_\pi^2(\Delta_3^S - \Delta - \tilde{t}_{33})\,. \tag{IV.4.76}$$

In this way, we obtain the contributions to the coefficients c_1–c_{10} in eq. (IV.4.1) from the s-wave parts of $\mathcal{M}_2$.

4.4.4 d-wave contributions

First, we consider the e_{tu} term, which gives a contribution of $[tu]_d$ to the NLO matrix element. Here, we must use the $\{k,\ell,m\}$ basis for $\mathcal{K}_{\mathrm{df},3}$ since the subtraction is given in

this basis. Specifically, the d-wave contribution to the subtraction of eq. (IV.4.54) is, after recombining with spherical harmonics,

$$\text{Re}\,\mathcal{D}^{(u,u)\text{NLO,OPE},d}(\{p_i\},\{k_i\})$$
$$= -\sum_{m'}\sqrt{4\pi}Y_{2m'}(\hat{a}_p^*)\,\text{Re}\big\{\mathcal{M}_{2,\ell'=2}^{\text{NLO}}(\boldsymbol{p}_3)\big\}\left(\frac{k_p^*}{q_{2,p}^*}\right)^2\frac{\sqrt{4\pi}Y_{2m'}(\hat{k}_p^*)}{\bar{b}^2}\,\mathcal{M}_{2s}^{\text{LO}}(\boldsymbol{k}_3)+\leftrightarrow\,.$$
(IV.4.77)

Here, $\leftrightarrow$ indicates the term in which the roles of the LO and NLO vertices are interchanged and the notation is as in section 4.4.2. The factor of $(k_p^*/q_{2,p}^*)^2$ arises from G^∞, eq. (IV.2.12). This contribution to $\mathcal{D}^{(u,u)}$ is to be subtracted from

$$\text{Re}\,\mathcal{M}_3^{(u,u)\text{NLO,OPE},d}(\{p_i\},\{k_i\})$$
$$= -\sum_{m'}\sqrt{4\pi}Y_{2m'}(\hat{a}_p^*)\,\text{Re}\big\{\mathcal{M}_{2,\ell'=2,\text{off}}^{\text{NLO}}(\boldsymbol{p}_3)\big\}\frac{\sqrt{4\pi}Y_{2m'}(\hat{k}_p^*)}{\bar{b}^2}\,\mathcal{M}_{2s,\text{off}}^{\text{LO}}(\boldsymbol{k}_3)+\leftrightarrow\,.\quad(\text{IV.4.78})$$

The key observation now is that, using the decomposition of the tu term given by eqs. (IV.4.63) and (IV.4.65),

$$\mathcal{M}_{2,\ell'=2,\text{off}}^{\text{NLO}}(\boldsymbol{p}_3)=-\frac{8}{15}q_{2,p}^{*2}k_p^{*2}\qquad\text{and}\qquad\mathcal{M}_{2,\ell'=2}^{\text{NLO}}(\boldsymbol{p}_3)=-\frac{8}{15}q_{2,p}^{*4}\,,\qquad(\text{IV.4.79})$$

implying that

$$\mathcal{M}_{2,\ell'=2}^{\text{NLO}}(\boldsymbol{p}_3)\left(\frac{k_p^*}{q_{2,p}^*}\right)^2=\mathcal{M}_{2,\ell'=2,\text{off}}^{\text{NLO}}(\boldsymbol{p}_3)\,.\qquad(\text{IV.4.80})$$

In other words, the barrier factor from G^∞ converts the on-shell amplitude appearing in the subtraction term into exactly the off-shell amplitude. Thus, the subtraction only picks out the difference between on- and off-shell values of $F_\pi^2\mathcal{M}_2^{\text{LO}}$, given by $\delta_s^{\text{LO}}(\bar{b}^2)=\bar{b}^2$, and simply cancels the pole. One therefore obtains the contribution

$$F_\pi^6\mathcal{K}_{\text{df},3}^{(u,u)\text{NLO,OPE},d}\supset -e_{tu}\big([t_2'u_2']_d+[t_2u_2]_d\big)\,.\qquad(\text{IV.4.81})$$

The two terms on the right-hand side contain $\{\ell',\ell\}=\{2,0\}$ and $\{0,2\}$, respectively. To convert to our standard basis, we use

$$[tu]_d=tu-[tu]_s\,,\qquad(\text{IV.4.82})$$

with $[tu]_s$ given by eq. (IV.4.66), and observe that $t_2'u_2'+t_2u_2=81M_\pi^4Q_{tu}$ [see eq. (IV.4.2)], so that

$$F_\pi^6\mathcal{K}_{\text{df},3}^{(u,u)\text{NLO,OPE},d}\supset e_{tu}\big(-81M_\pi^4Q_{tu}+[t_2'u_2']_s+[t_2u_2]_s\big)\,.\qquad(\text{IV.4.83})$$

The $[tu]_s$ terms can be expanded in powers of $\bar{s}_2'$, $\bar{s}_2$, and $\bar{b}^2$, and lead to additional contributions of the form of eq. (IV.4.75), with $g_i^s\to g_i^d$. These can then be converted to the standard variables of the threshold expansion as explained above and thus contribute to c_1-c_{10} in eq. (IV.4.1).

Next, we consider the e'_{tu} term, which has the form $(\bar{b}^2 - \bar{s}_2)[tu]_d$. The analysis for the $\bar{s}_2[tu]_d$ part is the same as for $[tu]_d$ alone, leading to

$$F_\pi^6 \mathcal{K}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE},d} \supset -e'_{tu} \frac{\bar{s}_2}{M_\pi^2} \left(-81 M_\pi^4 \mathcal{Q}_{tu} + [t'_2 u'_2]_s + [t_2 u_2]_s\right). \tag{IV.4.84}$$

Recalling that the subtraction has already been done, we note that all terms are of too high order to contribute.

Finally, we consider the $\bar{b}^2[tu]_d$ part of the NLO amplitude. Since this vanishes on shell, there is no subtraction term and we easily find

$$F_\pi^6 \mathcal{K}_{\mathrm{df},3}^{(u,u)\mathrm{NLO,OPE},d} \supset 2e'_{tu} \left([t'_2 u'_2]_d + [t_2 u_2]_d\right) + \dots, \tag{IV.4.85}$$

where the overall factor of 2 comes from the value of the LO amplitude at threshold. The remainder of the analysis is as for the $[tu]_d$ term above, except that $e_{tu} \to 2e'_{tu}$.

4.4.5 The full OPE contribution

Combining s- and d-wave contributions computed in sections 4.4.3 and 4.4.4, we end up with expressions for c_1 through c_{10} and c_{14} in terms of the coefficients d_i; there are no contributions to c_{11}, c_{12}, and c_{13}. We can now symmetrize using eqs. (IV.4.3a) to (IV.4.3e). We find that the contributions of the OPE diagrams at NLO are

$$\frac{F_\pi^6}{M_\pi^6} \mathcal{K}_0 \supset 25\kappa + 78L - 72(8\ell_1^{\mathrm{r}} + 6\ell_2^{\mathrm{r}} + \ell_3^{\mathrm{r}} - 2\ell_4^{\mathrm{r}}), \tag{IV.4.86a}$$

$$\frac{F_\pi^6}{M_\pi^6} \mathcal{K}_1 \supset \frac{6831}{20}\kappa + 372L - 18(74\ell_1^{\mathrm{r}} + 67\ell_2^{\mathrm{r}} - 14\ell_4^{\mathrm{r}}), \tag{IV.4.86b}$$

$$\frac{F_\pi^6}{M_\pi^6} \mathcal{K}_2 \supset \frac{230481}{280}\kappa + 576L - 108(10\ell_1^{\mathrm{r}} + 11\ell_2^{\mathrm{r}}), \tag{IV.4.86c}$$

$$\frac{F_\pi^6}{M_\pi^6} \mathcal{K}_A \supset -\frac{53199}{560}\kappa + 45L + \frac{27}{2}(14\ell_1^{\mathrm{r}} - 17\ell_2^{\mathrm{r}}), \tag{IV.4.86d}$$

$$\frac{F_\pi^6}{M_\pi^6} \mathcal{K}_B \supset \frac{54171}{140}\kappa + 216L - 324(2\ell_1^{\mathrm{r}} + \ell_2^{\mathrm{r}}). \tag{IV.4.86e}$$

For comparison, the LO results from the OPE diagram in fig. IV.1a are

$$\frac{F_\pi^4}{M_\pi^4} \mathcal{K}_0 \supset 36, \qquad \frac{F_\pi^4}{M_\pi^4} \mathcal{K}_1 \supset 63. \tag{IV.4.87}$$

Like eq. (IV.4.18), this is specific to our off-shell convention (see section 2.3).

5   Conclusions and outlook

This work presents the NLO ChPT result for the isospin-3 three-particle K-matrix, $\mathcal{K}_{\mathrm{df},3}$, which parametrizes three-particle interactions in the RFT three-particle finite-volume for-

malism [13, 14]. In particular, we have focused on the leading five terms in the threshold expansion. To determine $\mathcal{K}_{\mathrm{df},3}$, we have used the three-pion amplitude calculated in ref. [60] combined with the relation between this amplitude and $\mathcal{K}_{\mathrm{df},3}$ derived in ref. [14]. The main results of this work are summarized in section 3.1 and, in particular, in eq. (IV.3.1).

Various simplifications play an important role in obtaining these results. The first is the result of eq. (IV.2.27) that, at NLO in ChPT, $\mathcal{K}_{\mathrm{df},3}$ is simply given by $\mathrm{Re}\,\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO}}$, rather than requiring the solution of integral equations. The second simplification is that, although various contributions to $\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO}}$ can be singular at threshold, these singularities are absent in the real part of the total result. This allows us to obtain analytic results for the threshold expansion for almost all parts, the exception being the cutoff-dependent parts of the integrals appearing in the BH subtraction. The latter turn out to be numerically small.

One of the motivations for this work was to address the substantial discrepancy between lattice results for $\mathcal{K}_{\mathrm{df},3}$ and the LO ChPT prediction [44, 47, 52]. Focusing on the results for the first two terms in the expansion of $\mathcal{K}_{\mathrm{df},3}$, namely $\mathcal{K}_0$ and $\mathcal{K}_1$, we find that the NLO corrections are able to resolve the large disagreement between lattice QCD and LO ChPT, thus increasing confidence on the extractions of $\mathcal{K}_{\mathrm{df},3}$ from lattice calculations. We observe, however, that NLO effects are somewhat large in these two quantities. Regarding the term in the threshold expansion of $\mathcal{K}_{\mathrm{df},3}$ that couples to d-waves, $\mathcal{K}_{\mathrm{B}}$, we find a sign disagreement between the lattice QCD result and NLO ChPT result. While we do not have a definitive answer, we stress that the NLO ChPT contribution is the leading effect for $\mathcal{K}_{\mathrm{B}}$. Potentially, NNLO effects could be large and account for the discrepancy.

Since we are using an expansion of $\mathcal{K}_{\mathrm{df},3}$ about threshold, it is important to verify its convergence. This is necessary since at NLO in ChPT, $\mathcal{K}_{\mathrm{df},3}$ has contributions to all orders in the threshold expansion due to the presence of loop integrals. To do so, in section 3.3, we address the validity of the truncation of $\mathcal{K}_{\mathrm{df},3}$ at quadratic order. We find that for a pion mass of $M_\pi = 340$ MeV (the heaviest used in ref. [52]), the corrections beyond quadratic order account only for a 20% of the total at the 5π inelastic threshold. The corrections are even smaller for lighter masses. Similarly, two-pion partial waves with $\ell > 2$ in OPE diagrams, which do not enter the threshold expansion at quadratic order, only add a negligible contribution to the full $\mathcal{K}_{\mathrm{df},3}$. We conclude that truncating the threshold expansion at quadratic order provides a good approximation to $\mathcal{K}_{\mathrm{df},3}$.

Another timely question that we address is the cutoff dependence of $\mathcal{K}_{\mathrm{df},3}$. All lattice QCD calculations using the RFT formalism have adopted the same choice of cutoff function. However, this choice is not unique. In appendix A, we discuss how the NLO ChPT result varies for different cutoffs. Overall, we find that for a wide set of cutoff functions, the dependence is small, provided that the function does not drop to zero very rapidly below the two-pion threshold.

The results of this work can be extended to other systems, higher orders, or other EFTs. For instance, in preparation for future lattice QCD calculations, $\mathcal{K}_{\mathrm{df},3}$ for other three-pion isospin channels could be derived since ref. [60] provides results for the six-pion amplitude for general isospin. A potential issue is that, due to the presence of resonances in two-particle subchannels and in the three-particle channel itself, the convergence of ChPT

might be poor. Examples include σ and ρ in two-particle and ω and h_1 in three-particle channels. $\mathcal{K}_{\text{df},3}$ could also be computed for other systems of mixed mesons at maximal isospin, such as $\pi^+\pi^+K^+$ and $K^+K^+\pi^+$. In this case, the full amplitude is not available yet, and the results from ref. [60] would need to be extended to SU(3) ChPT. This is a very compelling follow-up in light of the recent lattice QCD results for such systems [79] and the observed tension with the LO ChPT prediction.

As shown by this work, the combination of EFTs and lattice QCD continues to be a potent tool for studying the hadron spectrum. This synergy has already yielded valuable insights into the three-hadron problem and will certainly keep contributing in the future.

Acknowledgments

A   Dependence on the cutoff

The cutoff function $H(x)$ is arbitrary so long as it is smooth and satisfies $H(x) = 0$ if $x \leq 0$ and $H(x) = 1$ if $x \geq 1$. Throughout this paper, we have been using the standard version

$$H(x) = \exp\left[-\tfrac{1}{x}\exp\left(-\tfrac{1}{1-x}\right)\right], \quad \text{when } 0 < x < 1, \tag{IV.I.I}$$

but ref. [15] and others consider a generalization thereof, corresponding to the replacement

$$x \to 1 + \frac{4}{3-\alpha}(x-1), \qquad -1 \leq \alpha < 3, \tag{IV.I.2}$$

with $\alpha = -1$ corresponding to eq. (IV.I.I). Larger α give sharper cutoffs, with the limit $\alpha \to 3$ being a step function, $H(x) \to \theta(x-1)$. There is also the symmetric version

$$H(x) = \left[1 + \exp\left(\tfrac{1}{x} - \tfrac{1}{1-x}\right)\right]^{-1}, \quad \text{when } 0 < x < 1, \tag{IV.I.3}$$

which commonly appears in other areas, and which is more numerically well-behaved due to the lack of nested exponentials.

At NLO in ChPT, $\mathcal{K}_{\mathrm{df},3}$ depends on $H(x)$ through the coefficients $\mathcal{D}_X$, defined in eq. (IV.4.36). These are effectively the remainders of $\mathcal{D}^{\mathrm{BH}}$ after removing the analytic approximation obtained with $H(x) = \theta(x)$. With the stardard cutoff choice, eq. (IV.1.1), we obtain the values in eq. (IV.3.2), which we restate here:

$$\mathcal{D}_0 \approx -0.0563476589\,, \qquad \mathcal{D}_1 \approx 0.129589681\,, \qquad \mathcal{D}_2 \approx 0.432202370\,,$$
$$\mathcal{D}_{\mathrm{A}} \approx 9.07273890 \cdot 10^{-4}\,, \qquad \mathcal{D}_{\mathrm{B}} \approx 1.62394747 \cdot 10^{-4}\,. \tag{IV.1.4}$$

For comparison, with eq. (IV.1.3) one instead obtains

$$\mathcal{D}_0 \approx -0.0470650424\,, \qquad \mathcal{D}_1 \approx 0.107630347\,, \qquad \mathcal{D}_2 \approx 0.583361673\,,$$
$$\mathcal{D}_{\mathrm{A}} \approx -0.118643915\,, \qquad \mathcal{D}_{\mathrm{B}} \approx -0.0400284275\,. \tag{IV.1.5}$$

In fig. IV.11 we show the result for these quantities as a function of α in eq. (IV.1.2).

A few features can be noted. Using the standard H, eq. (IV.1.1), the analytic approximation is very good, in the sense that $|\mathcal{D}_X| \ll |\mathcal{K}_X|$. It is especially good for $X = \mathrm{A}, \mathrm{B}$. With larger α, the approximation of the quadratic order in the threshold expansion ($X = 2, \mathrm{A}, \mathrm{B}$) quickly grows worse, but in the leading orders ($X = 0, 1$) it grows better: Near $\alpha = 0.875$, $\mathcal{K}_0$ and $\mathcal{K}_1$ are almost exactly approximated. All $\mathcal{D}_X$ diverge as $\alpha \to 3$, corresponding to extremely sharp cutoffs.

B Loop integrals

Let us bring up some basic definitions. Our notation, which follows ref. [60], differs from the standard Passarino–Veltman integrals by extra factors of $\kappa = 1/(16\pi^2)$. For the bubble integrals we use

$$B(q^2) = \frac{1}{i} \int \frac{\mathrm{d}^d \ell}{(2\pi)^d} \frac{1}{(\ell^2 - M_\pi^2)\,[(\ell - q)^2 - M_\pi^2]} \tag{IV.2.1}$$
$$= \kappa \frac{1}{\tilde{\epsilon}} - \kappa - L + \bar{J}(q^2)\,.$$

Following ChPT conventions, dimensional regularization in $4 - 2\epsilon$ dimensions uses

$$\frac{1}{\tilde{\epsilon}} \equiv \frac{1}{\epsilon} - \gamma_{\mathrm{E}} + \log 4\pi - \log \mu^2 + 1\,, \tag{IV.2.2}$$

where γ_{E} is the Euler–Mascheroni constant. We employ the standard definition for $\bar{J}(q^2)$:

$$\bar{J}(q^2) \equiv \kappa \left(2 + \beta \log \frac{\beta - 1}{\beta + 1} \right)\,, \tag{IV.2.3}$$
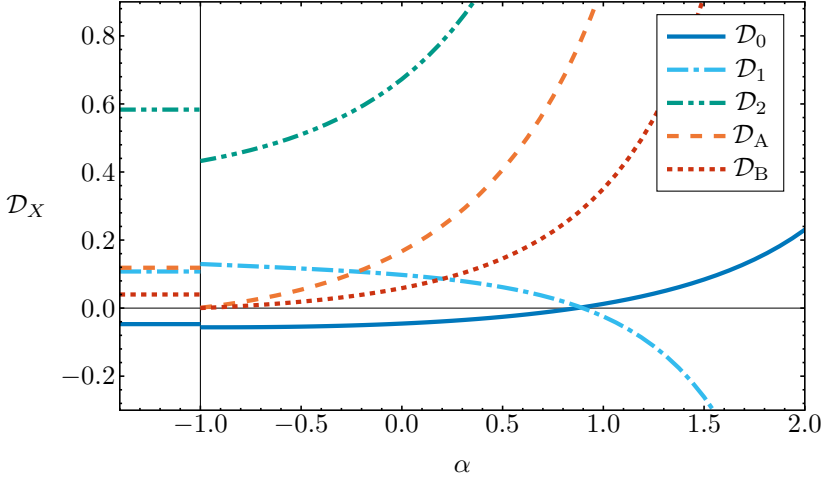
Figure IV.11: The numerical remainders $\mathcal{D}_X$ as functions of the parameter α, employing eq. (IV.1.1) together with eq. (IV.1.2). All $\mathcal{D}_X$ diverge as $\alpha \to 3$, while the standard $H(x)$ and the numerical values of eq. (IV.3.2) are recovered at $\alpha = -1$. For comparison, the values obtained with the symmetryic H, eq. (IV.1.3), are shown to the left of $\alpha = -1$. The line $\mathcal{D}_X = 0$ corresponds to the analytic approximation in section 4.3.1 being exact.

with $\beta \equiv \beta(q^2) = \sqrt{1 - \frac{4M_\pi^2}{q^2}}$. Regarding the scalar triangle integrals, we have

$$C(p_1, p_2, \ldots, p_6) = \frac{1}{i} \int \frac{\mathrm{d}^d \ell}{(2\pi)^d} \frac{1}{(\ell^2 - M_\pi^2)[(\ell - q_1)^2 - M_\pi^2][(\ell + q_2)^2 - M_\pi^2]}, \quad \text{(IV.2.4)}$$

where $q_1 = p_1 + p_2$, $q_2 = p_3 + p_4$; for completeness, we also define $q_3 = p_5 + p_6$. The expressions for $\mathcal{M}_3^{\mathrm{NLO}}$ also include the tensor integrals C_{11} and C_{21} (and C_3, which does not contribute at $I = 3$); their definitions and properties are found in appendix A of ref. [60]. In all cases, $\overline{C}_X$ denotes the UV-finite part of C_X, although among these integrals, only C_{21} is UV-divergent.

For the numerical evaluation of the triangle integrals we use *LoopTools* [75]. We also use the following analytic expression, which can be utilized once the tensor integrals are reduced to the scalar ones through the Passarino–Veltman reduction. The one-loop triangle function C (related to the standard C_0) can be written in terms of 12 dilogarithms as

$$\frac{1}{\kappa} C(p_1, p_2, \ldots, p_6) = C_0(q_1^2, q_2^2, q_3^2, M_\pi^2, M_\pi^2, M_\pi^2)$$

$$= \sum_{\alpha_1, \alpha_2 \in \{-1, 1\}} \frac{\alpha_1}{\xi_1 \xi_2} \mathrm{Li}_2 \left\{ \left(1 - \alpha_1 \xi_2\right) \left[\frac{1}{1 - \alpha_2 \beta_1 \xi_2} + i \alpha_2 \, \mathrm{sgn}\{q_1^2\} \epsilon \right] \right\} + \text{cycl.}, \quad \text{(IV.2.5)}$$

with $\lambda \equiv \lambda(q_1^2, q_2^2, q_3^2) > 0$ the Källén triangle function, $\beta_1 = \sqrt{1 - \frac{4M_\pi^2}{q_1^2}}$, $\xi_1 \equiv q_1^2 - q_2^2 - q_3^2$, $\xi_2 \equiv \sqrt{1 - 4q_2^2 q_3^2 / \xi_1^2}$, and "cycl." standing for cyclic permutations in $\{q_1^2, q_2^2, q_3^2\}$.

C   Cancellation of imaginary parts

To confirm that the formalism is working as it should, we would like to check that $\mathcal{K}_{\mathrm{df},3}^{\mathrm{NLO}}$, defined by eqs. (IV.2.24) and (IV.2.26), is indeed real, so that it is correct to use the master equation (IV.2.27). This should be possible at a diagram-by-diagram level, with each small set of contributions to $\mathcal{M}_3^{\mathrm{NLO}}$ matched with corresponding subtraction terms in $\mathcal{M}_{\mathrm{df},3}$ and corresponding ρ terms in eq. (IV.2.24). In the following subsections, we show that this holds as long as the contributing amplitudes, both on and off shell, satisfy unitarity. This is the case if a consistent off-shell convention is used throughout (see section 2.3).

All cancellations proven here have been verified using direct calculation at a selection of kinematic configurations detailed in appendix D.

C.1   Bull's head diagram

The BH diagram contributing to $\mathcal{M}_3$ (fig. IV.12 and its crossings) leads to imaginary contributions due to the presence of two two-particle cuts in the s-channel. Here, we address the issue of how these imaginary contributions are canceled in the transition to $\mathcal{K}_{\mathrm{df},3}$. This requires both the subtraction term $\mathcal{D}^{(u,u)\mathrm{NLO}}$ and the ρ terms in eq. (IV.2.24), and can be broken down diagrammatically in such a way that the cancellation is straightforward.

We consider the BH diagram in fig. IV.9a as a contribution to the unsymmetrized amplitude, $\mathcal{M}_3^{(u,u)\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k})$, since the cancellation can be seen before symmetrization. We start with the NLO subtraction. The relevant part of this term is the one corresponding to the BH diagram,

$$\mathcal{D}^{(u,u)\mathrm{NLO}}(\boldsymbol{p},\boldsymbol{k}) \supset \mathcal{D}_{ss}^{\mathrm{BH}}(\boldsymbol{p},\boldsymbol{k}) \equiv \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) \int_r \left\{ G_{ss}^{\infty}(\boldsymbol{p},r)\mathcal{M}_{2s}^{\mathrm{LO}}(r)G_{ss}^{\infty}(r,\boldsymbol{k}) \right\} \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k})\,.$$

$$(\mathrm{IV.3.1})$$

In the main text, all quantities were implicit matrices in ℓ, m space. Here, since $\mathcal{M}_2^{\mathrm{LO}}$ is purely s-wave, we can drop the ℓ, m indices in all quantities and leave it implicit that $\mathcal{D}_{ss}^{\mathrm{BH}}$ is nonzero only for $\ell' = \ell = 0$ (as indicated by the subscripts). This means that G_{ss}^{∞} is different here than in eq. (IV.2.20):

$$G_{ss}^{\infty}(\boldsymbol{p},\boldsymbol{r}) = \frac{H(x_p)H(x_r)}{b_{pr}^2 - M_\pi^2 + i\epsilon}\,, \qquad b_{pr} = (P - p - r)^2\,. \qquad (\mathrm{IV.3.2})$$

Note that in our calculation, $H(x_p) = 1$ because p is a momentum for which both particles in the interacting pair can be on shell. Since $\mathcal{M}_2^{\mathrm{LO}}$ is real, the imaginary part of $\mathcal{D}_{ss}^{\mathrm{BH}}$ arises only from the $i\epsilon$ in G_{ss}^{∞}, and can be pulled out using the standard Cauchy principal value,

$$\frac{1}{z + i\epsilon} = \mathcal{P}\frac{1}{z} - i\pi\delta(z)\,. \qquad (\mathrm{IV.3.3})$$

To proceed, we assume that $\boldsymbol{p}$ and $\boldsymbol{k}$ are chosen such that the poles in $G_{ss}^{\infty}(\boldsymbol{p},\boldsymbol{r})$ and $G_{ss}^{\infty}(\boldsymbol{r},\boldsymbol{k})$ occur for non-overlapping values of $\boldsymbol{r}$. Then we get two distinct contributions

to the imaginary part, one from the left-hand cut and the other from the right-hand cut. The case where the poles overlap can be handled with a more careful application of the principal value, following ref. [77], and we have done so as a cross-check. However, it can be circumvented with the following argument: An infinitesimal deformation of p and k is enough to remove the overlap, so by the smoothness of $\mathcal{K}_{df,3}$, if the imaginary parts cancel in the deformed case, they must also cancel without the deformation.

Thus, assuming that the poles do not overlap, we can change to the pair CMF variables r_p^* and r_k^* (in the notation of section 2.1) *separately* for the two delta-function contributions, and it is then straightforward to evaluate them. We focus on the left-hand (LH) cut contribution, as all the following holds separately for the right-hand (RH) cut.

Although the explicit form of the imaginary part is not needed in the following, it is still instructive to compute it. We find that the LH cut contribution to the imaginary part is

$$i \operatorname{Im} \mathcal{D}_{ss}^{\mathrm{BH,LH}}(\boldsymbol{p}, \boldsymbol{k}) = \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) \frac{(-i\pi)}{(2\pi)^3} \frac{q_{2,p}^*}{4E_{2,p}^*} \int \mathrm{d}\Omega(\hat{\boldsymbol{r}}_p^*) \left\{ \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{r}_{\mathrm{on}}) G_{ss}^\infty(\boldsymbol{r}_{\mathrm{on}}, \boldsymbol{k}) \right\} \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}) , \tag{IV.3.4}$$

where $\boldsymbol{r}_{\mathrm{on}}$ is the on-shell projection of $\boldsymbol{r}$ obtained using the prescription of ref. [13]. The integral in this equation runs over the directions of $\boldsymbol{r}_p^*$, and both $\mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{r}_{\mathrm{on}})$ and $G_{ss}^\infty(\boldsymbol{r}_{\mathrm{on}}, \boldsymbol{k})$ depend on this direction. The no-overlap assumption implies that the latter quantity is real, i.e., we do not have simultaneous contributions from both delta functions. We now observe that

$$\frac{q_{2,p}^*}{E_{2,p}^*} = 16i\pi\rho(\boldsymbol{p}) , \tag{IV.3.5}$$

so the contribution from the LH cut has the form of an $\mathcal{M}_2 \rho \mathcal{K}_{df,3}$ term, namely

$$i \operatorname{Im} \mathcal{D}_{ss}^{\mathrm{BH,LH}}(\boldsymbol{p}, \boldsymbol{k}) = 2\mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p})\rho(\boldsymbol{p}) \int \frac{\mathrm{d}\Omega(\hat{\boldsymbol{r}}_p^*)}{4\pi} \left\{ \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{r}_{\mathrm{on}}) G_{ss}^\infty(\boldsymbol{r}_{\mathrm{on}}, \boldsymbol{k}) \right\} \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}) . \tag{IV.3.6}$$

Next, we consider the following unsymmetrized term appearing in eq. (IV.2.24), which has the same $\mathcal{M}_2 \rho \mathcal{K}_{df,3}$ form,

$$\frac{1}{3} \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) \, \rho(\boldsymbol{p}) \, \mathcal{M}_{df,3}^{\mathrm{LO}}(\boldsymbol{p}, \boldsymbol{k}) . \tag{IV.3.7}$$

Again, this is purely s-wave (since $\mathcal{M}_2^{\mathrm{LO}}$ and $\mathcal{M}_{df,3}^{\mathrm{LO}}$ are), so we just need to consider the $\ell = \ell' = 0$ part. Only a subset of the terms in $\mathcal{M}_{df,3}^{\mathrm{LO}}$ contribute to the cancellation of the imaginary part of the BH diagram. These are all contained in the OPE contribution

$$\mathcal{M}_{df,3}^{\mathrm{LO,OPE}}(\boldsymbol{p}, \boldsymbol{k}) = -\mathcal{S} \left\{ \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}) \frac{1}{b_{pk}^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k}) - \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) G_{ss}^\infty(\boldsymbol{p}, \boldsymbol{k}) \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}) \right\} . \tag{IV.3.8}$$

Here, as in the main text, $b_{pk} = P - p - k$ and the subscript "off" indicates that the b_{pk} leg is off shell. Note it is important that a consistent off-shell convention is used throughout the calculation.

We now note that the initial-state symmetrization in eq. (IV.3.8) (i.e., that over k_1, k_2, k_3) will be repeated when the $\mathcal{M}_2 \rho \mathcal{K}_{\mathrm{df},3}$ term is symmetrized. Thus, one can drop the initial-state symmetrization in eq. (IV.3.8) and remove the factor of $1/3$ in eq. (IV.3.7). The final-state symmetrization then yields three terms,

$$\mathcal{M}_{\mathrm{df},3}^{\mathrm{LO,OPE}}(\boldsymbol{p},\boldsymbol{k}) \supset \mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{p},\boldsymbol{k}) + \mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{a}_p,\boldsymbol{k}) + \mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{b}_p,\boldsymbol{k}), \qquad \text{(IV.3.9)}$$

where $-\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{p},\boldsymbol{k})$ is the expression in braces in eq. (IV.3.8) and $\boldsymbol{a}_p, \boldsymbol{b}_p$ are the momenta of the final-state interacting-pair particles. Note that while the symmetrized $\mathcal{M}_{\mathrm{df},3}^{\mathrm{LO,OPE}}$ is a function of the total center-of-mass energy alone, the individual terms are not. The first term, $\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{p},\boldsymbol{k})$, will contribute to the cancellation of the imaginary part of the NLO OPE diagram (to be discussed in the following subsection), while the other two contributions, $\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{a}_p,\boldsymbol{k})$ and $\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{b}_p,\boldsymbol{k})$, are needed for the cancellation of the imaginary part of the BH diagram.

In fact, since $\mathcal{M}_2^{\mathrm{LO}}(\boldsymbol{p})$ is purely s-wave, there is an implicit angular integral over the first argument in both these BH contributions, arising from the projection onto $\ell' = 0$. It then follows from $\boldsymbol{a}_p^* = -\boldsymbol{b}_p^*$ that $\mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{a}_p,\boldsymbol{k}) = \mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{LO}}(\boldsymbol{b}_p,\boldsymbol{k})$. We thus keep only one of these two terms and multiply by a factor of 2. This leads to the final contribution to the unsymmetrized left-hand cut part of $[\mathcal{M}_2 \rho \mathcal{K}_{\mathrm{df},3}]_{\mathrm{BH}}$:

$$-2\mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p})\rho(\boldsymbol{p}) \int \frac{\mathrm{d}\Omega(\hat{\boldsymbol{a}}_p^*)}{4\pi} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{a}_{p,\mathrm{on}}) \frac{1}{b_{kr}^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k})$$

$$+ 2\mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p})\rho(\boldsymbol{p}) \int \frac{\mathrm{d}\Omega(\hat{\boldsymbol{a}}_p^*)}{4\pi} \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{a}_{p,\mathrm{on}}) G_{ss}^\infty(\boldsymbol{a}_{p,\mathrm{on}},\boldsymbol{k}) \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}). \qquad \text{(IV.3.10)}$$

This is purely imaginary, as the pole in the b_{kr} propagator is not crossed, given our assumptions about p and k. Since $\mathcal{D}$ enters with a minus sign, we see that the second term in eq. (IV.3.10) exactly cancels the imaginary part of $\mathcal{D}^{\mathrm{BH,LH}}(\boldsymbol{p},\boldsymbol{k})$ given in eq. (IV.3.6). This leaves the first term in eq. (IV.3.10), which itself exactly cancels the LH cut contribution to the imaginary part of the full BH diagram in the amplitude $\mathcal{M}_3$ (using the cutting rules).

Exactly analogous arguments hold for the RH cut part, in which one must use the $\mathcal{K}_{\mathrm{df},3}\rho\mathcal{M}_2$ term from the relation between $\mathcal{M}_3$ and $\mathcal{K}_{\mathrm{df},3}$. Thus, altogether, we have seen how the imaginary parts must cancel in the full BH contributions to $\mathcal{K}_{\mathrm{df},3}$.

C.2   OPE diagrams

We now consider the OPE diagrams in which the initial interaction is of NLO in ChPT and the final one of LO. The arguments are identical for the "flipped" time ordering.

Since we are here interested in the imaginary part, several simplifications occur. By construction, the OPE pole is canceled in $\mathcal{M}_{\mathrm{df},3}$, so the only source of a imaginary part is $\mathcal{M}_2^{\mathrm{NLO}}$, and this is present only in the s-wave, so only the ss part of G^∞ contributes. The

contribution to the unsymmetrized $\mathcal{M}_{\mathrm{df},3}$ is thus given by

$$
i \operatorname{Im} \mathcal{M}_{\mathrm{df},3}^{(u,u)\mathrm{NLO}}(\boldsymbol{p}, \boldsymbol{k}) \supset -i \operatorname{Im}\left[\mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(\boldsymbol{p})\right] \frac{1}{b_{pk}^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k})
$$
$$
+ i \operatorname{Im}\left[\mathcal{M}_{2s}^{\mathrm{NLO}}(\boldsymbol{p})\right] G_{ss}^\infty(\boldsymbol{p}, \boldsymbol{k}) \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}), \qquad \text{(IV.3.11)}
$$

where, as usual, $\boldsymbol{p}$ and $\boldsymbol{k}$ are final and initial spectator momenta, respectively, and "off" indicates that the b_{pk} leg is off shell. As above, here we are using the notation without implicit ℓm indices. To this must be added the contribution from the $\mathcal{M}_2 \rho \mathcal{K}_{\mathrm{df},3}$ term in eq. (IV.2.24), in which the OPE part of $\mathcal{K}_{\mathrm{df},3}^{\mathrm{LO}} = \mathcal{M}_{\mathrm{df},3}^{\mathrm{LO}}$, i.e., the first term of eq. (IV.3.9), is included. This contribution, which is purely imaginary, is

$$
-\mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) \rho(\boldsymbol{p}) \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}) \frac{1}{b_{pk}^2 - M_\pi^2 + i\epsilon} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{k})
$$
$$
+ \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) \rho(\boldsymbol{p}) \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) G_{ss}^\infty(p, k) \mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{k}). \qquad \text{(IV.3.12)}
$$

Now, we use unitarity and cutting rules to obtain

$$
i \operatorname{Im} \mathcal{M}_{2,\mathrm{off}}^{\mathrm{NLO}}(\boldsymbol{p}) = -\mathcal{M}_{2s}^{\mathrm{LO}}(\boldsymbol{p}) \rho(\boldsymbol{p}) \mathcal{M}_{2,\mathrm{off}}^{\mathrm{LO}}(\boldsymbol{p}), \qquad \text{(IV.3.13)}
$$

which applies also for the on-shell amplitude. Using this, we find that the sum of eqs. (IV.3.11) and (IV.3.12) vanishes.

C.3 Remaining diagrams

The remaining diagrams with an imaginary part involve a LO six-point vertex and an s-channel loop closed by a LO four-point vertex, either in the initial or final state. An example is shown in fig. IV.8. These diagrams are divergence-free by themselves. Thus, the imaginary parts must be canceled by $\mathcal{M}_2 \rho \mathcal{K}_{\mathrm{df},3}$-like terms appearing in the six-point vertex contribution to $\mathcal{K}_{\mathrm{df},3}$. That this is the case follows from unitarity, which introduces a factor of $-\rho$, and the double symmetrization, which cancels the $1/3$. No off-shell amplitudes appear in the imaginary parts, so the cancellation is independent of the off-shell convention.

D Threshold expansion using single-parameter kinematic configurations

In this appendix, we explain a method that we use to cross-check several of the calculations presented in the main text, and also for plotting the numerical behavior of the full NLO contribution in section 3.3. In particular, it provides an alternative analytic approach for obtaining the contributions of A_J, A_π, A_L, and A_l to $\mathcal{M}_{\mathrm{df},3}^{\mathrm{NLO,non\text{-}OPE}}$, which are discussed in section 4.2. It has also been use to perform a numerical check of all other contributions to $\mathcal{K}_{\mathrm{df},3}$.

Table IV.2: The five families used in the calculations, labeled by a. We use $\omega_p \equiv \sqrt{p^2 + M_\pi^2}$, and, for brevity, $\Omega_3 \equiv \omega_{\sqrt{3}p}$, $\Omega_5 \equiv \omega_{\sqrt{5}p}$ and $\Omega_7 \equiv \omega_{\sqrt{7}p/2}$. All momenta are on-shell with invariant mass M_π, and in all cases, $\boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3 = \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 = \boldsymbol{0}$. For compactness, p_3 and k_3 have been omitted but are easily inferred using $\boldsymbol{P} = \boldsymbol{0}$ and the on-shell condition.

a	$p_1^{(a)}(p)$	$p_2^{(a)}(p)$	$k_1^{(a)}(p)$	$k_2^{(a)}(p)$
1	$(\,\omega_p,\ p,\ \ 0, 0\,)$	$(\,\omega_p, -\tfrac{1}{2}p, \tfrac{\sqrt{3}}{2}p, 0\,)$	$(\,\omega_p,\ 0, 0,\ -p\,)$	$(\,\omega_p, \tfrac{\sqrt{3}}{2}p,\ \ 0,\ \tfrac{1}{2}p\,)$
2	$(\,\omega_p,\ p,\ \ 0, 0\,)$	$(\,\omega_p, -\tfrac{1}{2}p, \tfrac{\sqrt{3}}{2}p, 0\,)$	$(\,\omega_p, -p, 0,\ \ 0\,)$	$(\,\omega_p, \tfrac{1}{2}p, \tfrac{\sqrt{3}}{2}p,\ \ 0\,)$
3	$(\omega_{2p}, 2p,\ \ 0, 0\,)$	$(\,\Omega_7,\ -p, \tfrac{\sqrt{3}}{2}p, 0\,)$	$(\omega_{2p},\ 0, 0, -2p\,)$	$(\,\Omega_7, \tfrac{\sqrt{3}}{2}p,\ \ 0,\ \ p\,)$
4	$(\omega_{2p}, 2p,\ \ 0, 0\,)$	$(\,\Omega_3,\ -p, \sqrt{2}p, 0\,)$	$(\omega_{2p},\ 0, 0, -2p\,)$	$(\,\Omega_3, \sqrt{2}p,\ \ 0,\ \ p\,)$
5	$(\,\Omega_5,\ p, -2p, 0\,)$	$(\omega_{2p},\ \ 0,\ \ 2p, 0\,)$	$(\,\Omega_5,\ 0, p,\ \ 2p\,)$	$(\omega_{2p},\ \ 0,\ \ 0, -2p\,)$

As described in the main text, $\mathcal{K}_{\mathrm{df},3}$ is a function of eight kinematic degrees of freedom, so it is not straightforward to explore its general momentum dependence. Near threshold, however, its behavior is characterized by a few parameters, five if we work to quadratic order in Δ [see eq. (IV.2.2)]. In order to determine these coefficients from a given contribution to $\mathcal{K}_{\mathrm{df},3}$, one approach is to use families of momenta, each of which is a one-dimensional projection of the full momentum dependence. If one uses enough such families and controls the momentum dependence of $\mathcal{K}_{\mathrm{df},3}$ to high-enough order for each family, then the coefficients $\mathcal{K}_X$ of the threshold expansion can be determined.

Each family depends on a single parameter p that has dimension of momentum and, by design, vanishes at threshold. We use the five families listed in table IV.2. While this is more than the minimum number of families needed for our applications, using this number provides redundancy and cross-checks. Family 1 is the one used in ref. [60], while family 2 is a variant thereof, with the momenta arranged as equilateral triangles. Families 3–5 use isosceles triangles instead.

To use the families to determine the coefficients in the threshold expansion, we note that

$$\mathcal{K}_{\mathrm{df},3}(\mathcal{F}_a) = c_0^a + c_1^a p^2 + c_2^a p^4 + \mathcal{O}(p^6)\,, \tag{IV.4.1}$$

where a labels the family and $\mathcal{F}_a \equiv \{p_i^{(a)}(p), k_i^{(a)}(p)\}$. The coefficients c_i^a can be determined numerically or analytically. We also need the expansions

$$\Delta(\mathcal{F}_a) = d_1^a\, p^2 + d_2^a\, p^4 + \mathcal{O}(p^6)\,,$$
$$\Delta_{\mathrm{A}}(\mathcal{F}_a) = d_{\mathrm{A}}^a\, p^4 + \mathcal{O}(p^6)\,, \qquad \Delta_{\mathrm{B}}(\mathcal{F}_a) = d_{\mathrm{B}}^a\, p^4 + \mathcal{O}(p^6)\,. \tag{IV.4.2}$$

A closely related approach replaces the expansion in p^2 with one in E^{*2}. We have used both and checked that the results agree.

A single family is sufficient to determine

$$\mathcal{K}_0 = c_0^a\,, \qquad \mathcal{K}_1 = c_1^a/d_1^a\,, \tag{IV.4.3}$$

with other families providing cross-checks. To obtain the quadratic constants, we need three families of momenta, from which we can construct the matrix

$$Q = \begin{pmatrix} (d_1^{a_1})^2 & d_A^{a_1} & d_B^{a_1} \\ (d_1^{a_2})^2 & d_A^{a_2} & d_B^{a_2} \\ (d_1^{a_3})^2 & d_A^{a_3} & d_B^{a_3} \end{pmatrix} . \tag{IV.4.4}$$

We also collect the quadratic coefficients of $\mathcal{K}_{df,3}$ into a vector and subtract the p^4 term arising from $\mathcal{K}_1$:

$$V = \begin{pmatrix} c_2^{a_1} - c_1^{a_1} d_2^{a_1} / d_1^{a_1} \\ c_2^{a_2} - c_1^{a_2} d_2^{a_2} / d_1^{a_2} \\ c_2^{a_3} - c_1^{a_3} d_2^{a_3} / d_1^{a_3} \end{pmatrix} . \tag{IV.4.5}$$

Then,

$$\begin{pmatrix} \mathcal{K}_2 \\ \mathcal{K}_A \\ \mathcal{K}_B \end{pmatrix} = Q^{-1} V . \tag{IV.4.6}$$

This assumes that Q is invertible, which is true for some triplets of families. In particular, for the expansion of the matrix element, it is convenient to use the triplets of families $\{1, 2, 3\}$ and $\{1, 2, 4\}$. For numerical cross-checks, on the other hand, the triplet $\{1, 4, 5\}$ turned out to be the most convenient.

E   An integration method for less well-behaved $\mathcal{M}_3$

Section 4.3.4 covers the method used to calculate $\mathcal{D}^{(u,u)\mathrm{BH}}$ without first performing a threshold expansion. The applicability of this calculation relies on the finiteness of $\mathcal{D}^{(u,u)\mathrm{BH}}$, which in turn follows from the finiteness of the corresponding part of the amplitude, $\mathcal{M}_3^{(u,u)\mathrm{BH}}$. If that were not the case, one would have to regularize the divergences on both sides in a consistent way before the subtraction can take place, and it is not obvious how to do that. The same situation also invalidates the threshold expansion of sections 4.3.1 to 4.3.3 unless the divergent parts can be isolated first.

In this appendix, we present an alternative approach, which computes the difference $\mathcal{M}_3^{(u,u)\mathrm{BH}} - \mathcal{D}^{(u,u)\mathrm{BH}}$ without explicitly dealing with the individual terms. In the present case of NLO scattering at maximum isospin, this is nothing but an overly complicated cross-check procedure (and, in earlier stages, a contingency in case the finiteness turned out to be wrong), but it is conceivable that when the scope is generalized, one will eventually encounter a sufficiently pathological subtraction that this approach becomes worthwhile. Despite this more general outlook, we will present it as it would be applied to the present calculation, so that the technical details can be shown in full.

The goal is to write $\mathcal{M}_3^{(u,u)\mathrm{BH}}$ as similarly as possible to $\mathcal{D}^{(u,u)\mathrm{BH}}$, and then manipulate its expression to obtain a piece identical to $\mathcal{D}^{(u,u)\mathrm{BH}}$ plus compensatory terms, which must then equal $\mathcal{M}_3^{(u,u)\mathrm{BH}} - \mathcal{D}^{(u,u)\mathrm{BH}}$. Since this quantity is always divergence-free, its evaluation should be unproblematic. To specify $\mathcal{M}_3^{(u,u)\mathrm{BH}}$, which is parametrization-dependent,
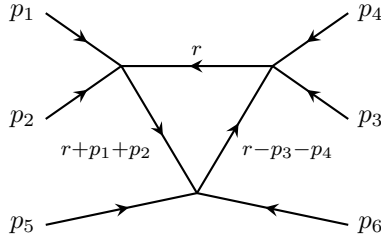
Figure IV.12: The bull's head diagram, showing the momentum routing used in eq. (IV.5.4). All momenta are ingoing, following the conventions of ref. [60]; the routing of fig. IV.9 is straightforwardly obtained through crossing.

it turns out that the most convenient parametrization for our purposes is the 5th among the ones presented in ref. [60], for which the $O(N+1)/O(N)$ Lagrangian is

$$\mathcal{L} = \frac{F_\pi^2}{2}\, \partial_\mu \Phi^\mathsf{T} \partial^\mu \Phi + F_\pi^2 \chi^\mathsf{T} \Phi \,, \qquad \Phi = \Phi_5 = \frac{1}{1 + \frac{1}{4}\frac{\boldsymbol{\phi}^\mathsf{T}\boldsymbol{\phi}}{F_\pi^2}} \left(1 - \frac{1}{4}\frac{\boldsymbol{\phi}^\mathsf{T}\boldsymbol{\phi}}{F_\pi^2},\ \frac{\boldsymbol{\phi}^\mathsf{T}}{F}\right)^\mathsf{T}, \quad \text{(IV.5.1)}$$

where $\boldsymbol{\phi}$ is a real vector of fields transforming linearly under the unbroken part of the symmetry group. In this representation, the 4-point vertex for flavors f_i and incoming off-shell momenta p_i reads[22]

$$\begin{aligned}
F_\pi^2 \mathcal{M}_{(5)}^{\mathrm{LO}}(p_1, f_1; p_2, f_2; p_3, f_3; p_4, f_4) = {}& \delta_{f_1 f_2}\delta_{f_3 f_4}\left(M_\pi^2 + p_1 \cdot p_2 + p_3 \cdot p_4\right)\\
& + \delta_{f_1 f_3}\delta_{f_2 f_4}\left(M_\pi^2 + p_1 \cdot p_3 + p_2 \cdot p_4\right)\\
& + \delta_{f_1 f_4}\delta_{f_2 f_3}\left(M_\pi^2 + p_1 \cdot p_4 + p_2 \cdot p_3\right).
\end{aligned} \quad \text{(IV.5.2)}$$

Now, we let $\mathcal{M}_3^{(u,u)\mathrm{BH}}$ be precisely the contribution from the diagram fig. IV.12 in this parametrization. Assembling the diagram and fixing all external flavors to

$$|\pi^+\rangle = |\pi^-\rangle^* = \frac{|1\rangle + i|2\rangle}{\sqrt{2}}\,, \quad \text{(IV.5.3)}$$

we get considerable cancellation among the terms in the sum over internal flavors, leaving

$$\begin{aligned}
\mathcal{M}_3^{(u,u)\mathrm{BH}} = {}& \frac{1}{F_\pi^6}\,(2p_1 \cdot p_2)(2p_5 \cdot p_6)\\
& \times \frac{1}{i}\int \frac{d^4 r}{(2\pi)^4} \frac{r^2 - M_\pi^2 - 2r \cdot p_3}{\left[r^2 - M_\pi^2\right]\left[(r + p_1 + p_2)^2 - M_\pi^2\right]\left[(r - p_3 - p_4)^2 - M_\pi^2\right]}\,.
\end{aligned} \quad \text{(IV.5.4)}$$

---

[22]With the momenta on-shell, this of course reduces to the parametrization-independent LO 4-particle amplitude, but the off-shell form presented here is precisely the 4-point vertex in the Φ_5 parametrization. Note that this is a different off-shell convention than anywhere else in this paper. They are compatible, however, since our main convention only separates OPE from non-OPE, whereas the one presented here separates the BH contribution from the rest of the non-OPE part (and is more convenient for that purpose).

Replacing $\{p_1, \ldots, p_6\}$ by $\{k_1, k_2, k_3, -p_1, -p_2, -p_3\}$ as described above eq. (IV.2.35), this bears a striking resemblance to eq. (IV.4.23). Indeed, comparing to eq. (IV.4.24) reveals

$$\mathcal{M}_3^{(u,u)\mathrm{BH}} = \frac{\Omega}{F_\pi^6} \frac{1}{i} \int \frac{d^4r}{(2\pi)^4} \frac{G(r; p_+, k_+)}{r^2 - M_\pi^2 + i\epsilon}, \qquad \mathcal{D}^{(u,u)\mathrm{BH}} = -\frac{\Omega}{F_\pi^6} \int_r H^2(x_r) G(r; p_+, k_+),$$
(IV.5.5)

where $\Omega \equiv (2p_1 \cdot p_2)(2k_1 \cdot k_2)$ for brevity.

The key step is now to place the r^0 integral of $\mathcal{M}_3^{(u,u)\mathrm{BH}}$ in the complex plane, and close the integration contour from below. This picks up three poles, one for each of the three propagators going on-shell with positive energy (closing it above would pick up the negative-energy poles). The residue at $r^2 = M_\pi^2$ contributes $-\Omega F_\pi^{-6} \int_r G(r; p_+, k_+)$; that is, it is precisely $\mathcal{D}^{(u,u)\mathrm{BH}}$ except that there is no $H(x_r)$. However, this absence is not straightforward to compensate for, since without a cutoff the on-shell integrals become UV-divergent. Let us therefore look closer at G, whose numerator is

$$(P - r)^2 - 2M_\pi^2 = (r^2 - M_\pi^2) + (P^2 - M_\pi^2) - 2P \cdot r. \tag{IV.5.6}$$

The first term on the right-hand side vanishes on-shell, while in $\mathcal{M}_3^{(u,u)\mathrm{BH}}$ it cancels one propagator and gives a simple, UV-divergent B integral [see eq. (IV.2.3)]. The second term is UV-finite. The third is also UV-finite for the purposes of $\mathcal{M}_3^{(u,u)\mathrm{BH}}$, but under an $\int_r$ integral it diverges logarithmically. These divergences must ultimately cancel, but those cancellations are very difficult to handle numerically. Therefore, we apply tensorial Passarino–Veltman reduction, replacing r with p_+ and k_+:

$$\begin{aligned}
-2P \cdot r \quad \longrightarrow \quad & \left([(p_+ - r)^2 - M_\pi^2] - [r^2 - M_\pi^2] - p_+^2 \right) \xi(p_+, k_+) \\
& + \left([(k_+ - r)^2 - M_\pi^2] - [r^2 - M_\pi^2] - k_+^2 \right) \xi(k_+, p_+),
\end{aligned}$$
(IV.5.7)

where we define ξ to also handle the case $p_+ = k_+$:

$$\xi(p, k) \equiv \begin{cases} \dfrac{p^2(P \cdot k) - (p \cdot k)(P \cdot p)}{p^2 k^2 - (p \cdot k)^2}, & \text{if } p \neq k, \\ \dfrac{P \cdot p}{2p^2}, & \text{if } p = k. \end{cases} \tag{IV.5.8}$$

Each term in square brackets in eq. (IV.5.7) cancels a propagator in $\mathcal{M}_3^{(u,u)\mathrm{BH}}$ and gives another B integral.

With this in mind, we define the fully UV-safe integrand

$$\widetilde{G}(r; p, k) \equiv \frac{P^2 - M_\pi^2 - p_+^2 \xi(p_+, k_+) - k_+^2 \xi(k_+, p_+)}{[(p - r)^2 - M_\pi^2 + i\epsilon][(k - r)^2 - M_\pi^2 + i\epsilon]} \tag{IV.5.9}$$

(note that it is p, k in the denominator but p_+, k_+ in the numerator; this will be important later). Compensating for the modified numerator results in

330

$$\mathcal{M}_3^{(u,u)\mathrm{BH}} = \frac{\Omega}{F_\pi^6}\left\{\frac{1}{i}\int\frac{d^4r}{(2\pi)^4}\frac{\widetilde{G}(r;p_+,k_+)}{r^2-M_\pi^2+i\epsilon}+\big[1-\xi(p_+,k_+)-\xi(k_+,p_+)\big]B\big((p_+-k_+)^2\big)\right.$$
$$\left.+\xi(p_+,k_+)B(k_+^2)+\xi(k_+,p_+)B(p_+^2)\right\},\quad\text{(IV.5.10)}$$

where it is now manifest what carries the UV divergence [namely $B\big((p_+-k_+)^2\big)$, which matches the result obtained by evaluating eq. (IV.5.4) the standard way], while

$$\mathcal{D}^{(u,u)\mathrm{BH}} = -\frac{\Omega}{F_\pi^6}\int_r H^2(x_r)\left\{\widetilde{G}(r;p_+,k_+)+\frac{\xi(p_+,k_+)}{(k_+-r)^2-M_\pi^2+i\epsilon}+\frac{\xi(k_+,p_+)}{(p_+-r)^2-M_\pi^2+i\epsilon}\right\}.$$
$$\text{(IV.5.11)}$$

Here, each term in the integral is individually convergent: We have moved the problematic cancellations to the extra B's in eq. (IV.5.10), where they are no problem at all. The extra ξ terms have at most simple poles.

Now, we apply the contour integration discussed above eq. (IV.5.6) to the first term in eq. (IV.5.10). The $r=M_\pi$ residue now contributes $-\Omega F_\pi^{-6}\int_r\widetilde{G}(r;p_+,k_+)$, which mostly cancels against the first term in $\mathcal{D}^{(u,u)\mathrm{BH}}$, while the remaining residues (from $\widetilde{G}$) can either come from two simple poles or one double pole. Covering both cases, their contribution is[23]

$$G_P(r)\equiv\begin{cases}\widetilde{G}(r;-k_+,p_+-k_+)+\widetilde{G}(r;k_+-p_+,-p_+)\,, & \text{if } p_+\neq k_+\,,\\[2mm]-\widetilde{G}(r;-p_+,-p_+)\Big[1+\dfrac{2r\cdot p_++p_+^2+2\omega_r p_{+0}}{2\omega_r^2}\Big]\,, & \text{if } p_+=k_+\,,\end{cases}\quad\text{(IV.5.12)}$$

so that

$$\frac{\Omega}{F_\pi^6}\int\frac{d^4r}{(2\pi)^4}\frac{\widetilde{G}(r;p_+,k_+)}{r^2-M_\pi^2+i\epsilon}-\mathcal{D}^{(u,u)\mathrm{BH}}=-\frac{\Omega}{F_\pi^6}\int_r\big[G_H(r)+G_P(r)\big]\,,\quad\text{(IV.5.13)}$$

where we defined

$$G_H(r)\equiv\widetilde{G}(r;p_+,k_+)-H^2(x_r)G(r;p_+,k_+)$$
$$=\big[1-H^2(x_r)\big]\widetilde{G}(r;p_+,k_+)-\frac{H^2(x_r)\xi(p_+,k_+)}{(k_+-r)^2-M_\pi^2+i\epsilon}-\frac{H^2(x_r)\xi(k_+,p_+)}{(p_+-r)^2-M_\pi^2+i\epsilon}\,.$$
$$\text{(IV.5.14)}$$

Equation (IV.5.13), along with eq. (IV.5.10), is the master formula for this subtraction.

The right-hand side of eq. (IV.5.13) can be dealt with by modifying the methods of section 4.3.4. First, we numerically evaluate

$$\int_{r<R}^{(\diamond)}\big[1-H^2(x_r)\big]\left\{\widetilde{G}(r;p_+,k_+)+\frac{\xi(p_+,k_+)}{(k_+-r)^2-M_\pi^2+i\epsilon}+\frac{\xi(k_+,p_+)}{(p_+-r)^2-M_\pi^2+i\epsilon}\right\},$$
$$\text{(IV.5.15)}$$

---

[23]There is no kinematic configuration that gives a triple pole; the $r=M_\pi$ pole is always separate from the others. In particular, the $p_+=k_+$ version of $\widetilde{G}_P$ is appropriate at threshold.

where ($\diamond$) indicates that the integral is evaluated in the Breit frame, i.e., $\int^{(\diamond)} \widetilde{G}(r; p, k)$ is taken in the CMF of $p + k$. By design, the integrand is entirely free from singularities. Then,

$$
\int_r G_H(r) = (\text{IV.5.15}) + \int_{r>R}^{(\diamond)} \widetilde{G}(r; p_+, k_+)
$$
$$
- \int_{r<R}^{(\diamond)} \left\{ \frac{\xi(p_+, k_+)}{(k_+ - r)^2 - M_\pi^2 + i\epsilon} + \frac{\xi(k_+, p_+)}{(p_+ - r)^2 - M_\pi^2 + i\epsilon} \right\}, \quad (\text{IV.5.16})
$$

where the angles can be integrated out of the remaining integrals as in eq. (IV.4.44). Note, however, that $a_{1,2,3}$ are now different than in eq. (IV.4.43), while for ξ we get

$$
\int_{r<R}^{(\diamond)} \frac{\xi(p, k)}{(k - r)^2 - M_\pi^2 + i\epsilon} = \int_0^R \frac{2\pi r^2 \mathrm{d}r}{2\omega_r (2\pi)^3} \, \xi(p, k) \, g_\xi(k^2 - 2k_0^\diamond \omega_r + i\epsilon; c), \quad (\text{IV.5.17})
$$

where c is given by eq. (IV.4.43), and

$$
g_\xi(b; c) \equiv \begin{cases} \dfrac{2}{b}, & \text{if } c = 0, \\ \dfrac{1}{c} \big[\log(b - c) - \log(b + c)\big] & \text{otherwise}. \end{cases} \quad (\text{IV.5.18})
$$

Likewise, the angles can be integrated out of $\int_r G_P$. Note, however, that each individual $\widetilde{G}$ in eq. (IV.5.12) is integrated in a different frame, as described below eq. (IV.5.15). With that clarified, and with eq. (IV.4.43) used in each frame separately,

$$
\int_r^{(\diamond)} \widetilde{G}(r; p, k) = \int \frac{2\pi r^2 \mathrm{d}r}{2\omega_r (2\pi)^3} \, N \, g(1, 0; b_1, b_2; c), \quad (\text{IV.5.19a})
$$
$$
\int_r^{(\diamond)} \widetilde{G}(r; -p_+, -p_+) N' = \int \frac{2\pi r^2 \mathrm{d}r}{2\omega_r (2\pi)^3} \, N \, g\left[1 + \frac{p_+^2 + 4p_{+0}^\diamond \omega_r}{2\omega_r^2}, -\frac{2rq}{2\omega_r^2}; b_1, b_2; c\right], \quad (\text{IV.5.19b})
$$

where [taken from eqs. (IV.5.9) and (IV.5.12), respectively]

$$
N \equiv P^2 - M_\pi^2 - p_+^2 \xi(p_+, k_+) - k_+^2 \xi(k_+, p_+),
$$
$$
N' \equiv 1 + \frac{2r \cdot p_+ + p_+^2 + 2\omega_r p_{+0}}{2\omega_r^2}. \quad (\text{IV.5.20})
$$

Most of the integrands in eqs. (IV.5.16) and (IV.5.19) have singularities, but they are, in a sense, less severe than those encountered in section 4.3.4, and all give finite integrals at threshold (this is obvious for g_ξ but quite subtle for G_P). Our efforts with the numerator also ensure that the integrals to infinity can be safely done using a suitable numerical method. Assembling all these pieces completes the subtraction.

References

[1] Briceño, R. A., Dudek, J. J. & Young, R. D. "Scattering processes and resonances from lattice QCD". *Rev. Mod. Phys.* **90,** 025001. arXiv: 1706.06223 [hep-lat] (2018).

[2] Hansen, M. T. & Sharpe, S. R. "Lattice QCD and Three-particle Decays of Resonances". *Ann. Rev. Nucl. Part. Sci.* **69,** 65–107. arXiv: 1901.00483 [hep-lat] (2019).

[3] Rusetsky, A. "Three particles on the lattice". *PoS* **LATTICE2019,** 281. arXiv: 1911.01253 [hep-lat] (2019).

[4] Hörz, B. "Spectroscopy and Hadron Interactions". *PoS* **LATTICE2021,** 006 (2022).

[5] Mai, M., Döring, M. & Rusetsky, A. "Multi-particle systems on the lattice and chiral extrapolations: a brief review". *Eur. Phys. J. ST* **230,** 1623–1643. arXiv: 2103.00577 [hep-lat] (2021).

[6] Mai, M., Meißner, U.-G. & Urbach, C. "Towards a theory of hadron resonances". *Phys. Rept.* **1001,** 1–66. arXiv: 2206.01477 [hep-ph] (2023).

[7] Romero-López, F. *Three-particle scattering amplitudes from lattice QCD* in *19th International Conference on Hadron Spectroscopy and Structure* (2021). arXiv: 2112.05170 [hep-lat].

[8] Romero-López, F. *Multi-hadron interactions from lattice QCD* in *39th International Symposium on Lattice Field Theory* (2022). arXiv: 2212.13793 [hep-lat].

[9] Beane, S. R., Detmold, W. & Savage, M. J. "n-Boson Energies at Finite Volume and Three-Boson Interactions". *Phys. Rev.* **D76,** 074507. arXiv: 0707.1670 [hep-lat] (2007).

[10] Detmold, W. & Savage, M. J. "The Energy of n Identical Bosons in a Finite Volume at $\mathcal{O}(L^{-7})$". *Phys. Rev.* **D77,** 057502. arXiv: 0801.0763 [hep-lat] (2008).

[11] Briceño, R. A. & Davoudi, Z. "Three-particle scattering amplitudes from a finite volume formalism". *Phys. Rev.* **D87,** 094507. arXiv: 1212.3398 [hep-lat] (2013).

[12] Polejaeva, K. & Rusetsky, A. "Three particles in a finite volume". *Eur. Phys. J. A* **48,** 67. arXiv: 1203.1241 [hep-lat] (2012).

[13] Hansen, M. T. & Sharpe, S. R. "Relativistic, model-independent, three-particle quantization condition". *Phys. Rev. D* **90,** 116003. arXiv: 1408.5933 [hep-lat] (2014).

[14] Hansen, M. T. & Sharpe, S. R. "Expressing the three-particle finite-volume spectrum in terms of the three-to-three scattering amplitude". *Phys. Rev. D* **92,** 114509. arXiv: 1504.04248 [hep-lat] (2015).

[15] Briceño, R. A., Hansen, M. T. & Sharpe, S. R. "Relating the finite-volume spectrum and the two-and-three-particle S matrix for relativistic systems of identical scalar particles". *Phys. Rev.* **D95,** 074510. arXiv: 1701.07465 [hep-lat] (2017).

[16] Koenig, S. & Lee, D. "Volume Dependence of N-Body Bound States". *Phys. Lett.* **B779,** 9–15. arXiv: 1701.00279 [hep-lat] (2018).

[17] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three-particle quantization condition in a finite volume: 1. The role of the three-particle force". *JHEP* **09**, 109. arXiv: 1706.07700 [hep-lat] (2017).

[18] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three particle quantization condition in a finite volume: 2. general formalism and the analysis of data". *JHEP* **10**, 115. arXiv: 1707.02176 [hep-lat] (2017).

[19] Mai, M. & Döring, M. "Three-body Unitarity in the Finite Volume". *Eur. Phys. J. A* **53**, 240. arXiv: 1709.08222 [hep-lat] (2017).

[20] Briceño, R. A., Hansen, M. T. & Sharpe, S. R. "Numerical study of the relativistic three-body quantization condition in the isotropic approximation". *Phys. Rev.* **D98**, 014506. arXiv: 1803.04169 [hep-lat] (2018).

[21] Briceño, R. A., Hansen, M. T. & Sharpe, S. R. "Three-particle systems with resonant subprocesses in a finite volume". *Phys. Rev.* **D99**, 014516. arXiv: 1810.01429 [hep-lat] (2019).

[22] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "Implementing the three-particle quantization condition including higher partial waves". *JHEP* **03**, 106. arXiv: 1901.07095 [hep-lat] (2019).

[23] Pang, J.-Y., Wu, J.-J., Hammer, H.-W., Meißner, U.-G. & Rusetsky, A. "Energy shift of the three-particle system in a finite volume". *Phys. Rev.* **D99**, 074513. arXiv: 1902.01111 [hep-lat] (2019).

[24] Jackura, A. W. *et al.* "Equivalence of three-particle scattering formalisms". *Phys. Rev. D* **100**, 034508. arXiv: 1905.12007 [hep-ph] (2019).

[25] Briceño, R. A., Hansen, M. T., Sharpe, S. R. & Szczepaniak, A. P. "Unitarity of the infinite-volume three-particle scattering amplitude arising from a finite-volume formalism". *Phys. Rev.* **D100**, 054508. arXiv: 1905.11188 [hep-lat] (2019).

[26] Romero-López, F., Sharpe, S. R., Blanton, T. D., Briceño, R. A. & Hansen, M. T. "Numerical exploration of three relativistic particles in a finite volume including two-particle resonances and bound states". *JHEP* **10**, 007. arXiv: 1908.02411 [hep-lat] (2019).

[27] Hansen, M. T., Romero-López, F. & Sharpe, S. R. "Generalizing the relativistic quantization condition to include all three-pion isospin channels". *JHEP* **07**, 047. arXiv: 2003.10974 [hep-lat] (2020).

[28] Blanton, T. D. & Sharpe, S. R. "Alternative derivation of the relativistic three-particle quantization condition". *Phys. Rev. D* **102**, 054520. arXiv: 2007.16188 [hep-lat] (2020).

[29] Blanton, T. D. & Sharpe, S. R. "Equivalence of relativistic three-particle quantization conditions". *Phys. Rev. D* **102**, 054515. arXiv: 2007.16190 [hep-lat] (2020).

[30] Pang, J.-Y., Wu, J.-J. & Geng, L.-S. "DDK system in finite volume". *Phys. Rev. D* **102**, 114515. arXiv: 2008.13014 [hep-lat] (2020).

[31] Romero-López, F., Rusetsky, A., Schlage, N. & Urbach, C. "Relativistic N-particle energy shift in finite volume". *JHEP* **02**, 060. arXiv: 2010.11715 [hep-lat] (2021).

[32] Blanton, T. D. & Sharpe, S. R. "Relativistic three-particle quantization condition for nondegenerate scalars". *Phys. Rev. D* **103,** 054503. arXiv: 2011.05520 [hep-lat] (2021).

[33] Müller, F., Rusetsky, A. & Yu, T. "Finite-volume energy shift of the three-pion ground state". *Phys. Rev. D* **103,** 054506. arXiv: 2011.14178 [hep-lat] (2021).

[34] Blanton, T. D. & Sharpe, S. R. "Three-particle finite-volume formalism for $\pi^+\pi^+K^+$ and related systems". *Phys. Rev. D* **104,** 034509. arXiv: 2105.12094 [hep-lat] (2021).

[35] Müller, F., Pang, J.-Y., Rusetsky, A. & Wu, J.-J. "Relativistic-invariant formulation of the three-particle quantization condition". arXiv: 2110.09351 [hep-lat] (2021).

[36] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "Implementing the three-particle quantization condition for $\pi^+\pi^+K^+$ and related systems". *JHEP* **02,** 098. arXiv: 2111.12734 [hep-lat] (2022).

[37] Jackura, A. W. "Three-body scattering and quantization conditions from S matrix unitarity". arXiv: 2208.10587 [hep-lat] (2022).

[38] Müller, F., Pang, J.-Y., Rusetsky, A. & Wu, J.-J. "Three-particle Lellouch-Lüscher formalism in moving frames". *JHEP* **02,** 214. arXiv: 2211.10126 [hep-lat] (2023).

[39] Draper, Z. T., Hansen, M. T., Romero-López, F. & Sharpe, S. R. "Three relativistic neutrons in a finite volume". *JHEP* **07,** 226. arXiv: 2303.10219 [hep-lat] (2023).

[40] Beane, S. R. *et al.* "Multi-Pion Systems in Lattice QCD and the Three-Pion Interaction". *Phys. Rev. Lett.* **100,** 082004. arXiv: 0710.1827 [hep-lat] (2008).

[41] Detmold, W. & Smigielski, B. "Lattice QCD study of mixed systems of pions and kaons". *Phys. Rev. D* **84,** 014508. arXiv: 1103.4362 [hep-lat] (2011).

[42] Mai, M. & Döring, M. "Finite-Volume Spectrum of $\pi^+\pi^+$ and $\pi^+\pi^+\pi^+$ Systems". *Phys. Rev. Lett.* **122,** 062503. arXiv: 1807.04746 [hep-lat] (2019).

[43] Hörz, B. & Hanlon, A. "Two- and three-pion finite-volume spectra at maximal isospin from lattice QCD". *Phys. Rev. Lett.* **123,** 142002. arXiv: 1905.04277 [hep-lat] (2019).

[44] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "$I = 3$ Three-Pion Scattering Amplitude from Lattice QCD". *Phys. Rev. Lett.* **124,** 032001. arXiv: 1909.02973 [hep-lat] (2020).

[45] Mai, M., Döring, M., Culver, C. & Alexandru, A. "Three-body unitarity versus finite-volume $\pi^+\pi^+\pi^+$ spectrum from lattice QCD". *Phys. Rev. D* **101,** 054510. arXiv: 1909.05749 [hep-lat] (2020).

[46] Culver, C., Mai, M., Brett, R., Alexandru, A. & Döring, M. "Three pion spectrum in the $I = 3$ channel from lattice QCD". *Phys. Rev. D* **101,** 114507. arXiv: 1911.09047 [hep-lat] (2020).

[47] Fischer, M. *et al.* "Scattering of two and three physical pions at maximal isospin from lattice QCD". *Eur. Phys. J. C* **81,** 436. arXiv: 2008.03035 [hep-lat] (2021).

[48] Hadron Spectrum, Hansen, M. T., Briceño, R. A., Edwards, R. G., Thomas, C. E. & Wilson, D. J. "Energy-Dependent $\pi^+\pi^+\pi^+$ Scattering Amplitude from QCD". *Phys. Rev. Lett.* **126,** 012001. arXiv: 2009.04931 [hep-lat] (2021).

[49] NPLQCD, QCDSF, Beane, S. R. *et al.* "Charged multihadron systems in lattice QCD+QED". *Phys. Rev. D* **103,** 054504. arXiv: 2003.12130 [hep-lat] (2021).

[50] Alexandru, A. *et al.* "Finite-volume energy spectrum of the $K^-K^-K^-$ system". *Phys. Rev. D* **102,** 114523. arXiv: 2009.12358 [hep-lat] (2020).

[51] Brett, R. *et al.* "Three-body interactions from the finite-volume QCD spectrum". *Phys. Rev. D* **104,** 014501. arXiv: 2101.06144 [hep-lat] (2021).

[52] Blanton, T. D. *et al.* "Interactions of two and three mesons including higher partial waves from lattice QCD". *JHEP* **10,** 023. arXiv: 2106.05590 [hep-lat] (2021).

[53] GWQCD, Mai, M. *et al.* "Three-Body Dynamics of the a1(1260) Resonance from Lattice QCD". *Phys. Rev. Lett.* **127,** 222001. arXiv: 2107.03973 [hep-lat] (2021).

[54] Weinberg, S. "Phenomenological Lagrangians". *Physica A* **96,** 327–340 (1979).

[55] Gasser, J. & Leutwyler, H. "Chiral Perturbation Theory to One Loop". *Annals Phys.* **158,** 142 (1984).

[56] Flavour Lattice Averaging Group (FLAG), Aoki, Y. *et al.* "FLAG Review 2021". *Eur. Phys. J. C* **82,** 869. arXiv: 2111.09849 [hep-lat] (2022).

[57] ETM, Helmes, C. *et al.* "Hadron-hadron interactions from $N_f = 2 + 1 + 1$ lattice QCD: isospin-2 $\pi\pi$ scattering length". *JHEP* **09,** 109. arXiv: 1506.00408 [hep-lat] (2015).

[58] Mai, M., Culver, C., Alexandru, A., Döring, M. & Lee, F. X. "Cross-channel study of pion scattering from lattice QCD". *Phys. Rev. D* **100,** 114514. arXiv: 1908.01847 [hep-lat] (2019).

[59] Fu, Z. & Chen, X. "$I = 0$ $\pi\pi$ s-wave scattering length from lattice QCD". *Phys. Rev. D* **98,** 014514. arXiv: 1712.02219 [hep-lat] (2018).

[60] Bijnens, J. & Husek, T. "Six-pion amplitude". *Phys. Rev. D* **104,** 054046. arXiv: 2107.06291 [hep-ph] (2021).

[61] Bijnens, J., Husek, T. & Sjö, M. "Six-meson amplitude in QCD-like theories". *Phys. Rev. D* **106,** 054021. arXiv: 2206.14212 [hep-ph] (2022).

[62] Jackura, A. W., Briceño, R. A., Dawid, S. M., Islam, M. H. E. & McCarty, C. "Solving relativistic three-body integral equations in the presence of bound states". *Phys. Rev. D* **104,** 014507. arXiv: 2010.09820 [hep-lat] (2021).

[63] Dawid, S. M., Islam, M. H. E. & Briceño, R. A. "Analytic continuation of the relativistic three-particle scattering amplitudes". arXiv: 2303.04394 [nucl-th] (2023).

[64] Scherer, S. & Schindler, M. R. *A Primer for Chiral Perturbation Theory* 1st ed. ISBN: 978-3-642-19253-1 (Springer-Verlag, 2012).

[65] Pich, A. *Effective Field Theory with Nambu-Goldstone Modes* in *Les Houches summer school: EFT in Particle Physics and Cosmology* (2018). arXiv: 1804.05664 [hep-ph].

[66] Bijnens, J. & Lu, J. "Meson-meson Scattering in QCD-like Theories". *JHEP* **03**, 028. arXiv: 1102.0172 [hep-ph] (2011).

[67] Colangelo, G., Gasser, J. & Leutwyler, H. "$\pi\pi$ scattering". *Nucl. Phys. B* **603**, 125–179. arXiv: hep-ph/0103088 (2001).

[68] MILC, Bazavov, A. *et al.* "Results for light pseudoscalar mesons". *PoS* **LATTICE2010** (ed Rossi, G.) 074. arXiv: 1012.0868 [hep-lat] (2010).

[69] Borsanyi, S. *et al.* "SU(2) chiral perturbation theory low-energy constants from 2+1 flavor staggered lattice simulations". *Phys. Rev. D* **88**, 014513. arXiv: 1205.0788 [hep-lat] (2013).

[70] Budapest-Marseille-Wuppertal, Dürr, S. *et al.* "Lattice QCD at the physical point meets SU(2) chiral perturbation theory". *Phys. Rev. D* **90**, 114504. arXiv: 1310.3626 [hep-lat] (2014).

[71] Boyle, P. A. *et al.* "Low energy constants of SU(2) partially quenched chiral perturbation theory from $N_f = 2 + 1$ domain wall QCD". *Phys. Rev. D* **93**, 054502. arXiv: 1511.01950 [hep-lat] (2016).

[72] Beane, S. R. *et al.* "SU(2) Low-Energy Constants from Mixed-Action Lattice QCD". *Phys. Rev. D* **86**, 094509. arXiv: 1108.1380 [hep-lat] (2012).

[73] Vermaseren, J. A. M. "New features of FORM". arXiv: math-ph/0010025 (2000).

[74] Bijnens, J. "CHIRON: a package for ChPT numerical results at two loops". *Eur. Phys. J. C* **75**, 27. arXiv: 1412.0887 [hep-ph] (2015).

[75] Hahn, T. & Pérez-Victoria, M. "Automatized one-loop calculations in four and D dimensions". *Comput. Phys. Commun.* **118**, 153–165. arXiv: hep-ph/9807565 (1999).

[76] Galassi, M. *et al.* *GNU scientific library* http://www.gnu.org/software/gsl/ (Network Theory Limited Godalming, 2002).

[77] Davies, K. T. R., Glasser, M. L., Protopopescu, V. & Tabakin, F. "The mathematics of principal value integrals and applications to nuclear physics, transport theory, and condensed matter physics". *Math. Models Methods Appl. Sci.* **06**, 833–885 (1996).

[78] Costin, O. & Friedman, H. M. "Foundational aspects of singular integrals". *Journal of Functional Analysis* **267**, 4732–4752. ISSN: 0022-1236. eprint: 1401.7045 (2014).

[79] Draper, Z. T. *et al.* "Interactions of πK, $\pi\pi K$ and $KK\pi$ systems at maximal isospin from lattice QCD". *JHEP* **05**, 137. arXiv: 2302.13587 [hep-lat] (2023).

Paper V

V

The three-particle K-matrix at NLO in ChPT

Jorge Baeza-Ballesteros,[1] Johan Bijnens,[2] Tomáš Husek,[2,3]
Fernando Romero-López,[4] Stephen R. Sharpe[5] and Mattias Sjö[2]

[1]IFIC, CSIC-Universitat de València, 46980 Paterna, Spain

[2]Department of Physics, Lund University, Box 118, SE 22100 Lund, Sweden

[3]Institute of Particle and Nuclear Physics, Charles University,
    V Holešovičkách 2, 180 00 Prague, Czech Republic

[4]CTP, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[5]Physics Department, University of Washington, Seattle, WA 98195-1560, USA

Abstract: The three-particle K-matrix, $\mathcal{K}_{\mathrm{df},3}$, is a scheme-dependent quantity that parametrizes short-range three-particle interactions in the relativistic-field-theory three-particle finite-volume formalism. In this work, we compute its value for systems of three pions in all isospin channels through next-to-leading order (NLO) in Chiral Perturbation Theory, generalizing previous work done at maximum isospin. We obtain analytic expressions through quadratic order (or cubic order, in the case of zero isospin) in the expansion about the three-pion threshold.

Note: This paper, although mostly complete, is still in preparation for publishing. The results presented here have not been fully cross-checked and may therefore contain errors, and the conclusions drawn from them are still preliminary. The reader is directed to the finished version of the paper, which will likely be published soon after this thesis.

1 Introduction

First-principles studies of three-hadron physics from Quantum Chromodynamics (QCD) are finally becoming possible after a number of theoretical, numerical and algorithmic developments [1–33]. While only simple three-meson systems at maximal isospin have been studied using lattice QCD [34–46], it is to be expected that more complicated ones will be investigated soon. The scattering of three generic pions constitutes a potential next milestone for lattice QCD, since some relevant low-lying resonances such as the $\omega(782)$ can be found.

The extraction of three-particle scattering amplitudes from lattice QCD utilizes the three-particle finite-volume formalism, which connects finite-volume energies obtained in lattice QCD to the three-particle scattering amplitude. Mainly following three different approaches, the formalism has been developed for a number of relevant three-hadron systems. The approach that we will consider in this work, the so-called relativistic-field-theory (RFT) three-particle formalism [5, 6], has been frequently used in the literature for numerical studies [37, 40, 41, 45]. In the RFT formalism, the central object parametrizing short-range three-particle interactions is the three-particle K-matrix, $\mathcal{K}_{\mathrm{df},3}$.

The interface between lattice QCD and Chiral Perturbation Theory (ChPT) has proven to be a valuable source of insights for first-principles predictions of multi-pion quantities. A recent example is the comparison between lattice QCD results and ChPT for three-pion results [47], which has provided useful understanding of the chiral dependence of three-pion quantities. In particular, in ref. [47] we computed the three-pion maximum-isospin K-matrix at next-to-leading order (NLO) in ChPT. We showed that the previously observed tension between leading-order (LO) predictions and lattice QCD results for $\mathcal{K}_{\mathrm{df},3}$ was significantly reduced when comparing against the NLO prediction. This improved agreement was also an important check of the RFT formalism itself.

In the present work, we aim at generalizing the NLO ChPT results of ref. [47] to the case of three pions in any possible isospin channel. This result will be useful for future lattice QCD calculations, either by providing constraints in the near-threshold energy region of $\mathcal{K}_{\mathrm{df},3}$ or by inspiring parametrizations of the three-particle K-matrix. Note, however, that the presence of resonances when the isospin is not maximal will reduce the energy range of validity with respect to the maximal-isospin case.

The strategy followed in this work is similar to that followed for the computation at maximal isospin [47]. We make use of the $3\pi \to 3\pi$ amplitude computed in refs. [48, 49] at NLO in ChPT. We relate this amplitude to the K-matrix of the three-pion generalization of the RFT formalism, derived in ref. [20]. Several complications due to the presence of nonidentical pions are present in this calculation. These include additional structures in the threshold expansion of $\mathcal{K}_{\mathrm{df},3}$, the presence of odd partial waves in certain channels, a more complicated symmetrization procedure needed to account for all diagrammatic contributions, and the presence of a s-channel diagram of the form $3\pi \to \pi \to 3\pi$, which contributes to the isospin-1 3π K-matrix.

This paper is organized as follows. In section 2, we present the structure of $\mathcal{K}_{\mathrm{df},3}$, the three-

pion states in the various isospin channels, and the structure of the threshold expansion. Most details of the relation between $\mathcal{K}_{\mathrm{df},3}$ and the ChPT amplitude are deferred to ref. [47], since they are the same there as in the present cases. In section 3, we go through the calculation of $\mathcal{K}_{\mathrm{df},3}$, first at LO and then at NLO. We present the results in section 4, and state some preliminary conclusions in section 5. This paper contains 2 appendices, detailing the cutoff dependence of $\mathcal{K}_{\mathrm{df},3}$ (appendix A) and giving some insight into the threshold expansion based on group theory (appendix B).

2 Theoretical background

2.1 The three-particle K-matrix from ChPT
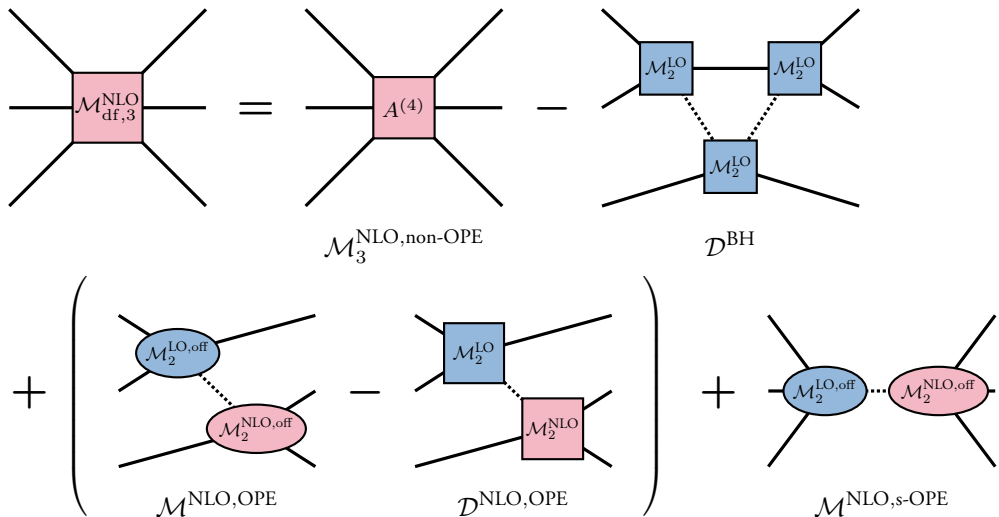
Figure v.1: Sketch of eq. (v.2.1) (compare fig. 2 in ref. [47]). Solid lines represent on-shell pions, while dotted lines are off-shell propagators. Square boxes indicate fully on-shell amplitudes, while oval boxes have one leg off shell (factors of G^∞ ensure only on-shell amplitudes are needed in $\mathcal{D}$). Finally, blue and pink filling indicate, respectively, LO and NLO quantities. We leave implicit that we add (LO $\leftrightarrow$ NLO) for OPE and s-OPE, and take the real parts of all quantities.

In order to compute $\mathcal{K}_{\mathrm{df},3}$ at NLO we will use the same master equation as in ref. [47], $\mathbf{K}_{\mathrm{df},3} = \operatorname{Re}\mathbf{M}_{\mathrm{df},3}$, with the main exception being that quantities in boldface denote matrices in isospin space, as described in the next section. In the case of generic three-pion isospin, this calculation is split in several parts:

$$\operatorname{Re}\mathbf{M}_{\mathrm{df},3} = \operatorname{Re}\mathbf{M}_{\mathrm{df},3}^{\mathrm{non\text{-}OPE}} - \operatorname{Re}\mathbf{D}^{\mathrm{BH}} + \left(\operatorname{Re}\mathbf{M}_{\mathrm{df},3}^{\mathrm{OPE}} - \operatorname{Re}\mathbf{D}^{\mathrm{NLO,OPE}}\right) + \operatorname{Re}\mathbf{M}_{\mathrm{df},3}^{\mathrm{s\text{-}OPE}}. \quad \text{(v.2.1)}$$

This decomposition, schematically shown in fig. v.i, is similar to the one in ref. [47], except that the we explicitly account for the 's-channel OPE' (s-OPE) contribution, which is only present at $I = 1$ since the entire 3-particle isospin is transferred to a single pion. The other contributions are the non-OPE part that does not require subtraction, the bull's head subtraction and the OPE part with its subtraction.

2.2   States and channels

Single pions are typically presented in the isospin or charge basis, which is the one used to define multi-pion states, as they can be easily combined using Clebsh-Gordan coefficients. However, one usually needs to relate them to the flavor basis (in which the matrix of pseudo-Nambu–Goldstone boson states is $|\phi\rangle = \sum_i \sigma^i |i\rangle$ with σ^i being the Pauli matrices) in order to determine scattering amplitudes from effective models. The two bases are related through

$$|\pi^\pm\rangle = \mp \frac{|1\rangle \pm i|2\rangle}{\sqrt{2}} , \qquad |\pi^0\rangle = |3\rangle . \tag{v.2.2}$$

where we use the Condon–Shortley sign convention.

To study three-pion states, we will follow the approach presented in ref. [20] and use states with zero electric charge, since these occur for all three-particle isospins.[1] Assuming isospin is an exact symmetry, exactly the same results will be obtained from states of different charge within the same isospin multiplet. Thus, in the *charge basis* we order the 7 zero-charge states as

$$|\pi\pi\pi\rangle = \begin{pmatrix} |\pi^-\pi^0\pi^+\rangle \\ |\pi^0\pi^-\pi^+\rangle \\ |\pi^-\pi^+\pi^0\rangle \\ |\pi^0\pi^0\pi^0\rangle \\ |\pi^+\pi^-\pi^0\rangle \\ |\pi^0\pi^+\pi^-\rangle \\ |\pi^+\pi^0\pi^-\rangle \end{pmatrix} . \tag{v.2.3}$$

This implies that all quantities appearing in this derivation, unless otherwise stated, are 7×7 matrices in the space of three-pion states.

In many parts of the calculation, it is more useful to rotate to a basis of states with definite three-particle isospin. This rotation is not unique, however; the choice made for the most part in ref. [20] is to let the first two particles form states of definite two-particle isospin, which we label $|\sigma\rangle$ ($I_{\pi\pi} = 0$, the channel where the σ resonance is present), $|\rho\rangle$ ($I_{\pi\pi} = 1$, where the ρ resonance is present) and $|\Pi\rangle$ ($I_{\pi\pi} = 2$, with no resonances). In this *isospin*

---

[1]We have used higher-charge states for some cross-checks, including of course the maximum-isospin $|\pi^+\pi^+\pi^+\rangle$ state studied in ref. [47].

basis,

$$
|\pi\pi\pi\rangle =
\begin{pmatrix}
|\Pi\pi\rangle_3 \\
|\Pi\pi\rangle_2 \\
|\rho\pi\rangle_2 \\
|\Pi\pi\rangle_1 \\
|\rho\pi\rangle_1 \\
|\sigma\pi\rangle_1 \\
|\sigma\pi\rangle_0
\end{pmatrix}
=
\begin{pmatrix}
\frac{1}{\sqrt{5}}\left(|\Pi^+\pi^-\rangle + \sqrt{3}|\Pi^0\pi^0\rangle + |\Pi^-\pi^+\rangle\right) \\
\frac{1}{\sqrt{2}}\left(|\Pi^+\pi^-\rangle - |\Pi^-\pi^+\rangle\right) \\
\frac{1}{\sqrt{6}}\left(|\rho^+\pi^-\rangle + 2|\rho^0\pi^0\rangle + |\rho^-\pi^+\rangle\right) \\
\frac{1}{\sqrt{10}}\left(\sqrt{3}|\Pi^+\pi^-\rangle - 2|\Pi^0\pi^0\rangle + \sqrt{3}|\Pi^-\pi^+\rangle\right) \\
\frac{1}{\sqrt{2}}\left(|\rho^+\pi^-\rangle - |\rho^-\pi^+\rangle\right) \\
|\sigma\pi^0\rangle \\
\frac{1}{\sqrt{3}}\left(|\rho^+\pi^-\rangle - |\rho^0\pi^0\rangle + |\rho^-\pi^+\rangle\right)
\end{pmatrix},
\qquad \text{(v.2.4)}
$$

where the subscripts indicate three-particle isospin, and the specific two-pion states are

$$
|\sigma\rangle = \frac{|\pi^+\pi^-\rangle + |\pi^-\pi^+\rangle - |\pi^0\pi^0\rangle}{\sqrt{3}}
\qquad
|\Pi^0\rangle = \frac{|\pi^+\pi^-\rangle + |\pi^-\pi^+\rangle + 2|\pi^0\pi^0\rangle}{\sqrt{6}}
$$

$$
|\rho^0\rangle = \frac{|\pi^+\pi^-\rangle - |\pi^-\pi^+\rangle}{\sqrt{2}}
\qquad
|\Pi^\pm\rangle = \frac{|\pi^\pm\pi^0\rangle + |\pi^0\pi^\pm\rangle}{\sqrt{2}}
\qquad \text{(v.2.5)}
$$

$$
|\rho^\pm\rangle = \pm\frac{|\pi^\pm\pi^0\rangle - |\pi^0\pi^\pm\rangle}{\sqrt{2}}
\qquad
|\Pi^{\pm\pm}\rangle = |\pi^\pm\pi^\pm\rangle.
$$

$\mathbf{K}_{df,3}$ and all other relevant quantities block-diagonalize in this basis, as is described in detail in ref. [20].

Yet another basis, which relates more directly to the threshold expansion, aligns the states with irreps of the S_3 group describing permutations of the three particles, still with definite isospin. States are denoted $|\chi_s\rangle$ for the trivial (symmetric) irrep, $|\chi_a\rangle$ for the alternating irrep, and $|\chi_1\rangle, |\chi_2\rangle$ for the two-dimensional standard irrep; the details of the irreps are given in appendix C of ref. [20]. Of these irreps, $I = 3$ is in the trivial, $I = 0$ in the alternating, $I = 2$ in the standard, and $I = 1$ in a direct sum of the trivial and standard. Thus, only for $I = 1$ does the isospin basis differ from this *symmetric basis*, where[2]

$$
|\pi\pi\pi\rangle =
\begin{pmatrix}
|\chi_s\rangle_3 \\
|\chi_1\rangle_2 \\
|\chi_2\rangle_2 \\
|\chi_s\rangle_1 \\
|\chi_1\rangle_1 \\
|\chi_2\rangle_1 \\
|\chi_a\rangle_0
\end{pmatrix}
=
\begin{pmatrix}
|\Pi\pi\rangle_3 \\
|\Pi\pi\rangle_2 \\
|\rho\pi\rangle_2 \\
\frac{2}{3}|\Pi\pi\rangle_1 + \frac{\sqrt{5}}{3}|\sigma\pi\rangle_1 \\
|\rho\pi\rangle_1 \\
-\frac{\sqrt{5}}{3}|\Pi\pi\rangle_1 + \frac{2}{3}|\sigma\pi\rangle_1 \\
|\chi_a\rangle_0
\end{pmatrix}.
\qquad \text{(v.2.6)}
$$

The rotation from the charge basis to the irrep and symmetric bases is given by the orthogonal matrices $\mathcal{C}_I$ and $\mathcal{C}_S$, respectively (the former stated in eq. (2.60) of ref. [20]), given

---

[2]This is eqs. (C.11) to (C.19) of ref. [20], and eq. (v.2.4) is eqs. (C.1) to (C.9). Note the ordering of the $I = 1$ states, which is consistent with most of ref. [20] but not with eqs. (C.11) to (C.19) thereof. Note also the differing sign conventions.

by

$$
\mathcal{C}_{\mathrm{I}} = \mathcal{C}
\begin{pmatrix}
1 & 1 & 1 & 2 & 1 & 1 & 1 \\
-1 & -1 & 0 & 0 & 0 & 1 & 1 \\
-1 & 1 & -2 & 0 & 2 & -1 & 1 \\
3 & 3 & -2 & -4 & -2 & 3 & 3 \\
\sqrt{3} & -\sqrt{3} & 0 & 0 & 0 & -\sqrt{3} & \sqrt{3} \\
0 & 0 & 2 & -2 & 2 & 0 & 0 \\
-1 & 1 & 1 & 0 & -1 & -1 & 1
\end{pmatrix},
$$

$$
\mathcal{C}_{\mathrm{S}} = \mathcal{C}
\begin{pmatrix}
1 & 1 & 1 & 2 & 1 & 1 & 1 \\
-1 & -1 & 0 & 0 & 0 & 1 & 1 \\
-1 & 1 & -2 & 0 & 2 & -1 & 1 \\
2 & 2 & 2 & -6 & 2 & 2 & 2 \\
-1 & -1 & 2 & 0 & 2 & -1 & -1 \\
\sqrt{3} & -\sqrt{3} & 0 & 0 & 0 & -\sqrt{3} & \sqrt{3} \\
-1 & 1 & 1 & 0 & -1 & -1 & 1
\end{pmatrix},
$$

(v.2.7)

with $\mathcal{C} = \mathrm{diag}\left(\frac{1}{\sqrt{10}}, \frac{1}{2}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{60}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{6}}\right)$ pulling out common coefficients.

2.3 The threshold expansion

Here, we write down the parametrization of the threshold expansion. The fundamental building blocks of this parametrization are, following ref. [14],

$$
\tilde{t}_{ij} \equiv \frac{(p_i - k_j)^2}{9M_\pi^2}, \qquad \Delta_i \equiv \frac{(P - k_i)^2 - 4M_\pi^2}{9M_\pi^2}, \qquad \Delta \equiv \frac{P^2 - 9M_\pi^2}{9M_\pi^2}, \qquad \text{(v.2.8)}
$$

plus Δ'_i, which is the analogue of Δ_i obtained by substituting $k_i \to p_i$.[3] All of these are considered to be $\mathcal{O}(\Delta)$ in the expansion. They are related through

$$
\Delta = -\tfrac{1}{2} \sum_{i,j} \tilde{t}_{ij}, \qquad \Delta_j = \Delta + \sum_i \tilde{t}_{ij}, \qquad \Delta'_i = \Delta + \sum_j \tilde{t}_{ij},
$$

where all sums run from 1 to 3.

In the following, we restate the threshold expansions derived in refs. [14, 20], plus a few terms not present in those publications. Similarly to ref. [47], we somewhat simplify the notation for the $\mathcal{K}$ coefficients, and furthermore depart from ref. [20] in defining all operators (Δ, $\vec{\xi}$, etc.) to be dimensionless.

2.3.1 $I_{\pi\pi} = 3$

Through quadratic order in Δ, we have the five terms computed in ref. [47]:

$$
M_\pi^2 \mathbf{K}_{\mathrm{df},3}^{[I=3]} = \mathcal{K}_0 + \mathcal{K}_1 \Delta + \mathcal{K}_2 \Delta^2 + \mathcal{K}_A \Delta_A + \mathcal{K}_B \Delta_B + \mathcal{O}(\Delta^3), \qquad \text{(v.2.9)}
$$

---

[3]Throughout the following, a prime always refers to this substitution.

where

$$\Delta_{\mathrm{A}} \equiv \sum_i \left(\Delta_i^2 + \Delta_i'^2 \right) - \Delta^2 \,, \qquad \Delta_{\mathrm{B}} \equiv \sum_{i,j} \tilde{t}_{ij}^2 - \Delta^2 \,. \tag{v.2.10}$$

2.3.2 $I_{\pi\pi} = 2$

This channel involves a two-dimensional flavor space, so all operators need to be doublets that transform under the standard representation of S_3. Following the basis choice of ref. [20], the initial-state doublet at linear order in momenta is

$$\vec{\xi}^\mu = \left(\xi_1^\mu, \xi_2^\mu \right), \qquad \text{where} \quad \xi_1 \equiv \frac{2k_3 - k_1 - k_2}{\sqrt{6} M_\pi} \,, \quad \xi_2 \equiv \frac{k_2 - k_1}{\sqrt{2} M_\pi} \,. \tag{v.2.11}$$

Still following ref. [20], quadratic order introduces three Lorentz tensor doublets, of which only the following two are relevant here:

$$\vec{\xi}(S)^{\mu\nu} \equiv \frac{\vec{\xi}^\mu P^\nu + \vec{\xi}^\nu P^\mu}{M_\pi} \,, \qquad \vec{\xi}(\bar{S})^{\mu\nu} \equiv \left(\xi_2^\mu \xi_2^\nu - \xi_1^\mu \xi_1^\nu, \ \xi_1^\mu \xi_2^\nu + \xi_2^\mu \xi_1^\nu \right). \tag{v.2.12}$$

There is also one Lorentz scalar doublet,

$$\vec{\xi}^{(2)} = \left(\xi_1^{(2)}, \xi_2^{(2)} \right), \qquad \text{where} \quad \xi_1^{(2)} \equiv \frac{2\Delta_3 - \Delta_1 - \Delta_2}{\sqrt{6}} \,, \quad \xi_2^{(2)} \equiv \frac{\Delta_2 - \Delta_1}{\sqrt{2}} \,, \tag{v.2.13}$$

which has the property $\vec{\xi}^{(2)} = -\frac{2}{9M_\pi} \vec{\xi}^\mu P_\mu$. The ξ's and their primed counterparts form four independent tensors in isospin space (one at linear order and three at quadratic); here, we label these using

$$\begin{aligned}
\boldsymbol{\Xi}_1 &\equiv \vec{\xi}'^\mu \otimes \vec{\xi}_\mu \,, & \boldsymbol{\Xi}_2 &\equiv \vec{\xi}'^{(2)} \otimes \vec{\xi}^{(2)} \,, \\
\boldsymbol{\Xi}_3 &\equiv \frac{1}{\sqrt{6}} \left[\vec{\xi}'(\bar{S})^{\mu\nu} \otimes \vec{\xi}(S)_{\mu\nu} + \vec{\xi}'(S)^{\mu\nu} \otimes \vec{\xi}(\bar{S})_{\mu\nu} \right], & \boldsymbol{\Xi}_4 &\equiv \vec{\xi}'(\bar{S})^{\mu\nu} \otimes \vec{\xi}(\bar{S})_{\mu\nu}
\end{aligned} \tag{v.2.14}$$

where $\otimes$ indicates a tensor product like

$$\vec{\xi}'^\mu \otimes \vec{\xi}_\mu = \begin{pmatrix} \xi_1' \cdot \xi_1 & \xi_1' \cdot \xi_2 \\ \xi_2' \cdot \xi_1 & \xi_2' \cdot \xi_2 \end{pmatrix} = \boldsymbol{\Xi}_1 \tag{v.2.15}$$

and we have pulled out a factor of $\sqrt{6}$ in the definition of $\boldsymbol{\Xi}_3$, since this would otherwise appear in all our results. This allows the threshold expansion to be written

$$M_\pi^2 \mathbf{K}_{\mathrm{df},3}^{[I=2]} = \left(\mathcal{K}_0^{\mathrm{T}} + \mathcal{K}_1^{\mathrm{T}} \Delta \right) \boldsymbol{\Xi}_1 + \sum_{n=2,3,4} \mathcal{K}_n^{\mathrm{T}} \boldsymbol{\Xi}_n + \mathcal{O}(\Delta^3) \,, \tag{v.2.16}$$

where 'T' stands for 'isotensor'.

2.3.3 $I_{\pi\pi\pi} = 1$

Here, the flavor space is three-dimensional. Following ref. [20], we decompose the states into a singlet and a doublet, transforming under the trivial and standard representation of S_3, respectively, and put the singlet as the first component. Thus, in block form, we have

$$\mathbf{K}_{\text{df},3}^{[I=1]} = \begin{pmatrix} \mathbf{K}_{\text{df},3}^{[I=1,\text{SS}]} & \mathbf{K}_{\text{df},3}^{[I=1,\text{SD}]} \\ \mathbf{K}_{\text{df},3}^{[I=1,\text{DS}]} & \mathbf{K}_{\text{df},3}^{[I=1,\text{DD}]} \end{pmatrix}. \tag{v.2.17}$$

The singlet-singlet (SS) sector is similar to the $I = 3$ case, eq. (v.2.9):

$$M_\pi^2 \mathbf{K}_{\text{df},3}^{[I=1,\text{SS}]} = \mathcal{K}_0^{\text{SS}} + \mathcal{K}_1^{\text{SS}}\Delta + \mathcal{K}_2^{\text{SS}}\Delta^2 + \mathcal{K}_A^{\text{SS}}\Delta_A + \mathcal{K}_B^{\text{SS}}\Delta_B + \mathcal{O}(\Delta^3), \tag{v.2.18}$$

whereas the doublet-doublet (DD) sector is similar to the $I = 2$ case, eq. (v.2.16):

$$M_\pi^2 \mathbf{K}_{\text{df},3}^{[I=1,\text{DD}]} = \left(\mathcal{K}_0^{\text{DD}} + \mathcal{K}_1^{\text{DD}}\Delta\right)\mathbf{\Xi}_1 + \sum_{n=2,3,4} \mathcal{K}_n^{\text{DD}}\,\mathbf{\Xi}_n + \mathcal{O}(\Delta^3). \tag{v.2.19}$$

At $\mathcal{O}(\Delta)$, the sole operator that fits the singlet-doublet (SD) mixing sector is $\vec{\xi}^{(2)}$, defined in eq. (v.2.13). At $\mathcal{O}(\Delta^2)$, new operators are needed that are not included in ref. [20]. They are constructed by taking the following building blocks, which are singlets under permutations of the final-state momenta,

$$\Delta\Delta_i\,, \qquad \Delta_i\Delta_j\,, \qquad \sum_i \tilde{t}_{ij}\tilde{t}_{ik}\,, \tag{v.2.20}$$

and forming doublets under permutations of the initial-state momenta. From $\Delta\Delta_i$, we simply obtain $\Delta\xi^{(2)}$, while $\Delta_i\Delta_j$ and $\sum_i \tilde{t}_{ij}\tilde{t}_{ik}$ yield two operators each:[4]

$$\vec{\xi}^{(4,n)} = \left(\xi_1^{(4,n)},\, \xi_2^{(4,n)}\right), \qquad \text{where} \quad n = 2, 3, 4 \quad \text{and}$$

$$\begin{cases} \xi_1^{(4,2)} \equiv \dfrac{2\Delta_3^2 - \Delta_1^2 - \Delta_2^2}{\sqrt{6}}\,, & \xi_2^{(4,2)} \equiv \dfrac{\Delta_2^2 - \Delta_1^2}{\sqrt{2}}\,, \\[2ex] \xi_1^{(4,3)} \equiv \displaystyle\sum_i \dfrac{(\tilde{t}_{i1} + \tilde{t}_{i2})\tilde{t}_{i3} - 2\tilde{t}_{i1}\tilde{t}_{i2}}{\sqrt{6}}\,, & \xi_2^{(4,3)} \equiv \displaystyle\sum_i \dfrac{(\tilde{t}_{i2} - \tilde{t}_{i1})\tilde{t}_{i3}}{\sqrt{2}}\,, \\[2ex] \xi_1^{(4,4)} \equiv \displaystyle\sum_i \dfrac{2\tilde{t}_{i3}^2 - \tilde{t}_{i1}^2 - \tilde{t}_{i2}^2}{\sqrt{6}}\,, & \xi_2^{(4,4)} \equiv \displaystyle\sum_i \dfrac{\tilde{t}_{i2}^2 - \tilde{t}_{i1}^2}{\sqrt{2}}\,. \end{cases} \tag{v.2.21}$$

Thus,

$$\frac{M_\pi^2}{\sqrt{30}}\mathbf{K}_{\text{df},3}^{[I=1,\text{SD}]} = \left(\mathcal{K}_0^{\text{SD}} + \mathcal{K}_1^{\text{SD}}\Delta\right)\vec{\xi}^{(2)} + \sum_{n=2,3,4} \mathcal{K}_n^{\text{SD}}\vec{\xi}^{(4,n)} + \mathcal{O}(\Delta^3). \tag{v.2.22}$$

---

[4]Throughout, we use the sign convention established by ref. [20] with $\vec{\xi}^{(2)}$, so that the format of $\xi_2^{(4,n)}$ is '$2 - 1$' rather than '$1 - 2$'. It is also possible to form the analog of $\vec{\xi}^{(4,3)}$ using $\Delta_i\Delta_j$, but that is equal to $\vec{\xi}^{(2)}\Delta - \vec{\xi}^{(4,2)}$ and is therefore redundant.

$\mathbf{K}_{\mathrm{df},3}^{[I=1,\mathrm{DS}]}$ is obtained from this by exchanging $p_i \leftrightarrow k_i$ and taking the transpose. We have pulled out a factor of $\sqrt{30}$ in the definition of $\mathbf{K}_{\mathrm{df},3}^{[I=1,\mathrm{SD}]}$, since this would otherwise appear in all our results.

2.3.4 $I_{\pi\pi\pi} = 0$

Here, all operators must be totally antisymmetric under permutations of the momenta, which puts the leading order at $\mathcal{O}(\Delta^2)$ and makes $\mathcal{O}(\Delta^3)$ contributions simple enough to include, unlike in the other channels (see appendix B). The threshold expansion is

$$M_\pi^2 \mathbf{K}_{\mathrm{df},3}^{[I=0]} = \left(\mathcal{K}_0^{\mathrm{AS}} + \mathcal{K}_1^{\mathrm{AS}}\Delta\right)\Delta_{\mathrm{AS}}^{(2)} + \mathcal{K}_3^{\mathrm{AS}}\Delta_{\mathrm{AS}}^{(3)} + \mathcal{K}_4^{\mathrm{AS}}\Delta_{\mathrm{AS}}^{(4)} + \mathcal{O}(\Delta^4), \qquad (\text{v.2.23})$$

where 'AS' stands for 'antisymmetric', and the operators are

$$\Delta_{\mathrm{AS}}^{(2)} \equiv \sum_{\substack{i,j,k \\ m,n,r}} \epsilon_{ijk}\epsilon_{mnr}\tilde{t}_{im}\tilde{t}_{jn},$$

$$\Delta_{\mathrm{AS}}^{(3)} \equiv \sum_{\substack{i,j,k \\ m,n,r}} \epsilon_{ijk}\epsilon_{mnr}\tilde{t}_{im}\tilde{t}_{jn}\tilde{t}_{kr}, \qquad \Delta_{\mathrm{AS}}^{(4)} \equiv \sum_{\substack{i,j,k \\ m,n,r}} \epsilon_{ijk}\epsilon_{mnr}\tilde{t}_{im}\tilde{t}_{jn}\left(\tilde{t}_{im} + \tilde{t}_{jn}\right), \qquad (\text{v.2.24})$$

of which $\Delta_{\mathrm{AS}}^{(4)}$ was missed in the analysis of ref. [20].

3 The calculation of $\mathcal{K}_{\mathrm{df},3}$

Here, we describe the explicit calculation of $\mathcal{K}_{\mathrm{df},3}$. It largely follows the same lines as that performed in ref. [47], and we refer the reader there for most of the procedural details.

3.1 Leading-order calculation

We start with the calculation at LO. Some of the details are the same for the NLO case, but it is illustrative to show them explicitly in the simpler case. In particular, the bull's head subtraction in eq. (v.2.1) is absent at LO. We thus split the calculation in three-parts, each with a subsection: the OPE contribution, the s-channel OPE contribution, and the non-OPE part.

3.1.1 OPE contributions

In terms of flavor indices, the two-particle amplitude for $\phi^a(k_1)\phi^b(k_2) \rightarrow \phi^c(p_3)\phi^d(b)$ decomposes as [50]

$$\mathcal{M}_2(s,t,u) = \delta^{ab}\delta^{cd}A(s,t,u) + \delta^{ac}\delta^{bd}A(t,u,s) + \delta^{ad}\delta^{bc}A(u,s,t), \qquad (\text{v.3.1})$$

where

$$s \equiv (k_1 + k_2)^2, \qquad t \equiv (k_1 - p_3)^2, \qquad u \equiv (k_2 - p_3)^2, \qquad \text{(v.3.2)}$$

and the function $A(s, t, u)$ is symmetric in its last two arguments. The leg with momentum b may be off-shell, in which case we use the off-shell convention of ref. [48] to express A. Since the choice of the first argument of A completely fixes A, we abbreviate: $A(t) \equiv A(t, u, s)$, and similarly for $A(s)$ and $A(u)$.

Letting $\mathcal{M}_2$ act on the first two particles in a three-pion state with the third particle as spectator, considering which particle exchanges are valid reveals a block-diagonal form in the charge basis, eq. (v.2.3):

$$\mathbf{M}_2 = \begin{pmatrix} \mathbf{M}_2^+ & & \\ & \mathbf{M}_2^0 & \\ & & \mathbf{M}_2^- \end{pmatrix}, \qquad \text{with} \qquad \mathbf{M}_2^+ = \mathbf{M}_2^- = \begin{pmatrix} A(t) & A(u) \\ A(u) & A(t) \end{pmatrix} \qquad \text{(v.3.3)}$$

$$\text{and} \qquad \mathbf{M}_2^0 = \begin{pmatrix} A(s) + A(t) & -A(s) & A(s) + A(u) \\ -A(s) & A(s) + A(t) + A(u) & -A(s) \\ A(s) + A(u) & -A(s) & A(s) + A(t) \end{pmatrix}.$$

(a) OPE diagram   **(b)** Contact diagram   **(c)** s-channel OPE diagram
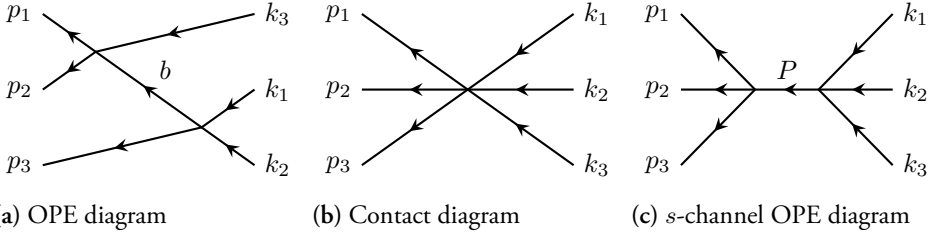
Figure v.2: Feynman diagrams contributing to $\mathcal{M}_3$ at LO. For diagram (a), there are an additional eight diagrams corresponding to the symmetrization of initial and final momenta. Diagram (c) only contributes at $I = 1$.

At LO, the OPE contributions come from the symmetrization of fig. v.2a, and from the matching subtraction term. Thus, the OPE contribution to the unsymmetrized and divergence-free amplitude, $\mathbf{M}_{\mathrm{df},3}^{(u,u)}$, is

$$\mathbf{M}_{\mathrm{df},3}^{(u,u),\mathrm{OPE}} = -\mathbf{M}_{2,\mathrm{off}} \frac{\mathbf{T}_G}{\bar{b}^2 + i\epsilon} \mathbf{M}_{2,\mathrm{off}} + \sum_{\ell'\ell} \mathbf{M}_{2,\mathrm{on}}^{\ell'} \mathbf{T}_G G_{\ell'\ell}^{\infty} \mathbf{M}_{2,\mathrm{on}}^{\ell}, \qquad \text{(v.3.4)}$$

where $\bar{b}^2 \equiv b^2 - M_\pi^2$, $\mathbf{M}_{2,\mathrm{on}}^{\ell}$ is the partial-wave projected on-shell scattering amplitude, and the definition of $G_{\ell'\ell}^{\infty}$ can be found in eq. (2.12) of ref. [47]. Note that in this equation, the momentum dependence of the two-pion amplitudes is left implicit: the amplitudes to the left[5] depend on the outgoing $\{p_i\}$ momenta, and those to the right on the incoming momenta $\{k_i\}$. We also emphasize that in $\mathbf{M}_{2,\mathrm{off}}$, the momentum of the exchanged particle

---

[5]Strictly speaking, it should appear transposed in eq. (v.3.4), but we have chosen our bases such that it is symmetric.

b may be off-shell, while in $\mathbf{M}_{2,\mathrm{on}}$ everything is kept on-shell. The matrix $\mathbf{T}_G$ indicates valid exchanges between states; in the charge basis, it is

$$
\mathbf{T}_G =
\begin{pmatrix}
\square & \square & \square & \square & \square & \square & \blacksquare \\
\square & \square & \square & \square & \blacksquare & \square & \square \\
\square & \square & \square & \square & \square & \blacksquare & \square \\
\square & \square & \square & \blacksquare & \square & \square & \square \\
\square & \blacksquare & \square & \square & \square & \square & \square \\
\square & \square & \blacksquare & \square & \square & \square & \square \\
\blacksquare & \square & \square & \square & \square & \square & \square
\end{pmatrix}, \qquad \square = 0, \quad \blacksquare = 1,
\tag{v.3.5}
$$

using squares rather than numbers for legibility. Expanding A to leading order in ChPT, we find

$$
\mathbf{M}_2^{\mathrm{LO}} = \mathbf{k}_0 + \mathbf{k}_2 \bar{s} + \mathbf{k}_2(t+u) + \mathbf{k}_3(t-u),
\tag{v.3.6}
$$

where $\bar{s} \equiv s - 4M_\pi^2$, and the coefficients $\mathbf{k}_i$ are straightforward to compute in the charge basis for a given A amplitude. In the subtraction, we need to separate the on-shell two-pion amplitude into partial waves. At LO, only s and p waves appear. The separation can thus be performed simply by dividing the amplitude into a symmetric and an antisymmetric part,

$$
\begin{aligned}
\mathbf{M}_2^s(s,t,u) &= \tfrac{1}{2}\left[\mathbf{M}_2(s,t,u) + \mathbf{M}_2(s,u,t)\right] = \mathbf{k}_0 + \mathbf{k}_2 \bar{s} + \mathbf{k}_2(t+u), \\
\mathbf{M}_2^p(s,t,u) &= \tfrac{1}{2}\left[\mathbf{M}_2(s,t,u) - \mathbf{M}_2(s,u,t)\right] = \mathbf{k}_3(t-u).
\end{aligned}
\tag{v.3.7}
$$

To the ChPT order we are working (and also at NLO) the p-wave amplitude is proportional to $t - u$, which can be expanded using the addition theorem for spherical harmonics to

$$
t - u = 4\boldsymbol{p}_k^* \cdot \boldsymbol{a}_k^* = 4p_k^* q_{2k}^* \left[\frac{4\pi}{3}\sum_m Y_{1m}^*(\hat{\boldsymbol{a}}_k^*)Y_{1m}(\hat{\boldsymbol{p}}_k^*)\right].
\tag{v.3.8}
$$

In our off-shell prescription, this implies that, as is the case for d-waves [47], that the difference between the on- and off-shell p-wave amplitudes is entirely made up of the barrier factors:

$$
\mathbf{M}_{2,\mathrm{off}}^p(\{k_i\}) = \mathbf{M}_{2,\mathrm{on}}^p(\{k_i\})\left(\frac{p_k^*}{q_{2k}^*}\right), \qquad
\mathbf{M}_{2,\mathrm{off}}^p(\{p_i\}) = \mathbf{M}_{2,\mathrm{on}}^p(\{p_i\})\left(\frac{k_p^*}{q_{2p}^*}\right).
\tag{v.3.9}
$$

In order to compute the subtracted result, we separate the s-wave part into an on-shell and an off-shell part, the latter of which is proportional to $\bar{b}^2$ as can be seen by applying the off-shell relation $t + u = 4M_\pi^2 - s + \bar{b}^2$. Taking this into account, we write

$$
\mathbf{M}_{2,\mathrm{off}} = \mathbf{M}_{2,\mathrm{on}}^s + \mathbf{M}_{2,\mathrm{off}}^p + \bar{b}^2\,\delta\mathbf{M}_2^s.
\tag{v.3.10}
$$

Then, the unsymmetrized divergence-free OPE amplitude is

$$
\mathbf{M}_{\mathrm{df},3}^{(u,u),\mathrm{OPE,LO}} = -\delta\mathbf{M}_2^s \frac{\bar{b}^2}{\bar{b}^2}\mathbf{T}_G \mathbf{M}_{2,\mathrm{off}} - \mathbf{M}_{2,\mathrm{off}} \frac{\bar{b}^2}{\bar{b}^2}\mathbf{T}_G \delta\mathbf{M}_2^s + \delta\mathbf{M}_2^s \frac{\bar{b}^4}{\bar{b}^2}\mathbf{T}_G \delta\mathbf{M}_2^s.
\tag{v.3.11}
$$

This can be computed easily, but we need not show this intermediate result.

To obtain the full contribution, the symmetrization procedure must be performed. In the case of general three-pion isospin, it is slightly more complicated than in ref. [47], since it must be done for momentum and flavor simultaneously. This is achieved by using

$$\mathbf{M}_{\mathrm{df},3} = \sum_{m=0}^{2} \sum_{n=0}^{2} \left(\mathbf{R}^m\right)^{\mathsf{T}} \mathbf{M}_{\mathrm{df},3}^{(u,u)} \left(R^m\{p_i\}, R^n\{k_i\}\right) \mathbf{R}^n \,, \tag{v.3.12}$$

where $R\{p_1, p_2, p_3\} = \{p_2, p_3, p_1\}$ is a cyclic permutation (due to the symmetry of the interacting pair, only the cyclic subgroup of S_3 needs to be considered) and $\mathbf{R}$ enacts the same permutation on the space of three-pion states. Following ref. [20], its form in the charge basis is

$$\mathbf{R} = \begin{pmatrix} \square & \square & \square & \square & \blacksquare & \square & \square \\ \square & \square & \square & \square & \square & \square & \blacksquare \\ \square & \blacksquare & \square & \square & \square & \square & \square \\ \square & \square & \square & \blacksquare & \square & \square & \square \\ \square & \square & \square & \square & \square & \blacksquare & \square \\ \blacksquare & \square & \square & \square & \square & \square & \square \\ \square & \square & \blacksquare & \square & \square & \square & \square \end{pmatrix} \,, \qquad \square = 0 \,, \quad \blacksquare = 1 \,. \tag{v.3.13}$$

In the symmetric basis, it instead takes the block-diagonal form

$$\mathbf{R} = \mathrm{diag}\left(1, \mathbf{R}_2, 1, \mathbf{R}_2, 1\right), \qquad \mathbf{R}_2 = \frac{1}{2} \begin{pmatrix} -1 & -\sqrt{3} \\ +\sqrt{3} & -1 \end{pmatrix} \,. \tag{v.3.14}$$

corresponding to the distribution of one- and two-dimensional irreps.

After symmetrization and conversion of the kinematical variables to $\tilde{t}_{ij}$, we can identify the terms in the threshold expansion. At LO, this can be done by inspection since there is only one term per order in the threshold expansion in each isospin sector; at higher orders, it requires the solution of systems of equations. The results are listed in table v.1. Note that most of the contributions are pure s-wave; all pure p-wave contributions cancel, and only $\mathcal{K}_0^{\mathrm{T}}$ and $\mathcal{K}_0^{\mathrm{DD}}$ get contributions from mixed s- and p-wave diagrams (amounting to 9 out of the total $21/2$ in both cases).

3.1.2 s-channel OPE contributions

The s-channel OPE diagram, fig. v.2c, needs no subtraction since the exchanged momentum $P = k_1 + k_2 + k_3$ is off-shell in the kinematic range of interest. This contribution appears only in the $I = 1$ channel, and for zero-charge states, the exchanged particle must be a π^0. Thus, the s-channel OPE amplitude can be factorized as follows:

$$\mathbf{M}_{\mathrm{df},3}^{\mathrm{s\text{-}OPE,LO}} = -\boldsymbol{v}_{\mathrm{LO}}\left(\{p_i\}\right) \frac{1}{P^2 - M_\pi^2} \boldsymbol{v}_{\mathrm{LO}}^\dagger\left(\{k_i\}\right), \tag{v.3.15}$$

where v is a column vector of $\pi\pi\pi \to \pi^0$ amplitudes from all seven states in the charge basis, with the π^0 off-shell. It can be computed from the amplitude introduced in eq. (v.3.1), giving

$$
\begin{aligned}
F_\pi^2 \mathcal{M}_2^{\text{LO}}[\pi^0(k_1)\pi^0(k_2)\pi^0(k_3) \to \pi^0(P)] &= s_{12} + s_{23} + s_{13} - 3M_\pi^2 \,, \\
F_\pi^2 \mathcal{M}_2^{\text{LO}}[\pi^+(k_1)\pi^0(k_2)\pi^0(k_3) \to \pi^0(P)] &= M_\pi^2 - s_{13} \,,
\end{aligned}
\tag{v.3.16}
$$

where $s_{ij} \equiv (k_i + k_j)^2$. After taking the relevant permutations, rotating to the symmetric basis and converting to threshold expansion parameters, we get in the $I = 1$ sector of the symmetric basis

$$
\boldsymbol{v}_{\text{LO}}(\{k_i\}) = \frac{M_\pi^2}{F_\pi^2} \begin{pmatrix} -3\sqrt{15}(1+\Delta) \\ 81\xi_1^{(2)}/\sqrt{2} \\ 81\xi_2^{(2)}/\sqrt{2} \end{pmatrix} \,,
\tag{v.3.17}
$$

where $\vec{\xi}^2$ is given in eq. (v.2.13). We also need the threshold expansion of the single-particle propagator,

$$
\frac{1}{P^2 - M_\pi^2} = \frac{1}{M_\pi^2(8 - 9\Delta)} = \frac{1}{8M_\pi^2}\left(1 - \frac{9}{8}\Delta + \frac{81}{64}\Delta^2 + \mathcal{O}(\Delta^3)\right) \,.
\tag{v.3.18}
$$

Note that this expansion formally sets the radius of convergence of the threshold expansion at $|\Delta| \leq 8/9$.

Using these expressions we can directly identify the coefficients in the threshold expansion after multiplication. The results up to quadratic order are listed in table v.1. Note that some terms will appear at higher orders in the threshold expansion, but we will not consider them here; work to check their effect numerically is ongoing.

3.1.3 non-OPE contributions

At leading order, the non-OPE part comes from the diagram fig. v.2b, and has a simple form given in ref. [48]. In the charge basis, the amplitude matrix follows by crossing from

$$
\begin{aligned}
\mathcal{M}_3[\pi^0\pi^0\pi^0 \to \pi^0\pi^0\pi^0] &= 27M_\pi^2 \,, \\
\mathcal{M}_3[\pi^0\pi^0\pi^0 \to \pi^+\pi^0\pi^-] &= 5M_\pi^2 - 3s'_{13} - t_{12} - t_{22} - t_{32} \,, \\
\mathcal{M}_3[\pi^+\pi^0\pi^- \to \pi^+\pi^0\pi^-] &= -6M_\pi^2 + s_{13} + s'_{13} + t_{11} + 2t_{22} + t_{33} \,,
\end{aligned}
\tag{v.3.19}
$$

where the momenta are $k_1, k_2, k_3 \to p_1, p_2, p_3$ (in that order), and

$$
s_{ij} = (k_i + k_j)^2 \,, \qquad s'_{ij} = (p_i + p_j)^2 \,, \qquad t_{ij} = (k_i - p_j)^2 \,.
\tag{v.3.20}
$$

The result is non-divergent, so no subtraction is needed. Identifying the coefficients of the threshold expansion yields the results, given in table v.1.

Table v.1: LO contributions to $\mathbf{K}_{\mathrm{df},3}$, and the contributions from different parts. There is no bull's head subtraction or cutoff dependence at this order. Note that 's-channel OPE' is not a part of 'OPE', but a separate contribution.

		Total	OPE	s-channel OPE	non-OPE
$I=3$	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_0$	18	36	0	-18
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_1$	27	63	0	-36
$I=2$	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_0^{\mathrm{T}}$	$\frac{9}{2}$	$\frac{21}{2}$	0	-6
$I=1$	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_0^{\mathrm{SS}}$	$-\frac{111}{8}$	-54	$-\frac{135}{8}$	57
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_1^{\mathrm{SS}}$	$-\frac{1137}{64}$	-27	$-\frac{945}{64}$	24
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_2^{\mathrm{SS}}$	$-\frac{135}{512}$	0	$-\frac{135}{512}$	0
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_0^{\mathrm{SD}}$	$-\frac{3}{8}$	-9	$-\frac{27}{8}$	12
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_1^{\mathrm{SD}}$	$\frac{27}{64}$	0	$\frac{27}{64}$	0
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_0^{\mathrm{DD}}$	$\frac{1}{2}$	$\frac{21}{2}$	0	-10
	$\frac{F_\pi^4}{M_\pi^4}\mathcal{K}_2^{\mathrm{DD}}$	$-\frac{81}{4}$	0	$-\frac{81}{4}$	0
$I=0$		(there are no $I=0$ contributions at this order)			

3.2 Next-to-leading-order calculation

Unlike at LO, the NLO amplitude depends on the low-energy constants (LECs) of ChPT. The four LECs that are relevant to our calculations are denoted ℓ_i^{r} for $i = 1, 2, 3, 4$, with the 'r' indicating that they are renormalized, as described more closely in ref. [47]. The renormalization scale μ appears through the quantity $L \equiv \kappa \log(M_\pi^2/\mu^2)$, with $\kappa \equiv 1/(16\pi^2)$.

3.2.1 OPE contributions

These contributions are calculated in a similar way to what is described in section 3.1.1, except that one $\pi\pi$ amplitude is promoted to NLO. Similar to eq. (v.3.6), it can be expressed as

$$
\begin{aligned}
\mathbf{M}_2^{\mathrm{NLO}} = \ &\mathbf{a}_1 + \mathbf{b}_1\bar{s} + \mathbf{b}_2(t+u) + \mathbf{b}_3(t-u) \\
&+ \mathbf{c}_1\bar{s}^2 + \mathbf{c}_2\bar{s}(t+u) + \mathbf{c}_3\bar{s}(t-u) + \mathbf{c}_4(t+u)^2 + \mathbf{c}_5(t+u)(t-u) + \mathbf{c}_6 ut \\
&+ \mathbf{d}_1\bar{s}^3 + \mathbf{d}_2\bar{s}^2(t+u) + \mathbf{d}_3\bar{s}^2(t-u) + \mathbf{d}_4\bar{s}(t+u)^2 + \mathbf{d}_5\bar{s}(t+u)(t-u) \quad \text{(v.3.21)} \\
&+ \mathbf{d}_6\bar{s}ut + \mathbf{d}_7(t+u)^3 + \mathbf{d}_8(t+u)^2(t-u) + \mathbf{d}_9(t+u)ut + \mathbf{d}_{10}(t-u)ut \\
&+ \mathcal{O}(\bar{s}^4, t^4, u^4)\,.
\end{aligned}
$$

All the coefficients $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i, \mathbf{d}_i$ can be computed analytically from ref. [50]. The central equation for this computation is eq. (v.3.4), with a contribution with the incoming amplitude at LO and the outgoing amplitude at NLO, and vice versa:

$$\mathbf{M}_{\mathrm{df},3}^{(u,u),\mathrm{OPE}} = -\mathbf{M}_{2,\mathrm{off}}^{\mathrm{LO}} \frac{\mathbf{T}_G}{\bar{b}^2 + i\epsilon} \mathbf{M}_{2,\mathrm{off}}^{\mathrm{NLO}} + \sum_{\ell'\ell} \mathbf{M}_{2,\mathrm{on}}^{\mathrm{LO},\ell'} \mathbf{T}_G G_{\ell'\ell}^{\infty} \mathbf{M}_{2,\mathrm{on}}^{\mathrm{NLO},\ell} + (\mathrm{LO} \leftrightarrow \mathrm{NLO}).$$

$$(\mathrm{v.3.22})$$

For simplicity, we will subdivide the calculation into multiple parts based on these coefficients and their contributions to different partial waves, up to and including $\ell = 3$.

Terms with $\mathbf{a}_1, \mathbf{b}_1$ and $\mathbf{c}_1$ are completely on-shell and pure s-wave. The only contribution that survives after subtraction comes from the $(t+u)$ part of the LO amplitude; specifically,

$$\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset -\left(\mathbf{a}_1 + \mathbf{b}_1\bar{s} + \mathbf{c}_1\bar{s}^2\right)\mathbf{G}\,\mathbf{k}_2\bar{b}^2 + (\mathrm{in} \leftrightarrow \mathrm{out})\,. \qquad (\mathrm{v.3.23})$$

Terms with $\mathbf{b}_2, \mathbf{c}_2$ and $\mathbf{d}_2$ contain both on and off-shell parts and are pure s-wave. All terms in the LO amplitude contribute, but the cubic term $\mathbf{d}_2$ only survives in combination with $\mathbf{k}_0$. Specifically,

$$\begin{aligned}
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset &- \left(\mathbf{b}_2 + \mathbf{c}_2\bar{s}\right)\bar{b}^2\,\mathbf{G}\,\left(\mathbf{k}_0 + \mathbf{k}_1\bar{s} + \mathbf{k}_3(t-u)\right) \\
&- \left(\mathbf{b}_2 + \mathbf{c}_2\bar{s}\right)\mathbf{G}\,\mathbf{k}_2\bar{b}^2\left(\bar{b}^2 - \bar{s} - \bar{s}'\right) - \mathbf{d}_2\bar{s}^2\bar{b}^2\,\mathbf{G}\,\mathbf{k}_0 \\
&+ (\mathrm{in} \leftrightarrow \mathrm{out})\,.
\end{aligned} \qquad (\mathrm{v.3.24})$$

Terms with $\mathbf{b}_3, \mathbf{c}_3$ and $\mathbf{d}_3$ are pure p-wave, and no off-shellness remains after accounting for barrier factors. Terms with $\mathbf{d}_3$ do not contribute at quadratic order, and those with $\mathbf{b}_3$ and $\mathbf{c}_3$ contribute only in combination with $\mathbf{k}_2$, leading to contributions with s-p wave mixing; specifically,

$$\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset - \left(\mathbf{b}_3 + \mathbf{c}_3\bar{s}\right)(t-u)\,\mathbf{G}\,\mathbf{k}_2\bar{b}^2 + (\mathrm{in} \leftrightarrow \mathrm{out})\,. \qquad (\mathrm{v.3.25})$$

The term with $\mathbf{c}_4$ contains both on- and off-shell parts, and is pure s-wave, with all terms in the LO amplitude contributing. Terms with $\mathbf{d}_4$ and $\mathbf{d}_7$ are pure s-wave and have off-shell parts that contribute only in combination with $\mathbf{k}_0$. Specifically,

$$\begin{aligned}
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset &- \mathbf{c}_4\bar{b}^2(\bar{b}^2 - 2\bar{s})\,\mathbf{G}\,\left(\mathbf{k}_0 + \mathbf{k}_1\bar{s} + \mathbf{k}_3(t-u)\right) \\
&- \mathbf{c}_4\bar{s}\,\mathbf{G}\,\mathbf{k}_2\bar{b}^2\left((\bar{b}^2)^2 - \bar{b}^2\bar{s}' + 2\bar{s}\bar{s}' - 2\bar{s}\bar{b}^2 + \bar{s}^2\right) \\
&- \mathbf{d}_4\bar{s}\bar{b}^2(\bar{b}^2 - 2\bar{s})\,\mathbf{G}\,\mathbf{k}_0 - \mathbf{d}_7\bar{b}^2\left((\bar{b}^2)^2 - 3\bar{b}^2\bar{s} - 3\bar{s}^2\right)\mathbf{G}\,\mathbf{k}_0 \\
&+ (\mathrm{in} \leftrightarrow \mathrm{out})\,.
\end{aligned} \qquad (\mathrm{v.3.26})$$

Terms with $\mathbf{c}_5, \mathbf{d}_5$ and $\mathbf{d}_8$ are pure p-wave and the contain off-shell parts even after accounting for barrier factors. They contribute as

$$\begin{aligned}
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset &- \mathbf{c}_5\bar{b}^2(t-u)\,\mathbf{G}\,\left(\mathbf{k}_0 + \mathbf{k}_1\bar{s} + \mathbf{k}_3(t-u)\right) \\
&- \mathbf{c}_5(t-u)\bar{b}^2\,\mathbf{G}\,\mathbf{k}_2(\bar{b}^2 - \bar{s} - \bar{s}') \\
&- \left(\mathbf{d}_5\bar{s} + \mathbf{d}_8(\bar{b}^2 - 2\bar{s})\right)\bar{b}^2(t-u)\,\mathbf{G}\,\mathbf{k}_0 + (\mathrm{in} \leftrightarrow \mathrm{out})\,.
\end{aligned} \qquad (\mathrm{v.3.27})$$

Terms with $\mathbf{c}_6$, $\mathbf{d}_6$ and $\mathbf{d}_9$ contain s and d waves and contribute

$$
\begin{aligned}
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset &-\mathbf{c}_6 \bar{b}^2 \left(\tfrac{1}{4}\bar{b}^2 - \tfrac{1}{3}\bar{s} - \tfrac{1}{48}\bar{s}\bar{b}^2\right) \mathbf{G}\left(\mathbf{k}_0 + \mathbf{k}_1\bar{s} + \mathbf{k}_3(t-u)\right) \\
&-\mathbf{c}_6[ut]_s\,\mathbf{G}\,\mathbf{k}_2(\bar{b}^2 - \bar{s}) + \mathbf{c}_6[ut]_s^{\mathrm{on}}\,\mathbf{G}\,\mathbf{k}_2(-\bar{s}) \\
&-\mathbf{c}_6[ut]_d\,\mathbf{G}\,\mathbf{k}_2(\bar{b}^2) - \mathbf{d}_9[ut]_d\,\mathbf{G}\,\mathbf{k}_0(\bar{b}^2) - \mathbf{d}_6\bar{s}\bar{b}^2\left(\tfrac{1}{4}\bar{b}^2 - \tfrac{1}{3}\bar{s}\right)\mathbf{G}\,\mathbf{k}_0 \\
&-\mathbf{d}_9\bar{b}^2\left(\tfrac{1}{4}(\bar{b}^2)^2 - \tfrac{7}{12}\bar{s}\bar{b}^2 + \tfrac{1}{2}\bar{s}^2\right)\mathbf{G}\,\mathbf{k}_0 + (\text{in} \leftrightarrow \text{out}),
\end{aligned} \tag{v.3.28}
$$

where $tu = [tu]_s + [tu]_d$,

$$
[tu]_s = \frac{1}{4}(\bar{s} - \bar{b}^2)^2 - \frac{4}{3}q_{2,p}^{*2}k_p^{*2}, \quad [tu]_d = q_{2,p}^{*2}k_p^{*2}\frac{8\pi}{15}\sum_m Y_{2m}^*(\hat{\boldsymbol{a}}_p^*)Y_{2m}(\hat{\boldsymbol{k}}_p^*), \tag{v.3.29}
$$

and $[tu]_s^{\mathrm{on}} = \tfrac{1}{4}(\bar{s})^2 - \tfrac{4}{3}q_{2,p}^{*4}$.

The term with $\mathbf{d}_{10}$ is cubic, so it only contributes in combination with $\mathbf{k}_0$, and since $\mathbf{k}_0$ terms are on-shell, only the off-shellness of the $\mathbf{d}_{10}$ term survives after subtraction. It contains both p- and f-waves ($\ell = 3$), which requires the decomposition

$$
(\boldsymbol{a}_p^* \cdot \boldsymbol{k}_p^*)^3 = q_{2,p}^{*3}k^{*3}\left[\frac{3}{5}\frac{4\pi}{3}Y_{1m}^*(\hat{\boldsymbol{a}}_p^*)Y_{1m}(\hat{\boldsymbol{k}}_p^*) + \frac{2}{5}\frac{4\pi}{7}Y_{3m}^*(\hat{\boldsymbol{a}}_p^*)Y_{3m}(\hat{\boldsymbol{k}}_p^*)\right], \tag{v.3.30}
$$

which is the f-wave counterpart of eq. (v.3.8). Guided by this, we split the coefficient as $(t-u)ut = [(t-u)ut]_p + [(t-u)ut]_f$, of which the latter cancels exactly in the subtraction. The remaining p-wave part is

$$
[(t-u)ut]_p = \frac{1}{4}\left((\bar{s} - \bar{b}^2)^2 - \frac{48}{5}q_{2,p}^{*2}k^{*2}\right)q_{2,p}^*k^*\frac{16\pi}{3}Y_{1m}^*(\hat{\boldsymbol{a}}_p^*)Y_{1m}(\hat{\boldsymbol{k}}_p^*). \tag{v.3.31}
$$

Analogously to eq. (v.3.9), the on-shell version is obtained by setting $\bar{b} = 0$ and multiplying by a factor of k_p^*/q_{2p}^*, which the G in the subtraction changes to $(k^*)^2$. Thus, the subtracted result is

$$
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset \tfrac{1}{4}\mathbf{d}_{10}\left(\bar{b}^2 - 2\bar{s} + \tfrac{6}{5}\bar{s}\right)\bar{b}^2(t-u)\,\mathbf{G}\,\mathbf{k}_0(\text{in} \leftrightarrow \text{out}). \tag{v.3.32}
$$

Lastly, the cubic-order terms for $I = 0$ require only the pure p-wave part of the LO amplitude, namely $\mathbf{k}_3(t-u)$, and we likewise require only the part of the NLO amplitude that is proportional to $(t-u)$, namely

$$
\mathbf{M}_2^{\mathrm{NLO}} \supset \mathbf{d}_3\bar{s}^2(t-u) + \mathbf{d}_5\bar{s}(t+u)(t-u) + \mathbf{d}_8(t+u)^2(t-u) + \mathbf{d}_{10}(t-u)ut, \tag{v.3.33}
$$

of which $\mathbf{d}_3$ vanishes after subtraction, and the coefficient $\mathbf{d}_5$ is found to be zero at NLO. The contributions from $\mathbf{d}_5$ and $\mathbf{d}_8$ are simple to evaluate:

$$
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset -\left(\mathbf{d}_5\bar{s} + \mathbf{d}_8(\bar{b}^2 - 2\bar{s})\right)\bar{b}^2(t-u)\,\mathbf{G}\,\mathbf{k}_3(t'-u') + (\text{in} \leftrightarrow \text{out}), \tag{v.3.34}
$$

while for $\mathbf{d}_{10}$ we obtain, after similar manipulations to those above,

$$
\mathbf{M}_{\mathrm{df},3}^{(u,u)} \supset \tfrac{1}{4}\mathbf{d}_{10}(\bar{b}^2 - 2\bar{s} + \tfrac{6}{5}\bar{s})\bar{b}^2(t-u)\,\mathbf{G}\,\mathbf{k}_3(t'-u') + (\text{in} \leftrightarrow \text{out}). \tag{v.3.35}
$$

The complete OPE contributions, including the cubic-order ones for $I = 0$, are listed in table v.2.

Table v.2: All NLO OPE contributions up to quadratic order in the threshold expansion (excluding the s-channel OPE). The cubic-order contributions to $I = 0$ are also included.

I		
$I = 3$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0$	$25\kappa + 78L - 576\ell_1^r - 432\ell_2^r - 72\ell_3^r + 144\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1$	$\frac{6831\kappa}{20} + 372L - 1332\ell_1^r - 1206\ell_2^r + 252\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2$	$\frac{230\,481\kappa}{280} + 576L - 1080\ell_1^r - 1188\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A$	$-\frac{53\,199\kappa}{560} + 45L + 189\ell_1^r - \frac{459}{2}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B$	$\frac{54\,171\kappa}{140} + 216L - 648\ell_1^r - 324\ell_2^r$
$I = 2$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{T}}$	$\frac{207\kappa}{40} - 2L - 210\ell_1^r - 15\ell_2^r + 42\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{T}}$	$\frac{351\,251\kappa}{3360} + \frac{125L}{2} - \frac{483}{2}\ell_1^r - \frac{267}{4}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{T}}$	$-\frac{47\,109\kappa}{160} - \frac{387L}{2} + \frac{837}{2}\ell_1^r + \frac{1485}{4}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{T}}$	$\frac{138\,043\kappa}{20\,160} + \frac{27L}{4} - \frac{45}{4}\ell_1^r - \frac{117}{8}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{T}}$	$-\frac{2693\kappa}{630} + \frac{11L}{3} - 17\ell_1^r - \frac{5}{2}\ell_2^r$
$I = 1$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SS}}$	$-\frac{1475\kappa}{6} + 303L - 96\ell_1^r - 312\ell_2^r - 132\ell_3^r - 216\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SS}}$	$-\frac{12\,773\kappa}{40} + 362L - 522\ell_1^r - 501\ell_2^r - 108\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SS}}$	$\frac{304\,767\kappa}{560} + 516L - 1170\ell_1^r - 963\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A^{\mathrm{SS}}$	$\frac{489\,117\kappa}{1120} + \frac{1365L}{4} - \frac{1917}{2}\ell_1^r - \frac{1089}{2}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B^{\mathrm{SS}}$	$\frac{95\,097\kappa}{280} + 351L - 648\ell_1^r - 729\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SD}}$	$\frac{154\kappa}{5} + 59L - 72\ell_1^r - 33\ell_2^r - 36\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SD}}$	$-\frac{53\,775\kappa}{896} - \frac{99L}{4} - \frac{171}{8}\ell_1^r + \frac{1359}{16}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SD}}$	$\frac{24\,123\kappa}{224} + \frac{237L}{4} - \frac{837}{4}\ell_1^r - \frac{585}{8}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{SD}}$	$\frac{3729\kappa}{320} + \frac{75L}{4} - \frac{351}{4}\ell_1^r - \frac{99}{8}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{SD}}$	$-\frac{61\,143\kappa}{1120} - \frac{3L}{2} - \frac{351}{4}\ell_1^r + \frac{387}{8}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{DD}}$	$-\frac{857\kappa}{120} - 18L - 126\ell_1^r - 9\ell_2^r + 42\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{DD}}$	$\frac{926\,543\kappa}{10\,080} + \frac{305L}{6} - \frac{309}{2}\ell_1^r - \frac{301}{4}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{DD}}$	$-\frac{134\,797\kappa}{1120} - \frac{93L}{2} - \frac{405}{2}\ell_1^r + \frac{963}{4}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{DD}}$	$\frac{398\,287\kappa}{60\,480} + \frac{149L}{36} + \frac{13}{4}\ell_1^r - \frac{337}{24}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{DD}}$	$\frac{37\,577\kappa}{7560} + \frac{13L}{3} - 17\ell_1^r - \frac{9}{2}\ell_2^r$
$I = 0$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{AS}}$	$\frac{693\kappa}{20} + 54L - 324\ell_1^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{AS}}$	0
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{AS}}$	$-\frac{1215\kappa}{32}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{AS}}$	$-\frac{26\,487\kappa}{1120}$

3.2.2 s-channel OPE contributions

This calculation is done as described in section 3.1.2, but instead of eq. (v.3.15), we use

$$\mathbf{M}_{\mathrm{df},3}^{\text{s-OPE,NLO}} = -\boldsymbol{v}_{\mathrm{NLO}}^{\dagger}(\{p_i\}) \frac{1}{P^2 - M_\pi^2} \boldsymbol{v}_{\mathrm{LO}}(\{k_i\}) + (\text{in} \leftrightarrow \text{out}), \qquad (\text{v.3.36})$$

where, in the $I = 1$ sector of the isospin basis,

$$\boldsymbol{v}_{\mathrm{NLO}}(\{k_i\}) = \frac{M_\pi^2}{F_\pi^2} \begin{pmatrix} c_0^{\mathrm{S}} + c_1^{\mathrm{S}}\Delta + c_2^{\mathrm{S}}\Delta^2 + c_{\mathrm{A}}^{\mathrm{S}}\Delta_{\mathrm{A}} \\ c_1^{\mathrm{D}}\xi_1^{(2)} + c_{21}^{\mathrm{D}}\Delta\xi_1^{(2)} + c_{22}^{\mathrm{D}}\xi_1^{(4,2)} \\ c_1^{\mathrm{D}}\xi_2^{(2)} + c_{21}^{\mathrm{D}}\Delta\xi_2^{(2)} + c_{22}^{\mathrm{D}}\xi_2^{(4,2)} \end{pmatrix}, \qquad (\text{v.3.37})$$

and

$$\begin{aligned} \frac{1}{\sqrt{15}}c_0^{\mathrm{S}} &= -\tfrac{1}{2}(19\kappa - 19L + 16\ell_1 + 16\ell_2 + 4\ell_3 + 12\ell_4), \\ \frac{1}{\sqrt{15}}c_1^{\mathrm{S}} &= -\tfrac{1}{12}(181\kappa - 168L + 288\ell_1 + 144\ell_2 + 72\ell_4), \\ \frac{1}{\sqrt{15}}c_2^{\mathrm{S}} &= \tfrac{201}{16}\kappa + 9L - 27\ell_1 - \tfrac{27}{2}\ell_2, \\ \frac{1}{\sqrt{15}}c_3^{\mathrm{S}} &= \tfrac{405}{16}\kappa + \tfrac{27}{2}L - 27\ell_1 - 27\ell_2, \\ \frac{1}{\sqrt{2}}c_1^{\mathrm{D}} &= \tfrac{125}{4}\kappa + 42L - 72\ell_1 - 36\ell_2 - 18\ell_4, \\ \frac{1}{\sqrt{2}}c_{21}^{\mathrm{D}} &= -\tfrac{9}{2}\kappa + 27L - 81\ell_2, \\ \frac{1}{\sqrt{2}}c_{22}^{\mathrm{D}} &= \frac{513}{8}\kappa - 162\ell_1 + 81\ell_2. \end{aligned} \qquad (\text{v.3.38})$$

Combining these results and expanding the propagator, the contributions up to quadratic order in the threshold expansion are summarized in table v.3.

3.2.3 non-OPE contributions

At NLO, the non-OPE contribution encompasses the large number of diagrams not covered by the OPE or s-channel OPE parts, including the "bull's head" triangle diagrams shown in fig. v.3. These contributions are all regular in the real part, so they can be added directly to $\mathbf{K}_{\mathrm{df},3}$ without considering any subtractions. The highly nontrivial threshold expansion of the loop integral functions is described in section 4.2 of ref. [47]; it is done in the same way here, although a larger number of cases must be considered. The complete contributions are summarized in table v.4.

3.2.4 Bull's head subtraction

The bull's head subtraction term, shown schematically in fig. v.1 and corresponding to the topology fig. v.3a, is given by

$$\mathbf{D}^{(u,u)\mathrm{BH}} = \int_r \mathbf{M}_{2,\mathrm{on}}^{\mathrm{LO}}(\boldsymbol{p}_3)\mathbf{T}_G G^\infty(\boldsymbol{p}_3, \boldsymbol{r})\mathbf{M}_{2,\mathrm{on}}^{\mathrm{LO}}(\boldsymbol{r})\mathbf{T}_G G^\infty(\boldsymbol{r}, \boldsymbol{k}_3)\mathbf{M}_{2,\mathrm{on}}^{\mathrm{LO}}(\boldsymbol{k}_3), \qquad (\text{v.3.39})$$

Table v.3: NLO s-OPE contributions up to quadratic order in the threshold expansion for the $I = 1$ channel. They are not present in the other channels.

$I=1$		
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SS}}$	$-\dfrac{855\kappa}{8} + \dfrac{855L}{8} - 90\ell_1^{\mathrm r} - 90\ell_2^{\mathrm r} - \dfrac{45}{2}\ell_3^{\mathrm r} - \dfrac{135}{2}\ell_4^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SS}}$	$-\dfrac{10005\kappa}{64} + \dfrac{9225L}{64} - \dfrac{1035}{4}\ell_1^{\mathrm r} - \dfrac{495}{4}\ell_2^{\mathrm r} + \dfrac{45}{16}\ell_3^{\mathrm r} - \dfrac{945}{16}\ell_3^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SS}}$	$\dfrac{75525\kappa}{512} + \dfrac{49455L}{512} - \dfrac{9045}{32}\ell_1^{\mathrm r} - \dfrac{4725}{32}\ell_1^{\mathrm r} - \dfrac{405}{128}\ell_3^{\mathrm r} - \dfrac{135}{128}\ell_4^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_A^{\mathrm{SS}}$	$\dfrac{18225\kappa}{64} + \dfrac{1215L}{8} - \dfrac{1215}{4}\ell_1^{\mathrm r} - \dfrac{1215}{4}\ell_1^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_B^{\mathrm{SS}}$	0	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SD}}$	$\dfrac{33\kappa}{32} + \dfrac{423L}{16} - 36\ell_1^{\mathrm r} - \dfrac{45}{2}\ell_2^{\mathrm r} - \dfrac{9}{4}\ell_3^{\mathrm r} - \dfrac{27}{2}\ell_4^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SD}}$	$-\dfrac{2073\kappa}{256} + \dfrac{1521L}{128} - \dfrac{27}{2}\ell_1^{\mathrm r} - \dfrac{513}{16}\ell_2^{\mathrm r} + \dfrac{81}{32}\ell_3^{\mathrm r} + \dfrac{27}{16}\ell_4^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SD}}$	$\dfrac{1539\kappa}{64} \qquad\quad - \dfrac{243}{4}\ell_1^{\mathrm r} + \dfrac{243}{8}\ell_2^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{SD}}$	0	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{SD}}$	0	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{DD}}$	0	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{DD}}$	0	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{DD}}$	$\dfrac{1125\kappa}{8} + 189L - 324\ell_1^{\mathrm r} - 162\ell_2^{\mathrm r} \qquad\quad - 81\ell_4^{\mathrm r}$	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{DD}}$	0	
$\dfrac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{DD}}$	0	

where $\int_r \equiv \int \mathrm{d}^3 r/[2\omega_r(2\pi)^3]$ is the Lorentz-invariant integral over the on-shell loop momentum r, and partial-wave indices are implicitly summed over.

Unlike the OPE contributions, the bull's head contribution lacks singularities in the real part, so there is no need to cancel $\mathbf{D}^{\mathrm{BH}}$ against a matching off-shell expression. Only leading-order two-particle amplitudes are present, so following the on-shell version of eq. (v.3.7), there is only the s-wave component $\mathbf{M}_{2,\mathrm{on}}^{\mathrm{LO},s}(\boldsymbol{q}) = \mathbf{k}_0 + (\mathbf{k}_1 - \mathbf{k}_2)(s_q - 4M_\pi)$ where

$$s_{k_3} = 2(M_\pi^2 + k_1 \cdot k_2)\,, \qquad s_{p_3} = 2(M_\pi^2 + p_1 \cdot p_2)\,, \qquad s_r = (P - r)^2\,, \qquad \text{(v.3.40)}$$

and the p-wave component $\mathbf{M}_{2,\mathrm{on}}^{\mathrm{LO},p}(\boldsymbol{q}) = \mathbf{k}_3(t - u)_q$ where [compare eq. (v.3.8)]

$$(t - u)_q = 4\boldsymbol{a}_q^* \cdot \boldsymbol{a}_q'^* = 4\big(q_{2q}^*\big)^2\left[\frac{4\pi}{3}\sum_m Y_{1m}^*(\hat{\boldsymbol{a}}_k^*)Y_{1m}(\hat{\boldsymbol{p}}_k^*)\right]. \qquad \text{(v.3.41)}$$

The $(q_{2q}^*)^2$ cancels the denominators of the barrier terms in G^∞, so the sum rule for spherical harmonics leaves expressions of the kind $\boldsymbol{p}_q^* \cdot \boldsymbol{k}_q^* = \frac{(p\cdot q)(k\cdot q)}{M_\pi^2} - p \cdot k$. Thus, the numerator of the integral $\mathbf{D}^{(u,u)\mathrm{BH}}$ is expressed entirely in terms of Lorentz products between r and the external momenta. The procedure for evaluating such an integral is described in detail in section 4.3 of ref. [47], and requires no modification here, although the number of terms
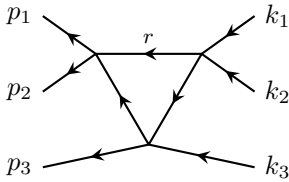
(a) The "bull's head" diagram.
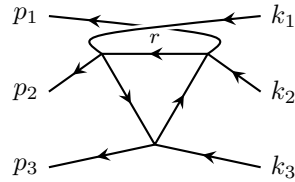
(b) The "crossed bull's head" diagram.

Figure v.3: Two configurations of the triangle-loop diagram. There are a total of 15 diagrams with the triangle topology, of which 9 correspond to the configuration (a) [so their sum corresponds to the symmetrization of (a)] and 6 to the configuration (b). Neither diagram is singular in the real part, and only (a) is singular in the imaginary part, which cancels against $\mathrm{Im}\,\mathbf{D}^{\mathrm{BH}}$.

in each component of $\mathbf{D}^{(u,u)\mathrm{BH}}$ or $\mathbf{D}^{\mathrm{BH}}$ is much larger than in the maximum-isospin case.[6] The integration is easier to perform in the symmetric basis, where the matrices are more sparse, but before the expressions are symmetrized, which limits the number of terms.

Following ref. [47], the result is expressed in terms of

$$
H_{m,n} \equiv \begin{cases} \dfrac{1}{\pi^2} \displaystyle\int_0^{1/\sqrt{3}} \mathrm{d}z\, \dfrac{\sqrt{1+z^2}}{z^m} \dfrac{\mathrm{d}^n}{\mathrm{d}x^n}\big[H^2(x)\big], & n > 0\,, \\[2ex] \displaystyle\int_0^{1/\sqrt{3}} \mathrm{d}z\, 6z f_m(z) \dfrac{\mathrm{d}}{\mathrm{d}x}\big[H^2(x)\big], & n = 0\,, \end{cases} \tag{v.3.42}
$$

where $x = 1 - 3z^2$ and

$$
\frac{\mathrm{d}}{\mathrm{d}z} f_m(x) = \frac{1}{\pi^2} \frac{\sqrt{1+z^2}}{z^m}\,. \tag{v.3.43}
$$

This is separated into

$$
H_{m,n} = \widetilde{H}_{m,n} + f_m(1/\sqrt{3})\delta_{n,0}\,, \tag{v.3.44}
$$

where only $\widetilde{H}_{m,n}$ is cutoff-dependent; setting $\widetilde{H}_{m,n} = 0$ corresponds to setting $H(x) = 1$ everywhere. $f_m(1/\sqrt{3})$ can be expressed entirely in terms of rational numbers and $\log 3$. Thus, each coefficient in the threshold expansion can be separated like

$$
\mathcal{K}_X = \mathcal{K}_X^{[\widetilde{H}_{m,n}=0]} + \mathcal{D}_X\,, \tag{v.3.45}
$$

where the cutoff-independent $\mathcal{K}_X^{[\widetilde{H}_{m,n}=0]}$ can be expressed similarly to the OPE and other contributions. The cutoff-dependent term $\mathcal{D}_X$ must be evaluated numerically, but is well-behaved and often relatively small. The contributions from the bull's head subtraction are listed in table v.5, and the cutoff dependence of $\mathcal{D}_X$ is studied in appendix A.

---

[6]Specifically, the method described in section 4.3.1-3 of ref. [47] works as-is also for general isospin. The complementary methods of section 4.3.4 and appendix E would require some modifications to account for the more complicated numerators. We do not use those methods here, sacrificing cross-check capability for simplicity, since we are by now confident in the correctness of the methods we use.

Table v.4: All NLO non-OPE contributions up to quadratic order in the threshold expansion, including cubic order for $I = 0$. Note that the non-OPE contributions do not include the bull's head subtraction.

I		Expression
$I = 3$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0$	$14\kappa + 33L + 288\ell_1^r + 36\ell_3^r - 72\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1$	$-\frac{35\kappa}{2} + 12L + 720\ell_1^r + 36\ell_2^r - 144\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2$	$-\frac{9747\kappa}{50} - 216L + 648\ell_1^r + 324\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A$	$\frac{576\kappa}{5} - 54L - 162\ell_1^r + 243\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B$	$-\frac{13\,797\kappa}{50} - 162L + 486\ell_1^r + 243\ell_2^r$
$I = 2$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^T$	$\frac{85\kappa}{12} + 2L + 120\ell_1^r + 6\ell_2^r - 24\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^T$	$-\frac{988\kappa}{25} - 36L + 144\ell_1^r + 36\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^T$	$\frac{2052\kappa}{25} + 108L - 324\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^T$	$\frac{501\kappa}{50}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^T$	$\frac{451\kappa}{150} - 2L + 12\ell_1^r$
$I = 1$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{SS}$	$\frac{1522\kappa}{3} - \frac{1129L}{2} + 528\ell_1^r + 840\ell_2^r + 126\ell_3^r + 228\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{SS}$	$545\kappa - 888L + 1440\ell_1^r + 1656\ell_2^r + 96\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{SS}$	$-\frac{30\,441\kappa}{25} - 846L + 1728\ell_1^r + 1674\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A^{SS}$	$-\frac{22\,461\kappa}{20} - 459L + 1188\ell_1^r + 783\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B^{SS}$	$-\frac{63\,039\kappa}{100} - 387L + 756\ell_1^r + 783\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{SD}$	$-\frac{23\kappa}{2} - 84L + 144\ell_1^r + 36\ell_2^r + 48\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{SD}$	$\frac{597\kappa}{10} + 108\ell_1^r - 54\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{SD}$	$-\frac{1041\kappa}{10} - 54L + 270\ell_1^r + 27\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{SD}$	$\frac{231\kappa}{20} + 108\ell_1^r - 54\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{SD}$	$-\frac{4179\kappa}{100} - 18L + 162\ell_1^r - 27\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{DD}$	$-\frac{239\kappa}{4} + 46L + 72\ell_1^r - 54\ell_2^r - 40\ell_4^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{DD}$	$-\frac{21\,158\kappa}{225} - \frac{20L}{3} + 80\ell_1^r - 20\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{DD}$	$-\frac{7607\kappa}{25} - 204L + 1152\ell_1^r + 36\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{DD}$	$-\frac{7897\kappa}{1350} + \frac{64L}{9} - \frac{64}{3}\ell_1^r + \frac{32}{3}\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{DD}$	$-\frac{6409\kappa}{1350} - \frac{14L}{9} + \frac{44}{3}\ell_1^r - \frac{8}{3}\ell_2^r$
$I = 0$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{AS}$	$\frac{1017\kappa}{5} - 54L + 162\ell_1^r + 81\ell_2^r$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{AS}$	$-\frac{972\kappa}{5}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{AS}$	$-\frac{14\,499\kappa}{70}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{AS}$	$\frac{88\,371\kappa}{2240}$

Table v.5: Contributions from the bull's head subtraction up to quadratic order in the threshold expansion, including cubic order for $I = 0$. The contributions are separated into a cutoff-independent analytic part (containing $L_3 \equiv \log 3$) and a cutoff-dependent numerical part according to eq. (v.3.45). The latter is computed using the standard cutoff choice, shown in eq. (v.1.1).

I		Analytic part	Numerical part
$I=3$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0$	$-36\kappa(\quad 4 + \quad\quad L_3)$	$+\ 0.056\,347\,6589$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1$	$-8\kappa(\quad 53 + \quad 12\,L_3)$	$-\ 0.129\,589\,681$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2$	$-\frac{27\kappa}{50}(\quad 363 + \quad 115\,L_3)$	$-\ 0.432\,202\,370$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A$	$\frac{9\kappa}{40}(\quad 1468 - \quad 75\,L_3)$	$-\ 0.000\,907\,273\,890$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B$	$\frac{9\kappa}{200}(\quad 404 - \quad 105\,L_3)$	$-\ 0.000\,162\,394\,747$
$I=2$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{T}}$	$\frac{\kappa}{108}(\quad 764 + \quad 1593\,L_3)$	$-\ 0.258\,397\,947$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{T}}$	$-\frac{\kappa}{7200}(\quad 658\,852 - 627\,165\,L_3)$	$-\ 0.211\,171\,132$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{T}}$	$-\frac{\kappa}{4800}(\quad 7\,561\,228 + 1\,085\,625\,L_3)$	$+\ 5.782\,919\,28$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{T}}$	$-\frac{\kappa}{10\,800}(\quad 451\,052 + \quad 13\,185\,L_3)$	$+\ 0.044\,568\,092\,9$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{T}}$	$-\frac{\kappa}{21\,600}(\quad 75\,148 + \quad 19\,935\,L_3)$	$+\ 0.002\,735\,915\,74$
$I=1$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SS}}$	$-\frac{3\kappa}{4}(\quad 1252 + \quad 423\,L_3)$	$+\ 1.717\,963\,97$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SS}}$	$-\frac{\kappa}{8}(\quad 7972 + \quad 9093\,L_3)$	$+\ 2.571\,033\,92$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SS}}$	$\frac{173\kappa}{1600}(\quad 27\,296 - \quad 13\,545\,L_3)$	$-\ 5.430\,530\,16$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A^{\mathrm{SS}}$	$\frac{63\kappa}{640}(\quad 3764 + \quad 5775\,L_3)$	$-\ 7.227\,961\,67$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B^{\mathrm{SS}}$	$\frac{27\kappa}{3200}(\quad 106\,228 - \quad 33\,985\,L_3)$	$+\ 5.445\,771\,06$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SD}}$	$\frac{\kappa}{16}(\quad 236 + \quad 225\,L_3)$	$-\ 0.270\,266\,343$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SD}}$	$\frac{\kappa}{16}(\quad 1004 - \quad 153\,L_3)$	$+\ 0.751\,221\,968$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SD}}$	$\frac{54\kappa}{5}(\quad 7 + \quad 20\,L_3)$	$+\ 1.896\,052\,35$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{SD}}$	$\frac{\kappa}{64}(\quad 6829 + \quad 297\,L_3)$	$+\ 0.023\,118\,1459$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{SD}}$	$\frac{\kappa}{1600}(\quad 356\,908 - \quad 108\,135\,L_3)$	$+\ 1.032\,569\,94$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{DD}}$	$\frac{\kappa}{108}(\quad 5564 + \quad 1593\,L_3)$	$-\ 0.114\,901\,505$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{DD}}$	$-\frac{\kappa}{2400}(\quad 6316 - \quad 5895\,L_3)$	$+\ 0.301\,860\,448$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{DD}}$	$-\frac{\kappa}{4800}(16\,008\,412 + 2\,096\,685\,L_3)$	$+\ 3.900\,275\,64$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{DD}}$	$-\frac{\kappa}{7200}(\quad 229\,132 - \quad 19\,215\,L_3)$	$-\ 0.129\,770\,638$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{DD}}$	$\frac{\kappa}{7200}(\quad 9436 - \quad 19\,395\,L_3)$	$+\ 0.054\,453\,6019$
$I=0$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{AS}}$	$-\frac{3\kappa}{2}(\quad 68 + \quad 27\,L_3)$	$+\ 0.301\,063\,917$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{AS}}$	$\frac{3\kappa}{320}(\quad 786\,604 + \quad 138\,105\,L_3)$	$-\ 30.496\,932\,4$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{AS}}$	$\frac{9\kappa}{640}(\quad 766\,996 + \quad 153\,495\,L_3)$	$-\ 46.550\,976\,1$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{AS}}$	$-\frac{81\kappa}{1280}(\quad 604 + \quad 2205\,L_3)$	$+\ 1.367\,998\,92$

4 Results

Our full results are stated in table v.6, supplemented by tables v.1 and v.5 for the LO contributions and cutoff-dependent remainders, respectively. A numerical comparison of the different contributions to $\mathcal{K}_{\mathrm{df},3}$ is given in table v.7. The results are plotted as functions of M_π in fig. v.4.

5 Conclusions and outlook

This work presents the NLO ChPT result for the three-particle K-matrix, $\mathbf{K}_{\mathrm{df},3}$, in all isospin channels, thereby extending the maximum-isospin results of ref. [47]. The main results are stated in tables v.1, v.6 and v.7 and fig. v.4. They are qualitatively reminiscent of those at maximum isospin, although several coefficients are numerically rather large. Similarly, NLO contributions are not necessarily small compared to LO contributions, especially at larger-than-physical pion masses.

With the list of K-matrix components now complete for purely pionic three-particle systems up to the given orders, we eagerly await the availability of lattice results against which our values can be compared; currently, no results exist beyond those already covered in ref. [47]. We also note that, although the technical complexity of the calculation has increased compared to that at maximum isospin, there have been no new hurdles that could not be dealt with effectively. This provides an encouraging outlook toward further developments within this ongoing program, for which next steps involve the inclusion of particles other than pions.

Table v.6: The full NLO results for $\mathcal{K}_{\mathrm{df},3}$ up to quadratic order in the threshold expansion (cubic for $I = 0$), combining the OPE, s-OPE, non-OPE and bull's head contributions, i.e., tables v.2 to v.5. The corresponding LO results are given in table v.1. For compactness, we use $L_3 \equiv \log 3$ and omit the cutoff-dependent remainders $\mathcal{D}_X$; they are given in table v.5 and further studied in appendix A.

		constant	L	$\ell_1^{\mathrm r}$	$\ell_2^{\mathrm r}$	$\ell_3^{\mathrm r}$	$\ell_4^{\mathrm r}$
$I=3$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0$	$-3\kappa(35+12L_3)$	$+\ 111L$	$-\ 288\ell_1^{\mathrm r}$	$-\ 432\ell_2^{\mathrm r}$	$-\ 36\ell_3^{\mathrm r}$	$+\ 72\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1$	$-\kappa\frac{1999+1920L_3}{20}$	$+\ 384L$	$-\ 612\ell_1^{\mathrm r}$	$-\ 1170\ell_2^{\mathrm r}$		$+\ 108\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2$	$207\kappa\frac{2923-420L_3}{1400}$	$+\ 360L$	$-\ 432\ell_1^{\mathrm r}$	$-\ 864\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A$	$9\kappa\frac{21\,809-1050L_3}{560}$	$-\ 9L$	$+\ 27\ell_1^{\mathrm r}$	$+\ \frac{27}{2}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B$	$27\kappa\frac{6698-245L_3}{1400}$	$+\ 54L$	$-\ 162\ell_1^{\mathrm r}$	$-\ 81\ell_2^{\mathrm r}$		
$I=2$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm T}$	$\kappa\frac{20\,879+15\,930L_3}{1080}$		$-\ 90\ell_1^{\mathrm r}$	$-\ 9\ell_2^{\mathrm r}$		$+\ 18\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm T}$	$-\kappa\frac{1\,335\,007-4\,390\,155L_3}{50\,400}$	$+\ \frac{53L}{2}$	$-\ \frac{195}{2}\ell_1^{\mathrm r}$	$-\ \frac{123}{4}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm T}$	$-\kappa\frac{8\,580\,514+1\,085\,625L_3}{4800}$	$-\ \frac{171L}{2}$	$+\ \frac{837}{2}\ell_1^{\mathrm r}$	$+\ \frac{189}{4}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm T}$	$-\kappa\frac{7\,528\,763+369\,180L_3}{302\,400}$	$+\ \frac{27L}{4}$	$-\ \frac{45}{4}\ell_1^{\mathrm r}$	$-\ \frac{117}{8}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm T}$	$-\kappa\frac{717\,748+139\,545L_3}{151\,200}$	$+\ \frac{5L}{3}$	$-\ 5\ell_1^{\mathrm r}$	$-\ \frac{5}{2}\ell_2^{\mathrm r}$		
$I=1$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SS}}$	$-\kappa\frac{6275+2538L_3}{8}$	$-\ \frac{1237L}{8}$	$+\ 342\ell_1^{\mathrm r}$	$+\ 438\ell_2^{\mathrm r}$	$-\ \frac{57}{2}\ell_3^{\mathrm r}$	$-\ \frac{111}{2}\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SS}}$	$-\kappa\frac{296\,689+363\,720L_3}{320}$	$-\ \frac{24\,439L}{64}$	$+\ \frac{2637}{4}\ell_1^{\mathrm r}$	$+\ \frac{4125}{4}\ell_2^{\mathrm r}$	$-\ \frac{225}{4}\ell_3^{\mathrm r}$	$-\ 12\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SS}}$	$\kappa\frac{217\,322\,699-131\,223\,960L_3}{89\,600}$	$-\ \frac{119\,505L}{512}$	$+\ \frac{2043}{16}\ell_1^{\mathrm r}$	$+\ 711\ell_2^{\mathrm r}$	$-\ \frac{405}{128}\ell_3^{\mathrm r}$	$-\ \frac{135}{128}\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_A^{\mathrm{SS}}$	$-9\kappa\frac{15\,458-282\,975L_3}{4480}$	$+\ \frac{273L}{8}$	$-\ 378\ell_1^{\mathrm r}$	$+\ \frac{477}{2}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_B^{\mathrm{SS}}$	$9\kappa\frac{1\,507\,124-713\,685L_3}{22\,400}$	$-\ 36L$	$+\ 108\ell_1^{\mathrm r}$	$+\ 54\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{SD}}$	$3\kappa\frac{1871+750L_3}{160}$	$+\ \frac{23L}{16}$	$+\ 36\ell_1^{\mathrm r}$	$-\ \frac{39}{2}\ell_2^{\mathrm r}$	$-\ \frac{9}{4}\ell_3^{\mathrm r}$	$-\ \frac{3}{2}\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{SD}}$	$\kappa\frac{486\,847-85\,680L_3}{8960}$	$-\ \frac{1647L}{128}$	$+\ \frac{585}{8}\ell_1^{\mathrm r}$	$-\ \frac{9}{8}\ell_2^{\mathrm r}$	$+\ \frac{81}{32}\ell_3^{\mathrm r}$	$+\ \frac{27}{16}\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{SD}}$	$27\kappa\frac{1713+3584L_3}{448}$	$+\ \frac{21L}{4}$		$-\ \frac{63}{4}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{SD}}$	$\kappa\frac{8314+297L_3}{64}$	$+\ \frac{75L}{4}$	$+\ \frac{81}{4}\ell_1^{\mathrm r}$	$-\ \frac{531}{8}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{SD}}$	$\kappa\frac{1\,418\,878-756\,945L_3}{11\,200}$	$-\ \frac{39L}{2}$	$+\ \frac{297}{4}\ell_1^{\mathrm r}$	$+\ \frac{171}{8}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{DD}}$	$-\kappa\frac{16\,603-15\,930L_3}{1080}$	$+\ 28L$	$-\ 54\ell_1^{\mathrm r}$	$-\ 63\ell_2^{\mathrm r}$		$+\ 2\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{DD}}$	$-\kappa\frac{79\,771-41\,265L_3}{16\,800}$	$+\ \frac{265L}{6}$	$-\ \frac{149}{2}\ell_1^{\mathrm r}$	$-\ \frac{381}{4}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_2^{\mathrm{DD}}$	$-\kappa\frac{121\,601\,602+14\,676\,795L_3}{33\,600}$	$-\ \frac{123L}{2}$	$+\ \frac{1251}{2}\ell_1^{\mathrm r}$	$+\ \frac{459}{4}\ell_2^{\mathrm r}$		$-\ 81\ell_4^{\mathrm r}$
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{DD}}$	$-\kappa\frac{3\,133\,679-269\,010L_3}{100\,800}$	$+\ \frac{45L}{4}$	$-\ \frac{217}{12}\ell_1^{\mathrm r}$	$-\ \frac{593}{24}\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{DD}}$	$\kappa\frac{77\,296-135\,765L_3}{50\,400}$	$+\ \frac{25L}{9}$	$-\ \frac{7}{3}\ell_1^{\mathrm r}$	$-\ \frac{43}{6}\ell_2^{\mathrm r}$		
$I=0$	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_0^{\mathrm{AS}}$	$3\kappa\frac{907-270L_3}{20}$		$-\ 162\ell_1^{\mathrm r}$	$+\ 81\ell_2^{\mathrm r}$		
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_1^{\mathrm{AS}}$	$3\kappa\frac{765\,868+138\,105L_3}{320}$					
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_3^{\mathrm{AS}}$	$9\kappa\frac{5\,246\,968+1\,074\,465L_3}{4480}$					
	$\frac{F_\pi^6}{M_\pi^6}\mathcal{K}_4^{\mathrm{AS}}$	$-81\kappa\frac{496+3087L_3}{1792}$					

Table v.7: Numerical values of the expressions given in tables v.1 to v.6. Note that only the "Total" column depends on the ratio M_π/F_π (here evaluated at the physical point, $M_\pi/F_\pi \approx 1.50$), and only the "BH" (bull's head) column depends on the cutoff [here using the standard choice, eq. (v.1.1)]. Numbers in parentheses indicate errors; entries without errors are exact up to rounding. Identically zero entries are left blank.

|		Total	LO×$\frac{F_\pi^4}{M_\pi^4}$	NLO×$\frac{F_\pi^6}{M_\pi^6}$	NLO×F_π^6/M_π^6			
					OPE	s-OPE	non-OPE	BH
$I=3$	$\mathcal{K}_0$	61.6(3.0)	18	−2.65(26)	0.50(53)		−2.04(28)	−1.11
	$\mathcal{K}_1$	33.4(5.3)	27	−9.04(46)	−1.8(1.0)		−3.75(61)	−3.48
	$\mathcal{K}_2$	−67.4(2.8)		−5.79(24)	−5.11(58)		1.43(37)	−2.11
	$\mathcal{K}_A$	25.77(18)		2.21(2)	−2.76(15)		3.00(14)	1.97
	$\mathcal{K}_B$	1.4(1.1)		0.12(9)	−0.22(37)		0.25(28)	0.08
$I=2$	$\mathcal{K}_0^T$	27.56(88)	4.5	0.38(8)	1.05(18)		−0.56(10)	−0.11
	$\mathcal{K}_1^T$	−0.12(67)		−0.01(6)	−0.16(14)		−0.02(9)	−0.18
	$\mathcal{K}_2^T$	−79.9(3.0)		−6.86(26)	1.33(23)		−2.42(7)	−5.77
	$\mathcal{K}_3^T$	−2.91(7)		−0.25(1)	−0.085(6)		−0.064	−0.23
	$\mathcal{K}_4^T$	−0.63(3)		−0.054(3)	−0.04(1)		0.014(7)	−0.03
$I=1$	$\mathcal{K}_0^{SS}$	−105.1(2.8)	−13.88	−2.90(24)	−9.21(63)	−3.06(20)	15.81(72)	−6.44
	$\mathcal{K}_1^{SS}$	−139.5(4.2)	−17.77	−4.14(36)	−9.10(42)	−3.71(16)	20.32(82)	−11.65
	$\mathcal{K}_2^{SS}$	62.3(1.6)	−0.26	5.47(14)	−4.80(65)	0.44(27)	6.75(94)	3.07
	$\mathcal{K}_A^{SS}$	−29.2(3.1)		−2.51(27)	−1.77(55)	0.70(38)	−0.51(67)	−0.93
	$\mathcal{K}_B^{SS}$	90.04(72)		7.73(6)	−4.19(35)		2.80(41)	9.13
	$\mathcal{K}_0^{SD}$	−3.90(30)	−0.38	−0.17(3)	−0.99(11)	−0.52(4)	1.42(16)	−0.08
	$\mathcal{K}_1^{SD}$	14.03(54)	0.42	1.02(5)	0.40(3)	−0.31(1)	−0.16(7)	1.08
	$\mathcal{K}_2^{SD}$	45.54(4)		3.910(4)	0.11(12)	0.45(4)	−0.53(17)	3.88
	$\mathcal{K}_3^{SD}$	3.11(28)		0.27(2)	−0.00(5)		−0.46(7)	0.73
	$\mathcal{K}_4^{SD}$	17.79(51)		1.53(4)	0.13(6)		−0.58(10)	1.97
	$\mathcal{K}_0^{DD}$	−4.51(35)	0.5	−0.61(3)	1.00(14)		−1.92(13)	0.31
	$\mathcal{K}_1^{DD}$	−6.34(47)		−0.54(4)	−0.04(9)		−0.81(5)	0.30
	$\mathcal{K}_2^{DD}$	−376.0(5.2)	−20.25	−23.34(44)	1.51(16)	−2.51(30)	−2.09(72)	−20.26
	$\mathcal{K}_3^{DD}$	−6.11(11)		−0.52(1)	−0.087(5)		−0.12(1)	−0.31
	$\mathcal{K}_4^{DD}$	−0.22(2)		−0.019(2)	−0.002(10)		−0.06(1)	0.04
$I=0$	$\mathcal{K}_0^{AS}$	19.6(1.3)		1.68(11)	0.36(20)		1.95(9)	−0.63
	$\mathcal{K}_1^{AS}$	279.317		23.979			−1.231	25.21
	$\mathcal{K}_3^{AS}$	410.220		35.216	−0.240		−1.312	36.77
	$\mathcal{K}_4^{AS}$	2.974		0.255	−0.150		0.250	0.16
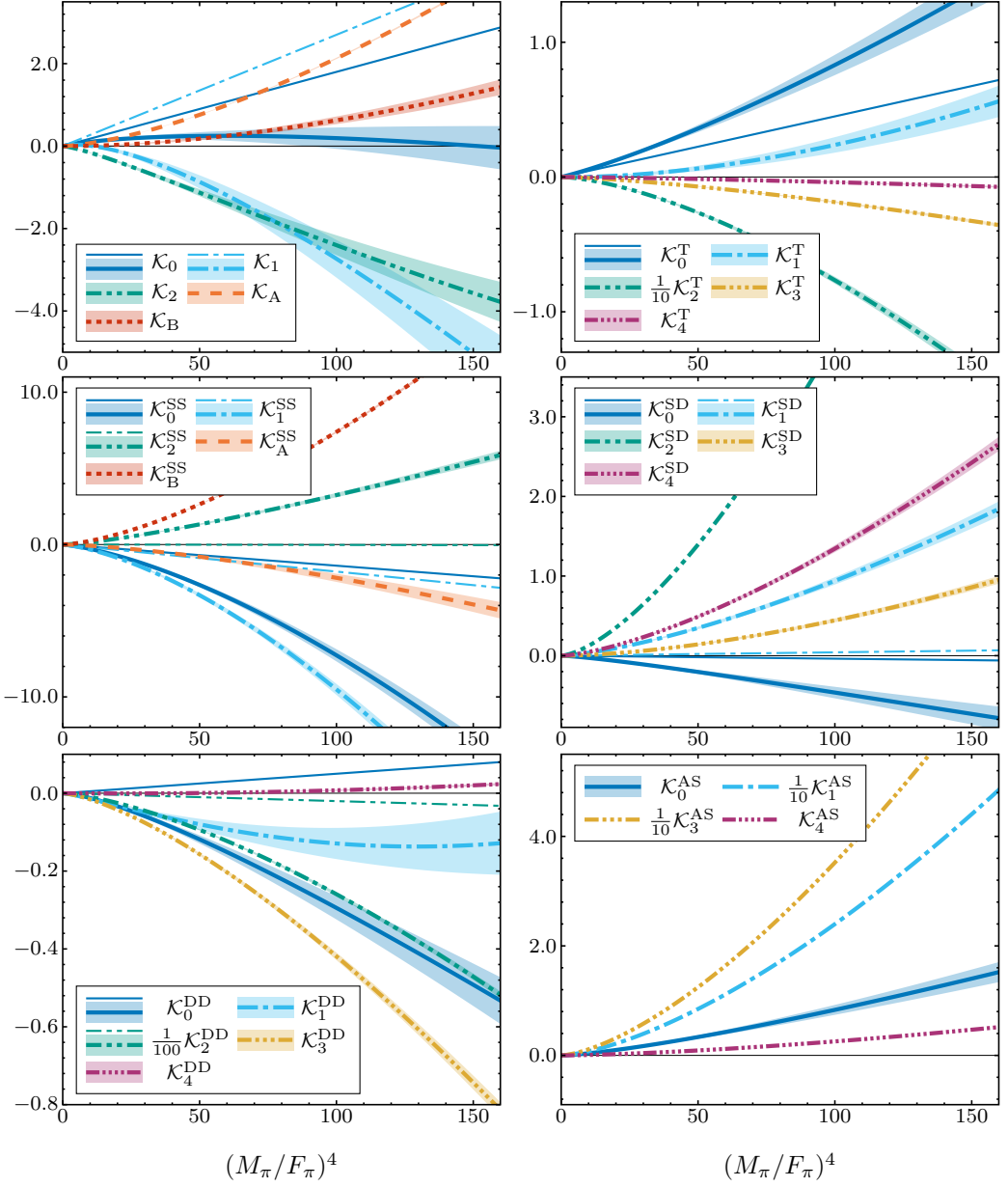
Figure v.4: LO+NLO ChPT predictions for $\mathcal{K}_{\mathrm{df},3}$, plotted as functions of $(M_\pi/F_\pi)^4$; for reference, the physical point is at $(M_{\mathrm{phys}}/F_{\mathrm{phys}})^4 \approx 5.25$. Colored bands represent uncertainties inherited from the LECs (see ref. [47]), and thin lines represent LO-only contributions when present. The coefficients are grouped by isospin, and the lines are drawn so that the number of dots reflects the numeric index on $\mathcal{K}_X$ when applicable. Some coefficients have been rescaled for legibility. Note that the top left panel corresponds to figs. 3 and 4 in ref. [47], although no fits to lattice data are shown.
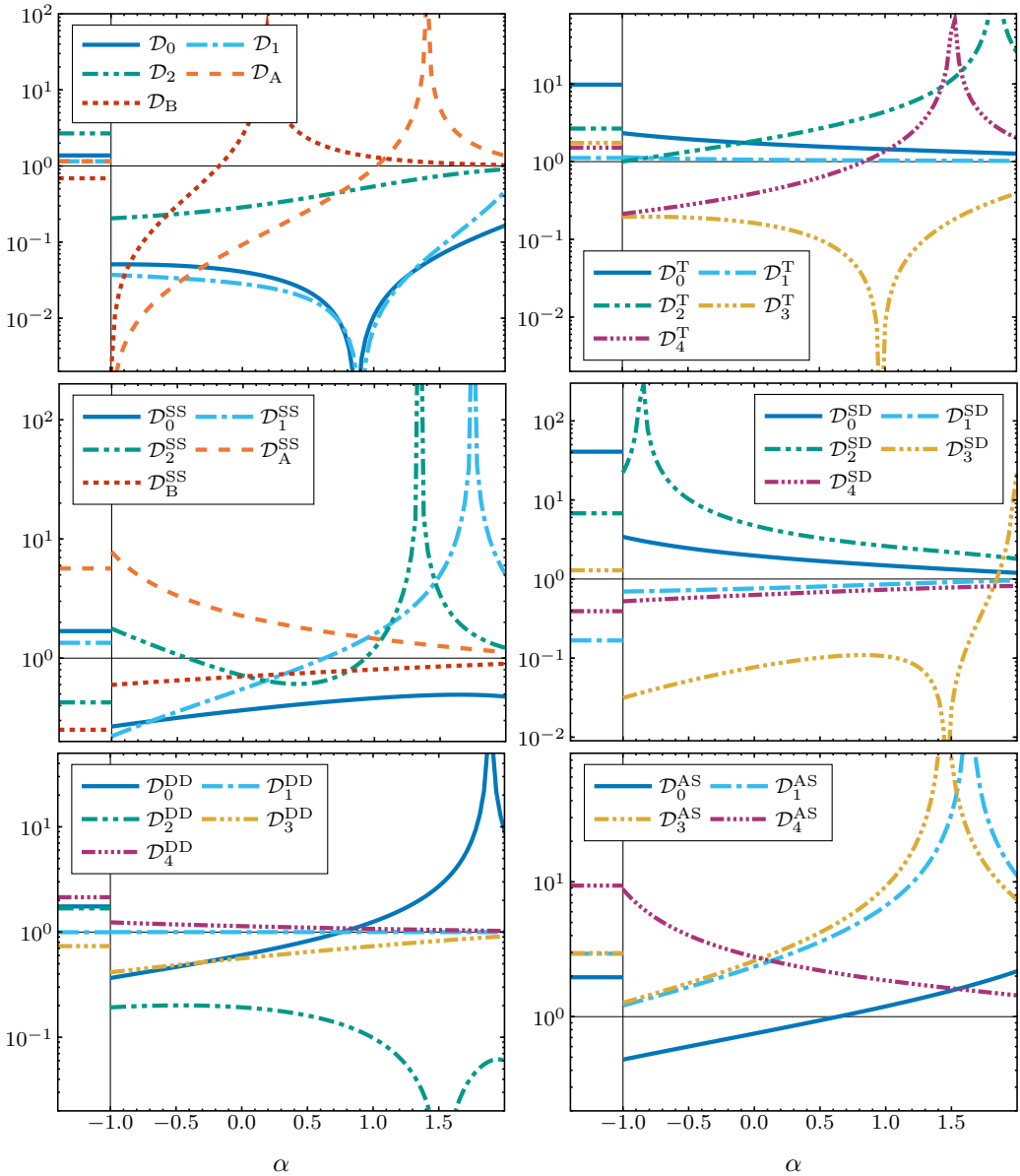
Figure v.5: Illustration of the size of the numerical bull's head remainders $\mathcal{D}_X$. Each line shows $\mathrm{abs}(\mathcal{D}_X/\mathcal{K}_X^{\mathrm{BH}})$, where $\mathcal{K}_X^{\mathrm{BH}}$ is the complete bull's head contribution to $\mathcal{K}_X$, plotted logarithmically as a function of the parameter α defined in eq. (v.1.2). The standard cutoff eq. (v.1.1) is recovered at $\alpha = -1$; the value using eq. (v.1.3) is shown to the left of that. The coefficients are grouped and displayed similarly to fig. v.4. Horizontal lines are drawn at the ratio 1, roughly indicating the border between 'small remainders' and 'large remainders'.

A   Dependence on the cutoff

The cutoff function $H(x)$ is arbitrary as long as it smoothly interpolates between $H(x \leq 0) = 1$ and $H(x \geq 1) = 0$, with the standard choice for $\mathcal{K}_{\mathrm{df},3}$ being

$$H(x) = \exp\left[-\tfrac{1}{x} \exp\left(-\tfrac{1}{1-x}\right)\right], \qquad 0 < x < 1\,, \tag{v.1.1}$$

A generalization corresponds to the replacement [7]

$$x \to 1 + \frac{4}{3-\alpha}(x-1)\,, \qquad -1 \leq \alpha < 3\,, \tag{v.1.2}$$

with $\alpha = -1$ recovering eq. (v.1.1). In ref. [47], we also used the symmetric function

$$H(x) = \left[1 + \exp\left(\tfrac{1}{x} - \tfrac{1}{1-x}\right)\right]^{-1}, \qquad 0 < x < 1\,. \tag{v.1.3}$$

The numerical bull's head remainders $\mathcal{D}_X$ defined in eq. (v.3.45) are the only cutoff-dependent terms in the threshold expansion of $\mathbf{K}_{\mathrm{df},3}$. Figure v.5 shows their dependence on the cutoff choice; the upper left panel displays similar information to fig. 11 in ref. [47].

Note that, unlike the monotonous α-dependence seen at maximum isospin, some coefficients such as $\mathcal{D}_2^{\mathrm{SS}}$ and $\mathcal{D}_3^{\mathrm{SD}}$ have local minima or maxima; $\mathcal{D}_2^{\mathrm{DD}}$ even has one of each (near $\alpha = 0.6$ and $\alpha = 1.9$). This is likely related to the larger size of the remainders compared to the maximum-isospin case: bull's head integrands involving p-waves typically have larger powers of r in the numerator compared to those that are pure s-wave, making the integral more sensitive to regions far away from $r = 0$. On one hand, this makes the analytic approximation worse, since these are the regions where $H(x_r)$ differs the greatest from 1, and on the other hand, it allows more interesting behavior as varying α cut off larger or smaller parts of these regions, with cancellations causing the slope of $\mathcal{D}_X$ to change sign. Interestingly, though, no coefficients show as strong a discrepancy between eqs. (v.1.1) and (v.1.3) as the maximum-isospin $\mathcal{K}_{\mathrm{A}}$, for which we have no explanation.

B   Group-theoretical enumeration of operators

In this appendix we describe how group-theoretical considerations can be used to determine the number of operators at each order in the threshold expansion.

As described in the main text, operators in this expansion can be written as products of the quantities $t_{ij} = (p_i - k_j)^2$, where p_i and k_j are, respectively, the final and initial momenta. We are interested here in operators that are linear, quadratic and cubic in the t_{ij}. Such products can be decomposed into irreps of the group $S_3' \times S_3$, where S_3' and S_3 act, respectively, on the outgoing and incoming particle momenta. As explained in ref. [20], and recalled in the main text, operators of a given isospin lie in (in general a sum of) particular irreps of $S_3' \times S_3$. By counting the number of different irreps that appear in products of the t_{ij}, we can determine the number of independent operators for each

Table v.8: Character table of S_3.

Class	1	(12)	(231)
Dim	1	3	2
1	1	1	1
-1	1	-1	1
D	2	0	-1

isospin at each order in the threshold expansion. This is a more systematic approach than an explicit enumeration, and, indeed, has led to the discovery of additional operators, as noted in the main text.

In fact, there is an additional symmetry that must be considered, namely the PT symmetry that interchanges initial and final momenta (and which holds exactly in QCD). Thus the operators must be decomposed into irreps of the group $G \equiv (S_3' \times S_3) \rtimes Z_2$, which involves a semidirect product. To see this, we consider the defining representation, which acts on the vectors $\{p_1, p_2, p_3, p_1', p_2', p_3'\}$. The matrices forming the individual subgroups are, in block form,

$$S_3 \to \begin{pmatrix} S_3 & 0 \\ 0 & 1 \end{pmatrix}, \quad S_3' \to \begin{pmatrix} 1 & 0 \\ 0 & S_3 \end{pmatrix}, \quad Z_2 \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{v.2.1}$$

Thus we have

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} S_3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & S_3 \end{pmatrix}, \tag{v.2.2}$$

showing that the Z_2 acts nontrivially. Our tasks are thus to determine the character table of G, and then to decompose operators of a given order in the t_{ij} into irreps using the standard character decomposition.

We first recall some results for the permutation group. The character table of S_3 is given in Table v.8. Here we label the irreps 1, -1 (the sign irrep), and D, the standard or doublet irrep. The character table of $S_3' \times S_3$ is then given in the standard way for tensor products, leading to 9 classes and 9 irreps. Classes are given simply by combining classes for the individual S_3's, e.g. $\{(12)', (231)\}$, while irreps are products of the individual irreps, e.g. $(-1)' \otimes D$. Characters or product irreps are simply the products of the characters of the individual irreps.

The inclusion of Z_2, which interchanges S_3' and S_3, leads to some of the conjugacy classes of $S_3' \times S_3$ being combined, and introduces additional classes. The combined classes are

$$(12)_S = (12)', \ 1 + 1, \ (12), \tag{v.2.3}$$
$$(231)_S = (231)', \ 1 + 1, \ (231), \tag{v.2.4}$$
$$(12)(231)_S = (12)', \ (231) + (231)', \ (12), \tag{v.2.5}$$

which reduces the 9 classes of $S_3' \times S_3$ down to 6. There are three additional classes, which involve the Z_2 element in combination with other transformations. In the defining irrep,

these are represented by elements of the form

$$z_1 = \left\{ \begin{pmatrix} 0 & S_3 \\ S_3^{-1} & \end{pmatrix} \right\}, \tag{v.2.6}$$

$$z_2 = \left\{ \begin{pmatrix} 0 & \text{even} \\ \text{odd} & \end{pmatrix}, \begin{pmatrix} 0 & \text{odd} \\ \text{even} & \end{pmatrix} \right\}, \tag{v.2.7}$$

$$z_3 = \left\{ \begin{pmatrix} 0 & \text{even}^{-1} \\ \text{even}' & \end{pmatrix}, \begin{pmatrix} 0 & \text{odd}^{-1} \\ \text{odd}' & \end{pmatrix} \right\}. \tag{v.2.8}$$

Here S_3 means any element of the group, "even" and "odd" refer to an arbitrary even and odd element, and the prime indicates a different element. In particular, in the class z_3, "even" and "even'" are arbitrary but different even elements. These three classes have 6, 18, and 12 elements, respectively.

The character table is given in Table v.9. The notation for irreps is as follows: SD is a combination of a singlet from S_3' and doublet from S_3, together with the PT-conjugate; DD is the combination of two doublets; the superscript $\pm$ indicates the sign obtained under the action of Z_2; and the subscript $\pm$ indicates the sign obtained if the combined parity of the $S_3' \times S_3$ permutation is odd. The notation SD and DD mirrors that used in the main text, while the singlet (here called "1") is also denoted SS in the main text.

The mapping from isospin irreps to those of G has been explained in ref. [20], and is recalled in section 2.2. Operators with $I = 3$ lie in the singlet irrep, 1_+^+, those with $I = 2$ lie in the DD_+^+ irrep, those with $I = 1$ lie in the singlet, SD_+^+ and DD_+^+ irreps, and those with $I = 0$ lie in the 1_-^+ irrep.

We now decompose operators composed of the t_{ij}. At linear order, there are 9 such operators, and the character vector is $\{9, 3, 0, 1, 0, 0, 3, 1, 0\}$, which decomposes as

$$(1_+^+) + (SD_+^+) + (DD_+^+). \tag{v.2.9}$$

It follows that, at this order, there is a single contribution to $I = 3$ [that given by the $\mathcal{K}_1$ term in Equation (v.2.9)], a single contribution to $I = 2$ [that given by the $\mathcal{K}_0^T$ term in Equation (v.2.16)], and three contributions to $I = 1$ [given by the $\mathcal{K}_1^{SS}$ term in Equation (v.2.18), the $\mathcal{K}_0^{DD}$ term in Equation (v.2.19), and the $\mathcal{K}_0^{SD}$ term in Equation (v.2.22)]. There are no contributions to $I = 0$ at this order.

Moving now to quadratic order, there are $9 \times 10/2 = 45$ distinct terms of the form $t_{ij}t_{k\ell}$. We find the character vector to be $\{45, 9, 0, 5, 0, 0, 9, 1, 0\}$, which decomposes as

$$3(1_+^+) + (1_+^-) + (1_-^+) + 4(SD_+^+) + 4(DD_+^+) + (SD_-^+) + (DD_+^-). \tag{v.2.10}$$

Thus there are three singlets, leading to the the $\mathcal{K}_2$, $\mathcal{K}_A$, and $\mathcal{K}_B$ ($I = 3$) terms in Equation (v.2.9), and the corresponding three $I = 1$ SS terms in Equation (v.2.18). Similarly, there are four DD terms in $I = 2$ and $I = 1$, given by the $\mathcal{K}_{1,2,3,4}^T$ terms in Equation (v.2.16), and the corresponding terms in Equation (v.2.19). There are four SD terms in $I = 1$, given by the $\mathcal{K}_{1,2}^{SD}$ terms in Equation (v.2.22). Finally, there is a single $I = 0$ contribution, with coefficient $\mathcal{K}_0^{AS}$ in Equation (v.2.23).

Table v.9: Character table of $(S'_3 \times S_3) \rtimes Z_2$.

Class	$1', 1$	$(12)_S$	$(231)_S$	$(12)', (12)$	$(12)(231)_S$	$(231)^{(')}$	z_1	z_2	z_3
Dim	1	6	4	9	12	4	6	18	12
1^+_+	1	1	1	1	1	1	1	1	1
1^-_+	1	1	1	1	1	1	−1	−1	−1
1^+_-	1	−1	1	1	−1	1	1	−1	1
1^-_-	1	−1	1	1	−1	1	−1	1	−1
2^+_+	2	0	2	−2	0	2	0	0	0
SD^+_+	4	2	1	0	−1	−2	0	0	0
SD^+_-	4	−2	1	0	1	−2	0	0	0
DD^+_+	4	0	−2	0	0	1	2	0	−1
DD^-_+	4	0	−2	0	0	1	−2	0	1

Moving lastly to cubic order, there are $9 \times 10 \times 11/6 = 165$ distinct terms that are cubic in the $\tilde{t}_{ij}$. With some effort, one finds that the character vector is $\{165, 19, 3, 5, 1, 3, 19, 1, 1\}$. For example, in the class $(12)(231)_S$, picking the element $(12)'(231)$, only the term $\tilde{t}_{31}\tilde{t}_{32}\tilde{t}_{33}$ is invariant. The most tricky case is the class z_3. Picking the element where even $= 1$ and even$' = (231)$, the single invariant term is $\tilde{t}_{12}\tilde{t}_{31}\tilde{t}_{23}$.

The decomposition of this character vector is

$$7(1^+_+) + 3(1^-_+) + 3(1^+_-) + 4(2^+_+) + 12(\mathrm{SD}^+_+) + 12(\mathrm{DD}^+_+) + 6(\mathrm{SD}^+_-) + 6(\mathrm{DD}^-_+) . \quad \text{(v.2.11)}$$

Given the large numbers of irreps, we focus on the $I = 0$ case, for which we learn that there are 3 independent 1^+_- irreps, and thus three coefficients at this order. These correspond to the coefficients $\mathcal{K}^{\mathrm{AS}}_1$, $\mathcal{K}^{\mathrm{AS}}_3$, and $\mathcal{K}^{\mathrm{AS}}_4$, in Equation (v.2.23). The final coefficient was missed in the enumeration of ref. [20].

References

[1] Beane, S. R., Detmold, W. & Savage, M. J. "n-Boson Energies at Finite Volume and Three-Boson Interactions". *Phys. Rev.* **D76**, 074507. arXiv: 0707.1670 [hep-lat] (2007).

[2] Detmold, W. & Savage, M. J. "The Energy of n Identical Bosons in a Finite Volume at $\mathcal{O}(L^{-7})$". *Phys. Rev.* **D77**, 057502. arXiv: 0801.0763 [hep-lat] (2008).

[3] Briceño, R. A. & Davoudi, Z. "Three-particle scattering amplitudes from a finite volume formalism". *Phys. Rev.* **D87**, 094507. arXiv: 1212.3398 [hep-lat] (2013).

[4] Polejaeva, K. & Rusetsky, A. "Three particles in a finite volume". *Eur. Phys. J. A* **48**, 67. arXiv: 1203.1241 [hep-lat] (2012).

[5] Hansen, M. T. & Sharpe, S. R. "Relativistic, model-independent, three-particle quantization condition". *Phys. Rev. D* **90,** 116003. arXiv: 1408.5933 [hep-lat] (2014).

[6] Hansen, M. T. & Sharpe, S. R. "Expressing the three-particle finite-volume spectrum in terms of the three-to-three scattering amplitude". *Phys. Rev. D* **92,** 114509. arXiv: 1504.04248 [hep-lat] (2015).

[7] Briceño, R. A., Hansen, M. T. & Sharpe, S. R. "Relating the finite-volume spectrum and the two-and-three-particle S matrix for relativistic systems of identical scalar particles". *Phys. Rev.* **D95,** 074510. arXiv: 1701.07465 [hep-lat] (2017).

[8] Koenig, S. & Lee, D. "Volume Dependence of N-Body Bound States". *Phys. Lett.* **B779,** 9–15. arXiv: 1701.00279 [hep-lat] (2018).

[9] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three-particle quantization condition in a finite volume: 1. The role of the three-particle force". *JHEP* **09,** 109. arXiv: 1706.07700 [hep-lat] (2017).

[10] Hammer, H.-W., Pang, J.-Y. & Rusetsky, A. "Three particle quantization condition in a finite volume: 2. general formalism and the analysis of data". *JHEP* **10,** 115. arXiv: 1707.02176 [hep-lat] (2017).

[11] Mai, M. & Döring, M. "Three-body Unitarity in the Finite Volume". *Eur. Phys. J. A* **53,** 240. arXiv: 1709.08222 [hep-lat] (2017).

[12] Briceño, R. A., Hansen, M. T. & Sharpe, S. R. "Numerical study of the relativistic three-body quantization condition in the isotropic approximation". *Phys. Rev.* **D98,** 014506. arXiv: 1803.04169 [hep-lat] (2018).

[13] Briceño, R. A., Hansen, M. T. & Sharpe, S. R. "Three-particle systems with resonant subprocesses in a finite volume". *Phys. Rev.* **D99,** 014516. arXiv: 1810.01429 [hep-lat] (2019).

[14] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "Implementing the three-particle quantization condition including higher partial waves". *JHEP* **03,** 106. arXiv: 1901.07095 [hep-lat] (2019).

[15] Pang, J.-Y., Wu, J.-J., Hammer, H.-W., Meißner, U.-G. & Rusetsky, A. "Energy shift of the three-particle system in a finite volume". *Phys. Rev.* **D99,** 074513. arXiv: 1902.01111 [hep-lat] (2019).

[16] Jackura, A. W. *et al.* "Equivalence of three-particle scattering formalisms". *Phys. Rev. D* **100,** 034508. arXiv: 1905.12007 [hep-ph] (2019).

[17] Briceño, R. A., Hansen, M. T., Sharpe, S. R. & Szczepaniak, A. P. "Unitarity of the infinite-volume three-particle scattering amplitude arising from a finite-volume formalism". *Phys. Rev.* **D100,** 054508. arXiv: 1905.11188 [hep-lat] (2019).

[18] Hörz, B. & Hanlon, A. "Two- and three-pion finite-volume spectra at maximal isospin from lattice QCD". *Phys. Rev. Lett.* **123,** 142002. arXiv: 1905.04277 [hep-lat] (2019).

[19] Romero-López, F., Sharpe, S. R., Blanton, T. D., Briceño, R. A. & Hansen, M. T. "Numerical exploration of three relativistic particles in a finite volume including two-particle resonances and bound states". *JHEP* **10,** 007. arXiv: 1908.02411 [hep-lat] (2019).

[20] Hansen, M. T., Romero-López, F. & Sharpe, S. R. "Generalizing the relativistic quantization condition to include all three-pion isospin channels". *JHEP* **07,** 047. arXiv: 2003.10974 [hep-lat] (2020).

[21] Blanton, T. D. & Sharpe, S. R. "Alternative derivation of the relativistic three-particle quantization condition". *Phys. Rev. D* **102,** 054520. arXiv: 2007.16188 [hep-lat] (2020).

[22] Blanton, T. D. & Sharpe, S. R. "Equivalence of relativistic three-particle quantization conditions". *Phys. Rev. D* **102,** 054515. arXiv: 2007.16190 [hep-lat] (2020).

[23] Pang, J.-Y., Wu, J.-J. & Geng, L.-S. "DDK system in finite volume". *Phys. Rev. D* **102,** 114515. arXiv: 2008.13014 [hep-lat] (2020).

[24] Romero-López, F., Rusetsky, A., Schlage, N. & Urbach, C. "Relativistic N-particle energy shift in finite volume". *JHEP* **02,** 060. arXiv: 2010.11715 [hep-lat] (2021).

[25] Blanton, T. D. & Sharpe, S. R. "Relativistic three-particle quantization condition for nondegenerate scalars". *Phys. Rev. D* **103,** 054503. arXiv: 2011.05520 [hep-lat] (2021).

[26] Müller, F., Rusetsky, A. & Yu, T. "Finite-volume energy shift of the three-pion ground state". *Phys. Rev. D* **103,** 054506. arXiv: 2011.14178 [hep-lat] (2021).

[27] Blanton, T. D. & Sharpe, S. R. "Three-particle finite-volume formalism for $\pi^+\pi^+K^+$ and related systems". *Phys. Rev. D* **104,** 034509. arXiv: 2105.12094 [hep-lat] (2021).

[28] Müller, F., Pang, J.-Y., Rusetsky, A. & Wu, J.-J. "Relativistic-invariant formulation of the three-particle quantization condition". arXiv: 2110.09351 [hep-lat] (2021).

[29] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "Implementing the three-particle quantization condition for $\pi^+\pi^+K^+$ and related systems". *JHEP* **02,** 098. arXiv: 2111.12734 [hep-lat] (2022).

[30] Jackura, A. W. "Three-body scattering and quantization conditions from S matrix unitarity". arXiv: 2208.10587 [hep-lat] (2022).

[31] Garofalo, M., Mai, M., Romero-López, F., Rusetsky, A. & Urbach, C. "Three-body resonances in the φ^4 theory". *JHEP* **02,** 252. arXiv: 2211.05605 [hep-lat] (2023).

[32] Müller, F., Pang, J.-Y., Rusetsky, A. & Wu, J.-J. "Three-particle Lellouch-Lüscher formalism in moving frames". *JHEP* **02,** 214. arXiv: 2211.10126 [hep-lat] (2023).

[33] Draper, Z. T., Hansen, M. T., Romero-López, F. & Sharpe, S. R. "Three relativistic neutrons in a finite volume". *JHEP* **07,** 226. arXiv: 2303.10219 [hep-lat] (2023).

[34] Beane, S. R. *et al.* "Multi-Pion Systems in Lattice QCD and the Three-Pion Interaction". *Phys. Rev. Lett.* **100,** 082004. arXiv: 0710.1827 [hep-lat] (2008).

[35] Detmold, W. & Smigielski, B. "Lattice QCD study of mixed systems of pions and kaons". *Phys. Rev. D* **84,** 014508. arXiv: 1103.4362 [hep-lat] (2011).

[36] Mai, M. & Döring, M. "Finite-Volume Spectrum of $\pi^+\pi^+$ and $\pi^+\pi^+\pi^+$ Systems". *Phys. Rev. Lett.* **122,** 062503. arXiv: 1807.04746 [hep-lat] (2019).

[37] Blanton, T. D., Romero-López, F. & Sharpe, S. R. "$I = 3$ Three-Pion Scattering Amplitude from Lattice QCD". *Phys. Rev. Lett.* **124,** 032001. arXiv: 1909.02973 [hep-lat] (2020).

[38] Mai, M., Döring, M., Culver, C. & Alexandru, A. "Three-body unitarity versus finite-volume $\pi^+\pi^+\pi^+$ spectrum from lattice QCD". *Phys. Rev. D* **101,** 054510. arXiv: 1909.05749 [hep-lat] (2020).

[39] Culver, C., Mai, M., Brett, R., Alexandru, A. & Döring, M. "Three pion spectrum in the $I = 3$ channel from lattice QCD". *Phys. Rev. D* **101,** 114507. arXiv: 1911.09047 [hep-lat] (2020).

[40] Fischer, M. *et al.* "Scattering of two and three physical pions at maximal isospin from lattice QCD". *Eur. Phys. J. C* **81,** 436. arXiv: 2008.03035 [hep-lat] (2021).

[41] Hadron Spectrum, Hansen, M. T., Briceño, R. A., Edwards, R. G., Thomas, C. E. & Wilson, D. J. "Energy-Dependent $\pi^+\pi^+\pi^+$ Scattering Amplitude from QCD". *Phys. Rev. Lett.* **126,** 012001. arXiv: 2009.04931 [hep-lat] (2021).

[42] NPLQCD, QCDSF, Beane, S. R. *et al.* "Charged multihadron systems in lattice QCD+QED". *Phys. Rev. D* **103,** 054504. arXiv: 2003.12130 [hep-lat] (2021).

[43] Alexandru, A. *et al.* "Finite-volume energy spectrum of the $K^-K^-K^-$ system". *Phys. Rev. D* **102,** 114523. arXiv: 2009.12358 [hep-lat] (2020).

[44] Brett, R. *et al.* "Three-body interactions from the finite-volume QCD spectrum". *Phys. Rev. D* **104,** 014501. arXiv: 2101.06144 [hep-lat] (2021).

[45] Blanton, T. D. *et al.* "Interactions of two and three mesons including higher partial waves from lattice QCD". *JHEP* **10,** 023. arXiv: 2106.05590 [hep-lat] (2021).

[46] GWQCD, Mai, M. *et al.* "Three-Body Dynamics of the a1(1260) Resonance from Lattice QCD". *Phys. Rev. Lett.* **127,** 222001. arXiv: 2107.03973 [hep-lat] (2021).

[47] Baeza-Ballesteros, J. *et al.* "The isospin-3 three-particle K-matrix at NLO in ChPT". *JHEP* **05,** 187. arXiv: 2303.13206 [hep-ph] (2023).

[48] Bijnens, J. & Husek, T. "Six-pion amplitude". *Phys. Rev. D* **104,** 054046. arXiv: 2107.06291 [hep-ph] (2021).

[49] Bijnens, J., Husek, T. & Sjö, M. "Six-meson amplitude in QCD-like theories". *Phys. Rev. D* **106,** 054021. arXiv: 2206.14212 [hep-ph] (2022).

[50] Bijnens, J. & Lu, J. "Meson-meson Scattering in QCD-like Theories". *JHEP* **03,** 028. arXiv: 1102.0172 [hep-ph] (2011).

List of abbreviations

Abbreviation	Meaning	Introduced
1PI	One-particle irreducible	p. 250
1PR	One-particle reducible	p. 250
BCFW	Britto–Cachazo–Feng–Witten	p. 66
BH	Bull's head	p. 292
BS	Bethe–Salpeter	p. 78
BSM	Beyond the Standard Model	pp. 3, 188
ChPT	Chiral perturbation theory	pp. 3, 240
CM(F)	Center-of-momentum (frame)	pp. 77, 282
DFO	Diagrammatic flavor-ordering	p. 67
EFT	Effective field theory	pp. 41, 96, 146, 240
eV	Electronvolt (also kilo-, mega-, giga-, etc.)	pp. 6, 7
FVU	Finite-volume unitarity	p. 80
IR	Infrared (zero-energy limit)	pp. 75, 112
LEC	Low-energy constant	pp. 44, 146, 243, 286
LO	Leading order	pp. 28, 147, 240, 280
NLO	Next-to-leading order	pp. 28, 146, 240, 281
NNLO	Next-to-next-to-leading order	pp. 28, 146, 242
N^kLO	*similarly, with 'next-to' repeated k times*	pp. 28, 240
NGB	Nambu–Goldstone boson	p. 37
NLSM	Non-linear sigma model	pp. 42, 96, 147
NRFT	Nonrelativistic field theory	p. 80
OPE	One-particle exchange	p. 281
RFT	Relativistic field theory	pp. 80, 280, 342
SM	(the) Standard Model	p. 2
QCD	Quantum chromodynamics	pp. 2, 146, 240
QED	Quantum electrodynamics	p. 2
QFT	Quantum field theory	p. 2
UV	Ultraviolet (infinite-energy limit)	p. 75
VEV	Vacuum expectation value	p. 37
WZW	Wess–Zumino–Witten	p. 48

LUND
UNIVERSITY

9 789180 398657