WILSON LOOPS

TOM MERTENS

Mathematical Foundations with applications in Quantum Chromodynamics.

Dec. 2014 – version 1.0

Tom Mertens: *Wilson Loops* , Mathematical Foundations with applications in Quantum Chromodynamics., Phd , © Dec. 2014

SUPERVISORS:
Pierre Van Mechelen
Igor Cherednikov

LOCATION:
Antwerp

TIME FRAME:
Dec. 2014

ABSTRACT

It is generally accepted that there are four fundamental forces in nature: the electro-magnetic force, the weak force, the strong force and the gravitational force. In this thesis we will focus on the strong force, which is described by a gauge theory that is known as Quantum Chromo-Dynamics (QCD)[1].

Considering nucleons, of which protons and neutrons are the simplest examples, we know that they are composed of quarks and gluons whose interactions are described by QCD. On the other hand, the fact that quarks and gluons are **confined** in nucleons is still not very well understood. We know that at low-energy the strong force becomes, well... strong (i.e. the coupling constant, which describes how strongly particles interact with each other, becomes large ~ 1) but this alone is not enough to explain confinement. For the moment the question of how and if confinement can be derived from theory is still open, so we will not discuss this in any more detail in this text.

Focusing on low-energy phenomena, the fact that the coupling constant becomes large is conceptually easy to understand but, from a calculation point of view, making predictions in this regime is far from trivial. The main reason for this is due to the fact that when the coupling constant gets large we can no longer use standard perturbative methods to do calculations. In other words, we have entered the non-perturbative regime of the theory (QCD). This means that if we want to "calculate" or predict low-energy or large-distance phenomena, starting from the QCD Lagrangian without inserting anything by hand, we need to have access to non-perturbative calculation methods. There are several ongoing attempts to construct such methods of which we mention: Lattice QCD [1], Borel Resummation [2], AdS/CFT [3] and BRST inspired techniques [4], none of which are completely satisfactory for the moment.

In this text we will discuss another approach, which makes use of a special kind of loop space, where with loop space we refer to the space constructed of all possible loops in the (space-time) manifold under consideration or put in more mathematical terms: the space of all (not necessarily continuous) maps from the circle (S^1) to the manifold. The "special kind" will refer to the fact that we consider a **Generalized** loop space, for which a detailed construction

1 Some of the techniques can be extended to the other forces, but here we will only consider a QCD setting.

will be given. The interesting part about this approach is that it might give access to some restricted parts of the non-perturbative sector and has an extra motivation coming from another question in nuclear physics: How does the proton (nucleon) inherits it's spin from the constituent particles [5]? Naively one could answer this question by stating that a proton is a composition of three quarks, two up and one down, where each of these quarks has spin $1/2$, such that one could conclude that the spins of two quarks cancel and the remaining one gives the proton the known spin $1/2$. Problem solved..., but not quite. From scattering experiments in the second half of the 20th century, and from **renormalization** theory, we know that the **(collinear) distribution functions**[2] of quarks in a proton depend on the energy scale at which we do the experiments and on the "longitudinal fraction of momentum" the quarks are carrying with respect to the proton. As a result we conclude that the naive picture from above is too simple and fails to reproduce the results of the high-energy experiments beyond the collinear approximation. Now, also taking gluon contributions to the spin into account, it is far from trivial how the proton gets to have a fixed spin $1/2$ value. The situation even gets worse when we realize that the quarks and gluons are moving around in the proton (nucleon), such that there is possibly also an angular momentum contribution. So how can we try to solve this problem?

In order to investigate what is happening inside the proton we turn to scattering experiments, which in the last century have given us the insights we have today in particle physics. Dedicated scattering experiments to unravel the proton's structure, performed in the second half of the last century, mainly measured the **longitudinal Parton momentum Distribution Functions** giving us in some sense only **one-dimensional** information. These distributions describe, in some reference frame, the probability to find a parton (quarks, anti-quarks and gluons) inside the proton (or nucleon) with a certain momentum fraction at some energy scale. Using a Fourier transform, this momentum information is then transformed in position information.

This information is clearly inadequate to describe the full internal structure of the proton, for instance if we are interested in the angular momenta of the parton's contribution to the protons spin we obviously will need something more than just the information provided by the longitudinal distribution

2 In certain inertial frames one can associate to these distributions a number of quarks, but the interpretation of these distributions in this way is non-trivial and debatable. We don't go into the details of this interpretation in this text to avoid getting into a philosophical discussion.

functions. Several generalizations of the longitudinal Distribution Functions have been proposed. We mention the Generalized Parton Distributions that give us access to the three dimensional structure of the nucleon, and the Transverse Momentum Distributions describing the probability to find a parton in a nucleon with a certain momentum fraction and a certain transverse momentum with respect to the nucleon direction of motion. Many others exist (see for instance [6]), but in this text we will restrict ourselves to these Transverse Momentum Distributions and in the last chapter we will focus on the half-Fourier transform of such a distribution, a Transverse Distance Dependent (TDD) distribution function.

However, we point out that there is still an ongoing discussion about the correct definition for these transverse distributions and that in this text we will use one specific definition [7].

The problem with these distribution functions, both the longitudinal and the transverse, is that they cannot be calculated from theory because they originate in the non-perturbative sector of the theory, which is problematic as we discussed above. All we can do at this time is to measure them over as large as possible ranges of their parameter space. Having this information we can then investigate if we can find evolution equations with respect to the different parameters hopefully providing us with some more insight about the underlying physics. In other words, starting from a certain value of one of the parameters, one can go over to another value for this parameter by means of the evolution equations. We mention as examples of such evolution equations the famous Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations with respect the mass parameter μ [8–11] and the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation for evolution with respect to the momentum fraction x [12, 13]. Both equations, unfortunately, do not cover the entire ranges possible for their evolution parameters so there is still room for improvement. For the transverse-dependent case we mention the Collins-Soper equations [14, 15], which again have only a restricted application range with respect to their parameter space.

The quest of this Dissertation is to look for a consistent field-theoretically motivated and mathematically correctly formulated set of evolution equations with respect to all the scales (ultra-violet, rapidity, infra-red, if needed) for the Transverse Momentum Distributions (TMDs). As we will discuss, using the standard quantum field-theoretical renormalization techniques for these TMDs do not always work. In the hope of dealing with some of the renormalization issues we turned our attention to the Generalized Loop Space formalism [16, 17]. In this formalism we only considered the subset of Wilson loops, where we then applied a kind of geometrical renormalization to arrive at an evolution equation. These equations should in principle be testable at the Thomas Jefferson National Accelerator Facility (Newport New, VA), planned Electron-Ion Collider, LHC (CERN) and other machines.

PUBLICATIONS

[1] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken. Evolution of cusped light-like Wilson loops and geometry of the loop space. *Phys. Rev.*, D86 :085035, 2012, arXiv:1208.1631.

[2] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken. Loop space and evolution of the light-like Wilson polygons. *Int. J. Mod. Phys. Conf. Ser.*, 20:109–117, 2012, arXiv:1208.5410.

[3] F. F. Van der Veken, I. O. Cherednikov, and T. Mertens. Evolution and Dynamics of Cusped Light-Like Wilson Loops in Loop Space. *AIP Conf. Proc.*, 1523:272–275, 2012, arXiv:1212.4345.

[4] I. O. Cherednikov, T. Mertens, and F. Van der Veken. Cusped light-like Wilson loops in gauge theories. *Phys. Part. Nucl.* , 44:250–259, 2013, arXiv:1210.1767.

[5] I. O. Cherednikov, T. Mertens, P. Taels, and F. F. Van der Veken. Evolution of transverse-distance dependent parton densities at large-x_B and geometry of the loop space. *Int.J.Mod.Phys.Conf.Ser.*, 25:1460006, 2014, arXiv:1308.3116.

[6] T. Mertens and P. Taels. Evolution of Light-Like Wilson Loops with a Self-Intersection in Loop Space. *Phys.Lett.*, B727:563–567, 2013, arXiv:1308.5296.

[7] I. Cherednikov and T. Mertens. Fréchet derivative for light-like Wilson Loops. *submitted to Phys. Lett. B*, 2014, arXiv:1401.2721.

[8] I. O. Cherednikov and T. Mertens. Quadrilateral Wilson loop on the light-cone in the De Donder gauge. *In preparation*, 2014.

[9] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken. *Wilson lines in quantum field theory*. De Gruyter, Berlin, 2014.

[10] T. Mertens. Generalized loop space and tmds. *Proceedings of NTHEP conference, Alushta*, pages 220–228, 2013.

[11] I. Cherednikov and T. Mertens. On Geometric Scaling of Light-Like Wilson Polygons: Higher Orders in α_s. *Phys.Lett.*, B734:198–202, 2014, arXiv:1404.6713.

[12] T. Mertens. Generalized loop space and TMDs. *EPJ Web Conf.*, 73:02011, 2014.

[13] I. Cherednikov and T. Mertens. Generalized Loop Space and Evolution of the Light-Like Wilson Loops. *submitted to Int. J. Mod. Phys. Conf. Ser.*, 2014, arXiv:1410.2425.

[14] T. Mertens and J. Weeks. Maxwell's equations in a multi-connected universe. *submitted to American Mathematical Monthly*, 2014.

Eine neue wissenschaftliche Wahrheit pflegt
sich nicht in der Weise durchzusetzen,
daß ihre Gegner überzeugt werden und
sich als belehrt erklären,
sondern vielmehr dadurch,
daß ihre Gegner allmählich aussterben und
daß die heranwachsende Generation von vornherein
mit der Wahrheit vertraut gemacht ist.

— Max Planck [18]

ACKNOWLEDGMENTS

First of all I would like thank my promotors Prof. Pierre Van Mechelen and Dr. Igor Cherednikov for giving me the possibility to perform the research described in this text. I'm especially grateful that they allowed me to explore several different aspects of theoretical physics through the different schools and conferences I was able to attend during my Phd term. Also a word of gratitude to my fellow PhD students Frederik and Pieter for our, sometimes frustrating, but allows enlightening discussions. I would like to thank Frederik also for sharing some of his latex code and knowledge.

I'm also thankful to the mathematical department of our University, especially Prof. W. Lowen and Prof. B. Shoikhet, for discussions, seminars and suggested ideas.

Furthermore I would like to thank A. Prokudin for discussions during the JLab HUGS13 summer school, and for his notes on TMD evolution.

I'm also most grateful to Prof. P. Mulders at Nikhef for the invitations to the Nikhef seminars, the sharing of knowledge and ideas and for the useful discussion during conferences. In this respect I also need to thank Prof. D. Sivers for the many discussions we had in Rome and Chile.

Finally I would like to thank my family for their support during my studies, and especially my father for taking care of many practical things when I was abroad or too busy.

Thank you all!

CONTENTS

LIST OF FIGURES

ACRONYMS

RGE Renormalization Group Equations

TMDPDF Transverse Momentum Dependent Parton Distribution Function

RHIC Relativistic Heavy Ion Collider

EMC European Muon Collaboration

FNAL Fermi National Accelerator Laboratory

DGLAP Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

BFKL Balitsky-Fadin-Kuraev-Lipatov

TMD Transverse Momentum Distribution

TMDs Transverse Momentum Distributions

RHS Right Hand Side

LHS Left Hand Side

TDD Transverse Distance Dependent

PDF Parton Density/Distribution Function

YM Yang-Mills

SYM Super-Yang-Mills

SD Schwinger-Dyson

QCD Quantum Chromo-Dynamics

pQCD	Perturbative Quantum Chromo-Dynamics
LO	Leading Order
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
DIS	Deep Inelastic Scattering
QFT	Quantum Field Theory
SIDIS	Semi-Inclusive-Deep-Inelastic-Scattering
FF	Fragmentation Functions
DY	Drell-Yan
BDS	Bern-Dixon-Smirnov
UV	Ultra-Violet
IR	Infra-Red
SD	Schwinger-Dyson
GLS	Generalized Loop Space
MM	Makeenko-Migdal
TDD	Transverse Distance Dependent
SLAC	Stanford Linear Accelerator Center
MHV	Maximally Helicity Violating
CSW	Cachazo-Svrcek-Witten
IMF	Infinite Momentum Frame
QED	Quantum Electro-Dynamics
TVS	Topological Vector Space
SUSY	Super-Symmetry
NLMC	Nuclear Locally Multiplicative Convex

Part I

QUANTUM CHROMO-DYNAMICS, DISTRIBUTION FUNCTIONS AND WILSON LOOPS.

This part will give a short history and a comparison of Feynman's parton model with the Gell-Mann-Zweig quark constituent model ultimately unified by QCD. We also give a very brief (graphical) review of renormalization, the importance of infrared safety and factorization. A discussion of two scattering processes, Deep Inelastic Scattering and Semi-Inclusive Deep Inelastic Scattering, is then used to introduce the Transverse Momentum Dependent Parton Distribution Functions. At the end of this part we discuss the relation between Wilson loops and these Transverse Momentum Dependent Distribution Functions followed by a chapter that gives an extra motivation to study Wilson loops coming from Super-Symmetry.

1 INTRODUCTION

In this chapter I give a brief account of the discoveries and problems that lead to the introduction of Transverse Momentum Distribution (TMD)s. Although the main focus of this thesis is to study of Wilson Loops and their application to the evolution equations of these TMDs this allows me to place it in a broader context. More detailed explanations and definitions are given in the chapters following this introductory overview.

This Dissertation is part of a larger program which has the goal to describe and understand the internal structure of nucleons (e.g. protons, neutrons, deuterium nuclei,...). Translated in more technical terms, to understand QCD, the theory assumed to describe the strong or nuclear force, both at the perturbative and non-perturbative level. This so-called strong force is one of the four fundamental forces (gravitational, electro-magnetic, weak and strong) so far observed in nature. The first force, gravity, will not be discussed any further in this text, which technically reflects in the fact that we only consider flat (Minkowski) metrics for the space-time backgrounds in this thesis. The last three forces are considered to be accurately described by the Standard Model having as symmetry group

$$U(1) \times SU(2) \times SU(3).$$

Here we will only be concerned about the $SU(3)$ part of the Standard Model [19], where this group is the symmetry group of a non-Abelian gauge theory referred to as QCD, describing the force that keeps nucleons together. This force is also often referred to as the strong force reflecting the fact that the gauge theory coupling constant α_s (see for instance [19]) can be large (~ 1), rendering a perturbative expansion in the parameter useless. It should then not come as a surprise that QCD is permeated by non-perturbative effects. This is a big problem from a calculation point of view since until now we have no idea how to deal with this. There are some attempts like Lattice calculations [1] , AdS-CFT [3] , Ghosts Fields [4] and so on, but none of them is fully satisfactory at the moment.

The above fact is contrasted by the huge successes of Perturbative Quantum Chromo-Dynamics (pQCD) in the last quarter of the twentieth century. But how can we understand these successes in the background of the failure of perturbation theory? For pQCD to work as well as it does we need essentially two things: **Asymptotic Freedom [20, 21] and Factorization [22–24]**.

The first is provided by the discovery that non-Abelian (sometimes referred to as Yang-Mills theories) gauge theories are renormalizable [19, 24], which allowed to derive the QCD β-function describing the running of the coupling constant α_s with energy. It came as a great surprise that this β-function showed that at high-energies the QCD coupling constant becomes small, allowing for perturbation theory, at least in some energy region.

In the mean time, on the experimental side, it had become clear that nucleons are not fundamental but composite objects. In the late sixties of the previous century this had given rise to two models: **The constituent quark model (by Gell-mann/Zweig) [25, 26] and Feynman's parton model [27–30]** later to be unified in QCD [20, 21] (see next chapters for a discussion).

Combining these partons (quarks/gluons) with Lorentz boosts and the running of the coupling constant α_s eventually lead to the discovery of **Asymptotic Freedom**, where these partons can be considered as "free" (or non-interacting with other nucleon constituents) in high-energy scattering experiments.

The second thing we needed was **factorization**, which basically allows us to separate cross-section calculations for high-energy scattering experiments in a hard and a soft part. Asymptotic freedom shows that the coupling constant in the hard part is small ($<< 1$) such that this part can be calculated in perturbation theory. On the other hand in the soft part the coupling constant is large (~ 1) such that this cannot be calculated from theory. This leads to the introduction of PDFs and FFs, describing the soft part, which fortunately can be measured in experiments. These PDFs and FFs turned out to be universal objects in the sense that they are the same for different types of processes (e.g. Deep Inelastic Scattering (DIS), Drell-Yan (DY),...), a property that is highly desirable if one wants to make predictions for other processes.

Since QCD was shown to be renormalizable, these PDFs (FFs) also need to obey some renormalization group equations. These equations were written down by Dokschitzer,Gribov, Lipatov, Altarelli and Parisi, which are now referred to as the DGLAP equations[1] [8–11]. The DGLAP equations allow us to

1 Since these PDFs (FFs) not only depend on the energy scale but also on the momentum-fraction carried by the parton there exists also and evolution equation with respect to this parameter. These evolution equations are known as the BFKL evolution equations, but are beyond the scope of this text. For more details see for instance [12, 13]

make predictions at different energy scales, which combined with Asymptotic Freedom, Factorization and the universality of the PDFs (FFs) led to the predictive power of pQCD over a large range of energies (at very low energies, the non-perturbative nature of QCD takes over, destroying the predictive power of pQCD).

But then there was spin...

In the last decade of the twentieth century experiments started showing single spin asymmetries [31, 32] in scattering experiments with polarized beams and/or targets. The polarizations of the beams and/or targets themselves do not form much of an obstruction since they can be incorporated in polarization dependent PDFs (FFs). The problem is that even when including this polarization dependence pQCD predicts that these spin asymmetries should be zero, while the experiments clearly measured a non-zero result. Such asymmetries where first measured in 1988 in the European Muon Collaboration (EMC) [33] using longitudinal polarized protons. Around the same time, at Fermi National Accelerator Laboratory (FNAL), similar asymmetries where found for transversely polarized proton's [34–39] which recently have been confirmed at larger energies at the Relativistic Heavy Ion Collider (RHIC) facility (STAR, PHENIX and BRAHMS collaborations) [40–42].

Even more upsetting was that when one added all the known spin contributions of quarks (valence and sea) one only arrives at about 25-30 percent of the total spin of the proton (first measured by EMC [33] returning a quark spin contribution of $14 \pm 9 \pm 21$ percent to the proton's spin). Naively thinking of the proton to be built out of just three valence quarks, which are as we now know spin $1/2$ particles, we would expect for the proton's spin something like $\frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2}$, but this doesn't seem to be right. So, where is the spin of the proton coming from?

As a first idea, in an attempt to tackle this problem, one would think that maybe the gluons (spin 1 particles) contribute somehow. This idea is supported by the fact that the picture of the protons changes with the kinematical variables that are used to probe the proton wave function, for instance that for large x values (this is the momentum fraction of the proton carried by the parton) the proton seems to be more or less built from three quarks, while for low x values gluons seem to dominate. Unfortunately current experimental data [43, 44], which are very limited for the moment, seems to indicate that gluons indeed contribute to the proton's spin but in an insufficient way to explain the full spin of the proton.

A second idea is then to include the orbital momenta of the quarks and gluons, but unfortunately the decomposition of the proton spin in spin and angular momenta contributions is not unique. The specific structure of the composition depends on the chosen gauge and quantization scheme[2], making things even more complicated and obscuring the physical interpretations of the different contributing parts. Measurements for both are still underway, as are measurements to reduce the error bars on all the contributions mentioned above which are still large due to limited data.

The important step in this last idea is the emergence of the dependence on the (intrinsic) transverse momenta of quarks and gluons in the PDFs (FFs). There are several different extensions of the "collinear" PDFs (FFs) to include these transverse momenta (see for instance [6, 46, 47]), but here we will restrict ourselves to Transverse Momentum Dependent Parton Distribution Function (TMDPDF)s (FFs).

These TMDs turn out not only to depend on the transverse momentum but also on a rapidity parameter, which is captured by the Wilson line structure in the TMDs definition[3]. Because of this the renormalization group equations for TMDs will be more involved then in the collinear case (DGLAP), leading to combined energy-rapidity evolution equations.

This now brings us to the main subject of this thesis, Wilson loops and their relation to TMDs.

In this thesis we will discuss the link between the rapidity evolution of some specific TMDs and the area evolution of Wilson Loops in Generalized Loop Space (GLS). Using this link we where then able to derive evolution equations for these TMDs which was the main goal of this thesis.

The structure of the text is as follows:

- Part 1 starts by giving a more detailed motivation to introduce TMDs, followed by a short review of the history of pQCD, a brief introduction to TMDs, an explicit demonstration of how the area variation of Wilson Loops is related to rapidity evolution and ends with an additional motivation coming from Super-Symmetry (SUSY) to study Wilson Loops.
- Part 2 contains an overview of most of the mathematical terminology and theorems used throughout this thesis, readers with a strong math-

2 Most well-known are the Ji, Jaffe-Manohar, Hatta and Chen decomposition, see for instance [45] and references therein.

3 A more detailed account of this and how it is related to different TMD definitions will be discussed in the next chapters.

ematical background might skip this part. Readers unfamiliar with this material can also skip this and use reference links in the electronic version to jump to the relevant definitions when necessary.

- Part 3 gives a detailed introduction to GLS (following Tavares [16]) along with its algebraic structure and differential operators necessary to understand area evolution in this space. In this part we also review some aspects of the geometric representation of gauge theory by Principal Fiber Bundles, showing explicitly how Wilson Lines and Loops emerge in gauge theory as parallel transporters.

- Part 4 contains the main results of my research. We start by reviewing some Quantum Field Theory (QFT) leading to the Schwinger-Dyson (SD) equations. After a discussion on the large N_c-limit, the SD approach together with the area derivative operation is then used to derive the Makeenko-Migdal (MM) equations, as was originally done by Makeenko and Migdal. This is followed by the formal introduction of the quadrilateral Wilson loop on the light cone, for which we then explicitly show in Section 22.5 that the area derivative, as used by Makeenko and Migdal, is not well defined. This motivated us to introduce a new derivative in Section 22.6 [48] and in Section 22.11 we show that this derivative is equivalent, at least at Leading Order, to the Fréchet derivative [49]. We then use this derivative, in combination with the renormalization parameter derivative ($\frac{d}{d\ln\mu}$) to conjecture an evolution equation for the Wilson Loop quadrilateral on the light cone in Section 22.7 [48] and investigate our conjecture at leading order for several other Wilson Loops in the following Sections [50]. Before moving on to the Next-to-Leading Order (NLO) tests of our conjecture we briefly discuss the connection between the area derivative used by MM and Polyakov's loop derivative in Section 22.12. Testing our conjecture at NLO, forced us to modify our evolution equation to incorporate the running of the coupling constant, introducing an extra derivative with respect to the coupling constant multiplied with the QCD β-function next to the renormalization parameter derivative. We then confirm our modified evolution equation at NLO for the quadrilateral on the light cone and the Π-shaped Wilson Loops in Section 22.13 [11]. Next we show the validity of our conjectured evolution equation to all orders in Super-Yang-Mills (SYM) Theory [11]. We end with an application of our evolution equations to a special distribution function namely a TDD, conjecturing an evolution equation for this function that in principle can be tested after the upgrade at the JLab facility in Virginia.

- The last part contains a conclusion and a discussion on open questions and possible future research.

We also point out that throughout the entire text we work in natural units

$$\hbar = c = 1,$$

only writing these symbols explicitly when absolutely necessary.

2 QUARKS VS PARTONS

In the late 1960's there were two competing models to describe the atom nucleus, the constituent quark model (Gell-mann and Zweig) and Feynman's parton model. This chapter will give a very brief history that led to the introduction of these models and addresses the question if the two models can be related to each other.

2.1 HISTORY

In the nineteenth century people where trying to figure out what material objects where made of. The organizing of discovered elements in the Mendeleev table seemed to indicate that there was an underlying structure to the different types of matter. Dalton proposed that all these elements where made of atoms, but this atom concept was still quite raw and not well understood.

In 1897 Thomson discovered the electron which led to the idea of a plumb pudding model for atoms, where negative and positive charges where uniformly distributed in the atom. But then in 1911 Rutherford interpreted the Geiger-Marsden scattering experiments on gold atoms as a favoring of a planetary model for the atom, where one now assumed the positive charge to be at the center of the atom and a cloud of electrons surrounding that core, much like planets around the sun.

Although these scattering experiments where still rudimentary, they contained the basic concept of studying the structure of matter by scattering experiments.

Using similar scattering experiments Chadwick then discovered the neutron in 1932, which quickly led to the proposition by Ivanenko that the atom nucleus was composed of protons and neutrons.

During the 1940's the pion triplet was discovered followed by a whole zoo of new hadronic particles during the 1950's and 60's leading to the classification of hadrons into mesons and baryons. In 1955 Robert Hofstadter [51] discovered that the proton charge was smeared out, giving an indication that it might have an internal substructure. Combining this information with the observed symmetries ($SU(3)$ multiplets of these hadrons **Gell-Mann and**

	Massive/slow	Massless/fast
Energy/Momentum	$E = \frac{p^2}{2m}$	E=p
Hadrons	Quark Model	Parton Model

Feynman
Deep Inelastic Scattering
Fast moving hadrons
$\updownarrow$
Collection of ∞ partons with wide-spread momentum distribution

Gell-Mann/Zweig
Symmetry - Hadron spectrum
Slow moving Hadrons
$\updownarrow$
Bound states of quarks

Figure 1: Differences quark and parton model

Zweig proposed their constituent quark model, where now mesons are assumed to be made up of two quarks and baryons of three quarks. Using these multiplets the Ω^- particle was predicted in 1962 and was discovered in 1964, a success for the introduction of color as a new degree of freedom. As a result **quarks were considered as a pure mathematical trick**, since there was no observation of any individual quarks.

Then in 1967 the Stanford Linear Accelerator Center (SLAC) came online and with it the start of the age of DIS, ultimately leading to **Feynman's parton model** [27–29]. Naturally Feynman's partons were identified with **Gell-Mann and Zweig's quarks and gluons**, but is this really correct? A question we will try to answer in the next section of this chapter.

2.2 ARE QUARKS AND PARTONS THE SAME THING?

Let us first try to understand the difference between the two models. The quark model was proposed based on symmetry for stationary hadrons, thus in a low-energy regime. The parton model on the other hand was proposed based on scattering experiments at relatively large energies (a few GeV) and thus for highly energetic colliding hadrons. Another difference can be found in the number of quarks or partons that are assumed to be present in hadrons, the quark model states that mesons are made up of two quarks and baryons of three quarks. In contrast the parton model has an infinite number of partons. These statements are summarized in figure 1. We can now wonder if quarks can be related to partons by boosting the Gell-Mann/Zweig quarks. To study this let us consider a simple hadron, a di-quark. In this simple model we assume both quarks to be distributed along the z-axis and t-axis with a Gaussian distribution, representing the Heisenberg uncertainty principle for position (q-number) and time (c-number). Due to the Gaussian nature, the Fourier transform is also Gaussian for momentum and energy. Both distributions are graphically represented in figure 2. Boosting these distributions has

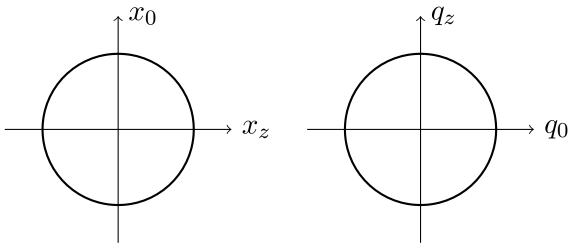


Figure 2: Di-quark hadron: quark coordinate and momentum distribution

the same effect on both of them, shown in figure 3. In time-position space the boost "weakens" the spring constant of the harmonic oscillator such that the quarks become almost free, which is consistent with the parton model. The energy-momentum boost shows that the momentum distribution of the quarks becomes wider again consistent with the parton model. But, there is an apparent violation of the uncertainty principle, for instance both momentum and position distribution become narrower. This violation is only apparent, because in the boosted frame z and q_z are not conjugate anymore, the correct conjugate variables are the light-cone (see A.2) position and momenta coordinates (x^+ and q^+). Also, there is still the problem of relating a finite number of quarks to an infinite number of partons. Boosting to the Infinite Momentum Frame (IMF), the frame where the hadron has infinite momentum along one of the light cone directions, and taking the massless approximation it follows from statistical mechanics that the number of particles is no longer conserved[1]. This now forms a bridge between the number of quarks and the number of partons. So, can we conclude from the above that partons are just boosted quarks? According to Feynman the answer is no, while according to Gell-Mann the answer is yes. As it turns out both are right, where the correct answer is provided by QCD, which will be discussed in the next section. Let

1 On light-cone massless particles can split and combine such that there is no longer a fixed number of them.

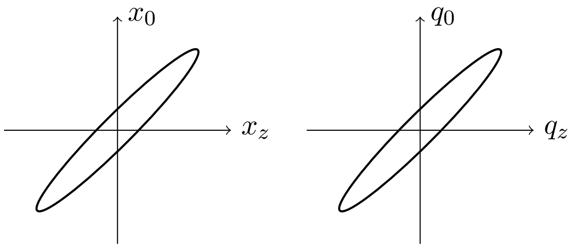


Figure 3: Di-quark hadron: boosted distributions

us end this chapter by giving again a comparison between quarks and partons but now from a QCD point of view.

Feynman	Gell-Mann/Zweig
Parton Model	Quark Model
"mass-less"	"massive"
> 3	Baryon : 3 Meson : 2
Perturbative QCD	Non-Perturbative QCD
Current quarks and gluons are fundamental DOF	Constituent quarks are "quasi" particles, dressed with gluons and quark anti-quark pairs
Point-like	Internal Structure

Figure 4: Partons vs Quarks

3

QCD AND NUCLEON STRUCTURE

This chapter will review, in a very elementary way, the most relevant concepts necessary to introduce TMDs and understand their relation to cross-sections of scattering processes.

3.1 QCD-LAGRANGIAN AND QUANTIZATION

As like any other gauge theory QCD is formulated in the Lagrangian formalism where the QCD Lagrangian is given by

$$\mathcal{L}_{\mathsf{QCD}}(\psi, A) = \sum_f \bar{\psi}_i^f(x)\Big[\Big(i\partial_\mu\delta_{ij} - gA_{\mu,a}(t_a)_{ij}\Big)\gamma^\mu - m_f\delta_{ij}\Big]\psi_j^f(x) \\ - \frac{1}{4}F^a_{\mu\nu}(x)F_a^{\mu\nu}(x), \tag{1}$$

and where the ψ_i^f are the quark fields with color index $i = 1, 2, 3 = N_c$ and flavor index $f = u, b, s, c, b, t$.

$$A_{\mu,a}(t_a)_{ij}$$

represent the color gauge field with (adjoint) color index $a = 1, 2, ..., 8 = N_c^2 - 1$.

Unfortunately if one wants to quantize this Lagrangian using the path integral method one runs into a problem, which leads to extra terms in the Lagrangian. We will here only give a brief description of this problem and how it is solved, a full treatment of this problem can be found in [19]. Let us start by writing a simple path integral for an action S only depending on the field A

$$\int \mathcal{D}Ae^{iS[A]}, \tag{2}$$

where the integral is over all field configurations. Now remember that in gauge theory there are an infinite number of field configurations that are related by gauge transformations, all representing the same physical situation. Due to this fact, the path integral in a sense over counts (physical) field configurations such that it is badly defined. Moreover, one cannot define a gluon propagator at this point (see [19] for details). Since we are ultimately only interested in physical configurations introduction of a gauge fixing condition (for instance $\partial_\mu A^\mu = 0$) allows us to separate the physical configurations from the gauge equivalent ones. This procedure leads to a factor in front of the path integral and an extra contribution, the gauge fixing term, to the Lagrangian (3). In Quantum Electro-Dynamics (QED) this extra term is independent of the field A and disappears after devision by the vacuum diagrams, in QCD on the other hand this term is field dependent and thus cannot be simply divided away. Fadeev and Popov solved this new problem by rewriting this factor as a new path integral over a set of new fields, which are now known as Fadeev-Popov ghosts. The introduction of these ghost fields adds another term, the ghost term, to the Lagrangian so that we finally end up with the following Lagrangian

$$\begin{aligned}\mathcal{L}_{\mathsf{QCD}}(\psi, A) = \sum_f \bar{\psi}_i^f(x)\Big[\Big(i\partial_\mu\delta_{ij} - gA_{\mu,a}(t_a)_{ij}\Big)\gamma^\mu - m_f\delta_{ij}\Big]\psi_j^f(x)\\ -\frac{1}{4}F^a_{\mu\nu}(x)F_a^{\mu\nu}(x) - \underbrace{\frac{1}{2\lambda}(\partial_\mu A_a^\mu)(\partial_\nu A_a^\nu)}_{\substack{\text{Covariant}\\ \text{Gauge Fixing}}}\\ + \underbrace{(\partial_\mu\bar{\eta}_a(x))(\partial^\mu\eta_a(x)) - gf^{abc}A_b^\mu(x)\eta_c(x)}_{\text{Ghosts}}\,.\end{aligned} \tag{3}$$

The parameter λ is called the gauge-fixing parameter and accounts for different choices of the covariant gauge fixing condition. What this new Lagrangian says is that if we want to use the path integral formalism to do quantum calculations in QCD in a covariant gauge, we need to take Feynman diagrams with ghost fields into account.

An alternative choice for the gauge fixing condition are the axial gauges, where

$$-\frac{1}{2\lambda}(\partial_\mu A_a^\mu)(\partial_\nu A_a^\nu)$$

is replaced by

$$\frac{1}{2\lambda}(n^\alpha A^\mu_{a\alpha})(n^\alpha A^\mu_{a\alpha}),$$

with the big advantage that no ghost fields are required, but with the disadvantage of resulting in a more complex gluon propagator. In a general covariant gauge the expression for the gluon propagator is given by

$$D^{ab}_{\mu\nu}(k) = \frac{i\delta^{ab}}{k^2}\Big(-g_{\mu\nu} + \frac{k_\mu k_\nu}{k^2}\Big(1 - \frac{1}{\lambda}\Big)\Big),$$

while in an axial gauge fixing it is given by

$$D^{ab}_{\mu\nu}(k) = \frac{-i\delta^{ab}}{k^2}\left(-g^{\mu\nu} - \frac{k_\mu n_\nu + k_\nu n_\mu}{kn} + k_\mu k_\nu \frac{n^2 + \lambda k^2}{(kn)^2}\right)$$

Most calculations in this thesis however, have been done in the Feynman gauge ($\lambda = 1$), such that the gluon propagator is well-defined and simplifies to

$$D^{ab}_{\mu\nu}(k) = \frac{i\delta^{ab}}{k^2}(-g_{\mu\nu}). \tag{4}$$

For an overview of the possible Feynman diagrams and their mathematical expressions we refer the reader to [19, 52].

3.2 RENORMALIZATION

3.2.1 *Renormalization as re-parametrization*

Calculating Feynman diagrams from the Lagrangian from the previous section one quickly runs into the problem that some of these diagrams are (Ultra-Violet (UV)) divergent. This problem was solved by the introduction of the concept of renormalization. In the renormalization procedure the divergent terms in a diagram calculations are first isolated through the introduction of a regulator (mass, dimensional,...) and then removed from the result. Of course the way this removal is done is not unique such that there exist different renormalization schemes, which we will not discuss in more detail here (see [52, 53] for a discussion).

The effect of this is the introduction of extra terms in the Lagrangian, known as **counter-terms** which depend on a renormalization (energy) scale (usually written as μ). This means that the theory itself depends on this parameter μ. The way to interpret this is that renormalization effectively re-sums terms in the perturbative calculation of a diagram in such a way that the diagram is finite order by order.

Instead of giving a long and complicated mathematical derivation, which falls outside the scope of this thesis, we follow the approach of J. Qui and use a simple example to show how this works. A formal mathematical treatment can be found in [53].

Consider the Left Hand Side (LHS) diagram in figure 5, which is UV divergent due to the large momenta that can run around in the (triangular) loop. Diagrams with these large momenta are sometimes referred to as high-mass states and when applying the uncertainty principle to them it is easy to see that they are "localized" diagrams due to the short existence time of such states. Since no experiment has an infinite resolution, high-mass states can only be seen up to a certain mass scale $\frac{1}{\mu}$. This means that the diagram on the LHS in figure 5 can be rewritten as in the Right Hand Side (RHS), separating low-mass states from high-mass states. The red blob represent a divergent term, such that the low-mass state becomes finite due to subtraction of this divergent term from the original (divergent) diagram.

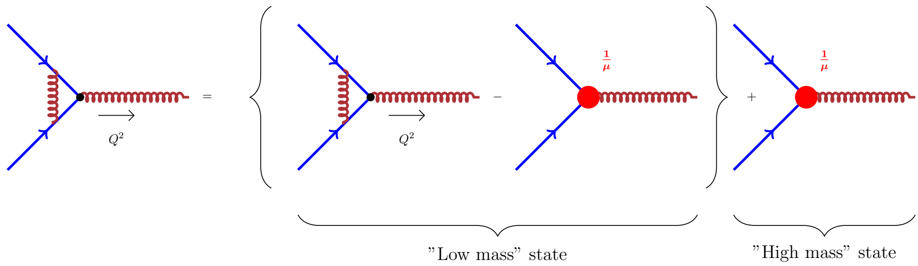


Figure 5: Example of a divergent diagram and renormalization

Combining now the different terms as in figure 6 results in a renormalized coupling constant $g(\mu)$ but with removed UV divergencies in the NLO diagram. This figure demonstrates that renormalization is indeed just a re-parametrization.

3.2.2 *β-function*

As discussed above the renormalization procedure introduces a parameter μ in the Lagrangian. This parameter is artificial in the sense that physical quantities cannot depend on it. As a direct consequence of this statement we have for instance for the cross-section σ that

$$\mu \frac{d}{d\mu} \sigma_{\text{phys}} \left(\frac{Q^2}{\mu}, g(\mu), \mu \right) = 0.$$

Such equations are referred to as Renormalization Group Equations (RGE).

As we have seen in the example above also the coupling constant gets renormalized for which in its turn the RGE lead to a running (with μ) of the coupling constant defining the Beta-function

$$\mu \frac{\partial g(\mu)}{\partial \mu} = \beta(g) = g^3 \frac{\beta_1}{16\pi^2} + \mathcal{O}(g^5),$$

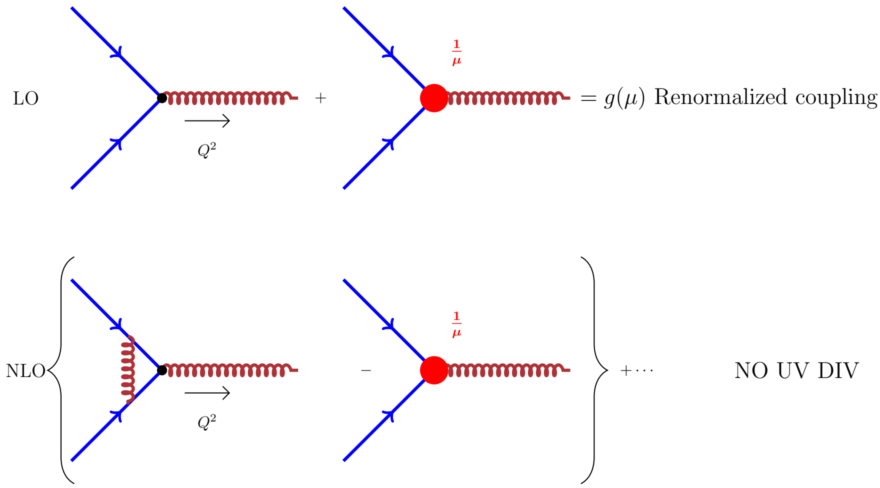


Figure 6: Re-parametrization of the perturbative series

where

$$\beta_1 = -\frac{11}{3}N_c + \frac{4}{3}\frac{1}{2}n_f \tag{5}$$

with N_c the number of colors and n_f the number of quark flavors. Notice that $\beta_1 < 0$ for $n_f \leq 16$. Thus the β-function describes how the coupling constant changes with the (energy) scale μ, and it's derivation only became possible after it was shown that Yang-Mills Theories are renormalizable. The exact form of the β-function then lead to the discovery of Asymptotic freedom.

3.3 ASYMPTOTIC FREEDOM, INFRARED SAFETY AND FACTORIZATION

3.3.1 *Asymptotic freedom*

Combining the expression

$$\alpha_s(\mu) = \frac{g^2(\mu)}{4\pi}$$

with the β-function from the previous section we can derive

$$\alpha_s(\mu_2) = \frac{\alpha_s(\mu_1)}{1 - \frac{\beta_1}{4\pi}\alpha_s(\mu_1)\ln\left(\frac{\mu_2^2}{\mu_1^2}\right)} \equiv \frac{4\pi}{-\beta_1 \ln\left(\frac{\mu_2^2}{\Lambda^2_{\text{QCD}}}\right)}. \tag{6}$$

Now taking the limit for $\mu_2 \to \infty$ the above expression goes to zero for $\beta_1 < 0$, an effect that is known as **Asymptotic freedom**. Asymptotic freedom states that the coupling constant becomes small at high energy ($\mu \gg \Lambda_{\text{QCD}}$), such that the partons become asymptotically free in this energy regime. The importance of this property becomes apparent in scattering experiments on hadrons where, due to this property, in first approximation we can neglect interactions of the constituent of the hadron we are scattering on with the other constituents, giving rise to the (generalized) parton model. Choosing the renormalization scale $\mu \sim Q$, the energy at which we perform the experiment and assuming that $Q \gg \Lambda_{\text{QCD}}$ form the background of pQCD[1].

3.3.2 *Infrared Safety*

In the previous sections we have seen that UV divergences induce renormalization, but there is another type of divergences that can occur in QCD namely Infra-Red (IR) divergences. This last type of divergences find their origin in the emission of soft (zero momentum) or collinear gluon emissions such that even in high-energy, short-distance regime, long-distance aspects of QCD cannot be ignored and are not described correctly by perturbation theory[2]. In this sense the study of IR divergencies is actually a study of which regions one can trust perturbation theory.

Divergent propagators are actually a sign of propagation of partons over long distances. In the case distances becomes comparable with the size of hadrons, the quasi-free partons of perturbative calculation are confined signaled by their hadronization which is a non-perturbative effect, and apparent divergences disappear.

In the previous section we found asymptotic freedom, which seemed to indicate that we would be able to perform perturbative calculations in a high-energy regime, but now we hit another wall or better another divergence. How can we deal with these new divergencies?

The solution is to restrict ourselves to observables for which perturbation theory is still valid. This statement seems somewhat empty but actually there are two classes of such observables

1 The freedom of choice of the scale μ is often used to optimize the accuracy of the perturbative expansion, see [52] for a discussion.

2 In the case we are considering the massless limit for quarks, also quark will contribute to IR divergencies.

(i) Infrared safe quantities: this are quantities that are insensitive to soft or collinear branching. This insensitivity can find its origin in the cancellation of IR divergencies between real and virtual contributions or are removed by kinematic factors. They are determined primarily by hard, short-distance physics. Long-distance effects show up as power corrections, suppressed by inverse powers of a large momentum scale $\mu \sim Q$.

(ii) Factorizable quantities (see also next section) : these are quantities in which infrared sensitivity can be absorbed into an overall non-perturbative factor, to be determined experimentally[3].

Cross-sections for scattering experiments are infrared sensitive such that we will need factorization theorems in order to use pQCD to calculate them.

3.3.3 *Factorization*

Hadrons are by definition non-perturbative objects such that we currently cannot derive their properties from the underlying theory. On the other hand we have seen in the previous sections that at high-energy the constituents of these hadrons become asymptotically free opening the door for perturbation theory to apply to scattering on such a constituent. Unfortunately the exact content or distribution of constituents is still non-perturbative making calculations impossible. This is the point where factorization comes in. Factorization tries, in any process containing hadrons, to separate the perturbative hard part (the scattering Feynman diagram) from the non-perturbative part (the hadron contents). Or put differently, separate the process into a perturbative and a non-perturbative part[4]. The non-calculable, non-perturbative part has to be described by a probability density function or PDF that gives the probability to find a parton with momentum fraction x in the parent hadron.

Unfortunately factorization has not been proven but for a small number of processes which include e^+e^--annihilation, DIS, SIDIS and DY (see further). PDFs, with as input a momentum fraction x, return the probability to hit a parton carrying this momentum fraction x when the hadron is struck by a photon. Again we point out that a PDF is not calculable from theory and thus needs to be determined from experiment.

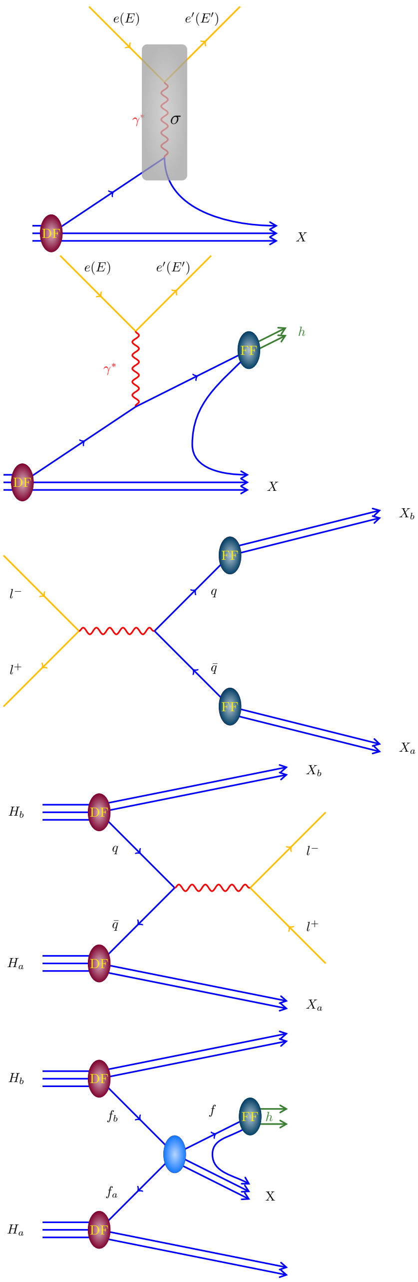
3 And essentially define the PDFs

4 In a sense factorization is the statement of the Axiom of Choice, claiming the existence of a choice function that allows one to select a particle inside the hadron that serves as the center for the hard scattering.

Next to PDFs we will also need FFs that describe how partons hadronize, in sense they describe the probability to find a hadron of a certain type inside a parton. Also these functions are non-perturbative in nature and thus not calculable, such that they also need to be determined experimentally.

We point out that the definition of both PDFs and FFs depends on the factorization scale μ_f, the scale that dictates which parts of the scattering belong to the hard part and which to the soft part. As a consequence the definition of these distribution functions is μ_f dependent, making them also susceptible to renormalization. In many cases the factorization scale is set equal to the renormalization scale i.e. $\mu_f = \mu$.

We end this section on factorization with a (graphical) overview of factorization for different types of processes where "DF" stands for Distribution Function and "FF" for Fragmentation Function. The hard part is represented by the gray blobs, and is only shown in the DIS diagram because it would make the drawings overloaded so we don't show it for the other processes. In the next chapter some of these processes will be discussed in more detail.



DIS : collinear factorization

$$\sigma^{ep\to eX} = \sum_q \mathrm{DF} \otimes \sigma^{eq\to eq}$$

SIDIS : very rich phenomenology, most explored

$$\sigma^{ep\to ehX} = \sum_q \mathrm{DF} \otimes \sigma^{eq\to eq} \otimes \mathrm{FF}$$

e^+e^- : Study Fragmentation

$$\sigma^{ee\to hhX} = \sum_q \sigma^{ee\to qq} \otimes \mathrm{FF} \otimes \mathrm{FF}$$

DY : Hard to measure in experiments

$$\sigma^{pp\to eeX} = \sum_q \mathrm{DF} \otimes \mathrm{DF} \otimes \sigma^{qq\to ee}$$

Hadron reactions : Challenge for theory due to Initial and Final state interactions

$$\sigma^{pp\to hX} = \sum_q \mathrm{DF} \otimes \mathrm{DF} \otimes \sigma^{qq\to qq} \otimes \mathrm{FF}$$

3.3.4 *pQCD summary*

We end this chapter with a short summary about pQCD.

- pQCD is a perturbative approach to the theory of QCD in the sense that it is a method to do calculation in full QCD under certain approximations.
  - We integrate out the UV region of momentum space by renormalization.
  - Match renormalized pQCD and QCD at the energy scale $\mu \sim Q$:

$$\sigma(Q, \alpha_s(0)) \to \sigma(\frac{Q}{\mu}, \alpha_s(\mu))$$

  - This operation then leads to the RGE.
- Collinear factorization in pQCD is an effective field theory of pQCD
  - We integrate out the transverse momenta of the active partons, returning the collinear factorization.
  - Match factorized form in pQCD at the energy scale $\mu_f \sim Q$ (with μ_f the factorization scale)

$$\sigma\left(\frac{Q}{\mu}, \alpha_s(\mu)\right) = \hat{\sigma}\left(\frac{Q}{\mu_f}, \frac{\mu}{\mu_f}, \alpha_s(\mu)\right) \otimes q(\mu_f, \alpha_s(\mu)) + \mathcal{O}\left(\frac{1}{Q}\right)$$

    where $q(\mu_f, \alpha_s(\mu))$ is the PDF associated to the quark of type q in the hadron.
  - μ_f independence of physical quantities leads to evolution equations for the PDFs that are known as the DGLAP equations. Note that this means that (the definition of) PDFs are scale dependent.
  - Going beyond leading order contributions introduces power corrections
    (i) generated by multi-parton correlation functions
    (ii) leading to modified evolution equations in μ_f (e.g. Bjorken scaling violation, PDFs become Q dependent)

4

NUCLEON STRUCTURE AND TRANSVERSE MOMENTUM DISTRIBUTIONS

4.1 HOW CAN WE ACCESS THE NUCLEON STRUCTURE?

The most straightforward way to access the structure of a nucleon is through interaction with an electromagnetic probe. In practice this is realized by scattering a lepton, usually an electron, off a nucleon. In this process the electron exchanges a virtual photon with four-momentum q_μ with the nucleon. If this probe is energetic enough it causes the breakup of the nucleon, a reaction

$$l + N \rightarrow l' + X$$

referred to as inelastic. From the uncertainty principle we know that the higher the momentum transfer q_μ the smaller distances we probe in the nucleon. In the case the virtual photons virtuality

$$Q^2 = -q^2$$

is large enough, the lepton is effectively interacting with the nucleon constituents, a process known as DIS.

In its simplest form we assume the photon to interact with only a single charge constituent in the nucleon, a model that we recognize from before, namely the parton model. It then doesn't come as a surprise that the nucleon constituents are often referred to as partons. It is then possible to distinguish if the parton we scatter on is a fermion or boson [54] such that it was quickly determined that the scattering happened on point-like charged fermions (quarks) and their (bosonic) gauge field (gluons).

Combining now the parton model with factorization, effectively allows us to determine the distribution of quarks in a nucleon parametrized by the PDFs. Although the PDFs are strongly modified by QCD corrections, using collinear factorization the concept of PDFs remains in a generalized parton model.

As already mentioned in the introduction, adding spin and transverse-momentum to these experiments reveals dynamics that cannot be captured fully by the collinear treatment. In these experiments emerged differential observables describing the azimuthal distributions of produced hadrons. One such type of experiment is SIDIS, where now in contrast to inclusive DIS not only the scattered lepton is measured but also one final-state hadron. This process

$$l + N \to l' + h + X$$

now also couples to the parton distribution functions, but where we now also have access to the transverse momentum of the active parton through the momentum of the measured final state hadron. Parametrizing these "new" distribution functions, that include the parton's transverse momentum dependence, require more extended PDFs that are called TMDs. Again these distributions get heavily modified by QCD corrections which are now captured by a TMD factorization process. A discussion of this factorization falls outside of the scope of this thesis, more details can be found for instance in [15, 52, 55, 56]. We only point out that the universality property is not fully retained, such that they become process dependent and require separate factorization theorems. It need not be said that this complicates things significantly.

Inclusion of spin not only affects the collinear PDFs but also the TMDs, introducing a rich collection of spin-momentum dependent distribution functions and correlation functions in cross-section calculations for SIDIS experiments with polarized beams and/or targets.

The next subsections will review both processes (DIS and SIDIS) in order to demonstrate how the different PDFs and TMDs can be parametrized, also showing how they contribute to the differential cross-section calculations. Of course a fully detailed treatment would take us to far, such that we will only sketch the general outline of the derivations. More details can then be found in the mentioned references.

4.1.1 *DIS,(spin dependent) PDFs and structure functions*

To discuss inclusive DIS in more detail we restrict ourselves to electrons for the scattering leptons on protons

$$e^- + P \to e^- + X.$$

In this type of experiments, shown graphically in figure 7 we only measure the scattered electron, or put in technical terms we integrate over all hadron final states. In order to describe this process analytically we need to chose which (kinematic) parameters we will use, which we require to be Lorentz invariant

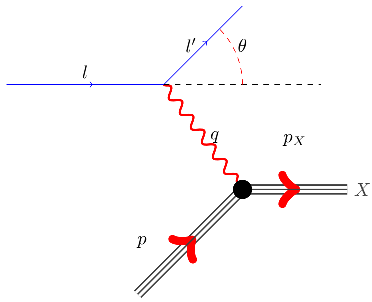


Figure 7: Proton - Lepton interaction

such that they are frame independent. In many cases these parameters are chosen to be the photon's virtuality Q^2 and the Bjorken variable

$$x_B \equiv \frac{Q^2}{2p \cdot q} = \frac{Q^2}{2M\nu},$$

M representing the mass of the proton and $\nu = E' - E$ the energy transfer. Working in the Bjorken limit, where $Q^2 \gg M^2$ we have the center-of-mass energy $\sqrt{s}$ (s is the usual Mandelstam variable) defined by

$$s \equiv (p+q)^2 = M^2 - Q^2 + 2p \cdot q \approx Q^2\left(\frac{1}{x_B} - 1\right),$$

where the approximation represents the Bjorken limit. Using the positive-definiteness of the energy this returns as kinematic limits for the Bjorken variable

$$0 \leq x_B \leq 1.$$

The photon emission and propagation can be described by perturbation theory and the coupling of this photon to the proton can be described by a matrix element of the electromagnetic current J_μ such that the Feynman amplitude can be expressed as

$$\mathcal{M} = \bar{u}(l', \lambda')\gamma^\mu u(l, \lambda)\frac{e^2}{Q^2}\left\langle P_X \left| J_\mu(0) \right| pS \right\rangle.$$

The λ, λ' and S represent the spin states of the electrons and of the incoming hadron respectively, l is the momentum of the incoming electron and $l' = l - q$ the momentum of the outgoing electron. P_X represents the unmeasured outgoing proton debris. Assuming the photon interacts with a quark, the hadronic current J_μ is given by

$$J_\mu = e_f \bar{\psi}(x)\gamma_\mu \psi(x),$$

where e_f is the charge of the quark of flavor f and a sum over flavors is assumed.

To calculate the cross-section we use the optical theorem, which is graphically represented in figure 8. This figure also shows the splitting of the differential cross-section in a leptonic part and a hadronic part represented respectively by the **Lepton Tensor** and the **Hadron Tensor** both defined through the square of the Feynman amplitude

$$|\mathcal{M}|^2 = \frac{e^4}{Q^4} L_{\mu\nu} H^{\mu\nu}.$$

The differential cross-section can then be written as

$$\begin{aligned} d\sigma &= \frac{1}{4p\cdot l}\frac{d^3 l'}{(2\pi)^3 2E_l'}\sum_X \int \frac{d^3 P_X}{(2\pi)^3 2E_X}(2\pi)^4 \delta^4(p+l-P_X-l')|\mathcal{M}|^2 \\ E_l' \frac{d^3\sigma}{dl'^3} &= \frac{2}{s-M^2}\frac{\alpha^2}{Q^4} L_{\mu\nu} W^{\mu\nu} \end{aligned} \tag{7}$$

where $L_{\mu\nu}$ represents the **leptonic tensor** given by (neglecting the lepton masses, see [57] for more details):

$$L_{\mu\nu}(l,\lambda;l',\lambda') = \delta_{\lambda\lambda'}\Big(2l_\mu l'_\nu + 2l_\nu l'_\mu - Q^2 g_{\mu\nu} + 2i\lambda\epsilon_{\mu\nu\rho\sigma} q^\rho l^\sigma\Big). \tag{8}$$

This expression also uses the Levi-Civita symbol $\epsilon_{\mu\nu\rho\sigma}$ where we use the convention $\epsilon_{0123} = 1$. The hadronic tensor is defined from (7) as

$$\begin{aligned} 2MW^{\mu\nu} &= 4\pi^3 \sum_X \int \frac{d^3 P_X}{(2\pi)^3 2E_X}\delta^4(p+q-P_X)\left\langle pS\left|J_\mu(0)\right|P_X\right\rangle \\ &\qquad \times \left\langle P_X\left|J_\nu(0)\right|pS\right\rangle \\ &= \frac{1}{4\pi}\int d^4x e^{iq\cdot x}\left\langle pS\left|J_\mu(x)J_\nu(0)\right|pS\right\rangle, \end{aligned} \tag{9}$$

where the translation operator was used

$$\left\langle pS\left|J_\mu(0)\right|P_X\right\rangle e^{i(p-P_X)\cdot x} = \left\langle pS\left|J_\mu(x)\right|P_X\right\rangle, \tag{10}$$

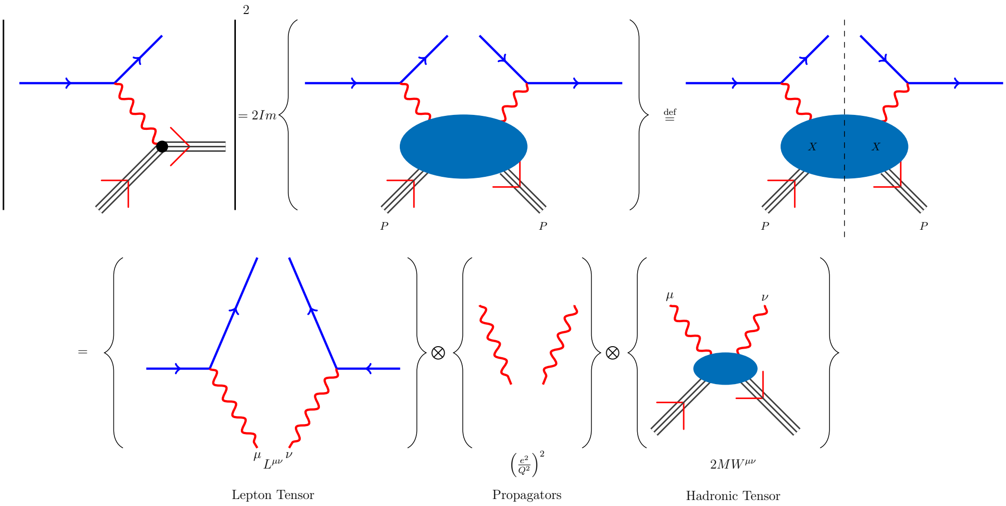


Figure 8: Optical Theorem : a graphical representation

and where we integrated out a complete set of states by use of the completeness relation:

$$\sum_X \int \frac{d^3 P_X}{(2\pi)^3 2E_X} \, |P_X\rangle \langle P_X| = \mathbb{1}. \tag{11}$$

Notice that the hadronic tensor is non-local, depending on two different space-time points and therefore non-gauge invariant. Using the operator product expansion (see [17] for details) the hadronic tensor for the handbag diagram shown in figure 9 can be rewritten as

$$W^{\mu\nu} = \sum_q e_q^2 \int d^4k \; \delta^+\Big((k+q)^2\Big) \mathrm{Tr}\Big[\phi_q \, \gamma^\mu \Big(\not{k} + \not{q}\Big)\gamma^\nu\Big] \tag{12}$$

$$\phi_{ij}^q(k,p,S) = \frac{1}{2}\int \frac{d^4x}{(2\pi)^4} e^{-ik\cdot x} \left\langle pS \left| \bar{\psi}_j(x)\psi_i(0) \right| pS \right\rangle \tag{13}$$

where ϕ is called the quark-quark correlator where a summation over color indices is assumed[1]. This correlator is a bi-local operator and thus obviously not gauge invariant. To **restore gauge invariance** one needs to introduce a **gauge link**, attached to each quark field (Mandelstam fields), making the correlator gauge invariant again. The gauge links are defined by $(\hbar = 1)$

$$U(x,y) = \mathcal{P} e^{ig \int_x^y dz^\mu \; A_\mu(z)},$$

1 The $\delta^+\big((k+q)^2\big) = \delta\big((k+q)^2\big)\theta((k+q)_0)$

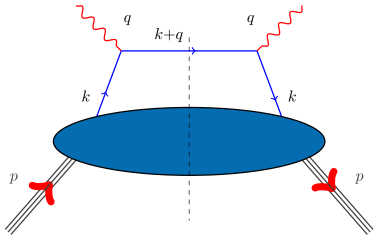


Figure 9: DIS handbag diagram.

which is in the perturbative sector[2] nothing more than the parallel transporter between the two relevant points 0 and x in the correlator (13), which will be introduced formally in the chapters on gauge theory and parallel transport. Notice already here that the **gauge links are path dependent**, which at leading order $\mathcal{O}\left(\frac{M}{Q}\right)$, can be determined from the process in hard scattering processes (or thus by the choice of **factorization scheme**). A derivation of the exact path for the gauge links falls outside of the scope of this text, the reader is referred to [7, 58–60] and reference therein for a detailed treatment. Including this gauge link (or Wilson line) in the definition of the quark-quark correlator we have

$$\phi^q_{ij}(k,p,S) = \frac{1}{2}\int \frac{d^4z}{(2\pi)^4}\, e^{-ik\cdot z}\left\langle pS\left|\bar{\psi}_j(z)U(0,z)\psi_i(0)\right|pS\right\rangle .$$

Although the introduction of the gauge links in the correlator seems random, it is **physically motivated** by the fact that after its introduction in the correlator, the correlator actually represents **a whole set of diagrams** that contribute at the same order. The situation is then similar to the case in QFT where a single diagram (like the correlator without the gauge link) is not gauge invariant but the combination of several diagrams is gauge invariant as shown in figure (10). In order to continue we change to light-cone coordinates

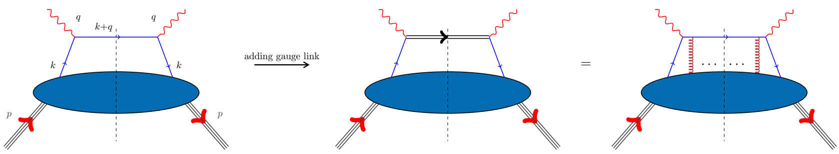


Figure 10: Effect of adding gauge link

2 In the non-perturbative sector the gauge potentials A_μ are no longer unique, such that the definition of the gauge link becomes degenerate.

(see A.2) and choose a specific frame, namely the frame where the incoming nucleon p defines the direction of the n^+ vector such that the kinematical vectors p, S, k can be written as

$$p = \frac{M^2}{2p^+}n^- + p^+n^+ \tag{14}$$

$$S = -\lambda\frac{M}{2p^+}n^- + \lambda\frac{p^+}{M}n^+ + S_\perp \tag{15}$$

$$k = k^-n^- + xp^+n^+ + k_\perp, \tag{16}$$

from which it is clear that $p^2 = M^2$ and $p \cdot S = 0$.

Returning to the correlator we see that it is constructed from a matrix element of Dirac spinors, thus is a Dirac matrix. This allows us to express it as a function of Lorentz vectors, pseudo-vectors (k, p, S) and Dirac Matrices[3]

$$B_D = (\mathbb{1}, \gamma^\mu, \gamma^\mu\gamma_5, i\sigma^{\mu\nu}, i\gamma_5),$$

with

$$\sigma^{\mu\nu} = \gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu.$$

Combining this with restrictions from Hermiticity, Parity and Time-Reversal, for which the analytical expressions are shown in table 1, returns the most general expression for the correlator (for more details see [61])

$$\begin{aligned}\phi(k,p,S) = {} & A_1 M + A_2 \not{P} + A_3 \not{k} + iA_4\frac{[\not{p},\not{k}]}{2M} + iA_5(k\cdot S)\gamma_5 \\ & + A_6 M \not{S}\gamma_5 + A_7\frac{(k\cdot S)}{M}\not{p}\gamma_5 + A_8\frac{(k\cdot S)}{M}\not{k}\gamma_5 \\ & + A_9\frac{[\not{p},\not{S}]}{2}\gamma_5 + A_{10}\frac{[\not{k},\not{S}]}{2}\gamma_5 + A_{11}\frac{(k\cdot S)}{M}\frac{[\not{p},\not{k}]}{2M}\gamma_5 \\ & + A_{12}\frac{\epsilon_{\mu\nu\rho\sigma}\gamma^\mu p^\nu k^\rho S^\sigma}{M}.\end{aligned} \tag{17}$$

Imposing **Time-reversal symmetry** will eliminate A_4, A_5 and A_{12}.

3 These matrices form a Clifford Algebra, more specifically they form the algebra of space-time.

$\phi^\dagger(k,p,S) = \gamma_0\phi(k,p,S)\gamma_0$	Hermiticity
$\phi(k,p,S) = \gamma_0\phi(k,\bar{p},-\bar{S})\gamma_0$	Parity
$\phi^*(k,p,S) = (-i\gamma_5 C)\phi(\bar{k},\bar{p},\bar{S})(-i\gamma_5 C)$	Time Reversal

Table 1: Quark-quark correlator consistency conditions [61]

For the DIS process we are interested in

$$\begin{aligned}
\phi_{ij}(x) &= \int d^2\boldsymbol{k}_\perp\, dk^- \;\; \phi_{ij}(k,p,S)|_{k^+=xp^+} \\
&= \int d^2\boldsymbol{k}_\perp\, dk^- \; \int dk^+ \delta\Big(k^+ - xp^+\Big)\phi_{ij}(k,p,S) \\
&= \int d^2\boldsymbol{k}_\perp\, dk^- \; \int dk^+ \frac{1}{2\pi}\int d\xi^- e^{ik^+\xi^-}\phi_{ij}(k,p,S) \\
&= \int \frac{d\xi^-}{2\pi} e^{ik\cdot\xi}\left\langle pS\left|\bar{\psi}_j(0)U(0,\xi)\psi_i(\xi)\right|pS\right\rangle\Big|_{\xi^+=\xi_\perp=0}. \qquad (18)
\end{aligned}$$

The matrix elements for which $\xi^+ = 0$, i.e. defined on the light cone, are called **light front matrix elements**. In this case the correlator is at equal light cone time (which is identified with the plus component in light cone coordinates), such that time ordering in these elements is automatic (see [62] and references therein). The correlator (18) can be parametrized as

$$\begin{aligned}
\phi(x) &= \frac{1}{2}\left(f_1(x)\not{n}^+ + \lambda g_1(x)\gamma_5\not{n}^+ + h_1(x)\frac{\gamma_5\left[\not{S}_\perp,\not{n}^+\right]}{2}\right) \\
&+\frac{M}{2p^+}\left(e(x) + g_T(x)\gamma_5\not{S}_\perp + \lambda h_L(x)\frac{\gamma_5\left[\not{n}^+,\not{n}^-\right]}{2}\right) \\
&+\frac{M}{2p^+}\left(-\lambda e_L(x)i\gamma_5 - f_T(x)\epsilon_T^{\rho\sigma}\gamma_\rho S_{\perp\sigma} + h(x)\frac{i\left[\not{n}^+,\not{n}^-\right]}{2}\right) \\
&+\frac{M^2}{2(p^+)^2}\left(f_3(x)\not{n}^- + \lambda g_3(x)\gamma_5\not{n}^- + h_3(x)\frac{\gamma_5\left[\not{S}_\perp,\not{n}^-\right]}{2}\right),
\end{aligned} \qquad (19)$$

where the factors

$$\frac{M}{p^+}$$

Leading twist (t=2)	Twist 3	Twist 4
$\phi^{\gamma^+}(x) = f_1(x)$	$\phi^1(x) = \frac{M}{p^+}e(x)$	$\phi^{\gamma^-}(x) = f_3(x)$
$\phi^{\gamma^+\gamma_5}(x) = \lambda g_1(x)$	$\phi^{i\gamma_5}(x) = \frac{M}{p^+}e_L(x)$	$\phi^{\gamma^-\gamma_5}(x) = \lambda g_3(x)$
$\phi^{i\sigma^{i+}\gamma_5}(x) = S^i_\perp h_1(x)$	$\phi^{\gamma^i}(x) = -\frac{M\epsilon_T^{i\rho}S_{\perp\rho}}{p^+}f_T(x)$	$\phi^{i\sigma i-\gamma_5}(x) = S^i_\perp h_3(x)$
	$\phi^{\gamma^i\gamma_5}(x) = \frac{MS^i_{T\rho}}{p^+}g_T(x)$	
	$\phi^{i\sigma^{+-}\gamma_5}(x) = \frac{M}{p^+}\lambda h_L(x)$	
	$\phi^{i\sigma^{ij}\gamma_5}(x) = \frac{M}{p^+}\epsilon_T^{ij}h(x)$	

Table 2: DIS distribution functions

assure **Lorentz invariance** but are also used to define the so called **operational twist** t of the distribution functions in the above parametrization, defined by

$$\left(\frac{1}{p^+}\right)^{t-2}.$$

The functions e_L, f_T and h vanish because of Time-Reversal Symmetry as they only depend on the amplitudes A_4, A_5 and A_{12} in (17). The distribution functions at different twist order can be extracted by projecting on the Dirac basis B_D according to

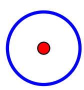
$$\phi^\Gamma \equiv \frac{1}{2}\mathrm{Tr}(\phi\Gamma),\ \Gamma \in B_D.$$

For example

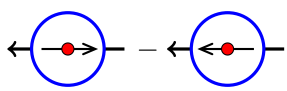
$$f_1 = \frac{1}{2}\mathrm{Tr}\Big(\phi\gamma^+\Big)$$

They are summarized in table 2. Before we move on to the SIDIS case we give a short graphical summary of the leading twist PDFs for the DIS process.

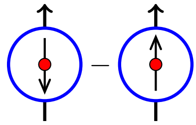
- Non-polarized PDF: $f_1^q(x, Q^2)$



- Longitudinal spin distribution : $\Delta f_1^q \equiv f^{q\rightarrow}_{\rightarrow} - f^{q\leftarrow}_{\rightarrow} \equiv g_1^q(x, Q^2)$



- Transverse spin distribution : $\delta f_1^q \equiv f_\uparrow^{q\uparrow} - f_\uparrow^{a\downarrow} \equiv h_1^q(x, Q^2)$



In the graphical representation, the red disc is the parton and the blue circle is the hadron. The drawings are assumed to move from left to right, the arrows in side the the hadron representing the polarization of the parton and the arrow outside the blue circle represents the polarization of the hadron[4]. In all the cases shown here the parton is assumed to be collinear, in the sense that it does not carry any transverse momentum $\boldsymbol{k}_\perp$. This concludes our short review of DIS and we are now ready to move on to the SIDIS case.

4.1.2 *SIDIS-TMDs*

A graphical representation of the process was given at the end of section 3.3.3 and a similar discussion to the DIS case about correlators exists, where now we are not only interested in the fully k-integrated correlators but also in the ones that are not integrated over $\boldsymbol{k}_\perp$ (the transverse component of the quark with momentum k in the frame where p has no transverse components). As stated in the introduction,we have access to $\boldsymbol{k}_\perp$ in the SIDIS process through the transverse momentum of the measured hadron. Thus the correlator of interest is now

$$\begin{aligned}\phi_{ij}(x, \boldsymbol{k}_\perp) &= \int dk^- \; \phi_{ij}(k, p, S)|_{k^+ = xp^+, k_\perp} \\ &= \int \frac{d\xi^- d^2\boldsymbol{\xi}_\perp}{(2\pi)^3} e^{ik\cdot\xi} \left\langle pS \left| \bar{\psi}_j(0) U(0,\infty) U(\infty,\xi) \psi_i(\xi) \right| pS \right\rangle \Big|_{\xi^+=0} .\end{aligned} \tag{20}$$

This introduces a dependence on the transverse separation between the quarks $\boldsymbol{\xi}_\perp$, explaining the designation of TMDs for the associated PDFs. Due to this dependence there will also be **additional gauge links** (see [7, 58–60] for details). Figures 11 and 12 show a graphical representation of these gauge links for the correlators and the fragmentation functions (see further) respectively. Restricting to twist two and twist three contributions the correlator (20) can

4 More specifically it means that when the arrows are horizontal we have longitudinal polarization, while in the vertical case we have transverse polarization.

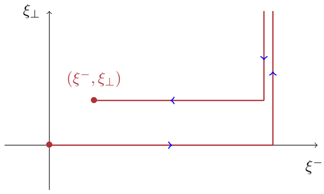


Figure 11: PDF gauge link structure SIDIS

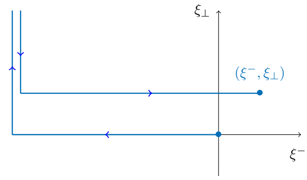


Figure 12: FF gauge link structure SIDIS

be expanded as (see [6, 47, 61] for derivations and a more detailed discussion)

$$
\begin{aligned}
\phi(x, \boldsymbol{k}_\perp) = \frac{1}{2}\Bigg(& f_1(x, \boldsymbol{k}_\perp)\not{n}^+ + f_{1T}^\perp(x, \boldsymbol{k}_\perp)\frac{\epsilon_{\mu\nu\rho\sigma}\gamma^\mu n^{+\nu} k_\perp^\rho S_\perp^\sigma}{M} \\
& + g_{1s}(x, \boldsymbol{k}_\perp)\gamma_5 \not{n}^+ + h_{1T}(x, \boldsymbol{k}_\perp)\frac{\gamma_5\left[\not{S}_\perp, \not{n}^+\right]}{2} \\
& + h_{1s}^\perp(x, \boldsymbol{k}_\perp)\frac{\gamma_5\left[\not{k}_\perp, \not{n}^+\right]}{2M} + h_1^\perp(x, \boldsymbol{k}_\perp)\frac{i\left[\not{k}_\perp, \not{n}^+\right]}{2M}\Bigg) \\
+ \frac{M}{2p^+}\Bigg(& e(x, \boldsymbol{k}_\perp) + f^\perp(x, \boldsymbol{k}_\perp)\frac{\not{k}_\perp}{M} - f_T(x, \boldsymbol{k}_\perp)\epsilon_\perp^{\rho\sigma}\gamma_\rho S_{\perp\sigma} \\
& - \lambda f_L^\perp(x, \boldsymbol{k}_\perp)\frac{\epsilon\rho\sigma_\perp\gamma_\rho k_{\perp\sigma}}{M} - e_s(x, \boldsymbol{k}_\perp) i\gamma_5 \\
& + g_T'(x, \boldsymbol{k}_\perp)\gamma_5 \not{S}_\perp + g_s^\perp(x, \boldsymbol{k}_\perp)\frac{\gamma_5 \not{k}_\perp}{M} + h_T^\perp(x, \boldsymbol{k}_\perp)\frac{\gamma_5\left[\not{S}_\perp, \not{k}_\perp\right]}{2M} \\
& + h_s(x, \boldsymbol{k}_\perp)\frac{\gamma_5\left[\not{n}^+, \not{n}^-\right]}{2} + h(x, \boldsymbol{k}_\perp)\frac{i\left[\not{n}^+, \not{n}^-\right]}{2}\Bigg) + \mathcal{O}\left(\left(\frac{M}{p^+}\right)^2\right),
\end{aligned}
\tag{21}
$$

a beastly expression where for notational convenience we introduced the notations

$$g_{1s}(x, \boldsymbol{k}_\perp) \equiv \lambda g_{1L}(x, \boldsymbol{k}_\perp) + g_{1T}(x, \boldsymbol{k}_\perp)\frac{(\boldsymbol{k}_\perp \cdot \boldsymbol{S}_\perp)}{M} \tag{22}$$

$$h_{1s}^\perp(x, \boldsymbol{k}_\perp) \equiv \lambda h_{1L}^\perp(x, \boldsymbol{k}_\perp) + h_{1T}^\perp(x, \boldsymbol{k}_\perp)\frac{(\boldsymbol{k}_\perp \cdot \boldsymbol{S}_\perp)}{M} \tag{23}$$

$$g_s^\perp(x, \boldsymbol{k}_\perp) \equiv \lambda g_L^\perp(x, \boldsymbol{k}_\perp) + g_T^\perp(x, \boldsymbol{k}_\perp)\frac{(\boldsymbol{k}_\perp \cdot \boldsymbol{S}_\perp)}{M} \tag{24}$$

$$h_s(x, \boldsymbol{k}_\perp) \equiv \lambda h_L(x, \boldsymbol{k}_\perp) + h_T(x, \boldsymbol{k}_\perp)\frac{(\boldsymbol{k}_\perp \cdot \boldsymbol{S}_\perp)}{M}. \tag{25}$$

The motivation of the restriction to twist three comes from the fact that factorization does not hold beyond this level for $\boldsymbol{k}_\perp$ dependent functions [61]. We also would like to point out that $f_{1T}^\perp, h_1^\perp, f_T, e_s, h$ and $f_l^\perp$ are Time-odd functions, which are kept to allow for the study of the spin-asymmetries in SIDIS measurements we mentioned in the introduction. Again the functions are extracted by projection on the Dirac basis B_D.

This overwhelming amount of functions are summarized in in a table 3, where the **bold functions survive integration over $\boldsymbol{k}_\perp$**, U stands for unpolarized, L for longitudinally polarized and T for Transversely polarized. For the twist two distribution functions there exists several names in the literature which we list below [6]

	Nucleon pol.	Quark pol. U	Quark pol. L	Quark pol. T
Twist 2	U	$\boldsymbol{f_1}$		$h_1^\perp$
	L		$\boldsymbol{g_1}$	$h_{1L}^\perp$
	T	$f_{1T}^\perp$	g_{1T}	$\boldsymbol{h_1}$ $h_{1T}^\perp$
Twist 3	U	$\boldsymbol{e}$		$f^\perp$
	L	$\boldsymbol{h_L}$		$g_L^\perp$
	T	h_T	g_T'	$\boldsymbol{g_T^\perp}$ $h_T^\perp$

Table 3: Twist 2 and Twist 3 TMDs

- f_1 : unpolarized TMD
- g_1 : helicity TMD
- h_1 : transversity TMD
- $f_{1T}^\perp$: Sivers TMD
- $h_1^\perp$: Boer-Mulders TMD
- $g_{1T}^\perp$: worm-gear or transversal helicity TMD
- $h_{1L}^\perp$: worm-gear, Kotzinian-Mulders or longitudinal transversity TMD

- $h_{1T}^{\perp}$: pretzelosity or quadrupole TMD.

Since in SIDIS we also measure a hadron we need more than just the distribution of quarks (gluons) in a hadron, since the hadron is formed out of fragments coming out of the struck hadron or from the active quark (as was shown at the end of section 3.3.3). In any case we need a function that describes how hadrons are formed from quarks (gluons), a process that is referred to as hadronization. The functions describing the probability to "find" a certain hadron in a quark (gluon) are called the fragmentation functions and look very similar to correlators.

This fragmentation functions depends on the quark momentum k, the momentum of the measured hadron p_h and the spin vector of this hadron S_h. They are defined by

$$\Delta_{ij}(k, P_h, S_h) = \sum_X \frac{1}{(2\pi)^4} \int d^4\xi \; e^{ik\cdot\xi} \times \langle 0 \left| U(0,\xi)\psi_i(\xi) \right| P_h, X \rangle \left\langle P_h, X \left| \bar{\psi}_j(0) \right| 0 \right\rangle \quad (26)$$

Again these functions can be expanded in terms of the vectors k, p_h, S_h and projected on the Dirac basis B_D to extract the different fragmentation functions for which we refer the reader to [6, 61]. Nevertheless we do want to show briefly how the correlators and fragmentation functions combine in the hadronic tensor for SIDIS through which they also enter the cross-section calculation where the hadronic tensor is contracted with the lepton tensor.

$$\begin{aligned}
2MW_{\mu\nu}(q, ps, P_h S_h) &= \frac{1}{(2\pi)^4} \int \frac{d^3 P_X}{(2\pi)^3 2P_X^0} (2\pi)^4 \delta^4(q + p - P_X - P_h) \\
&\quad \times \langle ps \left| J_\mu(0) \right| P_X; P_h S_h \rangle \langle p_X; P_h S_h \left| J_\nu(0) \right| ps \rangle \\
&= \frac{1}{(2\pi)^4} \int d^4x \; e^{iq\cdot x} \\
&\quad \times \left\langle ps \left| J_\mu(x) \sum_X \right| P_X; P_h S_h \right\rangle \left\langle P_X; p_h S_h \left| J_\nu(0) \right| ps \right\rangle \\
&\overset{\text{LO}}{=} \int d4p \; d^4k \; \delta^4(p + q - k) \mathrm{Tr} \left(\underbrace{\Phi(p)}_{\text{correlator}} \gamma_\mu \underbrace{\Delta(k)}_{\substack{\text{Fragmentation}\\ \text{Function}}} \gamma_\nu \right) \\
&\quad + \begin{Bmatrix} q \leftrightarrow -q \\ \mu \leftrightarrow \nu \end{Bmatrix}
\end{aligned} \quad (27)$$

Note that also the fragmentation functions depend on the gauge link structure, where this structure for the SIDIS fragmentation functions is shown in figure 12.

As a last part we would like to explain the relation between PDFs, FFs and Structure Functions, but before doing that we will give a summary of the leading twist TMDs in SIDIS and which observables are related to them. The kinematical parameters that can be used to construct observables from are given by

- Nucleon parameters
  (i) Direction $\vec{z}$ (assuming IMF)
  (ii) Helicity H
  (iii) Transverse Spin $\vec{S}_\perp$
- Parton parameters
  (i) Helicity h
  (ii) Transverse Spin $\vec{s}_\perp$
  (iii) Longitudinal Momentum Fraction x
  (iv) Transverse Momentum $\vec{k}_\perp$

which can be classified as shown in table 4. With these parameters the PDFs

Vectors	$\vec{k}_\perp, \vec{z}$
Axial Vectors	$\vec{s}_\perp, \vec{S}_\perp$
Pseudo - Scalars	h, H

Table 4: Kinematical parameters

and FFs can be summarized as shown in table 5, where in the graphics there are new arrows, compared to the DIS case, representing the transverse momentum of the parton inside hadron (yellow arrows and orange circles which use the usual convention for representing arrows pointing inwards or outwards of the page).

As a last part of this section on SIDIS we would like to demonstrate how structure functions are constructed from the convolution of PDFs and FFs in the differential cross-sections. Rather then giving a full detailed treatment we will restrict ourselves to one example of a differential cross-section and

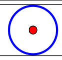
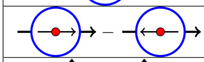
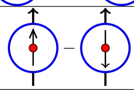
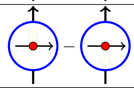
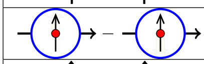
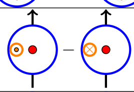
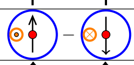
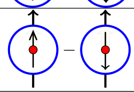
Graphics	PDF	Description	FF	Observables
	f_1	Momentum Distribution - Number Density	D_1	$\mathbb{1}$
	g_1	Helicity Distribution - Net Polarization	G_1	hH
	h_1	Transversity - Relativistic Effects Nucleon	H_1	$\vec{s}_\perp \cdot \vec{S}_\perp$
	g_{1T}	Worm-gear-T	G_{1T}	$h\left(\vec{k}_\perp \cdot \vec{S}_\perp\right)$
	$h_{1L}^\perp$	Worm-gear-L	$H_{1L}^\perp$	$H\left(\vec{k}_\perp \cdot \vec{s}_\perp\right)$
	$f_{1T}^\perp$	Sivers	$D_{1T}^\perp$	$\vec{S}_\perp \cdot \left(\vec{k}_\perp \times \vec{z}\right)$
	$h_1^\perp$	Boer-Mulders	$H_1^\perp$	$\vec{s}_\perp \cdot \left(\vec{k}_\perp \times \vec{z}\right)$
	$h_{1T}^\perp$	Prezelosity - Quadrupole	$H_{1T}^\perp$	$\left(\vec{s}_\perp \cdot \vec{k}_\perp\right)\left(\vec{s}_\perp \cdot \vec{k}_\perp\right) - \frac{1}{2}\vec{k}_\perp^2\left(\vec{s}_\perp \cdot \vec{s}_\perp\right)$

Table 5: Leading twist TMDs

explain in a graphical way how the structure functions emerge. In order to do that we first need to define some angles, which is also done in a graphical way in figure 13. The example cross-section is given by

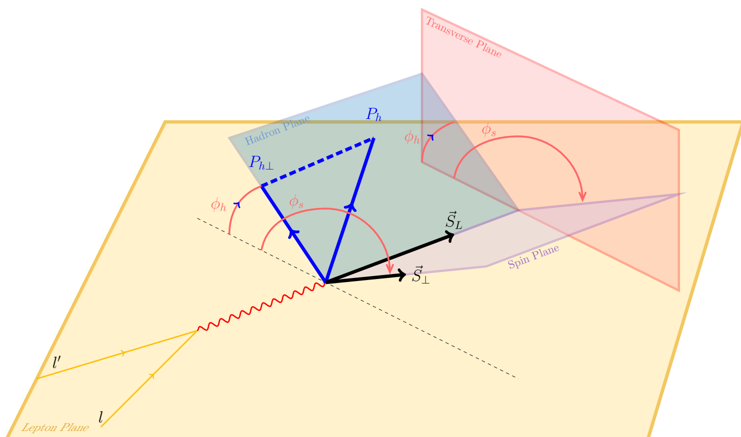


Figure 13: Defining SIDIS angles

$$\frac{d\sigma}{\delta\phi_h} = F_{UU} + \lambda_e S_L F_{LL} + \cos(2\phi_h) F_{UU}^{\cos(2\phi_h)} + S_L \sin(2\phi_h) F_{UL}^{\sin(2\phi_h)}$$
$$+ \lambda_e S_\perp \cos(\phi_h - \phi_S) F_{LT}^{\cos(\phi_h - \phi_S)}$$
$$+ S_\perp \Big[\sin(\phi_h - \phi_S) F_{UT}^{\sin(\phi_h - \phi_S)} + \sin(\phi_h + \phi_S) F_{UT}^{\sin(\phi_h + \phi_S)}$$
$$+ \sin(3\phi_h - \phi_S) F_{UT}^{\sin(3\phi_h - \phi_S)} \Big] + \text{ twist 3} \tag{28}$$

where the different F_i^j represent the different structure functions. Each of these functions is actually generate by a convolution of a PDF and a FF, to make this more clear consider the following examples

$$F_{UU} \sim \sum_a e_a^2 f_1^a \otimes D_1^a$$

$$F_{LL} \sim \sum_a e_a^2 g_{1L}^a \otimes D_1^a$$

$$F_{UU}^{\cos(2\phi_h)} \sim \sum_a e_a^2 h_1^{\perp a} \otimes H_1^{\perp a}$$

where the convolution integral is defined by

$$\otimes \to x_B \sum_q e_q^2 d^2 p_\perp d^2 k_\perp \delta(2) \Big(p_\perp - k_\perp - \frac{P_h}{z} \Big) w(p_\perp, k_\perp) f^a(x, p_\perp^2) D^a(z, p_\perp^2)$$

with

$$w(p_\perp, k_\perp)$$

an arbitrary weighting function. Following [6] this function is $\mathbb{1}$ for the first two cases and

$$-\frac{2(\hat{h} \cdot \boldsymbol{k}_\perp)(\hat{h} \cdot \boldsymbol{p}_\perp) - \boldsymbol{k}_\perp \cdot \boldsymbol{p}_\perp}{M M_h},$$

with $\hat{h}$ defined as

$$\hat{h} = \frac{\boldsymbol{P}_{h_\perp}}{|\boldsymbol{P}_{h_\perp}|}$$

for the last example. Similar relations hold for the other structure functions and details can be found in [6] and references therein.

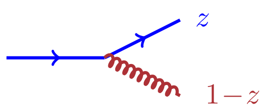
4.2 PDF AND TMD EVOLUTION: WHY WILSON LOOPS?

In the above discussion, restricting for a moment to the collinear case (no $\boldsymbol{k}_\perp$), we used factorization to split up the scattering process in its hard and soft part, where the distinction between them is set by a factorization scale μ_F. So far, although we did not say it explicitly, we only considered the leading order contribution in the hard part such that, considering higher order contributions, QCD corrections come into play. These corrections have as we know divergences (e.g. quark self-energy), such that a renormalization procedure is necessary. This ultimately, after tedious calculations, leads to a redefinition of the PDFs making them dependent on the scale μ_F (as discussed before). It is important to point out (again) that this means that **the definition of PDFs is scale dependent!**

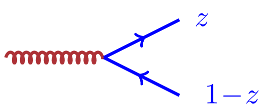
Considering that the structure functions, constructed out of these PDFs and FFs, are physical observables it is obvious that we will end up with RGE for the PDFs. For the collinear case these equations are known as the DGLAP equations [8–11] which can be summarized as

$$\frac{\partial}{\partial \ln \mu^2} \begin{pmatrix} q_i(x,\mu^2) \\ g(x,\mu^2) \end{pmatrix} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} \begin{pmatrix} P_{q_i q_j} & P_{q_i g} \\ P_{g q_j} & P_{gg} \end{pmatrix} \Bigg|_{\frac{x}{\xi}} \cdot \begin{pmatrix} q_j(\frac{x}{\xi},\mu^2) \\ g(\frac{x}{\xi},\mu^2) \end{pmatrix}. \tag{29}$$

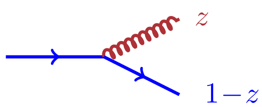
where the P_{ij} **are called the splitting functions or kernels** given at leading order by (table below is a courtesy of F. Van der Veken)



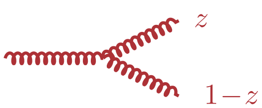
$$P_{qq}(z) = C_F \frac{1+z^2}{1-z}, \tag{30}$$



$$P_{qg}(z) = \frac{1}{2}\left(z^2 + (1-z)^2\right), \tag{31}$$



$$P_{gq}(z) = C_F \frac{1+(1-z)^2}{z}, \tag{32}$$



$$P_{gg}(z) = 2\,C_A \left(\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right). \tag{33}$$

We would now like to consider the evolution of TMDs which is unfortunately quite more involved. The reason for this is that the introduction of the gauge links (or Wilson lines), when laying on the light cone, introduced in the correlators and fragmentation matrix elements to keep them gauge invariant, makes them more divergent than usual Green functions. To deal with this several different definitions for TMDs have been put forward with respect to the gauge link structure in the correlators and fragmentation functions. Below we give an overview of the most relevant ones. In some of these definitions a new parameter, rapidity

$$\zeta^2 = \frac{(p \cdot v)^2}{|v^2|},$$

is introduced. This parameter essentially measures how much a gauge link is off-light cone and thus is can be interpreted as a kind of regulator to deal with the extra divergencies that emerge when putting the gauge links on the light cone. The divergence that emerges when taking the limit for the gauge link going from off the light cone to on the light cone is referred to as a rapidity divergence. This can be clearly seen from the definition of the rapidity parameter, where $v^2 \to 0$ if $v \to n$, with n a light cone direction.

As a side note we point out that in the papers [63–67] the projections on the Dirac basis have an index *unsub* which refers to the fact that the **soft factors** have not been factored out. These soft factors are designed to extract the rapidity divergences out of the TMDs into a multiplicative factor. We will not discuss these soft factors here in more detail since that would lead us to far, but a discussion on this subject can be found in [52, 68].

Before going into more detail let us first give an overview of the different TMD definitions.

1. Axial off-light cone (A_v):
   - Axial gauge fixing condition : $v \cdot A = 0$, where A is the gauge field and v is the (off-light cone) direction of the gauge link (such that in the gauge link integral $dz^\mu A_\mu = dt\ v \cdot A = 0$) and also the direction of the incoming hadron.
   - "Longitudinal" (along v) gauge links vanish
   - Transverse gauge links at infinity
   - Rapidity parameter dependence of the TMD used as cut-off
2. Covariant off-light cone (C_v):

- Covariant gauge fixing condition : $\partial \cdot A = 0$
- "Longitudinal" (along direction of incoming hadron) gauge links survive
- Transverse gauge links cancel
- Rapidity parameter dependence of the TMD used as cut-off
- Soft factor contains non light like gauge links

3. Axial on-light cone (A_n):
   - Axial gauge fixing condition : $n \cdot A = 0$, with $n^2 = 0$
   - "Longitudinal" (along v) gauge links vanish
   - Transverse gauge links at light cone infinity
   - Regularization parameter $\eta_{\text{LC}} = \frac{p \cdot n}{\eta}$ for the q^+ pole in the gluon propagator or regularization through Principal Value method

$$\frac{1}{[q^+]_\eta} = \frac{1}{2}\left(\frac{1}{q^+ + i\eta} + \frac{1}{q^+ - i\eta}\right)$$

   - Soft factor contains light like longitudinal and transverse gauge links

4. Covariant on-light cone (C_n):
   - Covariant gauge fixing condition : $\partial \cdot A = 0$
   - Longitudinal gauge links survive and are now on-light cone
   - Transverse gauge links cancel
   - Rapidity parameter dependence of the TMD in soft factor due to the presence of non-light like gauge links in this factor

All these definitions are a priori different and there need not be any relation between them, we only point out that integrating over the "parton" transverse momenta in the off light cone cases does not reduce to the collinear case which is rather disturbing. In this text we restrict ourselves to the unsubtracted (no extraction of soft factors) Covariant on the light cone case.

In this case for TMDs beyond the three approximation, at one loop level, the appearing divergences can be divided into three categories:

(i) Standard **Ultra-Violet poles** : removable by the usual renormalization procedure and dimensional regularization

(ii) Pure **rapidity divergences** : they depend on a rapidity cut-off but do not violate renormalizability as they can be re-summed using the Collins-Soper evolution equation [14, 15]

(iii) **Overlap divergencies** : these are the problematic ones since they contain UV and rapidity divergences simultaneously and **break renormalizability**. This type of divergencies are marked by the presence of terms proportional to

$$\frac{1}{\epsilon}\ln\zeta$$

Since we have chosen to work with light like gauge links, these overlapping divergencies will occur in our calculations. This means that if we want to apply the renormalization procedure by subtracting these "double" poles the RGE still contain a divergent part since only one pole is annihilated by the mass scale derivative

$$\mu\frac{\partial}{\partial\mu},$$

destroying the renormalization procedure. This means that we will somehow have to renormalize both divergencies at the same time, thus producing a combined evolution equation in energy and rapidity. Our approach to constructing such combined evolution equations is based on the observation that the (on light cone,divergent) rapidity evolution is related to the area variation of a Wilson loop laying on the light cone. A formal mathematical introduction and description of this area variation will be given in the last two parts of this thesis but for the moment let us just give a pictorial description to illustrate the relation with rapidity (figure 14). The Wilson loop is considered to have

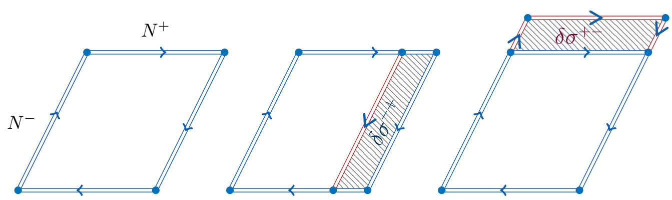


Figure 14: Light like Wilson Loop variation

all it's sides (all gauge links) on the light cone. Let now

$$\Sigma = N^+N^-$$

be the area of the loop, then the area variation is can be expressed as

$$\frac{\delta}{\delta \ln \Sigma} = \sigma^{\mu\nu} \frac{\delta}{\delta \sigma^{\mu\nu}}, \qquad \delta\sigma^{\mu\nu} = N^- \delta N^+ + N^+ \delta N^-.$$

On the other hand the (on light cone) rapidity is

$$Y = \lim_{\eta \to 0} \frac{1}{2} \ln \left(\frac{N^+ N^-}{\eta} \right),$$

and its variation is given by

$$\delta Y = \frac{1}{2} \frac{1}{N^+ N^-} \Big(N^+ \delta N^- + N^- \delta N^+ \Big).$$

From these expressions we derive that the Wilson loop area variation is related to rapidity variation

$$\frac{\delta}{\delta Y} \approx \frac{\delta}{\delta \ln \Sigma}.$$

This observation will allow us at the end of this Dissertation to conjecture an evolution equation for TDDs.

4.3 SUMMARY

In this Chapter we gave a short demonstration how considering the cross-section of DIS and SIDIS gives rise to PDFs and FFs as functions that parametrize nucleons. We gave a short overview of the lowest order contributions and how they combine in the physically observable structure functions. We ended by demonstrating how area variation of a Wilson loop lying on the light cone is related to rapidity evolution, an essential observation that will allow us to conjecture an evolution equation in the last chapters of this Dissertation.

5 SUPER YANG-MILLS, GLUON SCATTERING AMPLITUDES AND WILSON LOOPS

In this chapter we investigate the relationship between certain **gluon scattering amplitudes** in $\mathcal{N} = 4$ SYM and **Wilson loops**. We have no intention of explaining the details of SYM or even super-symmetry, for this we refer to the standard literature on these subjects. We will rather demonstrate how one can see that there is a well-established **duality between gluon scattering amplitudes and light-like Wilson loops**.

5.1 WHY $\mathcal{N} = 4$ SUPER YANG-MILLS?

SYM is a gauge theory which has an extended spectrum compared to Yang-Mills Theory without Super-Symmetry. The extra particles are two gluons with helicity ± 1, six scalars with helicity zero and eight gauginos with helicity $\pm\frac{1}{2}$.

Due to the Super-Symmetries the theory has only two free parameters, the t'Hooft coupling constant

$$\lambda = g_{YM}^2 N_c$$

and the number of colors N_c. In this theory at weak coupling the number of contributing Feynman integrals is very large compared to QCD but the final answer is much simpler, and at strong coupling by making use of the AdS/CFT correspondence it can be described by a weakly coupled string theory on a $AdS^5 \times S^5$ manifold [69–74].

5.2 SCATTERING AMPLITUDES

Scattering amplitudes have the property that they are **on-shell elements of the S-matrix** and that they are gauge independent. Not unlike **Wilson loops**, they are also non-trivial functions of the **generalized Mandelstam variables**

$$s_{ij} = (p_i + p_j)^2.$$

In SYM theory these amplitudes are easier but they still share many properties with the QCD amplitudes and they seem to have a remarkable structure that allows to make all-order conjectures about them. Furthermore, the large N_c limit of QCD is very similar to the $\mathcal{N} = 4$ SYM theory. Recently Arkani-Ahmed discovered the **amplituhedron** [75–77], which allows for full **all-order calculations** by calculating the **volume** of this new mathematical object. Perhaps, in the future, this object can be modified to a QCD setting providing new calculation methods for scattering amplitudes.

To show how the **Wilson loops** come out in quite a natural way let us investigate these scattering amplitudes in a little bit more detail by looking at **on-shell gluon scattering**. The quantum numbers describing these gluons are given by

$$|\, i\rangle = |\, p_i, h_i, a_i\rangle\,,$$

with $p_i^2 = 0$ the momentum, $h = \pm 1$ its helicity and a_i its color index. In SYM these scattering amplitudes suffer from IR **divergences** such that an IR **regulator** is needed and they look very similar to the gluon amplitudes in QCD. The amplitudes for **planar diagrams** can be parametrized by **partial color-ordered amplitudes**

$$\mathcal{A}_n = \text{Tr}\ \left[T^{a_1}\cdots T^{a_n}\right] A_n^{h_1,\cdots,h_n}(p_1,\cdots,p_n) + [\text{Bose-symm.}],$$

which can be classified according to their **helicity structure**. Due to supersymmetry relations we have that

$$A^{+\cdots+} = A^{-+\cdots+} = 0$$

from which we conclude that all four and five gluon scattering amplitudes are Maximally Helicity Violating (MHV). In general MHV amplitudes are amplitudes with n **external gauge bosons**, where $n-2$ gauge bosons have a **particular helicity** and the other two have the **opposite helicity**. These amplitudes are called MHV amplitudes, because at tree level, they **violate helicity conservation to the maximum extent** possible. These MHV amplitudes can be calculated efficiently by the **Parke-Taylor formula** [78]

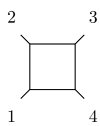
$$\mathcal{A}(1^+\cdots i^-\cdots j^-\cdots n^+) = i(-g)^{n-2}\frac{\langle i\ j\rangle^4}{\langle 1\ 2\rangle\langle 2\ 3\rangle\cdots\langle (n-1)\ n\rangle\langle n\ 1\rangle},$$

which was derived rigorously in [79]. This result was interpreted geometrically by Edward Witten when he introduced **Twistor String Theory** [80]. Using this geometrical interpretation Witten, together with Cachazo and Svrcek, introduced the **Cachazo-Svrcek-Witten (CSW)-rules** where MHV amplitudes are glued together to **build complex three diagrams**. They can be derived from the light cone Yang-Mills Lagrangian by performing a **canonical change of variables** [81]. These rules can be continued to quantum theory, i.e. constructing loops of MHV diagrams but it has problems for which we do not have the time and space here to discuss them any further (see [82] for a discussion). Together with Ruth Britto CSW generalized their approach to the **BCSW-rules** [83] which allows for **strong simplification** of calculating gluon scattering amplitudes due to these recursion relations, making them very relevant for LHC background calculations at higher orders in perturbation theory.

Before moving on the **singularity structure of the amplitudes** we would like to draw the readers attention to the **emergence of the twistor concept**. We will not explain this concept but point out that it is quickly gaining attention obviously in the **scattering amplitude community** but also in the **nucleon three-dimensional modeling community**. This is not such a big surprise since the **Penrose transform**, an integral transform used to express particle fields as a function of twistors, is very **similar to the Radon transform** which is used in **tomography** where one creates an image from the scattering data associated with cross-sectional scans of an object. More specifically, if a function f represents an unknown density, then the Radon transform represents the scattering data obtained as the output of a tomographic scan **(in nuclear physics this tomography is related to Generalized Parton Distributions)**. The inverse of the Radon transform allows then to reconstruct the original density from the scattering data, and thus it forms the mathematical basics for tomographic reconstruction, also known as image reconstruction.

Returning to the IR divergences, we first look at the following relation for the all-order four-gluon scattering amplitudes in SYM

$$\frac{\mathcal{A}_4}{\mathcal{A}_4^{\text{tree}}} = 1 + a$$



$$+ \mathcal{O}(a^2), \quad a = \frac{g_{\text{YM}}^2 N_c}{8\pi^2}. \tag{34}$$

which was derived in [84]. This result allows for the amplitude to be factorized in a finite and an **IR divergent** part

$$\mathcal{A}_4(s,t) = \text{Div}(s,t,\epsilon_{IR})\text{Fin}(\frac{s}{t}),$$

with

$$s = (p_1 + p_2)^2 \qquad \text{and} \qquad t = (p_1 + p_3)^2,$$

where the ϵ is the parameter associated with the **dimensional regularization operation**. As in every gauge theory ([85–87] and references therein) the **IR divergences exponentiate**

$$\text{Div}(s,t,\epsilon_{IR}) = \exp\left[-\frac{1}{2}\sum_{l=1}^{\infty} a^l \left(\frac{\Gamma^{(l)}_{\text{cusp}}}{(l\epsilon^2_{IR})} + \frac{G^{(l)}}{l\epsilon_{IR}}\right)\left\{(-s)^{l\epsilon_{IR}} + (-t)^{l\epsilon_{IR}}\right\}\right]. \tag{35}$$

In [88] it was then demonstrated that there exists a one-to-one correspondence between these IR divergences and UV divergences of cusped Wilson loops where in (35):

$$\Gamma_{\text{cusp}}(a) = \sum_l a^l \Gamma^{(l)}_{\text{cusp}} = \text{ cusp anomalous dim of Wilson loops} \tag{36}$$

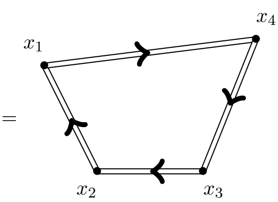
$$G(a) = \sum_l a^l G^{(l)}_{\text{cusp}} = \text{ collinear anomalous dim.} \tag{37}$$

Although interesting results on the finite part exist (for instance Bern-Dixon-Smirnov (BDS) conjecture [89]) we will not discuss it here further since we are interested in the singular behavior. Using the dual coordinates

$$p_i = x_i - x_{i+1}$$

where the p_i refer to the gluon momenta a new symmetry of gluon scattering amplitudes in SYM was discovered, namely the dual conformal invariance [90]. Combined with the IR-UV duality in [91] it was shown that the expectation value of a light-like Wilson loop possesses the same properties. More explicitly for four-gluon scattering we can write

$$W(C_4) = \frac{1}{N_c}\left\langle 0 \left| \text{Tr}\, \mathcal{P}\, e^{ig \oint_{C_4} dx^\mu \; A_\mu(x)} \right| 0 \right\rangle, \quad C_4 = \tag{38}$$



This contour is now composed of four **light-like segments**, has four light-like cusps generating the UV divergencies and where the **conformal symmetry** of SYM is mapped onto the conformal symmetry of the Wilson loop in terms of the coordinates x^μ. As we have shown in our paper [48], consistent with previous result from other authors and which will be derived in this thesis, that at one loop order the vacuum expectation value of this contour results in

$$\ln W(C_4) = \frac{g^2}{4\pi^2} C_F \left(-\frac{1}{\epsilon_{UV}^2} \left[\left(-x_{13}^2 \mu^2\right)^{\epsilon_{UV}} + \left(-x_{24}^2 \mu^2\right)^{\epsilon_{UV}} \right] + \frac{1}{2} \ln \left(\frac{x_{13}^2}{x_{24}^2} \right)^2 + \text{const} \right) + \mathcal{O}\left(g^4\right), \tag{39}$$

where from comparison with the gluon scattering amplitude at LO:

$$\ln \mathcal{A}(s,t) = \frac{g^2}{4\pi^2} C_F \left(-\frac{1}{\epsilon_{IR}^2} \left[\left(\frac{-s}{\mu_{IR}^2} \right)^{\epsilon_{IR}} + \left(\frac{-t}{\mu_{IR}^2} 2 \right)^{\epsilon_{IR}} \right] + \frac{1}{2} \ln \left(\frac{s}{t} \right)^2 + \text{const} \right) + \mathcal{O}\left(g^4\right), \tag{40}$$

we have the suggested identifications

$$x_{13}^2 \mu^2 := \frac{s}{\mu_{IR}^2}, \qquad x_{24}^2 \mu^2 := \frac{t}{\mu_{IR}^2}, \qquad \frac{x_{13}^2}{x_{24}^2} := \left(\frac{s}{t} \right). \tag{41}$$

It is now clear how the UV divergencies of the Wilson loop and the IR divergencies of the gluon amplitude map onto each other. We also mention that the finite terms coincide at least at the one loop order. From this Drummond, Henn, Korchemsky and Sokatchev proposed the duality between light-like Wilson loops and gluon amplitudes

$$\ln \mathcal{A}_4 = \ln W(C_4) + \mathcal{O}(\frac{1}{N_c^2}, \epsilon_{IR}). \tag{42}$$

The duality has been shown to hold at leading order in

$$\frac{1}{\sqrt{\lambda}} \qquad \text{(t'Hooft coupling)}$$

at strong coupling [70] using AdS/CFT to calculate the amplitudes in String Theory, at weak coupling it was verified up to two loop level [91]. The duality has also been generalized to $n \geq 5$ gluon MHV amplitudes, where the corresponding Wilson loop has now n cusps and n light-like segments. At weak coupling the BDS ansatz was confirmed [82] and for $n = 5$ the duality was confirmed at the two loop level.

Finally in [92, 93] the **Twistor concept** was used by David Skinner and collaborators to derive a **Super-Symmetric Twistor version of the MM** equations (see Chapter 21), which they then used to prove the **duality between Super-Symmetric Wilson loops and Correlation Functions (Amplitudes)**.

These are all very nice results, but in a Super-Symmetry setting. So what can we say when we consider QCD? Korchemsky showed in [94, 95] that for QCD the duality only holds in the **Regge limit**. Nevertheless Wilson loops, be it in regular Minkowski space, in Super-Symmetry, in String Theory or Twistor Theory, can provide us with a means to investigate symmetries and properties in a plethora of different theories. For us the main relevance remains that they have a singularity structure that is related to the singularity structure of TMDs, show up as soft factors in factorization schemes and ultimately they might allow us to recast QCD in a loop space setting. We also mention that the **cusp anomalous dimension** seems to show up in quite different settings, indicating its importance in understanding and studying not only QCD but any QFT.

5.3 SUMMARY

In this chapter we gave a very short overview of how Wilson loops are related to n-gluon scattering in $\mathcal{N} = 4$ Super-Yang-Mills theory. We explained how through several advances in calculation methods one was able to derive a duality between gluon scattering amplitudes in SYM and Wilson loops. A main attribute in this duality is the relation between singularities in vacuum expectation values of Wilson loops and those in gluon scattering amplitudes, giving an extra motivation to study the singularity structure of Wilson loops. It was also pointed out that the singularity structure of gluon scattering amplitudes is related to the cusp anomalous dimension, and since the singularity structure of TMDs is related to the ones in its Wilson line structure it will come as no surprise that in the evolution equation for TMDs we conjectured in our papers will depend strongly on the cusp anomalous dimension.

Combining the introduction on TMDs and on SYM the following diagram (figure 15) gives quite a good overview of the logic we applied to derive our evolution equation. The rest of this thesis will deal with the details of this diagram.

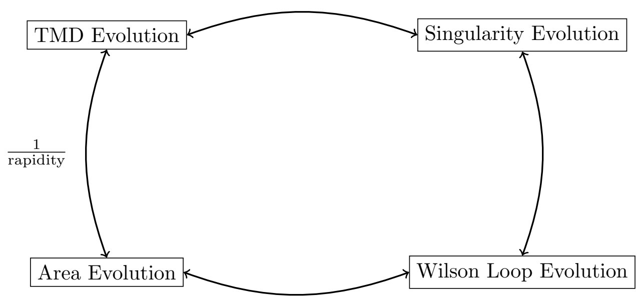


Figure 15: Overview diagram

Part II

MATHEMATICAL PRELIMINARIES

In this part we will review the necessary mathematical concepts needed to understand the next part of this thesis, where we introduce GLS. Topics covered are Topology, Differential Geometry, Algebra, Topological Algebras and Category Theory. In the review of Category theory we also introduce different categories of differentiations that will be used to generalize the concept of paths and loops, which are not standard material. We recommend readers familiar with most of the mathematical material to quickly browse through the Category Chapter before continue reading. We mention that the PDF version of this document has hyperlinks to most of the definitions, theorems,... given in this part to increase readability and reduce the necessity of looking up concepts in reference works.

6 GENERAL TOPOLOGY

6.1 BASIC DEFINITIONS

This chapter revises some of the basic concepts from general topology that will be used in this thesis. This review is far from complete and is not supposed to be an introduction to topology but rather a help for the reader, to avoid the need to lookup definitions in different textbooks.

Definition 6.1.1 (Topology)**.**
Let X be a set and $\mathcal{U}$ a collection of subsets of X. We call X a topological space provided that:

(i) $\emptyset, X \in \mathcal{U}$

(ii) $\mathcal{U}$ *is closed under finite intersections:*

$$U_1, \cdots, U_N \in \mathcal{U},\ N \in \mathbb{N} \Rightarrow \bigcap_{k=1}^{N} U_k \in \mathcal{U}$$

(iii) $\mathcal{U}$ *is closed under arbitrary (possibly uncountably infinite) unions:*

$$U_\alpha \in \mathcal{U}, \alpha \in A \Rightarrow \bigcup_{\alpha \in A} U_\alpha \in \mathcal{U}.$$

The sets $U \in \mathcal{U}$ are called open, their complements $X - U$ closed in X.

Defining a topology is stating which subsets of X are called open. From a topology we derive the concept of a **neighborhood of a point** $x \in X$. If $x \in X$ is a point and U an open set containing it, then U is called a neighborhood of x in X. When we have different topologies, defined on X, they can be compared with eachother, where a topology $\mathcal{U}$ is called stronger (finer) then a topology $\mathcal{U}'$ which then is weaker (coarser) if $\mathcal{U}' \subset \mathcal{U}$, as collections of subsets.

Considering the subsets of a space X, equipped with a topology, this naturally induces a topology on its subsets, referred to as the induced topology.

Definition 6.1.2 (Induced topology).
Let $(X,\mathcal{U}),(Y,\mathcal{V})$ be topological spaces such that $Y \subset X$. The relative or subspace topology $\mathcal{U}_Y$ induced on Y is given by defining the sets $U \cap Y$; $U \in \mathcal{U}$ to be open. We say that we have a topological inclusion, denoted by

$$Y \hookrightarrow X,$$

provided that the intrinsic topology $\mathcal{V}$ is stronger than the relative one ($\mathcal{U}_Y \subset \mathcal{V}$).

Different topologies can have varying properties, which will determine strongly how much structure is imposed on the considered sets. Some of these properties will be highly desirable from a physics point of view, allowing us to apply certain operations that would not be allowed on topological spaces not having these desired properties. One of the most relevant properties for us is the so called Hausdorff property:

Definition 6.1.3 (Hausdorff).
A topological space X is said to be Hausdorff iff for any two of its points $x \neq y$ there exist neighborhoods U, V of x, y respectively, which are disjoint.

This essentially allows one to separate points on the considered topological space. This will become highly relevant when considering limits.

6.2 TOPOLOGY AND BASIS

A common situation in physics is that one has a set X and a collection of subsets $\mathcal{U}$ of X on which we want to apply some operations, like for example take a derivative. To be able to define consistently such operations one usually needs quite a number of properties that are generated by choosing a topology on X such that all the elements of $\mathcal{U}$ are open. In many cases we want to restrict the imposed structure to a minimum which translates in finding an optimal topology with the above property. The fact that at least there exists[1] such a topology where all elements of $\mathcal{U}$ are open is treated by the following lemma.

Lemma 6.2.1

Let X be a set and $(\mathcal{U}_\beta)_{\beta\in B}$ a collection of topologies on X. Then $\mathcal{T} := \cap_{\beta\in B}\mathcal{U}_\beta$ is again a topology on X.

This topology can now be optimized due to the following proposition:

Proposition 6.2.1

Let X be a set and $\mathcal{D} \subseteq \mathcal{P}(\mathcal{X})$, a collection of subsets of X. Then there exists a coarsest, or weakest, topology on X with the property that all subsets $U \in \mathcal{D}$ are open. In other words, there exists a topology $\mathcal{T}$ such that:

(i) every $U \in \mathcal{D}$ is open in $\mathcal{T}$

(ii) if $\mathcal{U}$ is a topology on X such that every $U \in \mathcal{D}$ is open in $\mathcal{U}$, then $\mathcal{U}$ is finer then $\mathcal{T}$

Here $\mathcal{P}(X)$ represents the power set of X, i.e. the set of all subsets of X. Unfortunately proposition 6.2.1 does not give us an explicit method to determine such a topology $\mathcal{T}$. It will take some extra work before we can write an explicit form, so let us first consider a simpler case where the collection of sets $\mathcal{D}$ have an extra property.

Definition 6.2.1 (Topology basis)**.**
Let X be a set. A basis for a topology on X is a collection $\mathcal{B}$ of subsets of X with the properties:

(i) For every $x \in X$ there exists a $B \in \mathcal{B}$ such that $x \in B$

(ii) If $B_1, B_2 \in \mathcal{B}$, then there exists a $B_3 \in \mathcal{B}$ with $x \in B_3$ and $B_3 \subseteq (B_1 \cap B_2)$.

1 Choosing the power set $\mathcal{P}(X)$ for $\mathcal{U}$ makes all subsets of X open, such that there always exists a topology to start with.

If $\mathcal{T}$ is the coarsest (weakest) topology on X such that all $B \in \mathcal{B}$ are open in $\mathcal{T}$, then we call $\mathcal{B}$ a basis for $\mathcal{T}$ or we call $\mathcal{T}$ the by $\mathcal{B}$ generated topology.

Proposition 6.2.2

Let $\mathcal{B}$ be a basis for a topology on a set X and let $\mathcal{T}$ be the topology generated by this basis. If $U \subseteq X$ then the following properties are equivalent:

(i) U is open in $\mathcal{T}$

(ii) For every $x \in U$ there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$

(iii) U can be written as a union of sets B_α from the collection $\mathcal{B}$

In physics we will be mostly consider spaces that are equipped with a metric. This metric can be used to construct a topology where the open balls will form a basis for this topology. So let us first revise the definition of a metric on a set.

Definition 6.2.2 (Metric on a set)**.**
Let X be a set. A metric on X is a function

$$d : X \times X \to \mathbb{R}_{\geq 0},$$

with the following properties:

(i) $d(x, y) = 0$ *if and only if* $x = y$.

(ii) Symmetry :

$$d(x, y) = d(y, x), \ \forall x, y \in X.$$

(iii) Triangle inequality :

$$d(x, z) \leq d(x, y) + d(y, z) \ \forall x, y, z \in X.$$

A set with a metric is called a metric space. A metric is called an **ultra-metric** *if it satisfies the following stronger version of the* **triangle inequality** *where points can never fall* **between** *other points:*

$$\forall x, y, z \in X,\ d(x,z) \leq \max\left(d(x,y), d(y,z)\right)$$

A metric d on X is called **intrinsic** *if any two points $x, y \in X$ can be joined by a curve with length arbitrarily close to $d(x,y)$. For sets on which an addition*

$$+ : X \times X \rightarrow X$$

is defined, d is called a **translation invariant metric** *if:*

$$d(x,y) = d(x+a, y+a),\ \forall x, y, a \in X.$$

We now explicitly construct the topology induced by the metric. Define an open ball for $x \in X$ and a real number $r \geq 0$,

$$B(x,r) := \{y \in X \mid d(x,y) < r\}, \tag{43}$$

and a collection

$$\mathcal{B} := \{B(x,r) \mid x \in X,\ r > 0\}. \tag{44}$$

Simple calculations show that the B obey the conditions of the definition of a topology basis (6.2.1) and thus form the basis of a topology by definition. Note that the topological space $(X, \mathcal{T})$ is called metrizable if there exists a metric on X and that such a space is Hausdorff. Above we have constructed a topology starting from a given basis $\mathcal{B}$, but the goal is to construct a topology starting from a given collection of subsets, that not necessarily obey the properties of a basis. We need something more before we can realize such a construction, namely the concept of a sub-basis

Definition 6.2.3 (Subbasis of a topology)**.**
Let X be a set. A sub-basis of a topology on X is a collection $\mathcal{S}$ of subsets of X with the property that $\bigcup_{S \in \mathcal{S}} = X$.

Sub-bases have the nice property that they can be used to construct a basis.

Proposition 6.2.3

Let $\mathcal{S}$ be sub-basis for a topology on X. Then define the collection $\mathcal{B}$ of subsets $B \subseteq X$ that can be written as the intersection of a **finite** number of sets in the collection $\mathcal{S}$. Put differently, $B \in \mathcal{B}$ if and only if there exists $S_1, S_2, \cdots, S_n \in \mathcal{S}$ such that $B = S_1 \cap S_2 \cap \cdots \cap S_n$. Then $\mathcal{B}$ is a basis for a topology on X and the topology generated by $\mathcal{B}$ is the coarsest (weakest) topology on X with the property that every $S \in \mathcal{S}$ is open in this topology.

From proposition 6.2.3 it is now easy to construct a topology from a given collection of subsets. One just adds the set X to this given collection so that this new collection becomes a sub-basis for a topology on X. Proposition 6.2.3 then shows how to construct a basis and the coarsest topology for which the original collection of sets are now open. In this way we have achieved our initial goal of this section.

6.3 CONTINUITY

In this section we introduce the concept of continuity and some of its relations to topology and properties of topological spaces.

Definition 6.3.1 (Continuity).
A function

$$f:\ X \to Y$$

between topological spaces X, Y is said to be **continuous** *provided that the pre-image $f^{-1}(V)$ of any set $V \subset Y$ that is open in Y is open in X.*

The pre-image is defined by

$$f^{-1}(V) = \{x \in X;\ f(x) \in V\} \tag{45}$$

and despite the notation does not require f to be either an injection or a surjection. It is easy to demonstrate that f is continuous if it is continuous at each point $x \in X$, where f is continuous at $x \in X$ if for any open neighborhood V of $y = f(x)$ there exists an open neighborhood U of x such that

$$f(x') \in V,\ \forall x' \in U,$$

in other words $f(U) \subset V$. Strengthening the continuity conditions, by requiring that the inverse function is also continuous makes f into a **homeomorphism**.

Definition 6.3.2 (Homeomorphism)**.**
If f is a continuous bijection and also f^{-1} is continuous then f is called a **homeomorphism** *or a* **topological isomorphism***.*

One should think of a homeomorphism as an isomorphism between topological spaces. We point out that definition 6.3.1 for continuity is consistent with the usual definition of continuity in real calculus in the following way:

Corollary 6.3.1

Let (X, d_X) and (Y, d_Y) be metric spaces. Let $f : X \to Y$ be a map. Then f is continuous with respect to the metric topologies on X and Y if and only if:

$$\forall x \in X, \forall \epsilon > 0, \exists \delta > 0 : f(B(x, \epsilon)) \subseteq B(f(x), \delta).$$

Put differently f is continuous if and only if:

$$\forall x \in X, \forall \epsilon > 0, \exists \delta > 0 | d_X(x, \xi) < \delta \;\Rightarrow\; d_Y(f(x), f(\xi)) < \epsilon, \;\forall \xi \in X.$$

Interestingly continuity can be used to define a topology on products of sets.

Definition 6.3.3 (Product topology)**.**
Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of topological spaces, then consider the product set

$$Y := \prod_{\alpha \in A} X_\alpha.$$

Write

$$pr_\alpha : Y \to X_\alpha$$

for the projection on the factor X_α. We then define the product topology on Y as the coarsest (weakest) topology on Y where each of the projections pr_α is continuous.

Introducing extra conditions on a map allowed us to strengthen continuity to homeomorphism, putting now even more restrictions we can extend this further to define **open or closed maps**.

Definition 6.3.4 (Open/Closed map).
Let

$$f : X \to Y$$

by a map between topological space. Then f is called an open map if

$$\forall U \subseteq X$$

its image $f(U)$ is open in Y. Alternatively f is called a closed map if

$$\forall U \subseteq X$$

its image $f(U)$ is closed in Y.

In a certain way maps allow to transfer properties from one space to the next, usually this means that one can ask questions about the structure of the image of the source space when considered as part of the target space. Depending on how much structure is transferred by the map, maps get different names. For further reference, when we will discuss manifolds, we need such a specific map which is referred to as an **embedding**.

Definition 6.3.5 (Embedding).
A continuous map

$$i : X \to Y$$

is called an embedding if i is injective and is a homeomorphism from X to its image

$$i(X) \subset Y,$$

where $i(X)$ is equipped with the subspace (induced) topology.

An embedding thus has the following three properties:

(i) i is continuous

(ii) i is injective

(iii) $\forall U \subseteq X, \exists V \subseteq Y$ with $U = i^{-1}(V)$.

For the moment these properties might seem a bit random, but when we discuss maps between manifolds they will turn out to be very useful.

6.4 CONNECTEDNESS

In this section I will review the basic definitions of connectedness of a topological space.

Definition 6.4.1 (Connected)**.**
A topological space X is called disconnected if there exist non-empty open subsets

$$U, V \subset X$$

such that

$$U \cap V = \emptyset \qquad \text{and} \qquad U \cup V = X.$$

A topological space is called connected when it is not disconnected.

A bit strange is that according to this definition the empty set is connected. From the definition it also follows that X is only connected if and only if the only **clopen** (open and closed at the same time) subsets of X are $\emptyset$ and X. A property that is sometimes useful to check if a space is connected or not. From considering maps between topological spaces and restricting to continuous maps we get a generalization of the Mean Value Theorem.

Proposition 6.4.1: Mean Value

Let

$$f : X \to Y$$

be a continuous map. If X is connected then

$$f(X) \subseteq Y$$

is also connected.

The link with the usual Mean Value theorem from real calculus can be demonstrated by considering the map

$$f : X \to \mathbb{R}.$$

Under the assumption that this map is continuous we have from proposition 6.4.1 that is f is continuous we have that

$$f(X) \subseteq \mathbb{R}.$$

Put differently we have that if

$$x, y \in X$$

with

$$f(x) < f(y)$$

and c is a real number such that

$$f(x) < c < f(y),$$

then

$$\exists z \in X : f(z) = c.$$

Sometimes it will be useful to parametrize a space by its connected components, which can be considered as the classes of the equivalence relation induced by connectedness.

Definition 6.4.2 (Connected components)**.**
Let X be a topological space. The equivalence classes for the equivalence relation introduced by connectedness are called the connected components of X.

It follows that X is the disjunct union of its connected components. The name **connectedness** induces the image of having a path between points in a connected space. In topology this kind of connectedness is referred to as path-connectedness and is not necessarily the same as connectedness. To introduces path-connectedness we first need to define what we mean with a path in a topological space.

Definition 6.4.3 (Path and Loop in a topological space)**.**
Let X be a topological space. A path in X is a continuous map

$$\gamma : [0, 1] \to X.$$

$\gamma(0)$ is called the initial point and $\gamma(1)$ is called the terminal or endpoint of the path. If

$$\gamma(0) = \gamma(1)$$

then γ is called a loop.

This definition can then be used to define a path-connected topological space and a new equivalence relation introducing path-connected components, corresponding the naive image that every two points in a path-connected component can be connected by a path that completely lays within this component.

Definition 6.4.4 (Path-connected components)**.**
Let X be a topological space. The equivalence classes introduced by this equivalence relation are called the path-connected components of X.

Definition 6.4.5 (Path-connected)**.**
A topological space is called path-connected if every two points of X are equivalent under the above equivalence relation.

Naturally we have the following relation between path-connected and connected.

Corollary 6.4.1

A path-connected space is connected.

The concept of connectedness can be straightforwardly extended to more complicated spaces by using the product topology (definition 6.3.3).

Definition 6.4.6 (Connected products)**.**
Consider the topological spaces X and Y and let

$$X \times Y$$

have the product topology.

(i) If X and Y are connected then $X \times Y$ is also connected.

(ii) If X and Y are path-connected then $X \times Y$ is also path-connected.

As a consequence one also has

Corollary 6.4.2

Consider the topological spaces

$$X_1, \cdots, X_n.$$

(i) If every X_i is connected then

$$X_1 \times \cdots \times X_n$$

is also connected.

(ii) If every X_i is path-connected then

$$X_1 \times \cdots \times X_n$$

is also path-connected.

To demonstrate the relevance of these concepts in physics, we will give an example that is often used in physics courses for which we need the concept of **a topological group**.

Definition 6.4.7 (Topological Group)**.**
A topological group is a group G that has been augmented with a topology, such that the maps

$$\begin{aligned} &m : G \times G \to G, (g_1, g_2) \mapsto g_1 g_2 && \textit{Multiplication} \\ &i : G \to G,\ \ g \mapsto g^{-1} && \textit{Inverse} \end{aligned}$$

are continuous. Note that here the elements of the group are considered as the points of the topological space.

From this definition it is not so hard to prove that:

1. Left translations on G defined by

$$\forall a \in G, t_a : G \to G : t_a(g) = ag$$

   is a homeomorphism on G

2. A topological group G is Hausdorff if and only if the unit element $e \in G$ is a closed point.

3. Let $G^0 \subset G$ be the connected component of G containing the unit element e. Then G^0 is a subgroup of G.

4. If $H \subset G$ is a subgroup that is open, then H is also closed.

Let us now consider the following example.

Example 6.4.1. *Consider the group*

$$GL_2(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \; a, b, c, d \in \mathbb{R}, \; ad - bc \neq 0 \right\}. \tag{46}$$

We can consider $GL_2(\mathbb{R})$ as an open subset of $\mathbb{R}^4$ for the Eucledian topology, which then induces a topology on $GL_2(\mathbb{R})$. This is actually a topological group, moreover it has the structure of a C^∞-variety (a special type of manifold), making it an example of a Lie Group. This group cannot be connected. To see this consider the determinant map

$$\mathit{det} : GL_2(\mathbb{R}) \to \mathbb{R}^*$$

which is continuous and surjective while

$$\mathbb{R}^* = \mathbb{R} \setminus \{0\}$$

is not connected. By contraposition we thus have that the group cannot be connected.

6.5 LOCAL CONNECTEDNESS AND LOCAL PATH-CONNECTEDNES

This sections treats the local versions of the previous section, where locally refers to a neighborhood of some point of the topological space.

Definition 6.5.1 (Locally connected)**.**
A topological space X is called locally connected if

$$\forall x \in X$$

and for every open neighborhood U of x there exists a connected open neighborhood V of x such that

$$V \subseteq U.$$

In the same way we have for path-connected:

Definition 6.5.2 (Locally Path Connected)**.**
A topological space X is called locality path-connected if

$$\forall x \in X$$

and every open neighborhood U of x there exists a path-connected open neighborhood V of x such that

$$V \subseteq U.$$

The connection between connected and path-connected is stronger in the local versions, which is demonstrated by the following proposition:

Proposition 6.5.1

If X is locally path-connected then X is also locally connected.

As a warning it is important to keep in mind that if a space is **locally** (path-)connected it is not necessarily (path-)connected. Moreover the inverse is also not always true[2].

6.6 COMPACTNESS

Consider a topological space X and define an open cover of X as a collection

$$\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$$

of open subsets of X such that

$$X = \cup_{\alpha \in A} U_\alpha.$$

2 As a counter example consider the Dirac Comb.

Given such a cover $\mathcal{U}$ and a set

$$A' \subset A$$

this defines an open sub-cover $\mathcal{U}'$ if this itself is an open cover of X. Such covers will allow us to say when a topological space is compact or not, a property that will again show its usefulness when we are considering manifolds.

Definition 6.6.1 (Compact)**.**
A topological space X is called compact if **every** *open cover $\mathcal{U}$ of X (a collection of open sets of X whose union is all of X) has a* ***finite*** *sub-cover.*

Let us again give a familiar example from real analysis.

Example 6.6.1.
A closed interval $[a, b] \subset \mathbb{R}$ is compact in the Euclidean topology.

An alternative definition of compactness, continuity and closed can be given by using **nets**. We introduce them here since it will allow us to also introduce the Tychonov topology and Tychonoff's Theorem.

Definition 6.6.2 (Partially Ordered Set)**.**
A (non-strict) partial order is a binary relation $\leq$ over a set P which is reflexive, antisymmetric, and transitive, i.e., which satisfies

$$\forall a, b, c \in P :$$

(i) $a \leq a$ (reflexivity)

(ii) if $a \leq b$ and $b \leq a$ then $a = b$ (antisymmetry)

(iii) if $a \leq b$ and $b \leq c$ then $a \leq c$ (transitivity).

This is sometimes also called an **anti-symmetric preorder.**

Definition 6.6.3 (Directed Set)**.**
A directed set (or a **directed preorder** *or a* **filtered set***) is a nonempty set A together with a reflexive and transitive binary relation $\leq$ (that is,*

a preorder), with the additional property that every pair of elements has an upper bound:

$$\forall a, b \in A : \exists\, c \in A : a \leq c, b \leq c$$

Definition 6.6.4 (Net)**.**

(i) A net (x^α) in a topological space X is a map

$$\alpha \to x^\alpha$$

from a partially ordered and directed index set A (relation $\geq$) to X.

(ii) A net (x^α) converges to x, denoted

$$\lim_\alpha x^\alpha = x$$

if for every open neighborhood $U \subset X$ of x there exists $\alpha(U) \in A$ such that $x^\alpha \in U$ for every $\alpha \geq \alpha(U)$ (one says that (x^α) is eventually in U).

(iii) A subnet $\left(x^{\alpha(\beta)}\right)$ of a net (x^α) is defined through a map

$$B \to A;\ \beta \mapsto \alpha(\beta)$$

between partially ordered and directed index sets such that for any $\alpha_0 \in A$ there exists $\beta(\alpha_0) \in B$ with $\alpha(\beta) \geq \alpha_0$ for any $\beta \geq \beta(\alpha_0)$ (one says that B is **cofinal** *for A).*

(iv) A net (x^α) in a topological space X is called **universal** *if for any subnet*

$$Y \in X$$

the net (x^α) is eventually either only in Y or only in

$$X - Y.$$

Notice that for a subnet there is no relation between the index sets A, B except that $\alpha(B) \subset A$ such that in particular the subnet of a sequence $(A = \mathbb{N})$ may not be a sequence any longer. The notions of closedness, continuity and compactness can now be reformulated in terms of nets. The fact that one uses nets instead of sequences is that lemma 6.6.1 is no longer true when $A = \mathbb{N}$ unless we are dealing with metric spaces.

Lemma 6.6.1: Closed, Continuous and Compact using nets

(i) A subset Y of a toplogical space X is closed if for every convergent net (x^α) in X with

$$x^\alpha \in Y, \ \forall \alpha$$

the limit actually lies in Y.

(ii) A function

$$f: \ X \to Y$$

between topological spaces is continuous if for every convergent net (x^α) in X, the net $(f(x^\alpha))$ is convergent in Y.

(iii) A topological space X is compact if every net has a convergent subnet. The limit point of the convergent subnet is called a **cluster (accumulation) point** of the original net.

From the above it should be clear that if a net converges in some topology, then it also converges in any coarser (weaker) topology. Returning to compactness, we first give some properties, before continuing to describe the Tychonov topology .

Proposition 6.6.1

If

$$f : X \to Y$$

is a continuous map and X is compact, then also the image

$$f(X) \subseteq Y$$

is compact.

Proposition 6.6.2

If X is compact and $Z \subseteq X$ is closed in X then Z is compact.

Proposition 6.6.3

If X is Hausdorff and $Z \subseteq X$ is compact, then Z is closed in X.

Definition 6.6.5 (Tychonov).
The Tychonov topology on the direct product

$$X_\infty = \prod_{l \in \mathcal{L}} X_l$$

of topological spaces X_l, $\mathcal{L}$ any index set, is the weakest topology such that all the projections

$$p_l : \; X_\infty \to X_l, \; (x_{l'})_{l' \in \mathcal{L}} \mapsto x_l$$

are continuous, that is, a net

$$x^\alpha = (x_l^\alpha)_{l \in \mathcal{L}}$$

converges to

$$x = (x_l)_{l \in \mathcal{L}}$$

iff

$$x_l^\alpha \to x_l$$

for every $l \in \mathcal{L}$ point-wise (not necessarily uniformly) in $\mathcal{L}$. Equivalently, the sets:

$$p_l^{-1}(U_l) = \left[\prod_{l' \neq l} X_{l'} \right] \times U_l,$$

are defined to be open and form a basis for the topology of X_∞ (any open set can be obtained from those by finite intersections and arbitrary unions).

The definition of this topology is motivated by the following theorem.

Theorem 6.6.1: Tychonov

Let $\mathcal{L}$ be an index set of arbitrary cardinality and suppose that for each $l \in \mathcal{L}$ a compact topological space X_l is given. Then the direct product space

$$X_\infty = \prod_{l \in \mathcal{L}} X_l$$

is a compact topological space in the Tychonov topology.

This theorem has a nice proof using universal nets, for which we refer the reader to [96]. As a consequence we have the well known fact from real calculus

Corollary 6.6.1

A subset $Z \subseteq \mathbb{R}^n$ is compact if and only if Z is closed and bounded.

6.7 COUNTABILITY AXIOMS AND BAIRE

In an onset to get to the separation properties of topological spaces we need to consider the concept of **countability**.

Definition 6.7.1 (Neighborhood basis)**.**
Let X be a topological space and $x \in X$. Let

$$\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$$

a collection of open neighborhoods of x. Then $\mathcal{U}$ is a neighborhood basis of x if for every open neighborhood V of x there exists an α such that $U_\alpha \subseteq V$.

This lets us define the first countability axiom defining $A1$-spaces .

Definition 6.7.2 (A1)**.**
A topological space X obeys the first countability axiom if each $x \in X$

has a **countable** *neighborhood basis. Such a topological space is said to be $A1$.*

Note that every metric space is $A1$, which can be seen by considering open balls of radius $\frac{1}{N}$, $N \in \mathbb{N}$. A stronger version of this is called the second countability axiom returning $A2$-spaces.

Definition 6.7.3 (A2)**.**
A topological space X obeys the second countability axiom if there exists a a countable basis for the topology on X. Such a topological space is said to be $A2$.

Proposition 6.7.1

Let X be a topological space that is $A2$.

(i) Every open cover of X has a countable sub-cover. A space with such a property is called a **Lindelöf space**.

(ii) There is a countable subset of X that is dense (Definition 6.7.4) in X.

Important to note is that:

$$A2 \;\Rightarrow\; A1. \tag{47}$$

Definition 6.7.4 (Dense)**.**
A subset A of a topological space X is called dense (in X) if every point $x \in X$ either belongs to A or is a limit point of A (see definition 6.8.1). Formally, a subset A of a topological space X is dense in X if for any point $x \in X$, any neighborhood of x contains at least one point from A. Put differently, A is dense in X if and only if the only closed subset of X containing A is X itself. This can also be expressed by saying that the **closure** *$\overline{A}$ of A is X, or that the interior of the complement of A is empty.*

Definition 6.7.5 (Meagre subset - First Baire Category)**.**
Let X be a topological space. We say that a subset $V \subseteq X$ is nowhere dense if the interior of the closure of V is empty. We call V **meagre** *if V is a countable union of nowhere dense subsets.*

Despite their name meagre subset have some nice properties.

Proposition 6.7.2

Let X be a topological space.

(i) A subset $V \subseteq X$ is nowhere dense if and only if the interior of the complement $X \setminus V$ is dense in X.

(ii) A finite union of nowhere dense subsets is again nowhere dense.

(iii) A countable union of meagre subsets is again meagre.

Lemma 6.7.1

The following properties of a topological space X are equivalent.

(i) Every countable intersection of dense open sets is again dense in X.

(ii) If $C_1, C_2, \cdots$ are closed subsets of X with empty interiors, then also their union $\bigcup_{i=1}^{\infty} C_i$ has an empty interior.

(iii) If $U \subseteq X$ is a non-empty open subset, then U is not meagre.

(iv) If $V \subseteq X$ is a meagre subset, then the complement $X \setminus V$ is dense in X.

With **interior of a set** C we mean the points $x \in C$ that have an open neighborhood $x \in U$ such that $U \subset C$. Spaces that have the above properties are called **Baire spaces**, and with them comes a theorem

Theorem 6.7.1: Baire Category Theorem

Every compact Hausdorff space is a Baire space.

Baire referred to a meagre subset as a subset of the First Baire Category and to a non-meagre subset as being of the Second Baire Category. Then a Baire space is a space where every non-empty set is of the second category. Note that here category has nothing to do with Category-Theory, we just mention this old terminology since it is sometimes used in older literature.

6.8 CONVERGENCE

We now define what it means for sequences in a topological space to converge. This will allow us to give meaning to limits of sequences.

Definition 6.8.1 (Convergence and Accumulation point)**.**
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of elements in a topological space X and let $\xi \in X$.

(i) We say a sequence (x_n) converges to ξ, or that ξ is the limit of the sequence (x_n), if for every open neighborhood U of ξ there is an index $N(U)$ such that

$$x_n \in U, \forall n \geq N(U).$$

A sequence is called convergent if it has a limit.

(ii) We call ξ an accumulation point of the sequence (x_n) if for every open neighborhood U of ξ there exist an infinite number of indices n such that $x_n \in U$.

One might expect that this is enough to state that if a sequence converges that it has a unique limit, but this is not true! To be able to make this statement the space also needs to be Hausdorff, where a sequence can have at most one limit.

Definition 6.8.2 (Countable compact)**.**
A topological space X is called countable compact if every countable open cover

$$X = \bigcup_{\alpha \in A} U_\alpha$$

has a finite sub-cover.

Notice the difference with compactness, where every cover needs a finite sub-cover and not only the countable ones. Moreover we have that if X compact then it is also countable compact, but the inverse is not always true.

Proposition 6.8.1

A topological space X is countable compact if and only if each sequence $(x_n)_{n\in\mathbb{N}}$ has an accumulation point.

Definition 6.8.3 (Sequentially compact)**.**
A topological space X is called sequentially compact if every sequence in X has a converging sub-sequence.

Lemma 6.8.1

Let X be a topological space that is $A1$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Let $\xi \in X$. Then the following two statements are equivalent.

(i) The sequence (x_n) has a subsequence that converges to ξ.

(ii) ξ is an accumulation point of the sequence (x_n).

Theorem 6.8.1

Let X be a topological space.

(i) If X is sequentially compact, then X is also countable compact.

(ii) If X is countable compact and $A1$ then X is also sequentially compact.

(iii) If X is countable compact and $A2$ then X is also compact.

Graphically this is represented as:

$$\text{compact} \underset{\Longleftarrow}{\overset{+A2}{\Longrightarrow}} \text{countable compact} \underset{\Longleftarrow}{\overset{+A1}{\Longrightarrow}} \text{sequentially compact}$$

Definition 6.8.4 (Cauchy sequence and Complete space).
Let (X, d) be a metric space.

(i) A sequence $(x_n)_{n\in\mathbb{N}}$ in X is called **Cauchy** *if $\forall\epsilon > 0$ there exists an index N such that*

$$d(x_m, x_n) < \epsilon, \ \forall m, n \geq N.$$

(ii) This metric space is called **complete** *if every Cauchy sequence in X is convergent.*

As a consequence, if a Cauchy sequence has a convergent sub-sequence than the original sequence is convergent. We also have that every metric space can be completed.

Definition 6.8.5 (Totally bounded).
A metric space (X, d) is called totally bounded if there exists $\forall\epsilon > 0$ a finite cover of X with open balls of radius ϵ.

Lemma 6.8.2

A totally bounded metric space (X, d) is $A2$.

Theorem 6.8.2

If X is a metric space, then X is compact for the metric induced topology if and only if X is complete and totally bounded.

Definition 6.8.6 (Uniform Continuity)**.**
Let (X, d_X) *and* (Y, d_Y) *be metric spaces. A map*

$$f : X \to Y$$

is uniformly continuous if

$$\forall \epsilon > 0, \exists \delta > 0$$

such that

$$\forall x, x' \in X$$

with

$$d_X(x, x') < \delta$$

we have that

$$d_Y(f(x), f(x')) < \epsilon.$$

Theorem 6.8.3

Let (X, d_X) and (Y, d_Y) be metric spaces. If X is compact, then every continuous map

$$f : X \to Y$$

is uniformly continuous.

6.9 SEPARATION PROPERTIES

We begin by stating the separation axioms. Roughly speaking they classify which basic objects can be seperated in a topological space.

Definition 6.9.1 (Seperation Axioms)**.**
Let X *be a topological space. We call* X

(i) T_1 *: if all 1-point sets* $\{x\}$ *are closed in* X*.*

(ii) T_2 *: if* X *is Hausdorff*

(iii) T_3 *: if for every point* $x \in X$ *and for every closed subset* $C \subset X$ *with* $x \notin C$*, there exist open neighborhoods* U *of* x *and* V *of* C *such that* $U \cap V = \emptyset$.

(iv) T_4 *: if for every couple of closed set* $C, D \subset X$ *with* $C \cap D = \emptyset$ *there exist open neighborhoods* U *of* C *and* V *of* D *such that* $U \cap V = \emptyset$.

The T refers to the German "Trennung".

Depending on which (combinations) of the axioms are valid for a topological space they acquire different designations.

Definition 6.9.2 (Regular and Normal)**.**
A topological space X *is called regular if it is* T_1 *and* T_3*. If it is* T_2 *and* T_4 *we call it normal.*

Certainly it is clear that some of the axioms induce others, we have that

Lemma 6.9.1

$$(T_4 + T_1) \implies (T_3 + T_1) \implies T_2 \implies T_1$$

or in words

$$\text{normal} \implies \text{regular} \implies \text{Hausdorff} \implies T_1.$$

Proposition 6.9.1

(i) We have that every metric space X is normal.

$$T_1 + T_3 + A2 \implies T_1 + T_4$$

(ii) If X is compact and Hausdorff, then X is normal.

Lemma 6.9.2: Urysohn

Let X be normal. If A and B are disjunct subsets of X, then there exists a continuous map

$$f : X \to \mathbb{R}$$

with

$$f(a) = 0, \ \forall a \in A$$

and

$$f(b) = 1, \ \forall b \in B.$$

Theorem 6.9.1: Tietze

Let X be a normal space and C a closed subset of X. Assume

$$f_C : C \to \mathbb{R}$$

is given. Then there exists a continuous function

$$f : X \to \mathbb{R}$$

with

$$f|_C = f_C.$$

Theorem 6.9.2: Urysohn's metrizability theorem

If X is a regular space that is $A2$ then X is metrizable.

In other words one can define a metric on X such that it induces a topology on X.

6.10 LOCAL COMPACTNESS AND COMPACTIFICATION

This section will become important when we consider Wilson lines that go out to infinity on the space-time manifold. Since infinity is actually not part of the space-time manifold one needs to add it (compactification). This can be done in different ways (see below) and if one adds more then one point at infinity the question arises to which infinity is the Wilson line going? A problem not solved today.

Definition 6.10.1.
Let A be a subset of a topological space X. Then the subset $B \subseteq X$ is called a neighborhood of A if A is contained in the interior of B.

Definition 6.10.2.
A topological space X is called locally compact if $\forall x \in X$ there is a compact neighborhood.

As discussed we want to compactify a topological space. The most economical way of doing this for a locally compact Hausdorff space is the following compactification

Theorem 6.10.1: Alexandroff Compactification

Let X be a locally compact Hausdorff space. Then there exists a compact Hausdorff space X^* and a point $P \in X^*$ such that X is homeomorphic with $X^* \backslash P$. Moreover the pair (X^*, P) is unique up to homeomorphism in the following sense, assume that we have compact Hausdorff spaces X_1^* and X_2^*, together with points $P_i \in X_i^*$ and homeomorphisms

$$f_1 : X^* \to X_1^* \setminus \{P_1\}$$

and

$$f_2 : X^* \to X_2^* \setminus \{P_2\}.$$

Then there is a unique homeomorphism

$$g : X_1^* \to X_2^*$$

with

$$g(P_1) = P_2 \text{ and } g \circ f_1 = f_2.$$

This compactification method is sometimes also called **one-point compactification**, because we add one point at infinity. Note that this adds an extra symmetry to the space. This is best seen in two dimensions. Adding one point at infinity turns this "plane" into a Riemann sphere, a projective space that has conformal symmetry. This simple example demonstrates clearly that one has to be very careful when applying compactifications, if one is not careful one might introduce extra structure that is not necessarily wanted. An alternative compactification method is given by Stone-Cech compactification, which we will not discuss here because it is much more involved.

6.11 QUOTIENT TOPOLOGY

When one introduces an equivalence relationship on a topological space naturally the question arises if the set of equivalence classes also has a topological structure. We will explain below that they indeed have such a structure, the quotient topology.

We start by investigating the connection between the concepts

(i) Equivalence relations on a set X.

(ii) Partitions of a set X.

(iii) Surjective maps $X \twoheadrightarrow Y$.

If we have an equivalence relation $\sim$ on X then the equivalence classes form a partition of X. On the other hand if we have a partition of X we can introduce an equivalence relation by stating that two elements $x, y \in X$ are equivalent if they are in the same subset of the partition. We conclude that we have a bijection

$$\text{equivalence relations on } X \sim \text{partitions of } X.$$

To see the relation with surjective maps, assume that we have an equivalence relation on X and write $X/\sim$ for the set of equivalence classes. Then we clearly have a map

$$q : X \twoheadrightarrow X/\sim$$

taking an element $x \in X$ to its equivalence class. We also refer to this map q as dividing out the equivalence relation. Inversely we get an equivalence relation on X from a surjective map

$$f : X \twoheadrightarrow Y$$

by calling two elements $x, y \in X$ equivalent if $f(x) = f(y)$. The fibers[3] of f are thus per definition the equivalence classes.

Definition 6.11.1 (Quotient Topology).
Let X be a topological space and let $\sim$ be an equivalence relation on X. We then define the quotient topology on $X/\sim$ as the weakest topology for which the map

$$q : X \to X/\sim$$

is continuous.

Definition 6.11.2 (Quotient map).

(i) Let X, Y be topological spaces and

$$p : \ X \to Y$$

a surjection. The map p is said to be a quotient map provided that $V \subset Y$ is open in Y if and only if $p^{-1}(V)$ is open in X.

(ii) If X is a topological space, Y a set and

$$p : \ X \to Y$$

a surjection then there exists a unique topology on Y with respect to which p is a quotient map.

(iii) Let X be a topological space and let $[X]$ be a partition of X (i.e. a collection of mutually disjoint subsets of X whose union is X). Denote by

$$[x], \ x \in X$$

the subset of X in that partition of X which contains x. Equip $[X]$ with the quotient topology induced by the map

$$[] : \ X \to [X]; \ x \mapsto [x].$$

3 Notice here the reference to fibers in the context of equivalence relations. This will resurface when we are discussing fibre bundle theory in the context of gauge theory, where we will have classes of physically equivalent theories.

Then $[X]$ is called the quotient space of X.

Notice that the requirement for p to be a quotient map is stronger than just being continuous which would only require that $p^{-1}(V)$ is open in X whenever V is open in Y (but not vice versa). From a geometrical point of view taking quotients is gluing topological spaces.

Quotient spaces naturally arise if we have a group action

$$\lambda: \ G \times X \to X; \ (g, x) \to \lambda_g(x) := \lambda(g, x)$$

on a topological space X and define

$$[x] := \{\lambda_g(x); \ g \in G\}$$

to be the orbit of x. The orbits clearly define a partition of X.

Lemma 6.11.1

Let X be a compact topological space, Y a set and

$$p: \ X \to Y$$

a surjection. Then Y is compact in the quotient topology.

Lemma 6.11.2: Hausdorff in quotient topology

Let X be a Hausdorff space and

$$\lambda: \ G \times X \to X$$

a continuous group action on X (i.e., λ_g defined by $\lambda_g(x) := \lambda(g, x)$ is continuous for any $g \in G$). Then the quotient space

$$X/G := \{[x]; \ x \in X\}$$

defined by the orbits

$$[x] = \{\lambda_g(x); \ g \in G\}$$

is Hausdorff in the quotient topology.

Theorem 6.11.1: Equivariance

Let X, Y be topological spaces and let G be a group acting (not necessarily continuously) on them via λ, λ' respectively. If

$$f : X \to Y$$

is a homeomorphism with respect to which the actions λ, λ' are equivariant (i.e. they commute with the group action) then f extends as a homeomorphism to the quotient spaces $X/G, Y/G$ in their respective quotient topologies.

6.12 FUNDAMENTAL GROUP

As a last subject in the topology introduction we would like to discuss the fundamental group, which is of course very related to loops in a manifold. Before we start we introduce the notation

$$I := [0, 1]$$

for the rest of this chapter unless stated otherwise.

Definition 6.12.1 (Homotopy).
Let

$$f_0, f_1 : X \to Y$$

be two continuous maps between topological spaces. A homotopy from f_0 to f_1 is a continuous map

$$F : X \times I \to Y$$

such that

$$F(x, 0) = f_0(x) \text{ and } F(x, 1) = f_1(x).$$

If such a map exists we call f_0 and f_1 homotopy equivalent

$$f_0 \simeq f_1.$$

This means that there is a continuous deformation between the two functions. Note that a homotopy also provides a set of intermediate functions, which are relevant in physics in the study of renormalization flows and in relating vacuum expectation values or vacua in different frames or with different Hamiltonians.

Lemma 6.12.1

Homotopy is an equivalence relation on the set $C(X, Y)$ of continuous maps

$$X \to Y.$$

We can extend the restrictions on homotopies by introducing extra conditions.

Definition 6.12.2 (Relative homotopy)**.**
Let X, Y be topological spaces and $A \subseteq X$. Consider the two continuous maps

$$f_0, f_1 : X \to Y$$

with

$$(f_0)|_A = (f_1)|_A.$$

Then a homotopy from f_0 to f_1 relative A is a continuous map

$$F : X \times I \to Y$$

such that

$$F(x, 0) = f_0(x) \text{ and } F(x, 1) = f_1(x)$$

for all $x \in X$ with the extra condition that also

$$F(a, t) = f_0(a), \forall a \in A, \forall t \in I.$$

Applying the definition of relative homotopy for loops with a fixed base point (thus the set A is a single point) in the manifold we have that two loops

$$\gamma_0, \gamma_1 : I \to X$$

are homotopy equivalent if there exists a homotopy

$$F : I \times I \to X$$

relative $\{0, 1\}$. This lead to the definition of **the fundamental group**.

Definition 6.12.3 (Fundamental Group)**.**
If X is a topological space and $x_0 \in X$ a base point, then we define:

$$\pi_1(X, x_0) := \{\textit{homotopy classes of loops in } X \textit{ with base point } x_0.\}$$

We call this the fundamental group of X with base point x_0.

The group operation is defined as the composition of loops and can be proven to define a group structure on the homotopy equivalence classes. Moreover it can be proven that

Proposition 6.12.1: Independent of the base point.

Let X be a path-connected space and let

$$x_0, x_0' \in X,$$

then the groups $\pi_1(X, x_0)$ and $\pi_1(X, x_0')$ are isomorphic.

We now investigate what happens to the fundamental groups if we consider maps between topological spaces. Suppose we have a topological space X with base point x_0 and a continuous map

$$f : X \to Y,$$

with Y another topological space with

$$Y \ni y_0 = f(x_0).$$

If now γ is a loop in X at x_0 then

$$f \circ \gamma : I \to Y$$

is a loop in Y at y_0. Assume γ homotopy equivalent to γ' then we also have that $f \circ \gamma$ is homotopy equivalent to $f \circ \gamma'$. Thus we have a map f_* defined as

$$f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0) \quad [\gamma] \to [f \circ \gamma],$$

a map between fundamental groups.

Lemma 6.12.2

Let X be a topological space with $x_0 \in X$ a base point. Consider the continuous maps

$$f : X \to Y$$

and

$$g : Y \to Z,$$

with

$$y_0 := f(x_0)$$

and

$$z_0 := g(y_0).$$

Then we have that:

$$(g \circ f)_* = g_* \circ f_* : \pi_1(X, x_0) \to \pi_1(Z, z_0).$$

Corollary 6.12.1

Let X and Y be path-connected spaces that are homeomorphic, then

$$\pi_1(X, x_0) \cong \pi_1(Y, y_0)$$

for every choice of base points x_0 and y_0.

The fact that the corollary 6.12.1 can be taken literally is a consequence of the following theorem.

Theorem 6.12.1

Let X be a topological space with $x_0 \in X$ a base point and

$$f, g : X \to Y$$

continuous maps that are homotopic. Assume

$$y_0 := f(x_0)$$

and

$$y_1 := g(x_0).$$

Chose a homotopy

$$F : X \times I \to Y$$

from f to g, and consider the path α in Y from y_0 to y_1 given by

$$\alpha(t) := F(x_0, t).$$

If

$$a : \pi_1(Y, y_1) \stackrel{\cong}{\to} \pi_1(Y, y_0),$$

is the isomorphism given by

$$[\gamma] \to [\alpha^{-1}\gamma\alpha]$$

then the homomorphisms

$$a \circ g_* : \pi_1(X, x_0) \to \pi_1(Y, y_0) \text{ and } f_* : \pi_1(X, x_0) \to \pi_1(Y, y_1)$$

are equal.

This allows us to define a quite strong relationship between topological spaces, and thus perform a classification according to homotopy equivalence.

Definition 6.12.4 (Homotopy equivalence of spaces.)**.**
A continuous map

$$f : X \to Y$$

is called a homotopy equivalence if there exists a continuous map

$$g : Y \to X$$

such that $g \circ f$ *is homotopic to* id_X *and* $f \circ g$ *to* id_Y*. Two spaces are called homotopy equivalent if there is a homotopy equivalence between them.*

Note that homotopy equivalent is stronger than homeomorphic, and is closer to the notion of topological spaces as rubber objects.

Corollary 6.12.2

Let X, Y be path-connected spaces. If they are also homotopy equivalent the fundamental groups are isomorphic:

$$\pi_1(X, x_0) \cong \pi_1(Y, y_0),$$

for every choice of the base points x_0 and y_0.

Definition 6.12.5 (Simply connected)**.**
A path-connected space X is called simply connected if

$$\pi_1(X, x_0) = 0$$

for an $x \in X$.

Definition 6.12.6 (Contractible)**.**
A topological space X is called contractible if and only if it is homotopy equivalent to a point.

With these two last definitions we end the part on topology, now fully prepared to move on to the differential geometry part.

7 DIFFERENTIAL GEOMETRY

In this chapter we will review some of the basic definitions of differential geometry necessary to understand the discussion on gauge theory from a principal fibre point of view and to extend the manifold concept to more complicated spaces eventually allowing for the construction of GLS in later chapters.

7.1 MANIFOLDS

We start be reviewing some of the definitions of manifolds.

Definition 7.1.1 (Manifolds)**.**

(i) *A topological space M is called an m-dimensional C^k manifold provided there is a family of pairs $(U_I, x_I)_{I\in\mathcal{I}}$ consisting of an open cover of M and homeomorphisms*

$$x_I : U_I \to x_I(U_I) \subset \mathbb{R}^m,\ p \mapsto x_I(p),$$

such that

$$\forall I, J \in \mathcal{I}$$

with

$$U_I \cap U_J \neq \emptyset$$

the map

$$\phi_{IJ} := x_J \circ x_I^{-1} :\ x_I(U_I \cap U_J) \to x_J(U_I \cap U_J),$$

is a C^k map between open subsets I of $\mathbb{R}^m$.

(ii) *The sets U_I are called charts, the functions x_I coordinates, the family of charts and coordinates form an atlas and m refers to the dimension of M. Two atlases $(U_I, x_I)_{I\in\mathcal{I}}, (V_I, x_J)_{J\in\mathcal{J}}$ for a topological space M are compatible if their union is again an atlas.*

Compatibility defines an equivalence relation on atlases and such an equivalence class is called a differentiable C^k structure.

(iii) A topological space M is said to be a manifold with a boundary ∂M if each of the U_I is homeomorphic to an open subset of the negative half-space

$$H_- = \{x \in \mathbb{R}^m; x^1 \leq 0\}.$$

The smoothness condition now demands that the ϕ_{IJ} are C^k on open subsets of $\mathbb{R}^m$ containing $x_I(U_I \cap U_J)$. The boundary points have coordinates $x^1 = 0$, that is, they lie in

$$\partial H_- = \{x \in \mathbb{R}^m; x^1 = 0\}.$$

Definition 7.1.2 (Diffeomorphism)**.**
A map

$$\psi : M \to N$$

between C^k manifolds M, N is called C^k if for all pairs of charts U_I, V_J of atlases for M, N respectively for which

$$\psi(U_I) \cap V_J \neq \emptyset$$

the maps (where defined)

$$\psi_{IJ} := x_J \circ \psi \circ x_I^{-1} : x_I(U_I) \to x_J(V_J)$$

are C^k maps between open subsets of $\mathbb{R}^m, \mathbb{R}^n$ respectively. If all the ψ_{IJ} are invertible and also the inverses are C^k then ψ is called a C^k diffeomorphism. The diffeomorphisms of a manifold form a group which is denoted Diff(M).

Definition 7.1.3 (Paracompact)**.**
An atlas (U_I, x_I) is said to be locally finite provided that every $p \in M$ has an open neighbourhood in M intersecting only a finite number of the charts. A manifold M is called paracompact if each atlas (U_I, x_I)

admits a locally finite refinement (V_J, y_J) where each V_J is contained in some U_I .

Definition 7.1.4 (Sub-manifold)**.**
Let N be a subset of an m-dimensional manifold M. We can equip N with a manifold structure by making use of the induced topology and an induced (subspace) differentiable structure, given an atlas (U_I, x_I) for M, by the atlas $(V_I = N \cap U_I, y_I = (x_I)|_{V_I})$ for N. We thus have a differentiable structure under the condition that the maps

$$\phi_{IJ} = y_J \circ y_I^{-1}$$

for

$$V_I \cap V_J \neq \emptyset$$

have constant rank n.

Definition 7.1.5 (Immersion and embedding)**.**
Assume N an n-dimensional manifold and

$$\psi : N \to M$$

*is C^k. ψ is called **a local immersion** if each $q \in N$ has an open neighborhood V such that*

$$V \to \psi(V)$$

*is an **injection**. If ψ is a **global immersion**, that is,*

$$N \to \psi(N)$$

*is an injection (the image of N in M does not intersect itself), then ψ is called an **embedding**. If additionally for each V open in N the set $\psi(V)$ is open in the subset topology induced from M then ψ is called a **regular embedding** (the image of N does not come arbitrarily close to* *itself in M without ever self-intersecting). In the latter case we will say that N is an embedded sub-manifold of M. An embedded sub-manifold of dimension $n = m - 1$ is called a **hyper-surface**.*

Definition 7.1.6 (Orientability)**.**
A manifold M is said to be orientable if it admits an atlas such that:

$$\det\left(\frac{\partial \mathrm{x_J(p)}}{\partial \mathrm{x_I(p)}}\right) > 0, \ \forall \mathrm{p} \in \mathrm{U_I} \cap \mathrm{U_J}.$$

If M has a boundary then M induces an orientation on ∂M.

Definition 7.1.7 (Smooth and Real Analytic)**.**
*A manifold is called **smooth** if it is C^∞. A manifold is called **real analytic** or C^ω if the maps ϕ_{IJ} are real analytic.*

*A manifold of real dimension $2m$ is called **complex analytic or a holomorphic**, a manifold of complex dimension m, provided that the maps*

$$\phi_{IJ} = z_J \circ z_I^{-1} : \mathbb{C}^m \to \mathbb{C}^m$$

*satisfy the **Cauchy-Riemann** equations and where*

$$(x_I, y_I) \to z_I = x_I + i y_I$$

is the standard isomorphism between $\mathbb{R}^{2m}$ and $\mathbb{C}^m$.

Comparing homeomorphisms and diffeomorphisms of a topological space it can be shown that for $m < 4$ all homeomorphisms are also diffeomorphisms and for $m \geq 4$ this is not always the case. In that respect we mention that $\mathbb{R}^4$ has an infinite number of distinct differentiable structures while S^7 only has 28. Another interesting observation is that one can show that any smooth, paracompact manifold admits an analytic structure which is **unique** up to smooth diffeomorphisms. As we will discuss further in this text Wilson loops are constructed from integrals, so we better make sure the manifold we consider allows for an **integrable structure**, which is provided by the para-compactness property. Moreover a connected, finite-dimensional, Hausdorff manifold is para-compact if and only if it has a countable base.

7.2 DIFFERENTIAL CALCULUS

There are several differential objects that can be defined on a manifold. Here we will repeat the definition of some of these objects, restricting to only the most relevant for our objectives.

Definition 7.2.1 (Smooth function)**.**
A smooth function on a manifold M is a map

$$f : M \to \mathbb{C}$$

such that

$$f \circ x_I^{-1}$$

is smooth on

$$x_I(U_I) \subset \mathbb{R}^m.$$

The set of smooth functions $C^\infty(M)$ forms an Abelian $*$-**algebra** *(Definition 9.1.3) where operations are defined point-wise and the* **involution** *is given by complex conjugation.*

Definition 7.2.2 (Vector Field)**.**
A smooth vector field on M is a **derivation** *on $C^\infty(M)$. That is, it is a linear map*

$$v : C^\infty(M) \to C^\infty(M); f \mapsto v(f)$$

that obeys ***Leibniz***

$$v(fg) = v(f)g + fv(g),$$

and annihilates constants. Given an atlas (U_I, x_I) we can define special vector fields ∂_μ^I on U_I defined by the condition:

$$\partial_\mu^I(x_I^\nu) = \delta_\mu^\nu$$

for $p \in U_I$ where

$$x(p) = (x_1(p), \cdots, x_m(p)) \in \mathbb{R}^m.$$

This allows us to write a vector field v as

$$v(p) = v_I^\mu[x_I(p)]\partial_\mu^I(p),$$

where we assumed the Einstein summation convention. The Leibniz Rule induces the ***chain rule*** *such that we have:*

$$v_I^\mu[x_I(p)]\partial_\mu^I(p) = v_J^\mu[x_J(p)]\partial_\mu^J(p),$$

if

$$p \in U_I \cap U_J, x_J(p) = \phi_{IJ}(x_I(p)).$$

With this definition we can investigate the action of a vector field on a smooth function

$$v(f) = v_I^\mu \partial_\mu[f_I(x)]\big|_{x=x_I(p)},$$

with

$$f_I = f \circ x_I^{-1}.$$

The space of smooth vector fields on M is written as $T^1(M)$.

Definition 7.2.3 (Contra-variant Vector)**.**
A tangent vector or contra-variant vector, sometimes also just called a vector at $p \in M$, written symbolically as X, assigns to each coordinate patch $p \in (U, x)$ an n-tuple of real numbers

$$(X_U^i) = (X_U^1, ..., X_U^n)$$

such that if $p \in U \cap V$, then the **coefficients** *of a contra-variant transform as*

$$X_V^i = \sum_j \left(\frac{\partial x_V^i}{\partial x_U^j}(p) X_U^j\right) \tag{48}$$

$$X_V = c_{VU} X_U, \tag{49}$$

where c_{VU} is called **the transition function** *(Jacobian matrix). In a local coordinate system tangent vectors can be written as first-order differential operators*

$$\boldsymbol{X}_p = \sum_j X^j \frac{\partial}{\partial x^j}\Big|_p$$

Definition 7.2.4 (Linear Functional - Covector).
A (real) linear functional α on a vector space E is a real valued function

$$\alpha : E \to \mathbb{R}$$

from E to the one-dimensional vector space $\mathbb{R}$. We thus have

$$\alpha(a\boldsymbol{v} + b\boldsymbol{w}) = a\alpha(\boldsymbol{v}) + b\alpha(\boldsymbol{w}), \text{ Linearity}$$

for the real numbers a, b and vectors $\boldsymbol{v}, \boldsymbol{w}$. Such a linear functional is also called a covector or covariant vector or one-form. After introducing local coordinates and choosing a basis it can be expanded as

$$\alpha = \sum_j a_j(x)\; dx^j,$$

and it transforms under a coordinate change as

$$dx_V^i = \sum_j \left(\frac{\partial x_V^i}{\partial x_U^j}\right) dx_u^j,$$

such that its coefficients transform as

$$a_i^V = \sum_j a_j^U \left(\frac{\partial x_U^j}{\partial x_V^i}\right).$$

Notice the difference in the transformation rule with the transformation rule for contra-variant vectors! This difference generalizes to tensors, such that the transformation rule can be used to determine its contra-variant and covariant rank. In some literature these transformation rules are used to introduce objects that are then referred to as tensors.

Definition 7.2.5 (Dual Space).
The collection of all linear functionals α on a ***vector space*** *E form a new vector space E^*,* ***the dual space to*** *E, under the operations*

$$(\alpha + \beta)(\boldsymbol{v}) \;:=\; \alpha(\boldsymbol{v}) + \beta(\boldsymbol{w}), \quad a, \beta \in E^*,\ \boldsymbol{v} \in E \tag{50}$$

$$(c\alpha)(\boldsymbol{v}) \;:=\; c\alpha(\boldsymbol{v}), \quad c \in \mathbb{R} \tag{51}$$

Definition 7.2.6 (Tangent Bundle)**.**
The tangent bundle TM to a **differentiable** *manifold M is by definition the collection of all tangent vectors at all points of M. Note that if M is n-dimensional, then TM is 2n-dimensional and the differential refers to the differentiability of the atlas transition maps.*

Definition 7.2.7 (Interior Product)**.**
If v is a vector and α is a p-form, their interior product $(p-1)$-form $i_v\alpha$ is defined by

$$\begin{aligned} i_v\alpha^0 &= 0 && \text{if } \alpha \text{ is a } 0\text{-form} \\ i_v\alpha^1 &= \alpha(v) && \text{if } \alpha \text{ is a } 1\text{-form} \\ i_v\alpha^p(w_2, \cdots, w_p) &= \alpha^p(v, w_2, \cdots, w_p) && \text{if } \alpha \text{ is a } p\text{-form} \end{aligned}$$

Obviously we have that

$$i_{v+w} = i_v + i_w$$

and

$$i_{av} = ai_v$$

Definition 7.2.8 (Exterior product and Exterior Algebra)**.**
The exterior algebra $\bigwedge(V)$ over a vector space V over a field k is defined as the quotient algebra of the tensor algebra $T(V)$ by the two-sided ideal I (Definition 8.2.1) generated by all elements of the form $x \otimes x$ such that $x \in V$

$$\bigwedge(V) := T(V)/I.$$

The exterior product $\wedge$ of two elements of $\bigwedge(V)$ is defined by

$$\alpha \wedge \beta = \alpha \otimes \beta / I.$$

or written differently:

$$a \wedge \beta = \alpha \otimes \beta - \beta \otimes \alpha.$$

The algebra associated with this product is called the exterior algebra on M (notation: $\bigwedge^n M$) and is constructed from the vector space of one-forms on M.

Theorem 7.2.1: Interior product is an Anti-Derivation

$$i_v : \bigwedge^p \to \bigwedge^{p-1}$$

is an anti-derivation, that is

$$i_v(\alpha^p \wedge \beta^q) = [i_v\alpha^p] \wedge \beta^q + (-1)^p \alpha^p \wedge [i_v\beta^q]$$

Definition 7.2.9 (Differential of a map)**.**
Let

$$\phi : M \to N$$

be a smooth map of manifolds and $\phi(x) = y$. The differential ϕ_* is then the map between the tangent spaces $\phi_* : T_xM \to T_yN$ defined by:

$$\phi_*(v_x) = w_y,$$

where $v_x \in T_xM$ and $w_y \in T_yN$, elements of the respective tangent spaces at x and y.

Definition 7.2.10 (Pull-back)**.**
Let $\phi : M \to N$ be a smooth map of manifolds and let $\phi(x) = y$. Let

$$\phi_* : T_xM \to T_yN$$

be the differential of ϕ. The pull-back ϕ^* is the linear transformation taking covectors at y into covectors at x, $\phi^* : N^*(y) \to M^*(x)$ defined by

$$\phi^*(\beta)(v) := \beta(\phi_*(v)),$$

for all covectors β at y and vectors v at x.

Definition 7.2.11 (Local Section)**.**
A local section of P is a smooth map

$$s_I : U_I \to P$$

such that

$$\pi \circ s_I = id_{U_I},$$

where π is the projection on in P. A cross section is a global section, that is, defined everywhere on M.

Definition 7.2.12 (Push-Forward)**.**
Let ϕ be a smooth map

$$\phi : M \to N$$

and a let X be a vector field on M, it is not usually possible to define a push-forward of X by ϕ as a vector field on N. A demonstration of this is the example where the map ϕ is not surjective, then there is no natural way to define such a push-forward outside of the image of ϕ. Nevertheless, one can make this difficulty precise, using the notion of a vector field along a map.

A section (7.2.11) of ϕ^*TN over M is called a vector field along ϕ. Assume X to be a vector field on M, i.e., a section of TM. Then, applying the differential (7.2.9) pointwise to X yields the push-forward $\phi_* X$, which is a vector field along ϕ, i.e., a section of ϕ^*TN over M.

Any vector field Y on N defines a pullback section of ϕ^*TN with

$$(\phi^* Y)_x = Y_{\phi x}.$$

A vector field X on M and a vector field Y on N are said to be ϕ-related if

$$\phi_* X = \phi^* Y$$

as vector fields along ϕ. In other words

$$\forall x \in M, d\phi_x(X) = Y_{\phi(x)}.$$

Using the above introduced language we can give an alternate definition for an immersion.

Definition 7.2.13 (Immersion).
A smooth map of manifolds $\phi : M \to N$ *is an immersion and* $\phi(M)$ *is an immersed sub-manifold provided*

$$\phi_* : T_x M \to T_{\phi(x)} N,$$

is $1:1$ *or said otherwise* $\ker\ \phi_* = 0$ *at each* $x \in M$.

Definition 7.2.14 (Support (of a function)).
Let

$$f : M \to \mathbb{R}$$

be a real-valued continuous function. We thus have by the definition of continuity that the inverse image of every open set of $\mathbb{R}$ *is open in* M*. Now the set of non-zero real numbers form an open subset of* $\mathbb{R}$*, such that the subset of* M *where* $f \neq 0$ *is an open subset of* M : $f^{-1}(\mathbb{R} - 0)$*. The closure of this set is called the support of* f*. This definition can be extended to tensor fields on* M.

Definition 7.2.15 (Bump Function).
Given a point $p \in M$*, with* n *the dimension of* M*, we can easily construct an n-form with its support contained in an open* ϵ*-ball around the point* p.

$$\begin{aligned} \omega^n &:= f(\|x\|) dx^1 \wedge \cdots \wedge dx^n, \quad \text{for } x \text{ in the ball } \|x\| \leq \epsilon \\ \omega^n &:= 0, \quad \text{outside of the ball} \end{aligned}$$

This n-form is called a bump form or bump function if $n = 0$.

Definition 7.2.16 (Partition of Unity).
Let M *be a manifold of dimension* n *that can be covered by a finite number of coordinate patches (generalization is possible but one needs to take care). Given this covering* $\{U_\alpha\}$ *a partition of unity subordinate to this covering will return* n *real-valued differentiable functions*

$$f_\alpha : M \to \mathbb{R}$$

such that

1. $f_\alpha \geq 0, \ \forall \alpha, x$
2. *The support of* f_α *is a closed subset of the patch* U_α
3. $\sum_\alpha f_\alpha(x) = 1, \ \forall x \in M$

Such a partition always exists.

Note that when a manifold is compact, for every cover of this manifold there exists a finite sub-cover allowing thus a partition of unity. A very useful application of partition of unity is the construction of a Riemannian metric on a manifold M. Suppose this manifold is covered by the coordinate patches $\{U_\alpha, x^i_\alpha\}$, then we can introduce in each patch a metric

$$ds^2_\alpha = \sum_i (dx^i_\alpha)^2.$$

When introducing these metrics we have a problem in $U_\alpha \cap U_\beta$, since the metrics ds^2_α and ds^2_β need not to agree. But this can be solved by making use of a partition of unity, to get a "global" Riemannian metric

$$ds^2 = \sum_\alpha f_\alpha ds^2_\alpha.$$

7.3 STOKES THEOREM

Many derivations in loop space will heavily depend on the use of Stokes' theorem. Therefore we will here briefly review some of the material about Stokes.

Theorem 7.3.1: Stokes' Theorem

Let X be an oriented manifold of class C^2, dimension n, and let ω be a $(n-1)$-form on X, of class C^1. Assume that ω has compact support. Then

$$\int_X d\omega = \int_{\partial X} \omega$$

For a proof we refer the reader to [97]. The theorem can be extended for ω that have almost compact support, where almost compact support means that there exists a decreasing sequence of open sets U_k in X such that their intersection is empty together with a sequence of C^1 functions that satisfy

(i) $0 \leq g_k \leq 1$, $g_k = 1$ outside U_k and $g_k\omega$ has compact support.

(ii) if μ_k is the measure associated with $|dg_k \wedge \omega|$ on X then

$$\lim_{k\to\infty} \mu_k(\bar{U}_k) = 0$$

Stokes' theorem then becomes

Theorem 7.3.2: Stokes' Theorem 2

Let X be an oriented manifold of class C^2, dimension n, and let ω be a $(n-1)$-form on X, of class C^1. Assume that ω has almost compact support, and that the measures associated with $|d\omega|$ on X and $|\omega|$ on ∂X are finite. Then

$$\int_X d\omega = \int_{\partial X} \omega$$

Even more relevant for us is Stokes' theorem with singularities, meaning that we consider for instance in a two dimensional case polygons. In this case the problems occur at the vertices where ω becomes singular. Nevertheless under certain circumstances Stokes' theorem is still valid. For this version of the theorem we need the concept of **negligible subsets**.

Definition 7.3.1 (Negligible subset)**.**
Let S be a closed subset of $\mathbb{R}^n$. We will call S negligible for X if there exists an open neighborhood U of S in $\mathbb{R}^n$, a fundamental sequence (Cauchy sequence) of open neighborhoods $\{U_k\}$ of S in U, with (the closure) $\overline{U}_k \subset U$, and a sequence of C^1 functions $\{g_k\}$, having the properties:

(i) We have $0 \leq g_k \leq 1$ and $g_k = 0$ for x in some open neighborhood of S, and $g_k = 1$ for $x \notin U_k$.

(ii) If ω is an $(n-1)$-form of class C^1 on U, and μ_k is the measure associated with $|dg_k \wedge \omega|$ on $X \cap U$, then μ_k is finite for large k, and

$$\lim_{k\to\infty} \mu_k(U \cap X) = 0$$

The first condition says that $g_k\omega$ will vanish on an open neighborhood of S, in such a way that $g_k = 1$ on the complement of $\overline{U}_k$ such that $dg_k = 0$ there. Combining this with the second condition we see that the measure on X near the singular points $|dg_k \wedge \omega|$ will tend to zero if k goes to infinity because they are concentrated on shrinking neighborhoods. We can now write down Stokes' theorem with singularities.

Theorem 7.3.3: Stokes' Theorem 3

Let X be an oriented sub-manifold of class C^3 without boundary of $\mathbb{R}^n$. Let dim $X = n$. Let ω be a $(n-1)$-form on X, of class C^1 on an open neighborhood of $\overline{X}$ in $\mathbb{R}^n$, and with compact support. Assume that

(i) If S is the set of singular points in the frontier $\overline{X} - X$, then $S \cap$ supp ω is negligible for X.

(ii) The measures associated with $|d\omega|$ on X, and $|\omega|$ on $\partial\omega$, are finite.

Then

$$\int_X d\omega = \int_{\partial X} \omega$$

We conclude this section by giving two criteria for a set to be negligible [97]:

(i) Let S, T be compact negligible sets of a sub-manifold X of $\mathbb{R}^n$ (and assuming X without boundary). Then the union $S \cup T$ is negligible for X.

(ii) Let X be an open set, and let S be a compact subset in $\mathbb{R}^n$. Assume that there exists a closed rectangle R of dimension $m \leq n-2$ and a C^1 map $\sigma : R \to \mathbb{R}^n$ such that $S = \sigma(R)$. Then S is negligible for X.

8 ALGEBRA

In this chapter we revise some concepts from algebra theory where we start from the most basic definitions slowly building up the structures necessary to give a detailed discussion on the properties of the shuffle product which we will introduce in section 11.2.

8.1 RINGS AND MODULES

Definition 8.1.1 (Monoid)**.**
A monoid or semi-group with unit is a set, S, together with a binary operation "·" that satisfies the three axioms:

(i) Closure :

$$\forall a, b \in S : a \cdot b \in S$$

(ii) Associativity:

$$\forall a, b, c \in S : (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

(iii) Identity element:

$$\exists e \in S, \forall a \in S : (a \cdot e) = (e \cdot a) = a$$

Definition 8.1.2 (Ring)**.**
A ring is a set R equipped with two binary operations $+$ and "·" called "addition" and "multiplication", that map every pair of elements of R to a **unique element** *of R. These operations must satisfy the following properties called* **ring axioms**, *which must be true*

$$\forall a, b, c \in R$$

- *Addition is abelian:*
  - *(i)* $(a+b)+c = a+(b+c)$ *(+ is associative)*
  - *(ii) There is an element* $0 \in R$ *such that*
  $$0 + a = a$$
  (0 is the zero element)
  - *(iii)* $a + b = b + a$ *(+ is commutative)*
  - *(iv)* $\forall a \in R, \exists \ -a \in R | a + (-a) = (-a) + a = 0$ *($-a$ is the inverse element of a)*
- *Multiplication "$\cdot$" is associative:*
  - *(v)* $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

  Multiplication distributes over addition:
  - *(vi)* $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$
  - *(vii)* $(a+b) \cdot c = (a \cdot c) + (b \cdot c)$.
- *Sometimes the above structure is called a pseudo-ring, or a "ring", referring to the fact that some people believe that a ring should have an additional axiom:*
  - *(vii) Multiplicative identity:* $\exists 1 \in R | a \cdot 1 = 1 \cdot a = a$

Rings satisfying all of the above axioms are referred to as **unital rings.** *Although ring addition is commutative, such that $a + b = b + a$, ring multiplication is not required to be commutative ($a \cdot b \neq b \cdot a$). Rings that also satisfy commutativity for multiplication are called* **commutative rings.** *Some basic properties of a ring follow immediately from the axioms:*

- *The additive identity and the additive inverse are unique.*
- *The binomial formula holds for any commuting elements (i.e., $a \cdot b = b \cdot a$).*

Definition 8.1.3 (Field)**.** *(following [98])*
A field F is a set together with two composition laws:

$$F \times F \xrightarrow{+} F \qquad a, b \mapsto a + b \tag{52}$$
$$F \times F \xrightarrow{\times} F \qquad a, b \mapsto ab \tag{53}$$

called addition and multiplication satisfying the axioms:

(i) $(F, +)$ *is an abelian group*

(ii) $(F, \times)$ *is associative and commutative, making* $F \setminus \{0\}$ *into a group. The identity element is written as* 1*.*

(iii) Distributivity :

$$\forall a, b, c \in F, (a + b)c = ac + bc$$

Definition 8.1.4 (Vector Space)**.** *(following [98]))*
A vector space V over a field F is a set together with two composition laws:

$$V \times V \xrightarrow{+} V \qquad v, w \mapsto v + w \tag{54}$$
$$F \times V \xrightarrow{\times} V \qquad c, v \mapsto cv \tag{55}$$

called addition and scalar multiplication satisfying the axioms:

(i) $(V, +)$ *is an abelian group*

(ii) Scalar multiplication is associative with multiplication in F*:*

$$(ab)v = a(bv), \ \forall a, b \in F, \ v \in V \tag{56}$$

(iii) The element 1 *defined in 8.1.3 acts as identity*

$$1v = v, \forall v \in V.$$

(iv) Double Distributivity :

$$(a + b)v = av + bv$$

and

$$a(v + w) = av + aw,$$

$$\forall a, b \in F, \forall v, w \in V.$$

Sometimes the notion of a vector space is not adequate and one needs a generalization which is accomplished by the concept of modules. A module over a ring then generalizes the notion of vector space over a field, where the corresponding scalars are now the elements of an arbitrary ring instead of a field. Modules also generalize the notion of Abelian groups, which are modules over the ring of integers.

Definition 8.1.5 (Module)**.**
A left R-module M over the ring R consists of an abelian group $(M, +)$ and an operation

$$R \times M \to M$$

such that $\forall r, s \in R$ and $x, y \in M$, we have

(i) $r(x + y) = rx + ry$

(ii) $(r + s)x = rx + sx$

(iii) $(rs)x = r(sx)$

(iv) $1_R x = x$

In a similar way one can define a right module and a bimodule which is then a module which is a left module and a right module such that the two multiplications are compatible. If R is commutative, then left R-modules are the same as right R-modules and are simply called R-modules.

At this point we introduce the notation

$$Hom(A, B) = Hom_k(A, B)$$

for the set of morphisms between two k-modules A and B.

8.2 IDEALS

Now that we have introduced the concept of a ring we can define the ideal of a ring, which becomes relevant when considering homomorphisms and algebra morphisms due to theorem 8.2.2 where the relation between ideals and kernels of homomorphisms is stated.

Definition 8.2.1 (Ring Ideal [99])**.**
Let R be a ring. An ideal $\mathfrak{a}$ in R is a subset such that

(i) $\mathfrak{a}$ *is a subgroup of R regarded as a group under addition.*

(ii) $a \in \mathfrak{a}, r \in R \implies ra \in \mathfrak{a}$

For an arbitrary ring $(R, +, \cdot)$, where $(R, +)$ is the underlying additive group and I a subset, we call I a two-sided ideal (or simply an ideal) if it is an additive subgroup of R that "absorbs multiplication by elements of R". A subset S of R can generate an ideal as the intersection of all ideals $\mathfrak{a}$ containing S. The ideal consists of elements of the form

$$\sum r_i s_i$$

with

$$r_i \in R, s_i \in S.$$

Consider now two ideals $\mathfrak{a}$ and $\mathfrak{b}$ in R, then the set

$$\{a + b | a \in \mathfrak{a}, b \in \mathfrak{b}\}$$

is an ideal written as

$$\mathfrak{a} + \mathfrak{b}.$$

In the same way the set

$$\{ab | a \in \mathfrak{a}, b \in \mathfrak{b}\}$$

is an ideal denoted by

$$\mathfrak{a}\mathfrak{b}.$$

Note that

$$\mathfrak{a}\mathfrak{b} \subset \mathfrak{a} \cap \mathfrak{b}.$$

Next to these properties of ideals we also have that **the kernel of a homomorphism $A \to B$ is an ideal in A!**.

Theorem 8.2.1: Kernel is a subring

Let

$$\phi : (R_1, +_1, \circ_1) \to (R_2, +_2, \circ_2)$$

be a ring homomorphism. Then the kernel of ϕ is a subring of R_1.

proof 8.2.1

From the fact that a ring homomorphism of addition is a group homomorphism and the fact that the Kernel is a Subgroup we have

$$(\ker(\phi), +_1) \leq (R_1, +_1),$$

where $\leq$ denotes subgroup. Let now $x, y \in \ker(\phi)$, then we can write

$$\phi(x \circ_1 y) = \phi(x) \circ_2 \phi(y) = 0_{R_2} \circ_2 0_{R_2} = 0_{R_2}$$

Thus

$$x \circ_1 y \in \ker(\phi)$$

and the conditions for a subring are fulfilled, hence $\ker(\phi)$ is a subring of R_1.

Theorem 8.2.2: Kernel is an Ideal

Let

$$\phi : (R_1, +_1, \circ_1) \to (R_2, +_2, \circ_2)$$

be a ring homomorphism. Then the kernel of ϕ is an ideal of R_1.

proof 8.2.2

By theorem 8.2.1, $\ker(\phi)$ is a subring of R_1. Let

$$s \in \ker(\phi),$$

such that

$$\phi(s) = 0_{R_2}.$$

Suppose $x \in R_1$, then we have

$$\begin{aligned}\phi(x \circ_1 s) &= \phi(x) \circ_2 \phi(s) \\ &= \phi(x) \circ_2 0_{R_2} \qquad \text{as } s \in \ker(\phi) \\ &= 0_{R_2}\end{aligned}$$

and similarly for $\phi(s \circ_1 x)$ from which the theorem follows.

Definition 8.2.2 (Cokernel)**.**
Let

$$f : A \to B$$

be an R-module homomorphism. The cockerel is then defined as the quotient group $B/\mathrm{Im}(f)$. Thus f is injective iff its kernel is 0 and surjective iff its cockerel is 0.

Consider now the diagram in figure 16, where C is the kernel of the homomorphism f and i is the inclusion map. Let now

$$f \circ g = 0,$$

this means that there exists a unique module homomorphism

$$h : D \to C$$

such that g can be written as

$$i \circ h.$$

Since the set C contains essentially the same information as the kernel one could instead think of the kernel as the morphism i. In this view any map g that is mapped to 0 by f can be **factored** through

$$i : g = i \circ h.$$

This "factorization" will be essential to understand the diagrams that we will use to define algebraic paths (see 12). Interesting is that when the arrows in figure 16 are reversed we get the corresponding diagram for cokernels (figure 17). The map p is such that

$$C = B/\mathrm{im}(f).$$

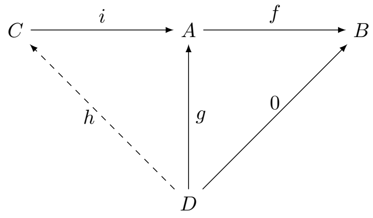


Figure 16: Kernel setup

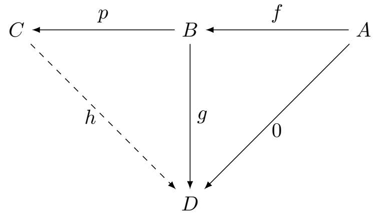


Figure 17: Cokernel setup

If now

$$g \circ f = 0,$$

then g can factored through h as

$$g = h \circ p.$$

Definition 8.2.3 (Prime Ideal [99])**.**
Let R be a ring. An ideal $\mathfrak{p}$ of R is prime if

$$\mathfrak{p} \neq R$$

and

$$ab \in \mathfrak{p} \Rightarrow a \in \mathfrak{p}$$

or

$$b \in \mathfrak{p}.$$

Note that this also means that $\mathfrak{p}$ is prime iff $R/\mathfrak{p}$ is nonzero and has the property

$$ab = 0,\ b \neq 0 \implies a = 0$$

which means that $R/\mathfrak{p}$ is an integral domain (see definition 8.2.8).

Definition 8.2.4 (Maximal Ideal [99])**.**
An ideal $\mathfrak{m}$ in R is maximal if it is maximal among the proper ideals, that is among the ideals that are strictly smaller than the whole ring R.

Thus $\mathfrak{m}$ is maximal iff $A/\mathfrak{m}$ is nonzero and has no proper nonzero ideals, and thus $\mathfrak{m}$ **is a field**.

Remark that if $\mathfrak{m}$ maximal $\implies$ $\mathfrak{m}$ prime. We point out for further use that ideals of $A \times B$ are always of the form $\mathfrak{a} \times \mathfrak{b}$ and that if $\mathfrak{a}$, $\mathfrak{b}$ are ideals in R with $\mathfrak{a} + \mathfrak{b} = R$ we call them **coprime**.

Definition 8.2.5 (Zero Divisor)**.**
A zero divisor a of a ring R is en element of R such that a is a left and right zero divisor :

$$\exists\, b \neq 0 \in R : a.b = 0 \quad \textit{Left zero divisor}$$
$$\exists\, c \neq 0 \in R : c.a = 0 \quad \textit{Right zero divisor}$$

Definition 8.2.6 (Non-zero divisor)**.**
$a \in R$ (ring) is a nonzero divisor if $ab \neq 0$ for all $b \neq 0$ otherwise a zero divisor. a is a **unit** *if there is a b such that $ab = 1$.*

Definition 8.2.7 (Domain)**.**
A non-zero ring R is a domain if every non-zero element is a non-zero divisor and a **field** *if every non-zero element is a* **unit**. *Clearly a field is a domain.*

Definition 8.2.8 (Integral Domain)**.**
An integral domain is a commutative ring that has no zero divisors.

Theorem 8.2.3: Chinese Remainder [99]

Let $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ be ideals in a ring R. If $\mathfrak{a}_i$ is coprime to $\mathfrak{a}_j$ whenever $i \neq j$, then the canonical map $A \to A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n$ is surjective with kernel $\prod \mathfrak{a}_i = \bigcap \mathfrak{a}_i$.

Definition 8.2.9 (Noetherian Ring [99]).
A ring R is called Noetherian if it satisfies the following equivalent conditions:

(i) every ideal in R is finitely generated

(ii) every ascending chain of ideals

$$\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots$$

eventually becomes constant, i.e. for some m

$$\mathfrak{a}_m = \mathfrak{a}_{m+1} = \cdots$$

(iii) every nonempty set of ideals in R has a maximal element (i.e. an element not properly contained in any other ideal in the set)

Definition 8.2.10 (Local Ring [99]).
A ring R is said to be local if it has exactly one maximal ideal $\mathfrak{m}$.

Corollary 8.2.1: Property maximal ideals in Noetherian ring [99]

Let R be a Noetherian ring with maximal ideal $\mathfrak{m}$, then regarding $\mathfrak{m}$ as an R-module, the action of R on $\mathfrak{m}/\mathfrak{m}^2$ factors through $k = R/\mathfrak{m}$.

The elements $a_1, \cdots, a_n$ of $\mathfrak{m}$ generate $\mathfrak{m}$ as an ideal iff their residues modulo $\mathfrak{m}^2$ generate $\mathfrak{m}/\mathfrak{m}^2$ as a vector space over k. In particular, the minimum number of generators for the maximal ideal is equal to the dimension of the vector space $\mathfrak{m}/\mathfrak{m}^2$.

Definition 8.2.11 (Height and Krull dimension [99]).
Let R be a Noetherian ring.

(i) The height ht($\mathfrak{p}$) of a prime ideal $\mathfrak{p}$ in R is the greatest length d of a chain of distinct prime ideals $\mathfrak{p} = \mathfrak{p}_d \supset \mathfrak{p}_{d-1} \supset \cdots \supset \mathfrak{p}_0$

(ii) The Krull dimension of R is sup$\{ht(\mathfrak{p})|\mathfrak{p} \subset R, \mathfrak{p}$ *prime*$\}$

Note that a field has Krull dimension 0 and conversely an integral domain (definition 8.2.8) of Krull dimension 0 is a field.

Definition 8.2.12 (Regular Local Noetherian Ring [99]).
A local Noetherian ring R of Krull dimension d is said to be regular if its maximal ideal can be generated by d elements.

Lemma 8.2.1

In a Noetherian ring, every set of generators for an ideal contains a **finite** generating subset.

8.3 ALGEBRAS

Definition 8.3.1 (Ring Algebra).
An algebra over a commutative ring is a generalization of the concept of an algebra over a field, where the base field k is replaced by a commutative ring R. So let R be a commutative ring. An R-algebra is an R-module M together with a binary operation called the M-multiplication $[\cdot,\cdot]$

$$[\cdot,\cdot] : M \times M \to M$$

that satisfies bi-linearity $\forall a, b \in R, \forall x, y, z \in M$:

$$[ax + by, z] = a[x, z] + b[y, z] \quad \text{and} \quad [z, ax + by] = a[z, x] + b[z, y]$$

An alternative definition is given by definition 8.3.2.

Definition 8.3.2 (k-Algebra).
A k-algebra is a k-vector space A together with two linear maps

$$m : A \otimes_k A \to A \quad \text{and} \quad u : k \to A$$

such that the maps are unital (see definition 8.3.3) and associative together with the conditions that both diagrams in figure 18 (the figure represent the diagrammatic statement of associativity and unit as is common in category theory [100, Chapter 1]) are commutative (where s represents the scalar multiplication) and the unit element in A is obtained as $1_A = u(1_k)$.

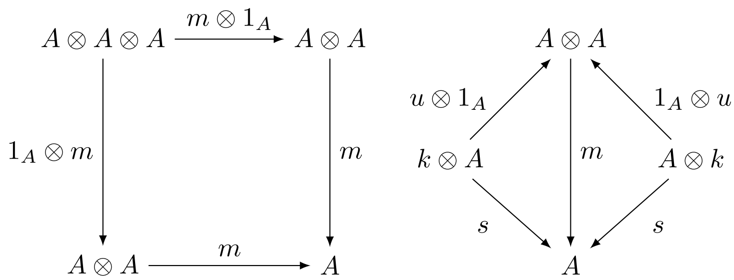


Figure 18: k-algebra commutative diagrams

Definition 8.3.3 (Unital Algebra)**.**
Let (A_R, m) be an algebra over the ring R. Then (A_R, m) is a unitary algebra if it has an identity element 1_A called a Unit of Algebra for m

$$\exists\, 1_A \in A_R : \forall a \in A_R : m(a, 1_A) = m(1_A, a) = a.$$

The unit is usually denoted $\mathbb{1}$ when there is no source of confusion with the identity elements of the underlying structures of the algebra. Another commonly used term is **unital algebra**.

Definition 8.3.4 (Graded Ring)**.**
A graded ring A is a ring that has a decomposition into (abelian) additive groups:

$$A = \bigoplus_{n \in \mathbb{N}} A_n = A_0 \oplus A_1 \oplus A_2 \oplus \cdots$$

such that the ring multiplication satisfies:

(i) $x \in A_s, y \in A_r \implies xy \in A_{s+r}$

(ii) $A_s A_r \subseteq A_{s+r}$.

Elements of any factor A_n *of the decomposition are known as* **homogeneous elements of degree n**. *An ideal or other subset* $\mathfrak{a} \subset A$ *is homogeneous if every element* $a \in \mathfrak{a}$ *is the sum of homogeneous elements that belong to* $\mathfrak{a}$. *For a given* a *these homogeneous elements are uniquely defined and are called the homogeneous parts of* a. *Equivalently, an ideal is homogeneous if for each* a *in the ideal, when*

$$a = a_1 + a_2 + \cdots + a_n$$

with all a_i *homogeneous elements, then all the* a_i *are in the ideal. If* I *is a homogeneous ideal in* A, *then* A/I *is also a graded ring, and has decomposition*

$$A/I = \bigoplus_{n \in \mathbb{N}} (A_n + I)/I.$$

Any (non-graded) ring A *can be given a gradation by letting* $A_0 = A$, *and* $A_i = 0$ *for* $i > 0$. *This is called* **the trivial gradation** *on* A.

Definition 8.3.5 (Graded Module)**.**
A graded module is a left module M over a graded ring A such that

(i) $M = \bigoplus_{i \in \mathbb{N}} M_i$,

(ii) $A_i M_j \subseteq M_{i+j}$.

Graded modules may be considered over non-graded rings by giving the trivial gradation to the ring. This allows one to consider a sequence of modules as a single graded module. This is used in homological algebra to extend to chain complexes some module constructions like direct sum or tensor product.

Definition 8.3.6 (Graded Algebra)**.**
An algebra A over a ring R *is a graded algebra if it is graded as a ring. In the usual case where the ring* R *is not graded (in particular if* R *is a field), it is given the trivial grading (every element of* R *is of grade* 0*). Thus* $R \subseteq A_0$ *and the* A_i *are* R*-modules. In the case where the*

ring R is also a graded ring, then one requires that $A_i R_j \subseteq A_{i+j}$ and $R_i A_j \subseteq A_{i+j}$.

Example 8.3.1 (Graded algebras)**.**

- *Polynomial rings. The homogeneous elements of degree n are exactly the homogeneous polynomials of degree n.*
- *The tensor algebra $T(V)$ of a vector space V. The homogeneous elements of degree n are the tensors of rank n, $T^n(V)$.*
- *The exterior algebra $\bigwedge V$ and symmetric algebra SV are also graded algebras.*
- *The cohomology ring H in any cohomology theory is also graded, being the direct sum of the H^n.*

Definition 8.3.7 (k-Algebra Homomorphism)**.**
Given k-algebras A and B, a k-algebra homomorphism is a ***k-linear map*** *$f : A \to B$ such that*

$$f(xy) = f(x)f(y), \forall x, y \in A.$$

The space of all k-algebra homomorphisms is frequently written as

$$\mathrm{Hom_k(A, B)}.$$

A k-algebra isomorphism is then a bijective k-algebra morphism.

Let now $\mathfrak{A}$ and $\mathfrak{B}$ be two commutative unitary k-algebras. Then the notation

$$\mathsf{Alg}(\mathfrak{A}, \mathfrak{B}) = \mathsf{Alg}_k(\mathfrak{A}, \mathfrak{B})$$

is used for the totality of k-algebra morphisms from $\mathfrak{A}$ to $\mathfrak{B}$ which **take the unit element of $\mathfrak{A}$ into the unit element of $\mathfrak{B}$**. Next to having an algebra structure we sometimes have extra structure that is captured by the co-algebra.

Definition 8.3.8 (Associative Coalgebra).
A k-coalgebra is a k-vector space C together with two linear maps

$$\Delta : C \to C \otimes C \quad \text{and} \quad \epsilon : C \to k$$

called the co-multiplication and co-unit respectively. Again we have axioms which are now referred to as co-associativity and co-unit giving rise to commutative diagrams as shown in figure 19 (in categorial language this is just the dual of the axioms of an algebra, where the arrows of the diagrams in figure 18 are reversed) such that

(i) $(1 \otimes \Delta) \circ \Delta = (\Delta \otimes 1) \circ \Delta$

(ii) $(1 \otimes \epsilon) \circ \Delta = (\epsilon \otimes 1) \circ \Delta$

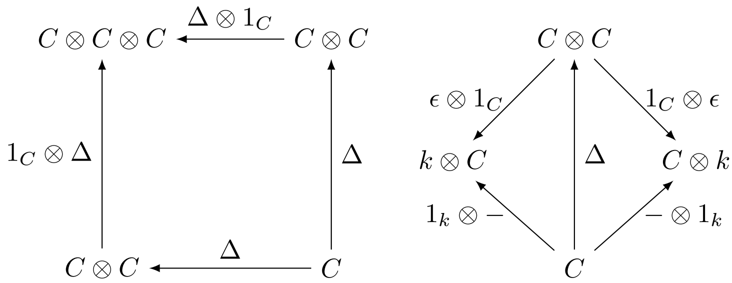


Figure 19: k-coalgebra commutative diagrams

8.4 HOPF ALGEBRA

In this section we discuss the concept of Hopf algebras, where the construction of such an algebra is based on combining an algebra and co-algebra into a bi-algebra.

Definition 8.4.1 (Bialgebra).
A bialgebra A is a k-vector space

$$A = (A, m, u, \Delta, \epsilon)$$

where (A, m, u) *is an algebra and* (A, Δ, ϵ) *is a co-algebra such that*

(i) Δ *and* ϵ *are algebra homomorphisms*

(ii) m *and* u *are coalgebra homomorphisms*

Assume we have two k-algebras A and B, then $A \otimes B$ is also a k-algebra with the multiplication defined by

$$(a \otimes b)(c \otimes d) = ac \otimes bd$$

such that

$$A \otimes B \otimes A \otimes B \overset{1_A \otimes \tau \otimes 1_B}{\longrightarrow} A \otimes A \otimes B \otimes B \overset{m_A \otimes m_B}{\longrightarrow} A \otimes B$$

where

$$\tau : B \otimes A \to A \otimes B$$

is a flipping operation. The unit of $A \otimes B$ is defined as

$$u_{A \otimes B} : k \cong k \otimes k \overset{u_A \otimes u_B}{\longrightarrow} A \otimes B$$
$$u_{A \otimes B}(1_k) = u_{A \otimes B}(1_k \otimes 1_k) = 1_A \otimes 1_B = 1_{A \otimes B}$$

and in the same way for the co-algebra we have

$$C \otimes D \overset{\Delta_C \otimes \Delta_D}{\longrightarrow} C \otimes C \otimes D \otimes D \overset{1_C \otimes \tau \otimes 1_D}{\longrightarrow} C \otimes D \otimes C \otimes D$$

where τ is again the flipping function with the co-unit now defined as

$$C \otimes D \overset{\epsilon_C \otimes \epsilon_D}{\longrightarrow} k \otimes k \cong k$$
$$\epsilon_C \otimes \epsilon_D(1_C \otimes 1_D) = \epsilon(1_C) \otimes \epsilon(1_D) = 1_k \otimes 1_k = 1_k$$

We have a special case of the above when $A = B$ and $C = D$. Notice that a bi-algebra morphism is both an algebra and co-algebra homomorphism.

Definition 8.4.2 (Bi-ideal)**.**
If

$$f : A \to B$$

is a bi-algebra homomorphism then $\ker\ f$ *is called a bi-ideal, meaning that* $\ker\ f$ *is both an ideal and a co-ideal.*

$$I \subset A$$

is a co-ideal when

$$\epsilon(I) = 0$$

and

$$\Delta(I) \subseteq C \otimes I + I \otimes C.$$

Definition 8.4.3 (Hopf Algebra)**.**
Given a commutative ring R, an R-algebra H is a Hopf algebra if it has additional structure given by R-algebra homomorphisms:

(i) comultiplication $\Delta : H \to H \otimes_R H$

(ii) counit $\epsilon : H \to R$

(iii) (antipode) R-module homomorphism $\lambda : H \to H$

satisfying

(iv) coassociativity: $(I \otimes \Delta)\Delta = (\Delta \otimes I)\Delta : H \to H \otimes H \otimes H$

(iiv) counitarity : $m(I \otimes \epsilon)\Delta = I = m(\epsilon \otimes I)\Delta$

(iiiv) antipode : $m(I \otimes \lambda)\Delta = \imath\epsilon = m(\lambda \otimes I)\Delta$

where I is the identity map on H,

$$m : H \otimes H \to H$$

is the multiplication in H, and

$$\imath : R \to H$$

is the R-algebra structure map for H, also called the unit map.

The symbol $\otimes_R$ stresses that the product is R equivariant, meaning that if either of factors is multiplied with a scalar (an element of R, since we are considering R-algebras), it can be taken out of the product symbolized by $\otimes$. Summarizing this symbolically we have

$$a \in R, h_1, h_2 \in H : \Delta(ah_1, h_2) = a\Delta(h_1, h_2).$$

Lemma 8.4.1

When I is a bi-ideal of a bi-algebra A, then the operations on A induce the structure of a bi-algebra on A/I, such that the bi-algebraic structure is preserved when projecting A on A/I.

Example 8.4.1 (Bialgebras)**.**

(i) *Group Algebras : Let G be any group and kG the group algebra. Defining*

$$\Delta : kG \to kG \otimes kG$$

by

$$\Delta(g) = g \otimes g, \forall g \in G$$

and

$$\epsilon(g) = 1, \forall g \in G$$

makes it into a bi-algebra.

(ii) *Tensor Algebras: Let V be a vector space and T(V) its tensor algebra*

$$(T(V) = \bigoplus_{n\geq 0}(V^{\otimes n})),$$

then defining

$$\Delta(V) = v \otimes 1_V + 1_V \otimes v \in T(V) \otimes T(V)$$

and

$$\epsilon(v) = 0, \forall v \in V$$

makes into a bi-algebra.

(iii) *Universal enveloping algebra U of a Lie Algebra $\mathfrak{g}$ is a vector space over a field k with a bilinear, anti-symmetric and "Jacobi Identity" obeying operation*

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}.$$

The universal enveloping algebra U *of* $\mathfrak{g}$ *is the factor algebra of the tensor algebra* $T(\mathfrak{g})$ *by the ideal*

$$I(\mathfrak{g}) = \langle [x,y] - xy + yx : x, y \in \mathfrak{g} \rangle$$

Note that there is a bijection between the left $U(\mathfrak{g})$*-modules and the representations of the Lie Algebra* $\mathfrak{g}$*, i.e. the Lie Algebra homomorphisms*

$$\rho : \mathfrak{g} \to \mathit{End}(V)$$

for k-vector spaces V.

- *Let*
$$n \geq 1, \mathfrak{g} = \sum_{i=1}^{n}{}^{\oplus} kx_i$$
with
$$[x_i, x_j] = 0, \forall i, j$$
then
$$U(\mathfrak{g}) \cong k[x_1, \cdots, x_n]$$
the commutative polynomial algebra.
- *Let* $n \geq 2$ *and* $\mathfrak{sl}(n,k)$ *the space of* $n \times n$ *matrices of trace* 0, $\mathfrak{sl}(2,k)$ *is 3-dimensional with the usual basis*

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

for which

$$[h,e] = 2e, \ [h,f] = -2f, \ [e,f] = h$$

such that we have for the ideal

$$I(\mathfrak{sl}(n,k)) = \langle he - e(h+2), \ hf - f(h-2), \ ef - fe - h \rangle$$

Now it is easy to see that $\{e^i f^j h^t : i, j, t \geq 0\}$ *forms a k-basis for* $U(\mathfrak{sl}(n,k))$.

Example (iii) above is a special case of the Poincaré-Birkhoff-Witt theorem (8.4.1), which we state below for completeness.

Theorem 8.4.1: Poincaré-Birkhoff-Witt

If $\mathfrak{g}$ is a Lie Algebra with k-basis $\{x_1, \cdots, x_n\}$ then $U(\mathfrak{g})$ has a k-basis $\{x_1^{t_1}, \cdots, x_n^{t_n} : t_i \geq 0\}$.

As we will see in the construction of GLS, a specific ideal will play a prominent role in its realization. One of the reasons that this ideal will turn out to be so important can be explained by some remarks on universal enveloping algebras of Lie algebras and their representations. Remember that a representation ρ assigns to any element x of a Lie algebra a linear operator $\rho(x)$. Due to this linearity these operators not only form a Lie algebra but also an associative algebra allowing the consideration of products like $\rho(x)\rho(y)$. In general the result of this product can depend on the chosen representation, but some properties seem to hold for any representation. The introduction of the universal enveloping algebra is then a way to single out these universal properties and study them. It should now be clear that $U(\mathfrak{g})$ is a bi-algebra for all Lie Algebras $\mathfrak{g}$ with

$$\Delta(x) = 1 \otimes x + x \otimes 1$$

and

$$\epsilon(x) = 0, \forall g \in \mathfrak{g}.$$

We then have that

$$U(\mathfrak{g}) = T(\mathfrak{g})/I(\mathfrak{g})$$

and

$$T(\mathfrak{g})$$

are bi-algebras with the same definitions for the co-unit and co-multiplication.

Definition 8.4.4 (Opposite Algebra and Co-algebra)**.**
The opposite algebra A^{op} of a k-algebra A is the same vector space as A but now with multiplication defined as

$$\forall a, b \in A : m'(a, b) := m(b, a).$$

In the same way we have for the co-algebra C, C^{op} which is defined on the same vector space but with

$$\Delta_{C^{op}} := \tau \circ \Delta_C$$

with τ the flipping operation as defined before in the discussion below definition 8.4.1.

Definition 8.4.5 (Co-commutative).
A co- or bi-algebra A is called co-commutative if

$$A^{op} = A$$

or thus

$$\Delta = \tau \circ \Delta.$$

To reduce notations further down this chapter we introduce **the Sweedler notation**. Take C a co-algebra and introduce the notation

$$c \in C : \Delta(c) = \sum c_1 \otimes c_2.$$

Considering the associativity axiom we have that

$$\begin{aligned} (1 \otimes \Delta) \circ \Delta(c) &= (1 \otimes \Delta)(\sum c_1 \otimes c_2) \\ &= \sum c_1 \otimes c_{21} \otimes c_{22} \\ &= \sum c_{11} \otimes c_{12} \otimes c_2 \end{aligned}$$

which we could write as

$$\sum c_1 \otimes c_2 \otimes c_3.$$

In the same way we can continue for more factors

$$\Delta_{n-1} : C \to C^{\otimes n}.$$

Using this notation and the right diagram in figure 19 it is easy to see that

$$c = \sum \epsilon(c_1)c_2 = \sum c_1\epsilon(c_2)$$

and that C is co-commutative iff

$$\Delta(c) = \sum c_1 \otimes c_2, \forall c \in C.$$

Definition 8.4.6 (Antipode)**.**
We start from the bi-algebra

$$A = (A, m, u, \Delta, \epsilon).$$

A linear endomorphism

$$S : A \to A$$

is called an antipode if the diagram in figure 20 commutes (see also definition 8.4.3 of a Hopf Algebra). Expressed in Sweedler notation this statement becomes

$$\forall a \in A, \epsilon(a) = \sum a_1 S(a_2) = \sum S(a_1) a_2.$$

A Hopf algebra is then a bi-algebra with an antipode. ***Hopf algebra morphisms are then antipode preserving bi-algebra morphisms****.*

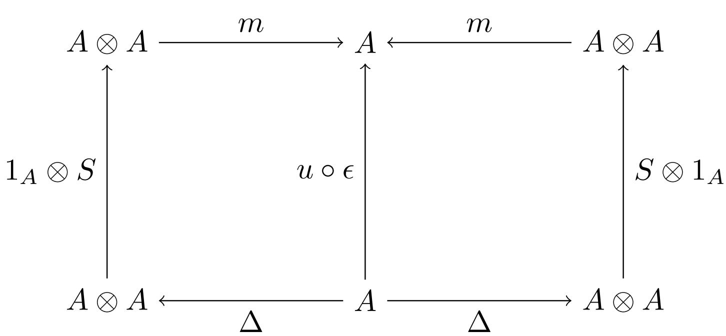


Figure 20: Hopf commutative diagram

Example 8.4.2.
Let G be a group and k any field. Then kG is a Hopf algebra if we define

$$S(g) = g^{-1}, \forall g \in G.$$

Since S is assumed linear by definition this translates in the Sweedler notation as

$$1 = \epsilon(g) = gg^{-1} = g^{-1}g, \forall g \in G$$

Definition 8.4.7 (Convolution Product)**.**
Let

$$A = (A, m, u)$$

be an algebra and

$$C = (C, \Delta, \epsilon)$$

a co-algebra over k. Then we can define [101] the convolution product "$*$" over

$$\mathrm{Hom}_k(C, A)$$

as

$$f, g \in \mathrm{Hom}_k(C, A), c \in C : (f * g)(c) = \sum f(c_1)g(c_2).$$

Proposition 8.4.1

$$(\mathrm{Hom}_k(C, A), *, u \circ \epsilon)$$

is an algebra.

The easiest way to see this is by noticing that

$$m \equiv * : \mathrm{Hom}_k(C, A) \to \mathrm{Hom}_k(C, A),$$

and

$$u \circ \epsilon : \mathrm{Hom}_k(C, A) \to \mathrm{Hom}_k(C, A),$$

an identity map. Now putting $C = A$ this becomes a bi-algebra with

$$(\mathrm{End}_k(A), *, u \circ \epsilon)$$

and for a bi-algebra we then have that the antipode S for A is an inverse of 1_A in

$$(\mathrm{End}_k(A), *, u \circ \epsilon)$$

which is **uniquely** determined due to the uniqueness of inverses.

Corollary 8.4.1

Let

$$C = (C, \Delta, \epsilon)$$

be any k-co-algebra, then

$$C^* = \mathrm{Hom}_k(C, k)$$

is an algebra with

$$(f * g)(c) = \sum f(c_1)g(c_2),$$

making C^* to be commutative if and only if C is co-commutative (8.4.5).

Theorem 8.4.2

Let

$$H = (H, m, \Delta, u, \epsilon, S)$$

be a Hopf algebra, making S a bi-algebra homomorphism from H to H^{opcop}

(i) $S(m(x, y)) = m(S(y), S(x)),\ S(1) = 1$

(ii) $(S \otimes S) \circ \Delta = S,\ \ \epsilon \circ S = \epsilon \longleftrightarrow \forall x, y \in H, S(x_2) \otimes S(x_1) = \sum (Sx)_1 (Sx)_2.$

Definition 8.4.8 (Anti-Homomorphism)**.**
An anti-homomorphism of rings is defined by the fact that if R, S are rings such that

$$\theta : R \to S$$

satisfies

$$\theta(rt) = \theta(t)\theta(r)$$

(notice the change of order !)*.*

Corollary 8.4.2: $S : A \to A^{\text{op}}$ is antipode

Let A be a bi-algebra and S be the map

$$S : A \to A^{\text{op}},$$

an algebra homomorphism. Then S is an antipode for A.

Example 8.4.3 (Hopf Algebra).
Let $\mathfrak{g}$ be a Lie Algebra and recall that

$$U(\mathfrak{g}) := T(\mathfrak{g})/I(\mathfrak{g})$$

then we see that

$$\forall x, y \in \mathfrak{g} : S([x,y] - xy + yx) = -[x,y] - (-y)(-x) + (-x)(-y) \\ = -([x,y] - xy + yx) \in I(\mathfrak{g}). \quad (57)$$

S is an anti-automorphism of $U(\mathfrak{g})$ making it into a Hopf algebra.

Definition 8.4.9 (Restricted Dual of a k-algebra).
Let A be any k-algebra, then the finite or restricted dual of A is defined by the set

$$A^\circ = \{f \in A^* : f(I) = 0 \text{ for some } I \triangleleft A,\ \dim_k(A/I) < \infty\},$$

where $I \triangleleft A$ indicates that I is an ideal of A.

We now extend the concept of a module over a ring to a module over a k-algebra and a k-coalgebra introducing also the concept of a co-module.

Definition 8.4.10 (Module over a k-algebra [101]).
The left module M over a k-algebra A is a k-vector space M with a k-linear map

$$\lambda : A \otimes M \to M$$

such that the diagrams shown in figure 21 commute, where again s is the scalar multiplication, u the algebra unit and m the algebra multiplication in A.

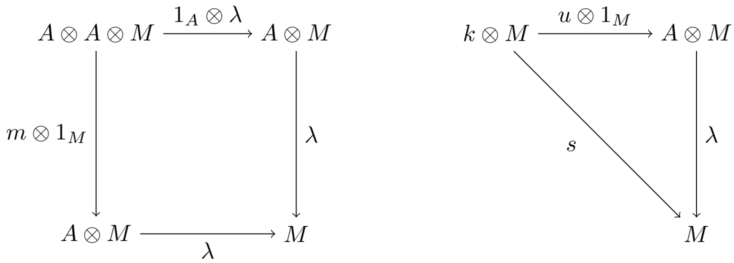


Figure 21: Module over a k-algebra

Definition 8.4.11 (Comodule [101]).
Starting with a k-coalgebra C, then a right co-module M over C is a k-vector space M with a k-linear

$$\rho : M \to M \otimes C,$$

such that the diagrams in figure 22 commute. Again Δ represents the co-algebra multiplication in C.

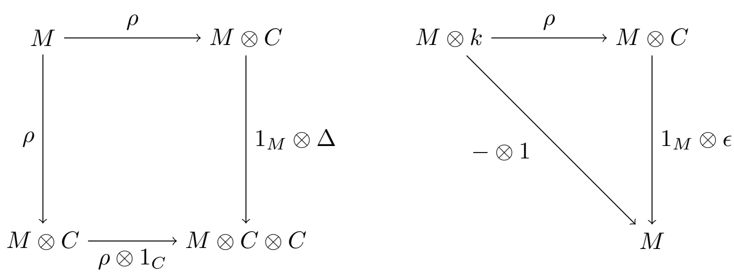


Figure 22: Comodule over a k-coalgebra

Proposition 8.4.2: Duality

1. Let M be a right co-module for the co-algebra C. Then M is a left module for
$$C^* = \mathrm{Hom}(C, k).$$

2. Let A be an algebra and M a left A-module. Then M is a right A°-comodule if and only if

$$\forall m \in M, \dim_k(Am) < \infty.$$

Definition 8.4.12 (Rational Module)**.**
An A-module M with $\dim_k(Am) < \infty, \forall m \in M$ is called ***rational****.*

We end this chapter with some comments on tensor products of modules and co-modules. Next to this we also briefly comment on homomorphism between modules.

Example 8.4.4 (Modules)**.**

- ***Tensor products of modules.***
  Let A be a bi-algebra, and let V and W be left A-modules. Then

  $$V \otimes W$$

  is a left A-module through

  $$a \cdot (v \otimes w) = \sum a_1 v \otimes a_2 w.$$

  If we now consider a third A-module X then co-associativity assures that
  $$(V \otimes W) \otimes X \cong V \otimes (W \otimes X).$$
  In the specific case of the trivial left A-module k ($a \cdot v = \epsilon(a)v, a \in A, v \in k$) it is clear that

  $$V \otimes k \cong V \cong k \otimes V,$$

  as left modules. If A is co-commutative then we also have that

  $$V \otimes W \cong W \otimes V,$$

  as left modules with the isomorphism given by the flip function

  $$\tau : v \otimes w \to w \otimes v.$$

- ***Homomorphism of modules.***
  Let H be a Hopf algebra and V, W left H-modules. Then

$$\mathrm{Hom}_k(V, W)$$

  is a left H-module with the action

$$(h \cdot f)(v) = \sum h_1 f((Sh_2)v), h \in H, f \in \mathrm{Hom}_k(V, W).$$

- ***Tensor products of comodules.***
  If B is a bi-algebra and V, W are right B-comodules, then

$$V \otimes W$$

  is a right co-module with

$$v \otimes w \mapsto \sum v_0 \otimes w_0 \otimes v_1 w_1.$$

This concludes our introduction to algebra theory, in the next chapter we will study algebras that also carry a topological structure forming topological algebras.

9

TOPOLOGICAL ALGEBRAS

The following definitions and theorems can be found in [102, 103], where much more information is available. Here we just state the definitions and properties relevant for the introduction of GLS (18).

9.1 TOPOLOGICAL ALGEBRA, C^*-ALGEBRAS AND BANACH ALGEBRAS

Definition 9.1.1 (Topological Algebra)**.**
A topological algebra is an algebra endowed with a non-trivial topology τ *which is* ***compatible with its linear structure*** *such that the map*

$$X \times X \to X, \; (x, y) \mapsto xy,$$

is continuous.

Definition 9.1.2 (Normed Algebra)**.**
An algebra A equipped with a norm is called a normed algebra if the norm is ***sub-multiplicative***

$$\|ab\| \leq \|a\| \, \|b\| \quad \forall a, b \in A.$$

A norm gives rise to a metric and from chapter 6 we know that this can be used to introduce a topology so if A is a normed algebra then the norm induces a metric on A which in its turn then induces the norm topology on A .

Lemma 9.1.1: Continuity in Normed Algebras

If A is a normed algebra, then all the algebraic operations are continuous in the norm topology on A.

proof 9.1.1

Let us for example consider the sub-multiplicativity of the norm, then we have that

$$\begin{aligned} \|xy - ab\| &\leq \|xy - xb\| + \|xb - ab\| \\ &\leq \|x\|\|y - b\| + \|x - a\|\|b\|, \end{aligned} \tag{58}$$

from which it clearly follows that multiplication is continuous from

$$A \times A \to A.$$

The same goes for addition and scalar multiplication.

Definition 9.1.3 (Involution on an Algebra)**.**
An involution on an algebra α is a map

$$* : \ A \to A | a \mapsto a^*$$

satisfying

(i) Conjugate linear

$$(za + z'b)^* = \overline{z}a^* + \overline{z}'b^*$$

(ii) Order reversing

$$(ab)^* = b^*a^*$$

(iii) Involutive

$$(a^*)^* = a$$

for all

$$a, b \in \alpha, \ \ z, z' \in \mathbb{C}.$$

An algebra with involution is called an $^*-$algebra.

Equipping a Banach algebra with an involution allows for the definition of a so-called C^*-algebra.

Definition 9.1.4 (C^*-algebra)**.**
A C^*-algebra A is a Banach algebra (Definition 9.1.6) equipped with an involution that satisfies the compatibility condition between the involutive and metrical structure

$$\|a^*a\| = \|a\|^2$$

A Banach space is a vector space X over the field of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$ which is equipped with a norm and which is complete with respect to that norm. Formally, the definition of a Banach space is

Definition 9.1.5 (Banach Space)**.**
*A normed space X is said to be a Banach space if for every **Cauchy sequence***

$$\{x_n\}_{n=1}^{\infty} \subset X,$$

there exists an element $x \in X$ such that

$$\lim_{n\to\infty} x_n = x.$$

Notice that the completeness refers to the existence of a limit for every Cauchy sequence such that this limit is an element of the space X. A little remark on the difference between norms and metrics. In metric spaces, the completeness is a property of the metric. It is not a property of the topological space itself. If you move on to an equivalent metric (that is a metric which induces the same topology), the completeness property can get lost. For norms on the other hand we have that two equivalent norms on a normed vector space, where one of them is complete, the other one is also complete. Therefore, in the case of normed vector spaces, the completeness is a property of the norm topology, which does not depend on the specific norm. In this sense the norm is a stronger concept than a metric.

Definition 9.1.6 (Banach Algebra)**.**
A normed algebra which is complete in the metric induced by the norm is called a Banach Algebra.

9.2 NUCLEAR MULTIPLICATIVE CONVEX HAUSDORFF ALGEBRAS AND THE GEL'FAND SPECTRUM.

In the above context a topological algebra isomorphism is an algebra morphism which is also a homeomorphism. This allows for huge structures to be built on such isomorphisms, of which we introduce some and discuss their properties.

Definition 9.2.1 (Filter Basis)**.**
A non-empty subset F of a partially ordered set $(P, \leq)$ is a filter if the following conditions hold:

(i) For every $x, y \in F$, there is some element $z \in F$ such that $z \leq x$ and $z \leq y$. (F is a filter base, or downward directed)

(ii) For every $x \in F$ and $y \in P, x \leq y$ implies that $y \in F$. (F is an upper set, or upward closed)

(iii) A filter is proper if it is not equal to the whole set P. This is sometimes omitted from the definition of a filter.

Definition 9.2.2 (Bases for compatible topologies)**.**
The filter basis $\mathcal{B}$ in the algebra X determines a basis at 0 for a compatible topology for X iff

(i) $\mathcal{B}$ is a neighborhood base at 0 for a topology which is compatible with X's linear structure.

(ii) For each $V \in \mathcal{B}$ there exists a $B \in \mathcal{B}$ such that[a] $BB \subset V$

a BB represents the product of two B's

The following type of induced topology will be very important in the topologization of the shuffle algebra.

Definition 9.2.3 (Initial Topology)**.**
Let X be an algebra, Y a topological algebra with neighborhood filter $V(0)$ at 0. Take

$$A : X \to Y$$

to be a homomorphism. Then the filter

$$A^{-1}(V(0))$$

defines a topology ***compatible*** *with X's* ***linear*** *structure. To see that it is compatible with the algebraic structure as well, we first note that for any*

$$V \in V(0)$$

there is a

$$B \in V(0)$$

such that

$$BB \subset V.$$

Hence

$$A^{-1}(B)A^{-1}(B) \subset A^{-1}(BB) \subset A^{-1}(V).$$

The topology determined by

$$A^{-1}(V(0))$$

is called the initial (inverse image, weak) topology induced by the homomorphism A.

Definition 9.2.4 (Final Topology)**.**

Suppose X is a topological algebra with filter of neighborhoods of 0 denoted by $V(0)$, Y an algebra, and

$$A : X \to Y$$

a homomorphism. It is easy to see that the collection $\mathcal{B}$ of subsets U of Y such that

$$A^{-1}(U) \in V(0)$$

forms a base at 0 for a topology compatible with X's linear structure. For any $U \in \mathcal{B}$ we may select $B \in V(0)$ such that

$$BB \in A^{-1}(U).$$

Thus

$$A(B)A(B) = A(BB) \subset A(A^{-1}(U)) \subset U.$$

Since

$$A^{-1}(A(B)) \supset B \in V(0)$$

it follows that

$$A^{-1}(A(B)) \in V(0)$$

i.e. that

$$A(B) \in \mathcal{B}$$

so that $\mathcal{B}$ is a base at 0 for a topology which is compatible with Y's algebraic structure. The topology generated by 0 is called the final topology for Y determined by the homomorphism A.

Suppose V is a vector space over k, a field or subfield of the complex numbers. A locally convex space is then defined either in terms of convex sets, or equivalently in terms of semi-norms.

Definition 9.2.5 (Convex space using convex sets)**.**
A subset C in V is called

(i) Convex if for each

$$x, y \in C,\ tx + (1-t)y \in C, \forall t \in [0,1].$$

In other words, C contains all line segments between points in C.

(ii) Circled if

$$\forall x \in C, \lambda x \in C$$

with

$$|\lambda| = 1.$$

If the underlying field k is the real numbers, this means that C is equal to its reflection through the origin. For a complex vector space V, it means for any $x \in C$, C contains the circle through x, centered on the origin, in the one-dimensional complex subspace generated by x.

(iii) A cone if for every $x \in C$ and $0 \leq \lambda \leq 1, \lambda x \in C$.

(iv) *Balanced* if

$$\forall x \in C, \lambda x \in C$$

with $|\lambda| \leq 1$. If the underlying field k is the real numbers, this means that if $x \in C$, C contains the line segment between x and $-x$. For a complex vector space V, it means for any $x \in C$, C contains the disk with x on its boundary, centered on the origin, in the one-dimensional complex subspace generated by x. Equivalently, a balanced set is a circled cone.

(v) *Absorbent or absorbing if the union of tC over all $t > 0$ is all of V, or equivalently for every $x \in V, tx \in C$ for some $t > 0$. The set C can be scaled out to absorb every point in the space.*

(vi) *Absolutely convex if it is both balanced and convex.*

More succinctly, a subset of V is absolutely convex if it is closed under linear combinations whose coefficients absolutely sum to ≤ 1. Such a set is absorbent if it spans all of V.

A locally convex topological vector space is a topological vector space in which the origin has a local base of absolutely convex absorbent sets. Because translation is (by definition of "topological vector space") continuous, all translations are homeomorphisms, so every base for the neighborhoods of the origin can be translated to a base for the neighborhoods of any given vector.

Definition 9.2.6 (Convex space using semi-norms).
A semi-norm on V is a map $p : V \to \mathbb{R}$ such that

(i) *p is positive or positive semidefinite*

$$p(x) \geq 0.$$

(ii) *p is positive homogeneous or positive scalable*

$$p(\lambda x) = |\lambda| p(x)$$

for every scalar λ. So, in particular, $p(0) = 0$.

(iii) p is subadditive. It satisfies the triangle inequality

$$p(x + y) \leq p(x) + p(y).$$

If p satisfies positive definiteness, which states that if $p(x) = 0$ then $x = 0$, then p is a norm. While in general semi-norms need not be norms, there is an analogue of this criterion for families of semi-norms defined below.

A locally convex space is then defined to be a vector space V along with a family of semi-norms $\{p_\alpha\}_{\alpha \in A}$ on V. The space carries a natural topology, the initial topology (9.2.3) of the semi-norms. In other words, it is the coarsest (weakest) topology for which all mappings

$$x \to p_\alpha(x - x_0), x_0 \in V, \alpha \in A,$$

are continuous. A base of neighborhoods of x_0 for this topology is obtained in the following way, for every finite subset "B" of "A" and every $\epsilon > 0$ let

$$U_{B,\varepsilon}(x_0) = \{x \in V : p_\alpha(x - x_0) < \varepsilon, \quad \alpha \in B\}.$$

That the vector space operations are continuous in this topology follows from properties (ii) and (iii) above. The resulting TVS (Topological Vector Space) is locally convex because each $U_{B,\varepsilon}(0)$ is ***absolutely convex and absorbent.***

Definition 9.2.7 (Multiplicative Convexity (m-convex))**.**
A subset U of an algebra X is called multiplicative (idempotent) if

$$U^2 = UU \subset U.$$

It is called multiplicatively-convex or m-convex if it is convex and multiplicative, absolutely m-convex if it is balanced and m-convex.

Example 9.2.1 (Multiplicative Convexity)**.**
An immediate example of multiplicative sets is afforded by the spheres, open or closed, of radius

$$1/n, n \in \mathbb{N}$$

about 0 in any normed algebra. As is apparent, each such sphere is absolutely m-convex as well.

Definition 9.2.8 (Multiplicative Semi-norm)**.**
A semi-norm p on an algebra X is multiplicative if

$$p(xy) \leq p(x)p(y), \forall x, y \in X.$$

We note that for a multiplicative semi-norm p to be non-trivial on an algebra X, it is necessary and sufficient that $p(e)$ be non-zero. The trivial semi-norm (i.e. identically zero) is multiplicative and generates the trivial topology.

Definition 9.2.9 (Locally m-convex Algebras and Fréchet Algebras)**.**
A topological algebra (X, τ) is a locally m-convex algebra (LMC algebra) if there is a basis of m-convex sets for $V(0)$. We also say that τ is locally m-convex or is an LMC- topology. X is a locally convex algebra if X is a topological algebra which carries a locally convex linear space structure. If, in addition to being locally m-convex, τ is Hausdorff, we say that X is an LMCH algebra, and τ to be LMCH. An LMC algebra which is a complete metrizable topological space is a ***Fréchet algebra****.*

Proposition 9.2.1

A topological algebra X is locally m-convex iff its topology is generated by a family of multiplicative semi-norms.

Using the above definitions we can extend the concept of real manifolds, modeled on $\mathbb{R}^n$ to spaces that are modeled on Banach, Hilbert or Fréchet spaces instead. We will discuss this in more detail when considering the Fréchet derivative on GLS. We just state here the definition for completeness.

The following space is a generalization of a Banach space, that is locally convex and complete with respect to a translation invariant metric. However the metric does not need to arise from a norm (a semi-norm suffices).

Definition 9.2.10 (Fréchet space 1)**.**
A topological vector space X is a Fréchet space if it has the following properties:

(i) *X is Hausdorff*

(ii) *The topology on X can be induced by a countable family of semi-norms*

$$\|\cdot\|_k, k = 0, 1, 2, \cdots .$$

This means that an open $U \subset X$ is open if and only if

$$\forall u \in U, \exists K \geq 0, \epsilon > 0 : \{\nu : \|\nu - u\|_k < \epsilon, \forall k \leq K\}$$

is a subset of U.

(iii) *X is complete with respect to the family of semi-norms.*

Definition 9.2.11 (Hilbert Space)**.**
In order to have a Hilbert space X we need:

(i) *A positive definite, sesquilinear form (the inner product) on the complex linear space X*

$$\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$$

such that

a) $\langle x, x \rangle \geq 0$ *and* $\langle x, x \rangle = 0$ *iff* $x = 0$.

b) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$

c) $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$

d) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

This makes X into a pre-Hilbert space.

(ii) *A collection of vectors* (x_n) *is said to be orthonormal iff* $\langle x_m, x_n \rangle = \delta_{mn}$

Definition 9.2.12 (Compact operator)**.**
An operator $\mathcal{L}$ *on a Hilbert space* $\mathcal{H}$*:*

$$\mathcal{L} : \mathcal{H} \to \mathcal{H}$$

is said to be a compact operator if it can be written in the form

$$\mathcal{L} = \sum_{n=1}^{N} \rho_n \langle f_n, \cdot \rangle g_n, \quad 1 \leq N \leq \infty$$

where

$$f_1, \ldots, f_N \text{ and } g_1, \ldots, g_N$$

are (not necessarily complete) orthonormal sets. Here,

$$\rho_1, \ldots, \rho_N$$

are a set of real numbers, the singular values of the operator, obeying $\rho_n \to 0$ *if* $N \to \infty$. *The bracket* $\langle \cdot, \cdot \rangle$ *is the scalar product on the Hilbert space; the sum on the right hand side must converge in the norm.*

Definition 9.2.13 (Singular Values of a Compact Operator)**.**
The singular values, or s-numbers of a compact operator

$$T : X \to Y$$

acting between Hilbert spaces X *and* Y, *are the square roots of the eigenvalues of the non-negative self-adjoint operator*

$$T^*T : X \to X$$

(where T^* *denotes the adjoint of* T*).*

Definition 9.2.14 (Nuclear Operator)**.**
*An operator that is compact as defined above is said to be **nuclear** or **trace-class** if*

$$\sum_{n=1}^{\infty} \rho_n < \infty$$

Proposition 9.2.2: Trace of Nuclear Operator on Hilbert Space

A nuclear operator on a Hilbert space has the important property that its trace may be defined so that it is finite and is independent of the basis. Given any orthonormal basis $\{\psi_n\}$ for the Hilbert space, one may define the trace as

$$\text{Tr } \mathcal{L} = \sum_n \langle \psi_n, \mathcal{L}\psi_n \rangle$$

since the sum converges absolutely and is independent of the basis. Furthermore, this trace is identical to the sum over the eigenvalues of $\mathcal{L}$ (counted with multiplicity).

Proposition 9.2.3: Trace of Nuclear Operator on Banach Space

Let A and B be Banach spaces, and A^* be the dual of A, that is, the set of all continuous or (equivalently) bounded linear functionals on A with the usual norm. Then an operator

$$\mathcal{L} : A \to B$$

is said to be **nuclear of order q** if there exist sequences of vectors

$$\{g_n\} \in B$$

with

$$\|g_n\| \leq 1,$$

functionals

$$\{f_n^*\} \in A^*$$

with

$$\|f_n^*\| \leq 1$$

and complex numbers $\{\rho_n\}$ with

$$\inf\left\{p \geq 1 : \sum_n |\rho_n|^p < \infty\right\} = q,$$

such that the operator may be written as

$$\mathcal{L} = \sum_n \rho_n f_n^*(\cdot) g_n$$

with the sum converging in the operator norm (see Definition 17.2.1). With additional steps, a trace may be defined for such operators when A = B. Operators that are nuclear of order 1 are called **nuclear operators**: these are the ones for which the series $\sum \rho_n$ is absolutely convergent. Nuclear operators of order 2 are called Hilbert-Schmidt operators.

More generally, an operator from a locally convex topological vector space A to a Banach space B is called nuclear if it satisfies the condition above with all f_n^* bounded by 1 on some fixed neighborhood of 0 and all g_n bounded by 1 on some fixed neighborhood of 0.

Definition 9.2.15 (Weak Topology).
The collection of all unions of finite intersection of sets of the form

$$f_i^{-1}(O_i)$$

for

$$f : X \to Y$$

where $i \in I$ and O_i is an open set in Y_i is a topology. It is called the weak topology on X generated by the

$$(f_i)_{i \in I}$$

and we denote it by

$$\sigma(X, (f_i)_{f_i \in I}).$$

***By definition**, the functions $(f_i)_{i \in I}$ are continuous for this topology.*

Definition 9.2.16 (Weak-$*$ Topology).
The weak-$*$ topology on X is the topology

$$\sigma(X, (f)_{f \in X^*}).$$

For convenience, it is simply denoted $\sigma(X, X^*)$.

Definition 9.2.17 (Vanish at infinity).
If X is a locally compact Hausdorff space, then a continuous function f on X is said to vanish at infinity if

$$\{x \in X : |f(x)| \geq \epsilon\}$$

is compact for all $\epsilon > 0$. The collection of all such f is denoted by $C_0(X)$.

Definition 9.2.18 (Gel'fand Space or Spectrum).
Let A be a commutative Banach algebra, then we write $\triangle(A)$ (or $\triangle$) for the collection of nonzero complex homomorphisms

$$h : A \to \mathbb{C}.$$

Elements of the Gel'fand space are called characters.

Note that definition 9.2.18 does not contain any reference to continuity or any other assumption on h.

Theorem 9.2.1

Suppose A is a commutative unital Banach Algebra.

(i) $\triangle \neq \emptyset$

(ii) J is a maximal ideal in A if and only if $J = \ker\, h$ for some $h \in \triangle$

(iii) $\|h\| = 1, \forall h \in \triangle$

(iv) $\forall a \in A : \sigma(a) = \{h(a) = h \in \triangle\}$

Lemma 9.2.1

If A is a unital Banach algebra, then every proper ideal is contained in a maximal ideal and every maximal ideal is closed.

Theorem 9.2.2: Gelfand-Mazur Theorem

A unital Banach Algebra in which every nonzero element is invertible (that is, a division ring) is isometrically isomorphic to $\mathbb{C}$.

Definition 9.2.19 (Gel'fand Transform)**.**
Let A be a commutative Banach algebra with $\triangle(A)$ nonempty. The Gel'fand transform of $a \in A$ is the function:

$$\hat{a} : \triangle(A) \to \mathbb{C} \big| h \mapsto \hat{a}(h) := h(a)$$

The space $\triangle(A)$ is then called the spectrum of A and if it has an identity is also called the ***maximal ideal space of A****.*

Definition 9.2.20 (Gel'fand topology)**.**
The Gel'fand topology on $\triangle(A)$ is the smallest topology making each $\hat{a}$ continuous.

Lemma 9.2.2: Gel'fand topology

The Gelfand topology on $\triangle(A)$ is the relative topology on $\triangle(A)$ viewed as a subset of A^* with the weak-$*$ topology.

9.3 REMARKS ON GEL'FAND REPRESENTATIONS

In this section we try to explain in a bit more detail the importance of the Gel'fand definitions and theorems, not only because of their relevance in the rest of this thesis but also because of their importance as a way to generalize the space concept which is in some sense an algebraic extension. We mention that extensions of these concepts are now used in a categorial setting to attempt to re-investigate the foundations (axioms) of quantum mechanics, QFT, computer science and even mathematics itself through either topos theory or homotopy type theory. We refer the interested readers to [104, 105] and references therein.

9.3.1 *Commutative Banach Algebras*

Starting from a commutative Banach algebra A, over the field of complex numbers we defined in the previous section the characters as the non-zero algebra homomorphism

$$\phi : A \to \mathbb{C}$$

and wrote $\triangle A$ for the set of these characters of A. It can be shown that every character on A is automatically continuous, making it a subset of the space A's dual space A^* of continuous linear functionals. Providing the characters with the relative weak-* topology, $\triangle A$ is also locally compact and Hausdorff (Banach-Alaoglu theorem). If the algebra A is now also unital the local compactness property generalizes to compactness. The above definition of $\triangle A$ and the topology on it ensures that $\hat{a}$ (from Definition 9.2.19) is continuous and vanishes at infinity, furthermore it ensures that $a \mapsto \hat{a}$ defines a norm-decreasing, unit-preserving algebra homomorphism from A to the continuous linear functionals on A.

Using the Gel'fand-Mazur theorem (Theorem 9.2.2) one can then demonstrate that there is a bijection between $\triangle A$ and the set of maximal proper ideals in A.

Example 9.3.1.

- $A = L^1(\mathbb{R})$, *the group algebra of* $\mathbb{R}$, *then* $\triangle A$ *is homeomorphic to* $\mathbb{R}$ *and the Gel'fand transform of* $f \in L^1(\mathbb{R})$ *is the Fourier transform* $\tilde{f}$.
- $A = L^1(\mathbb{R}^+)$, *the* L^1*-convolution algebra of the real half-line, then* $\triangle A$ *is homeomorphic to* $\{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\}$, *and the Gel'fand transform of an element* $f \in L^1(\mathbb{R}^+)$ *is the Laplace transform.*

9.3.2 C^*-*algebras*

Things get very interesting in the case that A is not just a commutative algebra, but a C^*-algebra. Consider as an example $A = C^0(X)$. In this case one can show that we may identify $\triangle A$ with X, not just as sets but as topological spaces. The Gel'fand representation is then an isomorphism!. Put differently we turned the algebra in a topological space, which has very nice properties we described in the commutative algebra case. If now A is a separable C^*-algebra, the weak-* topology is metrizable on bounded subsets. It then follows that the spectrum of a separable commutative C^*-algebra can be regarded as a **metric space**. So the topology can be characterized via convergence of sequences. It will be exactly this kind of construction we will need to define GLS.

The above statement can also be translated into a categorial setting where for C^*-algebras with unit, the spectrum map gives rise to a contra-variant functor from the category of C^*-algebras with unit and unit-preserving continuous *-homomorphisms, to the category of compact Hausdorff spaces and continuous maps.

We do not go deeper into this subject here, but we hope we have convinced the reader of the importance of the Gel'fand theorems.

10 CATEGORIES, FUNCTORS AND DIFFERENTIATIONS

10.1 CATEGORY

Although this chapter will at first seem a bit out of place, it will become very relevant when considering shuffle products (see Chapter 11) and their algebra allowing us to prove some strong properties.

Definition 10.1.1 (Category).
A category C consists of:

(i) a class $\mathrm{ob}(C)$ *of objects*

(ii) a class $\mathrm{Hom}(C)$ *of morphisms that one can interpret as arrows or maps between the objects. The morphism* f *has a unique source object* a *and target object* b *with*

$$a, b \in \mathrm{ob}(C) | f : a \to b.$$

(iii) a binary operation

$$\mathrm{Hom}(a, b) \times \mathrm{Hom}(b, c) \to \mathrm{Hom}(a, c)$$

called composition of morphisms

such that we have:

(iv) (associativity) if $f : a \to b, g : b \to c$ *and* $h : c \to d$ *then* $h \cdot (g \cdot f) = (h \cdot g) \cdot f$

(iiv) (identity) $\forall x \in \mathrm{ob}(C), \exists\ 1 \in \mathrm{ob}(C) |\ 1x : x \mapsto x$

With these properties it is easy to show that there is a unique identity map.

10.2 FUNCTORS

Definition 10.2.1 (Functor).
Let C and D be categories. A functor F from C to D is a mapping with the following properties

(i)

$$\forall\ X \in C, \exists\ Y \in D | F : X \mapsto Y = F(X)$$

(ii)

$$\forall\ f : X \mapsto Y \in C, \exists\ F(f) : F(X) \mapsto F(Y) \in D$$

for a covariant functor and

$$\forall\ f : X \mapsto Y \in C, \exists\ F(f) : F(Y) \mapsto F(X) \in D$$

for a contravariant functor.

such that the identity and composition of morphisms are preserved i.e.:

(a)

$$C \ni 1_C \mapsto F(1_C) = 1_D \in D$$

(b)

$$F(f \circ g) = F(f) \circ F(g)$$

for a covariant functor or

$$F(f \circ g) = F(g) \circ F(f)$$

for a contravariant functor.

A functor

$$F : C \to D$$

is ***full (resp. faithful, fully faithful)*** *if, for all objects a and b of C, the map*

$$\mathrm{H}om(a, b) \to \mathrm{H}om(F(a), F(b))$$

is surjective (resp. injective, bijective).

Definition 10.2.2 (Forgetful Functor).
Let C and D be categories such that the object $c \in C$ can be regarded as an object of D by ignoring some of the mathematical structure of c. A functor $U : C \to D$ which in its operation forgets about any imposed mathematical structure is called a forgetful functor.

Example 10.2.1 (Forgetful Functor).
The following are examples of forgetful functors:

- $U : \mathbf{Grp} \to \mathbf{Set}$ *takes groups into their underlying sets and group homomorphisms to set maps.*
- $U : \mathbf{Top} \to \mathbf{Set}$ *takes topological spaces into their underlying sets and continuous maps to set maps.*
- $U : \mathbf{Ab} \to \mathbf{Grp}$ *takes abelian groups to groups and acts as identity on arrows.*

10.3 DIFFERENTIATIONS

Definition 10.3.1 (k-module differentiation).
A differentiation of a k-module $\mathfrak{U}$ is a morphism of k-modules $\mathfrak{U}, \Omega$

$$d : \mathfrak{U} \to \Omega$$

where Ω is also a $\mathfrak{U}$-module such that d obeys the Leibniz rule

$$\forall f, g \in \mathfrak{U} : d(fg) = gdf + fdg.$$

Definition 10.3.2 (Category of k-module differentiations).
Let $\mathfrak{U}$ and $\mathfrak{U}'$ be k-algebras,

$$d : \mathfrak{U} \to \Omega$$

a differentiation of $\mathfrak{U}$ *and*

$$d' : \mathfrak{U}' \to \Omega'$$

a differentiation of $\mathfrak{U}'$. *Define*

$$\mathrm{D}iff(d, d')$$

as the totality of pairs

$$(\tilde{\phi}, \hat{\phi})$$

with

$$\tilde{\phi} \in \mathrm{A}lg(\mathfrak{U}, \mathfrak{U}')$$

and

$$\hat{\phi} \in \mathrm{H}om_k(\Omega, \Omega')$$

such that

$$d'\tilde{\phi} = \hat{\phi}d,$$

which makes the diagram representing these maps a commutative diagram, and

$$f \in \mathfrak{U}, w \in \Omega \Rightarrow \hat{\phi}(fw) = (\tilde{\phi}f)\hat{\phi}w.$$

We write $\mathcal{D}$ *for the category of differentiations of commutative unitary k-algebras with the category morphisms defined as above.*

Definition 10.3.3 (Pointed Differentiation)**.**
A pointed differentiation is a pair (d, p) *where*

$$d : \mathfrak{U} \to \Omega$$

is a differentiation and

$$p \in \mathrm{A}lg(\mathfrak{U}, k).$$

Write $\mathcal{PD}$ *for the category of pointed differentiations where the morphisms*

$$(d, p) \to (d', p')$$

are given by a pair

$$(\tilde{\phi}, \hat{\phi}) \in \mathrm{D}iff(d, d')$$

such that

$$p = p'\tilde{\phi}$$

and write for this category $\mathrm{D}iff(d, p; d'p')$. *The category morphisms then define* ***equivalences*** *of differentiations.*

When we will consider paths the fact that

$$p = p'\tilde{\phi},$$

will guarantee the uniqueness of the initial point of the path.

Definition 10.3.4 (Surjective Pointed Differentiation)**.**
We call a pointed differentiation (d, p) *is surjective if* d *is surjective.*

Definition 10.3.5 (Exact sequence [106])**.**
A sequence of homomorphisms of k-modules

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is said to be exact at B if

$$\mathrm{I}m\ f = \ker\ g.$$

A sequence of homomorphisms

$$A_0 \xrightarrow{f_0} A_1 \xrightarrow{f_1} A_2 \xrightarrow{f} \cdots \xrightarrow{f_{n-1}} A_n$$

is called an exact sequence if every term except the first and the last are exact. A ***five*** *term exact sequence*

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$$

is said to be ***short*** *exact.*

Remark 10.3.1 ([106]).

(i) *When $A = 0$, the sequence*

$$0 \xrightarrow{f} B \xrightarrow{g} C$$

is exact iff

$$\ker\ g = \mathrm{I}m\ f = 0$$

i.e., g is injective.

(ii) *Similarly, when $C = 0$, the sequence*

$$A \xrightarrow{f} B \xrightarrow{g} 0$$

is exact iff

$$\mathrm{I}m\ f = \ker\ g = B.$$

Definition 10.3.6 (Splitting Pointed Differentiation).
We call the pointed differentiation (d, p) splitting if as a k-module:

$$\mathfrak{U} = \ker\ d \oplus \ker\ p$$

which means that (d, p) is only splitting if and only if

$$\ker\ d \cap \ker\ p = 0.$$

(d, p) is splitting and surjective if and only if

$$0 \to k \xrightarrow{u} \mathfrak{U} \xrightarrow{d} \Omega \to 0$$

is a short exact sequence (u is the map $k \to A$ from the definition of an algebra A, see Definition 8.3.2). We write $\mathcal{SPD}$ for the subcategory of splitting surjective differentiations of the category $\mathcal{PD}$.

Writing this explicitly gives some intuition on how this works.

$$u_p \in \ker\ p, u_d \in \ker\ d : \begin{cases} d(u_p \oplus u_d) & = d(u_p) = d(p^{-1}(0_k)) \\ p(u_p \oplus u_d) & = p(u_d) = p(d^{-1}(0_\Omega)) \end{cases}$$

We already mention here that these differentiations were used by Chen [107] to introduce a generalization of the intuitive paths and loops in a manifold to algebraic paths and loops. These algebraic paths generalize the intuitive paths much like distributions generalize functions in real calculus. The fact that a mathematical consistent structure defining these algebraic paths exists is essential in the construction of GLS.

Part III

GENERALIZED LOOP SPACE

The general accepted mathematical model to describe gauge theories are **principal fibre bundles**, where the gauge fields (or potentials) are identified with sections of a connection one-form in the gauge bundle. The gauge potentials give rise to a **parallel transport equation** in the gauge bundle that can be solved by using product integrals, first introduced by Volterra. In this part we shall demonstrate that the solution of the parallel transport equation can be presented as a Wilson line. We will also demonstrate its relation to the standard covariant derivative in gauge theories.

An alternative way to construct a gauge theory is to use the holonomies in the gauge bundle instead of the gauge potentials. This approach is allowed by the Ambrose-Singer theorem that claims that the holonomies contain the same information as the gauge potential curvatures. However, in this setting there are issues with over-completeness, re-parametrization invariance and additional algebraic constraints coming from the matrix representation of the Lie algebra associated with the gauge group. These issues make straightforward application of the standard loop space approach to gauge field theories impossible. An interesting (partial) solution to these problems arises if one extends this setting to so-called **generalized loops**, first proposed by Chen and further studied by Gambini et al and Tavares [16, 108, 109]), thus introducing the GLS approach. We will follow the framework developed by these authors in our exposition of GLS.

11 THE SHUFFLE ALGEBRA

11.1 INTRODUCTION

For the moment it is sufficient to describe an n-dimensional *manifold* as a topological space, in which a neighborhood of each point is equivalent (strictly speaking, homeomorphic) to the n-dimensional Euclidean space. The fundamental geometrical object in a manifold we will be concerned about is a path. One has a natural intuitive idea of what a path or a loop in a manifold is. Mathematically one usually defines a path γ in a manifold M as the map

$$\gamma : [0,1] \to M,\ t \mapsto \gamma(t).$$

For closed paths, which are called loops, one just adds the extra condition that the initial and final points of the path coincide

$$\gamma(0) = \gamma(1) \in M.$$

The notions of paths and loops can be generalized to the so-called algebraic d-paths, where the d-paths are **algebraic objects** constructed in such a way to possess certain desirable properties. The resulting algebraic structure can then be equipped with a topology, turning it into a **topological algebra**. The topology is used to complete the algebraic properties with analytic ones, allowing one to introduce the necessary differential operators[1].

The previous part reviewed some mathematical concepts that we will now use and extend to introduce and study the above mentioned generalizations of paths and loops.

11.2 SHUFFLE ALGEBRA

Generalization of the concepts of paths and loops require a new algebra, the shuffle algebra. This algebra is constructed from a new product, the shuffle product, that is defined through the notion of $(k,l)-$shuffles. Let us start with defining these shuffles.

1 Most of the material in this Chapter is based on the original works by Chen [107, 108, 110, 111], where the proofs to a number of the stated theorems can also be found. We skip those proofs which do not bring more insight than needed into the subject of this text.

Definition 11.2.1 ((k,l)-Shuffle)**.**
A (k,l)-shuffle is a permutation σ of the $k+l$ letters such that

$$\sigma(1) < \cdots < \sigma(k)$$

and

$$\sigma(k+1) < \cdots < \sigma(k+l).$$

These shuffles can be traced back to shuffling cards and can also be used in the definition of the wedge product [106]

$$f \wedge g\ (v_1, \cdots, v_{k+l}) = \sum_{\sigma_{k,l}} (\operatorname{sgn} \sigma) f(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \cdots, v_{\sigma(k+l)}).$$

Using these (k,l)-shuffles we define a new type of product called the shuffle multiplication, symbolically represented by "$\bullet$".

Definition 11.2.2 (Shuffle Multiplication)**.**
Using the notations

$$\omega_1 \cdots \omega_k = \omega_1 \otimes \cdots \otimes \omega_k \in \bigotimes^k \bigwedge^1 M\ ,\ \ k \geq 1$$

and

$$\omega_1 \cdots \omega_k = 0 \text{ for } k = 0.$$

the shuffle multiplication is given by

$$\omega_1 \cdots \omega_k \bullet \omega_{k+1} \cdots \omega_{k+l} = \sum_{\sigma_{k,l}} \omega_{\sigma(1)} \cdots \omega_{\sigma(k+l)}$$

where $\sum_{\sigma_{k,l}}$ denotes the sum over all (k,l)-shuffles and $\bigwedge^1$ is the set of one-forms defined on the manifold M.

The examples below are quite instructive if one is not familiar with the shuffle product.

Example 11.2.1 (Shuffle Multiplication)**.**

- *Two objects:*

$$\omega_1 \bullet \omega_2 = \omega_1\omega_2 + \omega_2\omega_1$$

- *Three objects:*

$$\omega_1 \bullet \omega_2\omega_3 = \omega_1\omega_2\omega_3 + \omega_2\omega_1\omega_3 + \omega_2\omega_3\omega_1$$

- *Four objects:*

$$\begin{aligned}\omega_1\omega_2 \bullet \omega_3\omega_4 = &\ \omega_1\omega_2\omega_3\omega_4 + \omega_1\omega_3\omega_2\omega_4 + \omega_1\omega_3\omega_4\omega_2 \\ &+ \omega_3\omega_1\omega_2\omega_4 + \omega_3\omega_1\omega_4\omega_2 + \omega_3\omega_4\omega_1\omega_2\end{aligned} \tag{59}$$

The shuffle product $\bullet$, defined as multiplication on the set of forms of a manifold, can be considered as the symmetric version of the wedge product. Now let M be a manifold and

$$\Omega = \bigwedge^1 M$$

be the set of one-forms on M. We interpret Ω as a k-module, where for the moment we assume that k is a general ring of scalars with a multiplicative unity. Introducing the shuffle product on a k-module Ω defines the shuffle k-algebra, which is an alternative algebra to the tensor algebra, defined by the tensor product $\otimes$.

Definition 11.2.3 (Shuffle k-Algebra)**.**
Let Ω be a k-module (8.1.5) and let $T(\Omega)$ be the regular tensor algebra over k based on Ω. $T^r(\Omega)$ represents the degree r components of the algebra (note that $T^0(\Omega) = k$). Changing the regular tensor multiplication for the shuffle multiplication defined in definition 11.2.2 one gets a new algebra which is called the shuffle k-algebra $Sh(\Omega)$ based on the k-module Ω.

In this algebra the multiplication m (See Definition 8.3.2) is identified with the shuffle product such that we can write

$$m = \bullet : \mathrm{Sh} \to \mathrm{Sh}$$

and the algebra unit map u is defined by

$$u : k \to \mathrm{Sh}, \mathbb{1}_k \mapsto \mathbb{1}_{\mathrm{Sh}}.$$

We can now extend the algebraic structure of the k-shuffle algebra by introducing the k-linear maps ϵ, Δ.

Definition 11.2.4 (Co-unit and Co-multiplication).

$$\epsilon \in \mathrm{Alg}(\mathrm{Sh}(\Omega), k) : \begin{cases} \epsilon(1) = 1 & r = 0 \\ \epsilon(\omega_1 \cdots \omega_r) = 0 & r > 0 \end{cases} \tag{60}$$

$$\Delta : \mathrm{Sh}(\Omega) \to \mathrm{Sh}(\Omega) \otimes \mathrm{Sh}(\Omega) : \begin{cases} \Delta(1) = 1 & r = 0 \\ \Delta(\omega_1 \cdots \omega_r) = & \\ \quad \sum_{i=0}^{r} (\omega_1 \cdots \omega_i) \otimes (\omega_{i+1} \cdots \omega_r) & r > 0 \end{cases} \tag{61}$$

The map Δ is an associative co-multiplication

$$(1 \otimes \Delta)\Delta = (\Delta \otimes 1)\Delta,$$

but can also be considered as a k-module morphism. By defining the above operations on the k-shuffle algebra we have equipped it with a co-algebra structure, turning it into a bi-algebra (See Definitions (8.3.8) and (8.4.1)). This bi-algebra becomes a Hopf-algebra after introduction of the antipode map J, defined as

Definition 11.2.5 (Antipode).
The shuffle algebra antipode is a k-linear map

$$J : \mathrm{Sh} \to \mathrm{Sh}$$

such that

$$J(\omega_1 \cdots \omega_r) = (-1)^r \omega_r \cdots \omega_1. \tag{62}$$

It is easy to demonstrate that the antipode J has the following properties

$$\begin{aligned} &\bullet \circ (J \otimes \mathbb{1}) \circ \Delta = \bullet \circ (\mathbb{1} \otimes J) \circ \Delta = \eta \circ \epsilon \\ &J(\mathbf{u_1} \bullet \mathbf{u_2}) = J(\mathbf{u_2}) \bullet J(\mathbf{u_1}) \\ &J(\mathbb{1}) = \mathbb{1}\ ,\ J^2 = \mathbb{1}\ ,\ \epsilon \circ J = \epsilon \\ &\tau \circ (J \otimes J) \circ \Delta = \Delta \circ J \end{aligned} \tag{63}$$

$$\forall \mathbf{u_1}, \mathbf{u_2} \in \mathrm{Sh} \ ,$$

where

$$\bullet : \mathrm{Sh} \otimes \mathrm{Sh} \to \mathrm{Sh}$$

denotes shuffle multiplication and

$$\eta : k \to \mathrm{Sh}$$

the unit map (to avoid confusing with $\boldsymbol{u} \in \mathrm{Sh}$). The map

$$\tau : \mathrm{Sh} \otimes \mathrm{Sh} \to \mathrm{Sh} \otimes \mathrm{Sh}$$

is called the *transposition map* or *flipping operation* defined as

$$\tau(\mathbf{u_1} \otimes \mathbf{u_2}) = \mathbf{u_2} \otimes \mathbf{u_1} \ . \tag{64}$$

Therefore, we have the following Theorem

Theorem 11.2.1: $Sh(\Omega)$ is Hopf

The shuffle algebra $Sh(\Omega)$ is a Hopf k-algebra with comultiplication Δ and counit ϵ as defined in (60) and (61) .

Having now discussed the algebraic structure of the shuffle algebra we can change our focus to study the algebra morphisms $\mathrm{Alg}(\mathrm{Sh}(\Omega), k)$.

Definition 11.2.6 (Group Multiplication on $\mathrm{Alg}(\mathrm{Sh}(\Omega), k)$).
Let

$$\alpha_i \in \mathrm{Alg}(\mathrm{Sh}(\Omega), k)$$

be algebra homomorphisms and let us define a multiplication on them

$$\alpha_1\alpha_2 \in \mathrm{Alg}(\mathrm{Sh}(\Omega), k)$$

by

$$\alpha_1\alpha_2 := (\alpha_1 \otimes \alpha_2)\Delta.$$

For this multiplication we have

$$\epsilon\alpha_1 = \alpha_1\epsilon = \alpha_1$$

and

$$\begin{aligned}\alpha_1(\alpha_2\alpha_3) &= (\alpha_1 \otimes \alpha_2 \otimes \alpha_3)(1 \otimes \Delta)\Delta \\ &= (\alpha_1 \otimes \alpha_2 \otimes \alpha_3)(\Delta \otimes 1)\Delta \\ &= (\alpha_1\alpha_2)\alpha_3.\end{aligned}$$

We can represent this multiplication diagrammatically as shown in Fig. 23.

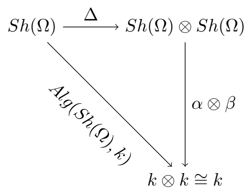


Figure 23: Multiplication of algebra morphisms.

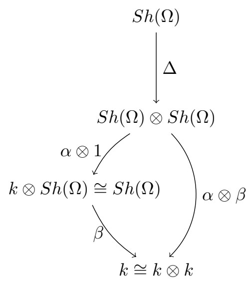


Figure 24: Multiplication of algebra morphisms.

Proposition 11.2.1

The multiplication defined in Definition 11.2.6 turns $\mathrm{Alg}(\mathrm{Sh}(\Omega), k)$ into a group.

Considering a categorial interpretation we have the following proposition:

Proposition 11.2.2

The map

$$\mathrm{Sh} : \Omega \to Sh(\Omega)$$

can be considered as a covariant functor to the category of Hopf k-algebras on the category of k-modules.

So far we have discussed algebra morphisms of the type $\mathrm{Alg}(Sh(\Omega), k)$. Let us now move on to consider morphisms of the type $\mathrm{Alg}(Sh(\Omega), Sh(\Omega))$. To this end, we define such a morphism which might look strange for the moment but will allow us to demonstrate the translational invariance of the group structure of GLS.

Definition 11.2.7 (L-operator).
For
$$\alpha \in \mathrm{Alg}(Sh(\Omega), k)$$
define
$$\tilde{L}_\alpha = (\alpha \otimes 1)\Delta \in \mathrm{Alg}(Sh(\Omega), Sh(\Omega)) \tag{65}$$
and
$$\hat{L}_\alpha = \tilde{L}_\alpha \otimes 1 \in \mathrm{Hom}(Sh(\Omega) \otimes \Omega, Sh(\Omega) \otimes \Omega). \tag{66}$$

Proposition 11.2.3: L-operator composition

For
$$\beta \in \mathrm{Alg}(Sh(\Omega), k)$$
we have by Definition 11.2.6 that
$$\beta \tilde{L}_\alpha = \alpha\beta$$
(see also figure 24) and
$$(\tilde{L}_{\alpha\beta}, \hat{L}_{\alpha\beta}) = (\tilde{L}_\beta \tilde{L}_\alpha, \hat{L}_\beta \hat{L}_\alpha).$$

proof 11.2.1: L-operator compostion

The above proposition is easily demonstrated by
$$\begin{aligned}
\tilde{L}_{\alpha\beta} &\overset{def}{=} (\alpha\beta \otimes 1)\Delta \\
&\overset{def}{=} ((\alpha \otimes \beta)\Delta \otimes 1)\Delta \\
&= (\alpha \otimes \beta \otimes 1)(\Delta \otimes 1)\Delta \\
&= (\alpha \otimes \beta \otimes 1)(1 \otimes \Delta)\Delta \\
&= (\alpha \otimes \tilde{L}_\beta)\Delta \\
&= \tilde{L}_\beta(\alpha \otimes 1)\Delta \\
&= \tilde{L}_\beta \tilde{L}_\alpha
\end{aligned}$$
such that also
$$\hat{L}_{\alpha\beta} = \tilde{L}_{\alpha\beta} \otimes 1 = \hat{L}_\beta \hat{L}_\alpha$$

Note that $\tilde{L}_\epsilon$ and $\hat{L}_\epsilon$ are identity morphisms on $Sh(\Omega)$ and $Sh(\Omega) \otimes \Omega$ respectively.

11.3 SHUFFLE DIFFERENTIATIONS

In the next Chapter we shall discuss the generalized or algebraic paths and loops, which are based on shuffle algebra morphisms. Because we are ultimately interested in the mathematically consistent formalism for variations of these paths and loops, we need the differentiations, introduced in Section 10.3, to be well-defined. To this end we study the application of the k-module differentiations to the shuffle algebra. Applying definition 10.3.1 with

$$\mathfrak{U} = Sh(\Omega)$$

and k-module Ω we get the following map

$$Sh(\Omega) \xrightarrow{d} Sh(\Omega) \otimes \Omega$$

where Ω now also is a $Sh(\Omega)$-module (shown explicitly in the map above) and the module conditions (Definition 8.1.5) can be written as:

(i) $\omega(w_1 +_\Omega w_2) = \omega w_1 +_\Omega \omega w_2$

(ii) $(\omega_1 +_{Sh(\Omega)} \omega_2)w = \omega_1 w +_\Omega \omega_2 w$

(iii) $(\omega_1 \bullet \omega_2)w = \omega_1 w \cdot_\Omega \omega_2 w$

(iv) $\mathbb{1}_{Sh(\Omega)} w = w$,

with

$$\omega_i \in Sh(\Omega), \forall i$$

and

$$w_i \in \Omega, \forall i.$$

Definition 11.3.1 (Surjective Shuffle Module Differentiation).
Considering now the k-module

$$Sh(\Omega) \otimes \Omega$$

as a $Sh(\Omega)$*-module*

$$u \bullet (v \otimes w) = (u \bullet v) \otimes w, \quad u, v \in Sh(\Omega), w \in \Omega \tag{67}$$

we define the surjective differentiation (Definition 10.3.6)

$$\delta \in \mathrm{Hom}(Sh(\Omega), Sh(\Omega) \otimes \Omega)$$

as

$$\delta(\mathbb{1}) = 0$$
$$\delta(uw) = u \otimes w$$

where

$$u \in Sh(\Omega) \text{ and } w \in \Omega.$$

The example below of a shuffle product of tensor products is very instructive and will be used to prove that δ is a differentiation.

Example 11.3.1 (Shuffle product of tensor products)**.**
Let

$$u_1, u_2 \in T^1(\Omega)$$

and

$$w_1, w_2 \in T^1(\Omega)$$

such that we have

$$\begin{aligned}(u_1 w_1) \bullet (u_2 w_2) &= u_1 w_1 u_2 w_2 + u_1 u_2 w_1 w_2 \\ &\quad + u_1 u_2 w_2 w_1 + u_2 u_1 w_2 w_1 \\ &\quad + u_2 w_2 u_1 w_1 + u_2 u_1 w_1 w_2 \\ &= (u_1 w_1 \bullet u_2) w_2 + (u_2 w_2 \bullet u_1) w_1\end{aligned} \tag{68}$$

Theorem 11.3.1

δ is a differentiation.

proof 11.3.1

Let
$$u_1, u_2 \in T(\Omega)$$
and
$$w_1, w_2 \in T^1(\Omega)$$
so that we have

$$(u_1 w_1) \bullet (u_2 w_2) = (u_1 w_1 \bullet u_2) w_2 + (u_2 w_2 \bullet u_1) w_1$$

by the properties of the shuffle multiplication as is discussed above. Applying δ results in

$$\begin{aligned} \delta\left((u_1 w_1) \bullet (u_2 w_2)\right) &= (u_1 w_1 \bullet u_2) \otimes w_2 + (u_2 w_2 \bullet u_1) \otimes w_1 \\ &= (u_1 w_1 \bullet (u_2 \otimes w_2) + (u_2 w_2 \bullet (u_1 \otimes w_1) \quad \text{by definition (11.3.1)} \\ &= (u_1 w_1) \bullet \delta(u_2 w_2) + (u_2 w_2) \bullet \delta(u_1 w_1) \qquad (69) \end{aligned}$$

Thus δ obeys the Leibniz rule showing that δ is indeed a differentiation, according to (Definition 10.3.1).

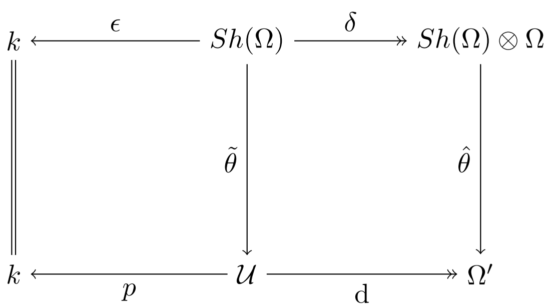


Figure 25: Splitting pointed differential.

For such differentiations we have the following Lemma.

Lemma 11.3.1: Splitting pointed differentiation homomorphism

Let (d, p) be a splitting pointed differentiation (Definition 10.3.6) and let there be a commutative diagram of k-module morphisms as shown in rigure 25 (the double line between the k's indicates that their values are equal, and the double arrow heads indicate that the differentiations are surjective). Defining

$$\tilde{\theta}(\mathbb{1}) = \mathbb{1}$$

and

$$\hat{\theta}(u \otimes w) = (\tilde{\theta}u)\hat{\theta}(\mathbb{1} \otimes w), \ \forall u \in Sh(\Omega), w \in \Omega$$

we obtain

$$(\tilde{\theta}, \hat{\theta}) \in \mathrm{Diff}(\delta, \epsilon; d, p).$$

Let now

$$\theta \in \mathrm{Hom}(\Omega, \Omega'),$$

from which we have an induced homomorphism between the tensor algebras $T(\Omega)$ and $T(\Omega')$ respectively, written symbolically as $T(\theta)$. Since the tensor algebra morphism is shuffle product preserving, we can also write it as $\mathrm{Sh}(\theta)$. Considering the category of k-modules we can define the following functor.

Definition 11.3.2 (Covariant Functor to $\mathcal{SPD}$)**.**
Write Δ_f for the covariant functor (Definition 10.2.1) to the category of Splitting Pointed Differentiations (Definition 10.3.6) on the category of k-modules such that

$$\Delta_f(\Omega) = (\delta, \epsilon) = (\delta(\Omega), \epsilon(\Omega))$$

and for

$$\theta \in \mathrm{Hom}(\Omega, \Omega') : \Delta_f(\theta) = (\mathrm{Sh}(\theta), \mathrm{Sh}(\theta) \otimes \theta).$$

Important here is to keep track of the (possibly) confusing notations. Here $\mathrm{Sh}(\theta)$ is the image of a morphism in the category of k-modules in the category of splitting pointed differentiations defined on that k-module under the functor Δ_f, thus $\mathrm{Sh}(\theta)$ is the first part of the morphism $(\mathrm{Sh}(\theta), \mathrm{Sh}(\theta) \otimes \theta)$ in the category of splitting pointed differentiations. A schematic representation of the above definition is shown in Fig. 26.

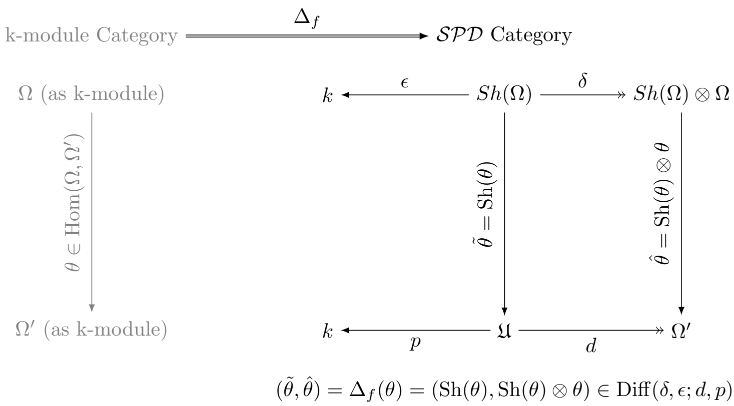


Figure 26: Covariant functor to the category of $\mathcal{SPD}$.

Theorem 11.3.2: Uniqueness

Let Ω be a k-module and (d', p') a splitting surjective pointed differentiation with

$$d' : \mathcal{U} \to \Omega'$$

then given

$$\theta \in \mathrm{Hom}(\Omega, \Omega'),$$

there exists a unique

$$(\tilde{\theta}, \hat{\theta}) \in \mathrm{Diff}(\delta, \epsilon; d', p')$$

such that

$$\forall w \in \Omega : \hat{\theta}(\mathbb{1} \otimes w) = \theta w.$$

This shows that Δ_f is an adjoint to the forgetful functor (Definition 10.2.2) to the category of k-modules on the category of Splitting Pointed Differentiations which assigns to each (d, p) the k-module Ω and to each

$$(\tilde{\phi}, \hat{\phi}) \in \mathrm{Diff}(d, p; d', p')$$

the morphism ϕ of k-modules. Taking it one step further by using the fact that

$$Sh(\Omega) \otimes Sh(\Omega) \otimes Sh(\Omega)$$

is a

$$Sh(\Omega) \otimes Sh(\Omega)$$

module and that

$$\epsilon \otimes \epsilon \in \mathrm{Alg}(Sh(\Omega) \otimes Sh(\Omega), k)$$

we arrive at the following Lemma.

Lemma 11.3.2

The morphism of k-modules

$$\mathbb{1} \otimes \delta : Sh(\Omega) \otimes Sh(\Omega) \to Sh(\Omega) \otimes Sh(\Omega) \otimes \Omega$$

is a differentiation, and

$$(\mathbb{1} \otimes \delta, \epsilon \otimes \epsilon)$$

is a splitting surjective pointed differentiation.

We end our treatment of shuffle algebras and their differentiations with property of the L-operator, defined in 11.2.3, with respect to the category of differentiations on the shuffle algebra.

Proposition 11.3.1

$(\tilde{L}_\alpha, \hat{L}_\alpha)$ is an equivalence in the category $\mathcal{D}$ ($=$ differentiations), i.e.

$$\hat{L}_\alpha \delta = \delta \tilde{L}_\alpha.$$

11.4 SUMMARY

In this chapter we used the concepts introduced in the mathematical preliminaries part of this text to first introduce an new product, the shuffle product and then derive some of its algebraic properties. In the last part we have studied the behavior of the algebraic structure generated by the shuffle product under application of some differentiations. These differentiations will be used in the next chapter to define d-paths and d-loops.

12 ALGEBRAIC PATHS

12.1 INTRODUCTION

In this chapter we introduce the so called **d-paths** as discussed by Chen in [107]. These algebraic paths are a generalization of the usual paths

$$p : [0, 1] \to \gamma \subset M$$

in a manifold M. Naturally d-loops are the generalization of the usual loops. As before we only included proofs of Lemmas and Theorems that we consider to add some insight into the subject.

12.2 ALGEBRAIC PATHS

The whole concept of the d-paths is captured in Figure 27, where the properties of the shuffle product and algebra on the k-module of one-forms on a manifold M are combined into a consistent mathematical structure that allows for the generalization of paths on M, algebraic paths or d-paths.

We first give a brief overview of the maps shown in Fig. 27 and a brief discussion on how they generate the d-paths. After this overview we will discuss their properties in more detail.

The essential part of this structure is a given pointed differentiation (d, p), which is mapped to the pointed differentiation (δ, ϵ) by the equivalence of differentiations we introduced in Definition 10.3.3. The δ in the Figure is the differentiation introduced in Definition 11.3.1, and ϵ is the co-unit from the co-algebra structure on $Sh(\Omega)$. The existence of the ideal $I = I(d, p)$ for $Sh(\Omega)$, which we will discuss below, and (δ_1, ϵ_1) the pointed differentiation induced by (δ, ϵ) after dividing out this ideal will allow us to factorize d-paths through this ideal (See the discussion of Figures 16 and 17). $\tilde{\chi}_0$ and $\hat{\chi}_0$ are the k-module morphisms as defined in Lemma 11.3.1 such that

$$\tilde{\chi}_0 f = pf + df$$

and

$$\hat{\chi}_0 w = \mathbb{1}_{\mathrm{Sh}} \otimes w.$$

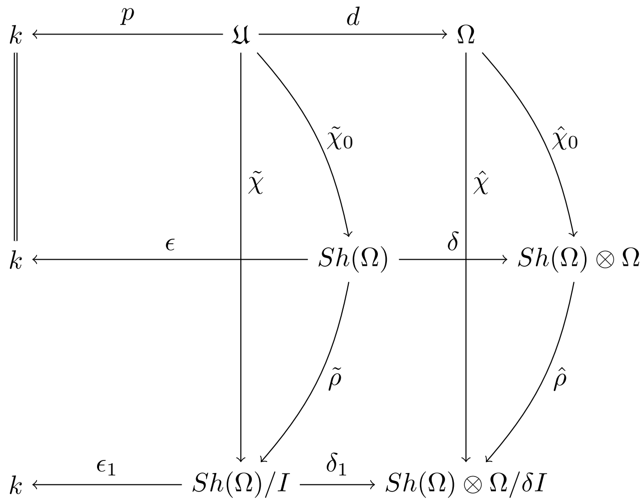


Figure 27: Path diagram.

$\tilde{\rho}$ and $\hat{\rho}$ are canonically induced by dividing out the ideal. The maps in the diagram defined as above, allow to define a d-path from p as a k-algebra morphism

$$Sh(\Omega) \to k$$

that can be factored through $Sh(\Omega)/I$ (or more correctly through $\tilde{\rho}$). From the diagram the importance of the ideal $I(d, p)$ should be clear. Before discussing this ideal in more detail we need to introduce one more type of differentiation that is defined using an ideal.

Definition 12.2.1 (d-closed differentiation)**.**
Let

$$d : \mathfrak{U} \to \Omega$$

be a differentiation. An ideal J of $\mathfrak{U}$ is called d-closed if dJ is a $\mathfrak{U}$-submodule of Ω and if

$$J\Omega \subset dJ.$$

If J is a d-closed ideal for $\mathfrak{U}$, then d induces a differentiation

$$d_J : \mathfrak{U}/J \to \Omega/d_J. \tag{70}$$

Let us explain this definition in more detail. The set $\mathfrak{U}$ is a k-module, such that $(\mathfrak{U}, +)$ is an Abelian group and we can use elements of k as "scalars" (see Definition 8.1.5) which we can multiply with elements of $\mathfrak{U}$. This multiplication can be written as the map

$$k \times \mathfrak{U} \to \mathfrak{U}.$$

Similarly $(\Omega, +)$ is an abelian group but also a $\mathfrak{U}$-module such that the elements of $\mathfrak{U}$ now act as "scalars". The differentiation d then takes the ideal J to a subset of Ω making it a $\mathfrak{U}$-(sub)module

$$\mathfrak{U} \times dJ \to \Omega.$$

The term "closed" refers then to the fact that

$$J\Omega \subset dJ \text{ in } \Omega$$

where we the elements of J act as "scalars" being multiplied to the elements of Ω.

We have seen before that kernels of homomorphism generate ideals. The next proposition relates the d-closed ideal to a homomorphism so that the ideal itself is generated by the kernel of the morphism.

Proposition 12.2.1

Let (d, p) and (d', p') be pointed differentiations such that (d, p) is surjective and (d', p'), splitting. If

$$\tilde{\phi}, \hat{\phi} \in \mathrm{Diff}(d, p; d', p')$$

then

$$\ker\ \tilde{\phi}$$

is a d-closed ideal of $\mathfrak{U}$.

Therefore J is generated by $\ker\ \tilde{\phi}$.

12.3 CHEN ITERATED INTEGRALS: INTRODUCTION

Having discussed ideals and their relation to homomorphisms, we would like to investigate the structure of such ideals. To study these ideals in more detail we need an extension of the definition of line integrals, which were introduced by Chen and thus are often referred to as **Chen Iterated Integrals.**

Definition 12.3.1 (Chen Iterated Integrals)**.**
The line integral along the path γ parametrized by s is given by

$$I_i(\gamma) = \int_a^b dx_i(s) = x_i(b) - x_i(a), \tag{71}$$

which we now extend recursively for

$$p \geq 2$$

by

$$I_{i_1 \cdots i_p}(\gamma) = \int_a^b dx_{i_p}(t)\ I_{i_1 \cdots i_{p-1}}(\gamma^t), \tag{72}$$

where γ^t represents the part of the path γ, for which the path parameter s runs from 0 to t (or, equivalently, the coordinates along the path vary from the point a to the point $\gamma(t)$).

However, this definition is coordinate dependent, which is not always desired. Therefore let us give an alternate definition that does not refer explicitly to coordinates, but only to the variable that parametrizes the path along which the integrals are evaluated.

Definition 12.3.2 (Chen Iterated Integrals without Coordinates)**.**
Let M be a smooth n-dimensional manifold and write $\mathcal{PM}$ for the set of piecewise smooth paths in M

$$\gamma : I \to M$$

where

$$I = [0, 1]$$

and

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M$$

are real-valued one-forms (respectively complex or $GL(n, \mathbb{C})$ valued). Using the notations

$$\omega_1 \otimes \cdots \otimes \omega_r = \omega_1 \cdots \omega_r,$$

$$\omega_k(t) \equiv \omega_k(\gamma(t)) \cdot \dot{\gamma}(t),$$

and

$$\gamma^t : I \to M, \ \gamma^t(s) \equiv \gamma(ts)$$

the iterated line integrals are defined in an inductive way:

$$\int_\gamma \omega_1 = \int_0^1 \omega_1(t) dt$$

$$\int_\gamma \omega_1 \omega_2 = \int_0^1 \left(\int_0^t \omega_1(s) ds \right) \omega_2(t) dt$$

$$\cdots$$

$$\int_\gamma \omega_1 \cdots \omega_r = \int_0^1 \left(\int_{\gamma^t} \omega_1 \cdots \omega_{r-1} \right) \omega_r(t) dt \tag{73}$$

The following Proposition claims that the Chen integrals are multiplication preserving, such that they can be interpreted as homomorphisms. Therefore we will be able to construct an ideal in the shuffle algebra $Sh(\Omega)$ by considering the kernel of the homomorphism

$$\omega_1 \cdots \omega_n \to \int_\gamma \omega_1 \cdots \omega_n.$$

Proposition 12.3.1: Chen Iterated Integrals preserve multiplication

Let γ be a piecewise linear path in the manifold M, i.e.

$$\gamma : [0, 1] \to M.$$

Let Ω be the set of real-valued (respectively complex or $GL(n, \mathbb{C})$ valued) one-forms on M. If we define γ as the map

$$\gamma : T(\Omega) \to \mathbb{R}(\mathbb{C}, GL(n, \mathbb{C})),$$

$$\omega_1 \cdots \omega_n \to \int_\gamma \omega_1 \cdots \omega_n,$$

then this map preserves multiplication:

$$\int_\gamma \omega_1 \cdots \omega_k \int_\gamma \omega_{k+1} \cdots \omega_{k+l} = \int_\gamma \omega_1 \cdots \omega_k \bullet \omega_{k+1} \cdots \omega_{k+l}. \tag{74}$$

In the case the one-forms are taking values in $\mathbb{C}$ or in $GL(n, \mathbb{C})$, the shuffle multiplication in the r.h.s. is replaced by the matrix multiplication, but where the multiplication of the matrix entries get multiplied by means of the shuffle multiplication. Notice that if we choose to represent a Lie algebra by a sub-algebra of the algebra of $GL(n, \mathbb{C})$ matrices, Chen integrals can be taken along horizontal lifts of paths in M when studying gauge theories in a principal fibre bundle setting (See Chapter 14). This will ultimately allow us to write down an expression for the parallel transporter in gauge theory.

Before continuing with the derivation of the shuffle ideal, let us give some examples of the above Proposition.

Example 12.3.1.

$$\begin{aligned}\int_\gamma \omega_1 \int_\gamma \omega_2 &= \int_\gamma \omega_1 \bullet \omega_2 \\ &= \int_\gamma \omega_1\omega_2 + \omega_2\omega_1 \end{aligned} \tag{75}$$

$$\begin{aligned}\int_\gamma \omega_1 \int_\gamma \omega_2\omega_3 &= \int_\gamma \omega_1 \bullet \omega_2\omega_3 \\ &= \int_\gamma \omega_1\omega_2\omega_3 + \omega_2\omega_1\omega_3 + \omega_2\omega_3\omega_2 \end{aligned} \tag{76}$$

Let us now continue with the derivation of the ideal. To this end consider the map from Proposition 12.3.1

$$\gamma : [0, 1] \to M$$

and let $f \in \mathfrak{U}$. From the definition of a line integral we can then write

$$f(\gamma(t)) = f(\gamma(0)) + \int_0^t df = pf + \int_0^t df,$$

where we defined the point evaluation map

$$pf \equiv f(\gamma(0)).$$

Using this result and partial integration we can derive the following line integral

$$\int_\gamma f\omega_1 = \int_\gamma \left(pf + \int_0^t df\right)\omega_1 = \int_\gamma df\ \omega_1 + pf\ \int_\gamma \omega_1.$$

Adding one more ω results in

$$\begin{aligned}\int_\gamma \omega_1(f\omega_2) &= \int_0^1 \left(\int_0^t \omega_1 \int_0^t df\right)\omega_2(\gamma(t))dt + pf\int_0^1\left(\int_0^t \omega_1\right)\omega_2(\gamma(t))dt \\ &= \int_\gamma (\omega_1 \bullet df)w_2 + pf\int_\gamma \omega_1\omega_2,\end{aligned}$$

where the integrals in the last line are Chen integrals. Extending this procedure, adding more and more one-forms, we arrive at the general expression

$$\begin{aligned}\int_\gamma \omega_1\cdots\omega_{i-1}(f\omega_i)\omega_{i+1}\cdots\omega_n = \int_\gamma ((\omega_1\cdots\omega_{i-1}) \bullet df)\omega_i\cdots\omega_n \\ + pf\int_\gamma \omega_1\cdots\omega_n, \qquad (77)\end{aligned}$$

where the integrals are again Chen iterated integrals as defined in Definition 12.3.2. Allowing for not only real-valued one-forms and also considering the line integrals from the endpoint $\gamma(1)$ we have the following Proposition.

Proposition 12.3.2

$$\forall f \in C^\infty M(resp., C^\infty M \otimes GL(n,\mathbb{C}))$$

and

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M$$

(resp., $\in \bigwedge^1 M \otimes GL(n,\mathbb{C})$):

$$\int_\gamma df \cdot \omega_1 \cdots \omega_r = \int_\gamma (f \cdot \omega_1)\omega_2 \cdots \omega_r - f(\gamma(0)) \cdot \int_\gamma \omega_1 \cdots \omega_r$$

$$\int_\gamma \omega_1 \cdots \omega_r \cdot df = \left(\int_\gamma \omega_1 \cdots \omega_r\right) \cdot f(\gamma(1)) - \int_\gamma \omega_1 \cdots \omega_{r-1} \cdot (\omega_r \cdot f)$$

$$\int_\gamma \omega_1 \cdots \omega_{i-1} \cdot (df) \cdot \omega_{i+1} \cdots \omega_r = \int_\gamma \omega_1 \cdots \omega_{i-1} \cdot (f \cdot \omega_{i+1}) \cdot \omega_{i+2} \cdots \omega_r - \int_\gamma \omega_1 \cdots (\omega_{i-1} \cdot f) \cdot \omega_i \cdots \omega_r$$

$$\int_\gamma \omega_1 \cdots \omega_{i-1} \cdot (f \cdot \omega_i) \cdot \omega_{i+1} \cdots \omega_r = f(\gamma(0)) \cdot \int_\gamma \omega_1 \cdots \omega_r + \int_\gamma ((\omega_1 \cdots \omega_{i-1}) \bullet df) \cdot \omega_i \cdots \omega_r$$

In Chapter 13 we will investigate more properties of these Chen integrals, but the above definitions and properties are enough to allow for the construction of the shuffle ideal. From Proposition 12.3.1 we know that the map

$$\gamma : T(\Omega) \to \mathbb{R}$$

is a homomorphism. Moving all the terms in Eq. 77 to the LHS, the RHS becomes zero. In this way the new LHS becomes an element of the kernel of the homomorphism γ. Proposition 8.2.2 now states that this homomorphism generates an ideal on $T(\Omega)$ and hence also on the shuffle algebra $Sh(\Omega)$ according to Proposition 12.3.1. This indicates that we have the ideal

Definition 12.3.3 (Shuffle Ideal).
Let

$$p \in \mathrm{Alg}(\mathfrak{U}, k)$$

and set

$$I = I(d, p)$$

the k-submodule of $Sh(\Omega)$ spanned by all

$$u(fw)v - (u \bullet df)wv - (pf)uwv \tag{78}$$

for

$$u, v \in T(\Omega),\ w \in T^1(\Omega), f \in \mathfrak{U}.$$

The following Lemma proves that this is indeed an ideal of the k-algebra $Sh(\Omega)$.

Lemma 12.3.1

The k-submodule $I(d,p)$ is an ideal of the k-algebra $Sh(\Omega)$.

See [107] for the proof. Note that $\mathbb{1} \notin I$ such that the factor algebra $Sh(\Omega)/I$ is again a commutative unitary k-algebra. Having now the ideal $I(d,p)$ on the shuffle algebra, we can define Chen's d-path [107]

Definition 12.3.4 (d-Path)**.**
A d***-path*** γ ***from*** p *is an element of* $\mathrm{Alg}(Sh(\Omega),k)$ *such that*

$$\gamma I(d,p) = \{0\}.$$

In other words a d-path is an algebra morphism from the shuffle algebra to the k-algebra that vanishes on the shuffle ideal. Returning for a moment to Chen integrals, it is clear from their relation to the shuffle algebra's ideal that if one takes such an integral over an element of the ideal $I(d,p)$, this will return zero. This not only demonstrates the link with the ideal being the kernel of the map

$$\int_\gamma$$

but that Chen integrals are also consistent with the definition of a d-path γ where one needs to have

$$\gamma(I(d,p)) = \int_\gamma (I(d,p)) = 0.$$

In other words, Chen integrals can be considered as d-paths. We shall come back to this point in Chapter 13.

Since homomorphisms preserve algebraic structure (shuffle algebra), the morphisms induced by the Chen iterated integrals can thus be considered as algebra morphisms. We write $\mathcal{A}_p$ for the resulting algebra. This leads us to the following remark, that will become relevant when introducing GLS.

Remark 12.3.1.
The kernel of the algebra map

$$Sh(\Omega) \to \mathcal{A}_p,$$

when considering closed d-paths (i.e. loops), does not only contains the ideal of the shuffle algebra but also

$$dC^\infty(M)$$

that we denote by

$$\langle dC \rangle.$$

This generates a new ideal in $Sh(\Omega)$ when considered on the space of closed paths at p

$$J_p = I_p + \langle dC \rangle$$

such that for d-loops we have the isomorphism

$$Sh(\Omega)/J_p \cong \mathcal{A}_p.$$

In Definition 12.2.1 we introduced the concept of a d-closed ideal. The next Proposition shows that the shuffle algebra ideal $I(d,p)$ is the least d-closed ideal.

Proposition 12.3.3: Least δ-Closed Ideal

I is the least δ-closed ideal of $Sh(\Omega)$ which is contained in $\ker \epsilon$ and contains all

$$fw - dfw - (pf)w \text{ for } f \in \mathfrak{U},\ w \in \Omega.$$

A proof of this Proposition can be found in [107].

We now introduce the following notation for the **canonical morphisms:**

$$\begin{aligned} \delta_I = \delta_1 = \delta_1(d,p) &: Sh(\Omega)/I \to (Sh(\Omega) \otimes \Omega)/\delta I \\ \tilde{\rho} &: Sh(\Omega) \to Sh(\Omega)/I \\ \hat{\rho} &: Sh(\Omega) \otimes \Omega \to (Sh(\Omega) \otimes \Omega)/\delta I \end{aligned} \tag{79}$$

Since $\epsilon I = 0$, ϵ has a unique factorization through the ideal (See Figure 28,

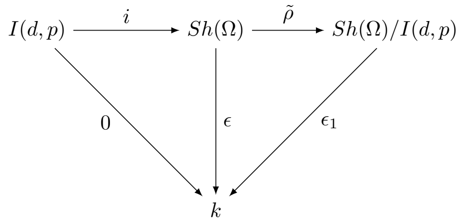


Figure 28: Factorization

where i is the inclusion map)

$$\epsilon = \epsilon_1 \tilde{\rho}, \ \epsilon_1 \in \mathrm{Alg}(\mathrm{Sh}(\Omega)/I, k).$$

Clearly,

$$\ker \delta_1 \cap \ker \epsilon_1 = 0$$

such that

$$(\delta_1, \epsilon_1) = (\delta_1(d,p), \epsilon_1(d,p))$$

is a splitting surjective pointed differentiation and

$$(\tilde{\rho}, \hat{\rho}) \in \mathrm{Diff}(\delta, \epsilon; \delta_1, \epsilon_1) \ .$$

The above combined with the definitions displayed in Figure 27 for $\tilde{\chi}_0, \hat{\chi}_0$ shows that

$$\delta \tilde{\chi}_0 = \hat{\chi}_0 d$$

which in general is not the case:

$$(\tilde{\chi}_0, \hat{\chi}_0) \notin \mathrm{Diff}(\delta, \epsilon; \delta_1, \epsilon_1).$$

This indicates that we will need an extra condition, which is given by the Theorem below.

Theorem 12.3.1

Let (d', p') be a splitting pointed differentiation, and let

$$(\tilde{\theta}, \hat{\theta}) \in \mathrm{Diff}(\delta, \epsilon; \delta_1, \epsilon_1)$$

then

$$(\tilde{\theta}\tilde{\chi}_0, \hat{\theta}\hat{\chi}_0) \in \mathrm{Diff}(d, p; d', p')$$

if and only if

$$I \subset \ker \tilde{\theta}.$$

Identifying $(\tilde{\theta}, \hat{\theta})$ with $(\tilde{\rho}, \hat{\rho})$ in diagram 27, it is clear that

$$I(d, p) \subset \ker \tilde{\rho}$$

such that the conditions of the above Theorem are satisfied. This is demonstrated by the next two Corollaries turning diagram 27 into a consistent mathematical construct for d-paths.

Corollary 12.3.1

If

$$\tilde{\chi} = \tilde{\rho}\tilde{\chi}_0$$

and

$$\hat{\chi} = \hat{\rho}\hat{\chi}_0,$$

then

$$(\tilde{\chi}, \hat{\chi}) \in \mathrm{Diff}(d, p; \delta_1, \epsilon_1).$$

Corollary 12.3.2

If (d', p') is a splitting pointed differentiation and if

$$(\tilde{\theta}, \hat{\theta}) \in \mathrm{Diff}(\delta, \epsilon; d', p')$$

such that

$$(\tilde{\theta}\tilde{\chi}_0, \hat{\theta}\hat{\chi}_0) \in \mathrm{Diff}(d, p; d', p'),$$

then there exists a unique

$$(\tilde{\Theta}, \hat{\Theta}) \in \mathrm{Diff}(\delta_1, \epsilon_1; d', p')$$

with

$$(\tilde{\Theta}\tilde{\chi}, \hat{\Theta}\hat{\chi}) = (\tilde{\theta}\tilde{\chi}_0, \hat{\theta}\hat{\chi}_0).$$

Using the ideal $I(d, p)$ of the shuffle algebra any d-path γ starting at p can be factorized through (see Figure 29)

$$\gamma' \in \mathrm{Alg}(Sh(\Omega)/I, k)$$

as

$$\gamma = \gamma' \tilde{\rho}.$$

With the aid of this factorization we obtain

$$q = \gamma \tilde{\chi}_0 = \gamma' \tilde{\chi} \in \mathrm{Alg}(\mathfrak{U}, k).$$

We call p and q the initial and terminal points of γ with γ being the d-path from p to q. If γ is a d-path from p to q, then

$$\gamma(df) = \gamma(\tilde{\chi}_0 f - pf) = qf - pf$$

that follows from the factorization through the ideal $I(d, p)$.

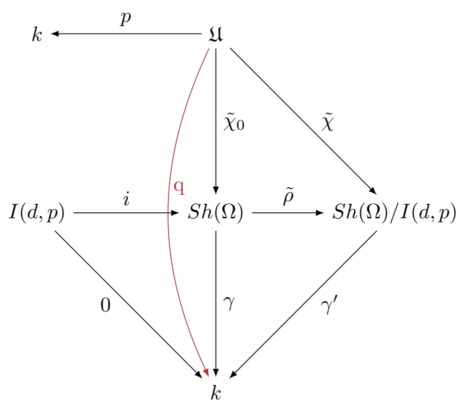


Figure 29: Factorization d-path

Proposition 12.3.4: Unique initial point

If k is an integral domain(8.2.8), then the initial point of a d-path

$$\gamma \neq \epsilon$$

is unique.

proof 12.3.1

Let γ be a d-path from p as well as from p'. Assuming

$$\gamma \neq \epsilon$$

there exist

$$w \in T^1(\Omega), \; v \in T(\Omega)$$

for which

$$\gamma(vw) \neq 0$$

Let now f be an element of $\mathfrak{U}$, we then have

$$\gamma((fw)v) = \gamma(dfwv) + (pf)\gamma(wv) \qquad = \gamma(dfwv) + (p'f)\gamma(wv)$$

From this it follows that

$$pf = p'f$$

such that we indeed have a unique initial point for the d-path γ.

From the exposition above one might be inclined to think that the algebraic structure might depend on the initial point $p \in M$. The following Lemma shows that this it not the case, i.e. the algebraic structure is preserved under translation of the path to another initial point.

Lemma 12.3.2

If γ is a d-path from p to q, then

$$\tilde{L}_\gamma I(d,p) = I(d,q) \tag{80}$$

proof 12.3.2

We know from Proposition 11.3.1 that $(\tilde{L}_\gamma, \hat{L}_\gamma)$ is an equivalence in the category of differentiations $\mathcal{D}$, such that $\tilde{L}_\gamma I(d,p)$ is indeed a δ-closed ideal of $Sh(\Omega)$. Hence

$$\tilde{L}_\gamma(fw - df\ w - (pf)w) = (\gamma \otimes 1)\Delta(fw - df\ w - (pf)w)$$

and we obtain

$$\begin{aligned} \Delta(fw) &= fw \otimes 1 + 1 \otimes fw \\ (\gamma \otimes 1)\Delta(fw) &= \gamma(fw) + fw \\ \Delta(df\ w) &= 1 \otimes df\ w + df\ w \otimes 1 + df \otimes w \quad \text{def } \Delta \\ (\gamma \otimes 1)\Delta(df\ w) &= df\ w + \gamma(df\ w) + \gamma(df)w \\ \Delta(pf\ w) &= pf\ w \otimes 1 + 1 \otimes pf\ w \\ (\gamma \otimes 1)\Delta(p\ fw) &= (pf)\gamma(w) + pf\ w \end{aligned}$$

summing all the above and using that γ is a d-path we get

$$\begin{aligned} &\tilde{L}_\gamma(fw - dfw - (pf)w) \\ &= \gamma(fw) + fw - dfw - \gamma(dfw) - \gamma(df)w - (pf)\gamma(w) - pfw \\ &= fw - \gamma(fw - dfw - pfw) - qfw - dfw + pfw - pfw \\ &= fw - dfw - (qf)w, \end{aligned}$$

where we used

$$\gamma(I) = 0,$$

such that by Proposition 12.3.3

$$I(d,q) \subset \tilde{L}_\gamma I(d,p).$$

Applying the same reasoning to $\tilde{L}_{\gamma^{-1}}$ we obtain

$$I(d,p) \subset \tilde{L}_{\gamma^{-1}} I(d,q)$$

such that

$$\tilde{L}_\gamma I(d,p) \subset I(d,q)\ .$$

This shows that

$$I(d,p) \equiv I(d,q)$$

This Lemma now gives an interpretation of the L-operator we introduced in Definition 11.2.3: it is the operator, associated to a path γ from p to q, that translates the algebra ideal $I(d,p)$ at p to the algebra ideal $I(d,q)$ at q, the endpoint of the d-path γ.

Having defined d-paths as the algebra morphism $\mathrm{Alg}(Sh(\Omega), k)$ that vanish on the ideal $I(d,p)$, we can wonder about the composition and inverses of these d-paths. The composition of paths can be interpreted as a multiplication in the space of d-paths where this multiplication is given by Definition 11.2.6 for

$$\gamma_1, \gamma_2 \in \mathrm{Alg}(Sh(\Omega), k).$$

As this multiplication turned the algebra homomorphisms into a group the same will also be true for d-loops, for d-paths this will turn out to be a semi-group[1].

Theorem 12.3.2

If γ_1 is a d-path from p to q and if γ_2 is a d-path from q to q', then $\gamma_1\gamma_2$ is a d-path from p to q', and γ_1^{-1} is a d-path from q to p.

proof 12.3.3

Since

$$\gamma_1\gamma_2 I(d,p) = \gamma_2 \tilde{L}_{\gamma_1} I(d,p) = \gamma_2 I(d,q) = 0 \tag{81}$$

$\gamma_1\gamma_2$ is a d-path from p. For

$$f \in \mathfrak{U}$$

$$(\gamma_1\gamma_2)(df) = \gamma_1(df) + \gamma_2(df) = q'f - pf, \tag{82}$$

meaning that q' is the end point of $\gamma_1\gamma_2$. Next to this we also have

$$\begin{aligned}\gamma_1^{-1} I(d,q) = \gamma_1^{-1}\tilde{L}_{\gamma_1} I(d,p) = \\ (\gamma_1^{-1}\gamma_1) I(d,p) = (\epsilon) I(d,p) = 0\end{aligned} \tag{83}$$

and

$$\gamma_1^{-1}(df) = -\gamma_1(df) = pf - qf \tag{84}$$

Hence γ_1^{-1} is a d-path from q to p.

12.4 CONNECTEDNESS

In the previous Sections we have formally introduced Chen's generalization of the intuitive idea of paths in a given space. Naturally, similarly to the case

1 This is due to the fact that the endpoint and initial points of two d-paths need not to coincide.

of paths in a topological space, we can wonder when a space is connected with respect to these d-paths. Spaces that turn out to be connected with respect to these d-paths are referred to as d-connected as compared to the path connected ones.

Definition 12.4.1 (d-Connectedness)**.**
If

$$\forall p, q \in \mathrm{Alg}(\mathfrak{U}, k)$$

there always exists a d-path from p to q, then $\mathfrak{U}$ is called d-connected.

In topology continuous maps transform path connected spaces to path connected spaces, can we make a similar statement for d-connected? The answer is provided by the following Proposition

Proposition 12.4.1: Maps Between d-Connected Spaces

Let

$$(\tilde{\phi}, \hat{\phi}) \in \mathrm{Diff}(d, d') \ .$$

If $\mathfrak{U}'$ is d'-connected and if $\tilde{\phi}$ induces a surjective map from $\mathrm{Alg}(\mathfrak{U}', k)$ onto $\mathrm{Alg}(\mathfrak{U}, k)$ then $\mathfrak{U}$ is d-connected.

d-paths are defined by means of a differentiation (d, p) that then returns the ideal $I(d, p)$ on which the d-path vanishes. Not surprisingly we have that if two points are points in a d-connected space, i.e. they can be connected by a d-path, their differentiations are in the same equivalence class.

Proposition 12.4.2: Equivalence of Differentiations

If $\mathfrak{U}$ is d-connected, then

$$\forall p, q \in \mathrm{Alg}(\mathfrak{U}, k)$$

the differentiations $\delta_1(d, p)$ and $\delta_1(d, q)$ are equivalent.

Having defined d-connectedness, we can introduce d-discrete points.

Definition 12.4.2 (d-Discrete Point)**.**
Any

$$p \in \mathrm{Alg}(\mathfrak{U}, k)$$

is said to be a d-discrete point if there exists no d-path

$$\gamma \neq \epsilon \text{ from } p.$$

From the definition of d-discrete points we can derive when

$$w \in \Omega$$

is trivial.

Definition 12.4.3 (P-trivial)**.**
Let

$$P \subset \mathrm{Alg}(\mathfrak{U}, k)$$

contain at least one d-nondiscrete point. We say that

$$w \in \Omega$$

is P-trivial if, for any d-path γ from any $p \in P$ and for any

$$u_1, u_2 \in T(\Omega)\ ,\ f \in \mathfrak{U}\ ,\ \gamma(u_1(fw)u_2) = 0.$$

We call

$$f \in \mathfrak{U}$$

P-trivial if df is P-trivial and if fw is P-trivial for any

$$w \in \Omega.$$

Notice that not every point of Ω is P-trivial, otherwise P would only consist of d-discrete points. Moreover, we see that

$$\mathbb{1} \in \mathfrak{U}$$

is not P-trivial.

Let $\mathfrak{U}_P$ be the quotient k-algebra of $\mathfrak{U}$ over the ideal of the P-trivial elements of $\mathfrak{U}$ and let Ω_P be the quotient of the $\mathfrak{U}$-module of Ω over the

$\mathfrak{U}$-submodule of the P-trivial elements of Ω. Obviously Ω_P, is also an $\mathfrak{U}_P$-module. The differentiation d maps the ideal of the P-trivial elements of $\mathfrak{U}$ into the submodule of the P-trivial elements of Ω and therefore induces the following differentiation:

Definition 12.4.4 (P-trivial differentiation)**.**
$d_P : \ \mathfrak{U}_P \to \Omega_P$

Let

$$\tilde{\pi}_P \in \mathrm{Alg}(\mathfrak{U}, \mathfrak{U}_P)$$

and let

$$\hat{\pi}_P \in \mathrm{Hom}(\Omega, \Omega_P)$$

be canonical projection homomorphisms. Then

$$\pi_P = (\tilde{\pi}_P, \ \hat{\pi}_P) \in \mathrm{Diff}(d, d_P).$$

Observe that $\mathfrak{U}_P, \Omega_P$ and π_P depend only on the d-nondiscrete points of P. The projection $\tilde{\pi}_P$ induces an injective map

$$\mathrm{Alg}(\mathfrak{U}_P, k) \to \mathrm{Alg}(\mathfrak{U}, k).$$

Proposition 12.4.3

Let k be an integral domain (See Definition 8.2.8). The totality of the d-non-discrete points of P is contained in the image of the **injective** map

$$\mathrm{Alg}(\mathfrak{U}_P, K) \to \mathrm{Alg}(\mathfrak{U}, k).$$

Being only interested in non-trivial elements, we introduce reduced spaces that only contain non-trivial elements.

Definition 12.4.5 (d-reduced)**.**
$\mathfrak{U}$ is called d-reduced if $\mathrm{Alg}(\mathfrak{U}, k)$ contains at least one d-non-discrete point and if the only $\mathrm{Alg}(\mathfrak{U}, k)$-trivial element of $\mathfrak{U}$ is zero.

Theorem 12.4.1

Let

$$\mathrm{Alg}(\mathfrak{U}, k)$$

be the disjoint union of P and Q satisfying the following conditions:

(i) Each of P and Q contains at least one d-non-discrete point.

(ii) There exists no d-path from any $p \in P$ to some $q \in Q$.

Let

$$\tilde{\theta} \in \mathrm{Alg}(\mathfrak{U}, \mathfrak{U}_P \oplus \mathfrak{U}_Q)$$

be given by

$$\tilde{\theta} f = \tilde{\pi}_P f + \tilde{\pi}_Q f, \tag{85}$$

If $\mathfrak{U}$ is d-reduced, then $\tilde{\theta}$ is injective, i.e. $\mathrm{Alg}(\mathfrak{U}, k)$ is a sub-direct sum of $\mathfrak{U}_P$ and $\mathfrak{U}_Q$.

The above concepts assures the mathematical consistency of the concept of path reduction that will be introduced in the next chapter.

12.5 d-loops

Before discussing the Chen iterated integrals in more detail we make a few comments on d-loops. Generalized loops or d-loops can be naturally defined as d-paths where the initial and end points coincide, but in this case the ideal on which they vanish needs to be extended with

$$\{d\mathfrak{U}\}.$$

This is obvious when we interpret Chen integrals as d-paths since they return zero over this set, such that the set

$$\{d\mathfrak{U}\}$$

needs to be added to the algebra ideal $I(d, p)$[2].

2 See also Remark 12.3.1

Definition 12.5.1 (d-loop).
A d-path from p to p will be called a d-loop from p. Then

$$\{d\mathfrak{U}\}$$

stands for the ideal of $Sh(\Omega)$ generated by

$$d\mathfrak{U} \subset T^1(\Omega).$$

Then

$$\gamma \in \mathrm{Alg}(Sh(\Omega), k)$$

is a d-loop from p if and only if γ annuls the ideal

$$I(d,p) + \{d\mathfrak{U}\}\ ,\ \text{of}\ Sh(\Omega).$$

To reduce notations we introduce

Definition 12.5.2 ($\mathrm{Shc}(d,p)$).
We define

$$\mathrm{Shc}(d,p)$$

to be the quotient k-algebra

$$Sh(\Omega)/(I + \{d\mathfrak{U}\}),$$

where

$$I = I(d,p).$$

The shuffle algebra $\mathrm{Shc}(d,p)$ is unitary and commutative with respect to the shuffle product. After introduction of the multiplication from Definition 11.2.6 and taking Theorem 12.3.2 into account we have that d-loops also form a group. In the Section 11.2 we have shown that $Sh(\Omega)$ is a Hopf-algebra, which induces a Hopf algebra structure on Shc:

Theorem 12.5.1

$\mathrm{Shc}(d,p)$ is a Hopf k-algebra with a co-multiplication Δ_c, a co-unit ϵ_c and antipode J_c , respectively, induced by Δ, ϵ and J.

When considering loops in topology, this is usually associated to the fundamental group, which is independent of the base point of the loops. In the case of d-loops we have similar properties, namely that the Hopf-algebra structure and the group structure of Shc is independent of the base point of the loops. Therefore we have the Proposition

Proposition 12.5.1

If γ is a d-path from p to q, then the Hopf k-algebras $\mathrm{Shc}(d, p)$ and $\mathrm{Shc}(d, q)$ are isomorphic.

From this proposition we naturally have

Corollary 12.5.1

If γ is a d-path from p to q, then the group of d-loops from p is isomorphic with the group of d-loops, from q.

12.6 SUMMARY

In the previous Sections we have introduced Chen's d-paths and d-loops as algebra morphisms. We discussed some of their properties, emphasizing on the importance of the ideals of algebra morphisms, relating ideals and algebra morphism kernels. The shuffle algebra ideal was constructed from Chen's generalization of line integrals. In the next Chapter we shall discuss the relation between d-paths and Chen integrals in more detail.

13

CHEN ITERATED INTEGRALS

13.1 INTRODUCTION

In this chapter we will revisit Chen's iterated integrals, following the papers by Chen [107, 108, 110, 111] and Tavares [16]. We will discuss in more detail how they can be interpreted as d-paths and d-loops.

13.2 d-LOOPS AND CHEN ITERATED INTEGRALS

In this Section we present some of the properties of Chen iterated integrals that will be used for introducing the group of generalized loops.

Before we do this, we start with explaining in more detail the relationship between Chen's integrals and the previously defined d-loops. Remark 12.3.1 demonstrated that the integral algebra $\mathcal{A}_p$ is isomorphic to the algebra $\mathrm{Shc}(d,p)$. A d-loop γ is then an algebra morphism $\mathrm{Alg}(Sh(\Omega),k)$ that vanishes on the ideal

$$I(d,p)+\{d\mathfrak{U}\}.$$

As we have discussed in the previous section, this ideal is also an ideal of the algebra of Chen iterated loop integrals $\mathcal{A}_p$, by definition. The fact that this is an isomorphism of algebras allows for the identification of a d-loop with an element of $\mathcal{A}_p^*$, the dual space of $\mathcal{A}_p$ formed by the real (complex,$GL(n,\mathbb{C})$) valued linear functionals on $\mathcal{A}_p$, such that we have

$$\mathrm{Alg}(\mathrm{Shc}(d,p),k)\ni\gamma\to\oint_\gamma\in\mathcal{A}_p^*. \tag{86}$$

The interesting part about this identification will emerge when we discuss the relation between Chen's integrals and the solution of the parallel transport equation in gauge theory (see Section 16). Keeping a principal fibre bundle setting in mind, we shall assume the one-forms, used in the functionals

$$\omega_1\cdots\omega_r\to X^{\omega_1\cdots\omega_r}=\int\omega_1\cdots\omega_r$$

to be Lie algebra-valued, where the integral is a Chen integral. Put differently, we shall assume

$$\omega_i \in \bigwedge^1 M \otimes \mathfrak{gl}(\mathfrak{g}),$$

where $\mathfrak{gl}$ is a matrix representation (i.e., an element of $GL(n, \mathbb{C})$) of the Lie algebra $\mathfrak{g}$, which explains the presence of

$$\omega_i \in \bigwedge^1 M \otimes GL(n, \mathbb{C})$$

in many of the previous and following definitions and properties of Chen iterated integrals.

13.3 CHEN ITERATED INTEGRALS: PROPERTIES

In Chapter 12 we introduced Chen iterated integrals (Definitions 12.3.1 and 12.3.2) discussing some of their properties. The main goal of this section is to extend them with properties that are relevant with respect to GLS.

We start by considering several elementary properties concerning the behavior of the Chen integrals, similar to the properties of regular line integrals. Firstly we investigate the behavior of these integrals with respect to intermediate points along the path of integration. The result is the following Lemma

Lemma 13.3.1: Intermediate Points

For

$$a \leq c \leq b \ , \ \gamma^c \sim \int_a^c \ , \text{ and } \gamma_c \sim \int_c^b$$

we have

$$\begin{aligned} I_{i_1 \cdots i_p}(\gamma) &= I_{i_1 \cdots i_p}(\gamma^c) + I_{i_1 \cdots i_{p-1}}(\gamma^c) I_{i_p}(\gamma_c) \\ &\quad + I_{i_1 \cdots i_r}(\gamma^c) I_{i_r \cdots i_p}(\gamma_c) + \cdots + I_{i_1 \cdots i_p}(\gamma_c) \end{aligned} \tag{87}$$

proof 13.3.1

The Lemma can be proved by induction. It is clear that

$$I_i(\gamma) = I_i(\gamma^c) + I_i(\gamma_c)$$

Using the induction hypotheses that Equation (87) is valid for $p-1$ we can derive

$$\begin{aligned} I_{i_1\cdots i_p}(\gamma) &= \int_a^c I_{i_1\cdots i_{p-1}}(\gamma^t)dx_{i_p}(t) + \int_c^b I_{i_1\cdots i_{p-1}}(\gamma^t)dx_{i_p}(t) \\ &= I_{i_1\cdots i_p}(\gamma^c) + \int_c^b \Big(I_{i_1\cdots i_{p-1}}(\gamma^c) + I_{i_1\cdots i_{p-2}}(\gamma^c) I_{i_{p-1}}(\gamma_c^t) \\ &\qquad + \cdots + I_{i_1\cdots i_{p-1}}(\gamma_c^t)\Big)\, dx_{i_p}(t) \\ &= I_{i_1\cdots i_p}(\gamma^c) + I_{i_1\cdots i_{p-1}}(\gamma^c) I_{i_p}(\gamma_c) \\ &\qquad + I_{i_1\cdots i_r}(\gamma^c) I_{i_r\cdots i_p}(\gamma_c) + \cdots I_{i_1\cdots i_p}(\gamma_c) \end{aligned}$$

Secondly we can ask how these integrals transform under a re-parametrization, which is answered by the Proposition

Proposition 13.3.1: Re-parametrization

$$\int_\gamma \omega_1 \cdots \omega_n$$

is invariant under orientation preserving re-parametrizations.

Remark 13.3.1.
I would like to point out at this point that there are approaches where one does not use a re-parametrization equivalence relation [112], meaning that in this case parametrizing a path with a parameter $t \in [0,1]$ is not equivalent to the same path parametrized by a parameter $q \in [0,2]$ for instance.

Having the above properties we consider Chen integrals over a composition of paths. Let therefore α and β be two paths with the endpoint of α equal to the starting point of β (let us call this point c). Composing these paths we form the path $\alpha\beta$ (that means β after α). Applying Lemma 13.3.1 to this composite path $\alpha\beta$ with the point c taken as an intermediate point we have the Lemma

Lemma 13.3.2: Combining paths

$$I_{i_1\cdots i_p}(\alpha\beta) = I_{i_1\cdots i_p}(\alpha) + I_{i_1\cdots i_{p-1}}(\alpha)I_{i_p}(\beta) + I_{i_1\cdots i_r}(\alpha)I_{i_r\cdots i_p}(\beta) + \cdots + I_{i_1\cdots i_p}(\beta) \quad (88)$$

Using the notations of Definition 12.3.2, this can be rewritten as

Proposition 13.3.2: Composition of Paths

If

$$\alpha, \beta \in \mathcal{PM},$$

the space of paths in the real smooth manifold M, i.e.

$$\alpha, \beta : [0,1] \to M$$

with

$$\alpha(1) = \beta(0),$$

then we can compose the paths using Equation (87) . When composing the paths, the Chen integrals change in the following way (where we introduced the notion of an inverse path):

$$\int_{\alpha\cdot\beta} \omega_1 \cdots \omega_r = \sum_{i=0}^{r} \int_\alpha \omega_1 \cdots \omega_i \cdot \int_\beta \omega_{i+1} \cdots \omega_r \quad (89)$$

$$\int_{\alpha^{-1}} \omega_1 \cdots \omega_r = (-1)^r \int_\alpha \omega_r \cdots \omega_1 \quad (90)$$

for

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M$$

and with the convention that

$$\int_\gamma \omega_1 \cdots \omega_r = 1, \text{ if } r = 0.$$

When applied to

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M \otimes GL(n, \mathbb{C})$$

(i.e. general linear group complex matrix-valued one-forms), Eq. (90) is replaced by

$$\int_{\alpha^{-1}} \omega_1 \cdots \omega_r = (-1)^r \int_\alpha [\omega_r^T \cdots \omega_1^T]^T \quad (91)$$

with ω^T the transpose of the matrix ω (associated with the matrix function).

The matrix valued one-forms in the Proposition are matrix functions, which we will define in Section 15.2. Again considering a principal fibre bundle approach to the formulation of gauge theories (Section 14) we shall identify the gauge potentials A_μ with such one-forms, where the matrices stem from a $GL(n, \mathbb{C})$ representation of the Lie algebra.

We introduce now the following notation to avoid overloading of equations in what follows

$$X^{\omega_1 \cdots \omega_r} : \mathcal{PM} \to \mathbb{C},\ X^{\omega_1 \cdots \omega_r}(\gamma) = \int_\gamma \omega_1 \cdots \omega_r = \gamma(\omega_1 \cdots \omega_r), \tag{92}$$

where $\int_\gamma$ is interpreted as a d-path and we considered the one-forms ω_i to be complex-valued,

$$\omega_i \in \bigwedge^1 M \otimes \mathbb{C}.$$

$\mathcal{PM}$ represents here the space of d-paths. An extension to complex matrix valued one-forms, and thus to Lie algebra valued one-forms as well is straightforward. The example below might be instructive to understand the notation 92.

Example 13.3.1.

$$X^{\omega_1 \omega_2}(\alpha) = \int_\alpha \omega_1 \omega_2,$$

where

$$\omega_1, \omega_2 \in \bigwedge^1 M \otimes GL(n, \mathbb{C})$$

are matrices of one-forms on M, is a matrix in $GL(n, \mathbb{C})$ with the elements given by:

$$\left(\int_\alpha \omega_1 \omega_2 \right)^i_j = \int_\alpha (\omega_1)^i_k \otimes (\omega_2)^k_j \tag{93}$$

Considering Chen integrals as d-paths and d-loops, we can derive some extra notions related to d-paths from the above properties.

Definition 13.3.1 (Elementary Equivalent Paths)**.**
Two paths are called elementary equivalent if

$$\alpha\beta\beta^{-1}\gamma = \alpha\gamma.$$

This equivalence induces an equivalence relation on the d-paths and thus also induces equivalence classes of paths $[\alpha\gamma]$.

Definition 13.3.2 (Piecewise Regular Paths)**.**
A piecewise regular path is a path in $\mathcal{PM}$ with non-vanishing tangent vectors.

Definition 13.3.3 (Reduced Paths)**.**
A path is called a reduced path if it is a piecewise regular path, which does not belong to the type

$$\forall\beta,\ \alpha\beta\beta^{-1}\gamma.$$

Combining these Definitions with Equations (89) and (90) allows to demonstrate that the functionals X defined in (92) depend only on the equivalence class and not on the specific path representing the class. As an example it is clear that γ and $\gamma\beta\beta^{-1}$ are representatives of the same class. From the composition of paths property and inverses of the Chen integrals we derive

$$X^{\omega_1\cdots\omega_n}(\gamma) = X^{\omega_1\cdots\omega_n}(\gamma\beta\beta^{-1}).$$

This is often graphically represented in the literature on loop spaces as in Figure 30, with functions X having the property

$$X^{\omega_1\cdots\omega_n}(\gamma) = X^{\omega_1\cdots\omega_n}(\gamma\beta\beta^{-1})$$

sometimes referred to as **Stokes functionals**. The above properties allow us to demonstrate that if we have a reduced piecewise regular path γ there always exist one-forms such that the functionals $X^{\omega_1\cdots\omega_r}(\gamma)$ are not zero.

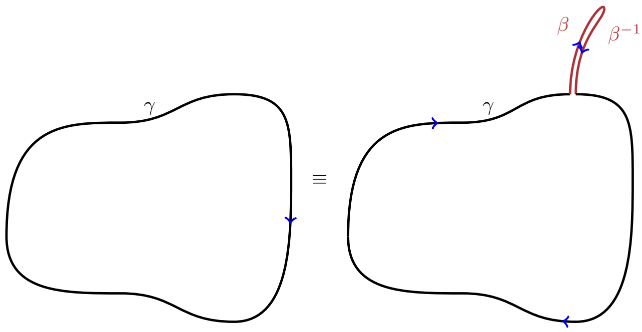


Figure 30: The property of path reduction.

Lemma 13.3.3: Non-Vanishing Chen Integrals

If

$$\gamma \neq \epsilon$$

(identity, co-unit), a reduced piecewise regular path in $\mathcal{PM}$, then one can find one-forms

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M,$$

$$\bigwedge^1 M \otimes GL(n, \mathbb{C}),$$

for $r \geq 1$, such that

$$X^{\omega_1 \cdots \omega_r}(\gamma) \neq 0$$

More importantly, this Lemma is used to prove that the functional $X^{\omega_1 \cdots \omega_r}$ can be used to separate points in the space of d-paths and d-loops.

Theorem 13.3.1: Separation property theorem

Two piecewise regular paths α, β are equivalent if and only if

$$X^{\omega_1 \cdots \omega_r}(\alpha) = X^{\omega_1 \cdots \omega_r}(\beta) \tag{94}$$

for any one-forms

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M \ , \ r \geq 1$$

In other words d-paths, defined by Chen integrals, are equivalent to exactly one reduced path and these paths can be distinguished from each other by the functionals $X^{\omega_1 \cdots \omega_r}$. Chen demonstrated in [108] the following Theorem, that states that if the d-paths of α and β respectively return the same value for the exponential homomorphism Θ, then α and β only differ by parametrization and left translation (Lie) provided they are reduced (see Definition 13.3.3):

Theorem 13.3.2

Let Θ be the (formal) exponential homomorphism:

$$\Theta(\alpha) = \mathbb{1} + \sum_{p=1}^{\infty} \sum \int_{\alpha} \omega_{i_1} \cdots \omega_{i_p} X_{i_1} \cdots X_{i_p}, \tag{95}$$

where the X_j are non-commutative indeterminates with respect to a base

$$\omega_1, \cdots, \omega_m$$

of the Maurer-Cartan forms $(g^{-1}dg)$ of a real Lie group G, and $\Theta(\alpha)$ is in an element of the G. Then one of two irreducible piecewise regular continuous paths α and β can be obtained from the other by left translation and change of parameter if and only if

$$\Theta(\alpha) = \Theta(\beta).$$

Identifying the exponential homomorphism with the parallel transporter or Wilson line, the above theorem strengthens the equivalence relation on d-paths and d-loops induced by path reduction from Definition 13.3.3. In other words the parallel transporter can be used to distinguish or separate d-paths and d-loops, a fact that will be quite helpful when topologizing the algebra $\mathcal{A}_p$.

13.4 SUMMARY

We have successfully made the connection between the shuffle product and the Chen Iterated Integrals, showing that there exists an algebra morphism between them that conserves the algebraic structures of the shuffle algebra. Applying the last theorem to loops shows the first and second step in the construction of the equivalence class that will deal with the over-completeness of loop space, while preserving the desired algebraic and differential structures. The first step was done by introducing re-parametrization equivalence and the second by identifying loops that have an equal exponential homomorphism.

14

GAUGE FIELDS AS CONNECTIONS ON A PRINCIPAL BUNDLE

A mathematical point of view on Quantum Field Theory suggests that the fundamental interactions between matter fields can be conveniently expressed in a geometrical setting by using principal fibre bundles. This Chapter introduces some of the basic concepts of **fibre bundle theory**, which will be used to derive the parallel transport equation. At a later stage in this text we will link the solution of this equation to the concept of Wilson lines. We comment that principal fibre bundles are not the only method to provide a geometrical description of physical interactions. Depending on the point of view of the user, other approaches, which in some sense are closer to the ideas of Quantum Mechanics, are provided by (Lie) algebroids (see for instance [113]) and non-commutative geometry (see for instance [114]).

Note that the exposition below is far from complete and much more detail can be found in the standard references [115–118] which we used as a basis for this section.

14.1 PRINCIPAL FIBRE BUNDLE, SECTIONS AND ASSOCIATED VECTOR BUNDLE.

We start with the definition of a principal (Yang-Mills) fibre bundle.

Definition 14.1.1 (Yang-Mills principal fibre bundel).
A Yang-Mills principal fibre bundle

$$P(M^4, G, \pi)$$

consists of the following elements:

1. *A differentiable manifold P, which is called* ***the total space***
2. *A Minkowskian manifold M^4, referred to as the* ***base space***
3. *A gauge (Lie) group G , which is called* ***the fibre***

4. *A surjective map*

$$\pi : P \to M^4,$$

referred to as ***the projection****, and its inverse image*

$$\pi^{-1}(x) \equiv G_x \cong G,$$

the fibre at x.

This geometrical structure has the following properties

(i) A Lie group G (in Yang-Mills theories it is $SU(N)$), called the structure group, acts on the fibres from the left

(ii) An open cover $\{U_i\}$ of M^4 together with the diffeomorphisms

$$\phi_i : U_i \times G \to \pi^{-1}(U_i)$$

such that

$$(\pi \circ \phi_i)(x, g) = x.$$

The ϕ_i are referred to as the local gauge or local trivialization since ϕ_i^{-1} maps $\pi^{-1}(U_i)$ onto the direct product $U_i \times G$

(iii) Using the notation

$$\phi_i(x, g) = \phi_{i,x}(g)$$

the map

$$\phi_{i,x} : G \to G_x$$

is a diffeomorphism. On

$$U_i \cap U_j \neq \emptyset,$$

it is required that

$$S_{ij}(x) \equiv \phi_{i,x}^{-1} \phi_{j,x} : G \to G$$

is an element of the structure group G. Both maps ϕ_i and ϕ_j are related by a smooth map

$$S_{ij} : U_i \cap U_j \to G$$

such that

$$\phi_j(x, g) = \phi_i(x, S_{ij}(x) g).$$

We refer to the S_{ij} as the transition functions or **passive** gauge transformations.

In such a principle fibre bundle the structure group is isomorphic to the fibre, inducing a right action of G on the fibre that **does not depend on the local gauges**[1]. This right action of G on $\pi^{-1}(U_i)$ is defined by

$$\phi_i^{-1}(pg) = (x, g_i g)$$

or

$$pg = \phi_i(x, g_i g) \ , \ \forall g \in G$$

and

$$p = \pi^{-1}(x).$$

To see that this is independent of local gauges, let us consider an

$$x \in U_i \cap U_j,$$

for which:

$$pg = \phi_j(x, g_j g) = \phi_j(x, S_{ji} g_i g) = \phi_i(x, g_i g), \tag{96}$$

demonstrating that it indeed does not depend on the choice of local gauge. In

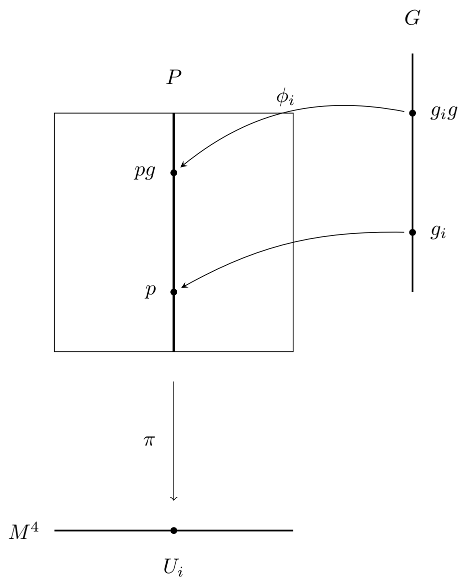


Figure 31: Right action of G on a fibre.

these principal fibre bundles we will define a kind of inverse of the projection maps called **sections**

1 Actually a fibre bundle is an equivalence class on the fibers.

Definition 14.1.2 (Section).
A section is a smooth map

$$s : M^4 \to P$$

such that

$$\pi \circ s = \mathbb{1}_{M^4}.$$

Clearly sections and local gauges are intimately related, as such given a section $s_i(x)$ over U_i, one can reconstruct the corresponding local gauge ϕ_i. To this end, let us consider

$$p \in \pi^{-1}(x),\ x \in U,$$

for which there exists a unique element

$$g_p \in G$$

such that

$$p = s_i(x) g_p.$$

Now we define ϕ_i through its inverse

$$\phi_i^{-1}(p) = (x, g_p).$$

Notice that in this specific gauge we get

$$s_i(x) = \phi(x, e),$$

where e is the identity element in the structure group G, often referred to as the **the canonical local trivialization** or **local gauge**.

The **gauge potentials**, used in gauge theory exist naturally in these principal fibre bundles that form a well-defined geometrical space. In this way we have a geometrical representation of the gauge potentials, of course we will also need a geometrical representation of the matter fields $\psi(x)$. This geometrization of matter fields however, requires an additional mathematical structure called **the associated vector bundle**

$$E(M^4, G, V, P, \pi_E).$$

This vector bundle E is constructed from a k-dimensional vector space V on which the gauge group G (from the principal fibre bundle) acts from the left. This left action is defined in the following way

$$g \in G, (p, v) \in P \times V : (p, v) \mapsto (pg, \rho^{-1}(g)v) \tag{97}$$

with ρ is the k-dimensional **unitary representation of** G.

The vector bundle

$$E(M^4, G, V, P, \pi_E)$$

is then an equivalence class

$$P \times V/G,$$

such that

$$(p, v) \equiv (pg, \rho^{-1}(g)v).$$

Notice that the action of $\rho^{-1}(g)$ is from the left on elements of the vector space. The bundle E now also has a fibre bundle structure

$$E = P \times_\rho V$$

where

$$\pi_E : E \to M^4, \pi_E(p, v) = \pi(p)$$

with local trivialization

$$\Phi_i : U_i \times V \to \pi_E^{-1}(U_i).$$

Similar to the principal fibre we have again transition functions, which are now $\rho(S_{ij}(x))$ and where the S_{ij} are transition functions on P.

A local section s_i on P then do not only determine a local gauge on P, but also on E:

$$\phi_{i,x}^{-1} \circ s_i(x) = \mathbb{1}_G \tag{98}$$

$$\Phi_{i,x}^{-1} \circ s_i(x) = \mathbb{1}_V, \tag{99}$$

with

$$\Phi_{i,x}^{-1} : \pi_E^{-1}(x) \to V, \ x \in U_i.$$

This associated vector bundle

$$E(M^4, G, V, P, \pi_E)$$

can be used to geometrize the matter fields.

Definition 14.1.3 (Matter field)**.**
A matter field of type (ρ, V) is defined as a section

$$\psi : M^4 \to E.$$

Being expressed in a gauge-independent way, it yields:

Definition 14.1.4 (Gauge Independent Definition Matter Field)**.**
A matter field of type (ρ, V) is defined as a map

$$\tilde{\psi} : P \to V$$

that is equivariant under the structure group

$$G : \tilde{\psi}(pg) = \rho(g^{-1})\tilde{\psi}(p), \ \forall p \in P, \ \forall g \in G.$$

where any explicit reference to points in M^4 have disappeared.

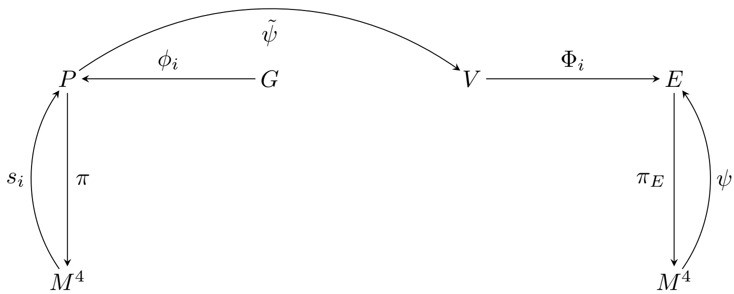


Figure 32: Definition matter field

14.2 GAUGE POTENTIAL AS A CONNECTION

In general it is agreed that in a gauge theory the gauge fields (or potentials) can be introduced as Lie algebra-valued one-forms on the principal fibre bundle associated to the gauge theory[2].

2 Let us emphaze that the identification of fields, as defined in quantum field theory in physics, with sections of the principal (gauge) fiber bundle is only valid in the perturbative sector. In the non-perturbative regime the situation becomes much more involved, see [119]

In this Section we will discuss this identification. A gauge field A_μ (also referred to as gauge potential) transforms under a gauge transformation $U(x)$ as

$$A_\mu \to U(x)A_\mu(x)U^\dagger(x) \mp \frac{i}{e_0}\partial_\mu U(x)U^\dagger(x),$$

with e_0 the coupling constant. This differs from the transformation law for vectors and looks more like the transformation of a connection

$$\omega \to g^{-1}\omega g + g^{-1}\partial_\mu g.$$

This is the initial motivation to the search for a connection that might be identified somehow with the gauge fields. Before we can look for such an identification we of course first need to know what a connection is. Therefore we will first define connections by giving two equivalent definitions of connection, the first of which is more used by mathematicians, while the second one is more favored by physicists.

Definition 14.2.1 (Connection (math)).
Let $P(M^4, G, \pi)$ be a principal fibre bundle, then a connection on P is a unique separation of the tangent space T_pP into the vertical space V_pP and the horizontal subspace H_pP, such that

(i)

$$T_pP = H_pP \oplus V_pP$$

(ii) A smooth vector field X on P can be split in

$$X^H \in H_pP$$

and

$$X^V \in V_pP,$$

such that

$$X = X^H + X^V$$

(iii) $H_{pg}P = R_{g*}H_pP$ for $p \in P,\ g \in G$

for a discussion on this subject. Sometimes one runs into problems of uniqueness, even in the perturbative sector. An example of this is, for instance, the $U(1)$-bundle of the sphere S^2 [115].

In this Definition the vertical space is considered to be tangent to the fibre G_x at p, which we shall discuss in more detail below. The last statement in the Definition says that the horizontal spaces $H_{pg}P$ and H_pP on the same fibre are related by a linear transformation induced by the right action of the gauge group G. To many people it is not immediately clear what these horizontal and vertical space are, therefore many physicist prefer the definition of a connection one-form due to Ehresmann. This definition requires some facts about Lie groups and Lie Algebras, so before giving the second definition of a connection we briefly revise some of these facts.

Consider a Lie group G to which we can associate a left (L_g) and a right action (R_g) defined respectively as

$$L_g h = gh$$

and

$$R_g h = hg$$

for $g, h \in G$. The left action L_g generates the map (push-forward, see Definition 7.2.12)[3]

$$L_{g*} : T_h(G) \to T_{gh}(G)$$

between tangent spaces at different points in the Lie group G. By requiring that

$$L_{g*} X|_h = X|_{gh}.$$

we can define a left-invariant vector field X . These left-invariant vector fields generate a Lie algebra of G, which we write as $\mathfrak{g}$. Now

$$X \in \mathfrak{g}$$

is specified by its value at the Lie groups unit element e, and vice versa. This means there exists a vector space isomorphism between the Lie algebra $\mathfrak{g}$ and the tangent space of G at the unit element, i.e.,

$$\mathfrak{g} \cong T_e G.$$

From Lie theory we learn that the Lie algebra $\mathfrak{g}$ has a set of generators $\{T_\alpha\}$ that also define the structure constants

$$f^\gamma_{\alpha\beta} : [T_\alpha, T_\beta] = f^\gamma_{\alpha\beta} T_\gamma.$$

3 Notice that this map is well-defined due to the fact that this action is an automorphism of G.

Next to the left and right action we can also define an **adjoint action** on Lie groups

$$\mathrm{ad}: \; G \to G, \; h \mapsto \mathrm{ad}_g h \equiv g h g^{-1},$$

which in its turn generates the adjoint map

$$\mathrm{Ad}_g : T_h(G) \to T_{ghg^{-1}}(G)$$

between tangent spaces. By choosing $h \in G$ in the adjoint map to be the unit element e, we immediately see that Ad_g maps

$$T_e(G) \cong \mathfrak{g}$$

onto itself.

The above facts about Lie groups and algebras, will now allow us to understand how to construct the vertical subspace V_pP, defined in Definition 14.2.1, of the tangent space of the principal fibre bundle T_pP. Let

$$A \in \mathfrak{g}$$

and $p \in P$, then the right action:

$$R_{\exp[tA]} p = p\mathrm{e}^{[tA]}, \tag{100}$$

defines a curve through p parametrized by t. Observe that

$$\pi(p) = \pi\left(p\mathrm{e}^{[tA]}\right) = x$$

implies that the curve lies in G_x, the fibre above

$$x \in M^4.$$

Using an arbitrary smooth function

$$f : P \to \mathbb{R}$$

we define the vector

$$A^\sharp \in T_pP$$

as

$$A^\sharp f(p) = \frac{d}{dt} f\left(p\mathrm{e}^{[tA]}\right)|_{t=0}. \tag{101}$$

This vector is tangent to P at p and by definition thus also tangent to G such that we have

$$A^\sharp \in V_pP.$$

Constructing such an $A^\sharp$ at each point of P, builds a vector field also noted as $A^\sharp$ referred to as the **fundamental vector field** generated by A. We obtain, therefore, the isomorphism

$$\sharp : \mathfrak{g} \to V_pP : A \mapsto A^\sharp,$$

and identify the complement of V_pP with H_pP from Definition 14.2.1.

We are now in a ready to define the **Ehresmann connection one-form**:

Definition 14.2.2 (Ehresmann Connection One-Form).
A connection one-form

$$\omega \in T^*P \otimes \mathfrak{g}$$

is a projection of T_pP onto the vertical component

$$V_pP \cong \mathfrak{g},$$

the Lie algebra of G. This one-form possesses the following properties:

(i)

$$\omega(A^\sharp) = A$$

with

$$A \in \mathfrak{g}$$

(ii)

$$R_g^*\omega = \mathrm{Ad}_{g^{-1}}\omega$$

or

$$X \in T_pP : R_g^*\omega_{pg}(X) = \omega_{pg(R_{g*}X)} = g^{-1}\omega_p(X)g$$

From this Definition, the horizontal subspace H_pP can be identified with the kernel of ω. Recall that the goal of this chapter is to relate A_μ to a Lie algebra-valued one-form. Now we have constructed a one-form, so the next question is how to relate the gauge fields A_μ and the Ehresmann connection one-form.

Take an open covering $\{U_i\}$ of M^4 and let s_i be a local section on each U_i. Using the Ehresmann connection ω define the Lie algebra-valued one-form A_i on U_i by[4]:

$$A_i \equiv s_i^* \omega \in \bigwedge^1 (U_i) \otimes \mathfrak{g}. \tag{102}$$

This shows how to get the gauge fields from the connection one-form. The inverse is also possible, given a gauge field and a section

$$s_i : U_i \to \pi^{-1}(U_i),$$

we can reconstruct a connection one-form ω by the following Theorem

Theorem 14.2.1

Given a $\mathfrak{g}$-valued one-form A_i on U_i and a local section

$$s_i : U_i \to \pi^{-1}(U_i) \ ,$$

there exists a connection one-form ω whose pull-back by

$$s_i^* \text{ is } A_i$$

Remark that the connection one-form ω can be defined globally, while the Lie algebra-valued one-form A_i cannot due to the **local** sections s_i. Theorem 14.2.1 is an existence Theorem for a connection one-form ω under a given gauge potential A_i in U_i but it does not say whether it is unique. The uniqueness requires an extra condition, called the **compatibility condition**. The condition follows from the fact that we need to accommodate that

$$\omega_i = \omega_j, \text{ on } U_i \cap U_j$$

with

$$\omega_i = \omega|_{U_i}.$$

This restriction clearly has something to do with the transition function associated to the change from U_i to U_j, so we can expect a statement that restricts the transition functions. The explicit form of this condition can derived by applying the connection one-form ω to Equation (103) in the following Lemma

4 The indices i refer to the covering and not to the space-time indices μ that accompany each A_i in U_i for a specific i.

Lemma 14.2.1

Let $P(M^4, G)$ be a principal fibre bundle and s_i, s_j local sections over U_i and U_j, respectively, such that

$$U_i \cap U_j \neq \emptyset.$$

For

$$X \in T_p M$$

with

$$p \in U_i \cap U_j,\ s_{i*}X,\ s_{j*}X$$

satisfy

$$s_{j*}X = R_{t_{ij}*}(s_{i*}X) + \left(t_{ij}^{-1} dt_{ij}(X)\right)^{\sharp}, \tag{103}$$

where

$$t_{ij} : U_i \cap U_j \to G$$

is the transition function.

After application of ω to Equation (103), and using the identity

$$\omega(s_{j*}) = s_j^* \omega$$

together with the second property of Definition 14.2.2, we obtain

$$A_j = t_{ij}^{-1} A_i t_{ij} + t_{ij}^{-1} dt_{ij}\ . \tag{104}$$

Now identifying the A_j with **gauge potentials** , we have for the components

$$A_{2\mu} = g^{-1}(p) A_{1\mu}(p) g(p) + g^{-1}(p) \partial_\mu g(p), \tag{105}$$

which is identical to a gauge transformation in gauge theory. In local coordinates it reads

$$A_i = \left(-ig A_\mu^a t^a dx_\mu\right)_i, \tag{106}$$

where g is now the coupling constant and t^a are the Lie algebra generators.

14.3 HORIZONTAL LIFT AND PARALLEL TRANSPORT

In the previous section we found that the tangent space TP of the principal fibre bundle $P(M^4, G, \pi)$ can be split in a horizontal and vertical part. This splitting allows us to define the so-called **horizontal lift** of a curve in the base manifold M^4.

Definition 14.3.1 (Horizontal lift)**.**
Let $P(M^4, G, \pi)$ *be a principal fibre bundle and*

$$\gamma : [0,1] \to M^4$$

a curve in M^4*. Then a curve*

$$\tilde{\gamma} : [0,1] \to P$$

is said to be a horizontal lift of γ *if the tangent vector to*

$$\tilde{\gamma}(t) \in H_{\tilde{\gamma}(t)}P.$$

From this Definition we derive the following Theorem

Theorem 14.3.1

Let

$$\gamma : [0,1] \to M^4$$

be a curve in M^4 and let

$$p \in \pi^{-1}(\gamma(0)).$$

Then there exist a unique horizontal lift $\tilde{\gamma}(t)$ in P such that

$$\tilde{\gamma}(0) = p$$

and corollary:

Corollary 14.3.1

Let $\tilde{\gamma}'$ be another horizontal lift of γ such that

$$\tilde{\gamma}'(0) = \tilde{\gamma}(0)g.$$

Then

$$\tilde{\gamma}'(t) = \tilde{\gamma}(t)g, \ \forall t \in [0,1]$$

The last statement of the Corollary makes the global gauge symmetry apparent, a global right action does not change the connection on the principal

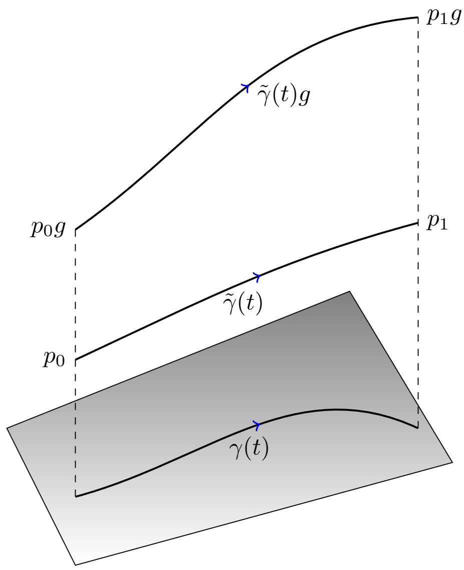


Figure 33: Horizontal lifts of a curve.

fibre. Let X be the tangent vector of $\gamma(t)$ at $\gamma(0)$, using the horizontal lift we have that

$$\widetilde{X} = \tilde{\gamma}_* X$$

is tangent to $\tilde{\gamma}$ at $p = \tilde{\gamma}(0)$. Given that this lifted tangent vector is horizontal by definition, we get

$$\omega(\widetilde{X}) = 0.$$

Using the fact that the transition functions are elements of G, Equation (103) can be rewritten as

$$\widetilde{X} = g_i^{-1}(t) s_{i*} X g_i(t) + \left(g_i^{-1}(t) dg_i(X)\right)^\sharp. \tag{107}$$

Application of the one-form ω to this result returns

$$0 = \omega(\widetilde{X}) = g_i^{-1}(t)\omega(s_{i*}X)g_i(t) + g_i^{-1}(t)\frac{dg_i(t)}{dt}, \tag{108}$$

the parallel transport equation. Expressing this result with gauge potentials by using

$$\omega(s_{i*}X) = s_i^* \omega(X) = A_i(X)$$

in Eq. (108) it follows that the parallel transport equation in the **local** form reads

$$\frac{dg_i(t)}{dt} = -A_i(X) g_i(t) \tag{109}$$

14.4 SUMMARY

After introducing principal fibre bundles and noticing that the gauge fields transform like a connection, we investigated the possibility of identifying gauge fields with such a connection. As it turns out this is possible, if we identify the gauge potentials with the pull-backs of sections of the Ehresmann one-form. Although this one-form is defined globally, i.e. the gauge symmetry applies all over the base manifold M^4, the gauge potentials can only be written down locally due to their dependence on the existence of local sections. As a second major result in this chapter we were able to use this relation between the gauge fields and sections of one-forms to derive the gauge theory variant of a parallel transport equation. In the next chapter we will introduce the mathematical tools that will eventually allow us to find, at least locally, a solution of this equation.

15

SOLVING MATRIX DIFFERENTIAL EQUATIONS BY CHEN ITERATED INTEGRALS

15.1 INTRODUCTION

In the previous chapter we have derived the parallel transport equation in gauge theory. Choosing a matrix representation for the Lie gauge group generators turns this equation into a matrix differential equation. The question then arises how to solve such an equation. Finding an answer to this question is the subject of this chapter. We will see that Chen iterated integrals play an important role in the set of solutions, moreover if the solution of the parallel transport equation is expressed with these integrals the Wilson line emerges which we will demonstrate explicitly in the next chapter.

15.2 DERIVATIVES OF A MATRIX FUNCTION

Assuming the reader is familiar with basic matrix theory, we will only define the derivative and product integral of a matrix function

$$A : [a, b] \to \mathbb{R}^{n\times n},$$

which is just a matrix-valued function. For now we restrict ourselves to real-valued matrices, but most definitions and properties can be straightforwardly extended to complex matrices.

If we want to define the derivative of a matrix function we first need to define what we mean with the **differentiability** of a matrix function.

Definition 15.2.1 (Differentiablility of a matrix function)**.**
A matrix function

$$A : [a, b] \to \mathbb{R}^{n\times n}$$

is called differentiable at a point $x \in (a, b)$ *if all its entries*

$$a_{ij},\ i, j \in 1, ..., n$$

are differentiable at $x \in [a, b]$*, where the entries are considered to be real-valued functions*

$$a_{ij} : [a, b] \to \mathbb{R} \; .$$

If the matrix function A *is differentiable we use the notation:*

$$A'(x) = \left\{ a'_{ij} \right\}_{i,j=1}^{n} \; . \tag{110}$$

Having a differentiable matrix function A we can now define not one, but two derivatives.

Definition 15.2.2 (Left and Right derivative of a matrix function)**.** *Let*

$$A : [a, b] \to \mathbb{R}^{n \times n}$$

be a differentiable and regular (single-valued and analytic) matrix function at $x \in (a, b)$ *then we define the left derivative of* A *at* x *as*

$$\frac{d}{dx} A(x) = A'(x) A^{-1}(x) = \lim_{\Delta x \to 0} \frac{A(x + \Delta x) A^{-1}(x) - I}{\Delta x} \; , \tag{111}$$

and similarly the right derivative as:

$$A(x) \frac{d}{dx} = A^{-1}(x) A'(x) = \lim_{\Delta x \to 0} \frac{A^{-1}(x) A(x + \Delta x) - I}{\Delta x} \; . \tag{112}$$

Similar to the case of scalar real valued functions the derivatives at the end-points of the interval $[a, b]$ are defined[1] by using the matrix entries a_{ij}. Both, the left and right derivatives of a matrix function share many properties with the common derivatives of functions, but care must be taken in some cases. To demonstrate this we just mention the application to a product.

Theorem 15.2.1

1 Here the left and right refer to approaching the endpoints of the interval from the left or the right and not to the derivatives of the matrix function.

Let

$$A_1, A_2 : [a_1, a_2] \to \mathbb{R}^{n\times n}$$

be differentiable and regular matrix functions at

$$x \in (a_1, a_2).$$

Hence one gets

$$\frac{d}{dx}(A_1A_2) = \frac{d}{dx}A_1 + A_1\Big(\frac{d}{dx}A_2\Big)A_1^{-1} = A_1\Big(A_1\frac{d}{dx} + \frac{d}{dx}A_2\Big)A_1^{-1} \quad (113)$$

$$(A_1A_2)\frac{d}{dx} = A_2\frac{d}{dx} + A_2^{-1}\Big(A_1\frac{d}{dx}\Big) = A_2^{-1}\Big(A_1\frac{d}{dx} + \frac{d}{dx}A_2\Big)A_2. \quad (114)$$

Theorem 15.2.2

Let

$$A_1, A_2 : [a, b] \to \mathbb{R}^{n\times n}$$

be differentiable and regular matrix functions at

$$x \in (a, b),$$

such that:

$$\frac{d}{dx}A_1 = \frac{d}{dx}A_2, \quad (115)$$

then there exists a constant matrix $A_3 \in \mathbb{R}^{n\times n}$ such that

$$A_2(x) = A_1(x)A_3,\ \forall\ x \in \{a, b\}.$$

15.3 PRODUCT INTEGRAL OF A MATRIX FUNCTION

Naturally, after having defined the derivatives of a matrix function, we are now interested in integrals of matrix functions. Let

$$A : [a, b] \to \mathbb{R}^{n\times n}$$

be a matrix function and D a partition of the interval $[a, b]$ defined as

$$D : a = t_0 \le \xi_1 \le t_1 \le \xi_2 \le \cdots \le t_{m-1} \le \xi_m \le t_m = b, \quad (116)$$

Next we introduce the notations:

$$\Delta t_i = t_i - t_{i-1},\ i = 1, ..., m \tag{117}$$

$$\nu(D) = \max_{1 \le i \le m} \Delta t_i, \tag{118}$$

and:

$$P(A, D) = \prod_{i=m}^{1} (I + A(\xi_i)\Delta t_i) = (I + A(\xi_m)\Delta t_m) \cdots (I + A(\xi_1)\Delta t_1) \tag{119}$$

$$P^*(A, D) = \prod_{i=1}^{m} (I + A(\xi_i)\Delta t_i) = (I + A(\xi_1)\Delta t_1) \cdots (I + A(\xi_m)\Delta t_m). \tag{120}$$

Volterra [120] then defined the left and right integral of the matrix function A as:

$$\int_a^b \{a_{ij}\} = \lim_{\nu(D)\to 0} P(A, D), \qquad \text{Left integral} \tag{121}$$

$$\{a_{ij}\} \int_a^b = \lim_{\nu(D)\to 0} P^*(A, D), \qquad \text{Right integral,} \tag{122}$$

where

$$\lim_{\nu(D)\to 0} M(D) = M, \tag{123}$$

is defined as

$$\forall\, \epsilon > 0,\ \exists \delta > 0,$$

such that

$$|M(D)_{ij} - M_{ij}| < \epsilon$$

for every partition D of $[a, b]$ as defined in Eq. (116). Starting from Volterra's definition for the integrals we can now define the left and right product integrals

Definition 15.3.1 (Left and Right Product Integrals)**.**
Let

$$A, B : [a, b] \to \mathbb{R}^{n\times n}$$

be a matrix function. If the limits

$$\lim_{\nu(D)\to 0} P(A,D) = \prod_a^b (I + A(t)dt), \tag{124}$$

$$\lim_{\nu(D)\to 0} P^*(A,D) = (I + A(t)dt)\prod_a^b \tag{125}$$

exist then they are called respectively the left and right product integral of A over $[a,b]$.

The relation with the usual Riemann integrals, becomes clear by noticing that a matrix function A is Riemann integrable if its matrix entries a_{ij} are Riemann integrable functions on $[a,b]$. In this case one has

$$\int_a^b A(t)dt = \left\{ \int_a^b a_{ij}(t)\ dt \right\}_{i,j=1}^n. \tag{126}$$

Assuming Riemann integrability we can expand the integrals of a matrix function which will allow us to relate them to the Chen iterated integrals. This expansion is given by the following Theorem:

Theorem 15.3.1

Let

$$A : [a,b] \to \mathbb{R}^{n\times n}$$

be a Riemann integrable matrix function, then the left and right product integrals exist and are given by[a]

$$\prod_a^x (I + A(t)dt) = \quad I + \sum_{k=1}^{\infty} \int_a^x \int_a^{t_k} \cdots \int_a^{t_2} A(t_k)\cdots A(t_1)dt_1\cdots dt_k, \tag{127}$$

$$(I + A(t)dt)\prod_a^x = \quad I + \sum_{k=1}^{\infty} \int_a^x \int_a^{t_k} \cdots \int_a^{t_2} A(t_1)\cdots A(t_k)dt_1\cdots dt_k, \tag{128}$$

where the series converge absolutely and uniformly for

$$x \in [a,b].$$

a Notice the ordering of the matrix functions under the integral signs.

The following Theorem 15.3.1 can then be proven

Theorem 15.3.2

Let

$$A : [a_1, b] \to \mathbb{R}^{n\times n}$$

be a Riemann integrable matrix function, and let

$$Y_1(x) = \prod_a^x (I + A(t)dt), \tag{129}$$

$$Y_2(x) = (I + A(t)dt) \prod_a^x, \tag{130}$$

then

$$\forall\ x \in [a_1, b]$$

they satisfy the integral equations

$$Y_1(x) = I + \int_a^x A(t)Y_1(t)\ dt \tag{131}$$

$$Y_2(x) = I + \int_a^x Y_2(t)A(t)\ dt. \tag{132}$$

15.4 CONTINUITY OF MATRIX FUNCTIONS

In order to continue towards our goal of finding solutions to the type of differential matrix equation that was generated in the derivation of the parallel transport equation in gauge theory we need to define the continuity of matrix functions. Similarly to the differentiability of the matrix function we define continuity through the matrix entries a_{ij}.

Definition 15.4.1. *Let*

$$A : [a, b] \to \mathbb{R}^{n\times n}$$

be a matrix function, then we call A continuous if the entries a_{ij} of A are continuous functions on $[a, b]$.

This definition now permits us to write down the types of differential equations we require, which are obtained by differentiating the integral equations of Theorem 15.3.2.

Theorem 15.4.1

Let

$$A : [a, b] \to \mathbb{R}^{n\times n}$$

be a continuous matrix function, then

$$\forall\ x \in [a, b]$$

the product integrals

$$Y_1(x) = \prod_a^x (I + A(t)dt), \tag{133}$$

$$Y_2(x) = (I + A(t)dt) \prod_a^x \tag{134}$$

satisfy the conditions

$$Y_1'(x) = A(x)Y_1(x), \tag{135}$$

$$Y_2'(x) = Y_2(x)A(x). \tag{136}$$

Using the symbolic notations for the left and right derivatives defined in section 15.2 the Equations (135) and (136) can be rewritten as

$$\frac{d}{dx} \prod_a^x (I + A(t)dt) = A(x),$$

$$(I + A(t)dt) \prod_a^x \frac{d}{dx} = A(x). \tag{137}$$

Moreover, we have

Corollary 15.4.1

A function

$$Y : [a, b] \to \mathbb{R}^{n\times n}$$

is a solution of the equation:

$$Y'(x) = A(x)Y(x), \ x \in [a, b] \tag{138}$$

and satisfies

$$Y(a) = I$$

if and only if Y solves the integral equation

$$Y(x) = I + \int_a^x A(t)Y(t)dt, \ x \in [a, b] \tag{139}$$

From the above it is now evident that solutions of Eqs. (135) and (136) respectively can be written as

$$Y_1(x) = I + \sum_{k=1}^{\infty} \int_a^b \int_a^{x_k} \cdots \int_a^{x_2} A(x_k) \cdots A(x_1) dx_1 \cdots dx_k, \tag{140}$$

$$Y_2(x) = I + \sum_{k=1}^{\infty} \int_a^b \int_a^{x_1 k} \cdots \int_a^{x_2} A(x_1) \cdots A(x_k) dx_1 \cdots dx_k, \tag{141}$$

which will be compared to the expressions given in Example 15.5.1.

All the above properties and theorems can by readily extended to matrix functions

$$A : [a, b] \to \mathbb{C}^{n \times n},$$

such that this is not an obstacle when considering matrix representations of gauge groups such as, for example, $SU(N)$.

15.5 ITERATED INTEGRALS AND PATH ORDERING

In this Section we relate the product integrals from the previous Section with Chen's iterated integrals which will also make the link with Wilson lines apparent. To this end we start with an example from Chapter 4 of the book by Peskin and Schroeder [19]

Example 15.5.1.
Consider the Schrödinger equation for a quantum evolution operator (recall the discussion in the first Chapter) in the interaction representation

$$i\partial_t U(t) = H(t)U(t), \ U(0) = 1 \tag{142}$$

where $H(t)$ is the interaction Hamiltonian, an operator function acting on Hilbert space. This unitary operator can be treated as well as a complex-valued scalar matrix function

$$U(t) : [0, t] \to \mathbb{C}.$$

The iterated integrals which contribute to the solution of Eq. (142) can be rewritten as

$$\int_0^t \int_0^{t_1} \cdots \int_0^{t_{l-1}} H(t_1) \cdots H(t_l) dt_1 \cdots dt_l = \frac{1}{l!} \int_0^t dt_1 \cdots dt_l \mathbf{T}\{H(t_1) \cdots H(t_l)\}, \tag{143}$$

where $\mathbf{T}$ indicates the time-ordering operation for the Hamilton operator $H(t)$. That is this operator orders the $H(t)...H(t')$ in time.

The previous expression then allows for the formal notation for the unitary operator $U(t)$

$$U_\tau(t) \equiv \mathbf{P} \exp\left[-i \int_0^t dt' H(t')\right], \tag{144}$$

which could be interpreted as a parallel propagator along a path through the time axis

$$\tau = [0, t].$$

Consider now the matrix function

$$A : [0, 1] \to \mathbb{C}^{n \times n},$$

so that A can be written as

$$A = S \circ \phi$$

where

$$\phi : [0, 1] \to M$$

$$t \mapsto \phi(t) = x^\mu(t)$$

and

$$S : M \to \mathbb{C}^{n \times n}$$

$$x^\mu \mapsto S(x^\mu) = A(x(t)).$$

Applying the same reasoning as in Example 15.5.1, we see that the equation

$$Y'(t) = A(t)Y(t), \tag{145}$$

has a unique solution

$$Y(t) = \mathbf{T}\exp\left[\int_0^t dt'\ A(t')\right] = \mathcal{P}\exp\left[\int_0^y dx\ S(x)\right] \tag{146}$$

given the initial condition

$$Y(0) = 1$$

and where the time-ordering is now replaced with **path-ordering**, which orders the operators $S(x)$ along the path in the manifold M. In this case the variable that parametrizes the path acts as the time parameter t from the example.

We will return to this type of equations in the next Chapter, after a brief discussion on the relation between product integrals and the Chen integrals from Chapter 13.

Returning to Eq. (73), it is easy to see that the **operators** ω_i are ordered under the integral sign. So one might as well rewrite this as:

$$\int_0^1 \left(\int_{\gamma^t} \omega_1 \cdots \omega_{r-1}\right) \omega_r(t)dt = \mathcal{P}\left\{\int_\gamma \cdots \int_\gamma \omega_1 \cdots \omega_r\right\}, \tag{147}$$

where we considered the integrals between the braces as ordinary integrals and not as a Chen Iterated integrals. Using this result we can rewrite the function $Y(t)$ from equation (146) with Chen iterated integrals:

$$Y(t) = \mathcal{P}\exp\left[\int_0^y dx\ S(x)\right] = \exp\left[\int_\gamma S\right], \tag{148}$$

if one identifies the operator $S(x)\ dx$ (interpreted as a form) with the forms

$$\omega = \omega_1 = \cdots = \omega_r$$

from (73). Some care is necessary with this last statement about the ω_i. We can indeed all identify them with ω, which will still depend the coordinates x^μ after having chosen a coordinate chart. Consider the simple example

$$\omega_1\omega_2 \mapsto \omega(x_1)\omega(x_2)$$

to clarify this statement.

15.6 SUMMARY

Starting from the properties of the derivatives and integrals of matrix functions we were able to express the solution of a matrix differential equation of the type to which also the parallel transport equation belongs with Chen iterated integrals. We showed that, considering such solutions, path-ordering naturally emerges and is not introduced by hand. In the next chapter we will apply this acquired knowledge to write down (locally) an explicit solution for the parallel transport equation.

16 WILSON LINES, PARALLEL TRANSPORT AND COVARIANT DERIVATIVE

16.1 INTRODUCTION

This Chapter is devoted to showing the link between the local solution of the parallel transport equation and Wilson lines. Furthermore we will discuss how the covariant derivative follows from the parallel transport equation, also explicitly demonstrating the link with Wilson lines during that process. The relation between the field strength tensor and the curvature of the gauge potentials will be discussed. We end with defining the holonomy group, and stating the Ambrose-Singer theorem that ultimately allows for the gauge theory to be expressed in a loop space setting.

16.2 PARALLEL TRANSPORT AND WILSON LINES

Returning to the parallel transport equation in gauge theory Eq. (109)

$$\frac{dg_i(t)}{dt} = -A_i(X)g_i(t), \tag{149}$$

where A_i is a Lie algebra-valued (i.e. a complex matrix when considering matrix representations for the Lie algebra) one-form. Given the initial condition

$$g_i(0) = e,$$

where e is the unit element in the gauge group G, a solution can be expressed using product integrals or Chen integrals, yielding (locally) the formal solution in the form of a functional of an arbitrary path $\gamma(t)$

$$g_i[\gamma(t)] = \mathcal{P}\exp\left[-\int_0^t A_{i\mu}(x(t))\frac{dx^\mu}{dt}dt\right]_\gamma \tag{150}$$

$$= \mathcal{P}\exp\left[-\int_{\gamma(0)}^{\gamma(t)} A_{i\mu}(x)dx^\mu\right]$$

$$= \exp\left[-\int_\gamma A_i\right] \tag{151}$$

where

$$A_{i\mu} = ig\mathcal{A}^a_{i\,\mu}t^a$$

with horizontal lift

$$\tilde{\gamma}(t) = s_i[\gamma(t)]g_i[\gamma(t)]. \tag{152}$$

In Eq. (151) the integrals are interpreted as Chen iterated integrals. More specifically we find that if

$$g_0 \in \pi^{-1}[\gamma(0)],$$

then

$$g_1 \in \pi^{-1}[\gamma(1)]$$

is the parallel transport of g_0 along the curve γ

$$\Gamma(\tilde{\gamma}) : \pi^{-1}[\gamma(0)] \to \pi^{-1}[\gamma(1)],\ g_0 \mapsto g_1.$$

Introducing a coordinate chart we can thus write locally:

$$g_1 = s_i(1)\ \mathcal{P}\exp\left[-\int_0^1 A_{i\mu}\frac{dx^\mu}{dt}dt\right]. \tag{153}$$

The relation with **Wilson lines** is now straightforward when considering Eq. (150). Put differently, a Wilson line along a path γ is the trace of the parallel transporter along this path, this explains the term **gauge link** we used in the introduction. Using the properties of the principal fibre bundle formalism we obtain

$$R_g\Gamma(\tilde{\gamma})(g_0) = g_1 g$$

and

$$\Gamma(\tilde{\gamma})R_g(g_0) = \Gamma(\tilde{\gamma})(g_0 g),$$

which together with the fact that $\tilde{\gamma}(t)g$ is the horizontal lift through $g_0 g$ and $g_1 g$ returns that $\Gamma(\tilde{\gamma})$ commutes with the right action. We also mention the following properties for the parallel transporter, which are easily proved from the properties of Chen iterated integrals

(i) Inverses:

$$\Gamma(\tilde{\gamma}^{-1}) = (\Gamma(\tilde{\gamma}))^{-1}.$$

(ii) Composition: Let

$$\gamma_{1,2} : [0,1] \to M,$$

be two curves such that

$$\gamma_1(1) = \gamma_2(0),$$

then

$$\Gamma(\widetilde{\gamma_1\gamma_2}) = \Gamma(\tilde{\gamma_2}) \circ \Gamma(\tilde{\gamma_1}).$$

16.3 DIRECTIONAL DERIVATIVE, WILSON LINE AND COVARIANT DERIVATIVE

Having an expression for the parallel transporter, a covariant derivative can be defined in the usual way. To show how the Wilson line generates the covariant derivative consider the so-called **directional derivative** of a matter field $\psi(x)$ which transform under a local symmetry transformation $U(x)$ as

$$\psi(x) \mapsto U(x)\psi(x).$$

The directional derivative along a vector V^μ is given by

$$V^\mu \partial_\mu \psi(x) = \lim_{\Delta\to 0} \frac{(\psi(x+\Delta\cdot V) - \psi(x))}{\Delta}. \quad (154)$$

It is immediately clear that this derivative is not well-defined due the fact that we consider matter fields at different space-time points which transform differently because gauge transformations are local transformations depending on the space-time point. This issue also emerges in general relativity when comparing vector or tensor fields at different space-time points on a curved background. In the case of general relativity this is solved by parallel transporting the fields by using the Levy-Civita connection. We can now do the same thing in gauge theory, where the gauge fields themselves act as connections, as we have discussed in the previous sections. Therefor we can use the gauge version of the parallel transporter, the Wilson line, to transport one of the fields in Eq. (154) to the location of the other field. To demonstrate this explicitly, we start from the Wilson line

$$U_\gamma[x,y] = \mathcal{P}\exp\left[ig\int_x^y A_\mu(z)dz^\mu\right]_\gamma,$$

where γ is the path along which one parallel transports between the spacetime points x and y, again $\mathcal{P}$ denotes to the path ordering operator. From the above expression for the Wilson line it is easy to see that it transforms under gauge transformations generated by $U(x)$ as

$$U_\gamma[x,y] \mapsto U(y)U_\gamma[x,y]U^{-1}(x),$$

such that the parallel transported Dirac field $U_\gamma[x,y]\psi(x)$ now transforms exactly as $\psi(y)$. This makes it now possible to 'compare' the matter fields defined at different points, which in its turn allows us adapt the directional derivative to a well-defined derivative. The adapted directional derivative is given by

$$V^\mu D_\mu\psi(x) = \lim_{\Delta\to 0}\left(\psi(x+\Delta\cdot V) - U\left(x+\Delta\cdot V;x\right)\psi(x)\right), \tag{155}$$

where we dropped the path γ which now is an infinitesimal straight line along the vector V_μ. Expanding the parallel transporter and the field at $x+\epsilon n$ to first order in ϵ we have

$$U[x+\Delta\cdot V;x] = 1 + ig\epsilon V^\mu A_\mu(x) + \mathcal{O}(\Delta^2) \tag{156}$$

$$\psi(x+\Delta\cdot V) = \psi(x) + \Delta V^\mu\partial_\mu\psi(x) + \mathcal{O}(\Delta^2). \tag{157}$$

Inserting this result in (155) we finally arrive at

$$V^\mu D_\mu\psi(x) = V^\mu(\partial_\mu\psi(x) - igA_\mu(x)\psi(x)), \tag{158}$$

which is nothing more than the **covariant derivative.**

The above derivation can be physically interpreted in a way, similar to the situation in general relativity. The parallel transporter takes into account the interactions with the background, here a gauge field background[1]. Not surprisingly one will often find drawings in the literature that look like figure 34, where the double line is the Wilson line representing the interactions with a background gauge field as shown in the figure on the right. In Deep Inelastic Scattering for instance the source of this background field will originate from the nucleon on which we scatter with a lepton.

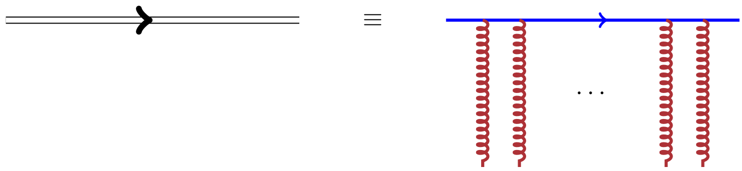


Figure 34: Graphical representation of a Wilson line associated with a quark.

16.4 HOLONOMY, CURVATURE AND THE AMBROSE-SINGER THEOREM

16.4.1 *Holonomy*

In the previous Section we have discussed the relation between Wilson lines, the parallel transport and how to construct a covariant derivative using these concepts. Turning our attention to Wilson loops, naturally the concept of holonomy and holonomy group surfaces.

Let $P(M, G)$ again be a fibre bundle and let γ_1 and γ_2 be two curves in M, such that

$$\gamma_1(0) = \gamma_2(0) = p_0$$

and

$$\gamma_1(1) = \gamma_2(1) = p_1.$$

If we consider the horizontal lifts of these curves for which

$$\tilde{\gamma}_1(0) = \tilde{\gamma}_2(0) = u_0,$$

then we do not necessarily get

$$\tilde{\gamma}_1(1) = \tilde{\gamma}_2(1).$$

1 Field strength tensor resembles gauge curvature $F_{\mu\nu} \sim$ curvature in general relativity.

This means that if we consider a loop γ in M, i.e.,

$$\gamma(0) = \gamma(1),$$

then, in general, the horizontal lift does not yield unavoidably

$$\tilde{\gamma}(0) \neq \tilde{\gamma}(1).$$

The loop γ thus induces a map

$$\tau_\gamma : \pi^{-1}(p) \to \pi^{-1}(p)$$

on the fibre at p. In the previous Section we have discussed that the horizontal lift $\Gamma(\tilde{\gamma})$ commutes with the right action such that we can write

$$\tau_\gamma(ug) = \tau_\gamma(u)g.$$

Considering now loops with fixed base-point in the manifold M, denoted by $C_p(M)$, τ_γ can only reach certain elements of G, combining them with a gauge transformation[2] we can reach all elements of G. The set of elements that can be reached starting from the point (p, u) in the principal fibre bundle form a subgroup of the structure group G and generate the **holonomy group** at u, where $\pi(u) = p$

$$\Phi_u = \{g \in G | \tau_\gamma(u) = ug,\ \gamma \in C_p M\}.$$

An interesting fact is that

$$\tau_{\gamma^{-1}} = \tau_\gamma^{-1}$$

inducing

$$g_{\gamma^{-1}} = g_\gamma^{-1}.$$

Since holonomy elements are generated by parallel transport around a loop, we have by the previous section that we can rewrite them as

$$g_\gamma = \mathcal{P} \exp\left[-\oint_\gamma A_{i\mu}(x) dx^\mu\right].$$

2 This fact determines that when we try to reconstruct the gauge potential from Wilson loops, we will only be able to determine them up to a gauge transformation.

16.4.1.1 *Curvature*

Having the picture of parallel transport around a loop on a curved space-time manifold in mind, naturally the idea of curvature follows. So let us introduce curvature in a principal fibre bundle setting. The previous Sections explained how the covariant derivative arises in gauge theory from parallel transport but, to introduce the curvature two-form we need the covariant derivative to be described in a more rigorous way.

Definition 16.4.1 (Covariant Derivative)**.**
Let V be a vector space of dimension k, and let $\{e_\alpha\}$ be a basis in V. Let

$$\phi \in \bigwedge^r(P) \otimes V$$

that is

$$\phi : \mathrm{TP} \wedge \cdots \wedge \mathrm{TP} \to V$$

and

$$X_1, \cdots, X_{r+1} \in T_uP \ .$$

The covariant derivative acting on

$$\phi = \sum_{\alpha=1}^{k} \phi^\alpha \otimes e_\alpha$$

is then defined as:

$$D\phi(X_1, \cdots, X_{r+1}) \equiv d_P\phi\Big(X_1^H, \cdots, X_{r+1}^H\Big),$$

with

$$d_P\phi \equiv d_P\phi^\alpha \otimes e_\alpha,$$

where d_P is the exterior differential for the fibre bundle P.

The **curvature** follows from this Definition:

Definition 16.4.2 (Curvature two-form)**.**
The curvature two-form Ω *is the covariant derivative of the Ehresmann connection one-form* ω

$$\Omega \equiv D\omega \in \bigwedge^2 P \otimes \mathfrak{g}.$$

Again considering a right action, the result of this action on curvature is expressed by the Proposition

Proposition 16.4.1

The curvature transforms under right action of an element of the structure group of the fibre bundle as

$$R_g^* \Omega = g^{-1} \Omega g, \;\; g \in G. \tag{159}$$

In gauge theory notations this becomes

$$R_g^* F_{\mu\nu} = g^{-1} F_{\mu\nu} g,$$

where $F_{\mu\nu}$ is the gauge-covariant field strength. The curvature is better known in its form expressed by **Cartan's structure equation** which will also be familiar when expressed with field strength tensors.

Theorem 16.4.1: Cartan's Structure Equation

Let

$$X_1, X_2 \in T_u P.$$

Therefore the curvature Ω and the Ehresmann connection ω satisfy the Cartan structure equation

$$\Omega(X_1, X_2) = d_P \omega(X_1, X_2) + [\omega(X_1), \omega(X_2)], \tag{160}$$

which can also be written as:

$$\Omega = d_P \omega + \omega \wedge \omega. \tag{161}$$

Using this equation the gauge curvature or field strength tensor $F_{\mu\nu}$ is defined as:

$$F_{\mu\nu} = d_P A_{\mu\nu} + A_\mu \wedge A_\nu = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \tag{162}$$

which should look more familiar.

Having defined the holonomy and curvature we are now ready to discuss the Ambrose-Singer Theorem.

16.4.2 *The Ambrose-Singer theorem.*

The connection of Wilson loops with holonomies is assumed to allow one, in principle, to recast gauge theory in the space of generalized loops. The **Ambrose-Singer theorem** forms the core of this program.

Theorem 16.4.1: Ambrose-Singer

Let $P(M,G)$ be a principal fiber bundle with connection ω, and curvature form Ω. Let $\Phi(u)$ be the holonomy group with reference point

$$u \in P(M,G)$$

and $P(u)$ the holonomy bundle of ω through u. Then the Lie algebra of $\Phi(u)$ is equal to the Lie sub-algebra of $\mathfrak{g}$, generated by all elements of the form $\Omega_p(v_1, v_2)$ for $p \in P(u)$ and v_1, v_2 horizontal vectors at p, where $\mathfrak{g}$ is the Lie algebra of G

In other words, this theorem states that all the information contained in the curvatures at a point in the principal fibre bundle P with connection ω can also be found in the holonomy group $\Phi(u)$ at that point. This means that, in principal, it should be possible to express all physical observables as functions of the holonomies, instead of as functions of the gauge potentials. Assuming the structure group G is sufficiently well-behaved, such that any element of G can be reached from an element of $\Phi(u)$ by a gauge transformation (right action of G onto itself), we can express all physical observables as functions of the holonomies[3]. Notice that the holonomy group is not invariant under gauge transformation such that real observables will need to be described by gauge invariant functions of these holonomies. In the next Chapters and Sections we will assume that Wilson Loop Functionals (see below) can play the role of these gauge invariant functions of the holonomies.

3 This should always be the case, otherwise there are elements of G or thus elements in the fibre that can not be reached by a gauge transformation which was the whole point of considering gauge theories.

16.4.2.1 *Wilson Loop Functional.*

For the last part of this Chapter we will summarize some of the properties of Wilson lines and loops from a gauge theory point of view and we introduce the gauge invariant Wilson loop functionals, which in the next Sections will be used to introduce and study GLS. Wilson lines are essentially solutions of the parallel transport equations which for a curve or open path is given by

$$U_\gamma = \mathcal{P}\exp\left[\int_\gamma A_\mu\right], \tag{163}$$

where we also introduced the commonly used notation for a Wilson line or loop U_γ. If γ is a loop this becomes

$$\Gamma_\gamma \equiv U_\gamma = \mathcal{P}\exp\left[\oint_\gamma A_\mu\right]. \tag{164}$$

Important to notice is that this infinite series, when one expands the exponential, converges (absolutely, see [16]) to an element $g \in G$. Since the parallel transporter acts as an operator on Hilbert space, due to this convergences becomes a bounded operator such that it will also make sense to introduce it's Trace (Definition 9.2.2). As we discussed before the gauge link is not gauge invariant but transform as

$$U_\gamma^g = g_y^{-1} U_\gamma g_x, \tag{165}$$

for a path γ from x to y or as

$$U_\gamma^g = g_x^{-1} U_\gamma g_x, \tag{166}$$

when γ is a loop with base point

$$x = \gamma(0).$$

Since observables are by definition gauge invariant, and as we will see, the advantage of using generalized loop space is its gauge invariance, we define the gauge invariant Wilson path/loop functional

$$W : \mathcal{LM} \to \mathbb{C}$$

by:

$$W(\gamma) = \frac{1}{N}\text{Tr } U_\gamma, \tag{167}$$

where $\mathcal{LM}$ represents the space of all loops in M. By continuity of the trace and the expansion of the exponential in Chen integrals we get

$$W(\gamma) = \frac{1}{N} \sum_{r \geq 0} \mathsf{Tr} \int_\gamma \underbrace{\omega\omega \cdots \omega}_{r} \tag{168}$$

with as before the convention that

$$\int_\gamma \underbrace{\omega\omega \cdots \omega}_{r} = Id,$$

if $r = 0$. Expressed with the gauge potentials A_μ this Wilson loop can be written as

$$W_\gamma = \frac{1}{N} \mathsf{Tr}\, \mathcal{P} \exp\left[\int_\gamma A_\mu\right], \tag{169}$$

for open paths and

$$W_\gamma = \frac{1}{N} \mathsf{Tr}\, \mathcal{P} \exp\left[\oint_\gamma A_\mu\right], \tag{170}$$

for loops. Both expressions are now gauge invariant, due to the traces. In terms of d-paths these Wilson loop functionals are complex-valued d-paths

$$W_\gamma \in \mathrm{Alg}(Sh(\Omega), \mathbb{C}),$$

i.e. they vanish on the ideal $I(d,p)$ defined in Chapter 12.

17
GENERALIZATION OF MANIFOLDS AND DERIVATIVES

17.1 INTRODUCTION

In the next Chapter we shall introduce the space of generalized loops. This space, as we will demonstrate, carries a manifold like structure, but not exactly. It differs from the common manifold in the sense that it is not **locally homeomorphic to** a Euclidean space $\mathbb{R}^n$, but to a **Banach space**. The goal of this chapter is to extend the manifold concept to such spaces, which are now allowed to be **infinite dimensional.** Having such a generalization of manifolds we can also extend the concept of derivatives, one such generalization is the Fréchet derivative that will turn out to be very useful for doing variational calculus on generalized loop space. This brief introduction is based on the works [97, 102, 121–123] where much more information is available. Here we restricted to the bare minimum for understanding the structures on generalized loop space.

17.2 MANIFOLD: FRÉCHET DERIVATIVE AND BANACH MANIFOLD

The definition of generalized manifolds as discussed in the introduction of this Chapter, requires the Fréchet derivative. So let us first discuss and define this derivative. This derivative is defined on Banach spaces and can be interpreted as a generalization of the derivative of a one parameter real-valued function to the case of a vector-valued function depending on multiple real values which is what we will need to define derivatives on the generalized loop space and is actually necessary to define the functional derivative in this space as we will see.

To give the definition of the Fréchet derivative we need a bounded linear operator, which is defined as

Definition 17.2.1 (Bounded Linear Operator)**.**
A bounded linear operator is a linear transformation L between normed vector spaces X and Y for which the ratio of the norm of $L(v)$ to that

of v is a bounded by the same number, over all non-zero vectors $v \in X$. Therefore

$$\exists\ M > 0,$$

such that

$$\forall\ v \in X:\ \ \|L(v)\|_Y \leq M\|v\|_X.$$

*The smallest M is called the **operator norm***

$$\|L\|_{\mathrm{op}} \text{ of } L.$$

A bounded linear operator is generally **not** a bounded function, which would require that the norm of $L(v)$ be bounded for all v, which is not possible unless Y is the zero vector space. In other words, a bounded linear operator is a locally bounded function. Recall, with respect to continuity, that a linear operator on a metrizable vector space is bounded if and only if it is continuous.

Having defined bounded linear operators we can now introduce the **Fréchet derivative**.

Definition 17.2.2 (Fréchet Derivative)**.**
Let X_1, X_2 be Banach spaces[a], and

$$U \subset X_1$$

an open subset. Then a function

$$F : U \to X_2$$

is called Fréchet differentiable at

$$x \in U$$

if there exists a bounded linear operator

$$A_x :\ \ X_1 \to X_2$$

such that

$$\lim_{\Delta \to 0} \frac{\|F(x+\Delta) - F(x) - A_x(\Delta)\|_{X_2}}{\|\Delta\|_{X_1}} = 0, \tag{171}$$

where the limit is defined as in the usual sense. If this limit exist, then

$$DF(x) = A_x$$

stands for the Fréchet derivative.

We call the function F, C^1 if

$$\mathbf{DF}: \ U \to B(X_1, X_2) \ ; \ x \mapsto DF(x) = A_x, \tag{172}$$

is continuous[b], B here highlights the fact that this is the space of bounded linear operators.

a Complete vector spaces with norm.

b Note the difference with the continuity of $DF(x)$.

Notice that the usual derivative is included in this definition. To demonstrate this, let us take

$$F : \mathbb{R} \to \mathbb{R},$$

such that $DF(x)$ is the function

$$t \mapsto t \ F'(x).$$

The Fréchet derivative can be extended to arbitrary Topological Vector Space (TVS)s. Where TVSs are vector spaces with a topology that makes the addition and scalar multiplication operations continuous, i.e. the topology is consistent with the linear structure of the vector space.

Definition 17.2.3 (Fréchet Derivative for Topological Vector Spaces)**.**

Let now X_1, X_2 be Topological Vector Spaces with

$$U \in X_1$$

an open subset that contains the origin and given a function

$$F : U \to X_2$$

preserving the origin

$$F(0) = 0.$$

To continue it is necessary to explain what it means for this function to have 0 as its derivative. We call the function F tangent to 0 if for every open neighbourhood

$$W \subset X_2, \text{ of } 0_{X_2},$$

there is an open neighbourhood

$$V \subset X_1, \text{ of } 0_{X_1},$$

together with a function

$$H : \mathbb{R} \to \mathbb{R}$$

such that

$$\lim_{\Delta \to 0} \frac{H(\Delta)}{\Delta} = 0$$

and

$$\forall \Delta : F(\Delta)V \subset H(\Delta)W.$$

This somewhat strange constraint can be removed by defining f to be Fréchet differentiable at a point

$$x_0 \in U$$

given that there exists a continuous linear operator

$$\lambda : X_1 \to X_2,$$

such that

$$F(x_0 + \Delta) - F(x_0) - \lambda\Delta,$$

considered as a function of Δ, is tangent to 0.

It can further be demonstrated that if the Fréchet derivative exists, then it is **unique**. Similarly to the usual properties of differentiable functions we find that

- if a function is Fréchet differentiable at a point it is necessarily continuous at this point;
- sums and scalar multiples of Fréchet differentiable functions are differentiable.

Hence we conclude that the space of Fréchet differentiable functions at some point x form a subspace of the functions that are continuous at that point x. Both the chain and Leibniz rule hold whenever Y is an algebra and a **topological vector space** in which multiplication is continuous. This will turn out to be exactly the case for the space of generalized loops, where the algebra multiplication is the shuffle product. Using the above generalization of derivative we can extend the manifold concept to that of a Banach manifold

Definition 17.2.4 (Banach Manifold)**.**
Let X be a set. An atlas of class

$$C^r, r \geq 0, \text{ on } X$$

is defined as a collection of pairs (charts)

$$(U_i, \phi_i'), \ i \in I,$$

such that

(i) for each

$$i \in I, \ U_i \subset X, \ \bigcup_i U_i = X$$

(ii) for each $i \in I$, ϕ_i is a bijection from U_i onto an open subset $\phi_i(U_i)$ of some Banach space E_i and

$$\forall \ i, j : \phi_i(U_i \cap U_j)$$

is open in E_i.

(iii) The crossover map

$$\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

is a smooth function r-times continuously differentiable function

$$\forall i, j \in I$$

meaning that the r-th Fréchet derivative

$$\mathrm{D}^r(\phi_j \circ \phi_i^{-1}) : \phi_i(U_i \cap U_j) \to \mathrm{Lin}\,(E_i^r; E_j)$$

exists and is a continuous function with respect to the E_i-norm topology on subsets of E_i and the operator norm topology (i.e., the topology induced by a norm on the space of bounded linear operators, Definition (17.2.1)) on the space of linear operators

$$\mathrm{Lin}\left(E_i^r; E_j\right),$$

where E_i^r represents the fact that the r-times iterated application of the linear operator defines the r-th Fréchet derivative.

It can be shown there is a **unique topology** on X such that $\forall i \in I,\ U_i$ is open and $\forall i \in I, \phi_i$ is a homeomorphism. This topological space is assumed to be a Hausdorff space in most cases, but this is not necessary from the point of view of the formal definition.

If all the E_i are equal to the same space E, the atlas is called an E-atlas. However, it is not necessary that the Banach spaces E_i be the same space, or even isomorphic as topological vector spaces. But, if two charts (U_i, ϕ_i) and (U_j, ϕ_j) are such that

$$U_i \cap U_j \neq \emptyset\ ,$$

it clearly follows from the derivative of the crossover map

$$\phi_j \circ \phi_i^{-1}:$$

$$\phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

that

$$E_i \cong E_j$$

thus they are isomorphic as topological vector spaces. It is important to realize that the set of points $x \in X$, for which there is a chart

$$(U_i, \phi_i) : x \in U_i$$

and E_i isomorphic to a given Banach space E, is both open and closed. Hence, one can assume that, on each connected component of X, the atlas is an E-atlas for some fixed E.

Similar to the usual differentiable manifolds, a new chart (U, ϕ) is called **compatible** with a given atlas

$$\{(U_i, \phi_i | i \in I\}$$

if the crossover map

$$\phi_i \circ \phi^{-1} : \phi(U \cap U_i) \to \phi_i(U \cap U_i)$$

is an r-times continuously differentiable function $\forall i \in I$. Two atlases are compatible when each chart in one atlas is compatible with the other atlas. Compatibility of atlases defines an equivalence relation on the class of all possible atlases on X.

As with real smooth manifolds, a C^r-manifold structure on X is defined as a choice of an equivalence class of atlases on X of class C^r. If all the Banach spaces E_i are isomorphic as topological vector spaces (as is guaranteed to be the case if X is connected), then an equivalent atlas can be found for which they are all equal to some Banach space E. X is then called an E-**manifold**, or one says that X is **modeled** on E.

We end this discussion by making the remark that a **Hilbert manifold** is a special case of a Banach manifold in which the manifold is locally modeled on Hilbert spaces.

17.2.1 *Fréchet Manifold*

A possible next step in the generalization of manifolds are **Fréchet spaces**, which are a special kind of topological vector spaces. Fréchet spaces are **locally convex spaces which are complete with respect to a translation invariant metric and their metric does not need to be generated by a norm.** Notice that this means that not every Fréchet space is a Banach space, which requires a norm. Typical examples are spaces of infinitely differentiable functions.

Fréchet spaces can be defined in two different ways, by translational invariant metrics or by a family of semi-norms, both are given below. We start with the definition via translational invariant metrics.

Definition 17.2.5 (Fréchet Spaces: via Translation-Invariant Metrics)**.**

A topological vector space X is a Fréchet space if and only if it satisfies the following three properties:

(i) there is a local basis for its topology at every point, i.e., it is locally convex

(ii) its topology can be induced by a translation invariant metric, meaning that a subset

$$U \subset X$$

is open if and only if

$$\forall u_1 \in U,\ \exists \epsilon > 0 | \{u_2 :\ d(u_2, u_1) < \epsilon\} \subset U.$$

(iii) it is a complete metric space[a]

a Notice the difference with Banach space, for which a norm is required.

The second definition is build on a family of semi-norms.

Definition 17.2.6 (Fréchet Spaces via Family of Semi-Norms)**.**
A topological vector space X is a Fréchet space if and only if it satisfies the following three properties:

1. *it is a Hausdorff space*
2. *its topology may be induced by a countable family of semi-norms*

   $$\| \cdot \|_l,\ l = 0, 1, 2, \cdots .$$

   This means that a subset

   $$U \subset X$$

   is open if and only if

   $$\forall u_1 \in U, \exists\ K \geq 0,$$

   $$\epsilon > 0\ |\{u_2 : \|u_2 - u_1\|_l < \epsilon, \forall l \leq K\} \subset U.$$
3. *it is complete with respect to the family of semi-norms*

A sequence

$$(x_n) \in X$$

converges to x in the Fréchet space defined by a family of semi-norms if and only if it converges to x with respect to each of the given semi-norms.

Note that every Banach space is a Fréchet space, as the norm induces a translation invariant metric and the space is complete with respect to this metric.

The following examples demonstrate how to build a family of semi-norms that can be used to topologize a space by semi-norms, turning it into a Fréchet space.

Example 17.2.1.
The vector space of infinitely differentiable functions $C^\infty([0,1])$

$$F : [0,1] \to \mathbb{R}$$

becomes a Fréchet space with the semi-norms

$$|F|_{(l)} = \sup \left\{ \left| \frac{d^l}{dx^l} F(x) \right| : x \in [0,1] \right\}, \forall\ \mathbb{N} \ni l \geq 0.$$

A sequence (F_n) of functions converges to

$$F \in C^\infty([0,1])$$

if and only if

$$\forall\ \mathbb{N} \ni l \geq 0,$$

the sequence $(F_n^{(l)})$ converges uniformly to $F^{(l)}$, where

$$F^{(l)} = \frac{d^l}{dx^l} F(x).$$

Example 17.2.2.
The space $C^\infty(\mathbb{R})$ of infinitely differentiable functions

$$F : \mathbb{R} \to \mathbb{R}$$

turns into a Fréchet space with the semi-norms $\forall\, \mathbb{N} \ni k, n \geq 0$

$$|F|_{k,n} = \sup\left\{\left|\frac{d^k}{dx^k}F(x)\right| : x \in [-n, n]\right\}.$$

Example 17.2.3.
The vector space $C^m(\mathbb{R})$ *of all* m*-times continuously differentiable functions*

$$f : \mathbb{R} \to \mathbb{R}$$

becomes a Fréchet space with the semi-norms:

$$|f|_{k,n} = \sup\{|f^{(k)}(x)| : x \in [-n, n]\}, \quad \forall\, \mathbb{N} \ni n \geq 0,\ k = 0, \cdots, m.$$

Example 17.2.4.
Let H *be the space of entire (everywhere holomorphic) functions on the complex plane. Then the family of semi-norms:*

$$|f|_n = \sup\{|f(z)| : |z| \leq n\}\ ,$$

makes H *into a Fréchet space.*

Example 17.2.5.
Let H *be the space of entire (everywhere holomorphic) functions of exponential type* τ*. Then the family of semi-norms:*

$$|f|_n = \sup_{z \in \mathbb{C}} \exp\left[-\left(\tau + \frac{1}{n}\right)|z|\right]|f(z)|,$$

makes H *into a Fréchet space.*

Example 17.2.6.
The set $C^\infty(M, B)$ *of infinitely differentiable functions, where* M *is a compact manifold and* B *is a Banach space,*

$$f : M \to B$$

becomes a Fréchet space with semi-norms to be the suprema of the norms of all partial derivatives.

Example 17.2.7.
Let M be a compact C^∞-manifold and V a vector bundle over M. Take

$$C^\infty(M,V)$$

the space of smooth sections of V over X. Choose Riemannian metrics and connections, which are guaranteed to exist, on the bundles TX and V. If s is a section, denote its j-th covariant derivative by $D_j s$. Then

$$|s|_n = \sum_{j=0}^{n} \sup_{x\in M} |D_j s| \ ,$$

(where $|\cdot|$ is the norm induced by the Riemannian metric) is a family of semi-norms making

$$C^\infty(M,V)$$

into a Fréchet space.

It is now easy to get confused when considering derivatives of maps between Fréchet spaces. Let X_1, X_2 be Fréchet spaces, then the set of all continuous linear maps $L(X_1, X_2)$

$$X_1 \to X_2$$

is **not** a Fréchet space in any natural manner.

At this point the theory of Banach spaces and that of Fréchet spaces strongly deviate and need a different definition for continuous differentiability of functions defined on Fréchet spaces, the **Gâteaux derivative**

Definition 17.2.7 (Gâteaux Derivative)**.**
Let X_1, X_2 be Fréchet spaces,

$$U \subset X_1$$

open, and

$$P : U \to X_2$$

a function,

$$x \in U, \text{ and } V \in X_1.$$

We say that P is differentiable at x in the direction V if the following limit exists

$$D_V[P(x)] = \lim_{\Delta \to 0} \frac{\Big(P(x + V\Delta) - P(x)\Big)}{\Delta}. \tag{173}$$

Then P is called continuously differentiable in U if

$$D[P] : \; U \times X_1 \to X_2, \tag{174}$$

is continuous.

Since the product of Fréchet spaces is again a Fréchet space, we can then differentiate $D[P]$ and define the higher derivatives of P in this fashion. The derivation operator

$$P : C^\infty([0,1]) \to C^\infty([0,1])$$

defined by

$$P(x) = x$$

is itself infinitely differentiable. The first derivative reads

$$D_V[P(x)] = V' \tag{175}$$

for any two elements

$$x, V \in C^\infty([0,1]).$$

This is an important advantage of the Fréchet space $C^\infty([0,1])$ as compared to the Banach space $C^k([0,1])$ for finite k. If

$$P : U \to X_2$$

is a continuously differentiable function, then the differential equation

$$x'(t) = P(x(t)), \; x(0) = x_0 \in U, \tag{176}$$

need not have any solutions, and even if it does, the solutions need not be unique, in strong contrast to the situation in Banach spaces[1].

1 We emphasise that the inverse function theorem does not hold in Fréchet spaces. A partial substitute to it is the *Nash-Moser theorem*, which extends the notion of an inverse function

One can now define Fréchet manifolds as spaces that locally look like Fréchet spaces, and one can then extend the concept of Lie groups to these manifolds, leading to a Fréchet Lie group . Such a Lie group is a group G which is also a manifold, but now a Fréchet manifold (infinite dimensional) such that the map:

$$G \times G \to G, (g, h) \mapsto gh^{-1} \tag{177}$$

is continuous. This is useful because for a given (ordinary) compact C^∞-manifold M, the set of all C^∞ diffeomorphisms

$$f : M \to M$$

forms a *generalized Lie group* in this sense, and this Lie group captures the symmetries of M. Some of the relations between Lie algebras and Lie groups remain valid in this setting, which will be used when studying the group structure and Lie algebra structure of generalized loop space. For a discussion on this see for instance [124].

from Banach spaces to a class of Fréchet spaces. In contrast to the Banach space case, in which the invertibility of the derivative (where the derivative is interpreted as a linear operator) at a point is sufficient for a map to be locally invertible, the Nash-Moser theorem requires the derivative to be invertible in a vicinity of a point. The theorem is widely used to prove local uniqueness for non-linear partial differential equations in spaces of smooth functions.

18 THE GROUP OF GENERALIZED LOOPS AND ITS LIE ALGEBRA

18.1 INTRODUCTION

In the previous Chapters we have already demonstrated that $Shc(d,p)$ forms a group with the multiplication introduced in Definition (11.2.6) and that this algebra is isomorphic to the (Chen) integral algebra $\mathcal{A}_p$ generated by all functionals $X^{\omega_1 \cdots w_r}$, such that d-loops can be identified with elements of the algebra morphisms $\mathrm{Alg}(\mathcal{A}_p, k)$. From now on we set

$$k \equiv \mathbb{C}.$$

The algebra $\mathrm{Shc}(d,p)$ can be supplied with a topology turning it into a topological algebra[1], more specifically into a **locally multiplicative convex** (LMC) algebra. This topology is built from **semi-norms**, a construction that is due to Tavares [16]. The topologizing of this algebra is an extensive process that we will explain in detail, to not get lost along the way we also provided a diagrammatic overview which can be read as a map of the different steps in this construction.

Equipped with such a topology $\mathrm{Shc}(d,p)$ turns into a Fréchet space and combined with the fact that the generalized loops form a group this will also return a Fréchet Lie group with an associated Lie algebra.

The algebraic properties combined with the differential operations from Chapter 11 and the fact that limits are well-defined in this new space allows us to extend differential calculus on manifolds to the generalized manifolds discussed in the previous Chapter. Several differential operators will be introduced in Section 19.2, which generate variations of the loops.

1 By the Gel'fand duality between a topological space and its algebra of functions any unital C^*-algebra describes a topological space. Furthermore this duality can be expanded to give an algebraic description of the geometry of a space.

18.2 THE SHUFFLE ALGEBRA OVER $\Omega = \bigwedge^1 M$, AS A NLMC HOPF ALGEBRA

The main advantage of d-paths is that they can be considered as algebraic paths, in the sense that they have a rich algebraic structure that can be used to derive many interesting properties which will be discussed in more detail in this Section.

We start by restating the co-multiplication and co-unit of the shuffle algebra and a summary of their properties with respect to the shuffle algebra operations:

$$\begin{aligned} \Delta(\omega \cdots \omega_r) &= \sum_{i=0}^{r} \omega_1 \cdots \omega_i \otimes \omega_{i+1} \cdots \omega_r \\ \epsilon(\omega_1 ... \omega_r) &= 0, \text{ if } r \geq 1 \\ &= 1, \text{ if } r = 0 \end{aligned} \tag{178}$$

Properties of the co-multiplication and co-unit are the following:

$$\begin{aligned} (\Delta \otimes 1) \circ \Delta = (1 \otimes \Delta) \circ \Delta & \qquad \text{(Coassociative law)} \\ (1 \otimes \epsilon) \circ \Delta = (\epsilon \otimes 1) \circ \Delta = 1 & \qquad \text{(counitary property)} \\ \Delta(u \bullet v) = \Delta(u) \bullet \Delta(v) & \qquad (\Delta \text{ is an algebra morphism}) \\ \epsilon(u \bullet v) = \epsilon(u) \bullet \epsilon(v) & \qquad (\epsilon \text{ is an algebra morphism}) \end{aligned}$$

$$\forall u, v \in \mathrm{Sh}.$$

A complete **Hopf algebra structure** is given by a multiplication (the shuffle product), a unit, a co-multiplication, a co-unit and an antipode. The antipode was defined in 11.2.5 as k-linear map

$$J : \mathrm{Sh} \to \mathrm{Sh}$$

and is restated here for convenience

$$J(\omega_1 \cdots \omega_r) = (-1)^r \omega_r \cdots \omega_1, \tag{179}$$

with the following operational properties:

$$\begin{aligned} s \circ (J \otimes \mathbb{1}) \circ \Delta = s \circ (\mathbb{1} \otimes J) \circ \Delta &= \eta \circ \epsilon \\ J(u_1 \bullet u_2) &= J(v) \bullet J(u) \\ J(\mathbb{1}) = \mathbb{1}, \ J^2 &= \mathbb{1} \\ \epsilon \circ J &= \epsilon \\ \tau \circ (J \otimes J) \circ \Delta &= \Delta \circ J \end{aligned} \tag{180}$$

$\forall u_1, u_2 \in \mathrm{Sh}$, where

$$s : \mathrm{Sh} \otimes \mathrm{Sh} \to \mathrm{Sh}$$

denotes shuffle multiplication and

$$\eta : k \to \mathrm{Sh}$$

the unit map. The map

$$\tau : \mathrm{Sh} \otimes \mathrm{Sh} \to \mathrm{Sh} \otimes \mathrm{Sh}$$

is called the transposition map or flipping operation defined by

$$\tau(u_1 \otimes u_2) = u_2 \otimes u_1$$

The above properties can be used to prove the following Proposition

Proposition 18.2.1

Let

$$\omega_i \in \bigwedge^1 M$$

(respectively

$$\bigwedge^1 M \otimes GL(n, \mathbb{C}),$$

ϵ be a co-unit as in Definition 178 and $\bullet$ denote the shuffle multiplication. We then have

$$\begin{aligned} \sum_{i=0}^{r} (-1)^i \omega_i \cdots \omega_1 \bullet \omega_{i+1} \cdots \omega_r &= \sum_{i=0}^{r} (-1)^{r-i} \omega_1 \cdots \omega_i \bullet \omega_r \cdots \omega_{i+1} \\ &= \epsilon(\omega_1 \cdots \omega_r) \end{aligned}$$

These definitions and properties describe the Hopf algebra structure of $Sh(\Omega)$, which can now be transferred to $\mathcal{A}_p$, the algebra generated by the functionals $X^{\omega_1 \cdots \omega_r}$, see Eq. (92). This transfer of the Hopf algebra structure is accommodated by use of Proposition 12.3.1 which turns the surjective map

$$Sh(\Omega) \to \mathcal{A}_p,$$

defined by

$$\mathbb{1} \mapsto \mathbb{1}, \text{ and } \omega_1 \cdots \omega_r \mapsto X^{\omega_1 \cdots \omega_r}$$

into a **homomorphism of algebras**. We know that algebra morphisms preserve the algebraic structures by definition, such that indeed $\mathcal{A}_p$ inherits a Hopf algebra structure from $Sh(\Omega)$.

Proposition 12.3.2 and Theorem 8.2.2 imply that the kernel of this morphism contains the ideal $I(d, p)$, which was given by

$$\omega_1 \cdots \omega_{i-1}(f\omega_i)\omega_{i+1} \cdots \omega_r - f(p)\omega_1 \cdots \omega_r - ((\omega_1 \cdots \omega_{i-1}) \bullet df)\omega_i \cdots \omega_r, \tag{181}$$

or in a reduced symbolic notation

$$u_1(f\omega)u_2 - (u_1 \bullet df)\omega u_2 - f(p)u_1\omega u_2 \tag{182}$$

where

$$u_1, u_2 \in \mathrm{Sh},\ \omega \in \bigwedge^1 M,\ f \in C^\infty M.$$

It now follows that d-paths can be identified with elements of the set of algebra morphisms

$$\mathrm{Alg}(\mathcal{A}_p, \mathbb{C})$$

that is a d-path is an algebra morphism

$$\gamma \in \mathrm{Alg}(\mathcal{A}_p, \mathbb{C})$$

where

$$\gamma : \mathcal{A}_p \to \mathbb{C}$$

vanishes on the ideal $I(d, p)$ by definition. In the case of d-loops, however, we need to extend the ideal to include $dC^\infty(M)$. However, in the integral algebra this is included by definition since

$$\int_\gamma df = 0$$

for

$$\gamma \in \mathcal{LM}$$

and thus

$$dC^\infty(M) \in \ker\ (Sh(\Omega) \to \mathcal{A}_p).$$

As before we denote this ideal by J_p

$$J_p = I(d, p) + \langle dC \rangle, \tag{183}$$

where $I(d,p)$ is the shuffle algebra ideal associated to the pointed differentiation (d,p). We have already discussed that this new ideal induces the algebra isomorphism

$$Sh(\Omega)/J_p \simeq \mathcal{A}_p \tag{184}$$

for d-loops.

We mentioned already that the algebra $\mathcal{A}_p$ inherits a Hopf algebra structure through the algebra morphisms, where the unit and multiplication follow from Proposition 12.3.1 and the co-multiplication, co-unit and antipode follow from these operations on $Sh(\Omega)$ as

$$\begin{aligned}
\Delta(X^{\omega_1\cdots\omega_r}) &= \sum_{i=0}^{r} X^{\omega_1\cdots\omega_i} \otimes X^{\omega_{i+1}\cdots\omega_r} \\
\epsilon(X^{\omega_1\cdots\omega_r}) &= \begin{cases} 0 & r \geq 1 \\ \mathbb{1} & r = 0 \end{cases} \\
J(X^{\omega_1\cdots\omega_r}) &= (-1)^r X^{\omega_r\cdots\omega_1}.
\end{aligned} \tag{185}$$

This explains the Hopf algebra structure, but the integral algebra can be equipped with a much richer structure, namely that of a **Nuclear Locally Multiplicative-Convex** (NLMC) algebra. This structure is induced from a topology, which we will discuss below, giving it the structure of a Fréchet space (Definition 17.2.4). Fig. 35 gives a diagrammatic overview of how different topologies are constructed on the involved algebras. We start with deriving a topology on the algebra

$$\bigotimes^{r} \bigwedge^{1} M,$$

which can be used to obtain a topology on $Sh(\Omega)$ consistent with its linear structure. We write

$$\Omega = \bigwedge^{1} M$$

as before. The construction of the topology will give us more than just a topology, it will endow $Sh(\Omega)$ with the structure of a Nuclear Locally Multiplicative Convex (NLMC) TVS, or Fréchet space, that is also Hausdorff, Banach and Hopf. We stress the Banach property, which will restrict the Fréchet Lie group of generalized loops to a Banach Lie group which is much better behaved [124].

The Riemannian metric and connection on M are the starting point for our topologizing process. The connection allows us to define a covariant derivative

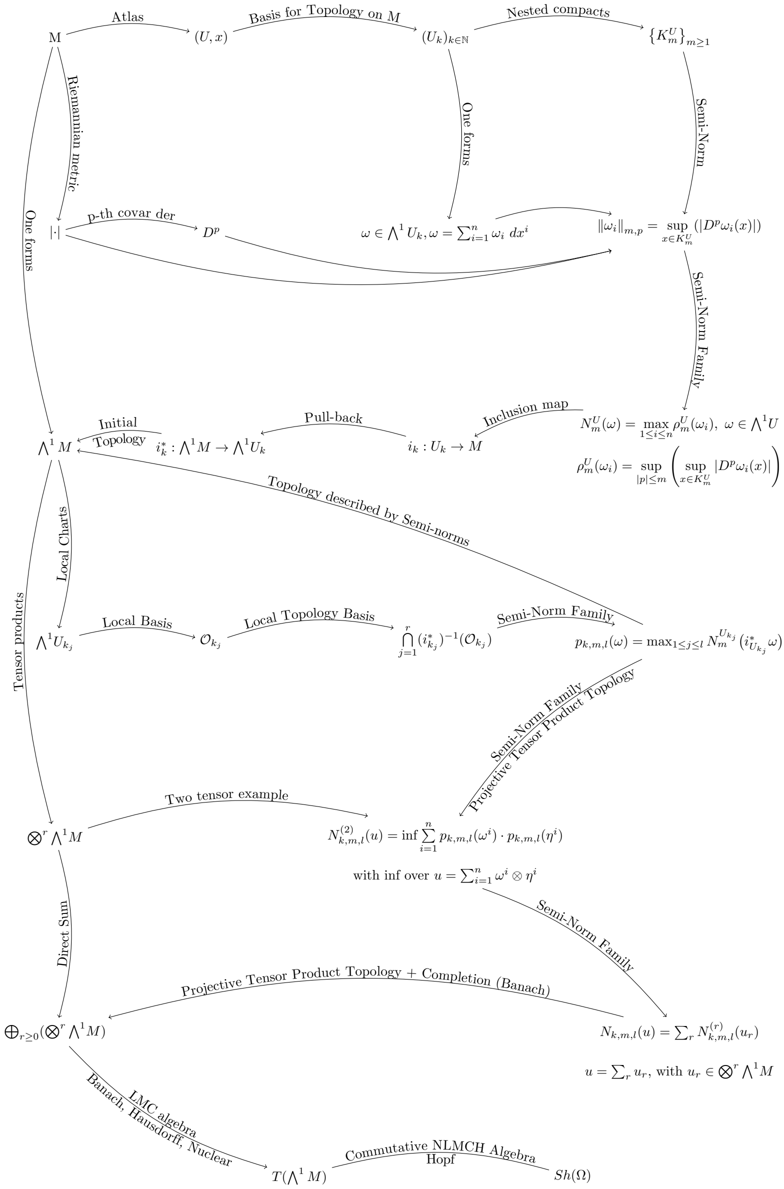


Figure 35: Topology on $Sh(\Omega)$.

D and the metric induces a norm $|\cdot|$. On the other hand, we know M as a manifold has a topology induced from its Riemannian metric. Combining this with the atlas of M we get local basis for this topology

$$(U_k)_{k\in\mathbb{N}}.$$

Using this local basis it is possible to construct a sequence of *nested compacts*

$$\left\{K_m^U\right\}_{m\geq 1}$$

in a local coordinate chart (U,x), such that

$$\bigcup_{m\geq 1} K_m^U = U.$$

We can then define a first family of semi-norms on (U,x) by using the Riemannian metric induced norm and covariant derivative

$$\|\omega_i\|_{m,p} = \sup_{x\in K_m^U}(|D^p\omega_i(x)|), \tag{186}$$

where

$$\omega_i \in C^\infty U$$

defined by the vectors

$$\omega = \sum_{i=1}^{n} \omega_i \; dx^i \in \bigwedge^1 M.$$

D^p denotes the p-th covariant derivative with respect to the (Riemannian) connection. A second family of semi-norms is now constructed from the first family of semi-norms

$$\|\omega\|_{m,p}$$

by

$$N_m^U(\omega) = \max_{1\leq i\leq n} \rho_m^U(\omega_i),\; \omega \in \bigwedge\nolimits^1 U,$$

where

$$\rho_m^U(\omega_i) = \sup_{|p|\leq m}\left(\sup_{x\in K_m^U}|D^p\omega_i(x)|\right).$$

As a result we obtain a family of semi-norms on the local coordinate chart (U,x). The next step is to extend this to the entire manifold M. This is realized by the inclusion map on the local basis for the (Riemannian) topology on M. To see this consider again the local basis

$$\{U_k\}_{k\in\mathbb{N}}$$

which can now also be interpreted as local charts. Define the map

$$i_k : U_k \to M$$

as the inclusion map, which embeds the local basis into M. The linear pull-back maps

$$i_k^* : \bigwedge^1 M \to \bigwedge^1 U_k$$

now define a map between the one-forms on this local basis and the same one-forms but now considered on M. Endowing

$$\bigwedge^1 M$$

with the initial topology defined by these maps, successfully equips it with a topology induced by semi-norms. Notice that by definition this topology is the weakest topology for which all the maps i_k^* are continuous, and a local topology basis consists of sets of the form

$$\bigcap_{j=1}^{r} (i_{k_j}^*)^{-1}(\mathcal{O}_{k_j}),$$

where the sets $\mathcal{O}_{k_j}$ run over a local basis of

$$\bigwedge^1 U_{k_j}.$$

In this way

$$\bigwedge^1 M$$

becomes a **Nuclear Locally-Convex topological vector space** (Fréchet space), whose topology can be described by the family of semi-norms

$$p_{k,m,l}(\omega) = \max_{1\leq j\leq l} N_m^{U_{k_j}}\left(i_{U_{k_j}}^* \omega\right)$$

From elementary calculus one learns that the definition of a differentiation depends on taking limits, which in its turn is defined by the convergence of a sequence. Given that we eventually will be interested in well-defined derivatives let us briefly consider convergence with the above family semi-norms. With such a family of semi-norms a sequence only converges if it converges with respect to all semi-norms in the family. In other words, a sequence of one-forms

$$(\omega_k)_{k\geq 1}, \text{ in } \bigwedge^1 M$$

converges to zero if and only if, in a neighborhood of each point of M, each derivative of each coefficient of ω_k converges uniformly to zero. The tensor powers

$$\bigotimes^r \bigwedge^1 M$$

now get a topology by the projective tensor product topology and becomes a Banach space when we also complete this space with respect to the semi-norms that describe this tensor topology. In other words, this topology is described by the semi-norms $N^{(r)}_{k,m,l}$ which are the tensor product of the above ones. To make this explicit, consider the example where $r = 2$.

Example 18.2.1. *Let*

$$u \in \bigwedge^1 M \otimes \bigwedge^1 M$$

for which we have

$$N^{(2)}_{k,m,l}(u) = \textit{inf} \sum_{i=1}^{n} p_{k,m,l}(\omega^i) \cdot p_{k,m,l}(\eta^i), \tag{187}$$

where inf is taken over all expressions of the element u in the form

$$u = \sum_{i=1}^{n} \omega^i \otimes \eta^i.$$

Extending now to elements in

$$\bigoplus_{r\geq 0} \Big(\bigotimes^r \bigwedge^1 M \Big),$$

which are finite sums

$$u = \sum_r u_r, \text{ with } u_r \in \bigotimes^r \bigwedge^1 M,$$

we get the semi-norms

$$N_{k,m,l}(u) = \sum_r N^{(r)}_{k,m,l}(u_r) \tag{188}$$

inducing a Nuclear Locally-Convex topology on

$$T\Big(\bigwedge^1 M\Big).$$

Due to the fact that all above topologies are consistent with the linear structures of the algebras, the shuffle product is a continuous map in this last topology. Moreover, the shuffle product is commutative so that $Sh(\Omega)$ inherits the structure of a commutative LMC algebra from

$$T\Big(\bigwedge^1 M\Big)$$

that also is Hopf, Banach and Hausdorff. We continue to write $Sh(\Omega)$ for this algebra.

We end this Section with the remark that the integral algebra $\mathcal{A}_p$ inherits the same structure through the isomorphism Eq. (184).

18.3 THE GROUP $\mathbf{L}\mathcal{M}_p$ OF LOOPS

The set of piecewise smooth loops with base point p, form a semi-group with respect to the product

$$\gamma_1 \cdot \gamma_2, \text{ for } \gamma_1, \gamma_2 \in \mathcal{L}\mathcal{M}_p$$

which also forms a loop space $\mathcal{L}\mathcal{M}_p$. Introducing the equivalence relation from Eq. (94), this semi-group can be turned into a group by defining a product $\mathcal{L}\mathcal{M}_p/\!\sim$ by

$$[\gamma_1] * [\gamma_2] = [\gamma_1 \cdot \gamma_2],$$

where the square brackets denote the equivalence classes under $\sim$. The inverses of elements

$$[\gamma] \in \mathcal{L}\mathcal{M}_p/\!\sim$$

are clearly given by

$$[\gamma]^{-1} = [\gamma^{-1}]$$

and the unit element is given by

$$\epsilon = [p],$$

the class of the constant loops equal to the point p. The group

$$(\mathcal{LM}_p/\sim, *),$$

now a **group of loops** on the manifold M based at p, is symbolically represented by

$$\mathbf{L}M_p.$$

18.4 THE GROUP $\mathbf{L}\widetilde{M}_p$ OF GENERALIZED LOOPS

In this Section we will use the algebraic paths and loops introduced by Chen to generalize the **group of loops** $\mathbf{L}M_p$ to the **Group of Generalized Loops**, equivalent to the space of d-loops as the algebra morphisms from $Sh(\Omega)$ to $\mathbb{C}$ that vanish on the ideal J_p. In previous sections we showed that the algebra $Shc(d,p)$ is equivalent to the algebra $\mathcal{A}_p$, such that to introduce the space of Generalized Loops we start by extending our study of this algebra. The main concept that we need in this extension is that of a **spectrum on a commutative Banach algebra**[2] (Definition 9.2.18). Considering the algebra[3] $\mathcal{A}_p$, writing $\triangle_p$ for the spectrum, we notice that

$$\phi \in \triangle_p$$

is also an element of the dual space $\mathcal{A}_p^*$ (the space of complex valued linear functionals) of $\mathcal{A}_p$. In Section 18.2 we constructed a topology on $Sh(\Omega)$, inducing a topology on $\mathcal{A}_p$. We would now like to have a topology on $\mathcal{A}_p^*$. Consider the dual space $\mathcal{A}_p^{**}$ of the dual space $\mathcal{A}_p^*$, then the following map

$$x \mapsto \Phi_x : \Phi_x(\phi) = \phi(x)$$

allows us to embed the original space $\mathcal{A}_p$ in $\mathcal{A}_p^{**}$. The maps Φ_x can now be used to define a coarsest topology on $\mathcal{A}_p^*$ such that all the Φ_x are continuous maps

$$\mathcal{A}_p^* \mapsto \mathbb{C}.$$

This topology is referred to as the **weak-$*$ topology**, in which the characters are now continuous by definition.

As mentioned before, we know from Section 18.2 that $\mathcal{A}_p$ inherits a semi-norm structure from $Sh(\Omega)$ such that by the Banach-Alaoglu theorem $\mathcal{A}_p$ is reflexive

$$\mathcal{A}_p^{**} \equiv \mathcal{A}_p.$$

2 Where the commutative refers to the shuffle product.
3 For the moment considering the one-forms to be complex valued.

It follows that every bounded sequence has a weakly converging subsequence, similar to the case in regular calculus. We remark that a sequence

$$\phi_n \in \mathcal{A}_p^*$$

converges in the weak-$*$ topology iff

$$\phi_n(x) \to \phi(x),\ \forall x \in \mathcal{A}_p.$$

This convergence is sometimes referred to as the point-wise convergence of linear functionals. The Hausdorff property of $\mathcal{A}_p$ can be understood from the separation property (13.3.1) of the functionals $X^{\omega_1 \cdots \omega_r}$, but is a direct consequence of the Gel'fand-Mazur Theorem (Theorem 9.2.2). It follows that the d-loops

$$\tilde{\gamma} : Sh(\Omega) \to \mathbb{C}$$

can be identified with elements of $\mathcal{A}_p^*$.

In the above discussion we restricted ourselves to complex valued one-forms, but in a gauge theory setting, and using the principal fibre bundle formalism, we will need to deal with Lie algebra valued one-forms. Choosing a matrix representation for the Lie algebra, the algebra elements form a sub-algebra of $GL(n, \mathbb{C})$. Therefore let us consider $GL(n, \mathbb{C})$ valued one-forms. The fact that the $Sh(\Omega)$ algebra is of the nuclear or trace class (NLMC), the traces of the matrices do not spoil the algebraic or topological structures such that convergence is still well-defined. Thus by adding the trace operator to the integrals in the functionals of $\mathcal{A}_p$ in the case of matrix valued one-forms we get again a set of continuous characters (complex valued!).

Notice that the nuclear property also assures that their exists a well-defined trace operator on the linear bounded operators used to define the Fréchet derivative in Definition 17.2.2. Moreover, it also assures that this trace is finite.

We can now safely identify the d-loops with the spectrum of $\mathcal{A}_p$ with the remark that if the

$$\omega \in \bigwedge^1 M$$

are $GL(n, \mathbb{C})$ valued, we need to take the trace to reduce the $GL(n, \mathbb{C}) - valued$ matrix to an element of $\mathbb{C}$. Let us now extend the previously introduced equivalence relation 13.3.1 on d-loops to

$$W_{\gamma_1} = \mathrm{Tr}\ U_{\gamma_1} = \mathrm{Tr}\ U_{\gamma_2} = \mathcal{W}_{\gamma_2}, \tag{189}$$

for two d-loops

$$\gamma_1, \gamma_2 \in M$$

with U and $\mathcal{W}$ defined in Eqs. (163) and (167), respectively.

By the continuity of the trace these form a subset of the d-loops, and also of the generalized loops (see below), that are still separable by Theorem 13.3.2. Weak-$*$ convergence is also still applicable due to the fact that convergence requires convergence for all elements in $\mathcal{A}_p$. We are now ready to define the Generalized Loops.

Definition 18.4.1 (Generalized Loop)**.**
*A **Generalized Loop** based at*

$$p \in M$$

is a character of the algebra $\mathcal{A}_p$ or, equivalently, a continuous complex algebra homomorphism

$$\tilde{\gamma} : Sh(\Omega) \to \mathbb{C}$$

that vanishes on the ideal J_p.

Important, for later purposes, is that convergence in the space of Generalized Loops is defined by the weak-$*$ topology convergence, as we discussed above. Having defined Generalized Loops we can ask how these are related to the (naive) loops from the previous section. The answer is provided by an embedding, called the **Dirac map**

$$\delta : \mathbf{L}M_p \to \triangle_p, \ [\gamma] \mapsto \delta_{[\gamma]} \tag{190}$$

defined by

$$\delta_{[\gamma]}(X^{\omega_1 \cdots \omega_r}) = X^{\omega_1 \cdots \omega_r}([\gamma]), \tag{191}$$

where

$$[\gamma] \in \mathbf{L}M_p.$$

This embedding is injective due to Theorem 13.3.1. Identifying $\mathbf{L}M_p$ with its image, under δ, in $\triangle_p$, it also inherits an induced topology.

Naturally we will want to be able to compose generalized loops like we combine normal loops in a manifold. The composition is realized by a multiplication, introduced as a convolution multiplication

$$\tilde{\alpha} \star \tilde{\beta}$$

of the two elements

$$\tilde{\alpha}, \tilde{\beta} \in \triangle_p$$

defined by

$$\tilde{\alpha} \star \tilde{\beta} = (\tilde{\alpha} \otimes \tilde{\beta}) \circ \Delta,$$

which gives $\triangle_p$ a group structure and where we used

$$k \otimes k \simeq k, \ \mathbb{C} \otimes \mathbb{C} \simeq \mathbb{C}.$$

The inverse in the group of

$$\tilde{\alpha} \in \triangle_p,$$

is given by

$$\tilde{\alpha} \circ J$$

that is

$$\tilde{\alpha}^{-1}(\omega_1 \cdots \omega_r) = (-1)^r \tilde{\alpha}(\omega_r \cdots \omega_1),$$

where one takes ϵ (co-unit) as the unit element. This defines **the Group of Generalized Loops**. Writing this out explicitly with the definition of Δ from Eq. (185) we have

$$\tilde{\alpha} \star \tilde{\beta}(X^{\omega_1 \cdots \omega_r}) = \sum_{i=0}^{r} \tilde{\alpha}(X^{\omega_1 \cdots \omega_i}) \cdot \tilde{\beta}(X^{\omega_{i+1} \cdots \omega_r}). \tag{192}$$

For loops in $\mathbf{L}M_p$, the Dirac map allows us to write Eq. (192) explicitly as

$$\begin{aligned}
\sum_{i=0}^{r} \tilde{\alpha}(X^{\omega_1 \cdots \omega_i}) \cdot \tilde{\beta}(X^{\omega_{i+1} \cdots \omega_r}) &= \sum_{i=0}^{r} \alpha(X^{\omega_1 \cdots \omega_i}) \cdot \beta(X^{\omega_{i+1} \cdots \omega_r}) \\
&= \sum_{i=0}^{r} (X^{\omega_1 \cdots \omega_i})(\alpha) \cdot (X^{\omega_{i+1} \cdots \omega_r})(\beta) \\
&= \sum_{i=0}^{r} \int_{\alpha} \omega_1 \cdots \omega_i \cdot \int_{\beta} \omega_{i+1} \cdots \omega_r \\
&= \int_{\alpha \cdot \beta} \omega_1 \cdots \omega_r,
\end{aligned} \tag{193}$$

which also shows that the convolution product defined on GLS makes sense as a composition of d-loops. Notice that this product is in general not commutative.

In Section 18.2 we discussed the topological structure on $Sh(\Omega)$, which induced a topology on the algebra $\mathcal{A}_p$ that in its turn by the weak-$*$ topology induced one on the spectrum. Since elements of the spectrum for loops vanish by definition on the ideal J_p we identified them with generalized loops, such that the group of generalized loops can also be considered as a **topological group**.

Definition 18.4.2 (Generalized Loop Space as Topological Group).

$$\tilde{\alpha} \star \tilde{\beta},\ \tilde{\alpha}^{-1} \text{ and } \epsilon$$

are generalized loops based at p, i.e., they are continuous characters on the algebra $\mathcal{A}_p$, or equivalently, continuous characters on the algebra $Sh(\Omega)$ that vanish on the ideal J_p. Moreover, $(\triangle_p, \star)$ has the structure of topological (Hausdorff and completely regular or Tychonoff) group.

Definition 18.4.3 (Group of Generalized Loops).
We call the above mentioned topological group

$$(\boldsymbol{\Delta}_p, \star),$$

the group of generalized loops of M based at

$$p \in M,$$

and we denote it by

$$\widetilde{\mathbf{L}\mathcal{M}_p}.$$

From the fact that the Dirac map preserves group operations, we conclude that $\mathbf{L}M_p$ is a topological subgroup of $\widetilde{\mathbf{L}M_p}$. The above discussion clearly shows that the **naive** loops are a subset of generalized loops, but they are not the same. To emphasize the difference we consider the following example.

Example 18.4.1.
Let

$$M = S^1.$$

Then it is easy to see that

$$\mathbf{L}S^1_p = \mathbb{Z}.$$

However,

$$\widetilde{\mathbf{L}S}^1_p = \mathbb{R}.$$

In fact, since

$$H^1(S^1, \mathbb{R}) = \mathbb{R},$$

any one-form

$$\omega \in S^1$$

is equal to a constant multiple of

$$\omega_0 \equiv d\theta$$

(the usual volume form in S^1), modulo an exact form

$$\omega = c\omega_0 + df, c \in \mathbb{R}.$$

$$\bigwedge^1 S^1 = \mathbb{R}\omega_0 \oplus dC^\infty(S^1). \tag{194}$$

From this fact, we can prove that $\mathcal{A}_p$ is isomorphic, as an Hopf algebra, to $\mathbb{R}[t]$, the polynomial ring in one variable

$$t \leftrightarrow X^{\omega_0}$$

(see [107]). The Hopf operations on $\mathbb{R}[t]$ are:

$$\Delta(t) = \mathbb{1} \otimes t + t \otimes \mathbb{1},\ J(t) = -t,\ \epsilon(t) = 0$$

Then

$$\widetilde{\mathbf{L}S}^1_p = \mathbb{R}$$

follows from Example 4.1 in [125]

In the above we have only discussed the generalization of loop, but a similar generalization applies to paths.

Remark 18.4.1 (Generalized Paths).
Consider the path space $\mathcal{PM}_p$ of paths based at

$$p \in M,$$

and the algebra $\mathcal{B}_p$ generated by all the functions $X^{\omega_1 \cdots \omega_r}$, considered now as functions on $\mathcal{PM}_p$. Similarly to the previous case, there exists an algebra isomorphism

$$Sh(\Omega)/I_p \simeq \mathcal{B}_p,$$

*which allows to consider $\mathcal{B}_p$ as an LMC algebra and define generalized paths, based at p, as continuous characters on $Sh(\Omega)$ that vanish on I_p. These generalized paths, however, do not form a group but only a **semi-group**.*

18.5 GENERALIZED LOOPS AND AMBROSE-SINGER THEOREM

As we have discussed in Section 16 the Ambrose-Singer theorem indicates that there exists a formulation of gauge theory that can be expressed in gauge-invariant variables and that a natural choice for these variables are Holonomies or Wilson Loops (without the trace!). Unfortunately the information in all Wilson loops is abundant, making a naive loop space over-complete. It follows that the variables will need to be constrained if one wants to be able to reconstruct the gauge potentials A_μ from the information in these gauge invariant variables. In this Section we will discuss these constraints based on Giles paper from 1981 [126].

The first constraint is an algebraic one and originates from the composition of parallel transports (Wilson lines/loops)

$$U_\gamma[[x, y] \circ [y, z]] = U_\gamma[x, y]U_\gamma[y, z]. \tag{195}$$

The second constraint comes from smoothness requirements. Put differently, the loops used in the functions of the gauge-invariant formulation of gauge theory need to be sufficiently smooth such that the reconstructed gauge potentials are finite. Expressed mathematically we require the paths and loops to be representable by a piecewise differential function $z\mu(t)$ such that for

$$U_\gamma[x, y]$$

$$z^\mu(0) = x \quad \text{and} \quad z^\mu(1) = y.$$

Furthermore if we have a piecewise differential function v^μ with

$$v^\mu(0) = v^\mu(1) = 0$$

then the homotopic path

$$z(\epsilon, t) = z^\mu(t) + \epsilon v^\mu(t)$$

is such that

$$U_\gamma[x,y](\epsilon) = U_\gamma[x,y] + \mathcal{O}(\epsilon^2). \tag{196}$$

Also if the path between the points x and y is straight we have from the expression for the gauge link that

$$U_\gamma[x,y](\epsilon) = 1 + (x^\mu - y^\mu)A_\mu + \mathcal{O}\Big((x^\mu - y^\mu)^2\Big). \tag{197}$$

For closed loops this becomes

$$U_\gamma[x,x](\epsilon) = U_\gamma[x,x] + \epsilon^\mu[U_\gamma[x,x], A_\mu] + \mathcal{O}(\epsilon^2), \tag{198}$$

which is now a constraint on translated loops. The reconstruction of the gauge potential can then be done from infinitesimal straight paths, that approximate the original path and are equal to it in taking the length of the infinitesimal paths to zero

$$\begin{aligned} U_\gamma[x,y] &= \lim_{M\to\infty} U_\gamma[x, x_{M-1}] \cdots U_\gamma[x_1, y] \\ &= \lim_{M\to\infty} \Big[1 + \Big(x^\mu - x^\mu_{M-1}\Big) A_\mu(x)\Big] \cdots [1 + (x^\mu_1 - y^\mu)A_\mu(x_1)]. \end{aligned}$$

Notice the structure of this last equation when comparing it with the expressions for the product integrals in Chapter 15. We also point out that for loops the gauge potential can only be determined up to gauge transformations. To see this let us consider the Wilson Loop

$$U_\gamma[x],$$

from which we first determine the open path phase-factors or Wilson lines. Assign to each open path π_{xy}

$$U_\gamma[x,y]$$

from x a phase-factor

$$\Phi(y) \equiv U(\pi_{yx}).$$

Consider a random path $P_{y,z}$ such that

$$\pi_{yx}^{-1} \circ P_{y,z} \circ \pi_{zx}$$

is a loop at x with corresponding phase-factor

$$U(P_{y,z}) = \Phi(y) U_\gamma[\pi_{yx}^{-1} \circ P_{y,z} \circ \pi_{zx}] \Phi^{-1}(z).$$

This phase-factor clearly satisfies the constraint shown in Eq. (195). The reason that the open phases are not **uniquely** determined follows from the freedom of choice for the paths π_{yx} and π_{zx} to form the loop

$$\pi_{yx}^{-1} \circ P_{y,z} \circ \pi_{zx}.$$

Choosing another path π'_{yx} for π_{yx}, it is easy to show that the with this new path associated open phase-factor for the path $P_{y,z}$ is related to the original one by a factor

$$\phi(y) = \Phi'(y) U_\gamma[\pi'^{-1}_{yx} \circ \pi_{yx}] \Phi^{-1}(y),$$

a gauge transformation. Due to the equivalence of fundamental groups at different points x, y in topological spaces we have that the closed-loop phases at each point describe the gauge-invariant content of field configurations, but the relation between the descriptions at different points is non-canonical. The equivalence of the holonomies at different points is non-canonical because for phase-factors each different choice for a path between the points x and y leads to a different equivalence, although the holonomy groups are related by an isomorphism.

Including now the traces in the Wilson Loops we change to the **Wilson Loop Variables**

$$W(\gamma_x) = \frac{1}{N} \text{Tr } U_\gamma[x]$$

which is now gauge invariant. Naturally $W(\gamma_x)$ is determined by $U_\gamma[x]$ in a unique way. The inverse problem is quite more involved, where again we have constraints but now on $W(\gamma_x)$. We need algebraic constraints that are the counterparts of the constraint in Eq. (195), where the first set of these is generated by the cyclicity of the Trace

$$W(\gamma_1 \circ \gamma_2) = W(\gamma_2 \circ \gamma_1).$$

A second set of constraints comes from the need for the complex values to be traces of some $SU(N)$ matrix. These constraints are highly non-linear and a description of them can be found in [127], where it is also demonstrated that in the **classical** case the equivalence relation from Theorem 13.3.2 is equivalent to the above set of constraints. The introduction of this equivalence relation has thus reduced the infinite dimensional holonomy algebra to a finite dimensional matrix algebra.

Formally this construction can be extended to the quantum field theory case by use of

$$W_n(C_1, \cdots, C_n) = \int_\mu dW \epsilon^{-S(W)} W(C_1) \cdots W(C_n),$$

where $S(W)$ is an (unknown) action on loop space and μ is over the characters of the representations of the loop group. This is a formal expression since no measure has been identified on μ, which has much to do with the complications of quantization of this loop space. Therefor this suggests that the Mandelstam constraints are sufficient to determine the (quantized) gauge potentials up to gauge transformation, but it is not a definite proof.

Notice also that, assuming that the above can be proved, we have exchanged gauge invariance for path dependence. This dependence has its own problems, but is manageable in some cases. For instance considering paths on the light cone in a certain light cone direction has only one option for a path such that the path dependence is not a real obstacle. In the general case however non-trivial phenomena might arise, an example of such an effect is the Aharanov-Bohm effect.

18.6 THE LIE ALGEBRA OF THE GROUP $\mathbf{L}\widetilde{\mathcal{M}}_p$

In Section 18.4 we explained that the space of generalized loops forms a topological group. Now we investigate the associated Lie algebra.

As we know from Section 14.1, Lie algebras have a close connection with right/left invariant vector fields. With this in mind we repeat here the definition of a **left invariant derivation** (respectively, **right invariant derivation**) on $\mathcal{A}_p$.

Definition 18.6.1 (Left Invariant Derivation).
A k-linear map

$$D : \mathcal{A}_p \to \mathcal{A}_p$$

is called a **left invariant derivation** *(respectively,* **right invariant derivation***) on $\mathcal{A}_p$ if D satisfies the following two conditions*

$$D(X^{\mathbf{u_1}} X^{\mathbf{u_2}}) = X^{\mathbf{u_1}} D(X^{\mathbf{u_2}}) + D(X^{\mathbf{u_1}}) X^{\mathbf{u_2}} \tag{199}$$

$$\Delta \circ D = (1 \otimes D) \circ \Delta. \tag{200}$$

(respectively, $\Delta \circ D = (D \otimes 1) \circ \Delta$), for all

$$\mathbf{u_1}, \mathbf{u_2} \in \mathrm{Sh}.$$

Using the topological group property of $\widetilde{\mathbf{L}\mathcal{M}_p}$ and the above invariant derivations we have the following definition from the general theory of affine k-groups

Definition 18.6.2 (Lie Algebra of $\widetilde{\mathbf{L}\mathcal{M}_p}$).
We define the Lie algebra of the group $\widetilde{\mathbf{L}\mathcal{M}_p}$ as the k-linear space $\widetilde{l\mathcal{M}_p}$ of all continuous left invariant derivations[a] *on $\mathcal{A}_p$. With the Lie bracket defined as*

$$[D_1, D_2] = D_1 D_2 - D_2 D_1. \tag{201}$$

a These are actually the left invariant vector fields on $\widetilde{\mathbf{L}\mathcal{M}_p}$.

Similarly to the situation with regular Lie groups it is possible to demonstrate that this Lie algebra is isomorphic the tangent space at the unit element $\widetilde{\mathbf{L}\mathcal{M}_p}$, i.e. with

$$T_\epsilon \widetilde{\mathbf{L}\mathcal{M}_p} \text{ at } \epsilon.$$

To justify this isomorphism and make the derivations more explicit, let us consider the convolution product

$$f_1 \star f_2$$

of two elements

$$f_1, f_2 \in \mathcal{A}_p^*,$$

the topological (weak) dual of $\mathcal{A}_p$

$$f_1 \star f_2(X^{\omega_1\cdots\omega_r}) = (f_1 \otimes f_2) \circ \Delta(X^{\omega_1\cdots\omega_r})$$
$$= \sum_{i=0}^{r} f_1(X^{\omega_1\cdots\omega_i}) \cdot f_2(X^{\omega_{i+1}\cdots\omega_r}). \tag{202}$$

The convolution product can be used to define left- and right-invariant endomorphisms on $\mathcal{A}_p^*$.

Lemma 18.6.1

$(\mathcal{A}_p^*, \star)$ is a topological k-algebra, isomorphic (resp., anti-isomorphic) to the topological algebra $\mathrm{End}^{LL}(\mathcal{A}_p)$ (respectively, $\mathrm{End}^{RL}(\mathcal{A}_p)$) of all left (or right) invariant k-linear endomorphisms of $\mathcal{A}_p$ (i.e., k-linear morphisms

$$\sigma : \mathcal{A}_p \to \mathcal{A}_p$$

that satisfy the left (right) invariance condition):

$$\Delta \circ \sigma = (1 \otimes \sigma) \circ \Delta \tag{203}$$

and, respectively,

$$\Delta \circ \sigma = (\sigma \otimes 1) \circ \Delta,$$

endowed with the topology of point wise convergence (weak-$*$). Moreover, each element of

$$\mathrm{End}^{LL}(\mathcal{A}_p)$$

commutes with each element of

$$\mathrm{End}^{RL}(\mathcal{A}_p).$$

Some instructive examples are given below.

Example 18.6.1.
Let

$$\alpha \in \mathbf{L}M_p$$

and

$$\delta_\alpha \in \mathcal{A}_p^*$$

be the Dirac map as defined before. Then Ψ_{δ_α} is the automorphism

$$X^u \to \alpha \cdot X^u$$

corresponding to the action of α, on $\mathbf{L}\mathcal{M}_p$ from the right[a]*. In fact, the right action of $\mathbf{L}\mathcal{M}_p$ on itself, through right translations*

$$r_{\alpha_1} : \alpha_2 \mapsto \alpha_2 \cdot \alpha_1,$$

induces a left action of $\mathbf{L}M_p$ on $\mathcal{A}_p$ by

$$(\alpha_1 \cdot X^u)(\alpha_2) \equiv X^u(\alpha_2 \cdot \alpha_1) \tag{204}$$

By the identification

$$\alpha_1 \to \delta_{\alpha_1},$$

we can write the RHS of the above equation in the form:

$$\begin{aligned} X^u(\alpha_2 \cdot \alpha_1) &= \delta_{\alpha_2 \cdot \alpha_1}(X^{\mathbf{u}}) = \delta_{\alpha_2} \star \delta_{\alpha_1}(X^{\mathbf{u}}) \\ &= \delta_{\alpha_2}((\mathbb{1} \otimes \delta_{\alpha_1})\Delta X^u) = \delta_{\alpha_2}\Big(\Psi_{\delta_{\alpha_1}}(X^u)\Big) \end{aligned} \tag{205}$$

while the LHS is simply

$$\delta_{\alpha_2}(\alpha_1 \cdot X^{\mathbf{u}}),$$

which allows the above mentioned identification

$$\Psi_{\delta_{\alpha_1}} \simeq \alpha_1 \cdot X^{\mathbf{u}}.$$

In the same way we can prove that $\Lambda_{\delta_{\alpha_1}}$ is the automorphism

$$X^u \to X^u \cdot \alpha_1$$

corresponding to the action of α_1, on $\mathbf{L}M_p$ from the left.

a $X^u(\beta) \to \alpha \cdot X^u(\beta) = \int_{\beta\alpha} X^u$ note the order change of the paths which makes it into a right action on $\mathbf{L}\mathcal{M}_p$ although it is written as a product from the left.

Taking now the σ, defined in the Lemma 18.6.1, to be a left invariant derivation

$$\sigma = D,$$

we can write

$$\Phi(D) = f_D = \epsilon \circ D \in \mathcal{A}_p^*$$

with:

$$\begin{aligned} f_D(X^{u_1}X^{u_2}) &= \epsilon D(X^{u_1}X^{u_2}) \\ &= \epsilon(X^{u_1}DX^{u_2} + DX^{u_1}X^{u_2}) \\ &= \epsilon(X^{u_1})f_D(X^{u_2}) + f_D(X^{u_1})\epsilon(X^{u_2}), \end{aligned} \tag{206}$$

demonstrating that the Lie algebra $\widetilde{l\mathcal{M}_p}$ is isomorphic, as k-linear space, to the subspace of $\mathcal{A}_p^*$ consisting of the pointed derivations (Definition 10.3.3) at ϵ

$$\widetilde{l\mathcal{M}_p} \cong \{\delta \in \mathcal{A}_p^* : \ \delta(X^{u_1}X^{u_2}) = \epsilon(X^{u_1})\delta(X^{u_2}) + \delta(X^{u_1})\epsilon(X^{u_2})\}.$$

The **tangent space** at ϵ

$$T_\epsilon\widetilde{\mathbf{L}\mathcal{M}_p},$$

is considered to be this k-linear space of pointed derivations at ϵ, just as desired. To motivate why we call this space the tangent space, consider $\tilde{\alpha}_\Delta$, a curve of generalized loops such that

$$\begin{aligned} \tilde{\alpha}_0 &= \epsilon \\ \lim_{\Delta\to 0} \tilde{\alpha}_\Delta &= \epsilon \\ \lim_{\Delta\to 0} \frac{\tilde{\alpha}_\Delta - \epsilon}{\Delta} &= \delta \in \mathcal{A}_p^* \end{aligned}$$

where the limits are taken in the weak (topology) sense

$$\lim_{\Delta\to 0} \tilde{\alpha}_\Delta(X^u) = \epsilon(X^u), \ \forall u \in Sh(\Omega)$$

in the weak-$*$ topology. As such δ is a tangent vector to this curve of loops. Applying this δ to X^uX^v returns:

$$\begin{aligned} \delta(X^{u_1}X^{u_2}) &= \lim_{\Delta\to 0} \frac{\tilde{\alpha}_\Delta(X^{u_1}X^{u_2}) - \epsilon(X^{u_1}X^{u_2})}{\Delta} \\ &= \lim_{\Delta\to 0} \left(\tilde{\alpha}_\Delta(X^{u_1})\frac{\tilde{\alpha}_\Delta(X^{u_2}) - \epsilon(X^{u_2})}{\Delta} + \frac{\tilde{\alpha}_\Delta(X^{u_1}) - \epsilon(X^{u})}{\Delta}\epsilon(X^v)\right) \\ &= \epsilon(X^{u_1})\delta(X^{u_2}) + \delta(X^{u_1})\epsilon(X^{u_2}), \end{aligned}$$

where we used the property that the Chen integrals preserve multiplicity (Proposition 12.3.1)

$$\tilde{\alpha}_t(X^{\mathbf{u}}X^{\mathbf{v}}) = (X^{\mathbf{u}}X^{\mathbf{v}})(\tilde{\alpha}_t) = X^{\mathbf{u}}(\tilde{\alpha}_t)X^{\mathbf{v}}(\tilde{\alpha}_t) = (\tilde{\alpha}_t)(X^{\mathbf{u}})(\tilde{\alpha}_t)(X^{\mathbf{v}}),$$

showing that δ is a pointed derivation at ϵ. We thus have a k-linear isomorphism:

$$T_\epsilon \widetilde{\mathbf{L}\mathcal{M}_p} \cong \widetilde{l\mathcal{M}_p},$$

given by

$$\delta \to D_\delta = (1 \otimes \delta) \circ \Delta.$$

The Lie bracket of

$$T_\epsilon \widetilde{\mathbf{L}\mathcal{M}_p}$$

for the pointed differentiations is defined by

$$[\delta, \eta] \equiv \epsilon \circ [D_\delta, D_\eta] = \delta \star \eta - \eta \star \delta. \tag{207}$$

Notice that for any pointed derivation δ, at ϵ

$$\delta(X^{\omega_1 \cdots \omega_r} X^{\omega_{r+1} \cdots \omega_{r+s}}) = 0, \; \forall r \geq 1, \forall s \geq 1,$$

which stems from the product properties of the X^u and the definition of $\epsilon(X^u)$ that show up when taking the δ of a product (see Proposition 12.3.1, Eqs. (185) and (207)). Combining these properties with the above we can conclude that

$$\delta^n(X^{\omega_1 \cdots \omega_r}) = 0, \; \forall n > r \geq 0, \tag{208}$$

where

$$\delta^n \equiv \delta^{n-1} \star \delta, \; \forall n \geq 1.$$

As with usual Lie theory we can define an exponential map

$$\exp[\delta]$$

for each tangent vector

$$\delta \in T_\epsilon \widetilde{\mathbf{L}\mathcal{M}_p} \cong \widetilde{l\mathcal{M}_p}$$

by

$$\exp[\delta] \equiv \epsilon + \sum\nolimits_{n \geq 1} \frac{\delta^n}{n!}$$

such that the map

$$X^{\omega_1 \cdots \omega_r} \to \exp[\delta(X^{\omega_1 \cdots \omega_r})]$$

is given by

$$\left(\epsilon + \sum\nolimits_{n \geq 1} \frac{\delta^n}{n!}\right)(X^{\omega_1 \cdots \omega_r}),$$

which is of course only valid under the assumption that the series converges. This is the case since by Eq. (208) this series is finite, such that $\exp \delta$ is well defined. Again, similar to the case in regular Lie theory, $\exp[\delta]$ represents a generalized loop. Moving on to the opposite case where we now have a given generalized loop

$$\tilde{\alpha} \in \widetilde{\mathbf{L}\mathcal{M}_p},$$

an Lie algebra element can be associated through the definition

$$\log \tilde{\alpha} \equiv \sum\nolimits_{n \geq 1} \frac{(-1)^{n-1}}{n} (\tilde{\alpha} - \epsilon)^n, \tag{209}$$

where

$$(\tilde{\alpha} - \epsilon)^n \equiv (\tilde{\alpha} - \epsilon)^{n-1} \star (\tilde{\alpha} - \epsilon), \quad \forall n \geq 1.$$

Again from the properties of the Hopf algebra operations we have that

$$(\tilde{\alpha} - \epsilon)^n (X^{\omega_1 \cdots \omega_r}) = 0, \ \forall n > r \geq 0,$$

such that $\log \tilde{\alpha}$ is well-defined and is an element of

$$T_\epsilon \widetilde{\mathbf{L}\mathcal{M}_p} \cong l\widetilde{\mathcal{M}_p}.$$

Continuing, similar to the reasoning in real calculus, expansion in formal power series allows to define

$$\exp[k \log \tilde{\alpha}] = \tilde{\alpha}^k, \ \forall k \in Z$$
$$\log[\exp \delta] = \delta,$$

which can be extended to define for each

$$\Delta \in k$$

$$\tilde{\alpha}^\Delta \equiv \exp[\Delta \log \tilde{\alpha}]$$

It is now not so hard to show that

$$\Delta \mapsto \tilde{\alpha}^\Delta$$

is a one-parameter subgroup of $\widetilde{\mathbf{L}\mathcal{M}_p}$, generated by $\log \tilde{\alpha}$, i.e.,

$$\tilde{\alpha}^0 = \epsilon$$
$$\tilde{\alpha}^\Delta \star \tilde{\alpha}^{\Delta'} = \tilde{\alpha}^{\Delta + \Delta'}$$
$$\lim_{\Delta \to 0} \frac{\tilde{\alpha}^\Delta - \epsilon}{\Delta} = \log \tilde{\alpha} = \delta,$$

such that

$$\tilde{\alpha}^{\Delta} = \exp[\delta],$$

a generalized loop and where in the last line the limit is taken in the weak (topology) sense.

Now that we have introduced the left- and right-invariant derivations, discussed their relation with the pointed derivations defined on the shuffle algebra and having defined a Lie algebra we can move on in the next Chapter to differential calculus on GLS.

18.7 SUMMARY

In this chapter we have introduced the concept of a generalized loop, demonstrated that it is a topological group and introduced its associated Lie algebra. We also discussed the equivalence relation that can be introduced on this space by the Wilson Loop Variables, still accommodating the separation of points, but reducing the infinite dimensional Lie algebra to a finite dimensional one. The relation between the Mandelstam constraints and this equivalence relation was highlighted and discussed. The next chapter will deal with differential calculus on GLS.

19

LOOP CALCULUS

19.1 INTRODUCTION

In this Chapter we will introduce differential operators defined on generalized path and loop space. More specifically we will discuss the path and area derivative that were used by Makeenko and Migdal to derive their loop equations. Finally we discuss the Fréchet derivative, introduced in Chapter 17, with respect to smooth diffeomorphisms and investigate it's link with Polyakov's infinitesimal loop variation.

19.2 PATH DERIVATIVES

The first class of differential operators we wish to introduce can act on both generalized paths and loops. These operators act on specific locations along the paths and loops, depending on if they operate at the initial or terminal point of the contours they are referred to as the initial- and terminal end-point derivatives. We point out that this class of derivatives depends on a (local) vector field, which for the terminal end-point derivative is assumed to exist in a neighborhood U of the terminal point $q = \gamma(1)$ of the generalized path γ (and in a neighborhood of the initial point for the initial end-point derivative). Writing

$$V(\gamma(1)) = v \in T_{\gamma(1)}M$$

for the vector field at q this local vector field v generates a local integral curve, starting at

$$q = \gamma(1),\ s = 0$$

which we symbolically write as

$$\eta_s^V = \Phi^V(s)(q).$$

We will write

$$\gamma_s = \gamma \cdot \eta_s^V$$

for the new path composed of the combination of the original path γ followed by the local integral curve induced by v and

$$q_s = \gamma_s(1)$$

for the varying end point of the combined path, this is graphically represented in the left panel of Fig. 36. The right panel shows that extending the original path

$$\gamma \to \gamma_s$$

returns a different path, i.e. a different point in the path space $\mathcal{PM}$, from which it is clear that the end point derivatives are actually directional derivatives in $\mathcal{PM}$. The **direction** of this directional derivative is determined by the local vector v, inducing a tangent vector δ_v in the path space $\mathcal{PM}$ which can be interpreted as the **direction** of derivation. Notice that we implicitly as-

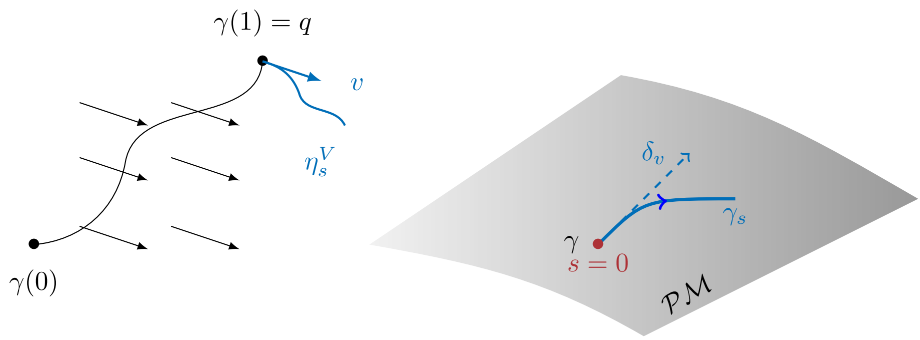


Figure 36: $\gamma_s = \gamma \cdot \eta_s^V$ and $q_s = \gamma_s(1)$

.

sumed a re-parametrization such that the parameter that describes the curve is still in the interval $[0, 1]$. Due to the identification between Chen integrals and generalized paths and loops, this re-parametrization invariance is naturally included. One could also introduce the invariance explicitly by dividing out the equivalence relation for paths that only differ by re-parametrization. In a quantum setting, using the path-integral formalism [128], this results in integrating over all re-parametrizations which gives rise to a constant factor. This factor then divides out if one divides by the vacuum diagrams in the calculation of an expectation value in quantum field theory[1]. With these no-

1 Notice that not in all path or loop spaces described in the literature this re-parametrization invariance is assumed (see for instance [112] for such an example).

tations and parametrizations we can now give the definition of the terminal covariant endpoint derivative.

Definition 19.2.1 (Terminal Covariant Endpoint Derivative).
Let U_γ be a path functional on $\mathcal{PM}$, with values in $\mathbb{R}$ (or $\mathbb{C}; GL(n, \mathbb{C})$). We define the ***terminal covariant endpoint Derivative***

$$\nabla_V^T(q_s)U_\gamma$$

of U_γ, at γ, in the direction of V, as the limit:

$$\nabla_V^T(q_s)U_\gamma = \lim_{\Delta \to 0} \frac{U_{\gamma_{s+\Delta}} - U_{\gamma_s}}{\Delta}. \tag{210}$$

Replacing γ_s by

$$\gamma_s = (\eta_s^V)^{-1} \cdot \gamma,$$

where V is now a vector at the initial end point of the original path γ, defines the ***initial covariant endpoint derivative***

$$\nabla_V^I(q_s)U_\gamma.$$

This is only well defined in a neighborhood of the endpoint

$$q = \gamma(1)$$

or better at

$$q_s = \gamma_s(1),$$

and moreover depends on the vector field V, which gives it the directional derivative like behavior. In the special case that $s = 0$, we can define the Terminal Endpoint Derivative.

Definition 19.2.2 (Terminal Endpoint Derivative).
Let U_γ be a path functional on $\mathcal{PM}$, with values in $\mathbb{R}$ (or in $\mathbb{C}; GL(n, \mathbb{C})$). We define the terminal endpoint derivative

$$\partial_v^T U_\gamma$$

of U_γ, in the direction of

$$v \in T_{\gamma(1)}M,$$

as the limit:

$$\partial_v^T U_\gamma = \lim_{\Delta \to 0} \frac{U_{\gamma_\Delta} - U_\gamma}{\Delta} \tag{211}$$

provided this limit exists independently of the choice of the vector field

$$V \in \mathcal{XM},$$

such that

$$V(\gamma(1)) = v.$$

The following example is instructive to understand the above definition.

Example 19.2.1.
Consider smooth function

$$F \in C^\infty M$$

and define a path functional U_F*, by*

$$U_F[\gamma] = F[\gamma(1)].$$

Applying the terminal covariant endpoint derivative returns

$$\nabla_V^T(q_s)U_F[\gamma] = V \cdot F[q_s] = dF[V_{q_s}]. \tag{212}$$

For the terminal endpoint derivative this reduces to

$$\partial_v^T U_F[\gamma] = V \cdot F[\gamma(1)] = dF[v] \tag{213}$$

which depends only on the vector v*, and not on the particular extension* V*.*

Inspired by this example one can also introduce the concept of a marked path functional, where the marked refers to the fact that it is determined by the evaluation of some function at a certain point along the path. Note that here it might make a difference if one assumes re-parametrization invariance or not!!!

Definition 19.2.3 (Marked Path Functional)**.**
Let U_γ be a path functional and

$$F \in C^\infty M.$$

We define the marked path functional $F \odot U_\gamma$, by:

$$(F \odot U_\gamma)[\gamma] = F[\gamma(1)]\ U_\gamma. \tag{214}$$

The following Lemma shows that the introduced endpoint derivatives are actual derivatives

Lemma 19.2.1: Leibniz Rule

Suppose that U_γ is a path functional for which the limit in Eq. (210) exists, and has the 'continuity condition'

$$\lim_{\Delta\to 0} U_{\gamma_{s+\Delta}} = U_{\gamma_s},\ \forall s \geq 0.$$

The covariant endpoint derivative then obeys the Leibniz rule:

$$\begin{aligned}\nabla_V^T(q_s)(F \odot U_\gamma)(\gamma) &= V \cdot F(q_s)U_{\gamma_s} + F[q_s]\nabla_V^T(q_s)U_\gamma \\ &= \nabla_V^T(q_s)F[\gamma]U_{\gamma_s} + F[\gamma_s]\nabla_V^T(q_s)U_\gamma \end{aligned} \tag{215}$$

where we have put

$$q_s = \gamma_s(1).$$

In particular, if

$$\partial_v^T U_\gamma$$

exists in the sense of Definition 19.2.2, then we have at the terminal endpoint

$$q = \gamma(1)$$

that

$$\partial_v^T(F \odot U_\gamma)[\gamma] = \partial_v^T F[\gamma]U_\gamma + F[\gamma]\partial_v^T U_\gamma \tag{216}$$

which depends only on the vector v, and not of the particular extension V.

Next we would like to investigate the application of the endpoint derivatives to the path functional $X^{\omega_1\cdots\omega_r}$. We start from the following lemma:

Lemma 19.2.2

Let

$$\eta_x = \eta_x^V.$$

Then:

$$\lim_{\Delta\to 0} \frac{\int_{\eta_\Delta} \omega}{\Delta} = \omega(v)$$

$$\lim_{\Delta\to 0} \frac{\int_{\eta_\Delta} \omega_1 \cdots \omega_r}{\Delta} = 0 \quad \forall r \geq 2$$

where

$$\omega, \omega_1, \cdots, \omega_r \in \bigwedge^1 M.$$

Respectively,

$$\omega, \omega_1, \cdots, \omega_r \in \bigwedge^1 M \otimes GL(n, \mathbb{C}).$$

This lemma is easy to prove after introducing local coordinates

$$\omega = A_\mu(x)\; dx^\mu.$$

proof 19.2.1

$$\begin{aligned}
\lim_{\Delta\to 0} \frac{\int_{\eta_s} \omega}{s} &= \lim_{\Delta\to 0} \frac{\int\limits_{\eta_s} A_\mu(x) dx^\mu}{\Delta} \\
&= \lim_{\Delta\to 0} \frac{\int\limits_0^\Delta \left(A_\mu(x(t)) \frac{dx^\mu}{dt}\right) dt}{\Delta} \\
&= \lim_{\Delta\to 0} \frac{\int\limits_0^\Delta \left(A_\mu(x(t)) v^\mu(t)\right) dt}{\Delta} \\
&= A_\mu(x(0)) v^\mu(0) = \omega(v),
\end{aligned} \tag{217}$$

which is valid under the assumption that there are no divergences in the kernel of the integral.

Returning to Lemma 19.2.2 and Eq. (89) with

$$\alpha_1 \cdot \alpha_2 = \gamma \cdot \eta_s$$

where γ is the path from $\gamma(0)$ to $\gamma(s))$ for

$$r \geq 1$$

returns a more useful expression for the covariant endpoint path derivative

$$\nabla_V^T(q_s) X^{\omega_1 \cdots \omega_r}(\gamma) = X^{\omega_1 \cdots \omega_{r-1}}(\gamma_s) \cdot \omega_r(V_{q_s}) \tag{218}$$

which for the terminal endpoint derivative ∂_v^T reduces to

$$\partial_v^T X^{\omega_1 \cdots \omega_r}(\gamma) = X^{\omega_1 \cdots \omega_{r-1}}(\gamma) \cdot \omega_r(v), \ \forall r \geq 1.$$

Equivalent expressions for the initial terminal point derivatives can be derived

$$\nabla_V^I(q_s) X^{\omega_1 \cdots \omega_r}(\gamma) = -\omega_1(V_{q_s}) \cdot X^{\omega_2 \cdots \omega_r}(\gamma_s), \ \forall r \geq 1$$

and

$$\partial_v^I X^{\omega_1 \cdots \omega_r}(\gamma) = -\omega_1(v) \cdot X^{\omega_2 \cdots \omega_r}, \ \forall r \geq 1$$

where again the form are considered to be

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M$$

or

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M \otimes GL(n, \mathbb{C}).$$

Keeping QFT in mind we can consider the commutator of two endpoint derivatives. Applying the commutator to $X^{\omega_1 \cdots \omega_r}$ returns the following result

$$\left[\partial_{u_1}^T, \partial_{u_2}^T\right] X^{\omega_1 \cdots \omega_r}(\gamma) = X^{\omega_1 \cdots \omega_{r-2}}(\gamma) \cdot (\omega_{r-1} \wedge \omega_r)(u_1 \wedge u_2). \tag{219}$$

This can be demonstrated by direct application of Eq. (218) combined with the Leibniz rule and the definitions of the Marked Path Functionals. The above results clearly show that

$$\nabla_V^T(q_s) X^{\omega_1 \cdots \omega_r}(\gamma)$$

given by Eq. (218), is a **marked path functional**, as defined in Definition 19.2.3 with

$$F = \omega_r(V).$$

It then follows that when we consider two vector fields U_1, U_2, locally defined around

$$q = \gamma(1),$$

we can apply the Leibniz rule (Lemma 19.2.1) and the identity ([116, Theorem 4.25])

$$d\omega(U_1, U_2) = U_1 \cdot \omega(U_2) - U_2 \cdot \omega(U_1) - \omega([U_1, U_2]),$$

to derive that at q we get

$$\begin{aligned}\left[\nabla^T_{U_1}(q), \nabla^T_{U_2}(q)\right] X^{\omega_1\cdots\omega_r}(\gamma) = X^{\omega_1\cdots\omega_{r-1}}(\gamma) \cdot d\omega_r(u_1 \wedge u_2) \\ + X^{\omega_1\cdots\omega_{r-2}}(\gamma).(\omega_{r-1} \wedge \omega_r)(u_1 \wedge u_2).\end{aligned} \tag{220}$$

Clearly only the dependence on the local vectors u, v matters, and not on the particular extensions U_1, U_2 such that we can write

$$[\nabla^T_{u_1}, \nabla^T_{u_2}] X^{\omega_1\cdots\omega_r}(\gamma).$$

In the previous Chapters we discussed the equivalence relation on generalized loop space induced by the Wilson loops. In this context we want to know the effects of the endpoint derivatives on a Wilson loop and Wilson line. Therefore we consider the gauge link as a path functional

$$U_\gamma : \mathcal{PM} \to GL(p)$$

which was introduced in Eq. (163). Application of the terminal endpoint derivative returns

$$\partial^T_{u_2} U_\gamma = U_\gamma \cdot \omega(u_2), \tag{221}$$

and for the initial endpoint derivative it returns:

$$\partial^I_{u_2} U_\gamma = -\omega(u_2) \cdot U_\gamma. \tag{222}$$

These results may seem not so spectacular at the moment, but their relevance will become clear when discussing the area derivative in the next Section. Especially the result of the application of the commutator of two endpoint derivatives to the gauge link, which is given by

$$\begin{aligned}[\nabla^T_{u_1}, \nabla^T_{u_2}] U_\gamma &= U_\gamma \cdot (d\omega + \omega \wedge \omega)(u_1 \wedge u_2) \\ &= U_\gamma \cdot \Omega(u_1 \wedge u_2),\end{aligned} \tag{223}$$

where Ω is the curvature of the connection one-form ω in Eq. (161), will become important. In a more familiar QFT notation this becomes

$$\left[\nabla^T_{u_1}, \nabla^T_{u_2}\right] U_\gamma = U_\gamma \cdot F_{\mu\nu}(u_1^\mu \wedge u_2^\nu), \tag{224}$$

We finish this Section by explaining the **covariant** part in the derivatives. The example below explicitly demonstrates where the name covariant endpoint derivative has its origin.

Example 19.2.2.
Given a path

$$\lambda \in \mathcal{PM}_p$$

and a function

$$F \in C^\infty M$$

and, respectively,

$$F \in C^\infty M \otimes GL(n, \mathbb{C}),$$

we define a (marked) path functional through:

$$Z^{\omega_1 \cdots \omega_r}_{(i)}(\lambda; F) \equiv X^{\omega_1 \cdots \omega_i}(\lambda) F(\lambda(1)) X^{\omega_{i+1} \cdots \omega_r}(\lambda^{-1})$$

where

$$\omega_1 \cdots \omega_r \in \bigwedge\nolimits^1 M$$

or

$$\omega_1 \cdots \omega_r \in \bigwedge\nolimits^1 M \otimes GL(n, \mathbb{C}).$$

Applying the Leibniz rule returns

$$\begin{aligned}\nabla^T_v Z^{\omega_1 \cdots \omega_r}_{(i)}(\lambda; F) =& X^{\omega_1 \cdots \omega_i}(\lambda) \cdot dF_q(v) X^{\omega_{i+1} \cdots \omega_r}(\lambda^{-1}) \\ &+ X^{\omega_1 \cdots \omega_{i-1}}(\lambda) \cdot \omega_i(v) \cdot F(q) \cdot X^{\omega_{i+1} \cdots \omega_r}(\lambda^{-1}) \\ &- X^{\omega_1 \cdots \omega_i}(\lambda) \cdot F(q) \cdot \omega_{i+1}(v) \cdot X^{\omega_{i+2} \cdots \omega_r}(\lambda^{-1}),\end{aligned} \tag{225}$$

where

$$q = \lambda(1)$$

Finally let us define, for a connection one-form ω, a (marked) path functional Ψ through

$$\Psi(\lambda; F) \equiv U_\lambda \cdot F(q) \cdot U_{\lambda^{-1}}$$

where

$$q = \lambda(1)$$

U *is the parallel transport operator of the connection* ω *and*

$$F \in C^\infty M \otimes GL(n, \mathbb{C}).$$

Then, using Eq. (225), we compute that

$$\begin{aligned} \nabla_v^T \Psi(\lambda; F) &= U_\lambda \cdot (dF_q(v) + [\omega, F](v)) \cdot U_{\lambda^{-1}} \\ &= U_\lambda \cdot D_q^\omega F(v) \cdot U_{\lambda^{-1}} \end{aligned} \tag{226}$$

where

$$D_q^\omega F(v) \equiv dF_q(v) + [\omega,\ F(v)]$$

stands for the usual covariant derivative of F*. This explains the name of the operator* ∇_v^T*, as terminal endpoint* ***covariant*** *derivative.*

19.3 AREA DERIVATIVE

Having discussed path variations in the previous Section we now turn our attention to the **area derivative**, graphically represented in Fig. 37. To be able

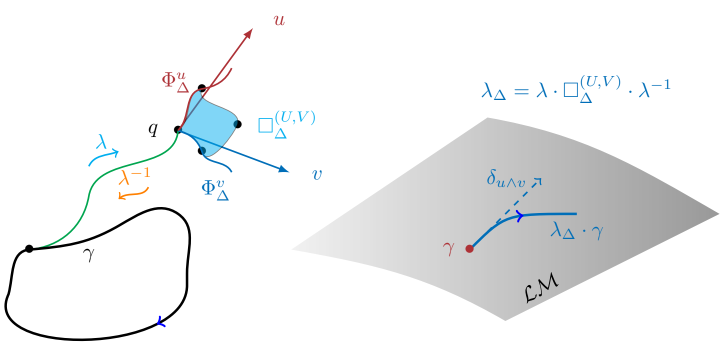


Figure 37: $\Delta_{\lambda; u_1 \wedge u_2}(q) X^{\omega_1 \cdots \omega_r}(\gamma)$

to define this area derivative we first need to define the area variation that we will consider, therefore we introduce the infinitesimal area extension.

Definition 19.3.1 (Ininitesimal Area Extension)**.**
Consider a loop

$$\gamma \in \mathcal{LM}_p,$$

a point

$$q \in M$$

and a path

$$\lambda \in \mathcal{PM}_p,$$

going from

$$p \text{ to } q = \lambda(1).$$

Given an ordered pair (u_1, u_2) *of tangent vectors*

$$u_1, u_2 \in T_qM,$$

we extend them by two commuting vector fields

$$U_1, U_2 \in \mathcal{XU},$$

defined in a small neighborhood $\mathcal{U}$ *of*

$$q = \lambda(1)$$

which is always possible for a smooth manifold. The infinitesimal area variation is then defined by the infinitesimal loop

$$\Box_\Delta^{(U_1,U_2)},$$

based at q*, defined by the local flows* Φ

$$\Box_\Delta^{(U_1,U_2)} = \Phi^{U_2}(-\Delta)\Phi^{U_1}(-\Delta)\Phi^{U_2}(\Delta)\Phi^{U_1}(\Delta)(q) \tag{227}$$

where Φ^{U_1,U_2} *is the local flow of* U_1, U_2*.*

We write λ_Δ for the (Δ-dependent) loop, see Fig. 37, where in the right panel we now have a curve of loops in $\mathcal{LM}$ generated by the infinitesimal loop,

$$\lambda \cdot \Box_\Delta^{(U_1,U_2)} \cdot \lambda^{-1}$$

for which we obtain, due to the path reduction property (Definition 13.3.3),

$$\lim_{\Delta \to 0} \lambda_\Delta = \epsilon,$$

where ϵ is the unity in the group $\boldsymbol{L}\mathcal{M}_p$ (of equivalence classes) of loops based at

$$p \in M.$$

In the classical case one can write

$$\lim_{\Delta \to 0} \lambda_\Delta(X^u) = \lim_{\Delta \to 0} X^u(\lambda_\Delta) = \epsilon(X^u), \ \forall u \in Sh(\Omega). \tag{228}$$

Given that

$$\lambda_\Delta \cdot \gamma$$

represents an infinitesimal deformation of the loop γ, in the topology of $\boldsymbol{L}\mathcal{M}_p$, the **area derivative** can be defined as follows

Definition 19.3.2 (Area Derivative)**.**
Given a loop functional

$$U_\gamma \text{ on } \boldsymbol{L}\mathcal{M}_p,$$

*with values in $\mathbb{R}$ (respectively, $\mathbb{C}; GL(n, \mathbb{C})$), we define its **area derivative**, denoted by*

$$\Delta_{\lambda;(u_1,u_2)}(q) \cdot U_\gamma,$$

as the limit

$$\Delta_{\lambda;(u_1,u_2)}(q) U_\gamma = \lim_{\Delta \to 0} \frac{U_{\lambda_\Delta \cdot \gamma} - U_\gamma}{\Delta^2} \tag{229}$$

provided this limit exists independently of the choice of the vector fields

$$U_1, U_2 \in \mathcal{XU}.$$

Similar to the endpoint derivatives, the area derivative can be interpreted as a directional derivative but now in generalized loop space. The "direction" of derivation is now the loop space tangent vector $\delta_{u \wedge v}$ (see figure 37) of which the exponential map

$$\exp{[\delta_{u \wedge v}]},$$

defined as in the previous Chapter, corresponds to the infinitesimal loop

$$\lambda \cdot \Box_\Delta^{(U_1,U_2)} \cdot \lambda^{-1}.$$

We will see that also for the area derivative we can define an initial and terminal version.

With the goal of applying the area derivative to Wilson Loop variables (for $SU(N)$ gauge theory)[2]

$$W_\gamma = \left\langle 0 \left| \frac{1}{N} \mathrm{Tr}\, \mathcal{P} \exp \left[ig \oint_\gamma A_\mu(x) dx^\mu \right] \right| 0 \right\rangle, \tag{230}$$

we investigate the application of this derivative to the Chen iterated integrals

$$X^{\omega_1 \cdots \omega_r} : \boldsymbol{L}\mathcal{M}_p \to \mathbb{R}$$

(respectively, $\mathbb{C}; GL(n, \mathbb{C})$). Similar to the path derivative, we can show that the area derivative of the functionals $X^{\omega_1 \cdots \omega_r}$ is well-defined considering that the kernels of the integrals in the following Lemma do not contain any divergences.

Lemma 19.3.1

Write $\square_\Delta$ for

$$\square_\Delta^{(U_1, U_2)}$$

as before.

$$\lim_{\Delta \to 0} \frac{\int_{\square_\Delta} \omega}{\Delta^2} = \int_V d\omega = d\omega(u_1 \wedge u_2) \tag{231}$$

$$\lim_{\Delta \to 0} \frac{\int_{\square_\Delta} \omega_1 \omega_2}{\Delta^2} = (\omega_1 \wedge \omega_2)(u_1 \wedge u_2) \tag{232}$$

$$\lim_{\Delta \to 0} \frac{\int_{\square_\Delta} \omega_1 \cdot \omega_r}{\Delta^2} = 0, \ \forall r \geq 3 \tag{233}$$

where

$$\omega, \omega_1, \cdots, \omega_r \in \bigwedge\nolimits^1 M$$

or

$$\omega, \omega_1, \cdots, \omega_r \in C; \bigwedge\nolimits^1 M \otimes GL(n, \mathbb{C}).$$

Introduction of coordinates, together with using Stokes' Theorem it is easy to proof this Lemma. However, one does need to make a remark with respect

2 Where we assume the transition from classical to quantum case is allowed, in other words the expectation values can be taken in a consistent way.

to the goal of applying the area derivative to the Wilson loop variables from Eq. (230). One subtle point can be highlighted by considering the integral in Eq. (232), but rewritten in a more familiar gauge theory notation:

$$\int_{\square_t} \omega_1\omega_2 = \int_{\square_t} A_\mu A_\nu,$$

where A_μ, A_ν are the gauge potentials. This integral is well-defined in the classical case but will become problematic in a field theory setting when taking vacuum expectation values, even when considering both one-forms to be locally constant the vacuum expectation value will give rise to a (divergent) tadpole. Fortunately this tadpole can be taken care of by a convenient regularization scheme such that in some cases we can deal with this problem. But, in the general case, more specifically in the case where the Wilson loops are lying entirely on the light cone, where we have additional light cone divergencies, one will not be able to resolve this issue. The question is then if one can interchange the integrals and the vacuum expectation values as one usually does.

For the remainder of this text we will assume that the integrals in Lemma 19.3.1 are well-defined and use the values shown there. Notice that the above defined area derivative introduces extra cusps along the contour, which is the main cause of divergencies of the integrals in Lemma 19.3.1 in a QFT setting. In the next Section we will discuss an alternative area derivative that in a sense avoids the introduction of extra cusps.

Again motivated by the gauge links, we are interested in the effects of the area derivative on the functionals $X^{\omega_1\cdots\omega_r}$. To investigate this we first define the following **derivation**

Definition 19.3.3.
For

$$u_1 \wedge u_2 \in \bigwedge^2 T_q M$$

define a derivation

$$D_{u_1\wedge u_2}(q),$$

in the algebra of iterated integrals, by

$$D_{u_1\wedge u_2}(q) X^{\omega_1\cdots\omega_r} = X^{\omega_1\cdots\omega_{r-1}} \cdot d\omega_r(u_1 \wedge u_2) \tag{234}$$

From the (algebraic) commutator

$$\left[\partial^T_{u_1}, \partial^T_{u_2}\right]$$

of two terminal endpoint derivatives at q, we also define the derivation

$$\mathcal{D}_{u_1 \wedge u_2}(q)$$

according to:

$$\mathcal{D}_{u_1 \wedge v_2}(q) = D_{u_1 \wedge u_2}(q) + \left[\partial^T_{u_1}, \partial^T_{u_2}\right], \tag{235}$$

for which we formulate the Lemma below, establishing its relationship to the area derivative

Lemma 19.3.2

Let

$$\Delta_{\lambda;(u_1,u_2)}(q)$$

be as introduced in Definition 19.3.2. Then

$$\Delta_{\lambda;(u_1,u_2)}(q) X^{\omega_1 \cdots \omega_r}(\epsilon) = \sum_{i=1}^{r} (\mathcal{D}_{u_1 \wedge u_2}(q) X^{\omega_1 \cdots \omega_i}(\lambda)) \Big(X^{\omega_{i+1} \cdots \omega_r}(\lambda^{-1}) \Big) \tag{236}$$

where

$$\omega_1 \cdots \omega_r \in \bigwedge\nolimits^1 M,$$

or, respectively,

$$\omega_1 \cdots \omega_r \in \mathbb{C}, \ \bigwedge\nolimits^1 M \otimes GL(n, \mathbb{C}).$$

We introduce the notation

$$\Delta_{(\lambda;u_1 \wedge u_2)}(q) X^{\omega_1 \cdots \omega_r}(\epsilon),$$

to emphasize that the derivative only depend on $u_1 \wedge u_2$ of the local vectors. The proof of this Lemma can be obtained by combining the properties of the Chen iterated integrals for products of loops in Definition 19.3.2 and comparing the resulting expression with the results of applying the derivative, Definition 19.3.3, to the functionals $X^{\omega_1 \cdots \omega_r}$.

Example 19.3.1.

$$\begin{aligned}\Delta_{(\lambda;u_1\wedge u_2)}(q)X^{\omega}(\epsilon) &= d\omega(u_1\wedge u_2)\\ \Delta_{(\lambda;u_1\wedge u_2)}(q)X^{\omega_1\omega_2}(\epsilon) &= d\omega_1(u_1\wedge u_2)X^{\omega_2}(\lambda^{-1})\\ &\quad + X^{\omega_1}(\lambda)\cdot d\omega_2(u_1\wedge u_2)\\ &\quad + (\omega_1\wedge\omega_2)(u_1\wedge u_2)\end{aligned}$$

and, more generally

$$\Delta_{(\lambda;u_1\wedge u_2)}(q)X^{\omega_1\cdots\omega_r}(\epsilon) = \sum_{i=1}^{r} X^{\omega_1\cdots\omega_{i-1}}(\lambda)\cdot d\omega_i(u_1\wedge u_2)X^{\omega_{i+1}\cdots\omega_r}(\lambda^{-1})$$
$$+\sum_{i=2}^{r} X^{\omega_1\cdots\omega_{i-2}}(\lambda)\cdot(\omega_{i-1}\wedge\omega_i)(u_1\wedge u_2)\cdot X^{\omega_{i+1}\cdots\omega_r}(\lambda^{-1}). \tag{237}$$

Keeping in mind that

$$\lim_{\Delta\to 0}\lambda_\Delta = \epsilon,$$

still assuming that the integrals in Lemma 19.3.1 are well-defined, one can show that

$$\lim_{\Delta\to 0}\frac{\lambda_\Delta-\epsilon}{\Delta} = 0$$

by introducing local coordinates. At the same time we find that

$$\lim_{\Delta\to 0}\frac{\lambda_\Delta-\epsilon}{\Delta^2}$$

exists and is actually the area derivative. We write now

$$\delta_{(\lambda;u\wedge v)}$$

for the operator in the algebra of iterated integrals $\mathcal{A}_p$, defined through the derivations from Eqs. (234) and (235):

$$\begin{aligned}\delta_{(\lambda;u_1\wedge u_2)}X^{\boldsymbol{u}} &\equiv \Delta_{(\lambda;u_1\wedge u_2)}(q)X^{\boldsymbol{u}}(\epsilon)\\ &= (\lambda\otimes\lambda)((\mathcal{D}_{u_1\wedge u_2}(q)\otimes J)\circ\Delta)X^{\boldsymbol{u}},\ \forall\boldsymbol{u}\in\mathrm{Sh},\end{aligned} \tag{238}$$

where in the last line J is the antipode and Δ co-multiplication of the Hopf algebra structure on $\mathcal{A}_p$.

The last equality can be demonstrated from combining the definitions of the operators written in the last line of Eq. (238) with the left action of a loop on the space of generalized loops. Since this is a topological group as we have seen in the previous Section, this action is well-defined. In a similar way it is possible to show that

$$\delta_{(\lambda;u_1\wedge u_2)}$$

is a **pointed derivation** at ϵ, such that we have

$$\delta_{(\lambda;u_1\wedge u_2)}(X^{u_1}X^{u_2}) = \delta_{(\lambda;u_1\wedge u_2)}(X^{u_1})\epsilon(X^{u_2}) + \epsilon(X^{u_1})\delta_{(\lambda;u_1\wedge u_2)}(X^{v}),$$
$$\forall u_1, u_2 \in Sh(\Omega).$$

From Eq.(238) we see that

$$\delta_{(\lambda;u_1\wedge u_2)} : \mathcal{A}_p \to k$$

is a linear map.

In Section 18.6 of the previous Chapter we described the Lie algebra on the generalized loop space and derived that the tangent space $T_\epsilon \boldsymbol{L}\mathcal{M}_p$, to the group $\boldsymbol{L}\mathcal{M}_p$, at ϵ, is the k-linear subspace of $\mathcal{A}_p^*$. Here we will demonstrate that this space is generated by all the

$$\delta_{(\lambda;u_1\wedge u_2)}.$$

We start by considering a loop

$$\gamma \in \boldsymbol{L}\mathcal{M}_p$$

and evaluate the area derivative (Fig. 38),

$$\triangle_{(\lambda;u_1\wedge u_2)}(q)X^{\omega_1\cdots\omega_r}(\gamma).$$

With these definitions we have the following Lemma

Lemma 19.3.3

Let

$$\gamma \in \boldsymbol{L}\mathcal{M}_p,$$

$$\lambda \in \mathcal{P}\mathcal{M}_p$$

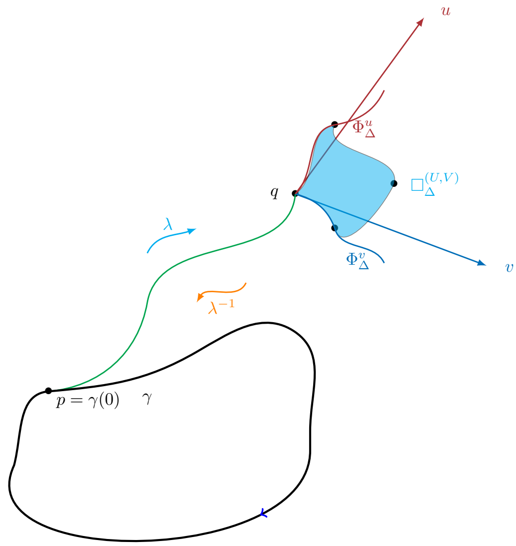


Figure 38: $\triangle_{\lambda;u_1\wedge u_2}(q)X^{\omega_1\cdots w_r}(\gamma)$

and

$$u_1 \wedge u_2 \in \bigwedge^2 T_{\lambda(1)}M.$$

Then

$$\triangle_{(\lambda;u_1\wedge u_2)}(q)X^{\omega_1\cdots\omega_r}(\gamma) = \sum_{i=1}^{r} \triangle_{(\lambda;u_1\wedge u_2)}(q)X^{\omega_1\cdots\omega_i}(\epsilon)X^{\omega_{i+1}\cdots\omega_r}(\gamma)$$
$$= \gamma \circ (\delta_{(\lambda;u_1\wedge u_2)} \otimes 1) \circ \Delta(X^{\omega_1\cdots\omega_r})$$

with

$$(\delta_{(\lambda;u_1\wedge u_2)} \otimes 1) \circ \Delta$$

the right invariant derivation on the algebra $\mathcal{A}_p$, which is associated to the tangent vector

$$\delta_{(\lambda;u_1\wedge u_2)}.$$

This Lemma allows us to use the notation

$$\triangle^R_{(\lambda;u_1\wedge u_2)} : \boldsymbol{L}\mathcal{M}_p \to \mathcal{A}^*_p,$$

given by

$$\triangle^R_{(\lambda;u_1\wedge u_2)}(\gamma) \equiv \gamma \circ (\delta_{(\lambda;u_1\wedge u_2)} \otimes 1) \circ \Delta$$

as its designation as **the right invariant 'vector field'** on $\boldsymbol{L}\mathcal{M}_p$, determined by

$$\delta_{(\lambda;u_1\wedge u_2)}.$$

In the special case that

$$\lambda = \epsilon,$$

we call

$$\triangle_{(\epsilon;u_1\wedge u_2)}(p),$$

similar to the path derivatives, the **initial endpoint area derivative** and denote it by

$$\triangle^I_{(\epsilon;u_1\wedge u_2)}(p),$$

as visualized in Fig. 39. In this situation the area derivative reduces to

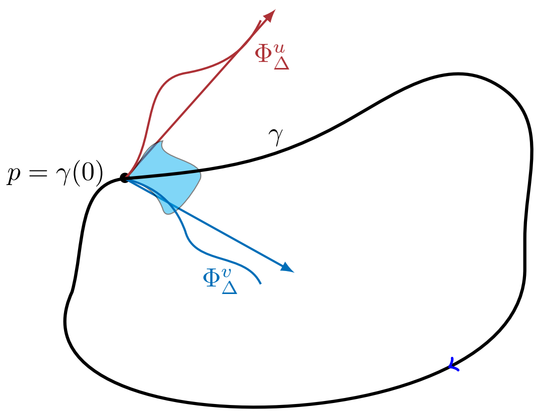


Figure 39: $\triangle^I_{(\epsilon;u_1\wedge u_2)}(p)$

$$\triangle^I_{(\epsilon;u_1\wedge u_2)}(p)X^{\omega_1\cdots\omega_r}(\gamma) = d\omega_1(u_1\wedge u_2)\cdot X^{\omega_2\cdots\omega_r}(\gamma) + (\omega_1\wedge\omega_2)(u_1\wedge u_2)\cdot X^{\omega_3\cdots\omega_r}(\gamma).$$

Another possibility, similar to the path derivatives, is to consider

$$\lambda = \gamma\cdot\eta,$$

$$\gamma \in \boldsymbol{L}\mathcal{M}_p,$$

$$\eta \in \mathcal{P}\mathcal{M}_p,$$

and

$$u_1 \wedge u_2 \in \bigwedge^2 T_{\eta(1)}M.$$

In this case, Fig. 40,

$$\begin{aligned}\lambda_t \cdot \gamma \equiv (\lambda \cdot \Box_t^{(U_1,U_2)} \cdot \lambda^{-1}) \cdot \gamma = \gamma \cdot \eta \cdot \Box_t^{(U_1,U_2)} \cdot \gamma \cdot \eta^{-1} \cdot \gamma \\ = \gamma \cdot (\eta \cdot \Box_t^{(U_1,U_2)} \cdot \eta^{-1}) \equiv \gamma \cdot \eta_t,\end{aligned}$$

such that we refer to this area derivative as the **terminal endpoint area derivative** which we write symbolically as

$$\triangle^E_{(\eta;u_1\wedge u_2)}(q).$$

An explicit equation for this derivative applied to the functionals $X^{\omega_1\cdots\omega_r}$ is given by

$$\begin{aligned}\triangle^E_{(\eta;u_1\wedge u_2)}(q)X^{\omega_1\cdots\omega_r}(\gamma) &= \sum_{i=1}^{r} X^{\omega_1\cdots\omega_i}(\gamma)\triangle_{(\eta;u_1\wedge u_2)}(q)X^{\omega_{i+1}\cdots\omega_r}(\epsilon) \\ &= \gamma \circ (1 \otimes \delta_{(\eta;u_1\wedge u_2)}) \circ \Delta(X^{\omega_1\cdots\omega_r}) \qquad (239)\end{aligned}$$

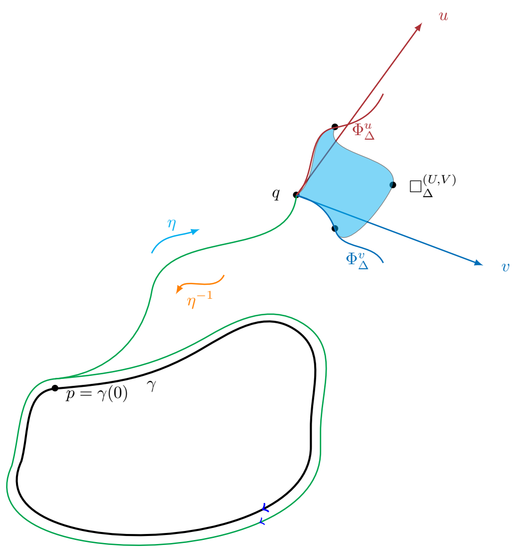


Figure 40: $\triangle^E_{(\eta;u_1\wedge u_2)}(q)$

Similar to the right invariant derivations, we can also introduce the left invariant derivation

$$(1 \otimes \delta_{(\eta;u_1 \wedge u_2)}) \circ \Delta$$

associated to $\delta_{(\eta;u_1 \wedge u_2)}$. Naturally

$$\triangle^L_{(\eta;u_1 \wedge u_2)} : \boldsymbol{L}\mathcal{M}_p \to \mathcal{A}^*_p,$$

given by

$$\triangle^L_{(\eta;u_1 \wedge u_2)}(\gamma) \equiv \gamma \circ (1 \otimes \boldsymbol{\delta}_{(\eta;u_1 \wedge u_2)}) \circ \Delta, \tag{240}$$

now referred to as **the left invariant "vector field"**, on

$$\boldsymbol{L}\mathcal{M}_p,$$

determined by

$$\delta_{(\eta;u_1 \wedge u_2)}.$$

As for the path derivatives we have the special case

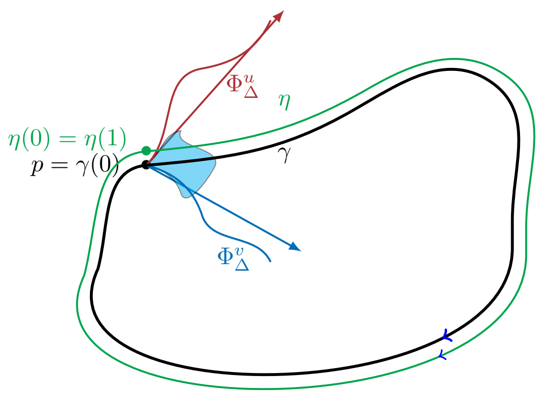


Figure 41: $\triangle^I_{(\epsilon;u_1 \wedge u_2)}(p)$

$$\eta = \epsilon$$

(see Fig. 41), such that the above formula simplifies to

$$\begin{aligned}\triangle^E_{(\epsilon;u_1 \wedge u_2)}(p) X^{\omega_1 \cdot \omega_r}(\gamma) &= \mathcal{D}_{u_1 \wedge u_2}(p) X^{\omega_1 \cdots \omega_r} \\ &= X^{\omega_1 \cdots \omega_{r-1}}(\gamma) \cdot d\omega_r(u \wedge v) \\ &\quad + X^{\omega_1 \cdots \omega_{r-2}}(\gamma) \cdot (\omega_{r-1} \wedge \omega_r)(u_1 \wedge u_2).\end{aligned} \tag{241}$$

As indicated in the previous Section, this last case is particularly interesting because we can relate the area derivative to the Lie bracket of terminal end point path derivations, Eq. (220)

$$\triangle_{(\epsilon;u_1\wedge u_2)}(q)X^{\omega_1\cdots\omega_r}(\gamma) = \left[\nabla_u^T, \nabla_v^T\right]X^{\omega_1\cdots\omega_r}(\gamma),$$

an interesting result. If, in this specific case, we consider not the functionals $X^{\omega_1\cdots\omega_r}$ but the holonomy U_γ instead, we obtain

$$\begin{aligned}\triangle^E_{(\epsilon;u_1\wedge u_2)}(p)U_\gamma &= U_\gamma \cdot (d\omega + \omega\wedge\omega)(u_1\wedge u_2)\\ &= U_\gamma \cdot \Omega(u_1\wedge u_2)\end{aligned} \tag{242}$$

where again Ω is the curvature of the connection ω. The fact that $\mathcal{A}_p$ is of trace class, we can apply this derivation to the Wilson loop W:

$$\begin{aligned}\triangle^E_{(\epsilon;u_1\wedge u_2)}(p)W(\gamma) &= \mathrm{Tr}((d\omega + \omega\wedge\omega)(u_1\wedge u_2)\cdot U_\gamma)\\ &= \mathrm{Tr}(\Omega(u_1\wedge u_2)\cdot U_\gamma).\end{aligned} \tag{243}$$

where these last equations are also referred to as the **Mandelstam formulas**. Since we are dealing with Lie algebras, it is not a surprise that we also have a **Bianchi identity**

Theorem 19.3.1: Bianchi Identity

$$\sum_{\mathrm{cycl}\{u_1,u_2,u_3\}} \nabla^T_{u_1}(\lambda(1))\delta_{(\lambda;u_2\wedge u_3)} = 0, \tag{244}$$

where

$$\sum_{\mathrm{cycl}\{u_1,u_2,u_3\}}$$

stands for the sum over the cyclic permutations of the vectors u_1, u_2, u_3.

Analogous to the path derivative case we can again consider the commutator of two area derivatives, which as elements of the Lie algebra, will allow the formal determination of the structure constants of this algebra

$$\left[\delta_{(\lambda;\mathbf{a_1}\wedge\mathbf{a_2})}, \delta_{(\eta;\mathbf{u_1}\wedge\mathbf{u_2})}\right] = \delta_{(\lambda;a_1\wedge a_2)}\star\delta_{(\eta;u_1\wedge u_2)} - \delta_{(\eta;u_1\wedge u_2)}\star\delta_{(\lambda;a_1\wedge a_2)}. \tag{245}$$

Using the definitions, introduced in this Section, of the area derivative this can be written as:

$$[\delta_{(\lambda;a_1\wedge a_2)}, \delta_{(\eta;u_1\wedge u_2)}]X^{\omega_1\cdots\omega_r} =$$
$$\sum_{i=0}^{r}\sum_{k=0}^{i}(\mathcal{D}_{a_1\wedge a_2}(\lambda(1))X^{\omega_1\cdots\omega_k}(\lambda))\Big(X^{\omega_{k+1}\cdots\omega_i}(\lambda^{-1})\Big)\delta_{(\eta;u_1\wedge u_2)}(X^{\omega_{i+1}\cdots\omega_r})$$
$$-\sum_{i=0}^{r}\sum_{k=0}^{i}(\mathcal{D}_{u_1\wedge u_2}(\eta(1))X^{\omega_1\cdots\omega_k}(\eta))\Big(X^{\omega_{k+1}\cdots\omega_i}(\eta^{-1})\Big)\delta_{(\lambda;a\wedge b)}(X^{\omega_{i+1}\cdots\omega_r}). \tag{246}$$

With this we end our introduction of the area derivative and move on to the variational derivative.

19.4 VARIATIONAL CALCULUS

In the previous Section we have introduced the area derivative, which depends on two independent local vector fields. A crucial problem with these derivatives in the context of QFT, while calculating perturbative matrix elements, vacuum expectation values etc., is that they may introduce extra cusps (angle-like obstructions) in the contours and, consequently, may generate extra singularities in the perturbative expansion. To handle this problem we need a derivative, that does not introduce extra singularities. In this Section, following Tavares [16], we shall introduce such a derivative, that is derived from area variations that are generated by **diffeomorphisms**, which can be related to the Fréchet derivative, a differential operator situated, in a sense, between the path- and an area-derivatives.

To introduce this derivative we start by considering $\mathrm{Diff}(M)$, the diffeomorphism group of M. Let now

$$\varphi \in \mathrm{Diff}(M)$$

be a diffeomorphism of M and

$$\gamma \in \mathcal{PM}$$

a path in M. Then we have

$$\varphi \cdot \gamma$$

for the image of the path γ under the diffeomorphism φ. From elementary manifold theory we see that the action of the diffeomorphism φ on the functionals $X^{\omega_1\cdots\omega_r}$ is given by

$$X^{\omega_1\cdots\omega_r}(\varphi\cdot\gamma)=X^{\varphi^*\omega_1\cdots\varphi^*\omega_r}(\gamma) \tag{247}$$

where, as before,

$$\omega_1,\cdots,\omega_r\in\bigwedge^1 M$$

or

$$\omega_1,\cdots,\omega_r\in\mathbb{C};\bigwedge^1 M\otimes GL(n,\mathbb{C})$$

and $\varphi^*\omega_i$ are the **pull-backs** of ω_i under the map φ. We now restrict ourselves to the diffeomorphisms that form a one-parameter group, infinitesimally generated by

$$Y\in\mathcal{X}M,$$

a vector field on M. This vector field generates a one parameter group of active diffeomorphisms $\psi(t)$ by the identification

$$\psi_t^Y(p):=c_p^Y(t), \tag{248}$$

where

$$t\to c_p^Y(t)$$

is the maximal integral curve in M starting at $p\in M$ with tangential vector field Y at each point of the curve. The situation is similar to the integral curves associated with directional derivatives. For a composition of diffeomorphisms we have

$$\psi_t^Y(p)\circ\psi_s^Y(p)=\psi_{s+t}^Y(p),$$

such that $\mathrm{Diff}(M)$ is indeed a group. The existence of a local form for the above vector field allows us to define a Lie derivative for any tensor field by the identification [116, Chapter 4]

$$(\mathcal{L}_Y(T))(p):=\left(\frac{d}{ds}\Big|_{s=0}\left(\psi_s^Y\right)^*T\right)(p), \tag{249}$$

which being applied to Eq. (247) results in

$$\begin{aligned}D_VX^{\omega_1\cdots\omega_r}(\gamma)&\equiv\frac{d}{ds}\Big|_{s=0}X^{\omega_1\cdots\omega_r}(\varphi_s\cdot\gamma)\\&=\sum_{i=1}^{r}X^{\omega_1\cdots\omega_{i-1}(\mathcal{L}_Y\omega_i)\omega_{i+1}\cdots\omega_r}(\gamma),\end{aligned} \tag{250}$$

where

$$\omega_1, \cdots, \omega_r \in \bigwedge^1 M$$

or

$$\omega_1, \cdots, \omega_r \in \mathbb{C}; \bigwedge^1 M \otimes GL(n, \mathbb{C})$$

and $\mathcal{L}_V \omega$ refers to the Lie derivative of the one-form ω, in the direction of V. From manifold theory we have Cartan's formula for the Lie derivative

$$\mathcal{L}_Y = \iota_Y d + d\iota_Y, \tag{251}$$

where

$$\iota_Y \omega^n(v_2, \cdots, v_n) := \omega^n(Y, v_2, \cdots, v_n)$$

is the interior product (Definition 7.2.7). Using this in Eq. (250) it reduces to[3]

$$\begin{aligned} D_V X^{\omega_1 \cdots \omega_r}(\gamma) = & \sum_{i=1}^{r} X^{\omega_1 \cdots \omega_{i-1} \cdot (\iota_Y d\omega_i) \cdot \omega_{i+1} \cdots \omega_r}(\gamma) \\ & + \sum_{i=2}^{r} X^{\omega_1 \cdots \omega_{i-2} \cdot \iota_Y(\omega_{i-1} \wedge \omega_i) \cdots \omega_{i+1} \cdots \omega_r}(\gamma) \\ & + \omega_r(V(1)) X^{\omega_1 \cdots \omega_{r-1}}(\gamma) - \omega_1(V(0)) X^{\omega_2 \cdots \omega_r}(\gamma). \end{aligned} \tag{252}$$

From Chapter 17 we know that, the Fréchet derivative applied to the path functionals $X^{\omega_1 \cdots \omega_r}$ is given by the linear map

$$D_V X^{\omega_1 \cdots \omega_r}(\gamma) : T_\gamma \mathcal{LM} \to \mathbb{C}$$

or respectively to $GL(n, \mathbb{C})$ defined by

$$D_V X^{\omega_1 \cdots \omega_r}(\gamma) \equiv \frac{d}{ds}\Big|_{s=0} X^{\omega_1 \cdots \omega_r}(\gamma_s).$$

In this way

$$D_V X^{\omega_1 \cdots \omega_r}(\gamma) \text{ associated with } V$$

is now a Fréchet derivative of

$$X^{\omega_1 \cdots \omega_r} \text{ at } \gamma,$$

3 Note the different limits of the summations.

in the direction of the *tangent vector*

$$V = Y \circ \gamma \in \gamma^* TM,$$

as introduced in Definition 17.2.2.

Restricting ourselves to the **pointed diffeomorphism group** $\mathrm{Diff}_p(M)$, the diffeomorphisms φ that fix the point p, and keeping in mind that this is also a topological group, we can consider it's 'Lie algebra' $\mathcal{X}_p(M)$. By definition this algebra consists of the vector fields Y that vanish on p. The action of the elements of this algebra on the algebra $\mathcal{A}_p$ can be naturally defined by making use of the pull-backs of the one-forms ω_i:

$$\begin{aligned}(\varphi, X^{\omega_1 \cdots \omega_r}) &\mapsto \varphi \cdot X^{\omega_1 \cdots \omega_r} \\ &\equiv X^{\varphi^* \omega_1 \cdots \varphi^* \omega_r}.\end{aligned} \tag{253}$$

Diffeomorphisms do not change the algebraic structure of the one-forms, thus it preserves the Hopf algebra structure and as such φ is an Hopf algebra automorphism, written explicitly

$$\begin{aligned}\varphi \cdot (X^u X^v) &= (\varphi \cdot X^u) \cdot (\varphi \cdot X^v) \\ \Delta \circ \varphi &= (\varphi \otimes \varphi) \circ \Delta.\end{aligned} \tag{254}$$

As a direct result φ induces an automorphism of $\widetilde{\boldsymbol{L}\mathcal{M}_p}$, through the identification:

$$\begin{aligned}\varphi \cdot \tilde{\alpha}(X^{\omega_1 \cdots \omega_r}) &\equiv \tilde{\alpha}(\varphi \cdot X^{\omega_1 \cdots \omega_r}) \\ &= \tilde{\alpha}(X^{\varphi^* \omega_1 \cdots \varphi^* \omega_r})\end{aligned} \tag{255}$$

where now φ, as an element of

$$\mathrm{Aut}(\widetilde{\boldsymbol{L}\mathcal{M}_p}),$$

has a differential (which maps the tangent space of the domain to the tangent space of the codomain, in this case to itself)

$$d\varphi : \widetilde{l\mathcal{M}_p} \to \widetilde{l\mathcal{M}_p},$$

defined as in standard differential geometry by

$$d\varphi(\delta)(X^{\omega_1 \cdots \omega_r}) \equiv \delta(X^{\varphi^* \omega_1 \cdots \varphi^* \omega_r}), \tag{256}$$

with δ a tangent vector (or derivation) of $\widetilde{\boldsymbol{L}\mathcal{M}}_p$. Using the differential

$$\varphi \mapsto d\varphi$$

produces a linear representation of the pointed diffeomorphism group

$$\mathrm{Diff}_p(M) \text{ on } \widetilde{l\mathcal{M}}_p,$$

similar to the situation with Lie groups where the Lie algebra forms a linear representation to the considered Lie group. The infinitesimal action of

$$Y \in \mathcal{X}_p(\mathcal{M}) \text{ on } \delta,$$

written as $Y \cdot \delta$, can be represented by

$$(Y \cdot \delta)(X^{\omega_1 \cdots \omega_r}) = \sum_{i=1}^{r} \delta\Big(X^{\omega_1 \cdots \omega_{i-1} \cdot (\mathcal{L}_Y \omega_i) \cdot \omega_{i+1} \cdots \omega_r}\Big), \tag{257}$$

where we used

$$Y(0) = 0 = Y(1).$$

Making use of Cartan's expression Eq. (251) for the Lie derivative and the expressions that defined the ideal J_p of $\mathcal{A}_p$, the above result can be reduced to

$$\begin{aligned}(Y \cdot \delta)(X^{\omega_1 \cdots \omega_r}) = &\sum_{i=1}^{r} \delta\Big(X^{\omega_1 \cdots \omega_{i-1} \cdot (\iota_Y d\omega_i) \cdot \omega_{i+1} \cdots \omega_r}\Big) \\ &+ \sum_{i=2}^{r} \delta\Big(X^{\omega_1 \cdots \omega_{i-2} \cdot \iota_Y (\omega_{i-1} \wedge \omega_i) \cdot \omega_{i+1} \cdots \omega_r}\Big),\end{aligned} \tag{258}$$

where ι_Y is again the interior product.

19.5 FRÉCHET DERIVATIVE IN GENERALIZED LOOP SPACE

In this Section we discuss the connection between the Fréchet derivative and diffeomorphisms in more detail, more specifically we will discuss how the diffeomorphism generating vector field V from the previous section becomes a variational vector field. Let

$$\gamma \in \mathcal{PM}_p$$

be a path, based at p, with $T_\gamma \mathcal{PM}_p$ the tangent space of $\mathcal{PM}_p$ at γ as visualized in Fig. 42. The vector fields along γ are defined throughout the

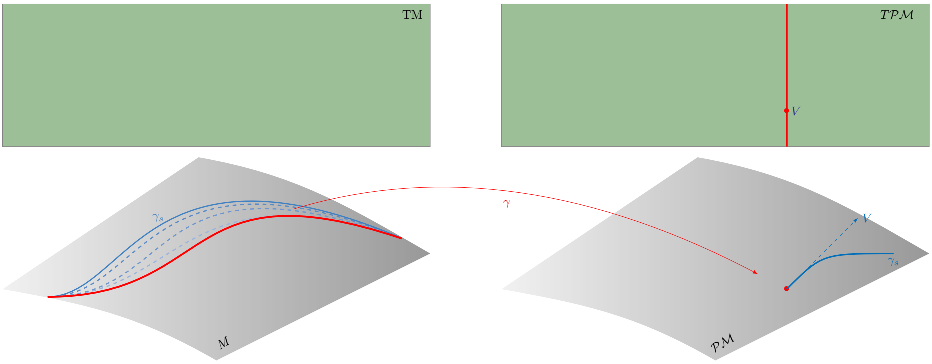


Figure 42: Diffeomorphism of a path

pull-back bundle $\gamma^*T\mathcal{M}$. Notice that these vanish at p since this point needs to stay fixed. Choose such a vector

$$V \in T_\gamma \mathcal{PM}_p.$$

Now defining

$$s \mapsto \gamma_s$$

as a curve of paths in $\mathcal{PM}_p$, starting at γ, in $s = 0$, with velocity V, we can write:

$$\gamma_0 = \gamma \tag{259}$$

$$V(t) = \frac{\partial}{\partial s}\Big|_{s=0} \gamma_s(t) \tag{260}$$

$$V(0) = 0. \tag{261}$$

The map

$$s \mapsto \gamma_s$$

is the variation of $\gamma = \gamma_0$, with associated variational vector field V. In the special case that the variation γ_s is induced by a diffeomorphism, like in the previous Section

$$\gamma_s = \varphi_s \circ \gamma,$$

and the vector field is the diffeomorphism generator

$$V = Y \circ \gamma$$

we can determine the Fréchet derivative of the path functionals $X^{\omega_1\cdots\omega_r}$, at

$$\gamma \in \mathcal{PM}_p.$$

This derivative was defined in Definition 17.2.2 as is the linear map

$$D_{V}\cdot X^{\omega_1\cdots\omega_r}(\gamma) : T_\gamma\mathcal{LM} \to \mathbb{R}$$

(respectively, $\mathbb{C}, GL(n, \mathbb{C})$). In the previous Section we concluded that in the case we are considering here it can be written as

$$D_V X^{\omega_1\cdots\omega_r}(\gamma) \equiv \frac{d}{ds}\Big|_{s=0} X^{\omega_1\cdots\omega_r}(\gamma_s)$$

The demonstration of this result depends on the following Lemma (see also [129, Chapter 12] or [116, Chapter 4])

Lemma 19.5.1

Let N be a manifold (in our case, N will be I or S^1),

$$\gamma : N \to M$$

an immersion (Definition 7.2.13), and ω a differential form in M. Assume that

$$\Gamma : N \times [0, \epsilon] \to M$$

is a smooth variation of γ, with **variational** vector field V. That is, putting

$$\gamma_s(t) = \Gamma(t, s),$$

$$\forall (t, s) \in N \times [0, \epsilon],$$

we have

$$\gamma_0 = \gamma$$

and

$$V(t) = \frac{\partial}{\partial s}\Big|_{s=0} \Gamma(t, s) = \Gamma_{*(t,0)}\left(\frac{\partial}{\partial s}\Big|_{(t,0)}\right),$$

$$\forall t \in N.$$

Then, as differential forms on

$$N = N \times \{0\}$$

we have:

$$\frac{d}{ds}\Big|_{s=0} {\gamma_s}^* \omega = \gamma^*(\iota_V d\omega + d(\iota_V \omega))$$
$$= \gamma^*(\iota_V d\omega) + d(\gamma^*(\iota_V \omega)). \tag{262}$$

In this Lemma $\iota_{V(t)}\omega$ is the interior product of the form $\omega(\gamma(t))$ with

$$V(t) \in T_{\gamma(t)}\mathcal{M}$$

and d is the usual differential operator. Consider now the case where

$$\gamma : I \to M$$

is an immersed (Definition 7.2.13) path, based at p and γ_s a variation of this path generated by the variational vector field V. The Fréchet derivative then becomes

$$\begin{aligned}\frac{d}{ds}\Big|_{s=0}\left(\int_{\gamma_s}\omega\right) = \frac{d}{ds}\Big|_{s=0}\left(\int_I {\gamma_s}^*\omega\right) &= \int_I \gamma^*(\iota_V d\omega + d(\iota_V\omega)) \\ &= \int_I \gamma^*(\iota_V d\omega) + \int_{\partial I} \gamma^* d(\iota_V \omega) \\ &= \int_I \gamma^*(\iota_V d\omega) + \omega(V(1)) - \omega(V(0)) \\ &= \int_\gamma \iota_V d\omega + \omega(V(1)) \end{aligned} \tag{263}$$

where we used the above Lemma (see [116, Eq. 4.33]) and where in the last equality we have used the identity

$$\int_\gamma \iota_V d\omega \ = \int_I \gamma^*(\iota_V \omega).$$

Applying this result to a loop

$$\gamma \in \mathcal{LM}_p$$

and using the fact that

$$V(0) = 0 = V(1)$$

results in

$$D_V X^\omega(\gamma) = X^{\iota_V d\omega}(\gamma) = \int_\gamma \iota_V d\omega \tag{264}$$

for the functional X^{ω}. Lemma 19.5.1 combined with an induction procedure results in the following expression for a general functional $X^{\omega_1\cdots\omega_r}$

$$\begin{aligned} D_V X^{\omega_1\cdots\omega_r}(\gamma) = &\sum_{i=1}^{r}\int_{\gamma}\omega_1\cdots\omega_{i-1}\cdot\iota_V(d\omega_i)\cdot\omega_{i+1}\cdots\omega_r \\ &+\sum_{i=2}^{r}\int_{\gamma}\omega_1\cdots\omega_{i-2}\cdot\iota_V(\omega_{i-1}\wedge\omega_i)\cdot\omega_{i+1}\cdots\omega_r \\ &+\Big(\int_{\gamma}\omega_1\cdots\omega_{r-1}\Big)\cdot\omega_r(V(1)), \end{aligned} \tag{265}$$

which is the equivalent to Eq. (252). For an immersed (Definition 7.2.13) loop

$$\gamma \in \mathcal{LM}_p,$$

we have only to consider variations V, that keep the base point p fixed

$$\mathcal{V}_p \equiv \{V \in \gamma^* TM : \quad V(0) = 0 = V(1)\} \tag{266}$$

Let us mention that any solution Ψ of the equation:

$$\begin{aligned} &D_\gamma \Psi(V) = 0, \\ &\forall V \in \mathcal{V}_p \end{aligned} \tag{267}$$

is called a relative **homotopy invariant** of the loop γ, which has its own interesting properties for instance in Chern-Simons theories or in String Theory.

Returning to our motivation for introducing the Fréchet derivative, we now see that if we consider smooth diffeomorphisms the number of cusps is preserved and we still have an area variation. In the next part, we will be only interested in the subgroup of diffeomorphisms that also preserve angles, i.e. the locally conformal diffeomorphisms. Despite this striking difference between the area derivative and the Fréchet derivative they are still very well related to each other. To make this relation explicit we define an element of $\widetilde{l\mathcal{M}}_p$ by

$$\Theta(\gamma; V) \equiv \int_0^1 \delta_{(\gamma_o^t; V(t)\wedge\dot{\gamma}(t))}(\gamma(t))dt \tag{268}$$

where

$$V \in \mathcal{V}_p$$

and γ_0^t stand for the part of γ, from $\gamma(0)$ to $\gamma(t)$. Using the operator from Definition 268 on the functionals

$$X^u, \; u \in Sh(\Omega)$$

results in

$$\Theta(\gamma; V)(X^{\boldsymbol{u}}) \equiv \int_0^1 \delta_{(\gamma_o^t; V(t)\wedge\dot{\gamma}(t))}(\gamma(t))(X^{\boldsymbol{u}})dt \tag{269}$$

if, of course, this is well defined. Tavares [16] demonstrated that:

$$D_V X^{\boldsymbol{u}}(\gamma) = \int_0^1 \triangle_{(\gamma_o^t; V\wedge\dot{\gamma})}(\gamma(t))X^u dt \tag{270}$$

$$= (\gamma \circ (\Theta(\gamma; V) \otimes 1) \circ \Delta)\ X^u. \tag{271}$$

This result shows that the Fréchet derivative associated with the variational vector field V can be considered as an integral along the path of area derivatives. This means that if one considers area variations induced by the area derivatives as little squares along the path, and integrate over them, we get a smooth area variation. The fact that this is possible is due to the overlapping sides of the little squares which are traversed in opposite direction such that they disappear due to the path reduction property and inverses. This canceling effectively eliminates the cusps introduced by every square such that in the end we have not introduced any new cusp and the result is a smoothly varied contour. Fig. 43 represents this idea graphically.

Naturally we are interested in the application of this result not only to the functionals

$$X^u,\ u \in Sh(\Omega)$$

but also to the holonomy or Wilson loop

$$U : \boldsymbol{L}\mathcal{M}_p \to GL(n, \mathbb{C})$$

of a connection ω. The result for the holonomy can be written as

$$D_V U_\gamma = U_\gamma \cdot \left(\int_0^1 U_{\gamma_0^t} \Omega_{\gamma(t)}(V(t) \wedge \dot{\gamma}(t)) \cdot U_{(\gamma_0^t)^{-1}} \right), \tag{272}$$

a formula that is also known as the **Non-Abelian Stokes theorem**, which for Wilson loop variables becomes (making use of the nuclear or trace class property)

$$\delta_{(\lambda; u\wedge v)} \langle 0 | \mathrm{Tr}\ U_\lambda | 0 \rangle = \langle 0 | \mathrm{Tr}\{U_\lambda \cdot F_{\mu\nu}(\lambda(1))(u^\mu \wedge v^\nu) \cdot U_{\lambda^{-1}}\} | 0 \rangle\,. \tag{273}$$

Similar computations show that:

$$\delta_{(\lambda; u_1\wedge u_2)} U_\lambda = U_\lambda \cdot F(\lambda(1))(u_1 \wedge u_2) \cdot U_{\lambda^{-1}}, \tag{274}$$

The above discussion shows that the Fréchet derivative induces a smooth variation of the (Wilson) loops that can be used to generate equations of motion in GLS (Fig. 44 shows such a variation, the effect on the holonomy and on the spectrum). In the next Part we will focus, as mentioned before, to angle preserving diffeomorphism. Of course we could also consider smooth diffeomorphisms, that still preserve the number of angles, but do not preserve the angle sizes. Investigation of such variations have not been done yet as far as we know, opening the door to extend contour variations to a bigger class of transformations.

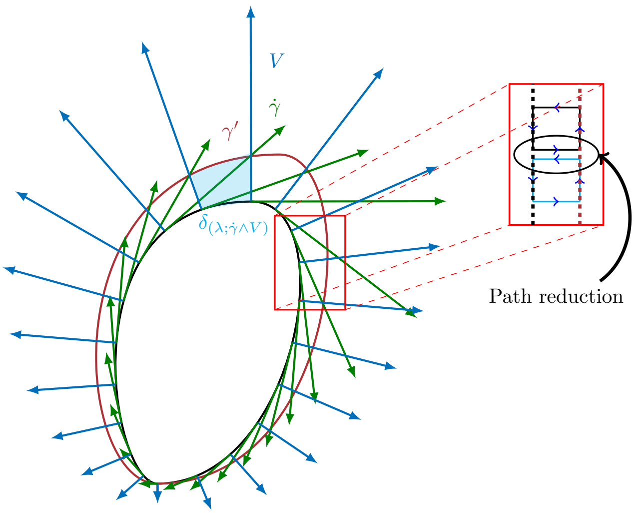


Figure 43: Variation induced by Fréchet derivative.

19.6 FRÉCHET DERIVATIVE AND POLYAKOV DERIVATIVE

We end our exposition with a brief remark on the connection between the Fréchet derivative and the Polyakov derivative. Polyakov proposed an area variation similar to area derivative but with two extra paths attached. What he effectively introduced was the infinitesimal version of the Fréchet derivative

$$U_\gamma \cdot \left(U_{\gamma_0^t} \Omega_{\gamma(t)}(V(t) \wedge \dot{\gamma}(t)) \cdot U_{(\gamma_0^t)^{-1}} \right), \tag{275}$$

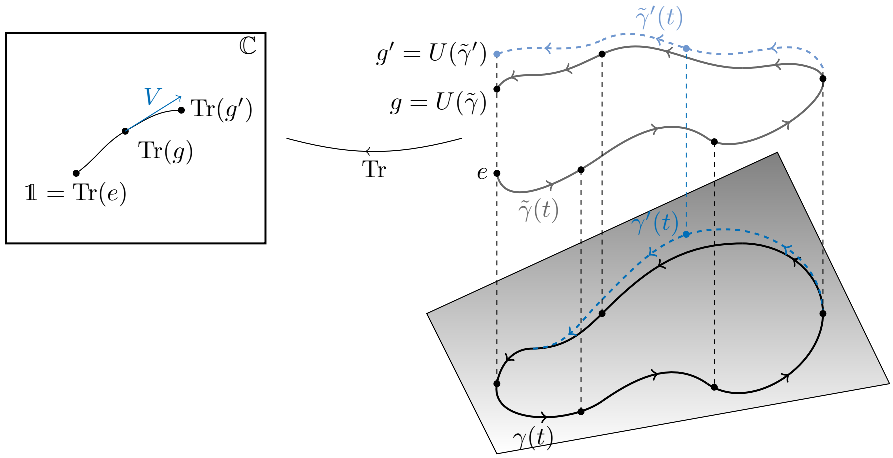


Figure 44: Loop variation and its effects on the holonomy and on the spectrum.

in the sense that his variation is the one induced by the integral kernel of the Fréchet derivative. Interestingly the Fréchet derivative unites two different derivatives that were heavily debated in the past, the area derivative preferred by Makeenko and Migdal contrasted with the Polyakov derivative. Here we clearly show that these are actually aspects of the same derivative, the Fréchet derivative.

19.7 SUMMARY

This Chapter introduced different differential operators that are well-defined on the generalized loop space. We started with the path- and area-derivative, used by Makeenko and Migdal in their loop equations (see next Part), and are well-known to most people in the field. Less known is the Fréchet derivative which, if related to a variational vector field generating a diffeomorphism in generalized loop space, introduces variations of the type that are shown in figure 43. Similar to the area-derivative the Fréchet derivative generates area variations, but of a completely different type as the area-derivative. It is exactly this difference that will allow us to resolve some issues associated with the area derivative, and ultimately will be used to demonstrate that the evolution equation we introduced in [48] has a well-defined mathematical basis.

Part IV

WILSON LOOPS ON THE LIGHT CONE

In the previous part we have defined GLS and introduced differential operators that are well-defined on this space. These operators generate geometric variations of the loops which in this Part will be used to derive an evolution equations. In order to discuss the MM equation we start by reviewing second quantization, path-integrals and introduce the SD equations, which we then use in combination with the differential operators from the previous Part to rederive the MM equations. Unfortunately these equations are only valid for contours without cusps, the contours we are mainly concerned with, which eventually forced us to define a different differential operator on loop space. This operator later turned out to be a special case of the Fréchet derivative, associated to certain diffeomorphisms of the base manifold. Restricting to smooth, angle preserving (conformal) diffeomorphisms, this "new" derivative in combination with the usual renormalization mass derivative allowed us to construct an evolution equation for some specific Wilson Loops. Applying this derivative then to gauge links showing up in PDFs and FF allows to derive evolution equations for TMDs.

20 PATH-INTEGRALS, SECOND QUANTIZATION AND SCHWINGER-DYSON EQUATIONS

20.1 INTRODUCTION

The goal of this section is to introduce the SD equations for pure Yang-Mills and QCD which in the next chapter will be used to derive the MM equations. In order to be able to derive the SD equations we will quickly review some aspects of the **path-integral formalism** where we follow the approach from [128].

20.2 OPERATOR CALCULUS

20.2.1 *Free Scalar Theory : propagator*

In the operator formalism the propagator of the free field

$$G(x-y) = \langle 0 \,|\boldsymbol{T}(\phi(x)\phi(y))|\, 0\rangle \,,$$

with $\boldsymbol{T}$ the time-ordering operator, obeys the **Klein-Gordon** equations

$$\begin{aligned} (-\partial^2 - m^2)\phi(x)\Big|0\Big\rangle &= 0 \\ \Big\langle 0\Big|(-\partial^2 - m^2)\phi(x) &= 0, \end{aligned} \tag{276}$$

together with the **canonical equal-time commutation relations**

$$\begin{aligned} \Big[\chi(t,\boldsymbol{x}),\dot{\phi}(t,\boldsymbol{y})\Big] &= i\delta^{d-1}((\boldsymbol{x})-\boldsymbol{y}) \\ [\chi(t,\boldsymbol{x}),\phi(t,\boldsymbol{y})] &= 0. \end{aligned}$$

such that

$$(-\partial^2 - m^2)G(x-y) = i\delta^d(x-y), \tag{277}$$

where d is the number of space-time dimensions. Fourier transforming Eq. (277) returns

$$G(x-y) = \int \frac{d^n k}{(2\pi)^n} e^{ik(x-y)} \frac{i}{k^2 - m^2 + i\epsilon},$$

the well-known **Feynman propagator** associated to the Klein-Gordon equations. In Eq. (276) we already made use of the the bra and ket notation [130] introduced by Dirac in 1939 (see appendix A.4). Using this bra and ket notation a **linear operator $\boldsymbol{O}$** this can be written as

$$\langle g\,|\boldsymbol{O}|\, f\rangle = \int d^dx \int d^dy\; g(y)\boldsymbol{O}f(x), \tag{278}$$

where we also have that

$$\langle y\,|1|\, x\rangle = \langle y\mid x\rangle = \delta^d(x-y), \tag{279}$$

such that we can write for the operator $\boldsymbol{O}$:

$$\langle y\,|\boldsymbol{O}|\, x\rangle = \boldsymbol{O}\;\delta^d(x-y), \tag{280}$$

where $\boldsymbol{O}$ operates on x. Considering the **Green function** $G(x-y)$ as the **resolvent** of the operator $(-\partial^2 - m^2)$ it can be expressed as the matrix element of the inverse operator, which in Dirac notation becomes

$$G(x-y) = \left\langle y \left| \frac{i}{\partial^2 - m^2} \right| x \right\rangle. \tag{281}$$

Wick rotating Eq. (277) by introducing $t = -ix_4$ results in the Euclidean version

$$(-\partial^2 + m^2)G(x-y) = \delta^d(x-y). \tag{282}$$

where the Green Function solution now is

$$G(x-y) = \int \frac{d^dk}{(2\pi)^d} e^{ik(x-y)} \frac{1}{k^2+m^2}. \tag{283}$$

Using a Wick rotation we are able to go to a Euclidean formulation of the QFT under consideration, where the path-integrals are well-defined. Note however that from a non-perturbative point of view there is no reason why the Minkowskian and Euclidean formulation should always be equivalent.

Before moving on to a discussion on the ordering of operators we have a quick look at the **Dirac γ-matrices** in the Euclidean formulation. In Minkowski space we have that the anti-commutator of the γ-matrices returns

$$\{\gamma_M^\mu, \gamma_M^\nu\} = 2g^{\mu\nu} I,$$

making γ^0 Hermitian and the spacial γ-matrices Anti-Hermitian. In Euclidean space on the other hand we have:

$$\{\gamma_E^\mu, \gamma_E^\nu\} = 2\delta^{\mu\nu} I,$$

now making all the γ-matrices Hermitian. Given the explicit representation:

$$\begin{aligned} \gamma_4 &= \gamma_0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \\ \gamma_j &= \begin{pmatrix} 0 & -i\sigma_j \\ i\sigma_j & 0 \end{pmatrix}, \end{aligned} \tag{284}$$

of the γ-matrices the Euclidean version of the Dirac equation becomes

$$\left(\partial\!\!\!/ + m\right)\psi = 0,$$

or in momentum space:

$$\left(p\!\!\!/ + m\right)\psi = 0,$$

with

$$p = \left(-i\partial_x, -i\partial_4\right).$$

20.2.2 *Path-ordering and the Path-integral*

Formally Eq. (282) can be expressed as

$$\boldsymbol{O}G(x-y) = \delta^d(x-y),$$

inverting the operator then results in

$$\begin{aligned} G(x-y) &= \boldsymbol{O}^{-1}\delta^d(x-y) \\ &= \frac{1}{-\partial^2+m^2}\delta^d(x-y) \\ &= \frac{1}{2}\int_0^\infty dt\; e^{\frac{1}{2}t(\partial^2-m^2)}\delta^d(x-y) \\ &= \frac{1}{2}\int_0^\infty dt e^{-\frac{1}{2}m^2 t}\mathcal{P}e^{\frac{1}{2}\int_0^t ds\partial^2(s)}\delta^d(x-y) \\ &= \frac{1}{2}\int_0^\infty dt e^{-\frac{1}{2}m^2 t}\int\limits_{z_\mu(0)=x_\mu} Dz_\mu(t)e^{-\frac{1}{2}\int_0^t ds\dot{z}_\mu^2(s)}\cdot\mathcal{P}e^{\int_0^t ds\dot{z}_\mu\partial_\mu(s)}\delta^d(x-y) \end{aligned} \tag{285}$$

where $\mathcal{P}$ now defines the path ordering of the operators along the path finding its origin in the fact that $e^{A+B} \neq e^A e^B$ when $[A,B] \neq 0$ leading to

$$e^{A+B} = \mathcal{P}e^{\int\limits_0^1 d\sigma' A(\sigma')}e^{\int\limits_0^1 d\sigma B(\sigma)} = e^A + \int\limits_0^1 d\sigma e^{(1-\sigma)A}Be^{\sigma A} + \cdots$$

(see [131] for the details). In the last line[1] of Eq. (285) we integrate over all the trajectories starting at x and where

$$\partial_\mu(t)\dot{z}_\nu = \frac{d}{dt}\delta_{\mu\nu} = 0.$$

We now recognize this integral over $z_\mu(t)$ as the path integral. Further simplification comes from the fact that

$$\mathcal{P}e^{\int_0^t ds\dot{z}^\mu(t)}$$

is a shift-operator

$$\begin{aligned} \mathcal{P}e^{\int_0^t ds\dot{z}^\mu(t)\partial_\mu(t)} &= e^{(z^\mu(t)-x^\mu)\frac{\partial}{\partial x^\mu}} \\ \mathcal{P}e^{\int_0^t ds\dot{z}^\mu(t)\partial_\mu(t)}\delta^d(x-y) &= \delta^d(z(t)-y). \end{aligned} \tag{286}$$

Hence for the propagator expressed with the **path-integral** we finally get

$$G(x-y) = \frac{1}{2}\int_0^\infty dt\ e^{-\frac{1}{2}tm^2}\int_{z_\mu(0)=x_\mu}^{z_\mu(t)=y_\mu} Dz_\mu(t)\ e^{-\frac{1}{2}\int_0^t ds\dot{z}_\mu^2(s)}. \tag{287}$$

20.2.3 *Re-parametrization invariance*

This is a good place to make a side note on re-parametrization invariance, since we have this property in the GLS. Note that Eq. (287) is NOT invariant under re-parametrization, to take this into account one also needs to integrate over the **re-parametrization group** which leads to a **group-volume factor**, due to the parametric invariance of the exponent. In general considering vacuum expectation values one divides out the vacuum bubble diagrams i.e.

$$\langle 0\,|\boldsymbol{O}|\,0\rangle := \frac{\left\langle 0\left|\int \mathcal{D}\Phi\ \boldsymbol{O}e^{-S}\right|0\right\rangle}{\langle 0\,|\int \mathcal{D}\Phi\ e^{-S}|\,0\rangle},$$

such that this volume factor cancels in the numerator and denominator.

1 The equivalence between the last line and the previous can be shown by filling in $z_\mu(t) \to z'_\mu(t) = z_\mu(t) + \int_0^t dt'\ \partial_\mu(t')$ in the Gaussian integral of the last line.

20.2.4 *Propagator in external field and parallel transporter*

Introducing an external field is the same as introducing a field $A^\mu(x)$ such that we need to replace the partial derivative by the covariant one, i.e.

$$\partial_\mu \to D_\mu = \partial_\mu - ieA_\mu.$$

This changes expression (285) to:

$$\begin{aligned} G(x,y;A) &\equiv \left\langle y \left| \frac{1}{-D_\mu^2 + m^2} \right| x \right\rangle \\ &= \frac{1}{2}\int_0^\infty dt\ e^{-\frac{1}{2}tm^2} \int_{z(0)=x}^{z(t)=y} Dz(t)\ e^{-\frac{1}{2}\int_0^t dt\ \dot{z}^2(t) + ie\int_0^t \dot{z}^\mu(t)A_\mu(z(t))} \\ &= \sum_{\Gamma_{yx}}{}' e^{ie\int_{\Gamma_{yx}} dz^\mu\ A_\mu(z)}, \end{aligned} \tag{288}$$

where

$$\sum_{\Gamma_{yx}}{}' = \sum_{\Gamma_{xy}} e^{-S_{\text{free}}[\Gamma_{yx}]}.$$

This is indeed a very interesting result, because it says that the **transition amplitude (i.e. propagator) of a quantum particle** in a classical electromagnetic field is a sum over paths of the Abelian phase factor

$$U[\Gamma_{ys}] = e^{ie\int_{\Gamma_{yx}} dz^\mu A_\mu(z)}. \tag{289}$$

As we have seen before, this is the parallel transporter associated to the gauge potentials $A_\mu(z)$, here the electro-magnetic (gauge) field. This shows that the phase factor depends on the specific path taken by the particle unless the field is a pure gauge field, or put otherwise, the field strength $F_{\mu\nu}(z)$ vanishes along the possible paths. Due to the simply connectedness (Section 6.4) of space-time this is usually the case, but sometimes some regions of space are unavailable for paths such that the path dependence of the phase-factor can have a physical effect. An example of such an effect is the Aharanov-Bohm effect, which we will not discuss further here.

Having quickly reviewed the relation between path-integrals, propagators and operator we turn our attention in the next Section to Second Quantization.

20.3 SECOND QUANTIZATION

20.3.1 *Introduction*

So far we have only considered first quantization where the coordinates and momentum were operators. Formulating this in a path-integral formalism, **first quantization** is related with **integrals over trajectories in coordinate space**. In **second quantization** the fields become the operators. Naturally one then wonders what is the associated integration in the path-integral formalism. The next Section shows that the **path-integral** will now integrate over **field configurations** instead of over trajectories in coordinate space.

20.3.2 *Partition function with fields*

Inspired by statistical mechanics let us introduce the partition function, defined for scalar fields, in the following way:

$$Z = \int \mathcal{D}\phi(x) e^{-S}, \tag{290}$$

with S the free action for the field ϕ

$$S_{\text{free}}[\phi] = \frac{1}{2}\int d^d x \, \left((\partial_\mu \phi)^2 + m^2\phi^2\right).$$

The measure of the integral in Eq. (290) can be defined as

$$\int \mathcal{D}\phi(x) \cdots = \prod_x \int_{-\infty}^{\infty} d\phi(x) \cdots$$

with x running all over space and the measure $d\phi$ is the Lebesgue measure [132]. Using the above definitions the Green function $G(x,y)$ is given by

$$G(x,y) = \langle \phi(x)\phi(y) \rangle \,, \tag{291}$$

and a general average is defined by[2]:

$$\langle F[\phi] \rangle = Z^{-1} \int \mathcal{D}\phi(x) e^{-S[\phi]} F[\phi]. \tag{292}$$

2 Note that here the square brackets are a sign of functional dependence!

From the Gaussian structure of the free action we have

$$G(x-y) = \left\langle y \left| \frac{1}{-\partial^2 + m^2} \right| x \right\rangle . \tag{293}$$

Remark that **Perturbation Theory** is well defined on Minkowski-space because of the definition of the Gaussian path integral

$$\int \mathcal{D}\phi e^{iS} = \int \mathcal{D}\phi e^{i \int d^dx \ \phi \boldsymbol{D} \phi}, \tag{294}$$

such that the identity is valid

$$\langle \phi_x \phi_y \rangle = \left\langle y \left| \frac{1}{i\boldsymbol{D}} \right| x \right\rangle . \tag{295}$$

We again point out that this does not need to be true in the non-perturbative region which is due to the presence of the complex weighting factor, possibly making the integrals divergent.

Until now we have only considered **scalar/bosonic fields**, we now wish to extend the treatment to **fermionic fields**. To this end we introduce **Grassmann variables** obeying the anti-commutation relations:

$$\{\psi_y, \psi_x\} = 0, \qquad \{\bar{\psi}_y, \bar{\psi}_x\} = 0, \qquad \{\bar{\psi}_y, \psi_x\} = 0,$$

from which it follows that these variables are **nilpotent** of order two

$$\psi_x^2 = \bar{\psi}_x^2 = 0. \tag{296}$$

This leads to the following path-integrals for fermi fields

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\int d^dx \bar{\psi} \boldsymbol{D} \psi} = \text{det } \boldsymbol{D}, \tag{297}$$

and for complex Bose fields:

$$\int \mathcal{D}\phi^\dagger \mathcal{D}\phi e^{-\int d^dx \phi^\dagger \boldsymbol{D} \phi} = (\text{det } \boldsymbol{D})^{-1}. \tag{298}$$

With the above we have that:

$$\left\langle \psi(x)\bar{\psi}(y) \right\rangle = \left\langle y \left| \boldsymbol{D}^{-1} \right| x \right\rangle . \tag{299}$$

using

$$\text{det } \boldsymbol{D} = e^{\text{Sp } \ln \boldsymbol{D}}$$

for fermions and similar for bosons we have

$$(\text{det } \boldsymbol{D})^{-1} = e^{-\text{Sp } \ln \boldsymbol{D}}$$

generating the minus sign that shows up in the logarithm of the partition function when considering closed fermion loops.

20.3.3 *Schwinger-Dyson equations*

The SD equations can be interpreted as the quantum version of the classical equations of motion, that generate the Feynman diagrams by iterated application. To derive the SD equations we consider the following field variation

$$\phi(x) \to \phi(x) + \delta\phi(x)$$

which is allowed because the path-integral measure is invariant under such variations. This immediately implies that the change in the path-integral should be zero, hence :

$$\int d^d x \delta\phi(x) \int \mathcal{D}\phi \; e^{-S[\phi]} \left[-\frac{\delta S[\phi]}{\delta\phi(x)} F[\phi] + \frac{\delta F[\phi]}{\delta\phi(x)} \right] = 0,$$

where $F[\phi]$ is for now an arbitrary functional. It immediately follows, due to the arbitrariness of $\delta\phi(x)$, that

$$\frac{\delta S[\phi]}{\delta\phi(x)} = \hbar \frac{\delta}{\delta\phi(x)},$$

where the explicit dependence on $\hbar$ is written. This last equation is said to be valid in the **weak sense** which means it is valid under the assumption that one takes averages of both sides of the equation leading to the functional equality

$$\left\langle \frac{\delta S[\phi]}{\delta\phi(x)} F[\phi] \right\rangle = \hbar \left\langle \frac{\delta F[\phi]}{\delta\phi(x)} \right\rangle . \tag{300}$$

As a simple example let us apply this to the free scalar field with $F[\phi] = \phi(y)$. In this case Eq. (300) becomes

$$(-\partial^2 + m^2) \langle \phi(x)\phi(y) \rangle = \hbar \left\langle \frac{\delta\phi(y)}{\delta\phi(x)} \right\rangle = \hbar\delta^n(x-y), \tag{301}$$

where the left hand side originates from

$$\frac{\delta S_{\text{free}}}{\delta\phi(x)} = (-\partial^2 + m^2)\phi(x).$$

The derivative on the right hand side serves as the conjugate momentum in the operator formalism.

Remark 20.3.1.
It is important to notice that taking this variational derivative in Euclidean space is the same as differentiating the T-product and applying the commutation relations in Minkowski space.

$$\langle F[\phi]\rangle = Z^{-1}\int \mathcal{D}\phi(x)e^{-S[\phi]}F[\phi] \leftrightarrow \langle 0\,|\boldsymbol{T}F[\phi]|\,0\rangle$$

In the **classical limit** $\hbar \to 0$ Eq. (301) reduces to

$$\begin{aligned} (-\partial^2 + m^2)\langle\phi(x)\phi(y)\rangle &= 0 \\ \frac{\delta S[\phi]}{\delta\phi} &= 0, \end{aligned} \tag{302}$$

the classical equation of motion for the field ϕ. To allow the introduction of a graphical representation of the SD equations, consider now the scalar interacting theory with action

$$S[\phi] = \int d^4x\left(\frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{3!}\phi^3\right),$$

Applying the SD approach with $F[\phi] = \phi(y)$ results in

$$(-\partial^2 + m^2)\underbrace{\langle\phi(x)\phi(y)\rangle}_{\text{propagator}} + \frac{\lambda}{2}\underbrace{\left\langle\phi^2(x)\phi(y)\right\rangle}_{\text{vertex}} = \hbar\delta^n(x-y). \tag{303}$$

Let us first consider the **vertex term**. Introducing the **3-point Fourier-transformed Green functions**

$$G_3(p,q,-p-q) = \int d^4x\; d^4y e^{-ipx-iqy}\langle\phi(x)\phi(y)\phi(0)\rangle\,, \tag{304}$$

and the notation

$$G_0(p) = \frac{1}{p^2+m^2}, \tag{305}$$

for the free propagator. Expanding Eq. (304) and writing only the lowest order term we have

$$G_3(p,q,-p-q) = -\lambda G_0(p)G_0(q)G_0(p+q) + \cdots\,. \tag{306}$$

From this expansion we derive the truncated part, i.e. we truncate the external legs, also known as the vertex function which is given by

$$\begin{aligned}\Gamma(p,q,-p-q) &= G_3(p,q,-p-q)G^{-1}(p)G^{-1}(q)G^{-1}(p+q) \\ &= -\lambda-\lambda^3\int\frac{d^nk}{(2\pi)^n}G_0(k-p)G_0(k)G_0(k+q)+\cdots,\end{aligned} \tag{307}$$

where we wrote the first terms of the perturbative expansion in the second line. A graphical representation of this expansion is shown in figure 45, where the bold dot represents the exact vertex. Applying a similar reasoning to

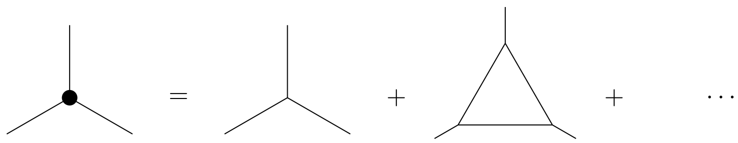


Figure 45: DS Vertex expansion

the Fourier-transformed of the propagator term leads to an expansion that is diagrammatically represented in figure 46 where now the bold line is the exact propagator. Combining the expressions for the exact vertex and propagator

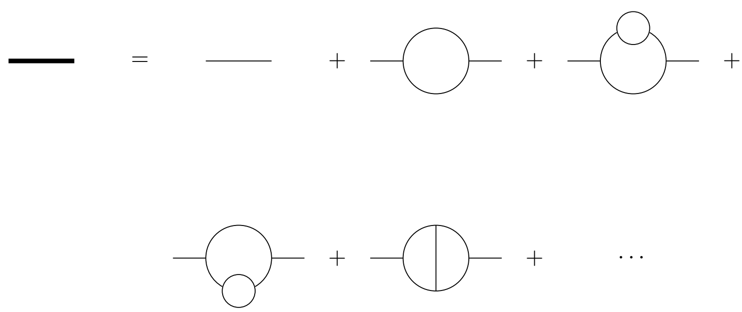


Figure 46: DS Propagator expansion

results in the graphical equation shown in figure 47, where the thick lines represent exact propagators and the dot on the right vertex is exact. The thin lines are the bare propagator $G_0(p)$. Analytically we have:

$$G_0(p)-G^{-1}(p)=-\frac{\lambda}{2}\int\frac{d^nq}{(2\pi)^n}G(q)\Gamma(-q,p,q-p)G(p-q). \tag{308}$$

If we multiply this with $G(p)$ and use the above definitions we get the Fourier transform of equation Eq. (303). It should now be clear that this equation is not closed, it relates the 2-point function to the 3-point function. Choosing now $F[\phi] = \phi(x_2)\cdots\phi(x_n)$ and apply the SD equations this results in

$$(-\partial^2 + m^2)G_n(x, x_2, \cdots, x_n) + \frac{\lambda}{2}G_{n+1}(x, x, x_2, \cdots, x_n) = \sum_{j=2}^{n} \delta^n(x - x_j)G_{n-2}(x_2, \cdots, \hat{x}_j, \cdots, x_n), \quad (309)$$

where the hat of $\hat{x}_j$ means that this parameter is omitted.

$$G_0(p) - G_0(p)G^{-1}(p)G_0(p) =$$

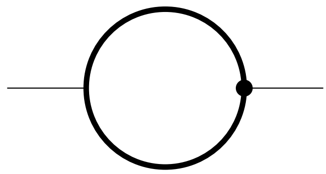


Figure 47: Graphical DS representation

20.3.4 *QCD : Schwinger-Dyson Equations*

20.3.4.1 *Pure Yang-Mills Theory*

The free QCD (or pure Yang-Mills) Lagrangian, without fermions, can be rewritten as [19]

$$\begin{aligned}
\mathcal{L}^{\text{QCD}}_{\text{free}} &= -\frac{1}{2}Tr\left[\left(F^a_{\mu\nu}\frac{\lambda^a}{2}\right)^2\right] \\
&= -\frac{1}{2}\left(F^a_{\mu\nu}\right)^2 Tr\left[\frac{\lambda^a}{2}\frac{\lambda^a}{2}\right] \\
&= -\frac{1}{2}\left(F^a_{\mu\nu}\right)^2\frac{1}{2} \\
&= -\frac{1}{4}\left(F^a_{\mu\nu}\right)^2
\end{aligned}$$

where we have neglected the **ghosts and gauge-fixing terms** since they cancel anyway (see [128] for a proof of this statement). This expression allows an easier variation of the Lagrangian with respect to the fields.

20.3.4.2 *Variation of the gauge field - SD equations*

The goal is to derive the SD equations for QCD. To this end we consider a variation of the gauge field, also inducing a variation in the derivative of the gauge field

$$\begin{aligned} A_\mu &\rightarrow A_\mu + \delta A_\mu \\ \partial_\nu A_\mu &\rightarrow \partial_\nu A_\mu + \delta \partial_\nu A_\mu . \end{aligned} \tag{310}$$

With these variations we can determine

$$\frac{\delta S}{\delta A_\mu},$$

derive the **Euler-Lagrange equations** for the free QCD Lagrangian and the SD equations. The detailed calculations can be found in appendix B.5 where we find for the Euler-Lagrange equations

$$\partial_\nu F^{\mu\nu,i} + g\ f^{ijk}\ F^{\mu\nu,k} A^j_\nu = 0, \tag{311}$$

where g is the coupling constant, f^{ijk} are the structure constants, A^j_ν the gauge potentials and $F^{\mu\nu,i}$ the "colored" (index i) field strength tensors (or curvatures of the gauge potentials). The SD equations (where the equalities are weak[3]) are given by

$$\left(\partial_\nu F^{\mu\nu,i} + g f^{ijk} F^{\mu\nu,k} A^j_\nu\right) \boldsymbol{F}[A] = D^{ab}_\nu F^{\mu\nu,b}(x) \boldsymbol{F}[A] = \hbar \frac{\delta}{\delta A^a_\nu(x)} \boldsymbol{F}[A], \tag{312}$$

where $\boldsymbol{F}[A]$ is the functional depending on the gauge fields and D^{ab}_μ represents the **covariant derivative in the adjoint representation.** Important to note here is that we have excluded the interaction with fermions, the motivation for the moment being that if we consider only the lowest order interactions on a quark Wilson loop we have **no fermion couplings.** The fermion couplings only emerge here at the second order where a gluon splits in a quark pair (i.e. gluon propagator correction). If one introduces the coupling with fermions Eqs. (510) and (511) get an extra contribution

$$g \frac{\lambda^a}{2} \bar{\psi} \gamma^\mu \psi,$$

which needs to be taken into account in the higher order contributions. So at lowest order, or suppressing fermion interactions, we can write formally

$$\left\langle D^{ab}_\nu F^{\mu\nu,b}(x) \boldsymbol{F}[A] \right\rangle = \hbar \left\langle \frac{\delta}{\delta A^a_\nu(x)} \boldsymbol{F}[A] \right\rangle . \tag{313}$$

3 They are only valid under taking averages.

20.4 SUMMARY

In this Chapter we have introduced the path-integral formalism at first quantization level after reviewing some of the definitions and notations from the operator formalism. With this formalism we had a short look at propagators and more specifically studied the propagator in an external field where the parallel transporter emerged as a phase-factor. We then moved on to the second quantization where the path-integral formalism turned out to be useful in deriving a set of quantum equations known as the Schwinger-Dyson equations. At the end of the chapter we applied these equations to pure Yang-Mills theory to derive Eq. (511) that we will use in the next chapter to derive the MM equations.

21 MAKEENKO-MIGDAL LOOP EQUATIONS

21.1 INTRODUCTION

After a short discussion on Fierz identies and an introduction to multi-color QCD we will use the SD equations derived in the previous chapter to derive the Makeenko-Migdal loop equations.

21.2 GENERALIZED FIERZ IDENTITIES

In the next section we will need to make use of a Fierz-type of identity to derive the MM equations. Here we will go through some of derivations demonstrated in [133] to derive general Fierz identities.

We start from an N-dimensional vector space V with orthonormal basis $\{e_i\}$ and inner product $(\cdot,\cdot)$ such that

$$(e_i, e_j) = \delta_{ij}.$$

In the canonical representation of these basis vectors we have that

$$(e_i)_j = \delta_{ij}.$$

Orthogonality can thus be represented as

$$e_i^T e_j = \delta_{ij}$$

and completeness by

$$\sum_i^N e_{i=1} e_i^T = \mathbb{1},$$

which is invariant under $O(N)$ respectively $U(N)$ transformations of the basis vectors. It should now be clear that the set of real or complex $N \times N$ matrices form a N^2-dimensional vector space with canonical basis $\{e^{ij}\}$ given by

$$(e^{ij})_{kl} = \delta^i_k \delta^j_l, \tag{314}$$

such that any matrix M can be written as

$$M = M_{ij} e^{ij}. \tag{315}$$

One can now define the following **inner product** on this vector space

$$(A, B) \equiv \text{Tr } (AB^T) \tag{316}$$

for which the canonical basis satisfies:

$$\text{Tr } [e^{ij} e^{kl^T}] = \delta^{ik} \delta^{jl}. \tag{317}$$

Note that in the complex case the transposition operation $\cdot^T$ needs to be replace by the hermitian conjugation operation $\cdot^\dagger$. If we drop the transposition (hermitian conjugate) in Eq. (316) we have instead defined a **bilinear function** on the $N \times N$ matrices. Using this bilinear we can define a dual basis

$$\{e_{ij}\} : e_{ij} \equiv e^{ji} = e^{ij^T},$$

in other words this **bilinear function can be interpreted as a metric** that can be used to raise or lower indices or to move between basis and dual basis

$$M_{ij} = \text{Tr } [M e_{ij}].$$

Inserting this in Eq. (315) this results in

$$(e_{ij})_{kl} (e^{ij})_{nm} = [(e^{ij})^T]_{kl} \delta^i_n \delta^j_m = \delta^i_l \delta^j_k \delta^i_n \delta^j_m = \delta_{km} \delta_{nl}. \tag{318}$$

This represents an identity in the space of general linear transformations over $M_N(\mathbb{R})$. A transformation in the general linear transformations on $M_N(\mathbb{R})$ can be written as

$$M \to (A \otimes B)M \equiv AMB^T = (A)_{ij} (B)_{lk} (M)_{jk} e^{il}.$$

Combining

$$(e_{ij} \otimes e^{ij})M = (e^{ij} \otimes e_{ij})M = M,$$

with Eq. (318) will generate **Fierz type identities.** We will focus here on $SU(2)$ and $SU(3)$ algebras, represented respectively by the Pauli and Gell-Mann matrices which form vector spaces with orthogonality relations[1]

$$\begin{aligned} \text{Tr } [\sigma_i \sigma_j] &= 2\delta_{ij}, \qquad (i, j \in [0, 3] \cap \mathbb{N}) \\ \text{Tr } [\lambda_a \lambda_b] &= 2\delta_{ab}, \qquad (a, b \in [0, 8] \cap \mathbb{N}). \end{aligned} \tag{319}$$

1 They are already orthogonal and hermitian

As we know from matrix theory the three Pauli matrices augmented with the unit matrix form a basis for the vector space $M_2(\mathbb{C})$ allowing us to write

$$M = M_0 \mathbb{1} + M_i \sigma^i, \qquad \sigma^i = \sigma_i, \tag{320}$$

where

$$M_0 = \frac{1}{2}\text{Tr }(M) \qquad \text{and} \qquad M_i = \frac{1}{2}\text{Tr }(M\sigma_i).$$

Expressed with individual elements this becomes

$$M_{ij} = \frac{1}{2} M_{kk}\delta_{ij} + \frac{1}{2} M_{lk}(\sigma_m)_{kl}(\sigma_m)_{ij}, \tag{321}$$

from which we read off the coefficients for the M_{lk} element

$$\delta_{il}\delta_{jk} = \frac{1}{2}\delta_{lk}\delta_{ij} + \frac{1}{2}(\sigma_m)_{kl}(\sigma_m)_{ij}.$$

Note that for Weyl spinors[2] this becomes

$$(\sigma_\mu)_{ij}(\tilde{\sigma}^\mu)_{kl} = 2\delta_{il}\delta_{kj},$$

where

$$\sigma^\mu = (\mathbb{1}, \boldsymbol{\sigma}) \qquad \text{and} \qquad \tilde{\sigma}^\mu = (\mathbb{1}, -\boldsymbol{\sigma}).$$

The equivalent expression of Eqs. (321) for $SU(3)$ becomes

$$\delta_{il}\delta_{jk} = \frac{1}{3}\delta_{lk}\delta_{ij} + \frac{1}{2}(\lambda_m)_{kl}(\lambda_m)_{ij}. \tag{322}$$

Generalizing these equations, with completeness relations

$$\text{Tr }[T_a T_b] = C\delta_{ab}$$

we finally get the general Fierz identities for $SU(N)$

$$\delta_{il}\delta_{jk} = \frac{1}{N}\delta_{lk}\delta_{ij} + \frac{1}{C}(T_a)_{kl}(T_a)_{ij}. \tag{323}$$

This last equation will be used to derive the MM equations in Section 21.4.

2 Using Minkowski to raise and lower indices!!

21.3 LARGE N_c EXPANSION : MULTI-COLOR QCD

21.3.1 *Introduction*

In 1974 Gerard t'Hooft [134, 135] found a new perspective on QCD, instead of expanding the theory in the coupling constant to do perturbation theory **he expanded the theory in terms of the number of colors** N_c. This effectively rearranges the Feynman diagrams according to their **topology** where the Leading Order (LO) of this expansion is topologically a **sphere** and is given by planar diagrams, simplifying calculations as will be explained below. Note that this topological expansion is equivalent to the expansion in the **string coupling in string models for the strong interaction**. Actually, when expanding in the number of colors, one can take **two different limits** for $N_c \to \infty$, the first is referred to as **t'Hooft limit** where the number of quark flavors N_f is kept constant [134] (and will lead to an effective expansion in $\frac{1}{N_c^2}$ if no internal quark lines are considered) and the second is known as the **Veneziano limit** [136] where now $\frac{N_f}{N_c}$ is kept fixed. We will mainly focus on the t'Hooft limit, and although this limit introduces **infinite dimensional matrices** the theory is simplified, not only due to the **reduction of number graphs in the leading order** contribution[3] but also due to the fact that correlators of **gauge invariant operators factorize.** The purpose of this section is to give a motivation for the study of Wilson loops and how they are related to QCD, as such we will not go through the full details but give more or less an overview of how things work where I mainly follow Makeenko's lectures [128] on this subject.

21.3.2 *Matrix-Field representation : introducing Index-Ribbon graphs*

Most readers are for sure familiar with this type of graphs but we will reintroduce them here for completeness and we will use them to explain some of the statements we made in the introduction. We start from the **matrix-field representation of the gauge fields**

$$\left[A_\mu^{ij}(x)\right] = \sum_a A_\mu^a(x)\,[t^a]^{ij}\,, \tag{324}$$

3 The number of planar graphs grows exponentially with the number of vertices instead of factorially.

where the matrix in the LHS is Hermitian and differs from the one t'Hooft uses

$$\mathcal{A}_\mu^{ij}(x) = igA_\mu^{ij}(x).$$

Using this notation and applying the normalization

$$\text{Tr}\left[t^a t^b\right] = \frac{1}{2}\delta^{ab}$$

(i.e. C from Eq. (323) is $\frac{1}{2}$) we have for the **matrix-field propagator**

$$\left\langle A_\mu^{ij}(x) A_\mu^{kl}(y) \right\rangle = \frac{1}{2}\left(\delta^{il}\delta^{kj} - \frac{1}{N_c}\delta^{ij}\delta^{kl}\right) D_{\mu\nu}(x-y), \tag{325}$$

with in **Feynman gauge** the propagator parametrized as

$$D_{\mu\nu}(x-y) = \frac{1}{4\pi^2}\frac{g_{\mu\nu}}{(x-y)^2}.$$

In the limit $N_c \to \infty$ the second term vanishes reducing the matrix propa-

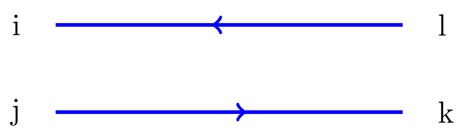


Figure 48: Matrix-field propagator

gator to

$$\left\langle A_\mu^{ij}(x) A_\mu^{kl}(y) \right\rangle \sim \delta^{il}\delta^{kj}, \tag{326}$$

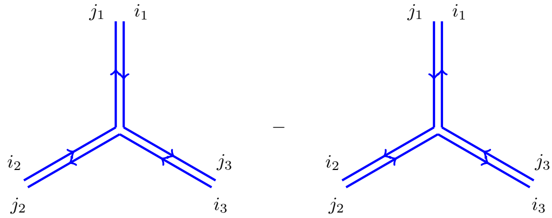
which can graphically be represented by the diagram in Figure 48. Notice the relation between the direction of the arrows and the order of the indices on the delta functions in Eq. (326), which is due to the fact that A_μ^{ij} is Hermitian and the off-diagonal elements are each others complex conjugate so that we can choose as independent fields

$$A_\mu^{ij} \in \mathbb{C}, i > j \qquad \text{and} \qquad A_\mu^{ii} \in \mathbb{R}$$

with

$$A_\mu^{ji} = \left(A_\mu^{ij}\right)^*.$$

The three gluon vertex in the matrix-field representation is shown diagrammatically in figure 49 and similar drawings can be made for the four-gluon vertex (but are not shown here). Keeping in mind that a three-gluon vertex



$$\sim g\left(\delta^{i_1 j_3}\delta^{i_3 j_2}\delta^{i_2 j_1} - \delta^{i_3 j_1}\delta^{i_1 j_2}\delta^{i_2 j_3}\right)$$

Figure 49: Ribbon graphs for three-gluon vertex

$\sim g$, a four-gluon vertex $\sim g^2$ and that when we have a **closed index loop in the ribbon graph** we get a factor N_c due to the sum over the indices, we can now analyze some graphs to demonstrate the advantage of using the large N_c expansion. Consider as a first example the one-loop **correction to**

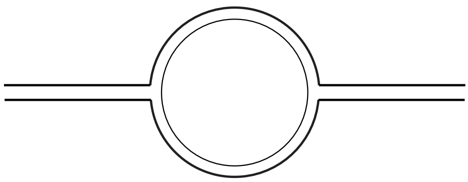


Figure 50: 1-loop correction matrix-field propagator

the matrix-field propagator, graphically represented in figure 50. In this diagram we have two three-gluon vertices resulting in a g^2 factor and one index loop contributing with a factor N_c, such that the diagram is proportional to

$$g^2 N_c.$$

If we want this diagram to have a non-trivial contribution in the large N_c limit it follows that

$$g^2 \sim \frac{1}{N_c},$$

where g is now **the bare coupling constant.** In a pure $SU(N_c)$ theory g^2 is given by the asymptotic-freedom formula (also known as the QCD β-function)

$$g^2 = \frac{24\pi^2}{11 N_c \ln\left(\frac{\Lambda}{\Lambda_{\text{QCD}}}\right)}, \tag{327}$$

where Λ is a UV cut-off parameter and Λ_{QCD} is the energy scale parameter from QCD renormalization. Eq. (327) is indeed consistent with what we wanted for the one-loop diagram in figure 50 such that this diagram is of order $\mathcal{O}(1)$ and thus has a non-trivial contribution in the $N_c \to \infty$ limit. To further demonstrate the advantage of the large N_c expansion we will now compare a planar and a non-planar diagram. The diagram in figure 51 is an

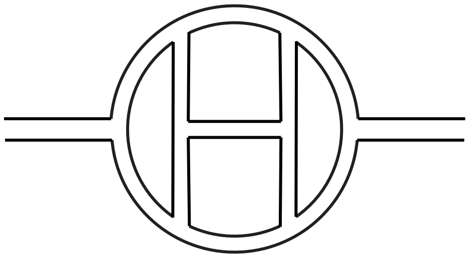


Figure 51: Planar diagram

example of a more complicated planar graph of which we would like to know the order of in the $\frac{1}{N_c}$ expansion. We now have eight three-gluon vertices and four closed index loops resulting in a total contribution proportional to

$$(g^2 N_c)^4 \sim 1,$$

showing that this graph contributes to the LO. If one now wants to add a loop and keep the diagram planar it is clear that one will also need to add two three-gluon vertices such that for a planar graph with n loops its order of contribution will go like $(g^2 N_c)^n \sim 1$ showing that **planar graphs all add at LO**. Consider now the diagram shown in figure 52 as an example of a

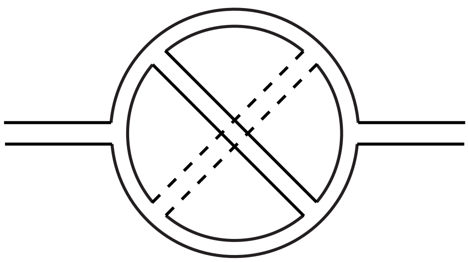


Figure 52: Non-planar diagram

non-planar diagram, where the dashed lines represent lines that go out of the plane. This diagram has now six three-gluon vertices and one closed index loop[4], which give a contribution order of $g^6 N_c \sim \frac{1}{N_c^2}$, a NLO contribution in the t'Hooft limit **if we discard virtual quark loops**, which will be introduced below. Also notice that restricting to the LO would give you an accuracy that is expected to be about the same as the meson widths, i.e. about 10-15 percent [128].

Before introducing the virtual quark loops to the game first a few comments. A first remark is on the topology of the graphs. **The non-planar diagram from figure 52 can be drawn without crossing, not on a plane or a sphere, but on a sphere with a handle or torus.** In general we have that a genus n diagram has a contribution order of

$$\left(\frac{1}{N_c^2}\right)^n .$$

This should make it more clear that moving from an expansion in the coupling constant to an expansion in $\frac{1}{N_c}$ reshuffles the Feynman diagrams according to their topology, this is the reason t'Hooft's expansion is sometimes referred to as a **topological or genus expansion.** This topology can be shown to be equivalent to the topology of the quantized dual string with quarks at the end [134]. A second, and final remark, is on the different contributions of the three-gluon vertex (Figure 49). If one **closes the external legs of the diagrams** (see figure 53) in a cyclic way, which **adds a trace of Non-Abelian phase factors**, it is clear that the two diagrams contribute at a different order in the t'Hooft expansion[5]. The **cyclically closed diagram** on the left in figure 53 can be drawn as a **planar graph** while the cyclically closed diagram on the right hand side will be non-planar, but again can be drawn without crossing on a torus. As discussed before this shows that they contribute at different orders. Let us now continue by allowing the **virtual quark loops** in the diagrams. The quark propagator

$$\left\langle \psi_i \bar{\psi}_j \right\rangle \sim \delta_{ij},$$

is graphically be represented by a single line in the Figures. We can now determine the order of graphs containing virtual quark loops. As an example

4 To see this just choose a starting point on one of the internal lines and follow the line, you will notice that after going through both the lines in the plane and out of the plane you will return to your starting point, so that there is indeed one closed index loop.

5 Physically this corresponds to consider the external legs to be attached to three valence quarks, represented by single lines in figure 53.

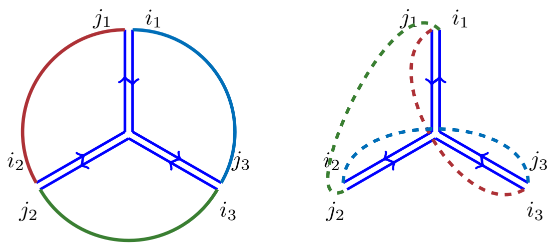


Figure 53: Ribbon graphs for three-gluon vertex with cyclic closing

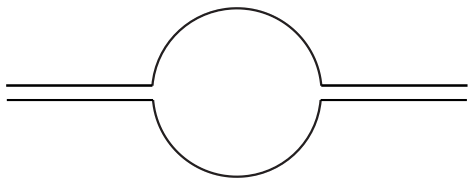


Figure 54: 1 virtual quark loop diagram

consider the diagram in figure 54, where we have two quark-gluon vertices and no closed index loops, the order of this diagram is thus given by

$$g^2 \sim \frac{1}{N_C}$$

making it an NLO diagram. In general a **diagram with L quark loops** will be of the order

$$\left(\frac{1}{N_C}\right)^{L+2n},$$

with n the genus of the graph. If we now also include external boundaries (i.e. valence quarks, which form the **boundary of the Riemann surfaces** generated by the graphs) and normalize the quark operator as

$$O = \frac{1}{N_C}\bar{\psi}\psi$$

to make it a leading order operator in the large N_c-expansion, then diagrams of genus n with L quark loops and B boundaries contribute at the order

$$\left(\frac{1}{N_C}\right)^{2n+L+2(B-1)}.$$

We end this subsection with the following remarks

Remark 21.3.1 (Link with Wilson Loop Variables).
Closing the planar graphs (like for example the left diagram in figure 53) cyclically by adding n boundary (valence) quarks gives rise to n-point Green's functions:

$$G^{(n)}_{\mu_1\cdots\mu_n}(x_1,\cdots,x_n) = \frac{(ig)^n}{N_c}\left\langle Tr[A_{\mu_1}(x_1)\cdots A_{\mu_n}(x_n)]\right\rangle,$$

which also naturally arises in the expansion of the Wilson loop variable:

$$\left\langle \frac{1}{N_c}\, Tr\, \mathcal{P}e^{\oint_\Gamma dx^\mu \mathcal{A}_\mu(x)}\right\rangle = \sum_{n=0}^{\infty}\oint_\Gamma dx_1^{\mu_1}\int_{x_1}^{x_2} dx_2^{\mu_2}\cdots\int_{x_1}^{x_{n-1}} dx_n^{\mu_n}\; G^{(n)}_{\mu_1\cdots\mu_n}(x_1,\cdots,x_n),$$

showing the link with these variables.

Remark 21.3.2 (Factorization of white operators).
A very important property of the large N_c-expansion is that the vacuum expectation values of white operators (singlets with respect to the gauge group) factorize in this limit (for a derivation see for instance [128]). Since Wilson loop variables are white (gauge invariant) operators we can apply this factorization such that in the large N_c-limit one can write:

$$\left\langle \frac{1}{N_c}\, Tr\, \mathcal{P}e^{\oint_{\Gamma_1} dx^\mu \mathcal{A}_\mu(x)}\cdots\frac{1}{N_c}\, Tr\, \mathcal{P}e^{\oint_{\Gamma_n} dx^\mu \mathcal{A}_\mu(x)}\right\rangle = \left\langle \frac{1}{N_c}\, Tr\, \mathcal{P}e^{\oint_{\Gamma_1} dx^\mu \mathcal{A}_\mu(x)}\right\rangle\cdots\left\langle \frac{1}{N_c}\, Tr\, \mathcal{P}e^{\oint_{\Gamma_n} dx^\mu \mathcal{A}_\mu(x)}\right\rangle$$

21.3.3 *Master Field*

Let us return to the Yang-Mills partition function

$$Z = \int \mathcal{D}A^a_\mu e^{-\int d^4x\ \frac{1}{4}F^a_{\mu\nu}F^a_{\mu\nu}}, \tag{328}$$

where due to the fact that $F^a_{\mu\nu}$ is of the order N_c the action is of order N_c^2 making it very large in the large N_c-limit. Next to that the entropy is also of order N_c^2 due to the $N_c^2 - 1$ integrations over A^a_μ such that

$$\mathcal{D}A^a_\mu \sim e^{N_c^2}.$$

From this it follows that the **saddle point equation for Yang-Mills theory** is **not** the classical saddle point Eq. (329).

$$\frac{\delta S}{\delta A^a_\mu} = (\nabla_\nu F_{\mu\nu})^a = 0. \tag{329}$$

The idea to deal with this problem is to rewrite the path integral over A_μ as one over **colorless composite fields** $\Phi(A)$, similar to the idea in **Sigma-Models** [137]. Applying this idea we expect the path integral to be of the form

$$Z \sim \int \mathcal{D}\Phi \, \frac{1}{\frac{\partial \Phi(A)}{\partial A^a_\mu}} e^{-N_c^2 S[\Phi]}, \tag{330}$$

where the **Jacobian** is given by

$$\frac{\partial \Phi(A)}{\partial A^a_\mu} \equiv e^{-N_c^2 J[\Phi]}$$

which is related to the old entropy such that we have

$$J[\Phi] \sim 1.$$

The original partition function is now given by

$$Z \sim \int \mathcal{D}\Phi \; e^{N_c^2 J[\Phi] - N_c^2 S[\Phi]}, \tag{331}$$

such that the saddle point equation becomes

$$\frac{\delta S}{\delta \Phi} = \frac{\delta J}{\delta \Phi} \tag{332}$$

$$\frac{\delta S}{\delta A^a_\mu} = (\nabla_\nu F_{\mu\nu})^a = \frac{\delta J}{\delta A^a_\mu}. \tag{333}$$

Note that the exact expression for $J[\Phi]$ depends on the new variable $\Phi(A)$. Witten [138] conjectured the existence of a classical solution A^{cl}_μ for multicolor QCD which was later referred to as the **master field** by Coleman [139]. But this solution is only determined up to gauge transformation, so that it is better to call the entire gauge orbit a solution, which is often referred to as being a solution in the strong sense and is necessary to preserve gauge invariance. If one assumes there is only a single solution then the path integral is saturated by this single configuration and the vacuum expectation values of gauge-invariant operators is given by their values at this configuration

$$\langle O\rangle = O\left(A_\mu^{cl}(x)\right), \tag{334}$$

such that the factorization property of Remark 21.3.2 is clearly satisfied. Since the LHS of Eq. (334) is Poincaré invariant so is the RHS such that space-time transformations can be undone by gauge transformations making the classical solution space-time independent. Haan [140] pointed out that assuming the existence of such as solution might be incorrect and indeed the assumption that there is a single solution seems restrictive. The point is that having multiple solutions is a tricky matter since one will need an additional averaging over these solutions. The large N_c-expansion in QCD assumption that the gauge-invariant objects behave like c-numbers instead of like operators seems to be incorrect and it would be better to think of $A_\mu^{cl}(0)$ as an operator in some Hilbert space. This follows from the fact that the $N_c \times N_c$ matrices become infinite matrices, hence operators on a Hilbert space. **A master field that is operator-valued is often referred to as a master field in the weak sense.** Although this concept of a master field is rather vague until a specific form for $\Phi(A)$ is given it is space-time independent[6]. A natural candidate for this composite operator $\Phi(A)$ is given by the Wilson loop variable [128, 141]:

$$W(\Gamma) := \frac{1}{N_c}\ \mathrm{Tr}\ \mathcal{P}e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu}, \tag{335}$$

where (matrix notation)

$$[\mathcal{A}_\mu] = ig\sum_a A_\mu^a(x)[t^a]^{ij}$$

and Γ is a closed loop. Notice that the new independent variables are now the loops.

Unfortunately up until now there has been no success in rewriting QCD in terms of these $\Phi(A)$ for finite N_c, but we would like to point out the paper [142] by Tsou. In this paper Tsou calculates the CKM and NMS matrices in an extension of the Standard Model which matches the experimental data quite

6 Notice how much this discussion resembles to problem of degenerate gauge potentials in the non-perturbative sector.

well. In his calculation he used Duality Operators, generalizing the Electro-Magnetic dual operators (Hodge dual) to Non-Abelian theories. The Non-Abelian dual operators were derived using loop space variables and Polyakov's derivative, showing a possible useful application of loop space within the Standard Model.

Coming back to the reason why we haven't been able to rewrite QCD in generalized loop space, this is mainly due to problems with self-intersections leading to the Mandelstam constraints which are taken into account by the equivalence relation used to introduce Generalized Loop Space in Chapter 18. So, the hope is that this formulation of loop space in the future might help understand and maybe solve this problem. First steps towards this goal will be discussed in chapter 22. Despite these problems Makeenko and Migdal were able to derive a reformulation using the (non-perturbative) SD equations resulting in the loop equations now referred to as the MM equations. In the next section we will derive these equations.

21.4 SD FOR WILSON LOOPS: MAKEENKO-MIGDAL EQUATIONS

We remind the reader of the definition of a Wilson loop used by Makeenko and Migdal [128, 141]:

$$W(\Gamma) := \frac{1}{N_c}\ \mathsf{Tr}\ \mathcal{P}e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu}, \tag{336}$$

where (matrix notation)

$$[\mathcal{A}_\mu] = ig\sum_a A^a_\mu(x)[t^a]^{ij}$$

and Γ is a closed loop. Using Eq. (313) with

$$\boldsymbol{F}[A] = W(\Gamma)$$

and the notation

$$\mathcal{F}_{\mu\nu} = \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu - [\mathcal{A}_\mu, \mathcal{A}_\nu]$$

(anti-Hermitian matrix of the Non-Abelian field strength) results in

$$\left\langle \frac{1}{N_c}\ \mathsf{Tr}\ \mathcal{P}D_\mu\mathcal{F}_{\mu\nu}(x)e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu}\right\rangle = \left\langle \frac{g^2}{2N_c}\ \mathsf{Tr}\ \frac{\delta}{\delta\mathcal{A}_\nu(x)}\mathcal{P}e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu}\right\rangle$$

where the D_μ is the covariant derivative in the **adjoint representation**

$$D^{\mathsf{adj}}_\mu \equiv \partial_\mu \cdot - [\mathcal{A}_\mu, \cdot]\,. \tag{337}$$

The variational derivative on the RHS can be calculated and follows from (using Tr $[t^a t^b] = \frac{1}{2}\delta^{ab}$):

$$\begin{aligned}\frac{\delta A^a_\mu(y)}{\delta A^b_\nu(x)} &= \delta_{\mu\nu}\delta^n(x-y)\delta^{ab} \\ &= \delta_{\mu\nu}\delta^n(x-y)2\text{Tr }[t^a t^b] \\ &= \delta_{\mu\nu}\delta^n(x-y)2\frac{1}{2}\left(\delta^{il}\delta^{jk} - \frac{1}{N_c}\delta^{ij}\delta^{jk}\right) \\ &= \delta_{\mu\nu}\delta^n(x-y)\left(\delta^{il}\delta^{jk} - \frac{1}{N_c}\delta^{ij}\delta^{jk}\right), \end{aligned} \tag{338}$$

resulting in [128]

$$\begin{aligned} N_c^2\text{Tr }\frac{\delta}{\delta\mathcal{A}_\nu(x)}\mathcal{P}e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu} = N_c^2 \oint_\Gamma dy_\nu\ \delta^n(x-y) \\ \times\left(\frac{1}{N_c}\text{Tr }\mathcal{P}\ e^{\int_{yx} dz^\mu\mathcal{A}_\mu}\frac{1}{N_c}\text{Tr }\mathcal{P}\ e^{\int_{xy} dz^\mu\mathcal{A}_\mu} - \frac{1}{N_c^3}\text{Tr }\mathcal{P}\ e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu}\right). \end{aligned} \tag{339}$$

For the LHS we can use Eq. (243) to rewrite it as

$$\begin{aligned}\triangle^E_{(\epsilon;u\wedge v)}(p)W(\Gamma) &\overset{\text{Not}}{\underset{\text{MM}}{=}} \frac{\delta}{\delta\sigma_{\mu\nu}(x)}\left(\frac{1}{N_c}\text{Tr }\mathcal{P}\ e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu}\right) \\ &= \frac{1}{N_c}\ \text{Tr }\mathcal{P}\ \mathcal{F}_{\mu\nu}(x)e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu}. \end{aligned} \tag{340}$$

Since the field strength tensor operates at a specific point along the contour this result is a **marked path functional** (Definition 19.2.3) on which we can apply the path derivative. By Eq. (226) this is the same as applying the covariant derivative in the adjoint representation (Eq. (337)) such that we have for the LHS

$$\partial^x_\mu\left(\frac{\delta}{\delta\sigma_{\mu\nu}(x)}\left[\frac{1}{N_c}\text{Tr }\mathcal{P}e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu}\right)\right] = \frac{1}{N_c}\ \text{Tr }\mathcal{P}\ D^{\text{adj}}_\mu\mathcal{F}_{\mu\nu}(x)e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu}. \tag{341}$$

Combining Eqs. (341) and (339) we arrive at the **Makeenko-Migdal equations**

$$\begin{aligned}\frac{1}{N_c}\ \text{Tr }\mathcal{P}\ D^{\text{adj}}_\mu\mathcal{F}_{\mu\nu}(x)e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu} = \frac{g^2N_c^2}{2N_c}\oint_\Gamma dy_\nu\ \delta^n(x-y) \\ \times\left(\frac{1}{N_c}\text{Tr }\mathcal{P}\ e^{\int_{\Gamma_{yx}} dz^\mu\mathcal{A}_\mu}\frac{1}{N_c}\text{Tr }\mathcal{P}\ e^{\int_{\Gamma_{xy}} dz^\mu\mathcal{A}_\mu} - \frac{1}{N_c^3}\text{Tr }\mathcal{P}\ e^{\oint_\Gamma dz^\mu\mathcal{A}_\mu}\right), \end{aligned} \tag{342}$$

where the equality is in the weak sense, rewriting this with the expectation values we have:

$$\left\langle \frac{1}{N_c} \text{ Tr } \mathcal{P} \; D_\mu^{\text{adj}} \mathcal{F}_{\mu\nu}(x) e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu} \right\rangle = \\ \frac{g^2 N_c}{2} \oint_\Gamma dy_\nu \; \delta^n(x-y) \left(\langle \phi(\Gamma_{yx}) \phi(\Gamma_{xy}) \rangle - \frac{1}{N_c^2} \langle \phi(\Gamma) \rangle \right), \tag{343}$$

where we used the notation

$$\phi(\Gamma) = \frac{1}{N_c} \text{Tr } \mathcal{P} \; e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu}$$

One of the problems with these equations is that they do not close, therefore in many cases one takes the large N_c limit such that the second term on the RHS vanishes and the first term factorizes as discussed in the previous section. This allows to simplify the Makeenko-Migdal equations, which we will not further discuss here. Details can be found in [128, 141].

21.5 SUMMARY

In this chapter, after reviewing some details on Fierz identities in $SU(N)$-theories, we introduced multi-color QCD. The Ribbon graphs where introduced and their relation with topology and Feynman diagrams was discussed. Using these graphs we commented on some advantages and properties of multi-color QCD in the t'Hooft or large N_c-limit. We also explained an approach to this multi-color QCD that is similar to the one used in Sigma-models, introducing the master field, and how this might open the door to rewrite QCD in a loop space representation by changing the variables (fields) in the action to composite colorless objects. A natural candidate for these composite objects are Wilson loop variables. Applying the SD equations, discussed in the previous Chapter, to vacuum expectation values of the Wilson loop variables then allowed us to re-derive the MM loop equations.

22

WILSON LOOP QUADRILATERAL ON THE LIGHT CONE

22.1 INTRODUCTION

In this Chapter we introduce a simple example of a Wilson loop, **the quadrilateral on the light cone**. We study the geometrical properties and behavior of this loop under the operation of different differential operators and relate it to its **ultraviolet and rapidity divergences.** The interesting fact about this loop is that it is not described by the Makeenko-Migdal loop equations due to the presence of **cusps**, as will become clear when we try to apply the area derivative to this loop. On the other hand, from a field theoretic point of view the study of these cusps is important to understand the underlying QFT since cusps are possible generators of anomalies and are related to the **cusp anomalous dimension**, an object that seems to be important in any QFT.

22.2 RENORMALIZATION OF LOOP FUNCTIONALS AND THE CUSP ANOMALOUS DIMENSION

This Section reviews some issues with the renormalization of loop functionals, which eventually lead to the introduction of the cusp anomalous dimension [143–148].

In [88] Ivanov, Korchemsky and Radyushkin discuss the renormalization properties of contour averages or said differently the renormalization properties of the vacuum expectation value of Wilson loop variables

$$W(\Gamma) = \mathsf{Tr}\ \left\langle 0 \left| \frac{1}{N_c} \mathcal{P} e^{ig \oint_\Gamma dz^\mu \ A_\mu(z)} \right| 0 \right\rangle .$$

They argue that in the case of a **smooth closed contour** the renormalization of such an object reduces to the **coupling constant renormalization**, but that this is **no longer true when the contour contains singular points.** This singularity can come from two possible scenarios:

(i) At the endpoint of an open contour

(ii) Cusps and Self-Intersections

In [143, 144, 146, 147] it was demonstrated that the first case leads to an extra multiplicative renormalization of the Wilson loop functional:

$$W(\Gamma, g(\mu), \mu) = Z_C\left(\frac{\mu}{\mu'}, g(\mu')\right) W(\Gamma, g(\mu'), \mu'),$$

where the μ and μ' refer to the renormalization parameters. For the second case it was found in [146, 147] that one needs an additional contribution (at LO)

$$W_{\text{cusp}}(\Gamma) = -\frac{\alpha_s}{\pi} C_F(\gamma \cot(\gamma) - 1) \ln\left(\frac{1}{a}\right),$$

where γ is the cusp angle and a **the ultraviolet regularization cut-off parameter.** In [148] it was then shown that together with the higher order contributions these extra singularities can be **rewritten as a multiplicative factor Z_{cusp}, from which the cusp anomalous dimension is defined as**

$$\Gamma_{\text{cusp}} = -\frac{\partial \ln Z_{\text{cusp}}}{\partial \ln \ \mu},$$

such that the for the renormalization of this type of Wilson loop functional one gets

$$W(\Gamma, g(\mu), \mu) = Z_{\text{cusp}}\left(\gamma_i, \frac{\mu}{\mu'}, g(\mu')\right) Z_C\left(\frac{\mu}{\mu'}, g(\mu')\right) W(\Gamma, g(\mu'), \mu'), \tag{344}$$

the γ_i representing the different cusps in the contour.

An **extra set of singularities mixes** with the above ones, when parts of or the entire contour is laying **on the light cone** [88, 149, 150]. In this case the singularities stemming from the light cone overlap with the previously discussed singularities destroying the simple multiplicative renormalization of Eq. (344). In our papers [48, 151–153], which we will discuss in detail in the following Sections, we demonstrated that by using a **geometrical approach** and combining it with the regular mass scale differential operator

$$\mu \frac{d}{d\mu}$$

we can write down **an evolution equation for contours on the light cone AND having cusps**. From the discussion above it will be no surprise that the cusp anomalous dimension will play an important role in this equation. Furthermore we give a simple example of a contour with **Self-Intersection** [50], which can be handled by the group structure of generalized loop space,

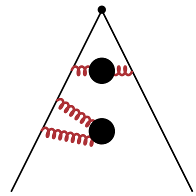


Figure 55: Simplest diagram with Γ_{cusp}

demonstrating that the evolution equation we conjectured is also valid in this case. We end this section with some details on the **two-loop expression of the cusp anomalous dimension** Γ_{cusp}. In [154] it was demonstrated that the cusp anomalous dimension depends only on one parameter, namely the cusp angle γ. From this fact they calculated the two-loop expression for the Γ_{cusp} using the simplest possible diagram shown in figure 55, where we have drawn some possible gluon exchanges. For the two loop calculation all two loop order corrections of the cusp needed to be calculated, which was done in [154].

The explicit two-loop expansions of the **cusp anomalous dimensions** for the case the **angle goes to infinity**, which is the case for the angle between two light-cone directions in Minkowski space due to the definition of the angle

$$\cosh\gamma = \frac{x \cdot y}{\sqrt{x^2 y^2}},$$

is given by

$$\Gamma_{\text{cusp}} = \frac{\alpha_s}{\pi} C_F + \left(\frac{\alpha_s}{\pi}\right)^2 C_F \left(C_A \left(\frac{67}{36} - \frac{1}{12}\pi^2\right) - \frac{5}{18} N_F\right), \tag{345}$$

where the term

$$\left(\frac{\alpha_s}{\pi}\right)^2 C_F \left(C_A \left(\frac{67}{36}\right)\right)$$

is a (renormalization) **scheme dependent** term stemming from the regularization of the different contributing diagrams to the cusp anomalous dimension (here $\overline{\text{MS}}$ was used).

22.3 PARAMETRIZATION OF PLANAR QUADRILATERAL WILSON LOOP

The Wilson loop we first considered is a quadrilateral on the light cone, shown in figure 56. This loop can be described using the four-vectors v_i where we assume each side of the loop is lying on a light cone direction such that

$$v_i^2 = 0, \forall\ i.$$

In order for the loop to be planar we also need to have that $v_1 = -v_3$

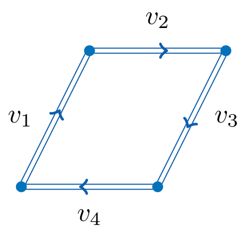


Figure 56: Loop Parametrization

and $v_2 = -v_4$. For further convenience we also introduce the Mandelstam variables

$$s = (v_1 + v_2)^2 \tag{346}$$
$$t = (v_2 + v_3)^2, \tag{347}$$

which in this specific case of the quadrilateral reduce to

$$s = 2v_1v_2 \qquad \text{and} \qquad t = 2v_2v_3.$$

These variables will allow us to write our results in a compact way.

22.4 LEADING ORDER RESULT

We now want to calculate the vacuum expectation value of the Wilson Loop variable from Eq. (335):

$$\left\langle 0 \left| \frac{1}{N_c} \text{Tr}\, \mathcal{P} e^{\oint_\Gamma dz^\mu \mathcal{A}_\mu} \right| 0 \right\rangle, \tag{348}$$

at leading order. This means that we will have to calculate the contributions of the diagrams shown in figure 57, where we also show the parametrizations of the involved segments. In our calculations we adopted the **Feynman**

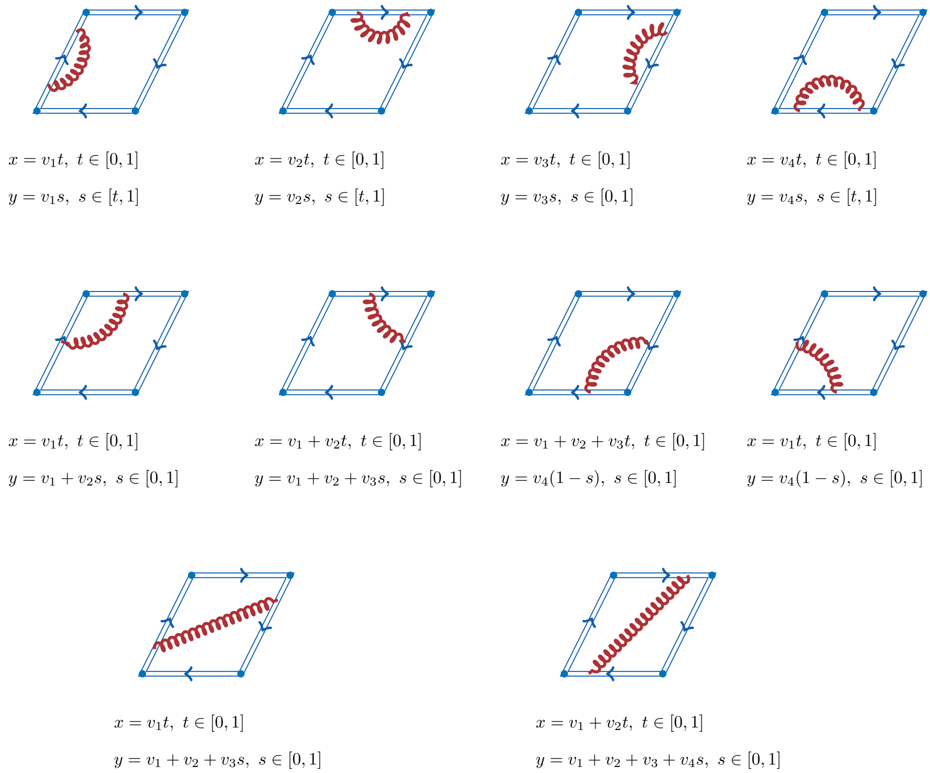


Figure 57: LO contributions

gauge, together with dimensional regularization, such that the **coordinate expression for the gluon propagator** given by

$$D_{\mu\nu}(x-y) = \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)g_{\mu\nu}\delta^{ab}\frac{1}{(-(x-y)^2)^{1-\epsilon}}, \qquad (349)$$

was used to evaluate the **Wick contractions of the gauge fields.** If one investigates the diagrams in Figure 57 closely then, considering the symmetries introduced by our choice of contour and its parametrizing vectors, it is obvious that we have three classes of diagrams represented by the rows of diagrams in this Figure. The diagrams in each row are related to each other by the relations between the parametrizing vectors v_i, such that we only need to calculate one diagram in each row. The first row of diagrams are the **Self-Energy (SE) diagrams** which in dimensional regularization all reduce to zero (see Appendix B.6 for the details of the calculation). An example of the calculation of a diagram with a cusp is shown in Appendix B.7, and will be shown in the total result. The diagrams in the **last row** of figure 57, although **finite**, are less easy to evaluate and eventually give rise to **poly-logarithms.** The method to solve this type of integrals emerging in such diagrams is explained in Appendix B.8, and the actual calculation is done in Appendix B.9. Combining all the above contributions results in

$$W_{\text{LO}}(\Gamma_{\square}) = 1 - \frac{\alpha_s C_F}{\pi}\left(2\pi\mu^2\right)^{\epsilon}\Gamma(1-\epsilon)\left[\underbrace{\frac{1}{\epsilon^2}\left(-\frac{s}{2}\right)^{\epsilon} + \frac{1}{\epsilon^2}\left(-\frac{t}{2}\right)^{\epsilon}}_{\text{Cusp}} \underbrace{-\frac{1}{2}\left(\ln^2\frac{s}{-t} + \pi^2\right)}_{\text{Cross}}\right] + \mathcal{O}(\alpha_s^2), \tag{350}$$

which we found to be consistent with the results found in [72, 155–157]. Since this Wilson Loop Variable is per definition a gauge invariant object it can be used to check consistency of different models and theories which was discussed, next to the prescription dependence of the extra pole emerging in the gluon propagator in axial gauges of Eq. (350), in [158].

22.5 FAILURE OF AREA DERIVATIVE

In this Section we investigate the behavior of the **area derivative** as defined on GLS in [16] on the **dimensionally regularized quadrilateral** from the previous section. From the paper by Tavares [16] it follows that there are two approaches for calculating this area derivative. One approach is to calculate the derivative directly in the sense that one calculates

$$\Phi_{\text{LO}}(\lambda_t \cdot \gamma)$$

explicitly in the calculation of

$$\Delta_{\triangle;(u,v)}(q)\Phi(\gamma) = \lim_{t\to 0}\frac{\Phi_{\text{LO}}(\lambda_t \cdot \gamma) - \Phi_{\text{LO}}(\gamma)}{t^2}, \tag{351}$$

where

$$\Delta_{\triangle;(u,v)}(q)\Phi(\gamma) \qquad \text{and} \qquad \lambda_t$$

are defined as in Chapter 19. Another approach is to make use of the group properties of GLS by considering the area variated loop as the product of the original loop with the loop representing the area variation to calculate

$$\Phi_{\text{LO}}(\lambda_t \cdot \gamma).$$

We used both approaches and found that the results were the same, reconfirming the validity of the group structure of GLS. Both approaches gave however a **divergent result**, when we considered area variations as shown in figure 58, which is due to the fact that **Stokes theorem seems to hold only in the classical case**, i.e. not in the quantum case[1]. This remark is quite profound (see also [159] for comments) and we will return to it after having shown that the result of the area derivative is actually divergent due to the presence of terms that have factors of the form

$$\frac{1}{a^{2-2\epsilon}\epsilon^2}$$

where the limit $a \to 0$ has to be taken indicating the **extra divergence** not taken care of by dimensional regularization.

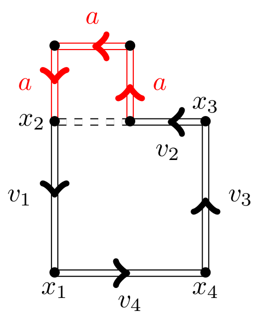


Figure 58: Diagram with infinitesimal area variation.

We also would like to point out that Makeenko and Migdal used this derivative [141, 160] to derive the loop equations as was demonstrated in chapter 21, making a discussion on the validity of this derivative or more specifically on the validity of Stokes' theorem more than relevant. As discussed in Section 19.3 one needs two independent vector fields to have a well defined infinitesimal area variation of the loop. It is easy to see that on the null-plane there are only two independent directions, the positive and negative light cone direction. Building an infinitesimal area variation out two such vectors, for which the local flows are straight paths on the light cone, leads to a diagram as shown in figure 58, where we have chosen the top left corner of our original diagram (the point x_2 of figure 58) as the point where the differential operator acts (see the discussion on marked path functionals in 19.2). In the direct calculation approach, we now get a new path where the double dashed line represents

1 Said differently: Quantum corrections break the application of the Stokes Theorem.

the part of the path that disappears due to path reduction [16, 50, 107] of the overlapping paths in the original diagram and the infinitesimal added "rectangle". The introduction of this area variation combined with the path

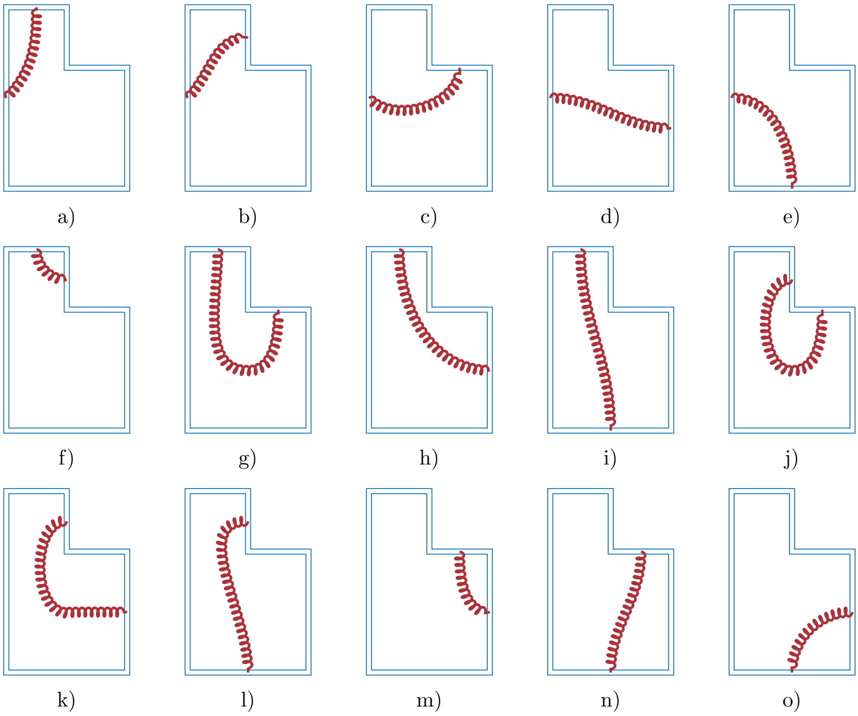


Figure 59: Table of generated diagrams.

reduction now generates new diagrams at the one-loop order which are shown in Figure 59. Direct calculation of these diagrams, for which the calculations are similar to the ones for the W_{LO} but with different integration boundaries, and subtracting the one loop contribution of the original loop results in Eq. (352). To arrive at the derivative we still need to divide by a^2 and take the $a \to 0$ limit, to be consistent with Eq. (351). The difference between the two diagrams is given by

$$
\begin{aligned}
&W_{\text{L.O.}}(\Gamma_{\square}^{E}) - W_{\text{L.O.}}(\Gamma_{\square}) = \\
&F\,(-s)^{\epsilon}\,\frac{1}{2}\left(\frac{1}{\epsilon}\right)^{2}\Big(-(1+a)^{\epsilon}a^{\epsilon} + ((-a)^{\epsilon}-1)\,(1-a^{\epsilon}) - ((a-1)^{\epsilon}-(-1)^{\epsilon})\,((1+a)^{\epsilon}-a^{\epsilon}) \\
&\qquad + (-a)^{\epsilon}((1-a)^{\epsilon}-(-a)^{\epsilon}) - ((a+1)^{\epsilon}-1)((a-1)^{\epsilon}-a^{\epsilon}) + 1\Big) \\
&+F\,(-t)^{\epsilon}\left(\frac{1}{\epsilon}\right)^{2}\Big(2-(1+a)^{\epsilon}-a^{2\epsilon}-(1-a)^{\epsilon}\Big) \\
&+F\int_0^{1+a} dt \int_0^{a} ds \frac{v_1v_3}{(-(2v_2v_3as+2v_1v_2a(1+a-t)+2v_1v_3s(1+a-t)))^{1-\epsilon}} \\
&+F\int_0^{1+a} dt \int_0^{1} ds \frac{v_1v_3}{(-(2v_1v_2(1-t)+2v_2v_3s+2v_1v_3s(1-t)))^{1-\epsilon}} \\
&+F\int_0^{a} dt \int_a^{1} ds \frac{v_2v_4}{(-(2v_1v_2at-2v_1v_4as-2v_2v_4st))^{1-\epsilon}} \\
&+F\int_0^{a} dt \int_0^{1} ds \frac{v_2v_4}{(-(2v_1v_2t(1+a)+2v_1v_4s(1+a)+2v_2v_4st))^{1-\epsilon}} \\
&+F\int_0^{a} dt \int_0^{1} ds \frac{v_1v_3}{(-(2v_1v_2(t-a)(1-a)+2v_2v_3(1-a)s+2v_1v_3(t-a)s))^{1-\epsilon}} \\
&+F\int_a^{1} dt \int_0^{1} ds \frac{v_2v_4}{(-(2v_1v_2t+2v_1v_4s+2v_2v_4st))^{1-\epsilon}} \\
&-F\int_0^{1} dt \int_0^{1} ds \frac{v_2v_4}{(-(2v_1v_2t+2v_1v_4s+2v_2v_4st))^{1-\epsilon}} \\
&-F\int_0^{1} dt \int_0^{1} ds \frac{v_1v_3}{(-(2v_1v_4t+2v_3v_4s+2v_1v_3st))^{1-\epsilon}},
\end{aligned}
\tag{352}
$$

where F represents a constant factor that is irrelevant in our discussion here.

If we now focus on the second term in the sum of Eq. (352) we notice that the highest power of a is 2ϵ, therefore contributing with a factor $a^{2-2\epsilon}$ in the denominator (next to the usual singular term ϵ^2, which is taken care of by dimensional regularization), leading to extra divergences in the limit $a \to 0$. **This factor makes the area derivative divergent even in the dimensionally regularized case.** This result differs from what one would expect from the approach taken by Makeenko and Migdal in their derivation of the Makeenko-Migdal equation.

To understand what is going on we have to discuss the definition of the area derivative in more detail. Looking back at Eqs. (242), (243) and the integrals in Lemma 19.3.1 we can ask where the factors involving the wedge product of the local vectors u and v come from. In fact they find their origin in the integrals shown in Eqs. (353-355), restated here for convenience,

$$\lim_{a\to 0} \frac{\int_{\square_a} \omega}{a^2} = \int_V d\omega = d\omega(u \wedge v) \quad \text{using generalized Stokes} \tag{353}$$

$$\lim_{a\to 0} \frac{\int_{\square_a} \omega_1\omega_2}{a^2} = (\omega_1 \wedge \omega_2)(u \wedge v) \tag{354}$$

$$\lim_{a\to 0} \frac{\int_{\square_a} \omega_1 \cdots \omega_r}{a^2} = 0, \qquad \forall r \geq 3, \tag{355}$$

where we refer to [16] and Part III of this text for the details on how these integrals emerge. The results in Eqs. (353-355) are indeed correct in a classical setting, in the sense that they do not contain any quantum state operators. But, in our calculation of the area derivative above we have introduced the **vacuum expectation value** taking us away from the classical setting into a quantum one. Introducing the vacuum expectation value is what generates the diagrams shown in figure 59, which introduces divergences due to the presence of cusps (in the Feynman gauge). Let us have a look in some more detail at the calculations to see where things might go wrong if we compare with the calculations done by Makeenko-Migdal. This investigation is the most straightforward in the second approach to calculate the area derivative where we used the group properties of GLS.

In GLS, adding the area variation (λ_a) at the top left corner, can be seen as multiplying the original loop ($\Gamma_\square$) with an extra infinitesimally small loop ($\lambda_a \cdot \Gamma_\square$) at that point (the top left corner becomes the loop space base point, which is allowed due to translation invariance caused by path reduction). Using the properties of the product in the topological group of generalized loops the leading order contribution can then be written as

$$W_{\text{LO}}(\lambda_a \cdot \Gamma_\square) = \int_{\lambda_a} \omega_1\omega_2 \cdot 1 \pm \int_{\lambda_a} \omega_1 \int_{\Gamma_\square} \omega_2 + 1 \cdot \int_{\Gamma_\square} \omega_1\omega_2, \tag{356}$$

where the $\pm$ indicates the two possible relative orientations between the loops. We will use the plus sign, where the orientations are the same so that the overlapping part of the loops cancel due to path reduction but choosing any of the signs will not change the singular structure, so the discussion is also valid for the minus sign. Subtracting the leading term of the original loop $W(\Gamma)$ from Eq. (356) will cancel the last term such that we are only left with two terms to check for convergence. To introduce the vacuum expectation value we will focus on the first term in Eq. (356). There are now two ways to introduce the vacuum expectation values

$$\left\langle 0 \left| \lim_{a\to 0} \frac{\int_{\lambda_a} \omega_1\omega_2}{a^2} \right| 0 \right\rangle, \tag{357}$$

$$\lim_{a\to 0} \frac{\int_{\lambda_a} \langle 0 |\omega_1\omega_2| 0\rangle}{a^2}. \tag{358}$$

Using the first approach, shown in Eq. (357), in combination with the result of Eq. (354) leads to a tadpole diagram at the top-left corner of the quadrilateral and can be treated by the usual regularizations schemes like dimensional regularization (see also comment in [159]). On the other hand in the second approach, shown in Eq. (358), the integral accumulates an extra divergence (as also demonstrated in the direct calculation approach of the area derivative). This can be seen from the fact that the first term in Eq. (356) is the same as the original loop with the Mandelstam variables s and t rescaled with a factor a^2 leading to the terms

$$-\frac{\alpha_s C_F}{\pi}\left(2\pi\mu^2\right)^\epsilon \Gamma(1-\epsilon)\left[\frac{1}{\epsilon^2}\left(\left(-\frac{a^2 s}{2}\right)^\epsilon + \left(-\frac{a^2 t}{2}\right)^\epsilon\right) - \frac{1}{2}\left(\ln^2\frac{s}{t} + \pi^2\right)\right] = \tag{359}$$

$$-\frac{\alpha_s C_F}{\pi}\left(2\pi\mu^2\right)^\epsilon \Gamma(1-\epsilon)\left[\frac{a^{2\epsilon}}{\epsilon^2}\left(\left(-\frac{s}{2}\right)^\epsilon + \left(-\frac{t}{2}\right)^\epsilon\right) - \frac{1}{2}\left(\ln^2\frac{s}{t} + \pi^2\right)\right]. \tag{360}$$

Dividing this result by a^2 and taking the limit $a \to 0$ already shows that this is divergent since ϵ is considered small (for sure smaller than 2) but non-zero in the dimensional regularization scheme we applied here to do the calculations. Applying a similar calculation to the second term of Eq. (356) gives rise to terms proportional to $(1+a)^\epsilon$ so that also this term after division by a^2 becomes divergent in the $a \to 0$ limit, when taking the vacuum expectation values as in Eq. (358). In the approach shown in Eq. (357) this would again return a finite result after the appropriate regularization. As can be seen from Eq. (243) these integrals appear at any order of the expansion of the exponential of the Wilson Loop Functional. From the discussion above **we are inclined to conclude that the integral and the vacuum expectation value does not commute with Stokes' theorem** as is assumed in many papers. This result is **not so unexpected if we take a diffeomorphism point of view.** It is not so hard to see that starting from a loop with let us say n cusps and deforming it to a loop with $n+1$ cusps can not by realized by a smooth diffeomorphism due to the break down of the continuity of the first order derivative. Introducing the area variation as in figure 58 shows that this introduces two more cusps in the loop. If we now follow the generally accepted idea that the same physics on a manifold can be related by diffeomorphisms, the fact that the original diagram and the diagram with the area variation included have a different number of cusps, hint of an idea that maybe **loops with different number of cusps represent different physical objects.** To sketch a picture in the mind (which might be wrong, but this is just to explain the idea), we could say that a loop with two cusps corresponds to a meson and one with three to a hadron for example.

22.6 GENERALIZED DERIVATIVE: A NEW DERIVATIVE

In the previous Section we showed that there are issues with applying the area derivative to the Wilson loop quadrilateral on the light cone. To solve this problem Dr. I. Cherednikov came up with an idea for a new differential operator which we introduced in our papers [151–153, 161] to derive an evolution equation conjecture which will be discussed in Section 22.9. Here we will give the definition of this differential operator.

The geometric operation we are interested in is area variation, which is also partially motivated by the area law [162] and its relation to confinement. The main purpose of the new differential operator is to result in an area variation after application to our quadrilateral loop from the above Sections. Taking into account that we would like that the resulting loop is still completely on the light cone the possible variations are quite limited. The only two possible variations are in the positive and negative light cone directions (see Figure 60). Considering these variations we can construct the corresponding differential

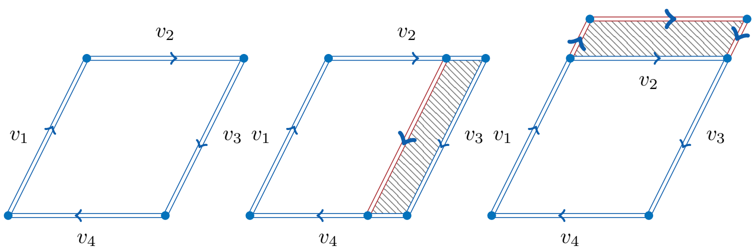


Figure 60: Two possible area variations on the light cone.

operators generating these variations:

$$\delta\sigma^{+-} = N^{+}\delta N^{-} \to v_2\delta v_1 = \frac{1}{2}\delta s\,, \quad \delta\sigma^{-+} = -N^{-}\delta N^{+} \to -v_1\delta v_2 = \frac{1}{2}\delta t\,, \tag{361}$$

from which we can construct a more general logarithmic differential operator, where the idea of taking the logarithmic derivation stems from a similar way of differentiating with respect to the renormalization parameter or mass scale parameter μ in renormalization schemes. This results in the differential operator:

$$\frac{\delta}{\delta \ln \sigma} \equiv \sigma_{+-}\frac{\delta}{\delta\sigma_{+-}} + \sigma_{-+}\frac{\delta}{\delta\sigma_{-+}}, \tag{362}$$

which we will relate to the **Fréchet derivative** (272) in Section 22.11. In the next Section we will apply this derivative to the quadrilateral loop.

22.7 APPLICATION TO THE QUADRILATERAL WILSON LOOP

Applying this new logarithmic differential operator (Eq. (362)) to the expanded logarithm of our LO result from Eq. (350) we get [161]:

$$\frac{\delta}{\delta\sigma}\ln\left[W(\Gamma)\right] = -\frac{\alpha_s C_F}{\pi}\frac{1}{\epsilon}\left(\left[s\mu^2 + i\cdot 0\right]^{\epsilon} + \left[t\mu^2 + i\cdot 0\right]^{\epsilon}\right). \tag{363}$$

Applying the renormalization derivative

$$\mu\frac{d}{d\mu},$$

followed by the limit $\epsilon \to 0$ to this result we end up with

$$\mu\frac{d}{d\mu}\frac{\delta}{\delta\sigma}\ln\left[W(\Gamma)\right] = -4\Gamma_{\text{cusp}}, \tag{364}$$

$$\Gamma_{\text{cusp}} = \frac{\alpha_s C_F}{\pi} + \mathcal{O}(\alpha_s^2), \tag{365}$$

where Γ_{cusp} is the **cusp anomalous dimension** we introduced in Section 22.2. A very important property of this differential operator is that it **lowers the degree of divergence**, which will be used in the next Sections to construct renormalization group equations for this type of loops.

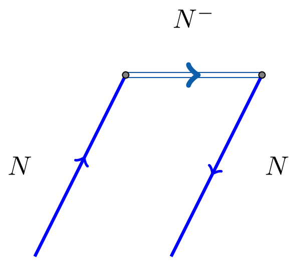


Figure 61: Pi-shaped contour

22.8 APPLICATION TO THE Π CONTOUR [151–153, 161]

To test [151–153, 161] if some similar equation holds for other loops than the quadrilateral on the light cone we consider here a **Pi-shaped contour** (Figure 61) where the two lines that run to infinity are **off the light cone** and the finite segment lays along the negative light cone direction [155]. The LO contribution of this contour is given by

$$W(\Gamma_\Pi) = 1 + \frac{\alpha_s N_c}{2\pi} + \left(-L^2(NN^-) + L(NN^-) - \frac{5\pi^2}{24}\right),$$

where

$$L(NN^-) = \frac{1}{2}\left(\ln\left(\mu NN^- + i\cdot 0\right) + \ln\left(\mu NN^- + i\cdot 0\right)\right)^2,$$

and where the area is now defined by the product NN^- of the non light like vector N and the light like vector N^-. Application of the differential operator (362) leads to the result

$$\mu\frac{d}{d\mu}\left[\frac{\delta}{\delta\sigma}\ln\left[W(\Gamma_\Pi)\right]\right] = -2\Gamma_{\text{cusp}}, \tag{366}$$

a similar result to Eq. (363).

22.9 CONJECTURE : A NEW EVOLUTION EQUATION

Eqs. (363) and (366) seem to indicate that the number of cusps in the loop determine the exact relation between the cusp anomalous dimension and the combined operation of the renormalization mass scale and our generalized area differentiation. This lead us to conjecture [161]

$$\mu\frac{d}{d\mu}\left[\frac{\delta}{\delta\sigma}\ln\left[W(\Gamma)\right]\right] = -\sum_{\text{cusps}}\Gamma_{\text{cusp}} \tag{367}$$

as an evolution equation for contours with cusps. Considering now the SD equations which we discussed in Section 20.3.3, the application of our new derivative in this context results in

$$\frac{\delta}{\delta\sigma}\langle\alpha|\beta\rangle = \frac{i}{\hbar}\left\langle\alpha\left|\frac{\delta S}{\delta\sigma}\right|\beta\right\rangle, \tag{368}$$

where S is the action operator in loop space that governs the variation of states. Up until now this action has been unknown, but if our conjecture would be valid to higher orders it would allow us to **reconstruct such an action**. This would open the door for quantum calculations on loop space, an alternative method to the regular gauge theory calculations. Moreover, the SD equations are also valid non-perturbatively such that an action in loop space might give us access to non-perturbative calculations.

22.10 APPLICATION TO A SYMMETRIC PRODUCT OF TWO WILSON LOOPS

To further explore the application range of the above conjecture we considered two symmetric extensions of the quadrilateral Wilson loop in our paper [50] (Shown in Figure 56). In a first case we added a copy of the original loop, but shifted it to the left resulting in a loop with overlapping paths. Due to the two possible relative orientations between the two constituting loops this gives rise to two possibilities which are shown in Figure 62, the dashed lines indicate that the two vertical paths actually overlap but we have drawn them separately enabling us to explain the cusp counting (see further). In a second case we also added a copy of the original loop but now we shifted it left and down, again taking into account the two possible relative orientations, this generates loops with a self-intersection shown in Figure 63. The question is

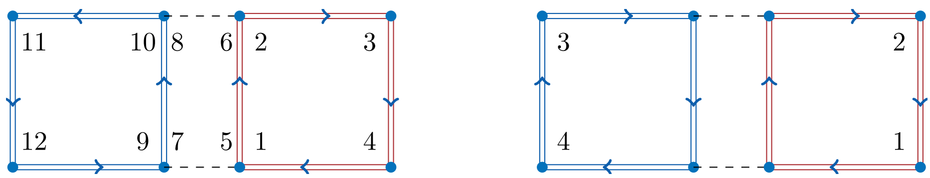


Figure 62: Overlapping paths

now if our conjecture in Eq. (367) is still valid for these diagrams if we consider the area variations as shown in Figures 64 and 65. Only these variations were considered because on one hand if one would vary asymmetrically the results

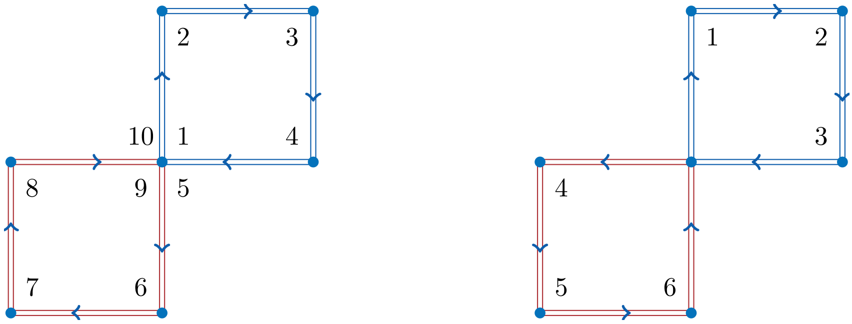


Figure 63: Self-intersecting paths

would need more parameters then just the Mandelstam variables s and t since the diagrams after variation would not be symmetric anymore and on the other hand some asymmetric variations lead to changes in the structure of the loops (for instance removal of the self-intersection). Moreover if one would consider the case where the two loops are not equal in size one would also need more parameters to describe the loops. Even in more general cases this is not a problem since formally one can extend the definition of the generalized derivative to incorporate the **generalized Mandelstam variables**

$$s_{ij} = (v_i + v_j)^2,$$

where the v_i, v_j are the vectors on which the polygonal loops are built. Investigating such cases would be a natural continuation of the research we discuss in this Section. To calculate the leading order contributions of these

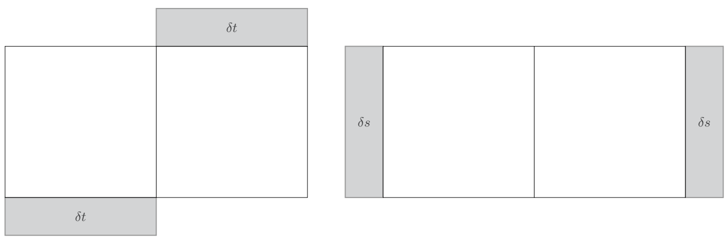


Figure 64: Variations of the first case

diagrams we expand the Wilson loop variable of Eq. (348) to first order

$$\mathcal{W}_1(\Gamma) = 1 - \frac{g^2}{2!}\text{Tr}(t^a t^b)\Big\langle 0\Big|\mathcal{T}\oint_\Gamma dz^\mu dz'^\nu A^a_\mu(z) A^b_\nu(z')\Big|0\Big\rangle \ , \tag{369}$$

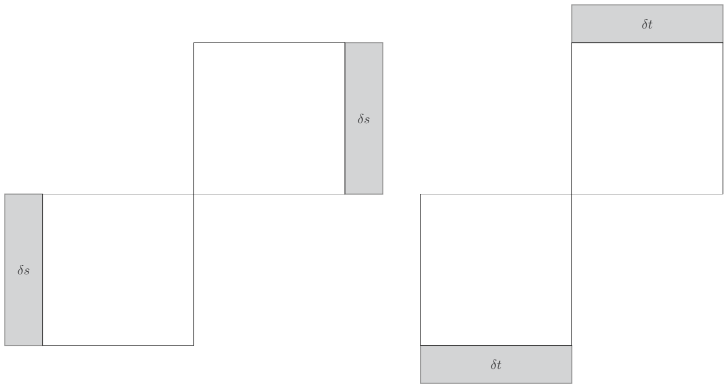


Figure 65: Variations of the second case

The generators and gauge connections in Eq. (369) are ordered along the loop by the time-ordering operation $\mathcal{T}$, where the "time" is represented by the path parameter $t \in [0, 1]$ such that

$$dz^{\mu} = \dot{z}^{\mu} dt.$$

Now, considering

$$\Gamma = \Gamma_1 \Gamma_2,$$

the group structure of GLS [16, 107, 108, 111] allows us to rewrite the integral in the second term of Eq. (369) as

$$\oint_{\Gamma_1\Gamma_2} \mathcal{A}_{\mu}\mathcal{A}_{\nu} = \oint_{\Gamma_1} \mathcal{A}_{\mu}\mathcal{A}_{\nu} \pm \oint_{\Gamma_1} \mathcal{A}_{\mu} \oint_{\Gamma_2} \mathcal{A}_{\nu} + \oint_{\Gamma_2} \mathcal{A}_{\mu}\mathcal{A}_{\nu} \ , \tag{370}$$

where $\mathcal{A}_{\mu}$ and $\mathcal{A}_{\nu}$ are again ordered along the path[2] and the $\pm$ represents the two different choices for the relative orientations of the loops where we also suppressed the integral measures. We now have three contributions (370), two coming from the loops considered independently, and one coming from the **interference terms.**

2 This means we do not need to consider the contribution $\int\limits_{\Gamma_1} \mathcal{A}_{\nu} \int\limits_{\Gamma_2} \mathcal{A}_{\mu}$.

As for the Wilson loop quadrilateral we use the dimensionally regularized gluon propagator in the Feynman gauge such that we can re-use the result for the single quadrilateral for two of the three contributions [163–166]

$$W_1(\Gamma_1) \quad = W_1(\Gamma_2) \approx 1 - \frac{C_F \alpha_s \pi^\epsilon}{\pi} \Gamma[1-\epsilon] \left(\frac{1}{\epsilon^2}(-s\mu^2)^\epsilon + \frac{1}{\epsilon^2}(-t\mu^2)^\epsilon - \frac{1}{2}\ln\left(\frac{s}{t}\right)^2 + \text{ finite} \right) + \mathcal{O}(\alpha_s^2), \tag{371}$$

and assumed that Γ_1 and Γ_2 are equal in size.

The interference terms ask for a bit more work but are also quite straightforward to evaluate, a calculation example of one such contribution can be found in [50]. The results for the interference term for the two diagrams with the different relative orientations are given below (the indices of the square brackets number the different possibilities).

$$\left[\left\langle 0 \left| \oint_{\Gamma_1\Gamma_2} A_\mu A_\nu \right| 0 \right\rangle\right]_1 \approx \frac{N_c \alpha_s (\pi\mu^2)^\epsilon}{2\pi} \Gamma[1-\epsilon] \frac{1}{\epsilon^2} \times \Big((-s)^\epsilon + (-t)^\epsilon - (2^\epsilon - 1)\left[(-t)^\epsilon + (-s)^\epsilon\right] \Big) + \text{finite} + \mathcal{O}(\epsilon) \tag{372}$$

$$\left[\left\langle 0 \left| \oint_{\Gamma_1\Gamma_2^{-1}} A_\mu A_\nu \right| 0 \right\rangle\right]_2 \approx -\frac{N_c \alpha_s (\pi\mu^2)^\epsilon}{2\pi} \Gamma[1-\epsilon] \frac{1}{\epsilon^2} \times \Big((-s)^\epsilon + (-t)^\epsilon - (2^\epsilon - 1)\left((-t)^\epsilon + (-s)^\epsilon\right) \Big) + \text{finite} + \mathcal{O}(\epsilon) \tag{373}$$

$$\left[\left\langle 0 \left| \oint_{\Gamma_1\Gamma_2^{-1}} A_\mu A_\nu \right| 0 \right\rangle\right]_3 \approx \frac{N_c \alpha_s (\pi\mu^2)^\epsilon}{2\pi} \Gamma[1-\epsilon] \frac{1}{\epsilon^2} \times \Big((-s)^\epsilon - 2(-s)^\epsilon(2^\epsilon - 1) + (-s)^\epsilon(2^\epsilon - 1)^2 \Big) + \text{finite} + \mathcal{O}(\epsilon) \tag{374}$$

$$\left[\left\langle 0 \left| \oint_{\Gamma_1\Gamma_2} A_\mu A_\nu \right| 0 \right\rangle\right]_4 \approx -\frac{N_c \alpha_s (\pi\mu^2)^\epsilon}{2\pi} \Gamma[1-\epsilon] \frac{1}{\epsilon^2} \times \Big((-s)^\epsilon - 2(-s)^\epsilon(2^\epsilon - 1) + (-s)^\epsilon(2^\epsilon - 1)^2 \Big) + \text{finite} + \mathcal{O}(\epsilon). \tag{375}$$

Applying now the generalized area variation operator (362) followed by the logarithmic mass scale operator

$$\mu\frac{d}{d\mu},$$

to the considered loop configurations, and then taking the $\epsilon \to 0$ limit results in

$$\lim_{\epsilon\to 0}\mu\frac{d}{d\mu}\frac{d\ln\left(\mathcal{W}_1(\Gamma)\right)_1}{d\ln\sigma} = -4\Gamma_{cusp} \tag{376a}$$

$$\lim_{\epsilon\to 0}\mu\frac{d}{d\mu}\frac{d\ln\left(\mathcal{W}_1(\Gamma)\right)_2}{d\ln\sigma} = -12\Gamma_{cusp} \tag{376b}$$

$$\lim_{\epsilon\to 0}\mu\frac{d}{d\mu}\frac{d\ln\left(\mathcal{W}_1(\Gamma)\right)_3}{d\ln\sigma} = -6\Gamma_{cusp} \tag{376c}$$

$$\lim_{\epsilon\to 0}\mu\frac{d}{d\mu}\frac{d\ln\left(\mathcal{W}_1(\Gamma)\right)_4}{d\ln\sigma} = -10\Gamma_{cusp}, \tag{376d}$$

for the LO where

$$\Gamma_{cusp} \approx \frac{\alpha_s C_F}{\pi} + \mathcal{O}(\alpha_s^2). \tag{377}$$

These results seem rather strange considering that naively counting the number of cusps in the different configurations yields a total of eight, in each configuration. If this would be true, the results contradict our conjecture (367) [161]: indeed, we would expect for all the configurations a value of $-8\Gamma_{\text{cusp}}$ in the RHS. So how can we then interpret the result from Eq. (376)? To understand this apparent contradiction, we need to take a closer look at how to count the number of **cusps effectively present** in the studied loop. Let us start with the contour shown in the right panel of Figure 62. Due to path reduction property of GLS [16, 107, 108, 111], this contour can be reduced to a single quadrilateral such that there are **effectively only four cusps** (hence the numbers shown in the right panel of figure 62). The numbers in the left hand panel of figure 62 demonstrate how to count for that contour, where the counting is motivated by the fact that the middle line is crossed twice in the same direction such that the color flow sees the attached cusps twice. This then results in a total of two times four cusps for the middle line[3]. Similar reasoning applies to the contours of figure 63, where we also indicated the counting.

3 Four because the cusps left and right of the middle path show up in the calculations of the interference contribution.

We have thus shown **our conjecture to be valid at LO** for the considered contours and variation, **with a caveat on how to count the number of cusps along the contour.**

22.11 GENERALIZED DERIVATIVE IS A FRÉCHET DERIVATIVE

In [49] we showed that the **generalized derivative** we introduced in [161] (see Section 22.6) coincides with a special case of the **Fréchet derivative** (Definition 17.2.2), at least at LO. From the geometric point of view this is not really a surprise since they generate the **same shape variation of the contour.** The shape variation generated by the generalized derivative is clear, so that if we want it to coincide with the Fréchet derivative we need to find the **associated diffeomorphism generating vector field.** In the case of our quadrilateral on the light cone this vector field is very easy to find and visualize (Figure 66). To demonstrate that the generalized derivative indeed coincides

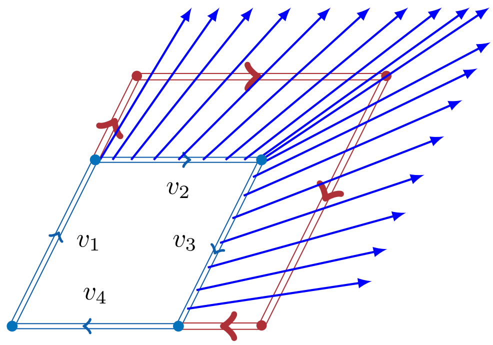


Figure 66: Variational vector field associated to the generalized derivative

with the Fréchet derivative associated to the vector field shown in figure 66 we start from the expression derived in [16] for the **Fréchet derivative applied to the Wilson Loop functional** (Eq. (272)) repeated here for convenience

$$D_V U(\gamma) = U_\gamma . \Big(\int_0^1 dt U_{\gamma_0^t} \mathcal{F}_{\mu\nu}(\gamma(t)) \big(V(t) \wedge \dot{\gamma}(t)\big) \cdot U_{(\gamma^t)^{-1}} \Big), \qquad (378)$$

where γ^t is the part of the contour where the path parameter is between 0 and t. $U(\gamma)$ can be expanded in the coupling constant g, such that we can compare the result of the Fréchet derivative with our generalized derivative order by order in g. In the case both derivatives give the same result, this would support our conjectured evolution equation [161] since we have a series expansion in the coupling constant on the LHS and RHS of Eq. (367) due to the fact that the **cusp anomalous dimension also has a series expansion in the coupling constant**.

To show the LO correspondence we need to extract the terms from Eq. (272) that contribute to this order. These contributions can written as

$$\begin{aligned}
D_V[\mathcal{W}_\gamma]_{\mathrm{LO}} = \mathbb{1} \cdot \oint_0^1 dt \Bigg[\Bigg(\oint_0^t \mathcal{A}_\sigma(x(s)) \frac{dx^\sigma}{ds} ds \\
\times \left(\partial_\mu \mathcal{A}_\nu(y(t)) - \partial_\nu \mathcal{A}_\mu(y(t)) \right) \left(V^\mu(y(t)) \wedge \dot{\gamma}^\nu(y(t)) \right) \cdot \mathbb{1} \Bigg) \\
- \Bigg(\mathbb{1} \cdot \left(\partial_\mu \mathcal{A}_\nu(y(t)) - \partial_\nu \mathcal{A}_\mu(y(t)) \right) \left(V^\mu(y(t)) \wedge \dot{\gamma}^\nu(y(t)) \right) \\
\times \oint_0^t \mathcal{A}_\lambda(x(u)) \, \frac{dx^\lambda}{du} du \Bigg) \Bigg] \\
+ \oint_0^1 \mathcal{A}_\sigma(x) \frac{dx^\sigma}{ds} ds \cdot \oint_0^1 dt \, \mathbb{1} \cdot \Bigg[\left(\partial_\mu \mathcal{A}_\nu(y(t)) - \partial_\nu \mathcal{A}_\mu(y(t)) \right) \\
\times \left(V^\mu(y(t)) \wedge \dot{\gamma}^\nu(y(t)) \right) \cdot \mathbb{1} \Bigg] ,
\end{aligned} \tag{379}$$

where the term with the minus in the first contribution originates from the inverse path $U_{(\gamma_0^t)^{-1}}$. In order to calculate **the vacuum expectation value** of this contribution we need the Wick contractions of the fields, leading to the propagators. We point out that the partial derivatives in Eq. (379) are with respect to the coordinate y. It is important to keep this in mind during the calculations. To simplify our calculations we split the variation generating vector field up into a sum of two vector fields, each one parallel to one of the light cone directions

$$V^\mu := (v_1^+ \sigma, 0^-, \mathbf{0}_\perp) + (0^+, v_2^- \sigma, \mathbf{0}_\perp) \ , \ \sigma \in [0, 1].$$

We also redefine the Mandelstam variable s and t we used before in the following way

$$S_{12}\frac{\delta}{\delta S_{12}} = (2v_1 \cdot v_2)\frac{\delta}{\delta(2v_1 \cdot v_2)} = v_1^+ \frac{\delta}{\delta v_1^+} \quad (380)$$

$$S_{23}\frac{\delta}{\delta S_{23}} = (2v_2 \cdot v_3)\frac{\delta}{\delta(2v_2 \cdot v_3)} = v_2^- \frac{\delta}{\delta v_2^-} , \quad (381)$$

with S_{ij} the adapted Mandelstam variables associated with the Wilson loop (with parametrization shown in figure 66). Restricting now the vector field to its first term

$$V^\mu := (v_1^+\sigma, 0^-, \mathbf{0}_\perp) \ , \ \sigma \in [0,1]$$

generates the shape variation we associated in our original paper to δs. With these notations and definitions we can now start discussing the **LO contribution of the Fréchet derivative.**

There are quite some terms in this contribution but as we will now show, that number can be reduced strongly. First of all **path-reduction eliminates** the terms between the square brackets in Eq. (379), such that we only need to worry about the last term. Secondly the **anti-symmetry of the wedge product** further restricts the contributions in this last term

- Along v_1:
$$V^\mu \wedge \dot{\gamma}^\nu = 0,$$
follows from the asymmetry of the wedge product and the fact that both vectors are parallel
- Along v_2:
$$V^\mu \wedge \dot{\gamma}^\nu = -v_1^+ v_2^- (\partial_+ \wedge \partial_-),$$
due to (anti-)linearity of the wedge product
- Along v_3:
$$V^\mu \wedge \dot{\gamma}^\nu = 0,$$
follows from the asymmetry of the wedge product and the fact that both vectors are parallel
- Along v_4:
$$V^\mu \wedge \dot{\gamma}^\nu = 0,$$
because we assume the vector field to be zero along this part of the path.

As a result we only need to consider the following Wick contractions

- $\overbrace{A^a_\sigma(x(\sigma))\partial_\mu A^b_\nu}(y(\sigma')) = \partial_\mu D^{ab}_{\sigma\nu}(x-y) = \delta^{ab}\partial_\mu D_{\sigma\nu}(x-y)$
- $\overbrace{A^a_\sigma(x(\sigma))\partial_\nu A^b_\mu}(y(\sigma')) = \partial_\nu D^{ab}_{\sigma\mu}(x-y) = \delta^{ab}\partial_\nu D_{\sigma\mu}(x-y)$,

with the remark that y is restricted to the top line in the diagram shown in Fig. 66. Each of these Wick contractions gives rise to four terms, one for each side of the quadrilateral in Fig. 66, which we will now calculate. To reduce the amount of typing we introduce the notation

$$K_\epsilon := \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon), \tag{382}$$

and we use the Feynman gauge, such that we have

$$\left\langle 0 \left| T[A^a_\mu(x)A^b_\nu(y)] \right| 0 \right\rangle = D^{ab}_{\mu\nu}(x-y) = \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)\ \frac{g_{\mu\nu}\delta^{ab}}{[-(x-y)^2]^{1-\epsilon}}\ , \tag{383}$$

for the propagator.

22.11.1 $\partial_\mu D_{\sigma\nu}(x-y) - \partial_\nu D_{\sigma\mu}(x-y)$ *term with* $x \in v_1$

Parametrizing the paths for x and y as (assuming that $x_1 = 0$):

$$x = \sigma v_1,\ \sigma \in [0,1] \tag{384}$$

$$y = v_1 + \sigma' v_2,\ \sigma' \in [0,1]\ , \tag{385}$$

we have:

$$\begin{aligned}
dx^\sigma &= \left(\frac{dx^\sigma}{d\sigma}\right) d\sigma = (v_1^+, 0^-, \mathbf{0}_\perp)d\sigma \\
dy^\nu &= \left(\frac{dy^\nu}{d\sigma'}\right) d\sigma' = (0^+, v_2^-, \mathbf{0}_\perp)d\sigma' = \dot\gamma(\sigma')d\sigma' \\
x-y &= (\sigma-1)v_1 - \sigma' v_2 \\
(x-y)^2 &= -2(\sigma-1)\sigma'\ (v_1^+ v_2^-)\ .
\end{aligned}$$

Calculating this contribution

$$
\begin{aligned}
&\int_0^1 d\sigma'\, d\sigma \frac{dx^\rho}{d\sigma}\left(\frac{\partial}{\partial y^\mu} D_{\rho\nu}(x-y) - \frac{\partial}{\partial y^\nu} D_{\rho\mu}(x-y)\right)\left[V^\mu(y)\wedge\dot{\gamma}^\nu(y)\right] \\
&= K_\epsilon \int_0^1 d\sigma'\, d\sigma \frac{dx^\rho}{d\sigma}\left[\left(\frac{dy^\nu}{d\sigma'}\frac{2(\epsilon-1)g_{\rho\nu}(x-y)_\mu V^\mu(\sigma')}{\left[-(x-y)^2\right]^{2-\epsilon}}\right)\right. \\
&\qquad\qquad \left. - \left(\frac{dy^\nu}{d\sigma'}\frac{2(\epsilon-1)g_{\rho\mu}(x-y)_\nu V^\mu(\sigma')}{\left[-(x-y)^2\right]^{2-\epsilon}}\right)\right] \\
&= K_\epsilon \left[\left(\frac{(1-\epsilon)}{2}(-S_{12})^\epsilon \int_0^1 \frac{d\sigma\; d\sigma'}{\sigma'^{1-\epsilon}(\sigma-1)^{2-\epsilon}}\right)\right. \\
&\qquad\qquad \left. - \left(\frac{(1-\epsilon)}{2}(-S_{12})^{\epsilon-1}(v_1)^2 \int_0^1 \frac{d\sigma\; d\sigma'}{t^{1-\epsilon}(s)^{2-\epsilon}}\right)\right] \\
&= \frac{1}{2}K_\epsilon \frac{S_{12}^\epsilon}{\epsilon}\,,
\end{aligned}
\tag{386}
$$

where S_{ij} represents the **Mandelstam variable** for the pair of vectors $v_{i,j}$. Which is exactly the same result as taking the derivative

$$v_1 \frac{\delta}{\delta v_1}$$

of the original integral

$$v_1 \frac{\delta}{\delta v_1} K_\epsilon \oint \frac{g_{\mu\nu}\, dx^\mu\, dy^\nu}{\left(-(x-y)^2\right)^{1-\epsilon}} = v_1 \frac{\delta}{\delta v_1} K_\epsilon \oint \frac{(v_1 v_2)\, d\sigma\, d\sigma'}{\left(-(2v_1v_2(\sigma-1)\sigma')^2\right)^{1-\epsilon}} = \frac{1}{2}K_\epsilon \frac{S_{12}^\epsilon}{\epsilon}\,. \tag{387}$$

22.11.2 $\partial_\mu D_{\rho\nu}(x-y) - \partial_\nu D_{\rho\mu}(x-y)$ *term with* $x \in v_2$

This term is trivial since it reduces to a Self-Energy on the light-cone which in dimensional regularization is formally zero [167].

22.11.3 $\partial_\mu D_{\rho\nu}(x-y) - \partial_\nu D_{\rho\mu}(x-y)$ *term with* $x \in v_3$

Making use of the symmetry

$$2v_1v_2 = -2v_2v_3 = S_{23},$$

where now S_{23} is the second Mandelstam variable (usually written as t) we can write down this contribution immediately

$$\int_0^1 d\sigma' \, d\sigma \frac{dx^\rho}{d\sigma} \left(\frac{\partial}{\partial y^\mu} D_{\rho\nu}(x-y) - \frac{\partial}{\partial y^\nu} D_{\rho\mu}(x-y) \right) [V^\mu(y) \wedge \dot{\gamma}^\nu(y)] = 0 \; . \tag{388}$$

which is again the same as taking the derivative

$$v_1 \frac{\delta}{\delta v_1}$$

since the original integral is **formally independent of v_1 thus resulting in zero**.

22.11.4 $\partial_\mu D_{\rho\nu}(x-y) - \partial_\nu D_{\rho\mu}(x-y)$ *term with* $x \in v_4$

This contribution is actually the most tricky to calculate, where the intricacies of the calculation are hidden in the combination of the integration and derivatives with respect to y. So here we will apply a slightly different approach. Instead of evaluating the integrals we will keep the integrals and show that taking the derivative

$$v_1 \frac{\delta}{\delta v_1}$$

results in the same integrals as when we take the Fréchet derivative. Using the parametrization

$$x = -(1-\sigma)v_4, \;\; \sigma \in [0,1] \; , \tag{389}$$

$$y = v_1 + \sigma' v_2 \; , \;\; \sigma' \in [0,1] \; , \tag{390}$$

we start by splitting up the calculations in the contributions

$$\partial_\mu D_{\rho\nu}(x-y) \qquad \text{and} \qquad -\,\partial_\nu D_{\rho\mu}(x-y).$$

For the first term we proceed as before resulting in

$$\int_0^1 d\sigma' \, d\sigma \frac{dx^\rho}{d\sigma} \left(\frac{\partial}{\partial y^\mu} D_{\rho\nu}(x-y) \right) [V^\mu(y) \wedge \dot{\gamma}^\nu(y)] =$$

$$-\,2(\epsilon-1) \int_0^1 d\sigma' \, d\sigma \;\; [v_1 \cdot (v_1 + \sigma' v_2 + (1-\sigma)v_4)] \, \frac{(v_2 \cdot v_4)}{(-(-(1-\sigma)v_4 - v_1)^2)^{2-\epsilon}} \; , \tag{391}$$

the second term is the tricky one. If we look at the index of the derivative with respect to y (i.e. ν) one can see that afterwards we integrate again over dy^ν, such that we might as well evaluate the original kernel

$$\frac{1}{\left(-(x-y)^2\right)^{1-\epsilon}}$$

between its boundary values as one would do by a normal integration. This results in

$$-\int_0^1 d\sigma'\, d\sigma \frac{dx^\rho}{d\sigma}\left(\frac{\partial}{\partial y^\nu} D_{\rho\mu}(x-y)\right)\left[V^\mu(y)\wedge\dot{\gamma}^\nu(y)\right] =$$
$$-\int_0^1 d\sigma (v_1\cdot v_4)\sigma'\left[\frac{1}{\left(v_1+v_2+(1-\sigma')v_4\right)^{2(1-\epsilon)}} - \frac{1}{\left(v_1+\sigma' v_4\right)^{2(1-\epsilon)}}\right] = 0 \;, \tag{392}$$

where we used

$$(v_2\cdot v_4) = 0 \qquad \text{and} \qquad v_1 v_2 = -v_1 v_4$$

making the two integrals equal which of course after subtraction results in the zero. Taking the

$$v_1\frac{\delta}{\delta v_1}$$

of the original integral results in

$$v_1\frac{\delta}{\delta v_1}\int_0^1 d\sigma'\, d\sigma \frac{dx^\sigma}{d\rho}\frac{dy^\mu}{d\sigma}\left(D_{\rho\mu}(x-y)\right) =$$
$$-2(\epsilon-1)\int_0^1 d\sigma'\, d\sigma\, \left[v_1\cdot(v_1+\sigma' v_2+(1-\sigma)v_4)\right]\frac{(v_2\cdot v_4)}{\left(-(v_1+\sigma' v_2+(1-\sigma)v_4)^2\right)^{2-\epsilon}} \;, \tag{393}$$

which is the same as Eq. (391) as desired.

Similar calculations with the variational vector field now chosen $(0^+, v_2^-, \mathbf{0}_\perp)$ and the point y restricted to the side v_3 of the quadrilateral (due to the anti-symmetry of the wedge product) result in the contribution

$$\frac{1}{2}K_\epsilon\frac{S_{23}^\epsilon}{\epsilon} - 2(\epsilon-1)\int_0^1 d\sigma'\, d\sigma\, \left[v_4\cdot(v_4+\sigma' v_1+(1-\sigma)v_3)\right]\frac{(v_1\cdot v_3)}{\left(-(v_4+\sigma' v_1+(1-\sigma)v_3)^2\right)^{2-\epsilon}} \;, \tag{394}$$

with

$$S_{23} = 2(v_2 \cdot v_3).$$

Taking the trace over the color matrices then adds the color factor C_F an using the linearity of the wedge product in the vector field V^μ we have the final result

Result : Equivalence Fréchet and Generalized Derivative (LO)

$$\left(v_1 \frac{\delta}{\delta v_1} + v_2 \frac{\delta}{\delta v_2} \right) \ [\mathcal{W}_\gamma] = D_V \ [\mathcal{W}_\gamma] \ ,$$

with

$$V^\mu = V_1^\mu + V_2^\mu = (v_1^+ \ , \ v_2^- \ , \ \mathbf{0}_\perp).$$

Taking into account the renormalization properties of the light-like Wilson quadrilateral loop [48, 91, 154], we come to our **final result** of this Section.

Result : Evolution Equation

$$\mu \frac{d}{d\mu} \ D_V \ [\mathcal{W}_\gamma] = -\sum \Gamma_{\text{cusp}} \ , \qquad (395)$$

where Γ_{cusp} is again the light-cone cusp anomalous dimension 22.2.

22.12 MAKEENKO-MIGDAL, POLYAKOV AND FRÉCHET: A DERIVATIVE CONNECTION

In this Section we will show that the **connection in loop space** $F_\mu(s, c)$**, introduced by Polyakov** in [146] is in fact **related to the Fréchet derivative**. We have already demonstrated the relation to the area derivative, used by Makeenko, but we will come back to this to show how the Fréchet derivative links Polyakov's approach to the one by Makeenko and Migdal.

Polyakov considers the **Wilson loop variable as a chiral field on loop space** and, inspired by the expression for a pure gauge field

$$A_\mu(x) = g^{-1}(x) \partial_\mu g(x),$$

introduces a connection on loop space by

$$F_\mu(s, C) = \frac{\delta U(C)}{\delta x^\mu(s)} U^{-1}(C). \qquad (396)$$

Polyakov continues by showing that

$$F_\mu(s,C) = \int_0^1 ds U(s,C)\mathcal{F}_{\mu\nu}(x(s))\frac{dx_\nu(s)}{ds}U^{-1}(s,C) \tag{397}$$

which can be written as

$$\frac{\delta U(C)}{\delta x_\mu(s)} = U(C)\int_0^1 ds U(s,C)\mathcal{F}_{\mu\nu}(x(s))\frac{dx_\nu(s)}{ds}U^{-1}(s,C). \tag{398}$$

If we contract this with the variational vector field V_μ we get the **Fréchet derivative!!!** Note that here the infinitesimal variations are not achieved by attaching some randomly chosen infinitesimal area somewhere around the contour, but is generated by the variational vector field. This makes the area variation well defined and thus not introduce additional cusps, in contrast with the usually introduced infinitesimal "squares" associated with area variation. Polyakov also demonstrates that the connection he introduced is actually flat

$$\frac{\delta F_\mu(s,C)}{\delta x^\nu(s_1)} - \frac{\delta F_\nu(s_1,C)}{\delta x^\mu(s)} + [F_\mu(s,C), F_n(S_1,C)] = 0, \tag{399}$$

and that the Yang-Mills equations can be written in a very simple way using this connection

$$\frac{\delta F_\mu(s,C)}{\delta x^\mu(s)} = 0. \tag{400}$$

Taking now the functional derivative of this connection results in

$$\frac{\delta F_\mu(s,C)}{\delta x_\mu(s^1)} + \Big[F_\mu(s^1), F_\mu(s)\Big] = U(s,C)\nabla_\mu F_{\mu\nu}(x(s))U^{-1}(s,C)\frac{dx_\nu(s)}{ds}, \tag{401}$$

where we clearly recognize the derivatives appearing in the Makeenko-Migdal equations.

We end this section with a remark for the more mathematical inclined readers. In [168] Zois relates the Hochschild homology of the associative algebra of differential forms $\Omega(M)$ to the de Rham cohomology of the standard loop space LM (infinite dimensional as we have discussed before) on M. This allows to re-express the variational derivation of Polyakov in a cohomology setting, allowing to study certain properties of loop spaces using cohomology. This approach might be useful when considering a (deformation) quantization approach to quantize GLS.

22.13 CONJECTURE AT NLO - RENORMALIZATION GROUP CORRECTION

In this Section we will investigate the validity of our **conjecture** at the NLO level for the **Pi-shaped contour and quadrilateral contour** from before. We will see that we slightly need to adapt our conjecture to take into account the **running of the coupling constant**, ultimately leading to an **extra contribution from the QCD Beta-function.** We will also discuss how our conjecture simplifies for SYM theory where the β-function disappears.

22.13.1 *Pi-shaped contour*

We start with showing that our conjecture is valid at NLO for the **Pi-shaped contour** with the finite part on the light cone, as before, if we take into account the **running of the coupling constant.** The fact that we need to introduce this modification is not surprising, since Γ_{cusp} depends on the coupling constant and hence is sensitive to its **renormalization** which is described by the β-function that for QCD at LO is given by

$$\beta(g) = -\left(\frac{11}{3} - \frac{2}{3}N_F\right)\frac{g^3}{16\pi^2}. \tag{402}$$

Taking this renormalization into account we propose our **adapted conjecture**

$$\left(\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g}\right)\left(\sum_i s_i \frac{\delta}{\delta s_i}\right)\ln W(\Gamma) = -\sum_{\text{cusps}} \Gamma_{\text{cusp}}. \tag{403}$$

To demonstrate the validity of this conjecture (up to NLO) we use the NLO result for the **Pi-shaped contour** from Korchemsky and Marchesini [155]. Without actually realizing it, they already proved Eq. (403) for this contour in this paper, with a bit of different notation[4]. We will here repeat part of that derivation by using their NLO expression for the Pi-shaped contour, the same strategy will then be followed to demonstrate that our updated conjecture is also valid for the quadrilateral on the light cone.

The renormalized NLO expression for the Pi-shaped contour is given by [155]

$$W_{ren.} = \left(\frac{\alpha}{\pi}\right)C_F\left(-L^2 + L - \frac{5}{24}\pi^2\right) + \left(\frac{\alpha}{\pi}\right)^2 C_F\left(BL^3 + CL^2 + DL + \mathcal{O}(L^0)\right) \tag{404}$$

4 See eq. (4.4) in [155] and the discussion below

where

$$\begin{aligned}
B &= -\frac{11}{18}C_A + \frac{1}{9}N_f,\\
C &= \left(\frac{1}{12}\pi^2 - \frac{17}{18}\right)C_A + \frac{1}{9}N_f,\\
D &= \left(\frac{9}{4}\zeta(3) - \frac{7}{18}\pi^2 - \frac{55}{108}\right)C_A + \left(\frac{1}{18}\pi^2 - \frac{1}{54}\right)N_f,\\
L &= \ln(i(\rho - i0)) + \gamma_{\mathrm{E}},\\
\rho &= (s\mu)\,.
\end{aligned}$$

If we now consider the series coefficients of L in this result, after application of the derivative

$$\left(\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right),$$

it is easy to see that the coefficients of L^3 and L^2 are multiplied with powers of $\left(\frac{\alpha}{\pi}\right)$ higher than two, due to the presence of the β-function as multiplicative factor. The coefficients of L on the other hand come from the L^2 and L^3 term in Eq. (404) by application of

$$\left(\mu\frac{\partial}{\partial\mu}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right)$$

and from

$$\left(\beta(g)\frac{\partial}{\partial g}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right)\left(\frac{\alpha}{\pi}\right)C_F(-L^2).$$

Their total contribution becomes

$$6B - \beta = -\left(\frac{11}{3} - \frac{2}{3}N_F\right) + \left(\frac{11}{3} - \frac{2}{3}N_F\right) = 0, \qquad (405)$$

where we used the notation

$$\beta(g) = \beta\frac{g^3}{16\pi^2}.$$

If our conjecture is now to hold, the constant terms should add up to the NLO expression for Γ_{cusp} (in the associated renormalization scheme). The first contribution to the constant terms originates from

$$\left(\frac{\alpha}{\pi}\right)C_F\left(-L^2\right)$$

and is given by

$$\left(\mu\frac{\partial}{\partial\mu}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right)\left(\frac{\alpha}{\pi}\right)C_F\left(-L^2\right) = -2\left(\frac{\alpha}{\pi}\right)C_F = -2\Gamma^{\mathsf{LO}}_{\mathsf{cusp}}. \tag{406}$$

The second contribution comes from the term

$$\left(\frac{\alpha}{\pi}\right)^2 C_F\left(CL^2\right)$$

which can be written as

$$\begin{aligned}&\left(\mu\frac{\partial}{\partial\mu}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right)\left(\frac{\alpha}{\pi}\right)^2 C_F\left(CL^2\right) = \\ &2C\left(\frac{\alpha}{\pi}\right)^2 C_F = 2\left(\left(\frac{1}{12}\pi^2 - \frac{17}{18}\right)C_A + \frac{1}{9}N_f\right)\left(\frac{\alpha}{\pi}\right)^2 C_F.\end{aligned} \tag{407}$$

Finally the third contribution originates from

$$\left(\frac{\alpha}{\pi}\right)C_F(L)$$

returning

$$\left(\beta(g)\frac{\partial}{\partial g}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right)\left(\frac{\alpha}{\pi}\right)C_F(L) = \frac{1}{2}\beta.$$

The first contribution already contributes in the correct way to the cusp anomalous dimension, so we only need to focus on the second and third term. Adding both contributions and extracting a -2 factor we get

$$\begin{aligned}&-2\left(\frac{\alpha}{\pi}\right)^2 C_F\left(-C - \frac{1}{2}\frac{1}{2}\beta\right) = \\ &-2\left(\frac{\alpha}{\pi}\right)^2 C_F\left(\left(-\frac{1}{12}\pi^2 + \frac{17}{18}\right)C_A - \frac{1}{9}N_f\right) + \left(\frac{11}{12} - \frac{2}{12}N_F\right) = \\ &-2\left(\frac{\alpha}{\pi}\right)^2 C_F\left(C_A\left(\frac{67}{36}(-\frac{1}{12}\pi^2\right) - \frac{5}{18}N_F\right) = -2\,\Gamma^{\mathsf{NLO}}_{\mathsf{cusp}}.\end{aligned} \tag{408}$$

Combining all the contributions shows that indeed Eq. (403) is valid at NLO for the Pi-shaped contour.

Result : Evolution Equation for Pi-shaped contour at NLO

$$\left(\mu\frac{\partial}{\partial\mu}+\beta(g)\frac{\partial}{\partial g}\right)\left(\sum_i s_i\frac{\delta}{\delta s_i}\right)\ln W_{\text{ren}}(\Gamma_{\text{Pi}}) =$$
$$-2\left(\left(\frac{\alpha}{\pi}\right)C_F+\left(\frac{\alpha}{\pi}\right)^2 C_F\left(C_A\left(\frac{67}{36}(-\frac{1}{12}\pi^2\right)-\frac{5}{18}N_F\right)\right) =$$
$$-2\mathbf{\Gamma}_{\text{cusp}} \tag{409}$$

where Γ_{cusp} is the light-cone cusp anomalous dimension given in Eq. (345).

22.13.2 *Quadrilateral*

Before investigating the validity of the conjectured evolution equation for the **Quadrilateral in a QCD setting**, we start first with a simpler case, namely the NLO verification in $\mathcal{N}=4$ SYM theory. To demonstrate the validity of the equation in this case we start from the discussion and main result in [69]. In this paper it is explained that the **Non-Abelian exponentiation Theorem for Wilson loops** [147, 169, 170], in the case of $\mathcal{N}=4$ SYM, allows for the Wilson loop to be expanded as

$$W(C_4)=1+\sum_{n=1}^{\infty}\left(\frac{g^2}{4\pi^2}\right)^n W^{(n)}=\exp\left[\sum_{n=1}^{\infty}\left(\frac{g^2}{4\pi^2}\right)^n c^{(n)}w^{(n)}\right]\,, \tag{410}$$

where $W^{(n)}$ are the **perturbative corrections** to the Wilson loop and $c^{(n)}w^{(n)}$ are given by the contribution to $W^{(n)}$ with the **"maximally non-abelian" color factors** $c^{(n)}$. We know that in lowest orders ($n=1,2,3$) these maximally non-abelian color factors are of the form $c^{(n)}=C_F N^{n-1}$. Combining this information with Eq. (410) we have

$$W^{(1)}=C_F w^{(1)}\,,\qquad W^{(2)}=C_F N w^{(2)}+\frac{1}{2}C_F^2\left(w^{(1)}\right)^2,\qquad \ldots \tag{411}$$

For a detailed discussion of this we refer the reader to [69] and references therein.

Using the above we can write for the (unrenormalized) two-loop expression of the Wilson loop:

$$\ln W(C_4) = \frac{g^2}{4\pi^2} C_F w^{(1)} + \left(\frac{g^2}{4\pi^2}\right)^2 C_F N w^{(2)} + \mathcal{O}(g^6). \tag{412}$$

From previous calculations we have for $w^{(1)}$

$$w^{(1)} = -\frac{1}{\epsilon^2}\left[\left(-x_{13}^2\,\mu^2\right)^{\epsilon} + \left(-x_{24}^2\,\mu^2\right)^{\epsilon}\right] + \frac{1}{2}\ln^2\left(\frac{x_{13}^2}{x_{24}^2}\right) + \frac{\pi^2}{3} + \mathcal{O}(\epsilon)\,, \tag{413}$$

where

$$x_{i,i+2}^2 = (x_{i+2} - x_i)^2 \to (v_i + v_{i+1})^2 = s_i$$

the Mandelstam variables. In [69] Drummond et al derive that $w^{(2)}$ is given by

$$w^{(2)} = \left[s\,\mu^2)^{2\epsilon} + (t\,\mu^2)^{2\epsilon}\right]\left\{\epsilon^{-2}\frac{\pi^2}{48} + \epsilon^{-1}\frac{7}{8}\zeta_3\right\} - \frac{\pi^2}{24}\ln^2\left(\frac{s^2}{t^2}\right) - \frac{37}{720}\pi^4 + \mathcal{O}(\epsilon)\,. \tag{414}$$

Applying our derivative Eq. (362) results in

$$\left(s\frac{\delta}{\delta s} + t\frac{\delta}{\delta t}\right) w^{(2)} = \left[s\,\mu^2)^{2\epsilon} + (t\,\mu^2)^{2\epsilon}\right]\left\{\epsilon^{-1}\frac{\pi^2}{24} + \frac{7}{4}\zeta_3\right\} + \text{Finite}. \tag{415}$$

If we now apply the mass scale differential operator to this and take the $\epsilon \to 0$ limit we finally get

$$\mu\frac{\delta}{\delta\mu}\left(\left(s\frac{\delta}{\delta s} + t\frac{\delta}{\delta t}\right)\right) w^{(2)} = 4 \cdot \frac{\pi^2}{12}, \tag{416}$$

which becomes

$$C_F C_A 4 \cdot \frac{\pi^2}{12}$$

when taking the **Non-Abelian color factors** into account and explicitly pulling out a factor 4. Combining this with the one-loop result (413), taking the correct color factors, we arrive at

Result : Evolution Equation for Quadrilateral (SYM) at NLO

$$\left(s\frac{\delta}{\delta s}+t\frac{\delta}{\delta t}\right)\ln\left(W(\Gamma)\right)=4\cdot\left(\left(\frac{\alpha_s}{\pi}\right)C_F+\left(\frac{\alpha_s}{\pi}\right)^2C_FC_A\frac{\pi^2}{12}\right)=-4\Gamma_{\text{cusp}},$$

consistent with our original conjecture Eq. (367) if one considers Γ_{cusp} as in [69], where one considers pure Yang-Mills (YM) theory (i.e. not considering fermions or ghost in the gluon Self-Energy corrections or put differently $N_F = 0$ in Eq. (345)).

We would now like to expand this result to the case of QCD. To this end we proceed, in the same way as for the Pi-shaped contour, now using the NLO results for the quadrilateral on the light cone derived by Korchemsky et al in [156]. In this paper it is shown that using the Non-Abelian exponentiation Theorem [147, 169, 170] the renormalized quadrilateral Wilson loop can be written as

$$W(s\mu^2,t\mu^2,g)=\exp\left(W^{\text{one-loop}}+W^{\text{two-loop}}\right), \tag{417}$$

where

$$W^{\text{one-loop}} = -\frac{\alpha_s}{2\pi}C_F\left(L^2\left(s\mu^2\right)+L^2\left(t\mu^2\right)\right) \tag{418}$$

$$\begin{aligned}W^{\text{two-loop}} = &-\left(\frac{\alpha_s}{\pi}\right)^2C_F\Big[w_1L^3\left(s\mu^2\right)+w_2L^2\left(s\mu^2\right)\\ &+w_3L\left(s\mu^2\right)L\left(t\mu^2\right)+w_4L\left(s\mu^2\right)\\ &+(s\to t)+\text{const}\Big]\end{aligned} \tag{419}$$

and with

$$\begin{aligned}L(x) &= \ln(x)\\ w_1 &= \left(\frac{11}{72}C_A-\frac{N_f}{36}\right)\\ w_2 &= \left(\left(\frac{67}{72}-\frac{\pi^2}{12}\right)C_A-\frac{5}{36}N_F\right)\\ w_3 &= \left(\frac{\pi^2}{24}C_A\right)\\ w_4 &= \left(\left(\frac{101}{54}-\frac{7}{4}\zeta(3)\right)C_A-\frac{7}{27}N_F\right).\end{aligned} \tag{420}$$

Eq. (418) is just another way to express our one loop result in **renormalized version.** To see how they are related write

$$\left(s\mu^2\right)^{\epsilon}$$

as

$$\exp\left(\epsilon \log\left(s\mu^2\right)\right) = 1 + \epsilon \log\left(s\mu^2\right) + \frac{1}{2}\epsilon^2 \log^2\left(s\mu^2\right) + \mathcal{O}\left(\epsilon^3\right)$$

and remember that this is multiplied with a factor

$$\frac{1}{\epsilon^2}$$

in the LO result Eq. (350). **Subtracting the poles** using the $\overline{\text{MS}}$ scheme [19, 52, 171] returns Eq. (418) if one applies the same transformation to

$$\left(t\mu^2\right)^{\epsilon}.$$

We point out that **applying our derivative followed by the mass scale derivative** to (418) gives again our conjecture at leading order

Result : Evolution Equation for Renormalized Quadrilateral (QCD)

$$\mu\frac{\delta}{\delta\mu}\left(\left(s\frac{\delta}{\delta s} + t\frac{\delta}{\delta t}\right)\left(W^{\text{one-loop}}\right)\right) = -8\frac{\alpha_s}{2\pi}C_F = -4\Gamma^{\text{LO}}_{\text{cusp}}. \tag{421}$$

For the NLO result Eq. (419) we follow the same strategy as for the Pi-shaped contour, where again we adapted the conjecture to Eq. (403). Simple calculation then shows that the L^2**-terms after application of the generalized or Fréchet derivative only contribute to higher orders of** $\frac{\alpha}{\pi}$, i.e. to Next-to-Next-to-Leading Order (NNLO) terms. Similarly to the Pi-shaped contour case the terms contributing to the L-terms cancel

$$\frac{1}{2}\beta + 2\cdot 6w_1 = -\frac{1}{2}\left(\frac{11}{3} - \frac{2}{3}N_F\right) + 12\left(\frac{11}{72}C_A - \frac{N_f}{36}\right) = 0.$$

The total contribution to the constant terms, after application of all the derivatives, is given by

$$
\begin{aligned}
&-4\left(\frac{\alpha}{\pi}\right)^2 C_F(2w_2+2w_3) = \\
&-4\left(\frac{\alpha}{\pi}\right)^2 C_F\left(2\left(\left(\frac{67}{72}-\frac{\pi^2}{12}\right)C_A-\frac{5}{36}N_F\right)+2\left(\frac{\pi^2}{24}C_A\right)\right) = \\
&-4\left(\frac{\alpha}{\pi}\right)^2 C_F\left(\left(C_A\left(\frac{67}{36}-\frac{1}{12}\pi^2\right)-\frac{5}{18}N_F\right)\right) = \\
&-4\Gamma^{\mathrm{NLO}}_{\mathrm{cusp}},
\end{aligned}
\tag{422}
$$

which combined with Eq. (421) demonstrates the validity of our adapted conjecture at the NLO level for the quadrilateral on the light cone. This result is now a second example where our adapted conjecture is valid beyond the LO level.

Result : Evolution Equation Renormalized Quadrilateral (QCD) at NLO

$$
\begin{aligned}
&\mu\frac{\delta}{\delta\mu}\left(\left(s\frac{\delta}{\delta s}+t\frac{\delta}{\delta t}\right)(W(\Gamma_{\mathrm{Quad}}))\right) = \\
&-4\left\{\left(\frac{\alpha}{\pi}\right)C_F+\left(\frac{\alpha}{\pi}\right)^2 C_F\left(\left(C_A\left(\frac{67}{36}-\frac{1}{12}\pi^2\right)-\frac{5}{18}N_F\right)\right)\right\} = \\
&-4\Gamma_{\mathrm{cusp}},
\end{aligned}
\tag{423}
$$

where Γ_{cusp} is the light-cone cusp anomalous dimension given in Eq. (345).

22.14 ALL ORDER CONFIRMATION OF OUR CONJECTURE IN $\mathcal{N}=4$ SYM THEORY

In [69] it is discussed that the **Wilson loop can be split up in a divergent and a finite part** given by

$$
\ln W_n = \ln Z_n + \frac{1}{2}\Gamma_{\mathrm{cusp}}(a)F_n + \mathcal{O}(\epsilon), \tag{424}
$$

where the **divergences are absorbed into the factor** Z_n and depends on the renormalization scale μ, ϵ and the vectors v_i forming the contour. In any gauge theory ([172–174] and references therein) this factor can be written as

$$\ln Z_n = -\frac{1}{2}\sum_{l=1}^{\infty} a^l \left(\frac{\Gamma^{(l)}_{\text{cusp}}}{(l\epsilon)^2} + \frac{\Gamma^{(l)}}{l\epsilon}\right)\sum_{i=1}^{n}(-x^2_{i,i+2}\,\mu^2)^{l\epsilon}, \tag{425}$$

where

$$\Gamma_{\text{cusp}}(a) = \sum_{l=1}^{\infty} a^l \Gamma^{(l)}_{\text{cusp}}$$

and

$$\Gamma(a) = \sum_{l=1}^{\infty} a^l \Gamma^{(l)},$$

where we have redefined the a^l to be consistent with the one-loop result. The term F_n refers to a finite contribution that is parametrized only by the x_i. Using the UV - IR Wilson loop duality, thus interchanging the x_i for generalized Mandelstam variables

$$x^2_{i,i+2} \to s_i$$

results in

$$\ln Z_n = -\frac{1}{2}\sum_{l=1}^{\infty} a^l \left(\frac{\Gamma^{(l)}_{\text{cusp}}}{(l\epsilon)^2} + \frac{\Gamma^{(l)}}{l\epsilon}\right)\sum_{i=1}^{n}(-s_i\,\mu^2)^{l\epsilon}, \tag{426}$$

Applying the generalized derivative of Eq. (362) to Eq. (424) returns

$$\left(\sum_{s_i} s_i \frac{\delta}{\delta s_i}\right)\ln W_n = -\frac{1}{2}\sum_{l=1}^{\infty} a^l \left(\frac{\Gamma^{(l)}_{\text{cusp}}}{(l\epsilon)} + \Gamma^{(l)}\right)\sum_{i=1}^{n}(-s_i\,\mu^2)^{l\epsilon} + \mathcal{O}(\epsilon),$$

and now taking the mass scale derivative of this result returns

$$\mu\frac{\delta}{\delta\mu}\left(\left(\sum_{s_i} s_i \frac{\delta}{\delta s_i}\right)\right)\ln W_n = -\sum_{l=1}^{\infty} a^l \left(\Gamma^{(l)}_{\text{cusp}} + l\epsilon\Gamma^{(l)}\right)\sum_{i=1}^{n}(-s_i\,\mu^2)^{l\epsilon} + \mathcal{O}(\epsilon). \tag{427}$$

In the limit $\epsilon \to 0$ we get the final result

Result : All order Evolution Equation for Quadrilateral in SYM

$$\mu\frac{\delta}{\delta\mu}\left(\sum_{s_i} s_i\frac{\delta}{\delta s_i}\right)\ln W_n = -n\sum_{l=1}^{\infty} a^l\left(\Gamma^{(l)}_{\text{cusp}}\right) = -n\Gamma_{\text{cusp}}, \qquad (428)$$

demonstrating that in SYM theory our conjecture holds to all orders.

Of course **this result depends strongly on the Non-Abelian exponentiation Theorem** and on the behavior of the **Sudakov form factor** in this theory. The question if our adapted conjecture holds in QCD for NNLO when using **renormalized Wilson loops** remains open for the moment.

We point out that from the results in Section 22.13 **we now understand how to transition from our conjecture for QCD to the one for SYM.** Since the **coupling constant does not run in SYM**, the β-function is zero, such that it eliminates all the contributions coming from this term in the derivative. Clearly this reduces our adapted conjecture to the original one, which we showed to be valid to all orders in SYM. This fact, combined with the two loop result gives us confidence that our adapted conjecture for QCD is also valid beyond the NLO level.

22.15 GENERALIZED LOOP SPACE, RENORMALIZATION, RAPIDITY AND TMDS REVISITED

In our papers [49, 175] we demonstrated **the relation between the Fréchet differentiation and rapidity differentiation**. We already briefly explained this in the introduction, we will repeat this discussion here for the readers convenience and relate it to the evolution of a new kind to TMD, namely a TDD.

From the RHS of our adapted conjecture in Eq. (403), keeping the SD approach in mind, it is clear that the **cusp anomalous dimensions** can be interpreted as a **fundamental ingredient of an effective quantum action** for the Wilson loops with simple obstructions. This fact demonstrates a relation between the geometrical properties (given in terms of the area/shape infinitesimal variations and corresponding differential equations) of the GLS, and the **renormalization-group behavior of the Wilson polygons with conserved angles between the light-like straight lines.** Put differently, **the dynamics in loop space is governed by the discontinuities of the path**

derivatives. These obstructions play the role of the sources within the Schwinger field-theoretical picture. We have shown, that the Schwinger quantum dynamical principle [176] is helpful in the investigation of certain classes of elements of the loop space, namely, the cusped Wilson polygons. It is worth noting that Eq. (403) suggests, in fact, **a duality relation between the rapidity evolution of certain correlation functions and the equations of motion in the GLS**. Rapidities associated with the light-like vectors $N^{\pm}$ are, of course, infinite, and are given by

$$Y^{\pm} = \frac{1}{2} \ln \left(\frac{(N^{\pm})^{+}}{(N^{\pm})^{-}} \right) = \lim_{\eta^{\pm} \to 0} \pm \frac{1}{2} \ln \left(\frac{N^{+}N^{-}}{\eta^{\pm}} \right) , \tag{429}$$

where $\eta^{\pm}$ is a cutoff and where we take into account the fact that plus- and minus- components of a vector a_{μ} are defined as the scalar products

$$a^{\pm} = (a \cdot N^{\mp}).$$

Eq. (429) suggests, clearly, that

$$\frac{d}{d \ln \sigma} \sim \frac{d}{dY} , \tag{430}$$

that is, **the rapidity evolution equation of a certain correlation function can be set dual to the area variation law of a properly chosen class of elements of the GLS.** In particular, we will argue below that Eq. (403) can be used to study the evolution of the $3D$-parton distribution functions. To this end, we will focus on the behavior of parton densities in the large Bjorken-x_B approximation.

In this context we will make some assumptions, **we assume that some reasonable TMD-factorization scheme can be formulated,** and that **an appropriate operator definition of the TMD quark distribution function exists** that has the same quantum numbers as the correlation function below, shown in Eq. (431).

We will also be mainly interested in the $3D$-correlation functions in the **large-x_B limit**, this regime is apparently easier to analyze within a simple factorization scheme and perfectly fits the Jefferson Lab 12 GeV kinematics. Furthermore, we will argue that the large-x_B approximation is an ideal natural laboratory for the study of applications of the GLS formalism in hadronic and nuclear physics.

Let us introduce the following TDD correlation function

$$\mathcal{F}\left(x, \boldsymbol{b}_\perp; P^+, n^-, \mu^2\right) = \int d^2k_\perp \ \mathrm{e}^{-ik_\perp \cdot b_\perp} \mathcal{F}\left(x, \boldsymbol{k}_\perp; p^+, n^-, \mu^2\right) =$$
$$\int \frac{dz^-}{2\pi} \Big\langle P \Big| \bar{\psi}(z^-, \boldsymbol{b}_\perp) W^\dagger_{n^-}[z^-, \boldsymbol{b}_\perp; \infty^-, \boldsymbol{b}_\perp] W^\dagger_{\boldsymbol{l}}[\infty^-, \boldsymbol{b}_\perp; \infty^-, \infty_\perp]$$
$$\times \gamma^+ W_{\boldsymbol{l}}[\infty^-, \infty_\perp; \infty^-, \boldsymbol{0}_\perp] \mathcal{W}_{n^-}[\infty^-, \boldsymbol{0}_\perp; 0^-, \boldsymbol{0}_\perp] \psi(0^-, \boldsymbol{0}_\perp) | \ P \Big\rangle \ . \tag{431}$$

This correlation function **allows us to extract information about the quark distribution in the longitudinal one-dimensional momentum space (x, P^+) and in the two-dimensional impact-parameter coordinate space ($\boldsymbol{b}_\perp$).** The semi-infinite Wilson lines occurring in the TDD definition are evaluated along a certain four-vector w_μ and are defined as

$$W_w[\infty; z] \equiv \mathcal{P} \exp\left[-ig \int_0^\infty d\tau \ w_\mu \ \mathcal{A}^\mu(z + w\tau)\right] \ , \tag{432}$$

where now, the vector w_μ can be either **longitudinal**

$$w_\mu = (w_L, \boldsymbol{0}_\perp),$$

or **transverse**

$$w_\mu = (0_L, \boldsymbol{l}_\perp).$$

Notice that the TDD is a **partial Fourier transform** of the standard TMD correlator

$$\mathcal{F}\left(x, \boldsymbol{k}_\perp; P^+, n^-, \mu^2\right).$$

The factorization and evolution of the gauge-invariant collinear PDFs in the large-x_B regime has been studied in [155] and here we propose to generalize this approach to the $3D$-PDF of Eq. (431).

The large-x_B regime requires some assumptions, for which we refer the reader to [175], that allow for the following factorization formula

$$\mathcal{F}\left(x, \boldsymbol{b}_\perp; P^+, n^-, \mu^2\right) = \mathcal{H}(\mu, P^2) \times \Phi(x, \boldsymbol{b}_\perp; P^+, n^-, \mu^2) \ , \tag{433}$$

where the contribution of incoming-collinear partons is summed up into the x-independent function, while the soft function Φ is given by[5]

$$\Phi(x, \boldsymbol{b}_\perp; P^+, n^-, \mu^2) = \\ P^+ \int dx \, \mathrm{e}^{-i(1-x)P^+z^-} \langle 0| \, W_P^\dagger[z; -\infty] W_{n^-}^\dagger[z; \infty] W_{n^-}[\infty; 0] W_P[0; \infty] \, |0\rangle \ , \tag{434}$$

with **two kinds of Wilson lines**

(i) incoming-collinear (non-light-like, $P^2 \neq 0$)

$$W_P ,$$

(ii) outgoing-collinear ($(n^-)^2 = 0$),

$$W_{n^-} .$$

The **associated rapidity and renormalization-group evolution equations** are given by

$$\mu \frac{d}{d\mu} \ln \mathcal{F}\left(x, \boldsymbol{b}_\perp; P^+, n^-, \mu^2\right) = \mu \frac{d}{d\mu} \ln \mathcal{H}(\mu^2) + \mu \frac{d}{d\mu} \ln \Phi(x, \boldsymbol{b}_\perp; P^+, \mu^2) \ , \tag{435}$$

$$P^+ \frac{d}{dP^+} \ln \mathcal{F}\left(x, \boldsymbol{b}_\perp; P^+, n^-, \mu^2\right) = P^+ \frac{d}{dP^+} \ln \Phi(x, \boldsymbol{b}_\perp; P^+, \mu^2) \ , \tag{436}$$

where the rapidity is introduced via

$$\ln P^+$$

with proper regularization [177]. The RHS of Eq. (435) is, in fact, $\boldsymbol{b}_\perp$-independent and contains only a single-log dependence on the rapidity [63–67, 178, 179]. As a consequence Eq. (436) corresponds to the **Collins-Soper-Sterman rapidity-independent kernel $\mathcal{K}_{\mathrm{CSS}}$.**

We are now ready to use of the **evolution equation Eq. (403).** We emphasize that the soft function $\mathcal{F}$ is a Fourier transform of an element of the GLS, the Wilson loop evaluated along the path, defined in Eq. (434). This enables us to consider the shape variations of this path, which are generated by

5 For simplicity, we work in covariant gauges, so that the transverse Wilson lines at infinity can be ignored.

the infinitesimal variations of the rapidity variable $\ln P^+$. The corresponding differential operator reads:

$$\frac{d}{d \ln \sigma} \sim P^+ \frac{d}{dP^+} \ , \tag{437}$$

given that

$$dP^+ = (dP \cdot n^-).$$

Therefore

Result : Evolution Equation for TDDs

$$\begin{aligned}
&\left(\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g}\right) \left(P^+ \frac{d}{dP^+} \ln \mathcal{F}\right) = \\
&\left(\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g}\right) \left(P^+ \frac{d}{dP^+} \ln \Phi\right) = \\
&- \sum_{\text{TDD}} \Gamma_{\text{cusp}}(\alpha_s) = \\
&\left(\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g}\right) \mathcal{K}_{\text{CSS}}(\alpha_s) \ .
\end{aligned} \tag{438}$$

Eqs. (435-438) should now be integrated to give a complete evolution of the TDD (431) in the large-x_B region. **A calculation we are still working on.** We end this chapter with pointing out that these evolution equations can be directly applied to the JLab 12 GeV phenomenology (see also [180] and references therein), again motivating the research done in this thesis.

22.16 SUMMARY

In this Chapter we studied the effect of the different differential operators, defined on GLS, on simple Wilson loops, the quadrilateral on the light cone and the Pi-shaped contour with a finite part on the light cone. Since the commonly used differential operators, like the area derivative, did not work on this quadrilateral we were forced to introduce a new differential operator. This new operator turned out to be a special case of the Fréchet derivative, making it mathematically well defined. Applying this "new" operator in combination with the usual renormalization mass scale differential operator to the

Wilson loop quadrilateral resulted in an equation that is quite simple and only seems to depend on the number of cusps along the contour and on the cusp anomalous dimension of the underlying theory. As a conjecture we generalized this result to Wilson loop polygons with n cusps, and tested this conjecture for a Pi-shaped contour for which it also turned out to be valid. Furthermore we demonstrated that our conjecture is not only valid at LO but also at NLO for QCD and to all orders in $\mathcal{N} = 4$ SYM theory where a huge simplification occurs due to the vanishing of the Beta-Function. We ended the Chapter with deriving a relation between our conjecture for Wilson loops and the renormalization group equations for a three dimensional parton density correlator, the TDD correlation function, introducing an evolution equation for them. These correlators can be immediately applied to the JLab 12 GeV phenomenology, making our conjecture testable in the upcoming JLab experiments and in future experiments at the EIC, RHIC and CERN.

23 THESIS SUMMARY AND OUTLOOK

23.1 SUMMARY

After giving a short introduction and motivation for the research done in this thesis in the first Part, we reviewed the basic mathematical concepts in the second Part. These mathematical concepts are needed to introduce Generalized Loop Space and its properties. More specifically we investigated different differential operators that generate different shape variations of the loops in the Generalized Loop Space.

A natural candidate for such shape variations was the area derivative that was used by Makeenko and Migdal to derive their famous loop equations. This operator had some issues with respect to the introduction of extra cusps, complicating its application to Wilson loops. Fortunately this is not the only well-defined shape variation inducing differential operator. Another, more complicated, differential operator is the Fréchet derivative associated to diffeomorphisms of the underlying base manifold on which the Generalized Loop Space is built.

The Fréchet derivate was then applied to different Wilson loops, relevant for TMD phenomenology. This led to our main results, which we summarize in the section below.

23.2 OVERVIEW OF MAIN RESULTS

As a first application of Generalized Loop Space we applied both the area derivative and the Fréchet derivative to a simple Wilson loop quadrilateral lying on the light cone, and demonstrated that indeed the area derivative is ill-defined due to the extra cusps it generates along the contour. It forced us to define an alternative differential operator that we showed to be equivalent with the Fréchet derivative associated to a specific diffeomorphism of the base manifold, at least at leading order. This "new" derivative, in combination with the well-known renormalization mass scale differential operator, allowed us to derive an evolution equation for these simple contours.

Inspired by this result we wanted to know if our evolution equation would also hold for other contours. Therefore we tested the evolution equation at leading order for the Π-shaped contour and for four contours which are symmetric extensions of the original quadrilateral contour.

At the moment our evolution equation seems to hold for all these contours, even for the contours with overlapping paths and self-intersections if one counts the number of cusps in the correct manner. Furthermore, these calculations also demonstrated that the group structure of loop space is able to handle contours with overlapping paths and self-intersections, a fact that is not so obvious.

Motivated by the above successes we conjectured an evolution equation for Wilson loops by generalizing the result for the quadrilateral. An important property of this evolution equation is that it only depends on the number of cusps along the contour and on the cusp anomalous dimension, a quantity that also shows up frequently when studying three dimensional parton distributions.

Having only a leading order result we then moved on to check if our conjecture also holds at higher orders (in α_s, the coupling constant). We checked that indeed our conjecture holds for the quadrilateral lying on the light cone to second order in $\mathcal{N} = 4$ SYM theory and in QCD, with the remark that in QCD one needs to take into account the running of the coupling constant. This dependence on running of the coupling constant is represented in our conjecture formula by an extra term constructed from the derivative with respect to the coupling constant multiplied with the β-Function acting on the Fréchet derivative of the logarithm of the Wilson loop under consideration. This is in contrast with our first result at leading order, where the term with β-function was not present. As a consequence we had to modify our evolution equation, adding this extra term. In the case of $\mathcal{N} = 4$ SYM this term is not present, as the coupling constant does not run in this theory, simplifying the evolution equation to our original conjecture.

As a direct consequence, we were able to show that in this special case of $\mathcal{N} = 4$ SYM theory our conjecture holds at all orders (under assumption of the Non-Abelian exponentiation theorem), giving us confidence that our conjecture also holds for higher orders in QCD.

Having JLab phenomenology in mind as an application for our evolution equation, we demonstrated a connection between rapidity evolution and the evolution induced by the Fréchet derivative. Through this connection it was then possible to propose an evolution equation for the introduced Transverse

Distance Dependent (TDD) three dimensional correlator, relevant for large x_B phenomenology **and** testable at the 12 GeV JLab upgrade.

The confidence in our TDD evolution is reinforced by the fact that it is remarkably similar to the Collin-Soper-Sterman evolution equation, which is known to be not completely exact. So our hope is that our evolution equation will do better and give a (more) accurate description for transverse momentum/distance dependent correlators.

23.3 OUTLOOK AND FUTURE RESEARCH

There are many opportunities to extend the research of this thesis, which are not necessarily only related to the main results of the thesis but also to the mathematical techniques and concepts used. We will try to give an overview of possible future research, splitting it up in research that builds on the results of this thesis on the one hand and in research that makes use of the concepts and mathematical techniques on the other.

Let us first consider the research that builds on the results of this thesis. A natural continuation of the research presented in this text would be to test our conjecture for different contours. A simple example would be to consider a quadrilateral with only three of its four sides on the light cone. This is an interesting case since such contours emerge in the soft factors that appear in the factorization of TMDs, contributing to the non-perturbative effects. Another example is the Π-shaped contour with transverse separation, which appears both as a soft-factor and as a Wilson line structure in TDDs.

In another line of research, one could try to reconstruct a quantum action on loop space by making use of the SD approach in combination with our conjecture, opening the door to non-perturbative calculations and to study an alternative representation of gauge theory. Associated with this, one should try to find a map between the gauge theory representation for observables and the loop space representation. Note that this will require some supergeometrical structure in order to be able to describe fermions.

On the diffeomorphism side, we have only considered angle preserving or conformal transformations. Extending to non-conformal diffeomorphisms can give more general evolution equations, where one would also get an additional evolution with respect to the cusp angles. It would be fascinating to see how such transformations would affect our conjectured evolution, and how the cusp anomalous dimensions behave in this situation.

On a more mathematical side it would be very interesting to study the cohomology of generalized loop space and how it responds to changes in topology of the base manifold, which was suggested to us by professor B. Shoikhet[1]

As a research topic which uses concepts and techniques from this thesis we mention the proposal by professor P. Mulders[2] to use the shuffle product, emerging when calculating corrections to the Drell-Yan process, to sum over all possible configurations to add gluons to interactions between three semi-infinite Wilson lines.

We also mention that, at the MENU13 conference in Rome, Dr. D.W. Sivers[3] suggested to use the generalized loop space structures to calculate the anomalous magnetic moment for electrons, and at "High-Energy physics in the LHC era"-conference in Chile we discussed the possibility of using generalized loop space to access the non-perturbative sector of the dualized Standard Model he presented there.

Recently the shuffle product also seems to emerge more often in calculations of scattering amplitudes in the context of recurrence relations, in the hopes of reducing the number of Feynman diagrams to be calculated when considering higher order corrections. Perhaps insights from this thesis might also help in this area of research.

One issue with Generalized Loop Space that has been haunting us during the writing of this thesis is the fact that there is no conclusive prove that the physical content of a gauge theory in a quantum field background can be recast in a loop space setting. A natural research topic that follows from this would be to investigate how to quantize Generalized Loop Space, where as a first attempt we would consider geometric quantization. This suggestion is motivated by the algebraic structure of Generalized Loop Space, showing many similarities to situations we encountered in papers on this subject.

As a final future research topic we mention that similar loop spaces have been used in the study of quantum loop gravity, such that also here there might be opportunities to explore the power of Wilson loops (see for instance [8] as a simple example).

---

1 Mathematics Department University of Antwerp.
2 NIKHEF
3 University of Michigan

Part V

APPENDIX

A NOTATIONS, CONVENTIONS AND DEFINITIONS

A.1 SPINORS AND GAMMA MATRICES

Any field with half-integer spin, i.e. a Dirac field, anti-commutes:

$$\left\{\psi_a(x), \psi_b^\dagger(y)\right\} = \delta(x-y)\delta_{ab},\ \{\psi_a(x), \psi_b(y)\} = \left\{\psi_a^\dagger(x), \psi_b^\dagger(y)\right\} = 0. \tag{439}$$

We define gamma matrices by the anti-commutation relations

$$\{\gamma^\mu, \gamma^\nu\} \equiv 2\, g^{\mu\nu}\, \mathbb{1}, \tag{440}$$

with the following additional property

$$(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0. \tag{441}$$

Although in particle physics we talk about γ matrices, these object are much more general as generators of a Clifford Algebra (here the algebra of space-time). The four-dimensional matrices used in particle physics are just a four dimensional representation, which can be extended to higher dimensions (see for instance [181]). A mentionable downside of the matrix representations for these objects is the fact that the pseudo-scalar of the Clifford Algebra of three-dimensional space (usually represented by the Pauli sigma matrices) and four-dimensional space-time are the same

$$\sigma_1\sigma_2\sigma_3 = \gamma_0\gamma_1\gamma_2\gamma_3,$$

which cannot be written down explicitly [182] with matrices. This as a side note. Moving on to the Dirac equation for a particle field ψ

$$(i\not{\partial} - m)\,\psi = 0, \tag{442}$$

where we also introduced the slash notation for

$$\not{p} = \gamma^\mu p_\mu. \tag{443}$$

We can identify an antiparticle field with $\bar{\psi}$ if we define

$$\bar{\psi} = \psi^\dagger\gamma^0, \tag{444}$$

which satisfies a slightly adapted Dirac equation:

$$(i\overleftarrow{\not\partial} + m)\,\bar{\psi} = 0. \tag{445}$$

We can expand Dirac fields in function of a set of plane waves:

$$\psi_1(x) = u^{\sigma}(p)\,\exp\left[-ip\cdot x\right] \qquad (p^2 = m^2, p^0 > 0), \tag{446}$$

$$\psi_2(x) = v^{\sigma}(p)\,\exp\left[+ip\cdot x\right] \qquad (p^2 = m^2, p^0 < 0), \tag{447}$$

where σ is a spin-index. If we define

$$\bar{u} = u^{\dagger}\gamma^0, \qquad \bar{v} = \gamma^0 v^{\dagger}, \tag{448}$$

we can find the completeness relations by summing over spin:

$$\sum_s u^s(p)\bar{u}^s(p) = \not{p} + m, \tag{449}$$

$$\sum_s \bar{v}^s(p) v^s(p) = \not{p} - m. \tag{450}$$

We will identify

- u with an incoming fermion,
- $\bar{u}$ with an outgoing fermion,
- $\bar{v}$ with an incoming anti-fermion,
- v with an outgoing anti-fermion.

If we define

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{i}{4!}\varepsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma, \tag{451}$$

$$\gamma^{\mu\nu} = \gamma^{[\mu}\gamma^{\nu]} = \frac{1}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right), \tag{452}$$

we can construct a complete Dirac basis:

$$\mathbb{1},\, \gamma^\mu,\, \gamma^{\mu\nu},\, \gamma^\mu\gamma^5,\, \gamma^5. \tag{453}$$

We will identify

- $\mathbb{1}$ with a scalar,
- γ^μ with a vector,

- $\gamma^{\mu\nu}$ with a tensor,
- $\gamma^{\mu}\gamma^{5}$ with a pseudo-vector,
- γ^{5} with a pseudo-scalar.

Furthermore, γ^5 has the following properties:

$$\left(\gamma^5\right)^\dagger = \gamma^5, \qquad \left(\gamma^5\right)^2 = \mathbb{1}, \qquad \left\{\gamma^5, \gamma^\mu\right\} = 0. \tag{454}$$

Let's list some contraction identities for gamma matrices in ω dimensions:

$$\gamma^\mu\gamma_\mu = \omega, \tag{455}$$

$$\gamma^\mu\gamma^\nu\gamma_\mu = (2-\omega)\gamma^\nu, \tag{456}$$

$$\gamma^\mu\gamma^\nu\gamma^\rho\gamma_\mu = 4\, g^{\nu\rho} + (\omega - 4)\gamma^\nu\gamma^\rho, \tag{457}$$

$$\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_\mu = -2\gamma^\sigma\gamma^\rho\gamma^\nu + (4-\omega)\gamma^\nu\gamma^\rho\gamma^\sigma. \tag{458}$$

And some trace identities:

$$\text{Tr}[\mathbb{1}] = \omega, \tag{459}$$

$$\text{Tr}[\text{odd number of } \gamma\text{'s}] = 0, \tag{460}$$

$$\text{Tr}[\gamma^\mu\gamma^\nu] = 4\, g^{\mu\nu}, \tag{461}$$

$$\text{Tr}[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma] = 4\left(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}\right) \tag{462}$$

A.2 LIGHT-CONE COORDINATES

In QCD it is common the work with so-called light-cone coordinates which are defined for a four-vector a^μ by

$$a^+ = \frac{1}{\sqrt{2}}\left(a^0 + k^3\right), \tag{463}$$

$$a^- = \frac{1}{\sqrt{2}}\left(a^0 - k^3\right), \tag{464}$$

$$\boldsymbol{a}^\perp = \left(a^1, a^2\right). \tag{465}$$

We will represent the plus-component first, i.e.

$$a^\mu = \left(a^+, a^-, \boldsymbol{a}^\perp\right). \tag{466}$$

The factor $\frac{1}{\sqrt{2}}$ is a normalization factor such that the Jacobian is unity, such that

$$d^4a = da^+ da^- d\boldsymbol{a}^\perp. \tag{467}$$

Using light cone coordinates the interior product can be expressed as

$$a\cdot\ b = a^+b^- + a^-b^+ - \boldsymbol{a}^\perp\cdot\boldsymbol{b}^\perp, \tag{468}$$

$$a^2 = 2a^+a^- - \left(\boldsymbol{a}^\perp\right)^2. \tag{469}$$

This implies that the metric becomes off-diagonal

$$g^{\mu\nu}_{\mathsf{LC}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{470}$$

Next we also define two light-like basis-vectors

$$n^\mu_+ = \left(1^+, 0^-, \mathbf{0}^\perp\right), \tag{471}$$

$$n^\mu_- = \left(0^+, 1^-, \mathbf{0}^\perp\right). \tag{472}$$

Care has to be taken when lowering the index, since this switches the light-like components because of the form of the metric

$$n_{+\,\mu} = \left(0^+, 1^-, \mathbf{0}^\perp\right), \tag{473}$$

$$n_{-\,\mu} = \left(1^+, 0^-, \mathbf{0}^\perp\right), \tag{474}$$

such that they project out the other light-like component of a vector:

$$a{\cdot}n_+ = a^-, \qquad\qquad a{\cdot}n_- = a^+. \tag{475}$$

Last we can define an anti-symmetric symbol:

$$\varepsilon^{\mu\nu}_\perp = \varepsilon^{+-\mu\nu} \tag{476}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{477}$$

where we adopt the convention $\varepsilon^{0123} = \varepsilon^{+-12} = +1$.

A.3 FOURIER TRANSFORMS

Below we give an overview of some useful functions with respect to Fourier transforms. We start with the Heaviside step function which is defined by

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}, \tag{478}$$

and is undefined for $x = 0$. Next we review the Dirac δ-function, the derivative of the above step function.

$$\delta(x) = \frac{d}{dx}\theta(x), \qquad \Longrightarrow \quad \int dx \delta(x) = 1, \tag{479}$$

and is zero everywhere, except at $x = 0$ Fourier transforms in this text are used in the following way

$$f(x) = \int \frac{d^4k}{16\pi^4} \tilde{f}(k) \ \exp[-ik \cdot x], \tag{480}$$

$$\tilde{f}(k) = \int d^4x f(x) \ \exp[ik \cdot x]. \tag{481}$$

A.4 DIRAC'S BRA AND KET NOTATION

We just give below a very short overview of some properties of the bra-ket notation.

$$\begin{aligned} \boldsymbol{x}_\mu \mid x \rangle &= x_\mu \mid x \rangle, \\ \int d^d x \ \ |x \rangle \langle x| &= 1, \\ \langle g \mid x \rangle &= g(x), \\ \langle x \mid f \rangle &= f(x), \\ |f \ \rangle &= \int d^d x \ f(x) \ |x\rangle, \end{aligned}$$

B

DETAILED CALCULATIONS

B.1 PROOF OF LEMMA 18.6.1 (FROM [16])

proof B.1.1

It should be clear that $(\mathcal{A}_p^*, \star)$ is a topological **k**-algebra. Define now **k**-linear maps $\Phi : End^{ll}(\mathcal{A}_p) \to \mathcal{A}_p^*$, $\Psi : \mathcal{A}_p^* \to End^{ll}(\mathcal{A}_p)$, and $\Lambda : \mathcal{A}_p^* \to End^{rl}(\mathcal{A}_p)$ respectivly by:

$$\Phi : \sigma \to \Phi(\sigma) \equiv f_\sigma \equiv \epsilon \circ \sigma \tag{482}$$

$$\Psi : f \to \Psi(f) \equiv \sigma_f \equiv (1 \otimes f) \circ \Delta \tag{483}$$

and

$$\Lambda : f \to \Lambda(f) \equiv \rho_f \equiv (f \otimes 1) \circ \Delta \tag{484}$$

Let us verify, for example, that Ψ is well defined and it's an algebra morphism. In fact:

$$\begin{aligned}(1 \otimes \Psi(f))\Delta &= (1 \otimes (1 \otimes f) \circ \Delta)\Delta \\ &= (1 \otimes 1 \otimes f)(1 \otimes \Delta)\Delta \\ &= (1 \otimes 1 \otimes f)(\Delta \otimes 1)\Delta \\ &= \Delta(1 \otimes f)\Delta = \Delta\Psi(f)\end{aligned} \tag{485}$$

which proves that $\Psi(f)$ is left invariant. On the other hand, composing with $1 \otimes g$, $g \in \mathcal{A}_p^*$, we obtain:

$$\begin{aligned}\Psi(g)\Psi(f) &= (1 \otimes g)\Delta\Psi(f) \\ &= (1 \otimes g)(1 \otimes \Psi(f))\Delta \\ &= (1 \otimes g\Psi(f))\Delta \\ &= (1 \otimes g \star f)\Delta \\ &= \Psi(g \star f)\end{aligned} \tag{486}$$

The rest of the proof follows directly from the definitions

□

B.2 PATH DERIVATIVE OF CHEN ITERATED INTEGRAL $X^{\omega_1\cdots\omega_r}(\gamma)$

$$
\begin{aligned}
&\nabla^T_V(q_s)X^{\omega_1\cdots\omega_r}(\gamma) = \\
&\lim_{s\to 0}\frac{\int_{\gamma\cdot\eta_s}\omega_1\cdots\omega_r - \int_\gamma \omega_1\cdots\omega_r}{s} = \\
&\lim_{s\to 0}\frac{\sum_{i=0}^{r}\int_\gamma \omega_1\cdots\omega_i\int_{\eta_s}\omega_{i+1}\cdots\omega_r - \int_\gamma \omega_1\cdots\omega_r}{s} = \\
&\lim_{s\to 0}\frac{\int_\gamma \omega_1\cdots\omega_r + \sum_{i=0}^{r-1}\int_\gamma \omega_1\cdots\omega_i\int_{\eta_s}\omega_{i+1}\cdots\omega_r - \int_\gamma \omega_1\cdots\omega_r}{s} = \\
&\lim_{s\to 0}\frac{\sum_{i=0}^{r-1}\int_\gamma \omega_1\cdots\omega_i\int_{\eta_s}\omega_{i+1}\cdots\omega_r}{s} = \\
&\sum_{i=0}^{r-1}\int_\gamma \omega_1\cdots\omega_i \lim_{s\to 0}\frac{\int_{\eta_s}\omega_{i+1}\cdots\omega_r}{s} = \\
&\int_\gamma \omega_1\cdots\omega_{r-1}\cdot\omega_r(V_{q_s}) = \\
&X^{\omega_1\cdots\omega_{r-1}}(\gamma_s)\cdot\omega_r(V_{q_s}) \qquad (487)
\end{aligned}
$$

B.3 AREA DERIVATIVE AS DERIVATIVES ON THE ITERATED INTEGRAL ALGEBRA

Let us show that indeed Eq. (238) is valid. We start by restating the left action of a path

$$
\begin{aligned}
(\alpha\cdot X^u)(\beta) &\equiv X^u(\beta\cdot\alpha) \\
\lambda\cdot X^u(\epsilon) &= X^u(\lambda\cdot\epsilon) = X^u(\lambda), \qquad (488)
\end{aligned}
$$

and the comultiplication on a functional at the group unit:

$$
\Delta X^u(\epsilon) = \sum_{i=0}^{r} X^{\omega_1\cdots\omega_i}(\epsilon)\cdot X^{\omega_{i+1}\cdots\omega_r}(\epsilon), \qquad (489)
$$

Now applying $\lambda\cdot J$ on the right part of the comultiplication returns:

$$
\begin{aligned}
\lambda\cdot J(X^{\omega_{i+1}\cdots\omega_r}(\epsilon)) &= \lambda\cdot(-1)^{r-i-1}X^{\omega_r\cdots\omega_{i+1}}(\epsilon) \\
&= (-1)^{r-i-1}X^{\omega_r\cdots\omega_{i+1}}(\lambda\cdot\epsilon) \\
&= X^{\omega_{i+1}\cdots\omega_r}(\lambda^{-1}). \qquad (490)
\end{aligned}
$$

Doing the similar calculation by applying $\lambda \cdot \mathcal{D}_{u\wedge v}(q)$ to the left part of the result of the co-multiplication in Eq.(489) we get:

$$\begin{aligned}
\lambda \cdot \mathcal{D}_{u\wedge v}(q) X^{\omega_1\cdots\omega_i}(\epsilon) &= \lambda \cdot \Big(D_{u\wedge v}(q) X^{\omega_1\cdots\omega_i}(\epsilon) + [\partial_u^T, \partial_v^T] X^{\omega_1\cdots\omega_i}(\epsilon) \Big) \\
\lambda \cdot \Delta_{u\wedge v}(q) X^{\omega_1\cdots\omega_i}(\epsilon) &= \lambda \cdot X^{\omega_1\cdots\omega_{i-1}}(\epsilon) \cdot d\omega_i(u\wedge v) \\
&= X^{\omega_1\cdots\omega_{i-1}}(\lambda) \cdot d\omega_i(u\wedge v) \\
\lambda[\partial_u^T, \partial_v^T] X^{\omega_1\cdots\omega_i}(\epsilon) &= \lambda \left(X^{\omega_1\cdots\omega_{i-2}}(\epsilon) \cdot (\omega_{i-1}\wedge\omega_i)(u\wedge v) \right) \\
&= X^{\omega_1\cdots\omega_{i-2}}(\lambda) \cdot (\omega_{i-1}\wedge\omega_i)(u\wedge v). \quad (491)
\end{aligned}$$

Combining the contributions above demonstrates the validity of (238).

B.4 COMMUTATOR OF AREA DERIVATIVE ON ALGEBRA OF ITERATED INTEGRALS

$$\begin{aligned}
&[\boldsymbol{\delta}_{(\lambda;a\wedge b)}, \boldsymbol{\delta}_{(\eta;u\wedge v)}] X^{\omega_1\ldots\omega_r} = \\
&\epsilon \circ \Big((1\otimes \boldsymbol{\delta}_{(\lambda;a\wedge b)}) \Delta (1 \otimes \boldsymbol{\delta}_{(\eta;u\wedge v)}) \Delta \Big) X^{\omega_1\cdots\omega_r} \\
&\qquad - \epsilon \circ \Big((1\otimes \boldsymbol{\delta}_{(\eta;u\wedge v)}) \Delta (1 \otimes \boldsymbol{\delta}_{(\lambda;a\wedge b)}) \Delta \Big) X^{\omega_1\cdots\omega_r} = \\
&\triangle^E_{(\lambda;a\wedge b)}(\lambda(1)) \Big((1\otimes \boldsymbol{\delta}_{(\eta;u\wedge v)}) \Delta X^{\omega_1\cdots\omega_r} \Big)(\epsilon) \\
&\qquad - \triangle^E_{(\eta;u\wedge v)}(\eta(1)) \Big((1\otimes \boldsymbol{\delta}_{\lambda;a\wedge b)}) \Delta X^{\omega_1\cdots\omega_r} \Big)(\epsilon) = \\
&\triangle^E_{(\lambda;a\wedge b)}(\lambda(1)) \Big(\sum_{i=0}^{r} X^{\omega_1\cdots\omega_i} \boldsymbol{\delta}_{(\eta;u\wedge v)}(X^{\omega_{i+1}\cdots\omega_r}) \Big)(\epsilon) \\
&\qquad - \triangle^E_{(\eta;u\wedge v)}(\eta(1)) \Big(\sum_{i=0}^{r} X^{\omega_1\cdots\omega_i} \boldsymbol{\delta}_{(\lambda;a\wedge b)}(X^{\omega_{i+1}\cdots\omega_r}) \Big)(\epsilon) = \\
&\sum_{i=0}^{r}\sum_{k=0}^{i} (\mathcal{D}_{a\wedge b}(\lambda(1)) X^{\omega_1\ldots\omega_k}(\lambda)) (X^{\omega_{k+1}\cdots\omega_i}(\lambda^{-1})) \boldsymbol{\delta}_{(\eta;u\wedge v)}(X^{\omega_{i+1}\cdots\omega_r}) \\
&- \sum_{i=0}^{r}\sum_{k=0}^{i} (\mathcal{D}_{u\wedge v}(\eta(1)) X^{\omega_1\cdots\omega_k}(\eta)) (X^{\omega_{k+1}\cdots\omega_i}(\eta^{-1})) \boldsymbol{\delta}_{(\lambda;a\wedge b)}(X^{\omega_{i+1}\cdots\omega_r}),
\end{aligned} \quad (492)$$

B.5 FREE QCD EULER-LAGRANGE AND SDYSON EQUATIONS

In the calculations below we have suppressed the color index to not overload the notations[1]. We start by deriving the variations of the pure Yang-Mills Lagrangian with respect to the derivatives of the gauge field.

$$\frac{\partial \mathcal{L}^{QCD}_{free}}{\partial\left(\partial_\mu A_\nu\right)} = \frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left[\left(\partial_\mu A_\nu - \partial_\nu A_\mu - ig\left[A_\mu, A_\nu\right]\right)\right.$$
$$\left.\left(\partial^\mu A^\nu - \partial^\nu A^\mu - ig\left[A^\mu, A^\nu\right]\right)\right]$$

This can be split up in three parts

$$\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu - \partial^\nu A^\mu - ig\left[A^\mu, A^\nu\right]\right)\right) \tag{493}$$
$$\frac{\partial}{\partial\left(\partial_\nu A_\mu\right)}\left(\left(-\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu - \partial^\nu A^\mu - ig\left[A^\mu, A^\nu\right]\right)\right) \tag{494}$$
$$\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(\left(-ig\left[A_\mu, A_\nu\right]\right)\left(\partial^\mu A^\nu - \partial^\nu A^\mu - ig\left[A^\mu, A^\nu\right]\right)\right) \tag{495}$$

The first part (493) splits up again in two parts when applying the derivative, the first part is equal to the field strength tensor with upper indices and color index a : $F^{\mu\nu,a}$. The second part requires some playing around with the indices:

$$\begin{aligned}
&\left(\partial_\mu A_\nu\right)\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(\partial^\mu A^\nu - \partial^\nu A^\mu - ig\left[A^\mu, A^\nu\right]\right) = \\
&\left(\partial_\mu A_\nu\right)\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(g^{\mu r} g^{\nu s}\left(\partial_r A_s - \partial_s A_r - ig\left[A_r, A_s\right]\right)\right) = \\
&\left(\partial^r A^s\right)\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(\left(\partial_r A_s - \partial_s A_r - ig\left[A_r, A_s\right]\right)\right) \underset{r\leftrightarrow\mu}{\overset{s\leftrightarrow\nu}{\equiv}} \\
&\partial^\mu A^\nu - \partial^\nu A^\mu
\end{aligned} \tag{496}$$

Note that the second term in Eq. (496) results from interchanging the role of μ and ν when applying the derivative. For the second part we have as before

1 they will give rise to δ-function over the color indices

that the first factor of (494) results again in a $F^{\mu\nu,a}$ term, where the second factor leads to

$$
\begin{aligned}
&(-\partial_\nu A_\mu)\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\partial^\mu A^\nu-\partial^\nu A^\mu-ig\left[A^\mu,A^\nu\right]\right)\overset{\mu\leftrightarrow\nu}{=}\\
&(-\partial_\mu A_\nu)\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(g^{\nu r}g^{\mu s}\left(\partial_r A_s-\partial_s A_r-ig\left[A_r,A_s\right]\right)\right)=\\
&(-\partial^s A^r)\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\left(\partial_r A_s-\partial_s A_r-ig\left[A_r,A_s\right]\right)\right)\underset{r\leftrightarrow\mu}{\overset{s\leftrightarrow\nu}{=}}\\
&-\partial^\nu A^\mu-(-\partial^\mu A^\nu)
\end{aligned}
\tag{497}
$$

where we again interchanged the roles of μ and ν to get the second term. Moving on to the third part of the derivative where now the first factor returns a zero contribution and where we have for the second factor:

$$
\begin{aligned}
&(-ig\left[A_\mu,A_\nu\right])\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\left(\partial^\mu A^\nu-\partial^\nu A^\mu-ig\left[A^\mu,A^\nu\right]\right)\right)=\\
&-ig\left[A_\mu,A_\nu\right]+ig\left[A_\nu,A_\mu\right]\\
&-ig\left[A_\mu,A_\nu\right]-ig\left[A_\mu,A_\nu\right]
\end{aligned}
\tag{498}
$$

Now combining all the terms we get the final result (499) where we re-introduced the color index.

$$
\frac{\partial\mathcal{L}^{QCD}_{free}}{\partial(\partial_\mu A_\nu)}=-F^{\mu\nu,a} \tag{499}
$$

Now we also need the variation of the Lagrangian with respect the gauge fields themselves. Since $\partial_\mu A_\nu$ is considered to be independent of A_μ we can restrict the calculation to the terms of the free QCD Lagrangian that contain terms in A_μ.

$$\frac{\partial \mathcal{L}^{QCD}_{Free}}{\partial A_\mu} =$$
$$\frac{\partial}{\partial A_\mu}\left[\left(\partial_\mu A_\nu\right)\left(-ig\left[A^\mu, A^\nu\right]\right) - \left(\partial_\nu A_\mu\right)\left(-ig\left[A^\mu, A^\nu\right]\right)\right.$$
$$+\left(\partial^\mu A^\nu\right)\left(-ig\left[A_\mu, A_\nu\right]\right) - \left(\partial^\nu A^\mu\right)\left(-ig\left[A_\mu, A_\nu\right]\right)$$
$$\left.+\left(-ig\left[A_\mu, A_\nu\right]\right)\left(-ig\left[A^\mu, A^\nu\right]\right)\right] =$$
$$\left(\partial_\mu A_\nu\right)\frac{\partial}{\partial A_\mu}\left(-ig\left[A^\mu, A^\nu\right]\right) - \left(\partial_\nu A_\mu\right)\frac{\partial}{\partial A_\mu}\left(-ig\left[A^\mu, A^\nu\right]\right) + s \tag{500}$$
$$\left(\partial^\mu A^\nu\right)\frac{\partial}{\partial A_\mu}\left(-ig\left[A_\mu, A_\nu\right]\right) - \left(\partial^\nu A^\mu\right)\frac{\partial}{\partial A_\mu}\left(-ig\left[A_\mu, A_\nu\right]\right) - \tag{501}$$
$$g^2\left\{\left[A_\mu, A_\nu\right]\frac{\partial}{\partial A_\mu}\left(A^\mu A^\nu\right) - \left[A_\mu, A_\nu\right]\frac{\partial}{\partial A_\mu}\left(A^\nu A^\mu\right) + \right. \tag{502}$$
$$\left.\left(\frac{\partial}{\partial A_\mu}\left(A_\mu A_\nu\right)\right)\left[A^\mu, A^\nu\right] - \left(\frac{\partial}{\partial A_\mu}\left(A_\nu A_\mu\right)\right)\left[A^\mu, A^\nu\right]\right\} \tag{503}$$

Note that as before the contributions with the upper (Eq. (501) and (503)) and lower indices (Eq. (500) and (502)) switched will give the same contribution after introducing the metric tensor to raise/lower the indices, renaming of some dummy indices and because:

$$\begin{aligned}
&\frac{\partial}{\partial A_\mu}\left(A_\mu A_\nu - A_\nu A_\mu\right)\left[A^\mu, A^\nu\right] = \\
&A_\nu\left[A^\mu, A^\nu\right] - A_\nu\left[A^\nu, A^\mu\right] = \\
&2A_\nu\left[A^\mu, A^\nu\right] = \\
&2A_\nu A^\mu A^\nu - 2A_\nu A^\nu A^\mu = \\
&2A^\nu A^\mu A_\nu - 2A^\mu A_\nu A^\nu = \\
&2A^\nu A^\mu A_\nu - 2A^\mu A^\nu A_\nu = \\
&2\left[A^\mu, A^\nu\right]A_\nu
\end{aligned} \tag{504}$$

Making the total contribution of Eq. (502) and Eq. (503):

$$-4g^2\left[A^\mu, A^\nu\right]A_\nu \tag{505}$$

For the other contributions we have:

$$\begin{aligned}
&\frac{\partial}{\partial A_\mu}\left(\left(\partial^\mu A^\nu\right)\left(-igA_\mu A_\nu + igA_\nu A_\mu\right)\right) = \\
&-ig\left(\partial^\mu A^\nu\right)\frac{\partial}{\partial A_\mu}\left(A_\mu A_\nu\right) + ig\left(\partial^\mu A^\nu\right)\frac{\partial}{\partial A_\mu}\left(A_\nu A_\mu\right) = \\
&-ig\left(\partial^\mu A^\nu\right)A_\nu + ig\left(\partial^\nu A^\mu\right)\frac{\partial}{\partial A_\mu}\left(A_\mu A_\nu\right) = \\
&-ig\left(\partial^\mu A^\nu\right)A_\nu + ig\left(\partial^\nu A^\mu\right)A_\nu \qquad (506)\\
&-\frac{\partial}{\partial A_\mu}\left(\left(\partial^\nu A^\mu\right)\left(-igA_\mu A_\nu + igA_\nu A_\mu\right)\right) = \\
&-ig\left(\partial^\mu A^\nu\right)A_\nu + ig\left(\partial^\nu A^\mu\right)A_\nu \qquad (507)
\end{aligned}$$

So that the total contribution becomes:

$$\begin{aligned}
&4ig\left(\left(\partial^\nu A^\mu\right) - \left(\partial^\mu A^\nu\right) + ig\left[A^\mu, A^\nu\right]\right)A_\nu = \\
&4ig\left(-F^{\mu\nu}\right)A_\nu, \qquad (508)
\end{aligned}$$

taking the factor $\frac{1}{4}$ into account this gives us $igF^{\mu\nu}A_\nu$. Now using the fact that

$$\left[\frac{\lambda^i}{2}, \frac{\lambda^j}{2}\right] = if^{ijk}\frac{\lambda^k}{2}$$

we can re-introduce the color [19] in Eq. (508)

$$igF^{\mu\nu}A_\nu \rightarrow gf^{ijk}F^{\mu\nu,k}A^j_\nu \qquad (509)$$

The Euler-Lagrange equations for the free QCD Lagrangian then become

$$\boxed{\partial_\nu F^{\mu\nu,i} + gf^{ijk}F^{\mu\nu,k}A^j_\nu = 0} \qquad (510)$$

And for the SD equations this becomes (where the equalities are weak[2])

$$\boxed{\left(\partial_\nu F^{\mu\nu,i} + gf^{ijk}F^{\mu\nu,k}A^j_\nu\right)\boldsymbol{F}[A] = D^{ab}_\nu F^{\mu\nu,b}(x)\boldsymbol{F}[A] = \hbar\frac{\delta}{\delta A^a_\nu(x)}\boldsymbol{F}[A]} \qquad (511)$$

where $\boldsymbol{F}[\phi]$ is the functional as earlier in the derivation of the SD equations and D^{ab}_μ represents the covariant derivative in the adjoint representation.

2 They are only valid under taking averages.

B.6 SELF-ENERGY CONTRIBUTION WILSON LOOP QUADRILATERAL ON THE LIGHT-CONE

Here we calculate the first SE diagram from figure 57. For this diagram we have that the coordinates x and y in the gluon propagator (349):

$$\begin{aligned} x &= v_1 t,\ t \in [0,1] \\ y &= v_1 s,\ s \in [t,1]. \end{aligned} \tag{512}$$

The total contribution of this diagram then can be written as:

$$\begin{aligned} W_{\mathsf{SE}} &= \int dx^\mu \int dy^\nu D_{\mu\nu}(x-y) \\ &= \int dx^\mu \int dy^\nu \,\frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon) g_{\mu\nu}\delta^{ab}\frac{1}{\left(-(x-y)^2\right)^{1-\epsilon}} \\ &= \int_0^1 v_1^\mu dt \int_0^t v_1^\nu ds \,\frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon) g_{\mu\nu}\delta^{ab}\frac{1}{\left(-(v_1(t-s))^2\right)^{1-\epsilon}} \\ &= (v_1^2)^\epsilon \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)\left(\frac{1}{1-2\epsilon}\right)\frac{1}{2\epsilon}(-1)^\epsilon \\ &\overset{\mathsf{DR}}{=} 0, \end{aligned} \tag{513}$$

showing that in the Feynman gauge, with dimensional regularization, these contributions vanish.

B.7 CUSP CONTRIBUTION WILSON LOOP QUADRILATERAL ON THE LIGHT-CONE

Here we calculate the first cusp (CU) diagram from figure 57. For this diagram we have that the coordinates x and y in the gluon propagator (349):

$$\begin{aligned} x &= v_1 t,\ t \in [0,1] \\ y &= v_1 + v_2 s,\ s \in [0,1]. \end{aligned} \tag{514}$$

The total contribution of this diagram then can be written as:

$$\begin{aligned} W_{\mathsf{CU}} &= \int dx^\mu \int dy^\nu D_{\mu\nu}(x-y) \\ &= \int dx^\mu \int dy^\nu \ \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon) g_{\mu\nu}\delta^{ab}\frac{1}{\left(-(x-y)^2\right)^{1-\epsilon}} \\ &= \int\limits_0^1 v_1^\mu dt \int\limits_0^1 v_2^\nu ds \ \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon) g_{\mu\nu}\delta^{ab}\frac{1}{\left(-(v_1(t-1)-v_2 s)^2\right)^{1-\epsilon}} \\ &= \int\limits_0^1 v_1^\mu dt \int\limits_0^1 v_2^\nu ds \ \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon) g_{\mu\nu}\delta^{ab}\frac{1}{\left(2v_1 v_2 s(t-1))^2\right)^{1-\epsilon}} \\ &= \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)\left(-\frac{1}{2}S^\epsilon\frac{1}{\epsilon^2}\right) \end{aligned} \tag{515}$$

showing that in the Feynman gauge, with dimensional regularization, has a double pole stemming from the overlap of a light cone divergence with a UV divergence.

B.8 WILSON LOOP : POLYLOG INTEGRAL

This Section reviews some aspects of the Polylog functions followed by the calculations of a specific integral that we encounter during the calculations of the vacuum expectation value of a light like quadrilateral Wilson loop.

B.8.1 *Summary of the Polylog function*

B.8.1.1 *Property 1*

$$Li_2[z] = -\int_0^z \frac{dt}{t}\ln[1-t] \tag{516}$$

$$= -\int_0^1 \frac{dt}{t}\ln[1-zt] \tag{517}$$

B.8.1.2 *Property 2*

$$Li_2[z] = -Li_2[1-z] - \ln[z]\ln[1-z] + \frac{\pi^2}{6} \tag{518}$$

B.8.1.3 *Property 3*

$$Li_2\left[0\right]=0 \tag{519}$$

B.8.1.4 *Property 4*

$$Li_2\left[1\right]=\zeta_2=\frac{\pi^6}{6} \tag{520}$$

B.8.1.5 *Property 5*

$$Li_2\left[1+z\right]+Li_2\left[1+\frac{1}{z}\right]=-\frac{1}{2}\ln^2\left[-z\right] \tag{521}$$

B.8.1.6 *Property 6*

$$\int_0^1\frac{\ln^n\left[y\right]}{y-\frac{1}{z}}dy=\left(-1\right)^{n+1}n!Li_{n+1}\left(z\right) \tag{522}$$

B.8.2 *Calculation of the integral*

The integral we want to calculate is given by:

$$\begin{aligned}
&\int_0^1\frac{dxdy\left(s+t\right)}{tx+sy-\left(s+t\right)xy}=\\
&-\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[\frac{s}{t}\right]+\ln\left[\frac{1-x}{x}\right]\right)=\\
&-\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[\frac{s}{t}\right]\right)-\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[1-x\right]\right)+\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[x\right]\right)=\\
&-\left(\ln\left[\frac{s}{t}\right]\left(\ln\left[\frac{t}{s+t}\right]-\ln\left[-\frac{s}{s+t}\right]\right)\right)-\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[1-x\right]\right)+\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[x\right]\right)
\end{aligned}$$

Now calculating the first remaining integral:

$$\begin{aligned}
-\int_0^1 dx\frac{1}{x-\frac{s}{s+t}}\left(\ln\left[1-x\right]\right) &= \int_0^1 dx\frac{1}{1-x+\frac{s}{s+t}-\frac{s+t}{s+t}}\left(\ln\left[1-x\right]\right)\\
&= \int_0^1 dx\frac{1}{1-x+\frac{s}{s+t}-\frac{s+t}{s+t}}\left(\ln\left[1-x\right]\right)\\
&= -\int_1^0 d\left(1-x\right)\frac{1}{\left(1-x\right)-\frac{t}{s+t}}\left(\ln\left[1-x\right]\right)\\
&= \int_0^1 d\left(1-x\right)\frac{1}{\left(1-x\right)-\frac{t}{s+t}}\left(\ln\left[1-x\right]\right)
\end{aligned}$$

Using Eq. (522)

$$
\begin{aligned}
-\int_0^1 dx \frac{1}{x-\frac{s}{s+t}}\left(\ln\left[1-x\right]\right) &= Li_2\left[\frac{1}{\frac{t}{s+t}}\right] \\
&= Li_2\left[\frac{s+t}{t}\right] \\
&= Li_2\left[1+\frac{s}{t}\right]
\end{aligned}
$$

Taking on the second remaining integral:

$$
\begin{aligned}
\int_0^1 dx \frac{1}{x-\frac{s}{s+t}}\left(\ln\left[x\right]\right) &= Li_2\left[\frac{s+t}{s}\right] \\
&= Li_2\left[1+\frac{1}{\frac{s}{t}}\right]
\end{aligned}
$$

Combining now the results for these integrals we have:

$$
\begin{aligned}
-\int_0^1 dx \frac{1}{x-\frac{s}{s+t}}\left(\ln\left[1-x\right]\right)+\int_0^1 dx \frac{1}{x-\frac{s}{s+t}}\left(\ln\left[x\right]\right) &= Li_2\left[1+\frac{s}{t}\right]+Li_2\left[1+\frac{1}{\frac{s}{t}}\right] \\
&= -\frac{1}{2}\ln^2\left[-\frac{s}{t}\right]
\end{aligned}
$$

Taking now a second look at the terms without the integrals:

$$
\begin{aligned}
-\left(\ln\left[\frac{s}{t}\right]\left(\ln\left[\frac{t}{s+t}\right]-\ln\left[-\frac{s}{s+t}\right]\right)\right) &= -\left(\ln\left[\frac{s}{t}\right]\left(\ln\left[\frac{t}{s+t}\right]-\ln\left[\exp\left[\imath\pi\right]\frac{s}{s+t}\right]\right)\right) \\
&= -\left(\ln\left[\frac{s}{t}\right]\left(\ln\left[\frac{t}{s+t}\right]-\ln\left[\frac{s}{s+t}\right]-\ln\left[\exp\left[\imath\pi\right]\right]\right)\right) \\
&= \ln\left[\frac{s}{t}\right]\left(-\ln\left[\frac{t}{s+t}\right]+\ln\left[\frac{s}{s+t}\right]+\ln\left[\exp\left[\imath\pi\right]\right]\right) \\
&= \ln\left[\frac{s}{t}\right]\left(\ln\left[\frac{s}{t}\right]+\imath\pi\right) \\
&= \ln^2\left[\frac{s}{t}\right]+\imath\pi\ln\left[\frac{s}{t}\right]
\end{aligned}
$$

Combining this with the result of the calculations with the integrals

$$\begin{aligned}\int_0^1 \frac{dxdy\,(s+t)}{tx+sy-(s+t)\,xy} &= \ln^2\left[\frac{s}{t}\right]+\imath\pi\ln\left[\frac{s}{t}\right]-\frac{1}{2}\ln^2\left[-\frac{s}{t}\right]\\ &= \ln^2\left[\frac{s}{t}\right]+\imath\pi\ln\left[\frac{s}{t}\right]-\frac{1}{2}\ln^2\left[\exp\left[\imath\pi\right]\frac{s}{t}\right]\\ &= \ln^2\left[\frac{s}{t}\right]+\imath\pi\ln\left[\frac{s}{t}\right]-\frac{1}{2}\left(\ln\left[\exp\left[\imath\pi\right]\frac{s}{t}\right]\right)^2\\ &= \ln^2\left[\frac{s}{t}\right]+\imath\pi\ln\left[\frac{s}{t}\right]-\frac{1}{2}\left(\imath\pi+\ln\left[\frac{s}{t}\right]\right)^2\\ &= \ln^2\left[\frac{s}{t}\right]+\imath\pi\ln\left[\frac{s}{t}\right]-\frac{1}{2}\left(-\pi^2+2\imath\pi\ln\left[\frac{s}{t}\right]+\ln^2\left[\frac{s}{t}\right]\right)\\ &= \frac{\pi^2}{2}+\frac{1}{2}\ln^2\left[\frac{s}{t}\right]\\ &= \frac{1}{2}\left(\pi^2+\ln^2\left[\frac{s}{t}\right]\right)\end{aligned}$$

B.8.3 *Summary*

The trick for solving this integral lies in the rewriting of a -1 factor in a logarithm as:

$$\begin{aligned}\ln\left[-a\right] &= \ln\left[\exp\left[\imath\pi\right]a\right]\\ &= \ln\left[a\right]+\ln\left[\exp\left[\imath\pi\right]\right]\\ &= \ln\left[a\right]+\imath\pi\end{aligned}$$

This trick enabled us to calculate the integral and express it in a simple form, making the squared logarithmic dependence clear.

B.9 CROSS DIAGRAM CONTRIBUTION WILSON LOOP

Here we calculate a crossed (CR) diagram from the last row of figure 57. For this diagram we have that the coordinates x and y in the gluon propagator Eq. (349):

$$\begin{aligned}x &= v_1 t,\ t\in[0,1]\\ y &= v_1+v_2+v_3 s,\ s\in[0,1].\end{aligned} \tag{523}$$

The total contribution of this diagram then can be written as:

$$
\begin{aligned}
&W_{\mathsf{CR}} = \\
&\int dx^\mu \int dy^\nu D_{\mu\nu}(x-y) = \\
&\int dx^\mu \int dy^\nu \; \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)g_{\mu\nu}\delta^{ab}\frac{1}{\left(-(x-y)^2\right)^{1-\epsilon}} = \\
&\int_0^1 v_1^\mu dt \int_0^1 v_3^\nu ds \; \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)g_{\mu\nu}\delta^{ab}\frac{1}{\left(-(v_1(t-1)-v_2 s)^2\right)^{1-\epsilon}} = \\
&\int_0^1 v_1^\mu dt \int_0^1 v_3^\nu ds \; \frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)g_{\mu\nu}\delta^{ab} \\
&\qquad\qquad \times \frac{1}{\left(2(-v_1 v_2(t-1)+v_2 v_3 s - v_1 v_3 s(t-1)))^2\right)^{1-\epsilon}} = \\
&\frac{(\mu^2\pi)^\epsilon}{4\pi^2}\Gamma(1-\epsilon)\frac{1}{2}\left(\ln\left(-\frac{s}{t}\right)^2 + \pi^2\right)
\end{aligned}
\tag{524}
$$

showing that in the Feynman gauge, with dimensional regularization, this contribution is finite.

BIBLIOGRAPHY

[1] M. Creutz. Confinement, chiral symmetry, and the lattice. *Acta Physica Slovaca*, 61:1–127, February 2011, arXiv:1103.3304.

[2] B. Ioffe, V. Fadin, and L. Lipatov. *Quantum Chromodynamics: Perturbative and Nonperturbative Aspects*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2010.

[3] J. D. Edelstein, J. P. Shock, and D. Zoakos. The AdS/CFT Correspondence and Non-perturbative QCD. *AIP Conf.Proc.*, 1116:265–284, 2009, arXiv:0901.2534.

[4] A. Slavnov. Gauge fields beyond perturbation theory. *PoS*, QCD-TNT-III:041, 2014, arXiv:1310.8164.

[5] R. Jaffe. Where does the proton really get its spin? *Phys.Today*, 48N9:24–30, 1995.

[6] A. Bacchetta. Transverse momentum distributions. Doctoral School Lectures, 2012.

[7] A. V. Belitsky, X. Ji, and F. Yuan. Final state interactions and gauge invariant parton distributions. *Nucl. Phys.*, B656:165–198, 2003, arXiv:0208038.

[8] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl.Phys.*, B126:298, 1977.

[9] Y. L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.

[10] L. Lipatov. The parton model and perturbation theory. *Sov.J.Nucl.Phys.*, 20:94–102, 1975.

[11] V. Gribov and L. Lipatov. e+ e- pair annihilation and deep inelastic e p scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:675–684, 1972.

[12] E. Kuraev, L. Lipatov, and V. S. Fadin. The Pomeranchuk Singularity in Nonabelian Gauge Theories. *Sov.Phys.JETP*, 45:199–204, 1977.

[13] I. Balitsky and L. Lipatov. The Pomeranchuk Singularity in Quantum Chromodynamics. *Sov.J.Nucl.Phys.*, 28:822–829, 1978.

[14] J. C. Collins and D. E. Soper. Back-To-Back Jets in QCD. *Nucl.Phys.*, B193:381, 1981.

[15] J. C. Collins, D. E. Soper, and G. F. Sterman. Factorization of Hard Processes in QCD. *Adv.Ser.Direct.High Energy Phys.*, 5:1–91, 1988, arXiv:0409313.

[16] J. N. Tavares. Chen integrals, generalized loops and loop calculus. *Int. J. Mod. Phys.*, A9:4511–4548, 1994, arXiv:9305173.

[17] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken. *Wilson lines in quantum field theory*. De Gruyter, 2014.

[18] M. Planck and M. von Laue. *Wissenschaftliche Selbstbiographie. Mit Einem Bildnis und Der Von Max Von Laue Gehaltenen Traueransprache*. 1948.

[19] M. Peskin and D. Schroeder. *An Introduction to Quantum Field Theory*. Advanced book classics. Addison-Wesley Publishing Company, 1995.

[20] D. J. Gross and F. Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys.Rev.Lett.*, 30:1343–1346, 1973.

[21] H. D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys.Rev.Lett.*, 30:1346–1349, 1973.

[22] S. B. Libby and G. F. Sterman. Jet and Lepton Pair Production in High-Energy Lepton-Hadron and Hadron-Hadron Scattering. *Phys.Rev.*, D18:3252, 1978.

[23] S. B. Libby and G. F. Sterman. Mass Divergences in Two Particle Inelastic Scattering. *Phys.Rev.*, D18:4737, 1978.

[24] G. Sterman. *An Introduction to Quantum Field Theory*. Cambridge University Press, 1993.

[25] M. Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys.Lett.*, 8:214–215, 1964.

[26] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. 1964.

[27] R. P. Feynman. Very high-energy collisions of hadrons. *Phys.Rev.Lett.*, 23:1415–1417, 1969.

[28] R. Feynman. Photon-hadron interactions. 1973.

[29] R. P. Feynman. Structure of the proton. *Science*, 183:601, 1974.

[30] R. Feynman, M. Kislinger, and F. Ravndal. Current matrix elements from a relativistic quark model. *Phys.Rev.*, D3:2706–2732, 1971.

[31] C. Bourrely and J. Soffer. Do we understand the single spin asymmetry for pi0 inclusive production in pp collisions? *Eur.Phys.J.*, C36:371–374, 2004, arXiv:0311110.

[32] U. D'Alesio and F. Murgia. Azimuthal and Single Spin Asymmetries in Hard Scattering Processes. *Prog.Part.Nucl.Phys.*, 61:394–454, 2008, arXiv:0712.4328.

[33] J. Ashman, B. Badelek, G. Baum, et al. A measurement of the spin asymmetry and determination of the structure function g1 in deep inelastic muon-proton scattering. *Phys. Lett.*, B206(2):364 – 370, 1988.

[34] D. Adams, N. Akchurin, N. Belikov, et al. Analyzing power in inclusive π^+ and π^- production at high x_f with a 200 gev polarized proton beam. *Phys. Lett.*, B264(3-4):462 – 466, 1991.

[35] Z.-t. Liang and C. Boros. Hyperon polarization and single spin left-right asymmetry in inclusive production processes at high-energies. *Phys.Rev.Lett.*, 79:3608–3611, 1997, arXiv:9708488.

[36] D. Adams, N. Akchurin, N. Belikov, et al. Comparison of spin asymmetries and cross sections in π^0 production by 200 gev polarized antiprotons and protons. *Phys. Lett.*, B261(1-2):201 – 206, 1991.

[37] D. Adams, N. Achurin, N. Belikov, et al. Measurement of single spin asymmetry for direct photon production in pp collisions at 200 gev/c. *Phys. Lett.*, B345(4):569 – 575, 1995.

[38] A. Bravar, D. L. Adams, N. Akchurin, et al. Single-spin asymmetries in inclusive charged pion production by transversely polarized antiprotons. *Phys. Rev. Lett.*, 77:2626–2629, Sep 1996.

[39] D. L. Adams, N. Akchurin, N. I. Belikov, et al. Single-spin asymmetries and invariant cross sections of the high transverse-momentum inclusive π^0 production in 200 gev/c pp and pp interactions. *Phys. Rev. D*, 53:4747–4755, May 1996.

[40] B. I. Abelev, M. M. Aggarwal, Z. Ahammed, et al. Measurement of transverse single-spin asymmetries for dijet production in proton-proton collisions at 200 gev. *Phys. Rev. Lett.*, 99:142003, Oct 2007.

[41] C. Bomhof, P. Mulders, W. Vogelsang, and F. Yuan. Single-Transverse Spin Asymmetry in Dijet Correlations at Hadron Colliders. *Phys.Rev.*, D75:074019, 2007, arXiv:0701277.

[42] S. S. Adler, S. Afanasiev, C. Aidala, et al. Measurement of transverse single-spin asymmetries for midrapidity production of neutral pions and charged hadrons in polarized $p+p$ collisions at 200 gev. *Phys. Rev. Lett.*, 95:202001, Nov 2005.

[43] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang. Evidence for polarization of gluons in the proton. *Phys. Rev. Lett.*, 113:012001, Jul 2014.

[44] E. R. Nocera, R. D. Ball, S. Forte, G. Ridolfi, and J. Rojo. A first unbiased global determination of polarized PDFs and their uncertainties. 2014, arXiv:1406.5539.

[45] C. Lorcé. The nucleon spin decomposition: news and experimental implications. *EPJ Web Conf.*, 73:02005, 2014, arXiv:1401.3954.

[46] A. Belitsky and A. Radyushkin. Unraveling hadron structure with generalized parton distributions. *Phys. Rept.*, 418:1–387, 2005, arXiv:0504030.

[47] A. Bacchetta, M. Diehl, K. Goeke, et al. Semi-inclusive deep inelastic scattering at small transverse momentum. *JHEP*, 0702:093, 2007, arXiv:0611265.

[48] I. Cherednikov, T. Mertens, and F. F. Van der Veken. Evolution of cusped light-like Wilson loops and geometry of the loop space. *Phys. Rev.*, D86:085035, 2012, arXiv:1208.1631.

[49] I. Cherednikov and T. Mertens. Fréchet derivative for light-like Wilson Loops. 2014, arXiv:1401.2721.

[50] T. Mertens and P. Taels. Evolution of light-like wilson loops with a self-intersection in loop space. *Phys. Lett.*, B727(4-5):563 – 567, 2013.

[51] R. Hofstadter. Electron scattering and nuclear structure. *Rev. Mod. Phys.*, 28:214–254, Jul 1956.

[52] J. Collins. *Foundations of Perturbative QCD*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2011.

[53] J. Collins and J. Collins. *Renormalization: An Introduction to renormalization, the Renormalization Group and the Operator-Product Expansion*.

[54] C. G. Callan and D. J. Gross. High-energy electroproduction and the constitution of the electric current. *Phys. Rev. Lett.*, 22:156–159, Jan 1969.

[55] J. P. Ralston and D. E. Soper. Production of dimuons from high-energy polarized proton-proton collisions. *Nucl. Phys.*, B152(1):109 – 124, 1979.

[56] J. C. Collins, D. E. Soper, and G. Sterman. Factorization for short distance hadron-hadron scattering. *Nucl. Phys.*, B261(0):104 – 142, 1985.

[57] A. V. Manohar. An Introduction to spin dependent deep inelastic scattering. 1992, arXiv:9204208.

[58] A. Efremov and A. Radyushkin. Field Theoretic Treatment of High Momentum Transfer Processes. 3. Gauge Theories. *Theor.Math.Phys.*, 44:774, 1981.

[59] X.-d. Ji and F. Yuan. Parton distributions in light cone gauge: Where are the final state interactions? *Phys.Lett.*, B543:66–72, 2002, arXiv:0206057.

[60] D. Boer, P. Mulders, and F. Pijlman. Universality of T odd effects in single spin and azimuthal asymmetries. *Nucl. Phys.*, B667:201–241, 2003, arXiv:0303034.

[61] P. Mulders. Transverse momentum dependence in structure functions in hard scattering processes. University Lectures, 2001.

[62] C.-R. Ji and A. T. Suzuki. Interpolating scattering amplitudes between the instant form and the front form of relativistic dynamics. *Phys.Rev.*, D87(6):065015, 2013, arXiv:1212.2265.

[63] I. Cherednikov and N. Stefanis. Renormalization, Wilson lines, and transverse-momentum dependent parton distribution functions. *Phys.Rev.*, D77:094001, 2008, arXiv:0710.1955.

[64] I. Cherednikov and N. Stefanis. Wilson lines and transverse-momentum dependent parton distribution functions: A Renormalization-group analysis. *Nucl.Phys.*, B802:146–179, 2008, arXiv:0802.2821.

[65] I. Cherednikov and N. Stefanis. Renormalization-group properties of transverse-momentum dependent parton distribution functions in the light-cone gauge with the Mandelstam-Leibbrandt prescription. *Phys.Rev.*, D80:054008, 2009, arXiv:0904.2727.

[66] N. Stefanis and I. Cherednikov. Renormalization-group anatomy of transverse-momentum dependent parton distribution functions in QCD. *Mod.Phys.Lett.*, A24:2913–2923, 2009, arXiv:0910.3108.

[67] I. Cherednikov, A. Karanikas, and N. Stefanis. Wilson lines in transverse-momentum dependent parton distribution functions with spin degrees of freedom. *Nucl.Phys.*, B840:379–404, 2010, arXiv:1004.3697.

[68] I. Cherednikov and N. Stefanis. Transverse-momentum parton densities: Gauge links, divergences and soft factor. pages 363–369, 2010, arXiv:1008.0725.

[69] J. M. Drummond, J. Henn, G. P. Korchemsky, and E. Sokatchev. On planar gluon amplitudes/Wilson loops duality. *Nucl. Phys.*, B795:52–68, 2008, arXiv:0709.2368.

[70] L. F. Alday and J. M. Maldacena. Gluon scattering amplitudes at strong coupling. *JHEP*, 0706:064, 2007, arXiv:0705.0303.

[71] Y. Makeenko. Light cone Wilson loops and the string / gauge correspondence. *JHEP*, 0301:007, 2003, arXiv:0210256.

[72] G. P. Korchemsky, J. M. Drummond, and E. Sokatchev. Conformal properties of four-gluon planar amplitudes and Wilson loops. *Nucl. Phys.*, B795:385–408, 2008, arXiv:0707.0243.

[73] L. F. Alday, B. Eden, G. P. Korchemsky, J. Maldacena, and E. Sokatchev. From correlation functions to Wilson loops. *JHEP*, 1109:123, 2011, arXiv:1007.3243.

[74] N. Beisert, C. Ahn, L. F. Alday, et al. Review of AdS/CFT Integrability: An Overview. *Lett. Math. Phys.*, 99:3–32, 2012, arXiv:1012.3982.

[75] N. Arkani-Hamed and J. Trnka. Into the Amplituhedron. 2013, arXiv:1312.7878.

[76] N. Arkani-Hamed and J. Trnka. The Amplituhedron. 2013, arXiv:1312.2007.

[77] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, et al. Scattering Amplitudes and the Positive Grassmannian. 2012, arXiv:1212.5605.

[78] S. J. Parke and T. Taylor. An Amplitude for n Gluon Scattering. *Phys.Rev.Lett.*, 56:2459, 1986.

[79] F. A. Berends and W. Giele. Recursive Calculations for Processes with n Gluons. *Nucl.Phys.*, B306:759, 1988.

[80] E. Witten. Perturbative gauge theory as a string theory in twistor space. *Commun.Math.Phys.*, 252:189–258, 2004, arXiv:0312171.

[81] P. Mansfield. The Lagrangian origin of MHV rules. *JHEP*, 0603:037, 2006, arXiv:0511264.

[82] A. Brandhuber and G. Travaglini. Quantum MHV diagrams. 2006, arXiv:0609011.

[83] R. Britto, F. Cachazo, B. Feng, and E. Witten. Direct proof of tree-level recursion relation in Yang-Mills theory. *Phys.Rev.Lett.*, 94:181602, 2005, arXiv:0501052.

[84] M. B. Green, J. H. Schwarz, and L. Brink. N=4 Yang-Mills and N=8 Supergravity as Limits of String Theories. *Nucl.Phys.*, B198:474–492, 1982.

[85] S. Ivanov and G. Korchemsky. The structure of infrared poles in QCD. (in Russian). *Proceedings Quarks 84, Tbilisi*, 2:145–146, 1984.

[86] G. Korchemsky and A. Radyushkin. Infrared Asymptotics of Perturbative QCD. Quark and Gluon Propagators. *Sov.J.Nucl.Phys.*, 45:127, 1987.

[87] G. Korchemsky and A. Radyushkin. Infrared Asymptotics of perturbative QCD: Renormalization properties of the Wilson loops in higher orders of perturbation theory. *Sov.J.Nucl.Phys.*, 44:877, 1986.

[88] S. Ivanov, G. Korchemsky, and A. Radyushkin. Infrared Asymptotics of Perturbative QCD: Contour Gauges. *Yad.Fiz.*, 44:230–240, 1986.

[89] Z. Bern, L. J. Dixon, and V. A. Smirnov. Iteration of planar amplitudes in maximally supersymmetric Yang-Mills theory at three loops and beyond. *Phys.Rev.*, D72:085001, 2005, arXiv:0505205.

[90] J. Drummond, J. Henn, V. Smirnov, and E. Sokatchev. Magic identities for conformal four-point integrals. *JHEP*, 0701:064, 2007, arXiv:0607160.

[91] J. Drummond, J. Henn, G. Korchemsky, and E. Sokatchev. Dual superconformal symmetry of scattering amplitudes in N=4 super-Yang-Mills theory. *Nucl.Phys.*, B828:317–374, 2010, arXiv:0807.1095.

[92] T. Adamo, M. Bullimore, L. Mason, and D. Skinner. A Proof of the Supersymmetric Correlation Function / Wilson Loop Correspondence. *JHEP*, 1108:076, 2011, arXiv:1103.4119.

[93] M. Bullimore and D. Skinner. Holomorphic Linking, Loop Equations and Scattering Amplitudes in Twistor Space. 2011, arXiv:1101.1329.

[94] I. Korchemskaya and G. Korchemsky. Evolution equation for gluon Regge trajectory. *Phys.Lett.*, B387:346–354, 1996, arXiv:9607229.

[95] G. Korchemsky. Integrable structures and duality in high-energy QCD. *Nucl.Phys.*, B498:68–100, 1997, arXiv:9609123.

[96] T. Thiemann. *Modern Canonical Quantum General Relativity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2007.

[97] S. Lang. *Introduction to Differentiable Manifolds*. Universitext (1979). Springer, 2002.

[98] M. Artin. *Algebra*. Prentice Hall, 1991.

[99] J. Milne. *Algebraic Geometry: V5.0*. Taiaroa Publishing, 2005.

[100] S. Lane. *Categories for the Working Mathematician*. Graduate Texts in Mathematics. Springer New York, 1998.

[101] K. Brown. Hopf algebras. Lecture Notes, University of Glasgow.

[102] E. Beckenstein, L. Narici, and C. Suffel. *Topological algebras*. Number v. 24 in Notas de matemática. Elsevier Science, 1977.

[103] D. Williams. Lecture Notes on C^*-algbras. Department of Mathematics, Dartmouth College, May 2011.

[104] C. Flori. *A first course in topos quantum theory*. Lecture Notes in Physics. Springer, Berlin, 2013.

[105] T. Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. http://homotopytypetheory.org/book, Institute for Advanced Study, 2013.

[106] L. Tu. *An Introduction to Manifolds*. Universitext Series. Springer, 2010.

[107] K. T. Chen. Algebraic paths. *J. Alg.*, 9:8–36, 1968.

[108] K. T. Chen. Integration of paths: A faithful representation of paths by noncommutative formal power series. *Tr. Am. Math. Soc.*, 89(2):395–407, 1958.

[109] C. Di Bartolo, R. Gambini, and J. Griego. The Extended loop group: An Infinite dimensional manifold associated with the loop space. *Commun. Math. Phys.*, 158:217–240, 1993, arXiv:9303010.

[110] K.-T. Chen. Iterated integrals and exponential homomorphisms. s3-4(1):502–512, 1954.

[111] K. T. Chen. Algebras of iterated path integrals and fundamental groups. *Tr. Am. Math. Soc.*, 156:359–379, 1971.

[112] H.-M. Chan and S. T. Tsou. Gauge Theories in Loop Space. *Acta Phys.Polon.*, B17:259, 1986.

[113] K. Mackenzie. *Lie Groupoids and Lie Algebroids in Differential Geometry*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1987.

[114] T. Masson. Gauge theories in non-commutative geometry. 1446:73–98, 2012, arXiv:1201.3345.

[115] M. Nakahara. *Geometry, Topology and Physics, Second Edition* . Graduate student series in physics. Taylor & Francis, 2003.

[116] T. Frankel. *The Geometry of Physics: An Introduction*. The Geometry of Physics: An Introduction. Cambridge University Press, 2011.

[117] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry Set*. Wiley Classics Library. Wiley, 2009.

[118] A. Guay. Geometrical aspects of local gauge symmetry. 2004.

[119] U. Schreiber. Quantization via Linear homotopy types. February 2014, arXiv:1402.7041.

[120] V. Volterra. *Sui fondamenti della teoria delle equazioni differenziali lineari*.

[121] A. Mallios. *Topological Algebras: Selected Topics*. North-Holland Mathematics Studies. Elsevier Science, 2011.

[122] N. Bourbaki and H. Eggleston. *Topological Vector Spaces: Chapters 1-5*. 20. Jh.-. Springer, 2002.

[123] H. Schaefer and M. Wolff. *Topological Vector Spaces*. Graduate Texts in Mathematics. Springer New York, 1999.

[124] R. Schmid. Infinite-dimensional Lie groups and algebras in mathematical physics. *Adv.Math.Phys.*, 2010:280362, 2010.

[125] A. E. *Hopf Algebras*. Cambridge University Press, 1977.

[126] R. Giles. The reconstruction of gauge potentials from Wilson Loops. *Phys. Rev.*, D24:2160, 1981.

[127] R. Giles. Reconstruction of gauge potentials from wilson loops. *Phys. Rev.*, D24:2160–2168, Oct 1981.

[128] Y. M. Makeenko. *Non-Perturbative Methods in Gauge Theory*. Lecture notes. 1995.

[129] R. Hermann. *Differential geometry and the calculus of variations*. Interdisciplinary mathematics. Math Sci Press, 1977.

[130] P. A. M. Dirac. A new notation for quantum mechanics. *Math. Proc. of the Cambridge Phil. Soc.*, 35:416–418, 7 1939.

[131] R. P. Feynman. An Operator calculus having applications in quantum electrodynamics. *Phys.Rev.*, 84:108–128, 1951.

[132] G. Folland. *Real analysis: modern techniques and their applications*. Pure and applied mathematics. Wiley, 1999.

[133] C. C. Nishi. Simple derivation of general fierz-type identities. *Am. J. Phys.*, 73:1160–1163, 2005, arXiv:0412245.

[134] G. 't Hooft. A Planar Diagram Theory for Strong Interactions. *Nucl.Phys.*, B72:461, 1974.

[135] E. Brézin and S. Wadia. *The Large N Expansion in Quantum Field Theory and Statistical Physics: From Spin Systems to 2-dimensional Gravity*.

[136] G. Veneziano. Some Aspects of a Unified Approach to Gauge, Dual and Gribov Theories. *Nucl.Phys.*, B117:519–545, 1976.

[137] G. 't Hooft. Large N. 2002, arXiv:0204069.

[138] The 1/n expansion in atomic and particle physics. In G. Hooft, C. Itzykson, A. Jaffe, et al., editors, *Recent Developments in Gauge Theories*, volume 59 of *NATO Advanced Study Institutes Series*. 1980.

[139] S. R. Coleman. 1/N. *Lecture Notes*, 1980.

[140] O. Haan. Large N as a thermodynamic limit. *Phys.Lett.*, B106:207–210, 1981.

[141] Y. Makeenko and A. A. Migdal. Exact Equation for the Loop Average in Multicolor QCD. *Phys. Lett.*, B88:135, 1979.

[142] S. T. Tsou. Electric magnetic duality and the dualized standard model. *Int.J.Mod.Phys.*, A18S2:1–40, 2003, arXiv:0110256.

[143] I. Y. Arefeva. Quantum contour field equations. *Phys.Lett.*, B93:347–353, 1980.

[144] J.-L. Gervais and A. Neveu. The Slope of the Leading Regge Trajectory in Quantum Chromodynamics. *Nucl.Phys.*, B163:189, 1980.

[145] A. M. Polyakov. String Representations and Hidden Symmetries for Gauge Fields. *Phys.Lett.*, B82:247–250, 1979.

[146] A. M. Polyakov. Gauge Fields as Rings of Glue. *Nucl.Phys.*, B164:171–188, 1980.

[147] V. Dotsenko and S. Vergeles. Renormalizability of Phase Factors in the Nonabelian Gauge Theory. *Nucl.Phys.*, B169:527, 1980.

[148] R. A. Brandt, F. Neri, and M.-a. Sato. Renormalization of Loop Functions for All Loops. *Phys.Rev.*, D24:879, 1981.

[149] G. Korchemsky and A. Radyushkin. Loop Space Formalism and Renormalization Group for the Infrared Asymptotics of QCD. *Phys.Lett.*, B171:459–467, 1986.

[150] G. Korchemsky. Double Logarithmic Asymptotics in QCD. *Phys.Lett.*, B217:330–334, 1989.

[151] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken. Loop space and evolution of the light-like Wilson polygons. *Int. J. Mod. Phys. Conf. Ser.*, 20:109–117, 2012, arXiv:1208.5410.

[152] F. F. Van der Veken, I. O. Cherednikov, and T. Mertens. Evolution and Dynamics of Cusped Light-Like Wilson Loops in Loop Space. *AIP Conf. Proc.*, 1523:272–275, 2012, arXiv:1212.4345.

[153] I. O. Cherednikov, T. Mertens, and F. Van der Veken. Cusped light-like Wilson loops in gauge theories. *Phys. Part. Nucl.*, 44:250–259, 2013, arXiv:1210.1767.

[154] G. Korchemsky and A. Radyushkin. Renormalization of the Wilson Loops Beyond the Leading Order. *Nucl.Phys.*, B283:342–364, 1987.

[155] G. Korchemsky and G. Marchesini. Structure function for large x and renormalization of Wilson loop. *Nucl. Phys.*, B406:225–258, 1993, arXiv:9210281.

[156] I. A. Korchemskaya and G. P. Korchemsky. On lightlike Wilson loops. *Phys. Lett.*, B287:169–175, 1992.

[157] A. Bassetto, I. Korchemskaya, G. Korchemsky, and G. Nardelli. Gauge invariance and anomalous dimensions of a light cone Wilson loop in lightlike axial gauge. *Nucl. Phys.*, B408:62–90, 1993, arXiv:9303314.

[158] R. Begliuomini. Wilson loops in two-dimensional Yang-Mills theories. 1999, arXiv:9906180.

[159] R. A. Brandt, A. Gocksch, M. Sato, and F. Neri. Loop Space. *Phys. Rev.*, D26:3611, 1982.

[160] Y. Makeenko and A. A. Migdal. Quantum Chromodynamics as Dynamics of Loops. *Nucl. Phys.*, B188:269, 1981.

[161] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken. Evolution of cusped light-like Wilson loops and geometry of the loop space. *Phys. Rev.*, D86:085035, 2012, arXiv:1208.1631.

[162] Y. Makeenko and A. A. Migdal. Selfconsistent area law in QCD. *Phys.Lett.*, B97:253, 1980.

[163] Y. Makeenko. A Brief Introduction to Wilson Loops and Large N. *Phys. Atom. Nucl.*, 73:878–894, 2010, arXiv:0906.4487.

[164] Y. Makeenko. Topics in Cusped/Lightcone Wilson Loops. *Acta Phys. Polon.*, B39:3047–3080, 2008, arXiv:0810.2183.

[165] Y. Makeenko. The First thirty years of large-N gauge theory. *ITEP conf. Proc.*, pages 14–27, 2004, arXiv:0407028.

[166] Y. Makeenko and A. A. Migdal. Exact Equation for the Loop Average in Multicolor QCD. *Phys. Lett.*, B88:135, 1979.

[167] G. Leibbrandt. Introduction to the Technique of Dimensional Regularization. *Rev.Mod.Phys.*, 47:849, 1975.

[168] I. P. Zois. On Polyakov's basic variational formula for loop spaces. *Rept.Math.Phys.*, 42:373–384, 1998, arXiv:9703069.

[169] J. Gatheral. Exponentiation of eikonal cross sections in nonabelian gauge theories. *Phys. Lett.*, B133(1-2):90 – 94, 1983.

[170] J. Frenkel and J. Taylor. Non-abelian eikonal exponentiation. *Nucl. Phys.*, B246(2):231 – 245, 1984.

[171] T. Cheng and L. Li. *Gauge Theory of Elementary Particle Physics*. Oxford Science Publications. Clarendon Press, 1984.

[172] A. H. Mueller. Asymptotic behavior of the sudakov form factor. *Phys. Rev.*, D20:2037–2046, Oct 1979.

[173] J. C. Collins. Algorithm to compute corrections to the sudakov form factor. *Phys. Rev.*, D22:1478–1489, Sep 1980.

[174] A. Sen. Asymptotic behavior of the sudakov form factor in quantum chromodynamics. *Phys. Rev.*, D24:3281–3304, Dec 1981.

[175] I. O. Cherednikov, T. Mertens, P. Taels, and F. F. Van der Veken. Evolution of transverse-distance dependent parton densities at large-x_b and geometry of the loop space. *Int. J. of Mod. Phys. Conf. Ser.*, 25:1460006, 2014, arXiv:1308.3116.

[176] J. S. Schwinger. The Theory of quantized fields. 1. *Phys.Rev.*, 82:914–927, 1951.

[177] H.-n. Li. Resummation in hard QCD processes. *Phys.Rev.*, D55:105–119, 1997, arXiv:9604267.

[178] I. Cherednikov. Layout of Wilson lines and light-cone peculiarities of transverse-momentum dependent PDFs. *PoS*, QNP2012:061, 2012, arXiv:1206.4212.

[179] N. Stefanis. Worldline techniques and QCD observables. *Acta Phys.Polon.Supp.*, 6:71–80, 2013, arXiv:1211.7218.

[180] A. Accardi, J. Owens, and W. Melnitchouk. The CJ12 parton distributions. *PoS*, DIS2013:040, 2013.

[181] D. Freedman and A. Van Proeyen. *Supergravity*. Cambridge University Press, 2012.

[182] C. D'Orangeville and A. Lasenby. *Geometric Algebra for Physicists*. Cambridge University Press, 2003.

INDEX