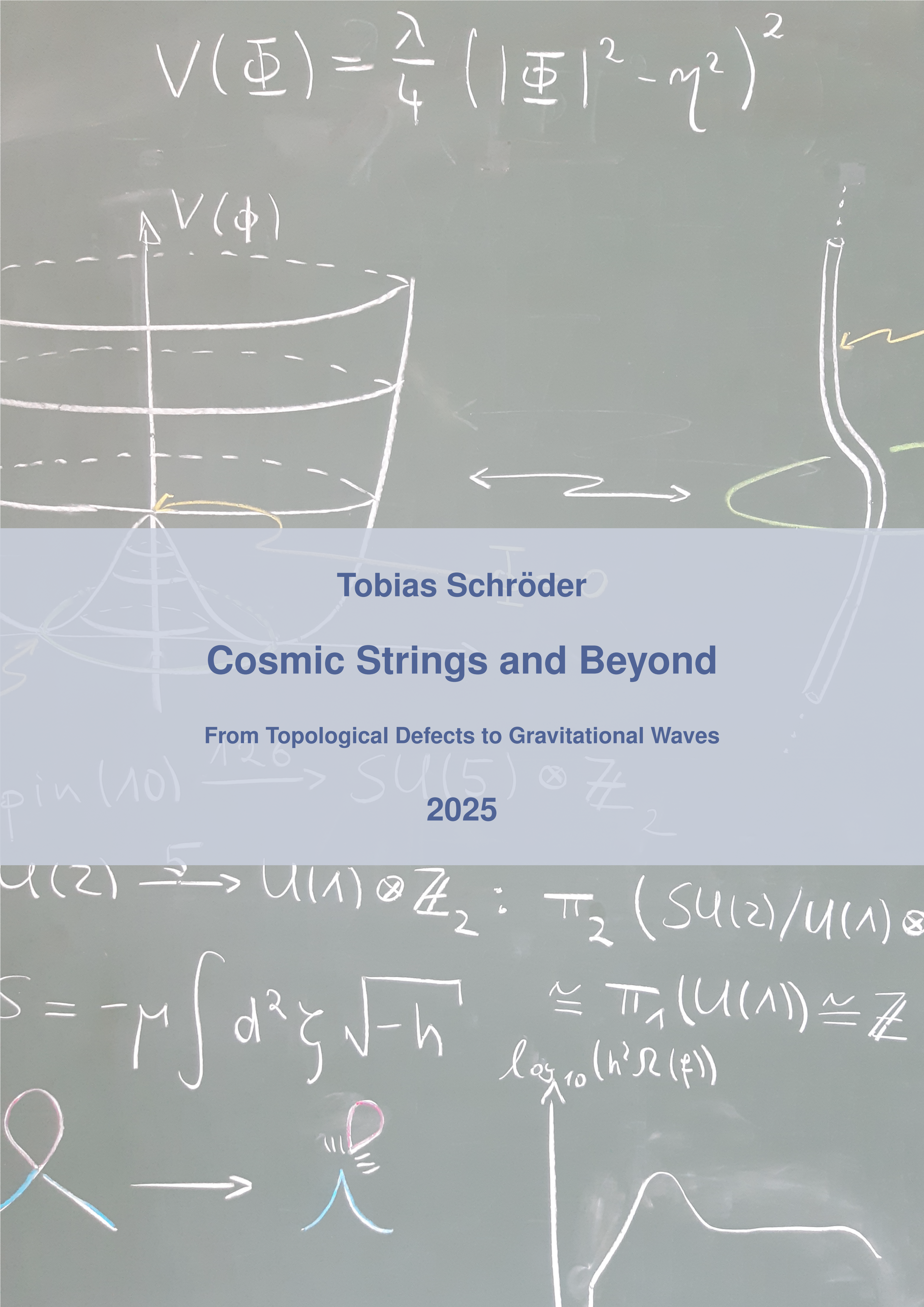


Tobias Schröder

Cosmic Strings and Beyond

From Topological Defects to Gravitational Waves

2025

Abstract

Cosmic defects are field configurations that are associated with non-vanishing energy but are prevented from decaying into the vacuum by topological or dynamical constraints. They arise naturally during cosmological phase transitions and are a generic product of symmetry-breaking chains in grand unified theories and other extensions of the Standard Model of particle physics. Cosmic defects are of great interest from a theoretical and observational perspective: They contain, on the one hand, plenty of information on the symmetries and the vacuum structure of particle physics theories, while, on the other hand, providing an extremely rich phenomenology. In particular, cosmic strings have gained increased attention as they are expected to produce a stochastic gravitational wave background (GWB) spanning many orders of magnitude in frequency. Such a GWB is attractive as it can be probed by many current and near-future gravitational wave observatories and, at the same time, contains information about the expansion history of the Universe from times when strings were formed until today. In fact, pulsar timing arrays (PTAs), including NANOGrav, have found strong evidence for a GWB.

We begin this thesis with a short introduction, providing an overview and motivation for all topics discussed.

The first part of this thesis starts by reviewing the physics of cosmological phase transitions, including a discussion of spontaneous symmetry breaking and high-temperature symmetry restoration, which we use to explain the origin of defects. We will then show in detail how different defects can be classified based on their vacuum topology. This framework can easily be extended to the more complicated metastable defects. Although the symmetry-breaking pattern and the resulting vacuum topology offer insights into the types and basic properties of defects, a complete understanding demands studying their field-theoretical structure. As an original contribution in this regard, we construct explicit field configurations for metastable cosmic strings containing monopoles with partially unconfined flux, providing a basis for future studies of string decay rates. Beyond topological and metastable defects, we comment on a variety of configurations that are stabilised in other ways, such as semilocal or melting defects. In particular, we propose a Standard Model extension that gives rise to plasma-stabilised embedded domain walls. The model is constructed such that electroweak symmetry remains unbroken inside the walls, even after the electroweak phase transition. We show that these walls can drive electroweak baryogenesis and compute the expected baryon-to-entropy ratio from a wall network. This mechanism is, over a wide range of parameters, able to explain the observed baryon asymmetry of the Universe.

The second part of this thesis is devoted to cosmic strings and their gravitational wave signatures. We begin by showing that, for cosmological purposes, (local) strings can be effectively described by the Nambu–Goto action. This exploits the fact that the width of strings is typically many orders of magnitude smaller than their curvature radius. We proceed to analyse the properties of solutions to the Nambu–Goto equations of motion, highlighting the origin of microscopic features known as kinks and cusps. Based on this treatment, we introduce the velocity-dependent one-scale (VOS) model to characterise the evolution of long-string networks and show how they can approach a scaling regime through energy loss into string loops. We describe how individual loops lose energy through the emission of gravitational radiation and combine this knowledge with the stochastic properties of the string network to derive the GWB spectrum from decaying loops. As a novel contribution, we will also discuss how the finite string width influences the GWB and at which frequency scales effects are to be expected.

Our analysis also uncovers a previously overlooked regime: low-scale cosmic strings. These are characterised by low string tensions and late formation times associated with large initial loop lengths. In contrast to the usually considered high-scale strings, none of the low-scale string loops will have fully decayed by the present time, resulting in a distinctive gravitational wave signature — an oscillating pattern in the GWB close to its peak, exhibiting local minima at integer multiples of

the first local minimum. We study in detail the new parameter regime in which low-scale strings arise. We also account for the possibility of loop decay into particle radiation and describe its effect on the loop evolution and the resulting spectra. We find that the features of low-scale-string GWB spectra are largely unaffected by particle decay, but the parameter space from which such strings arise may change drastically. The predicted spectra can, for parts of the parameter space, be probed by future experiments such as BBO and DECIGO.

For both low-scale and high-scale cosmic strings, computing the GWB signal is numerically very costly, especially in the case of large parameter scans. Such parameter scans are, however, necessary when evaluating models against observational data. Working with a simplified cosmological expansion history, we obtain fully analytical expressions for the resulting GWB spectra. These results, which represent a central original result of this thesis, are validated against numerical computations across a large region of parameter space. We also provide a detailed discussion of the GWB spectra and present explicit expressions for approximate power-law behaviours and characteristic break frequencies. Our results are valid at all frequencies and cover the entire conceivable range of parameters. Moreover, our computations are able to account for changes in the effective number of relativistic degrees of freedom.

Finally, we confront theoretical predictions with observational data from pulsar timing arrays. We use the NANOGrav 15-year data set, which shows evidence for a GWB in the nanohertz frequency range, to find preferred regions of parameter space and place competitive constraints on stable and metastable cosmic strings as well as on cosmic superstrings. In light of the data, we additionally compare all of the cosmic string models to a simplified model for a GWB from inspiraling supermassive black hole binaries.

We conclude the thesis with an overview, emphasising the original contributions of this work and pointing towards interesting future research directions.

Acknowledgements

I would like to thank those who supported me throughout the past three years of my PhD studies.

First and foremost, I would like to thank my advisor, Kai Schmitz, for offering me the opportunity to pursue a PhD in his back-then newly founded research group. It has been a privilege to witness the group's evolution, from two to around fifteen members now. I am especially grateful to Kai for creating a highly productive research environment and for his continuous support.

I would also like to thank Jochen Heitger for agreeing to serve as the second reviewer of my thesis and for being part of my defence committee. Similarly, my thanks go to Anton Andronic, who agreed to join this committee as well.

My sincere thanks go to the secretaries at the Institute for Theoretical Physics, Anne Volmer, Katharina Iker, and Petra Voß, who helped me navigate many formalities and administrative tasks.

I am grateful to all my friends and colleagues at the institute, who made both academic life and everyday work much more enjoyable. In particular, I want to thank Amin Abouibrahim, Adeela Afzal, Jan-Ravi Bade, Olivia Bitcon, Jonah Fabian, Adrian Finke, Katrin Greve, Doa Hashemi, Pia Leonie Jones, Jan-Christopher Knetsch, Justus Kuhlmann, Laura Moreno Valero, Vishnu Padmanabhan Kovilakam, Johannes Pirsch, Anne Poppe, Peter Risse, Rafael Lino dos Santos, Sasha Sobol, Luca Paolo Wiggering, and Jan Wissmann.
I am especially thankful to have had Richard von Eckardstein as my office mate for the past three years. We shared countless discussions — about physics, mathematics, and life. Over time, he has become a close friend.

My sincere gratitude also goes to Robert Brandenberger for hosting me at McGill University during a research visit that lasted over a month.
I am similarly thankful to J. J. Blanco-Pillado for supporting my participation in the 8th Cosmological Olentzero Workshop in Bilbao.
Furthermore, I am grateful to Huanchen Hu, Gregory Desvignes, and Michael Kramer for hosting me during a two-week research stay at the Max Planck Institute for Radio Astronomy in Bonn.

I would also like to thank my friends and colleagues who carefully proofread this thesis and gave me invaluable feedback on it: Jan-Ravi Bade, Olivia Bitcon, Richard von Eckardstein, Helia Etezadi, Jonah Fabian, Katrin Greve, Doa Hashemi, Justus Kuhlmann, Sasha Sobol, Luca Paolo Wiggering, and Jan Wissmann.

I thank my sister, my grandparents, and, above all, my parents. Without their support through my Bachelor's and Master's studies, I would never have been able to pursue a PhD.

To my friends outside of physics — Johannes, Kathi, Nicole, Steven, Tim, and Vanessa — thank you for your continued friendships, some of which have lasted now for more than twenty years.

Finally, I would like to thank Helia for her love and support during the final year of my PhD.

از تو آموختم چگونه در سختی‌ها لبخند بزنم.

List of publications

In the following, we provide a chronological list of the original papers that are included in this thesis. The author's contributions are summarised for each publication/preprint.

[1] NANOGrav Collaboration (A. Afzal et al.), *The NANOGrav 15 yr Data Set: Search for Signals from New Physics*, *Astrophys. J. Lett.* 951 (2023) 1, L11, Astrophys.J.Lett. 971 (2024) 1, L27 (erratum), [2306.16219]
This work was initiated and led by A. Mitridate and K. Schmitz, and was an effort of the NANOGrav collaboration. I contributed by cross-checking the GWB spectral data of the stable cosmic string models. I performed the data analysis for all cosmic string models (stable, metastable, and superstrings), which includes the generation of MCMC chains and their statistical evaluation. I also prepared the corresponding plots (except those showing median GWB spectra), which were subsequently finalised by A. Mitridate. Furthermore, I contributed to the software package PTArcade (see Ref. [9]), where I implemented different statistical methods (notably the K ratio) to derive parameter bounds from the MCMC chains and worked out sub-routines for plotting.

[2] T. Schröder and R. H. Brandenberger, *Embedded Domain Walls and Electroweak Baryogenesis*, *Phys. Rev. D 110, 043516*, [2404.13035]
This project was initiated by R. Brandenberger. I produced all original results for this paper. This includes the estimate of the baryon-to-entropy ratio from a network of domain walls with the corresponding plots, and the construction of the extension of the Standard Model that gives rise to plasma-stabilised embedded domain walls within which electroweak symmetry is restored. R. Brandenberger wrote the introduction and the conclusion and contributed through regular discussions. Together, we wrote the review section, while I wrote the text for the original results.

[3] K. Schmitz and T. Schröder, *Gravitational waves from low-scale cosmic strings*, *Phys. Rev. D 110, 063549*, [2405.10937]
I had the initial project idea, based on observations made when working on the project that led to Ref. [4]. I derived all expressions and formulae in the paper, did the numerical work, and created all the plots. K. Schmitz contributed through regular discussions on the project. Based on my notes of the computations, K. Schmitz wrote the manuscript, and together we finalised it.

[4] K. Schmitz and T. Schröder, *Gravitational waves from cosmic strings for pedestrians*, [2412.20907] (preprint)
K. Schmitz had the initial project idea. I derived all expressions and formulae in the paper. I carried out all numerical implementations and created all plots. K. Schmitz also contributed to the project through regular discussions. I wrote all parts of the manuscript and finalised it together with K. Schmitz.

[5] K. Schmitz and T. Schröder, *Gravitational waves from low-scale cosmic strings without scaling*, [2505.04537] (preprint)
K. Schmitz initiated the project. I derived all expressions and formulae in the paper, carried out all numerical implementations and created the plots. K. Schmitz also contributed through regular discussions on the project. I wrote the manuscript and completed it together with K. Schmitz.

The author, furthermore, contributed to the following articles that were published during his doctoral training, but are not part of this thesis:

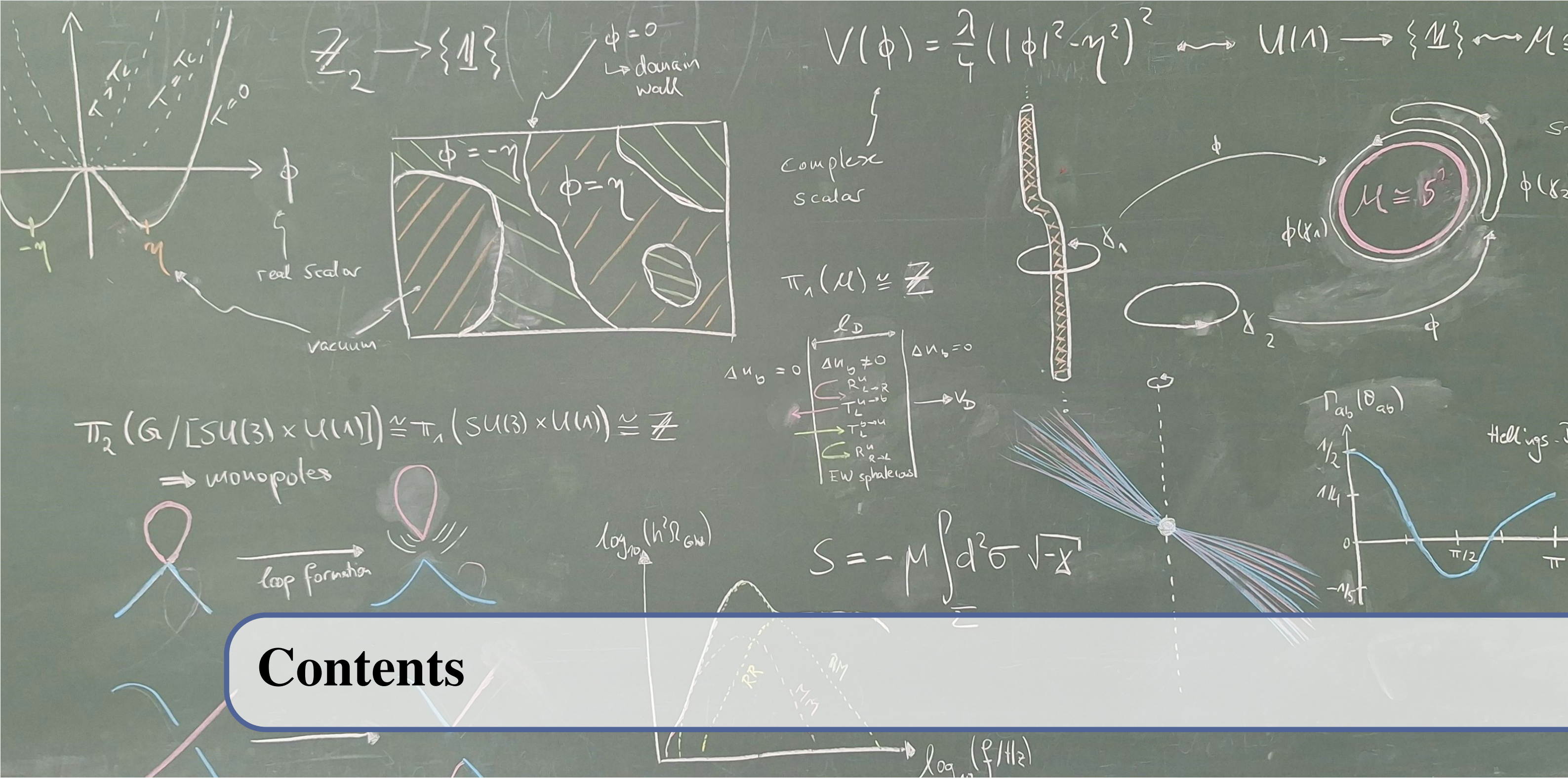
[6] M. E. Rojo and T. Schröder, *Asymptotic symmetries and memories of gauge theories in FLRW spacetimes*, *JHEP* 01 (2023) 11, arXiv: [2207.13726]

[7] N. Cordes, A. Mitridate, K. Schmitz, T. Schröder, K. Wassner, *On the overlap reduction function of pulsar timing array searches for gravitational waves in modified gravity*, Class. Quant. Grav. 42 (2025) 1, 015003, [2407.04464]

[8] NANOGrav Collaboration (G. Agazie et al.), *The NANOGrav 15 yr Data Set: Running of the Spectral Index*, *Astrophys.J.Lett.* 978 (2025) 2, L29 01 (2023) 11, [2408.10166]

The author also contributed to a software package for PTA data analyses, which was used in publication [1]:

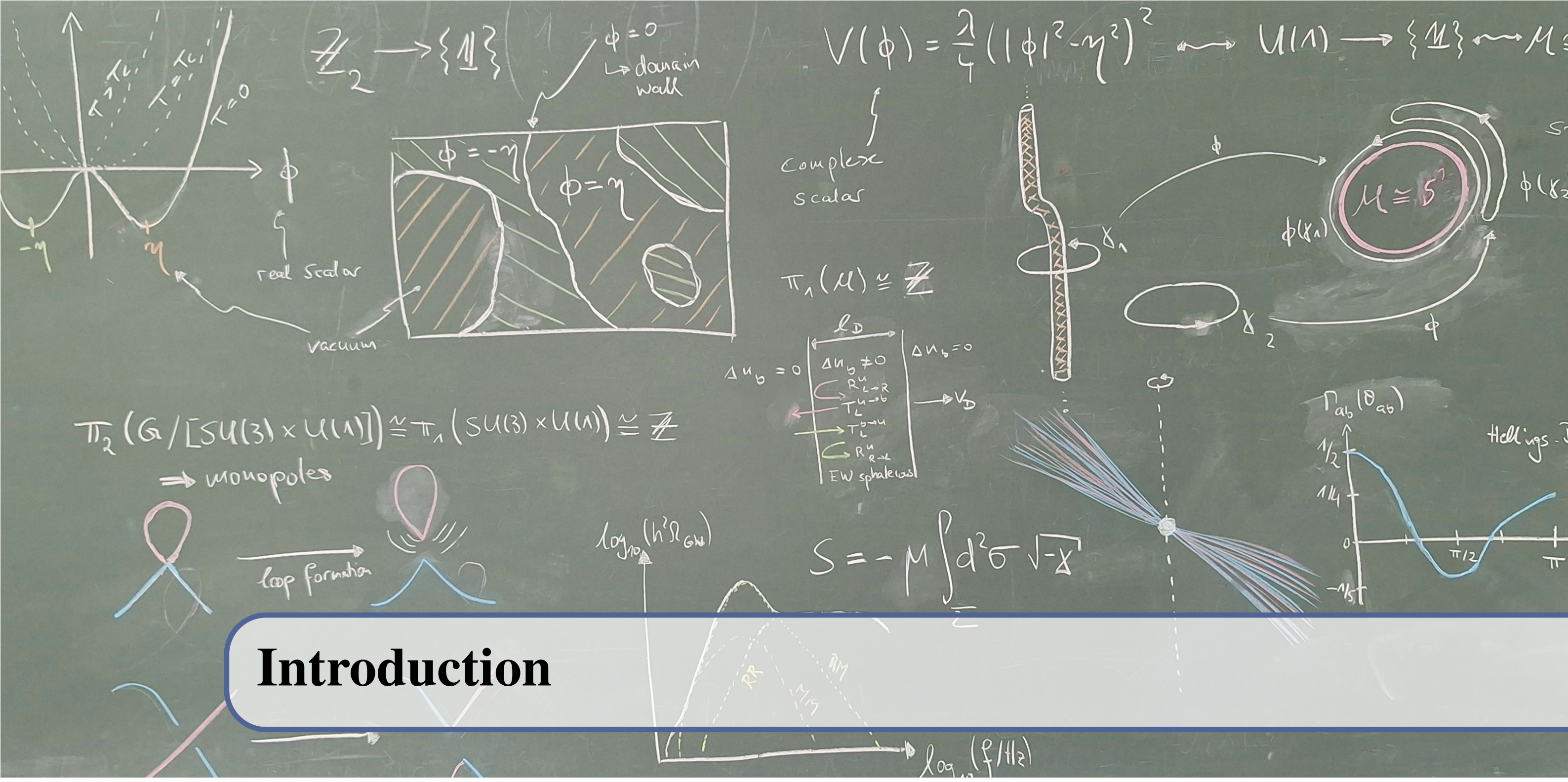
[9] A. Mitridate, D. Wright, R. v. Eckardstein, T. Schröder, J. Nay, K. Olum, K. Schmitz, T. Trickle, *PTArcade*, [2306.16377], Website: andrea-mitridate.github.io/PTArcade



Contents

I Cosmic Defects: Topology and Field Theory

II Gravitational Waves from Cosmic Strings



Introduction

The insight that our Universe is expanding dates back to the 1920s, when solutions to the general theory of relativity (GR) revealed that physical distances in a spatially homogeneous and isotropic spacetime evolve with a time-dependent scale factor [10, 11]. Support for this theoretical prediction soon followed with the discovery of a linear redshift–distance relation observed in nearby galaxies [12]. These results implied that an expanding Universe likely originated from a state of extreme temperature and density at a finite time in the past — a scenario later termed the “Big Bang”.

While modern cosmology often replaces the notion of an initial singularity with inflation, a bounce, or quantum gravitational effects beyond GR, it remains true that the early Universe was immensely hotter and denser than it is today. At temperatures as high as those in the early Universe, not only gravity is important, but also a detailed understanding of particle physics processes is essential. Our current best fundamental description of particle physics rests on the Standard Model (SM), formulated in the framework of relativistic quantum field theories.

Although the vast majority of predictions by the SM are experimentally verified with exceptional accuracy, the SM is known to be incomplete. For one, neutrinos in the SM are massless, which conflicts with the observation of neutrino oscillations [13]. Furthermore, for a successful description of observations of the cosmic microwave background (CMB) [14] dark matter and dark energy are needed, both of which lack a microscopic description in the SM.

With these discrepancies being only the most severe ones, it is clear that physics beyond the standard model (BSM) is needed. One possibility is to extend the SM to grand unified theories (GUTs) [15–17]. In these gauge theories, the SM gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$ is embedded in a larger simple gauge group G such as $SU(5)$, $\mathrm{Spin}(10)$, or E_6. These unifications are inspired by the observation that the running gauge couplings of the electroweak and strong interactions take on approximately the same value at high energies in the SM, and, more precisely, in supersymmetric extensions of the SM. Such a unification of couplings would naturally occur if there were an underlying GUT with a single gauge coupling. The energy scale of unification lies at around $\eta_{\mathrm{GUT}} \sim 10^{15} \ldots 10^{16}\,\mathrm{GeV}$, the GUT scale [18–21]. In addition, GUTs provide a natural reason for the observed quantisation of electric charges and allow to unify quarks and leptons within common representations of the gauge group G. While GUTs do not unify particle physics and gravity, they can still arise as low-energy effective theories of string theory.

Even if a fully unified theory turns out not to describe reality, the case for studying larger gauge groups that include the SM group remains compelling. These extended symmetries can impose additional structure on the low-energy theory, offering explanations for flavour hierarchies, the replication of fermion families, or dark matter candidates.

While such extended symmetry groups may be realised in Nature and restored at high temperatures in the early Universe [22], they must be spontaneously broken to the SM gauge group, possibly through a sequence of intermediate symmetry-breaking steps. It is during these symmetry-breaking phase transitions that cosmic defects — non-trivial, yet non-dissipative, field configurations — may form.

Defects are well known in condensed matter physics (see Ref. [23]), where they can actually be studied in a laboratory environment. A classical example is a ferromagnet cooled below its Curie temperature. To minimise its energy, the ferromagnet divides into domains of uniform magnetisation. However, the magnetisation will be randomly oriented in different domains. At the interface of two neighbouring domains, a simple type of defect forms, which is known as a domain wall. Other examples of defects in condensed matter physics notably include the magnetic flux lines in type-II superconductors [24], vortex lines in superfluid Helium-4 [25], or textures in Helium-3 [26, 27].

In cosmology, the first explicit and detailed discussions of defects date back to the 1970s in the form of Nielsen–Olesen vortices [28] or 't Hooft–Polyakov monopoles [29, 30]. These are non-trivial solutions to Abelian and non-Abelian Higgs models. Both configurations are associated with a non-vanishing energy density, which is, in the case of the vortex localised along a line and in the case of a monopole localised around a point. In 1974, it was furthermore realised that the spontaneous breaking of a discrete symmetry produces domains separated by domain walls, analogous to the magnetic domains in a ferromagnet [31].

A systematic understanding of the conditions under which topologically stable cosmic defects can arise, as well as a quantitative description of their formation, was developed in 1976 [32]. Remarkably, the question of whether stable defects form during a phase transition associated with the breaking of a symmetry Lie group G to a subgroup H is determined not by the specific dynamics of the theory, but solely by the topology of the vacuum manifold $\mathcal{M} \cong G/H$, more precisely, its homotopy structure. If the vacuum manifold is disconnected, the symmetry breaking leads to the formation of two-dimensional defects. These are the domain walls mentioned before. Likewise, a multiply-connected vacuum manifold implies the existence of tube-like strings, while a vacuum manifold whose second homotopy group is non-trivial stabilises point-like monopole configurations. Inside these defects, the phase transition cannot complete, and the symmetry remains unbroken. As a result, the defect carries a non-vanishing energy in a localised region of space.

The symmetry breaking from a larger symmetry group to the SM $G \to G_{\mathrm{SM}}$ does not have to occur in a single step. Indeed, in the case of GUTs, it is usually realised through a sequence of intermediate symmetry breakings, each associated with a phase transition that can give rise to defects. Defects originating from different stages of symmetry breaking may connect and destabilise each other [33]. However, if the energy scales separating the symmetry-breaking stages are sufficiently large, metastable defects can be long-lived or effectively stable. Both stable and metastable defects are a ubiquitous prediction of GUTs [34, 35]. It is clear that phenomenological implications will depend on the rates with which metastable defects decay. To accurately predict these rates from microphysics, one must first find the field configurations that describe the decay process of the metastable defect. We improve upon the literature by providing explicit field configurations describing the unwinding of metastable strings with unconfined flux in a specific model.

Stable and metastable defects do not comprise the entirety of cosmic defects. There is a variety of other non-topological or semi-topological defects, such as Q-balls, semilocal defects, or melting defects. While any specific SM extension may be speculative, it is the prevalence of their appearance in BSM models that strengthens the case for their relevance to cosmology. Furthermore, cosmic defects are associated with a wide range of observational signatures, including the emission of high-energy particles, imprints on the CMB and the large-scale structure, gravitational lensing effects, and the production of gravitational waves (GWs) (for an extensive overview, see Ref. [36]).

Despite their rich phenomenology, the observational signatures of topological defects come with a significant caveat: they can easily conflict with cosmological observations. A classical example is the monopole problem of GUTs, which generically predict the formation of monopoles with such high mass and abundance that they would overclose the Universe by many orders of magnitude [37, 38]. Several solutions have been proposed, most prominently cosmic inflation [39, 40], which dilutes the monopole abundance through a period of nearly exponential expansion. While domain walls are not as common in GUTs as monopoles, they pose a similar problem when they do arise. As the scale factor increases with time, the energy density of domain walls decreases more slowly than that of matter or radiation causing them to rapidly dominate the energy budget of the Universe [31]. Domain walls invoked to explain cosmological phenomena must therefore be unstable and have to decay at some point. This is typically achieved by introducing a bias term to the potential that lifts the degeneracy between the disconnected vacua, although other ways to obtain domain walls that are not completely stable have been explored in the literature (see, e.g., Refs. [41–44]). As another alternative, we provide in this thesis a specific extension of the SM and demonstrate the existence of domain walls that are only stabilised due to interactions with the early Universe plasma. Such walls must necessarily decay once plasma stabilisation becomes inefficient and can thereby avoid an overclosure problem. As an application, we demonstrate that a network of such walls can indeed be used to explain the observed overabundance of baryons over antibaryons in our Universe.

While models that involve the production of topologically stable domain walls and monopoles at the GUT scale require modifications or additional ingredients to avoid contradictions with observations, cosmic strings do not suffer from such difficulties. Their evolution approaches a scaling regime in which the total energy density in the string network remains a fixed fraction of the critical density, thereby avoiding any overclosure problem [45–47].

Cosmic strings themselves still give rise to a wide range of observational signatures. They can impact the CMB by altering the angular power spectrum of temperature anisotropies [48–51] through the introduction of step-like temperature discontinuities. This is known as the Gott–Kaiser–Stebbins effect [52–54]. In addition, they act as gravitational lenses [55–58], and can lead to an enhancement of the 21 cm hydrogen-line signal as a consequence of overdensities in string wakes [59, 60]. Beyond these gravitational effects, strings are able to produce cosmic rays [61–64]. If superconducting, strings can also give rise to CMB spectral distortions [65–67]. Moreover, cosmic strings may provide a mechanism for generating the observed baryon asymmetry of the Universe [68, 69], and have recently drawn renewed attention as potential seeds of structure formation, offering a possible explanation for the apparent overabundance of high-redshift galaxies that were observed by the James Webb Space Telescope [70]. Among the various observable signatures, GWs produced by cosmic strings offer one of the most compelling prospects for their detection.

Strings are topologically forbidden from having endpoints. Therefore, they must either be closed loops or infinitely long. While infinitely long strings are protected from decay as a consequence of winding number conservation, closed loops are not similarly protected. The decay channels of string loops depend on the nature of the considered symmetry breaking. For strings that arise from the spontaneous breaking of a global symmetry, the underlying model contains massless Nambu–Goldstone bosons. Although global strings inevitably produce gravitational waves [71–

73], they predominantly decay in Nambu–Goldstone radiation [74–78]. Strings that are due to the spontaneous breaking of a local symmetry primarily lose their energy through gravitational radiation [64, 79–81]. The focus of this thesis will be on the latter case.

Local strings associated with symmetry breaking at an energy scale η have an energy per unit length $\mu \propto \eta^2$, referred to as the string tension, and a width $\delta \propto \eta^{-1}$ [36]. Despite their narrowness, strings typically form with super-Hubble or even infinite length [82]. When they interact with each other or self-interact, they intercommute. In this process of exchanging string parts, long strings can chop off string loops of sub-Hubble size. These sub-Hubble loops oscillate under their tension and emit GWs, thereby gradually shrinking until they evaporate completely. On such closed loops, microstructures known as cusps and kinks can develop, which lead to short bursts of gravitational radiation with characteristic signatures [83–87]. Detecting strings in this way relies on a sufficiently nearby source whose beam of radiation furthermore needs to be directed towards the Earth. Instead of searching for such transient signals from a GW source that can be individually resolved, this thesis is concerned with the superposition of GWs from a network of cosmic string loops: a stochastic gravitational wave background (GWB). While the long-string network also emits gravitational radiation [73, 88–90], its contribution to the stochastic GWB is subdominant compared to that of loops [91].

The energy loss through loop production drives the long-string network towards the above-mentioned scaling regime, in which its energy density is a fixed fraction of the critical density. Since long strings cannot decay — this is in practice also true for finitely long super-Hubble loops as they are permanently stretched due to the cosmic expansion — the network continuously produces loops and, hence, GWs. The resulting stochastic GWB spans a vast range of frequencies: early-time contributions appear in the high frequency regime ($f \gtrsim$ MHz), while late-time emission dominates the low-frequency end ($f \lesssim$ nHz). Strings are therefore a particularly interesting source of gravitational radiation: On the one hand, their GWB spectrum encodes information about the expansion history of the Universe, which can reach from the time of string formation until today. On the other hand, their GW signature can be probed by a wide range of current and future GW experiments. These include ground-based interferometers such as LIGO [92], Virgo [93], and KAGRA [94, 95], space-borne detectors like LISA [96], DECIGO [97], and BBO [98], as well as pulsar timing arrays (PTAs) such as NANOGrav [99], EPTA [100], and SKA [101].

To establish a connection between GWB observations and cosmic strings as an underlying source, a proper theoretical understanding of the stochastic GWB from decaying string loops is necessary. In the present work, we make progress on that understanding through several advancements.

We show that the GWB signal from cosmic strings formed at a relatively low energy scale, $\eta \lesssim 10^9$ GeV, differs qualitatively from that produced by strings originating from higher energy scales. We refer to these strings as “low-scale cosmic strings”. The previously unrecognised features include a sharp cutoff frequency, f_{cut}, in the GW spectrum from the fundamental oscillation mode of closed string loops, and an oscillatory pattern in the total GW spectrum from summing over all modes, with local minima appearing at integer multiples of f_{cut}. These features arise because string loops with sufficiently low tension cannot shrink to vanishing size through GW emission alone within the age of the Universe. We explore in detail how these results are affected if not only gravitational decay but also particle decay influences the evolution of string loops, and demonstrate that for large regions of parameter space, this has minimal impact on our results. The obtained spectra open up the remarkable possibility of probing the discrete oscillation modes of closed string loops via GW observations. For cosmic strings formed at scales around $\eta \sim 10^9$ GeV, the distinctive spectral features lie within the sensitivity range of upcoming GW observatories such as BBO and DECIGO.

To interpret observed GWB spectra in terms of cosmic strings, a solid analytical understanding of the theoretical predictions across large parts of parameter space is essential. In this thesis, we derive analytic approximations for the total GW spectrum and provide comprehensive analytical expressions for key spectral features. These results extend beyond earlier works on this subject, capturing qualitatively distinct spectra across a wider range of parameters and all frequencies. In particular, our analytic expressions include the regime of low-scale cosmic strings. Beyond clarifying parameter dependencies of spectral features, our approach significantly reduces computational costs. While not always perfectly matching the precision of exact numerical methods, our fully analytical derivations avoid complicated integrals and sums, making them ideal for fast parameter scans with minimal accuracy loss.

Such parameter scans are already relevant since observational data are available: In 2023, PTA experiments found strong evidence for the existence of a stochastic GWB at nanohertz frequencies [102–105]. In this thesis, we confront theoretical predictions with the NANOGrav 15-year data set. We investigate different cosmic string models as a possible source of this signal and compare them to a simplified model of the standard interpretation: inspiraling supermassive black hole binaries. Our treatment places competitive constraints on stable and metastable strings as well as on cosmic superstrings.

As outlined in this introduction, cosmic defects, and cosmic strings in particular, connect early-Universe cosmology and GUT-scale physics with present and near-future observations, especially in the form of GWs. This thesis captures both aspects and naturally splits into two parts, which can be summarised as follows:

In Pt. I, we explore cosmic defects more broadly from a field-theoretical perspective. We begin in Ch. 1 with a review of cosmological phase transitions. We discuss spontaneously broken symmetries, which lays the foundation for our treatment of defects and, at the same time allows us to introduce notations and conventions. Following on this, we explain how symmetry restoration is achieved at high temperatures by discussing finite-temperature effective potentials. We conclude the first chapter with an overview of defect formation in the early Universe, including a discussion of the Kibble–Zurek mechanism. Having understood defect formation, Ch. 2 turns to the field-theoretical aspects and the topological stabilisation of cosmic defects. We start with a detailed study of domain walls, the simplest class of defects, and then generalise the discussion to classify topologically stable defects systematically. This is followed by a treatment of the field theory of cosmic strings and monopoles. Building on the topological arguments underlying stable defects, it is straightforward to introduce metastable defects. We first classify them and then explore metastable string field configurations in a specific particle physics toy model. We conclude the chapter with an overview of other types of defects such as composite defects, textures, and superconducting strings, all illustrated through toy models. Pt. I concludes with Ch. 3, where we examine how a network of domain walls can account for the observed baryon asymmetry of the Universe. Drawing on our previous discussions, we present an extension of the SM that gives rise to plasma-stabilised embedded domain walls capable of producing a net baryon number.

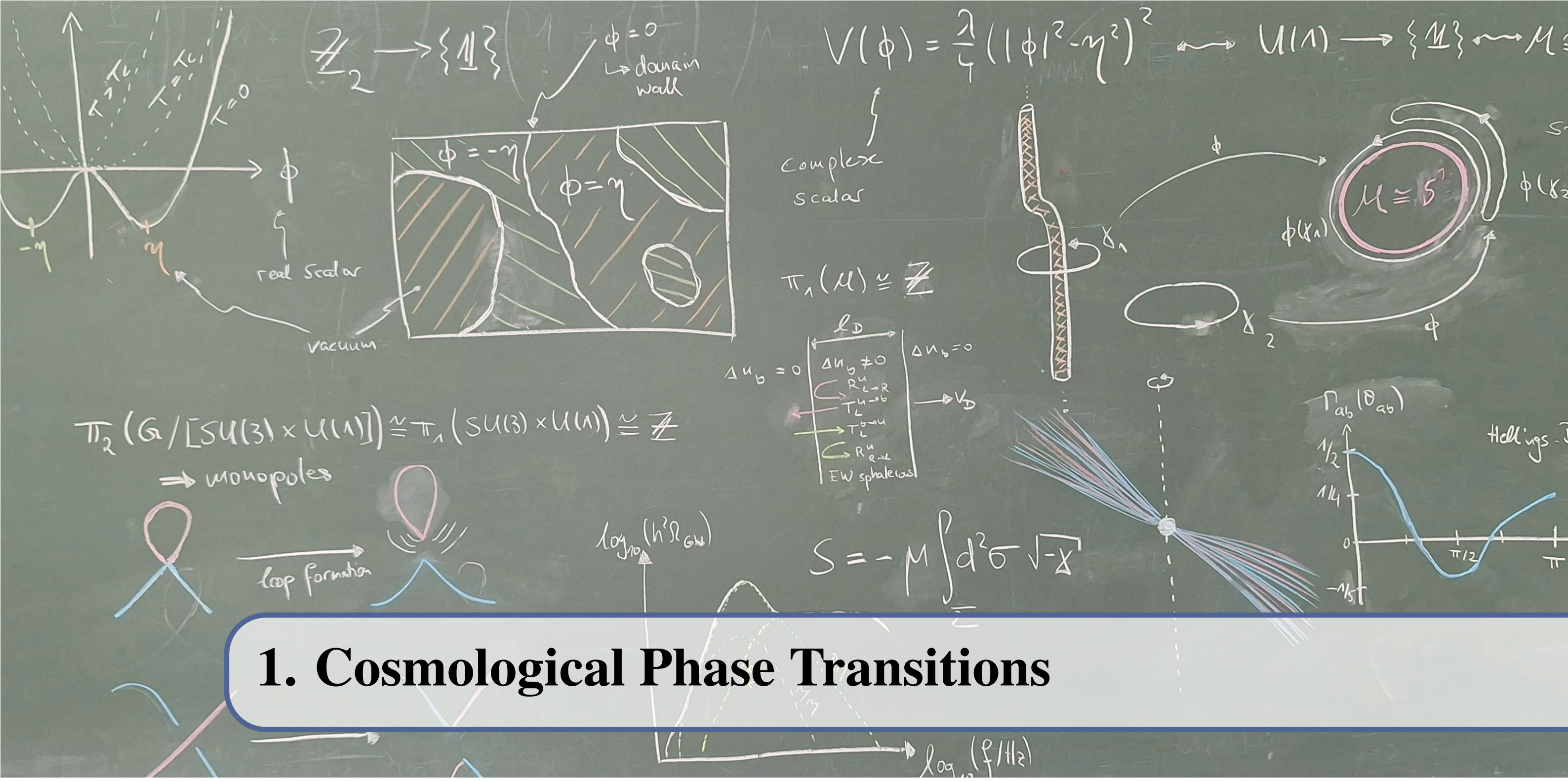
In Pt.II, our focus shifts to local cosmic strings as a potential source of a primordial GWB. In Ch. 4, we demonstrate that, on cosmological scales, the dynamics of local cosmic strings are effectively described by the Nambu–Goto action, derived from our preceding field-theoretical considerations. We then analyse in detail the equations of motion resulting from this action and study their solutions. Building on this, we describe the evolution of a cosmic string network, laying the groundwork for computing GWB spectra arising from a population of decaying string loops. Equipped with the necessary formalism, we introduce low-scale cosmic strings in Ch. 5. This chapter features an extended analysis of the characteristic oscillatory features in the GWB spectra produced by low-scale strings, both numerically and analytically. We critically examine the assumptions underlying these calculations and investigate the effects of particle emission from kink–kink collisions and cusps incorporated in nonscaling loop number densities. Ch. 6 returns to the GWB from cosmic strings in a more general context. We provide a step-by-step derivation of fully analytical expressions for the GWB spectrum. We begin by considering only the contribution of the fundamental oscillation mode of the loop to the spectrum, and then extend the analysis to include the contributions from all modes. At the end of the chapter, we provide a comprehensive overview of all analytical formulae needed to compute the GWB spectrum. In Ch. 7, we confront the theoretically predicted GWB spectra with observational data from the NANOGrav 15-year data set. Following a brief introduction to PTAs, we describe the data set and detail the data analysis employed. We then present the results, placing constraints on the parameter space of models of stable local strings, metastable strings, and superstrings.

The thesis concludes with a discussion of the original contributions and outlines promising directions for future research, as well as open questions that remain.

Notations and conventions: Throughout the thesis, we adapt units such that $\hbar = c = k_B = 1$, but keep Newton's constant G. We use the "mostly minus" $(+,-,-,-)$ signature for our metric. When considering cosmological spacetimes, we work with a spatially flat Friedmann–Lemaître–Robertson–Walker (FLRW) background spacetime, $\mathrm{d}s^2 = a^2\left(\eta\right)\left(\mathrm{d}\eta^2 - \mathrm{d}\boldsymbol{x}^2\right)$. For the Hubble constant, we use the value $H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, with reduced Hubble parameter $h = 0.674$ [14]. If not specified otherwise, the term "horizon" refers to the particle horizon that would occur in a Big Bang Cosmology model, but which is not necessarily the actual causal horizon if one assumes an early inflationary phase.

I

Cosmic Defects: Topology and Field Theory



1. Cosmological Phase Transitions

As the early Universe expanded and cooled, it most likely underwent a series of cosmological phase transitions, such as the electroweak or the QCD phase transition. In these phase transitions, an order parameter acquires a non-vanishing vacuum expectation value and spontaneously breaks the symmetry of the theory. In the context of this thesis, the order parameter corresponds to a scalar field referred to as the Higgs field. In Sec. 1.1, we review the spontaneous breaking of global and gauge symmetries by starting from a toy model and turning to a systematic approach afterwards. We use this as an opportunity to introduce some essential concepts, such as Goldstone's theorem and the Higgs mechanism, as well as some terminology and notation that we will heavily rely on in the remainder of Pt. I of this thesis. This discussion is completely based on classical field theory and furthermore neglects that fields in the early Universe do not evolve in a vacuum but in a thermal environment. We will address these issues in Sec. 1.2 by briefly introducing the concept of effective potentials. This will also allow us to understand the reason phase transitions occur at all. Finally, we want to sketch how phase transitions in the early Universe proceed. An important observation to understand the origin of defects is that phase transitions occur in the same way only on a certain length scale, the correlation length, and Higgs-field values become uncorrelated beyond that scale. We will explain how this insight can impose the formation of cosmic defects in Sec. 1.3.

1.1 Spontaneously broken symmetries

A symmetry of a theory is spontaneously broken if it does not leave a chosen ground state of this theory invariant. This has important consequences: It was first discovered by NAMBU in 1960 that the spontaneous breaking of a continuous global symmetry gives rise to massless modes [106]. GOLDSTONE then conjectured in 1961 that this is generally true and that such a symmetry breaking always leads to the occurrence of massless particles [107]. This statement, known as Goldstone's theorem, was subsequently proven by GOLDSTONE, SALAM, and WEINBERG in Ref. [108] in 1962, and the massless particles are now referred to as Nambu-Goldstone bosons.

When spontaneously breaking a local symmetry, the case is quite different. The massless modes of the global case will be unphysical as they can be removed by an appropriate gauge transformation from the theory. Instead, gauge fields will acquire an (effective) mass due to the spontaneous

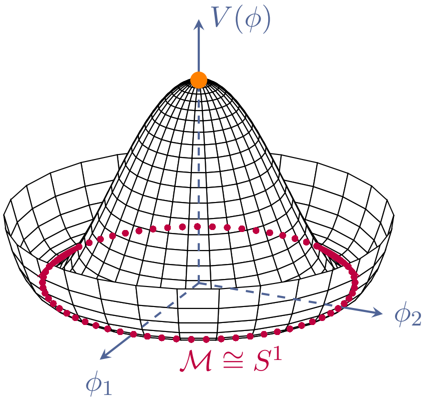


Figure 1.1: Plot showing the sombrero potential of Eq. (1.1). The orange dot marks the potential value in the unbroken phase $\phi = 0$. The degenerate minima of the potential, collectively referred to as the vacuum manifold $\mathcal{M}$, form a circle S^1, which is depicted as a red dotted line.

breaking of a gauge symmetry, which is nowadays referred to as the *Higgs mechanism* and was developed independently by three groups in 1964: HIGGS [109], ENGLERT and BROUT [110], as well as GURALNIK, HAGEN, and KIBBLE [111]. The original idea that gauge bosons can be effectively massive was already introduced by SCHWINGER in 1962 [112, 113] and shown in a non-relativistic setting by ANDERSON in 1963 [114].

A rigorous textbook treatment of the topic can be found in Ref. [115]. For a contemporary review covering many related topics that are subject to this thesis, see Ref. [116].

1.1.1 A toy model

Let us begin by considering a simple toy model of spontaneous symmetry breaking that already exhibits most of the features of more general models and is at the same time at the centre of the discussion of cosmic strings. Our toy model is the field theory of a complex scalar field $\phi(x) \in \mathbb{C}$ with the Lagrangian density

$$\mathcal{L} = (\partial_\mu \phi)^* (\partial^\mu \phi) - V(\phi) , \qquad \text{with} \qquad V(\phi) = \frac{\lambda}{4} \left(|\phi|^2 - \eta^2 \right)^2 , \tag{1.1}$$

and constants $\lambda, \eta > 0$. The potential V, known as the sombrero potential, is depicted in Fig. 1.1. For our discussion, it is often convenient to rewrite ϕ either in terms of a real and an imaginary part or in terms of an absolute value and a phase:

$$\phi = \frac{\phi_1 + i\phi_2}{\sqrt{2}} = \frac{1}{\sqrt{2}} \rho_\phi e^{i\vartheta_\phi} , \tag{1.2}$$

with $\phi_{1,2} \in \mathbb{R}$, $\rho_\phi \geq 0$, and $\vartheta_\phi \in [0, 2\pi)$.

Since the potential $V(\phi)$ only depends on the absolute value of the field $|\phi|$, it is invariant under phase shifts

$$\phi(x) \longmapsto e^{i\alpha(x)} \phi(x) . \tag{1.3}$$

By writing the complex field in its polar form, it is then obvious that the potential is independent of ϑ_ϕ and therefore independent of phase rotations $\vartheta_\phi \mapsto \vartheta_\phi + \alpha$. Since this shift in ϑ_ϕ corresponds to a (counter-clockwise) rotation by an angle α around the V-axis in Fig. 1.1, the potential must be invariant under such rotations. Note that for these transformations to be actual symmetries of the theory, they must not only leave the potential but also the kinetic terms of the Lagrangian invariant. While the potential is invariant under transformations for which α is spacetime-dependent, the kinetic term is not. Therefore, only transformations with $\alpha(x) = \mathrm{const.}$ are symmetries of the above theory, which correspondingly exhibits a *global* symmetry described by the group

$$U(1) = \left\{ e^{i\alpha} \;\middle|\; \alpha \in [0, 2\pi) \right\} . \tag{1.4}$$

We can now think about vacuum expectation values (VEVs) of the field ϕ. The energy density is clearly minimised if the field settles everywhere in space to the same value ϕ_0 with $|\phi_0| = \eta$, as only then the potential takes its minimum value $V(\phi_0) = 0$ and also the kinetic term and gradients vanish in this setting. Due to the symmetry, if ϕ_0 minimises the energy, then also $g\phi_0$ minimises the energy for any $g \in U(1)$. Let us, without loss of generality, assume that $\phi_0 = \eta$. Then the space of all minima we can obtain by acting with a $U(1)$ transformation on ϕ_0, i.e., its $U(1)$ orbit, is a circle:

$$\left\{ \eta e^{i\alpha} \;\middle|\; \alpha \in [0, 2\pi) \right\} \cong U(1) \cong S^1 . \tag{1.5}$$

This is nothing but the space of all minima of the potential $\left\{ (\phi_1, \phi_2) \in \mathbb{R}^2 \;\middle|\; \phi_1^2 + \phi_2^2 = 2\eta^2 \right\} \cong S^1$. Field values that minimise the potential, thereby representing ground states, are what we refer to as vacua, and the entirety of vacua forms the vacuum space, denoted $\mathcal{M}$. In Fig. 1.1, we show $\mathcal{M}$ as a red dotted circle at the bottom of the potential. The reason we consider this seemingly trivial vacuum space lies in its tremendous importance for cosmic defects as we will see in the next chapter.

Let us now consider the implications of the spontaneous breaking of the global $U(1)$ symmetry. Utilising the polar form of ϕ, we define a new field describing perturbations around the vacuum $\delta\rho_\phi = \rho_\phi - \sqrt{2}\eta$. By expressing the Lagrangian in terms of $\delta\rho_\phi$ and the canonically normalised field $\tilde{\vartheta}_\phi = \sqrt{2}\eta\vartheta_\phi$, we find

$$\mathcal{L} = \frac{1}{2}\partial_\mu\delta\rho_\phi\partial^\mu\delta\rho_\phi + \frac{1}{2}\partial_\mu\tilde{\vartheta}_\phi\partial^\mu\tilde{\vartheta}_\phi - \frac{\lambda\eta^2}{2}\delta\rho_\phi^2 + \mathcal{L}_{\mathrm{int}} , \tag{1.6}$$

where we absorbed interactions in $\mathcal{L}_{\mathrm{int}} = \partial_\mu\delta\rho_\phi\partial^\mu\tilde{\vartheta}_\phi - \lambda\eta/(4\sqrt{2})\delta\rho_\phi^3 - \lambda/16\delta\rho_\phi^4$. We observe that, due to spontaneous symmetry breaking, the above Lagrangian effectively describes an interacting field theory of a massless scalar field $\tilde{\vartheta}_\phi$ and another scalar field $\delta\rho_\phi$ with mass $m_s = \sqrt{\lambda}\eta$. The reason for this can also be understood heuristically: Perturbations of $\tilde{\vartheta}_\phi$ around the vacuum $\phi_0 = \eta$, i.e., $\tilde{\vartheta}_{\phi,0} = 0$, correspond to motions in field space along the $U(1)$ orbit of ϕ_0. A vacuum perturbed in this way will just be another vacuum and thus remain within $\mathcal{M}$. Since such a perturbation leads to no change in potential energy, we do not expect $\tilde{\vartheta}_\phi$ to be massive. In contrast to that, perturbations of $\delta\rho_\phi$ around the vacuum push the field in directions orthogonal to the orbit of $U(1)$, hence out of the vacuum, and a mass is expected.

The story is quite different when we promote the above global symmetry to a gauge symmetry. Such a model is known as the *Abelian Higgs model* discussed originally by HIGGS in his seminal article [109]. In this case, we need to introduce a gauge connection A_μ with covariant derivative $\mathcal{D}_\mu\phi = (\partial_\mu - ieA_\mu)\,\phi$ and field strength tensor $F_{\mu\nu} = 2\partial_{[\mu}A_{\nu]}$, where $e > 0$ denotes the coupling constant. The Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\mathcal{D}_\mu\phi)^*\,(\mathcal{D}^\mu\phi) - V(\phi) , \tag{1.7}$$

with the same potential as in Eq. (1.1) is then invariant under local $U(1)$ transformations for which α is allowed to be spacetime dependent. Here, the transformation of ϕ according to Eq. (1.3) is accompanied by a simultaneous change in the gauge field

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu\alpha . \tag{1.8}$$

Note that this symmetry gives us the freedom to choose a gauge in which we get rid of the field ϑ_ϕ: we simply need to perform a gauge transformation $\alpha(x) = -\vartheta_\phi(x)$ such that $\phi(x) = \eta + \delta\rho_\phi(x)/\sqrt{2}$. In the gauged theory, the one degree of freedom associated with $\tilde{\vartheta}_\phi$ is actually unphysical since we can remove it by a gauge transformation. Let us rewrite our Lagrangian in the given gauge:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\partial_\mu\delta\rho_\phi\partial^\mu\delta\rho_\phi - \frac{\lambda\eta^2}{4}\delta\rho_\phi^2 + e^2\eta^2 A_\mu A^\mu + \mathcal{L}_{\text{int}}\,, \tag{1.9}$$

where we absorbed the interactions in the term $\mathcal{L}_{\text{int}} = e^2 A_\mu A^\mu \delta\rho_\phi^2 - \lambda\eta/(4\sqrt{2})\delta\rho_\phi^3 - \lambda/16\delta\rho^4$.

The field $\delta\rho_\phi$ acquires the same mass as in the global case, whereas this time, we do not obtain a massless Goldstone boson, but the vector field A_μ acquires a mass $m_v = \sqrt{2}e\eta$ instead. Note that the spectrum of physical degrees of freedom consists of a massive scalar field $\delta\rho_\phi$ accounting for 1 degree of freedom and a massive spin-1 gauge field A_μ accounting for 3 degrees of freedom. The above mechanism by which gauge fields acquire an (effective) mass due to the spontaneous breaking of a gauge symmetry is a simple example of what is usually referred to as the Higgs mechanism.

1.1.2 General approach

Let us generalise the ideas that we illustrated by means of the simple Abelian model to a theory of n_Φ real scalar fields, collectively denoted by $\Phi = (\phi_1, \ldots, \phi_{n_\Phi})^T$ and described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\Phi\right)^T\left(\partial^\mu\Phi\right) - V(\Phi)\,, \tag{1.10}$$

with some arbitrary potential $V(\Phi)$. Let G be a compact n_G-dimensional Lie group that acts on these fields according to a real unitary (orthogonal) representation ρ of the symmetry group G such that any $g \in G$ acts on Φ via

$$\Phi \mapsto \rho(g)\Phi\,. \tag{1.11}$$

G is a (global) symmetry group of the above Lagrangian if $\forall g \in G$:

$$V(\rho(g)\Phi) = V(\Phi)\,. \tag{1.12}$$

Let us denote by Φ_0 one of the minima of V and define the subgroup $H \subset G$, which leaves this field value invariant:

$$H = \{\, g \in G \mid \rho(g)\Phi_0 = \Phi_0 \,\}\,. \tag{1.13}$$

This group is called the stabiliser subgroup (of G with respect to Φ_0). If $H \neq G$, we refer to the symmetry group G as being spontaneously broken to H. This statement can be rephrased as follows: A symmetry is spontaneously broken if the ground state of the theory (Φ_0) is not invariant under the symmetry group G of the Lagrangian but only under a proper subgroup H. Clearly, this is only possible if the field value in the vacuum is non-vanishing, i.e., $\Phi_0 \neq 0$. As in the $U(1)$ toy model, Φ_0 is only one of many vacua, and we can obtain the entire[1] vacuum space $\mathcal{M}$ as the orbit of Φ_0 under the symmetry group:

$$\mathcal{M} = \{\, \rho(g)\Phi_0 \mid g \in G \,\} \cong G/H\,, \tag{1.14}$$

---

[1] Here, and in the following, we will make the assumptions that there are no accidental symmetries as can arise from a fine-tuning of the parameters.

where the coset space $G/H = \{ \, gH \mid g \in G \, \}$ is the set of all left cosets $gH = \{ \, gh \mid h \in H \, \}$. The above statement is more formally known as the orbit-stabiliser theorem.[2] It will later be important that, since G and H are actually Lie groups, $\mathcal{M}$ is endowed with a unique smooth structure as a quotient manifold (for details see, e.g., Ref. [117]). These relations are very useful, as they allow us to find the vacuum manifold $\mathcal{M}$ by just knowing the symmetry breaking $G \to H$. The vacuum manifold is one of the most important concepts in our discussion, as its properties determine whether a spontaneous symmetry breaking leads to the formation of defects or not.

To find out more about the physical implications of the theory, let us make use of the fact that G is a Lie group and has a corresponding Lie algebra $\mathfrak{g}$. Let $T^a \in \mathfrak{g}$ be a set of n_G generators of the Lie algebra satisfying the Lie bracket relations

$$[T^a, T^b] = i f^{abc} T^c \, . \tag{1.15}$$

We refer to f^{abc} as the structure constants of G. Since G is compact, we can fix the normalisation of T^a and do so by requiring

$$\mathrm{tr}\left(T^a T^b\right) = \frac{1}{2}\delta^{ab} \, . \tag{1.16}$$

Any g in the universal cover of G can be written as

$$g = \exp(i\gamma_a T^a) \, , \tag{1.17}$$

where γ_a is a set of n_G real numbers. Let us denote by ρ' the Lie algebra representation corresponding to ρ.[3] We can then write for simplicity $\tau^a = \rho'(T^a)$ for the generators in this representation. Note that, since ρ is an orthogonal representation, these generators are antisymmetric and imaginary.

In this notation, Eq. (1.12) implies for all generators τ^a of the symmetry that

$$(\nabla_\Phi V)\tau^a \Phi = 0 \, . \tag{1.18}$$

Without loss of generality, we can also choose the n_G generators τ^a such that the first $n_H = \dim(H) \leq n_G$ generators τ^{a_u} with $a_u \in \{ \, 1, \ldots, n_H \, \}$ are the generators of the Lie algebra corresponding to H, and we give their indices a subscript u for "unbroken". Similarly, we denote the indices of the remaining generators $a_b \in \{ \, n_H + 1, \ldots, n_G \, \}$ with subscript b for "broken". The definition of H in Eq. (1.13) immediately implies for the unbroken generators

$$\tau^{a_u} \Phi_0 = 0 \, . \tag{1.19}$$

Let us now denote in analogy to the previous section $\Phi = \Phi_0 + \delta\Phi$. Assuming $\delta\Phi$ to be small, we can expand the Lagrangian around the minimum field value Φ_0 and obtain up to second order

$$\mathcal{L} = \frac{1}{2}\partial_\mu \delta\Phi^T \partial^\mu \delta\Phi - \frac{1}{2}\delta\Phi^T \mu^2(\Phi_0)\delta\Phi + \mathcal{L}_{\mathrm{int}} \, , \tag{1.20}$$

where we absorbed interaction terms of order $\mathcal{O}(\delta\Phi^3)$ in $\mathcal{L}_{\mathrm{int}}$. Without loss of generality, we assumed that we can choose $V(\Phi_0) = 0$ and that $\nabla_\Phi V(\Phi_0) = 0$ since Φ_0 is an extremum of V. The matrix $\mu^2(\Phi_0)$ is given by its elements

$$\mu^2_{ij}(\Phi_0) = \frac{\partial V}{\partial\phi_i \partial\phi_j}(\Phi_0) \, , \tag{1.21}$$

---

[2] For better intuition, we sketch the simple proof: Two points in $\mathcal{M}$ are equivalent if $\rho(g)\Phi_0 = \rho(g')\Phi_0 \Leftrightarrow \Phi_0 = \rho(g^{-1}g')\Phi_0$, which means that $g^{-1}g' \in H$ and hence $gH = g'H$. Therefore, the map $\mathcal{M} \to G/H$, $\rho(g)\Phi_0 \mapsto gH$ is an isomorphism.

[3] The correspondence is as follows. If $\rho : G \to GL(V)$, then $\rho' : \mathfrak{g} \to GL(V)$, $\rho'(X) = -i\frac{\mathrm{d}}{\mathrm{d}s}\rho\left(\exp\left(iXs\right)\right)|_{s=0}$

and its eigenvalues satisfy $\lambda_i = m_i^2$, with m_i being the eigenvalues of the mass eigenstates. The eigenvalues are guaranteed to be non-negative $\lambda_i \geq 0$ since Φ_0 is a local minimum of V. Differentiation of Eq. (1.18) and evaluation at Φ_0 implies

$$\mu^2 \tau^a \Phi_0 = 0 \,. \tag{1.22}$$

The $n_G - n_H$ broken generators τ^{a_b} satisfy by definition $\tau^{a_b}\Phi_0 \neq 0$ and the vectors $\tau^{a_b}\Phi_0$ are linearly independent (linear dependence would imply more unbroken generators and thus lead to a contradiction with our assumptions). This fact tells us via Eq. (1.22) (and the assumption of the absence of accidental symmetries) that $\dim(\ker\{\mu^2\}) = n_G - n_H$ and there are, therefore, exactly $n_G - n_H$ vanishing eigenvalues corresponding to $n_G - n_H$ massless fields associated with the broken generators. These are just the Goldstone bosons we encountered in the previous subsection and the general statement of Goldstone's theorem can now be summarised as follows: The spontaneous breaking of a continuous global n_G-dimensional symmetry (Lie) group G to an n_H-dimensional (Lie) subgroup H requires the spectrum of the theory to contain $n_G - n_H$ massless bosonic degrees of freedom. Note that this also tells us that the remaining $n_\Phi - n_G + n_H$ real fields will become massive.

After this discussion of a global symmetry, we again want to turn to local symmetries. We introduce a $\mathfrak{g}$-valued gauge field $A_\mu = A_\mu^a T^a$ and the covariant derivative

$$\mathcal{D}_\mu \Phi = \left(\partial_\mu - icA_\mu^a \tau^a\right)\Phi \,, \tag{1.23}$$

where $c > 0$ denotes the coupling constant. Furthermore, we will make use of the field strength tensor

$$F_{\mu\nu} = 2\partial_{[\mu}A_{\nu]} - ic[A_\mu, A_\nu] \,. \tag{1.24}$$

The Lagrangian we want to consider is of the form

$$\mathcal{L} = -\frac{1}{2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu}) + \frac{1}{2}\left(\mathcal{D}_\mu\Phi\right)^T\left(\mathcal{D}^\mu\Phi\right) - V(\Phi) \,. \tag{1.25}$$

It is invariant under gauge transformations $g(x) \in G$:

$$\Phi \longmapsto \rho(g)\Phi, \qquad\qquad A_\mu \longmapsto gA_\mu g^{-1} + \frac{i}{c} g\partial_\mu g^{-1} \,. \tag{1.26}$$

Playing the same game as in the toy model, we can get rid of the Goldstone degrees of freedom by choosing an appropriate gauge. To achieve this, we need a gauge transformation $\Phi \mapsto \tilde{\Phi} = \rho(g)\Phi$, such that

$$\tilde{\Phi}^T \tau^a \Phi_0 = 0 \,. \tag{1.27}$$

To understand why this is indeed the required condition, recall that the vectors $\tau^{a_b}\Phi_0 \neq 0$ spanned the $n_G - n_H$-dimensional vector space associated with Goldstone bosons. The above condition amounts to the statement that $\tilde{\Phi}$ is orthogonal to that space and, therefore, does not contain any Goldstone modes. That this so-called *unitarity gauge* in which Goldstone modes are absent can always be chosen is, e.g., shown in Ref. [115]. By fixing this gauge, we have removed $n_G - n_H$ unphysical degrees of freedom from the theory.

Let us now write $\tilde{\Phi} = \Phi_0 + \delta\Phi$ such that the antisymmetry of τ^a implies via Eq. (1.27) that $\delta\Phi^T\tau^a\Phi_0 = 0$. From the Lagrangian in Eq. (1.25), we find

$$\mathcal{L} = -\frac{1}{2}\mathrm{tr}\left(F_{\mu\nu}F^{\mu\nu}\right) + \frac{1}{2}M_{ab}^2(\Phi_0)A_\mu^a A_\mu^b + \frac{1}{2}\partial_\mu\delta\Phi^T\partial^\mu\delta\Phi - \frac{1}{2}\delta\Phi^T\mu^2\delta\Phi + \mathcal{L}_{\mathrm{int}} \,, \tag{1.28}$$

where we introduced the square mass matrix for the gauge fields,

$$M_{ab}^2(\Phi_0) = c^2 \Phi_0^T \tau^a \tau^b \Phi_0 \,, \tag{1.29}$$

and absorbed terms of higher than quadratic order in the fields $\delta\Phi$ and A_μ into $\mathcal{L}_{\text{int}}$. Note that the terms leading to a mixing between $\delta\Phi$ and A_μ dropped out as a consequence of Eq. (1.27).

A few remarks are in order. First, note that even though the terms in the gauge theory of Eq. (1.28) containing $\delta\Phi$ seem to be the same as in Eq. (1.20), where there was only a global symmetry, they are not. In the latter case, $\delta\Phi$ consisted of n_Φ independent real fields; in the former case, we have removed $n_G - n_H$ of these by imposing the gauge condition given in Eq. (1.27). Writing the Lagrangians out explicitly, in the global case, we would find n_Φ real fields $\delta\phi_i$, but in the gauged case, only $n_\Phi - n_G + n_H$. Second, let us consider the gauge field mass matrix in Eq. (1.29). Since Φ_0 has real components and the τ^a are imaginary and antisymmetric, the mass matrix M_{ab}^2 must be real, symmetric, and positive semidefinite as it should be. Furthermore, exactly n_H of its eigenvalues vanish.[4] This means that $n_G - n_H$ of the gauge bosons, namely those corresponding to broken generators, acquire a mass, while n_H massless gauge bosons remain. These are the gauge bosons of the unbroken subgroup H.

In summary, the particle spectrum of the theory after the spontaneous breaking of the gauge symmetry consists of

- $n_\Phi - n_G + n_H$ massive Higgs modes
- $n_G - n_H$ massive gauge bosons
- n_H massless gauge bosons.

Finally, we want to include, for completeness, fermions into our Lagrangian. This can be achieved by adding the term

$$\mathcal{L}_F = i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - \bar{\psi} Y \psi \Phi \,. \tag{1.30}$$

Here, ψ is a fermion multiplet corresponding some representation $\tilde{\rho}$ of G and its gauge-covariant derivative reads $\mathcal{D}_\mu \psi = \left(\partial_\mu - icA_\mu^a \tilde{\rho}'(T^a)\right)\psi$. As usual, γ^μ denote the Dirac matrices (adapted to the respective representation of ψ), and $\bar{\psi} = \psi^\dagger \gamma^0$ is the Dirac adjoint. The matrix Y is the Yukawa coupling matrix such that the fermion mass matrix is, after the symmetry breaking, given by

$$m_F(\Phi_0) = Y\Phi_0 \,. \tag{1.31}$$

Let us apply the above formalism to two simple examples that will become relevant when introducing cosmic defects.

■ **Example 1.1 — Global $SU(2) \to \{\,\mathbb{1}\,\}$.** We begin with the simple symmetry breaking of a global symmetry $SU(2) \to \{\,\mathbb{1}\,\}$. From Eq. (1.14), we know that the vacuum manifold will be a three-dimensional sphere:

$$\mathcal{M} \cong SU(2)/\{\,\mathbb{1}\,\} \cong S^3 \,. \tag{1.32}$$

Since we have $\dim(\mathcal{M}) = 3$, we will obtain 3 Goldstone bosons. The symmetry can be broken in the above way by using a field Φ associated with the fundamental representation of $SU(2)$, i.e., the doublet representation. We can express Φ in terms of four real scalar fields $\phi_1, \ldots, \phi_4$:

$$\Phi = (\Phi_1, \Phi_2)^T = \frac{1}{\sqrt{2}} \left(\phi_1 + i\phi_2, \phi_3 + i\phi_4\right)^T \,. \tag{1.33}$$

---

[4] This can be seen as follows. If we fix an index of the mass matrix corresponding to an unbroken generator then, by means of Eq. (1.19), the mass matrix vanishes $M_{a_u b}^2 = 0$, so the first n_H columns and rows of the mass matrix vanish, implying $\dim(\ker(M^2)) \geq n_H$. Since $\tau^{a_b}\Phi_0 \neq 0$ are linearly independent, the remaining non-vanishing block must be of full rank such that $\text{rank}(M^2) = n_G - n_H$, which forces, via the rank-nullity theorem, $\dim(\ker(M^2)) = n_H$. Hence, M^2 has exactly n_H vanishing eigenvalues.

This means that $n_\Phi = 4$ and, hence, we know that we will obtain one massive Higgs mode. The $SU(2)$-invariant Lagrangian we want to consider reads

$$\mathcal{L} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - \frac{\lambda}{4}\left(\Phi^\dagger\Phi - \eta^2\right)^2 . \tag{1.34}$$

In terms of the real scalar fields, the minimum of the potential is found to be

$$\mathcal{M} \cong \left\{ (\phi_1, \phi_2, \phi_3, \phi_4) \in \mathbb{R}^4 \,\middle|\, \sum_{i=1}^{4} \phi_i^2 = 2\eta^2 \right\} \cong S^3 . \tag{1.35}$$

This is just the description of a 3-sphere embedded in $\mathbb{R}^4$ and agrees with our statement in Eq. (1.32). Choosing $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = \sqrt{2}\eta$, we can easily find that the only non-vanishing eigenvalue of μ^2 is $m_s^2 = \lambda\eta^2$, such that m_s is the only non-vanishing mass and it is associated with the field $\delta\phi_3 = \phi_3 - \sqrt{2}\eta$, whereas $\phi_{1,2,4}$ are the massless Goldstone modes. Note that while the particle spectrum would be affected if we had considered a local symmetry breaking, the vacuum $\mathcal{M}$ would remain the same. ■

■ **Example 1.2 — Local $SU(2) \to U(1)$.** In our second example, we want to break the gauge group $SU(2)$, although not completely, but instead leave an unbroken Abelian subgroup $U(1)$. Such a symmetry breaking results in a vacuum manifold diffeomorphic to a two-sphere:

$$\mathcal{M} \cong SU(2)/U(1) \cong S^2 . \tag{1.36}$$

This will lead to $\dim(SU(2)) - \dim(U(1)) = 2$ massive gauge bosons and leave $\dim(U(1)) = 1$ massless gauge field. The symmetry breaking can be achieved by choosing Φ to belong to the adjoint representation of $SU(2)$, i.e., the triplet representation. Denoting by σ^a the three Pauli matrices, the $SU(2)$ generators satisfying the normalisation Eq. (1.16) are $T^a = \sigma^a/2$ and we can decompose Φ into three real scalar fields according to

$$\Phi = \phi_a T^a = \frac{1}{2}\begin{pmatrix} \phi_3 & \phi_1 - i\phi_2 \\ \phi_1 + i\phi_2 & -\phi_3 \end{pmatrix} . \tag{1.37}$$

Since $n_\Phi = 3$, the symmetry breaking will give rise to one massive Higgs mode. In this representation, local $SU(2)$ transformations $U(x) = \exp\left(i\alpha^a(x)T^a\right)$ with gauge parameters $\alpha^{1,2,3}(x)$ act on Φ according to

$$\Phi(x) \longmapsto U(x)\Phi(x)U^{-1}(x) . \tag{1.38}$$

Similarly, we can decompose the gauge field as $A_\mu = A_\mu^a T^a$, and the Lagrangian of our theory is

$$\mathcal{L} = -\frac{1}{2}\mathrm{tr}\left(F_{\mu\nu}F^{\mu\nu}\right) + \mathrm{tr}\left(\mathcal{D}_\mu\Phi\mathcal{D}^\mu\Phi\right) - \frac{\lambda}{4}\left(\mathrm{tr}\left(\Phi^2\right) - \eta^2\right)^2 , \tag{1.39}$$

where $F_{\mu\nu}$ is the field strength tensor associated with A_μ and $\mathcal{D}_\mu\Phi = \partial_\mu\Phi - ic[A_\mu, \Phi]$. The vacuum, as obtained by minimising the potential, is

$$\mathcal{M} \cong \left\{ (\phi_1, \phi_2, \phi_3) \in \mathbb{R}^3 \,\middle|\, \sum_{i=1}^{3} \phi_i^2 = 2\eta^2 \right\} \cong S^2 , \tag{1.40}$$

which is just a two-sphere embedded in $\mathbb{R}^3$ in agreement with Eq. (1.36). Without loss of generality, we can choose $\phi_1 = \phi_2 = 0$ and $\phi_3 = \sqrt{2}\eta$ and fix the unitarity gauge. For the mass of the Higgs $\delta\phi_3 = \phi_3 - \sqrt{2}\eta$, we find $m_s = \sqrt{\lambda}\eta$ and the gauge bosons A_μ^1 and A_μ^2 obtain the effective masses $m_v = \sqrt{2}c\eta$ while A_μ^3 remains the massless gauge boson of the unbroken $U(1)$ subgroup. ■

1.2 Finite-temperature effective potentials

In the previous section, we discussed the spontaneous breaking of a symmetry solely based on considering purely classical potentials. Fundamentally, the fields we are considering are quantum fields and will lead via (self) interactions to radiative corrections of the potential. Furthermore, at early times, the potential will also be strongly influenced by thermal fluctuations. All these effects can be taken into account by replacing the potential V with a (temperature-dependent) effective potential V_{eff}. An accessible textbook introduction to the topic can be found in Ref. [118]; a more technical and detailed one is presented in Ref. [119]. Classical reviews include Ref. [120], and the cosmology-focused Ref. [121]. For more recent reviews, see Refs. [122], and [123].

Let us begin at zero temperature and briefly describe how effective potentials come about. As usual in quantum field theories, we can start from the generating functional of full Green functions involving scalar fields $\Phi(x)$:

$$Z[J] = \left\langle 0 \middle| \mathcal{T} \left\{ \exp\left(i \int \mathrm{d}^4x \, J(x)\Phi(x) \right) \right\} \middle| 0 \right\rangle , \tag{1.41}$$

with an external source J. $\mathcal{T}$ denotes time-ordering, $|0\rangle$ is the full physical vacuum and Φ the field in the Heisenberg picture. We can then obtain expectation values as $\Phi_{\text{cl}}(x) \equiv \langle \Phi \rangle_J (x) = -i/Z[J] \delta Z[J]/\delta J(x)$. Introducing $W[J] = -i \ln(Z[J])$, which is the generating functional of connected Green functions and satisfies $\Phi_{\text{cl}}(x) = \delta W[J]/\delta J(x)$, one can define the so-called effective action as its Legendre transform, trading J for its conjugate Φ_{cl}, namely $\Gamma[\Phi_{\text{cl}}] = W[J] - \int \mathrm{d}^4x J(x)\Phi_{\text{cl}}(x)$. The effective action is the generating functional of one-particle irreducible Green functions. It is easy to see that

$$\frac{\delta\Gamma[\Phi_{\text{cl}}]}{\delta\Phi_{\text{cl}}(x)} = -J(x) \, , \tag{1.42}$$

and in the absence of an external source $J(x)$, Φ_{cl}, which is then the vacuum of the theory, is obtained by making the effective action stationary.

If our theory exhibits translational invariance, then Φ_{cl} will be a constant and the effective action can be written as $\Gamma[\Phi_{\text{cl}}] = -\int \mathrm{d}^4x V_{\text{eff}}(\Phi_{\text{cl}})$ where V_{eff} is called the effective potential. Stationarity of the action directly implies

$$\left. \frac{\partial V_{\text{eff}}}{\partial \Phi} \right|_{\Phi = \Phi_{\text{cl}}} = 0 \, , \tag{1.43}$$

and we can determine the vacuum expectation value of our field Φ in the absence of external sources by minimising V_{eff}. Effective potentials are usually obtained perturbatively in the coupling constant, and the result can be written as a power series in this constant:

$$V_{\text{eff}}(\Phi) = \sum_{n=0}^{\infty} V_n(\Phi) \, . \tag{1.44}$$

In this expansion, $V_0 = V$ represents the tree-level, i.e., the classical potential, and V_n corresponds to the contribution to the potential that is obtained from Feynman diagrams containing n closed loops. Note that with V_0 being the classical potential, our previous statement that VEVs are minima of the effective potential is justified, as it is not yet implied by Eq. (1.43).

The formalism of effective potentials was originally largely developed in Refs. [108, 124, 125], with the original explicit computation of the one-loop effective potential carried out by COLEMAN and E. WEINBERG in Ref. [126]. Given a Lagrangian containing scalars, gauge fields, and fermions as obtained from combining Lagrangians given in Eqs. (1.25) and (1.30), the result in the $\overline{\text{MS}}$ scheme reads [120, 127]

$$V_1(\Phi_{\text{cl}}) = \frac{1}{64\pi^2}\left\{\text{tr}\left(\mu^4 \ln\left(\frac{\mu^2}{\sigma^2}\right)\right) + 3\text{tr}\left(M^4 \ln\left(\frac{M^2}{\sigma^2}\right)\right) - 4\text{tr}\left(m_F^4 \ln\left(\frac{m_F^2}{\sigma^2}\right)\right)\right\}. \tag{1.45}$$

The above mass matrices for $\mu(\Phi_{\text{cl}})$, $M(\Phi_{\text{cl}})$, and $m_F(\Phi_{\text{cl}})$ are those defined in Eqs. (1.21), (1.29), and (1.31) for scalars, vectors, and spinors respectively.[5] Furthermore, σ denotes the renormalisation scale, which is an arbitrary scale on which the effective potential should not depend. This requirement can be translated into renormalisation group equations, which ensure the independence of σ by introducing a running of the coupling constants and masses. While for many models, logarithmic corrections leave the qualitative behaviour of the theory unchanged and merely affect the scaling behaviour close to a phase transition, there are actually cases in which they can induce spontaneous symmetry breaking not present in the classical theory. This is known as the *Coleman–Weinberg mechanism* [126] and the standard example is massless scalar quantum electrodynamics with a quartic potential $V(\phi) = \lambda|\phi|^4$. From a classical point of view, this allows only a minimum of the classical potential at $\phi = 0$. Radiative corrections will, however, shift the global minimum of the effective potential to $|\phi| \neq 0$ and thereby induce spontaneous symmetry breaking. As a rough estimate, we can set λ, g and y to be the typical Higgs, gauge, and Yukawa couplings such that $\mu(\Phi_{\text{cl}}) \sim \sqrt{\lambda}\Phi_{\text{cl}}$, $M(\Phi_{\text{cl}}) \sim g\Phi_{\text{cl}}$ and $m_F(\Phi_{\text{cl}}) \sim y\Phi_{\text{cl}}$. Radiative corrections to the potential are expected to be small as long as [36]

$$y^4, g^4 \ll \lambda \ll 1\,. \tag{1.46}$$

For the following discussions, we will indeed assume that radiative corrections are small and we can actually work with the classical potential.

The above treatment works under the assumption of scattering events occurring in empty space and is therefore bound to fail in the early Universe, where scattering occurs in an environment of high densities and temperatures. The formalism based on effective potentials taking finite-temperature effects into account was developed by S. WEINBERG [131], BERNARD [132], DOLAN and JACKIW [133], as well as KIRZHNITS and LINDE [134]. To understand it, recall that for a system with Hamiltonian H, we can introduce in a thermal equilibrium with temperature T the canonical density operator $\rho(T) \propto \exp(-H/T)$ with the normalisation $\text{tr}(\rho) = 1$.[6] Expectation values of operators $\mathcal{O}$ can then be obtained as $\langle\mathcal{O}\rangle_T = \text{tr}(\mathcal{O}\rho(T))$ with the trace running over a complete basis of states. In the imaginary time formalism of finite-temperature field theory, our generating functional gets replaced by the temperature-dependent generating functional

$$Z_T[J] = \left\langle \mathcal{T}_E\left\{\exp\left(i\int_0^{1/T} \mathrm{d}\tau \int \mathrm{d}^3\boldsymbol{x}\, J(\tau,\boldsymbol{x})\Phi(\tau,\boldsymbol{x})\right)\right\}\right\rangle_T. \tag{1.47}$$

The integration over time got replaced by an integration over Euclidean time $\tau = it \in [0, 1/T]$ and $\mathcal{T}_E$ imposes ordering with respect to Euclidean time. From this point on, we can proceed in the same way as before and obtain $W_T[J]$ and $\Gamma_T[\Phi_{\text{cl}}]$ analogous to the zero-temperature case.

---

[5] Note that while V_{eff} is not generally gauge-invariant and can furthermore be complex, this should not affect the gauge-invariant vacuum. Nevertheless, it can affect the derived masses. For more recent discussions on this and related points, see Refs. [128–130].

[6] We assume here that all chemical potentials are negligibly small $\mu \ll T$ which is true under the assumption of small asymmetries in the early Universe. If chemical potentials are relevant, one rather needs to work with the grand-canonical density operator.

Assuming translational invariance of Φ_{cl}, we can also write $\Gamma_T[\Phi_{\text{cl}}] = -\int \mathrm{d}\tau \int \mathrm{d}^3\boldsymbol{x} V_{\text{eff}}^{(T)}(\Phi_{\text{cl}})$ and obtain

$$\left.\frac{\partial V_{\text{eff}}^{(T)}}{\partial \Phi}\right|_{\Phi=\Phi_{\text{cl}}} = 0\,. \tag{1.48}$$

The formal equivalence between the zero-temperature and the finite-temperature case allows us to compute the effective potential perturbatively by replacing zero-temperature with finite-temperature Feynman rules (for a comparison, see e.g. Ref. [121]). The tree-level result is, of course, again the classical potential. The one-loop contribution, neglecting the Coleman–Weinberg correction of Eq. (1.45), is of the general form

$$V_1(\Phi_{\text{cl}}, T) = \frac{T^4}{2\pi^2}\left[\sum_B n_B J_-\left(\frac{m_B^2(\Phi_{\text{cl}})}{T^2}\right) - \sum_F n_F J_+\left(\frac{m_F^2(\Phi_{\text{cl}})}{T^2}\right)\right]\,, \tag{1.49}$$

where

$$J_\mp\left(\frac{m^2}{T^2}\right) = \int_0^\infty \mathrm{d}k\, k^2 \ln\left(1 \mp \exp\left(-\frac{\sqrt{k^2+m^2}}{T}\right)\right)\,. \tag{1.50}$$

The sums run over all bosons B and fermions F with n_B and n_F denoting the number of bosonic and fermionic degrees of freedom, respectively. At temperatures sufficiently far above all relevant masses, the effective potential is then well-approximated by

$$V_{\text{eff}}^{(T)}(\Phi_{\text{cl}}) = V(\Phi_{\text{cl}}) + \frac{1}{24}\mathcal{M}^2(\Phi_{\text{cl}})T^2 - \frac{\pi^2}{90}\mathcal{N}T^4\,, \tag{1.51}$$

with

$$\mathcal{N} = \mathcal{N}_B + \frac{7}{8}\mathcal{N}_F\,, \qquad \mathcal{M}^2 = \text{tr}\left(\mu^2\right) + 3\text{tr}\left(M^2\right) + \frac{1}{2}\text{tr}\left(\gamma^0 m_F \gamma^0 m_F\right)\,, \tag{1.52}$$

where μ^2, M^2, and m_F are the scalar, vector, and spinor mass matrices defined in Eqs. (1.21), (1.29), and (1.31), $\mathcal{N}_B = \sum_B n_B$ is the total number of bosonic degrees of freedom, and $\mathcal{N}_F = \sum_F n_F$ is the total number of fermionic degrees of freedom.

The above high-temperature expansion of $V_{\text{eff}}^{(T)}$ provides an important insight. Recall from Sec. 1.1.1 that spontaneous symmetry breaking is usually achieved using a Higgs potential with negative square masses, which shift the minimum of the potential away from the symmetric phase. This is, e.g., the case in the model of electroweak interactions [135–138]. At finite temperatures, we get, however, additional contributions $\sim \mathcal{M}^2T^2$ to the mass term. These contributions can assume at high temperatures large positive values, making the effective square mass term positive and, thus, preventing spontaneous symmetry breaking.

For a more intuitive understanding, it is helpful to notice that in equilibrium the finite-temperature effective potential is nothing but the free energy density (for details see, e.g., Refs. [131, 139]). At low temperatures, the free energy $F = E - TS$ is determined by the energy E while the contribution due to entropy S is negligible. With rising temperatures, the entropic contribution becomes increasingly relevant. A minimisation of the free energy then amounts to a maximisation of the entropy. An increase in entropy can be achieved by an increase in the available phase space volume and thus by a decrease in the particle masses. As we have seen, particle masses grow with the expectation value of the Higgs field, so at high temperatures, there is a natural tendency to decrease the Higgs field expectation value.

The possibility of symmetry restoration at high temperatures was, in the context of particle physics, first recognised by KIRZHNITS [22].[7] The transition from the symmetry-restored regime with $\Phi_{\text{cl}} = 0$ to the regime with spontaneously broken symmetry $\Phi_{\text{cl}} \neq 0$ is referred to as a *cosmological phase transition* with the two regimes referred to as unbroken and broken phase respectively. The temperature at which the $V_{\text{eff}}^{(T)}$ for both phases is equal is referred to as the critical temperature T_c, such that the symmetry is broken for $T < T_c$ and restored for $T > T_c$. Already on dimensional grounds, we can estimate that the critical temperature will be of the same order as the VEV of the Higgs field $T_c \sim |\Phi_0|$. Generally, the values of the coupling constants will, however, also affect the critical temperature. To see this and understand the above general statements better, let us take a look at two specific examples.

■ **Example 1.3 — $\mathbb{Z}_2 \to \{\, \mathbb{1} \,\}$.** The simplest non-trivial case we can consider is a canonically normalised real scalar field ϕ with the potential

$$V(\phi) = \frac{\lambda}{4}\left(\phi^2 - \eta^2\right)^2 . \tag{1.53}$$

This theory is clearly invariant under $\mathbb{Z}_2$ transformations $\phi \mapsto \pm\phi$. Classically, there are two possible vacua, namely $\phi_0 = \pm\eta$, which are both not invariant under the symmetry, and the model exhibits spontaneous symmetry breaking $\mathbb{Z}_2 \to \{\, \mathbb{1} \,\}$ at zero temperature. Let us now see what happens if we include finite-temperature effects. To apply Eq. (1.51), note that we have $\mathcal{N}_B = 1$, $\mu^2 = \lambda\left(3\phi_{\text{cl}}^2 - \eta^2\right)$ and, trivially, $\mathcal{N}_F = 0$, $M^2 = 0$ and $m_F = 0$, so we find for the effective potential

$$V_{\text{eff}}^{(T)}(\phi_{\text{cl}}) = \frac{1}{2}\lambda\left(\frac{T^2}{4} - \eta^2\right)\phi_{\text{cl}}^2 + \frac{\lambda}{4}\phi_{\text{cl}}^4 - \Delta V(T)\,, \tag{1.54}$$

where we omitted terms independent of ϕ_{cl} and T and with $\Delta V(T) = \lambda\eta^2T^2/24 + \pi^2T^4/90$. The

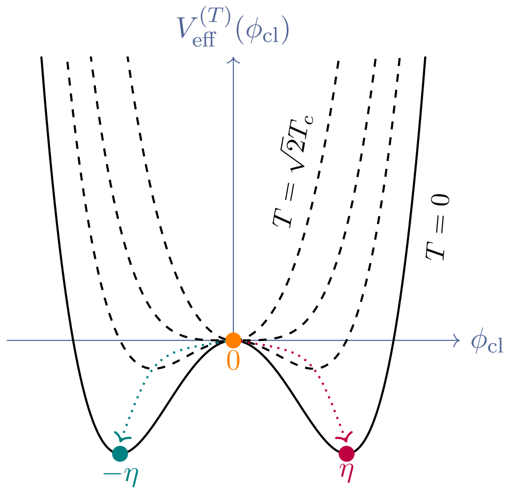


Figure 1.2: Plot of the finite-temperature effective potential given in Eq. (1.54) after subtraction of all field-independent terms $-\Delta V(T)$ for four different temperatures $T/T_c \in [0, 1/\sqrt{2}, 1, \sqrt{2}]$ ordered from outermost to innermost curve. The symmetric phase at $\phi_{\text{cl}} = 0$ is depicted by an orange dot, the degenerate minima of the zero-temperature potential at $\phi_{\text{cl}} = \pm\eta$ corresponding to the broken phase are depicted by purple and teal dots.

---

[7] The argument of symmetry-restoration does not hold generally true. See, e.g., Refs. [131, 140] or Ref. [141] for more recent work in the context of electroweak symmetry breaking. In Sec. 2.4, we will also see a simple example of symmetry non-restoration in the context of melting defects.

minima of $V_{\text{eff}}^{(T)}$ can be computed as

$$\phi_{\text{cl}} = \begin{cases} 0 & \text{if } T > T_{\text{c}} \\ \pm\sqrt{\eta^2 - T^2/4} & \text{else} \end{cases}, \qquad \text{with} \qquad T_c = 2\eta\,. \tag{1.55}$$

We note that the qualitative behaviour of the potential changes at T_c. For $T > T_c$, there is only one minimum corresponding to the symmetric phase. For $T < T_c$, the symmetric phase becomes a local maximum, and two degenerate minima both associated with the broken phase arise, signalling a phase transition at the critical temperature T_c. This is also clearly visible in Fig. 1.2, where we plot the potential for different temperatures. As we will see later, the above model leads to the formation of stable domain walls. ■

■ **Example 1.4 — $U(1) \to \{\,\mathbb{1}\,\}$.** Another model that we will investigate in more detail later on is the Abelian Higgs model given in Eq. (1.7). To apply the above treatment, we decompose our complex scalar field into two canonically normalised real scalar fields $\phi_{1,2}$ such that $\phi_{\text{cl}} = (\phi_1 + i\phi_2)/\sqrt{2}$. For the non-vanishing mass matrices, we obtain

$$\mu_{ij}^2 = \frac{\lambda}{2}\left(|\phi_{\text{cl}}|^2 - \eta^2\right)\delta_{ij} + \frac{\lambda}{2}\phi_i\phi_j \qquad \text{and} \qquad M^2 = 2e^2|\phi_{\text{cl}}|^2\,, \tag{1.56}$$

and, moreover, $\mathcal{N} = \mathcal{N}_B = 4$. The finite-temperature effective potential reads then according to Eq. (1.51)

$$V_{\text{eff}}^{(T)}(\phi_{\text{cl}}) = V(\phi_{\text{cl}}) + \frac{\lambda + 3e^2}{12}|\phi_{\text{cl}}|^2 T^2 - \Delta V(T)\,, \tag{1.57}$$

with $\Delta V(T) = \lambda\eta^2 T^2/24 + 2\pi^2 T^4/45$. The model exhibits a critical temperature

$$T_c = \sqrt{\frac{6\lambda}{\lambda + 3e^2}}\,\eta\,, \tag{1.58}$$

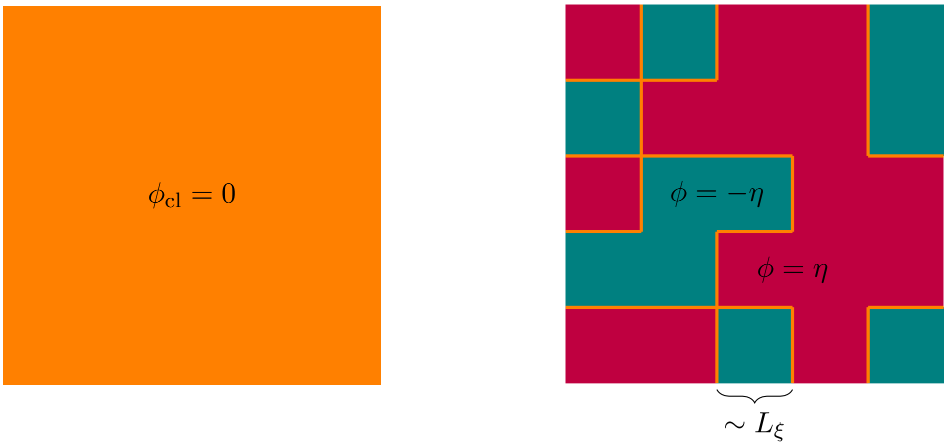
such that $V_{\text{eff}}^{(T)}$ has for $T > T_c$ one minimum at $\phi_{\text{cl}} = 0$, and the symmetry is restored, while for $T < T_c$ the effective potential has a circle-worth of degenerate minima at

$$|\phi_{\text{cl}}| = \sqrt{1 - \frac{T^2}{T_c^2}}\,\eta\,, \tag{1.59}$$

and the symmetry is spontaneously broken. In this example, we can see that the critical temperature not only depends on the Higgs VEV, but also on its coupling constants. The results for T_c and the minima of the potential can be carried over to the case in which the symmetry is not gauged by setting $e = 0$. As we will see soon, the above phase transition leads both in the global and the gauged case to the formation of cosmic strings. In the former case, these are referred to as global strings; in the latter case, they are referred to as local strings. Local strings will be a topic of great interest for this thesis, particularly in Pt. II. ■

1.3 From phase transitions to defect formation

The above discussion indicates that for systems with spontaneously broken symmetry at low temperatures, the symmetry is restored for $T > T_c$. Since in an expanding Universe like ours, the temperature decreases with time t, we conclude that we start from the symmetric phase at early times and spontaneous symmetry breaking occurs only later. While effective potential techniques are useful to indicate high-temperature symmetry restoration and to compute the temperature scale at which a phase transition occurs, one must keep in mind that these computations are based on the



(a) $T > T_c$ **(b)** $T = 0$

Figure 1.3: Illustration of the distribution of order parameter values, i.e., Higgs field expectation values, for the $\mathbb{Z}_2$-symmetry breaking of Ex. 1.3 in 2 spatial dimensions on a 5×5 square lattice (compare also to Fig. 1.2). **Left:** The expectation value of the Higgs field is above the critical temperature everywhere $\phi_{\rm cl} = 0$ and depicted in orange. **Right:** Distribution of Higgs expectation values at $T = 0$. We chose the side length of each cell to be L_ξ such that we can populate the cells with equal probability with either $\phi_{\rm cl} = -\eta$ (teal) or $\phi_{\rm cl} = \eta$ (purple). Along the line dividing regions of different vacua, continuity imposes $\phi_{\rm cl} = 0$, which is shown in orange.

assumption of a homogeneous expectation value of the order parameter $\Phi_{\rm cl}$ (see discussion above Eq. (1.48)) and on the assumption of thermal equilibrium. During a phase transition, the system departs from thermal equilibrium, and the transition also proceeds in an inhomogeneous manner. Effective potentials are hence incapable of determining the dynamics of phase transitions.

In cosmology, one usually distinguishes roughly between first-order and continuous phase transitions.[8] In a continuous phase transition, the value of the order parameter evolves continuously from the symmetric to the broken phase, a process known as spinodal decomposition. This is the case for both the examples 1.3 and 1.4 we saw in the previous section. The opposite is true for a first-order phase transition, where the order parameter changes discontinuously. The reason for this difference can be found in the shape of the effective potential below T_c. For a first-order phase transition, the symmetric phase corresponds to a local, but not global, minimum of the potential and is often referred to as the false vacuum. The broken phase is associated with one of possibly many degenerate global minima, referred to as true vacua. This means that false and true vacua are separated by a potential barrier, and the transition must happen via thermal fluctuations or quantum tunnelling. In physical space, the phase transition occurs then at a temperature below the critical temperature by bubbles of true vacuum nucleating within the false vacuum and afterwards expanding until they percolate. For continuous phase transitions, the symmetric phase corresponds at low temperatures rather to a local maximum of the potential, and there is no separating barrier.

While we do not need to know all details about the phase transition to make statements about cosmic defects, a very important concept is the so-called correlation length L_ξ. We saw that, in general, in a phase transition, the Higgs expectation value changes from $\Phi_{\rm cl} = 0$ to another non-vanishing value in the vacuum manifold $\Phi_{\rm cl} \in \mathcal{M}$. Due to the symmetry underlying the theory, all points in $\mathcal{M}$ are degenerate and, thus, chosen with the same probability. Field values will only be correlated on a certain characteristic length scale, and this is just what we refer to as the correlation length. We illustrated this in Fig. 1.3 by means of the symmetry breaking of Ex. 1.3 in two spatial dimensions. Choosing L_ξ as the length of each lattice site, we can populate the lattice by drawing for each lattice cell the value $-\eta$ or η with probability $1/2$.

---

[8] For a more detailed discussion, see, e.g., Refs. [120, 142] and references therein.

For first-order phase transitions, one can roughly estimate the correlation length at the end of the phase transition by noting that the transition terminates when all space is filled with true vacuum, i.e., when all bubbles have grown until they collide. L_ξ is then just the typical bubble size at percolation [143]. For a bubble nucleation rate Γ per unit volume and a bubble wall velocity of v, we can estimate [144]

$$L_\xi \sim \left(\frac{v}{\Gamma}\right)^{1/4} . \tag{1.60}$$

Note that even though the above expression might seem simple, computing v and Γ can be very hard on their own.

For continuous phase transitions, we can distinguish between two cases: crossovers and second-order phase transitions. In the SM, the electroweak and the QCD phase transition belong to the former class (see Refs. [145] and [146]). L_ξ is just the thermal correlation length at a cross-over in spontaneously broken local symmetries [147–149]. For a second-order phase transition, obtaining a precise estimate of L_ξ is much harder. In 1976, KIBBLE gave a rough estimate of L_ξ based on causality in Ref. [32] — the same article that introduced cosmic defects for the first time in a systematic way.[9] If a phase transition starts at a conformal time τ_i and terminates at time τ_f, then the maximum distance that information about the field can propagate is imposed by causality to the value $\tau_f - \tau_i$. While we do not know these times exactly, this distance is clearly smaller than the particle-horizon distance d_h, such that we obtain the upper bound[10]

$$L_\xi(t) < d_h(t) = a(t) \int_0^t \frac{dt'}{a(t')} . \tag{1.61}$$

The phase transitions we are interested in usually occur deep in the radiation-dominated era, and the above expression reduces to

$$L_\xi(t) \lesssim 2t . \tag{1.62}$$

It was later realised by ZUREK that similar arguments could be applied to condensed matter systems [151–153]. The key idea was roughly to use the correlation length under the assumption of thermal equilibrium $\bar{L}_\xi$ before the phase transition and connect it via a freeze-out during the phase transition to L_ξ. Let us briefly summarise the argument. Close to a second-order phase transition, the equilibrium correlation length diverges algebraically:

$$\bar{L}_\xi = L_0 |\epsilon|^{-\nu} , \qquad \text{with} \qquad \epsilon = \frac{T - T_c}{T_c} . \tag{1.63}$$

Here, L_0 is a length determined by the microphysics, and ν is known as the correlation length critical exponent. For times very close to T_c, we can assume that the temperature changes linearly with time:

$$T = T_c \left(1 - \frac{t}{\tau_Q}\right) , \tag{1.64}$$

where τ_Q is known as the quench time, which determines the rate at which the system is cooled and thus at which rate T is lowered. Another important quantity is the equilibrium relaxation

---

[9] The causality argument goes back to Ref. [31]. Ref. [32] contains another estimate of L_ξ based on the Ginzburg criterion. Its relevance for defect formation is, however, unclear, and we will not review it here (see also [150]).

[10] Note that d_h will become vast if there was an early inflationary phase, rendering this estimate meaningless. Nevertheless, for a phase transition occurring after the end of inflation, $\tau_f - \tau_i$ will still be smaller than $\tilde{d}_h$, the would-be-horizon distance, if there was no inflationary phase. Since we do not consider the effect of inflation on defects in this thesis, we will always implicitly work with the latter scale when talking about particle horizons.

time τ_r, the characteristic time over which equilibrium fluctuations decay. At T_c, it also diverges algebraically:

$$\tau_r = \tau_0 |\epsilon|^{-z\nu} , \tag{1.65}$$

where z is called the dynamical critical exponent and τ_0 is a time scale determined by microphysics. At temperatures high above the critical temperature, τ_r is very small compared to the time remaining until reaching the critical temperature, and the system evolves adiabatically. Closer to T_c, the equilibrium relaxation time grows arbitrarily large, and the system eventually drops out of thermal equilibrium and essentially freezes. This roughly happens when the time remaining until the critical point becomes equal to the relaxation time, i.e., at

$$t_{\text{freeze}} = \left(\tau_0 \tau_Q^{z\nu}\right)^{\frac{1}{1+z\nu}} . \tag{1.66}$$

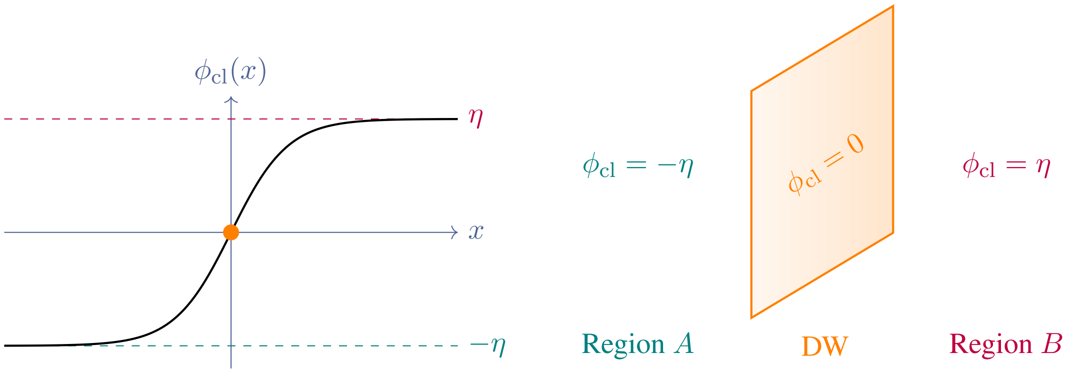
The resulting estimate for a characteristic size of a domain is then given as

$$L_\xi = \bar{L}_\xi(t_{\text{freeze}}) = L_0 \left(\frac{\tau_Q}{\tau_0}\right)^{\frac{\nu}{1+z\nu}} . \tag{1.67}$$

Note that, while L_0 and τ_0 are determined by the microphysics of the model, the critical exponents ν and z are completely determined by the symmetry properties of the theory, i.e., its underlying universality class (for a review, see Ref. [154]).

The above discussion is known as the *Kibble–Zurek mechanism* and its predictions have been confirmed by many experiments in condensed-matter systems (for a review, see Ref. [155]). Finally, note that the estimate of L_ξ in Eq. (1.61) is most often used as it is not only simple to compute but also valid independently of whether the phase transition is first-order, second-order or a smooth crossover.

Let us now see how the above discussion of L_ξ relates to the formation of defects in a phase transition — a connection made by KIBBLE in 1976 [32]. To this end, consider again the simple $\mathbb{Z}_2$-symmetry breaking of Ex. 1.3. Let A and B be two spatial regions such that in region A, $\phi_{\text{cl}}|_A = -\eta$ and in region B we have $\phi_{\text{cl}}|_B = \eta$. We are interested in what happens between these two regions. From our two-dimensional illustration in Fig. 1.3b, A and B would correspond to a teal and a purple region. Our interest lies in the neighbourhood of the line which separates these two regions. For simplicity, we can even restrict to one spatial dimension for which we can choose a coordinate x such that A is reached for $x \ll 0$ and B for $x \gg 0$. Our field must be continuous as finite changes in ϕ over arbitrarily small distances would lead to arbitrarily large gradient contributions to the field's energy $\sim (\partial_x \phi_{\text{cl}})^2$. This means that between $x < 0$ and $x > 0$, the field must at least once take every value between $-\eta$ and η, including $\phi_{\text{cl}} = 0$. The simplest configuration we can imagine is one for which every value is only taken once and $\phi_{\text{cl}}(x = 0) = 0$, as depicted in the left panel of Fig. 1.4a. The key insight is the following: Due to continuity, we are forced to have $\phi_{\text{cl}} = 0$ somewhere between regions A and B, but this corresponds to the unbroken phase and, as is easy to see from Fig. 1.2, to the local maximum of the zero-temperature potential. Therefore, as a result of the different spatial regions choosing different VEVs, the phase transition cannot happen everywhere, and similarly, the field cannot assume its vacuum everywhere. We have obtained a non-vacuum configuration, called a cosmic defect. The particular defect we encountered is known as a ϕ^4 *kink* and we can realise it in different spatial dimensions. In d spatial dimensions, the equation $\phi_{\text{cl}}(\boldsymbol{x} = 0) = 0$ implicitly describes a hypersurface of codimension 1. In two dimensions, this corresponds to a one-dimensional line, and we depict it by orange lines separating different vacua in Fig. 1.3b. In three dimensions, the condition describes a two-dimensional surface separating the vacua as illustrated in Fig. 1.4b. Since we live in a Universe

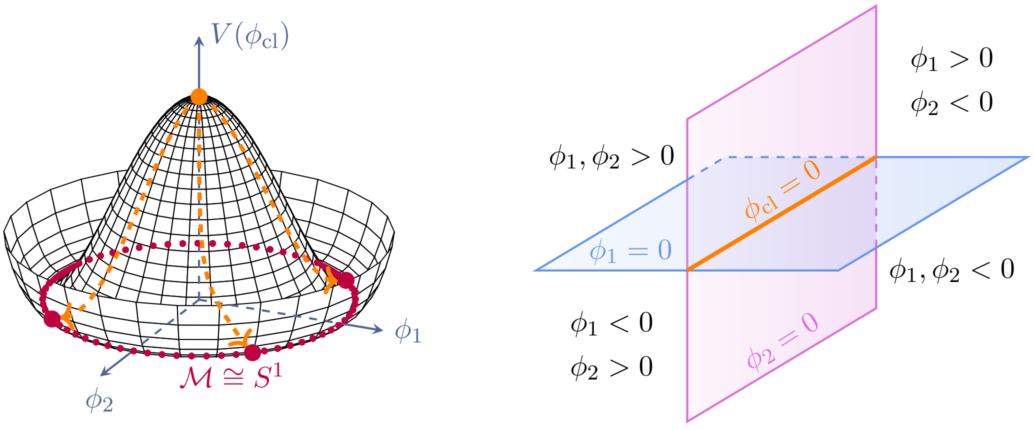


(a) **(b)**

Figure 1.4: Left: Plot of a field configuration corresponding to a single $\mathbb{Z}_2$ domain wall arising from the model studied in Ex. 1.3 in one spatial dimension. The plotted function is explicitly given in Eq. (2.3). The point at which $\phi_{\rm cl} = 0$ is marked by an orange dot corresponding to the symmetric phase and the local maximum of the zero-temperature potential of Eq. (1.53) depicted in Fig. 1.2. **Right:** Illustration of the domain wall on which $\phi_{\rm cl} = 0$ as a separation of regions A and B with $\phi_{\rm cl}|_A < 0$ and $\phi_{\rm cl}|_B > 0$.

with three spatial dimensions, defect configurations of codimension 1 are accordingly referred to as *domain walls*.

Other defects form in essentially the same way. Consider the Abelian Higgs model discussed in Ex. 1.4. As before, for regions separated by a length L_ξ, the phase transition from $\phi_{\rm cl} = 0$ to $\phi_{\rm cl} \in \mathcal{M} \cong S^1$ can occur independently, as depicted in Fig. 1.5a. Consider that in a region A both real and imaginary part of the field $\phi_{\rm cl} = (\phi_1 + i\phi_2)/\sqrt{2}$ are smaller than 0 while in region B they are larger: $\phi_{1,2}|_A < 0$, $\phi_{1,2}|_B > 0$. The continuity of ϕ_1 and ϕ_2 implies that between region A and region B we must have somewhere $\phi_1(\boldsymbol{x}) = 0$ and $\phi_2(\boldsymbol{x}) = 0$. This defines two hypersurfaces of codimension 1 which generically intersect (assuming $d > 1$) in a codimension-2 hypersurface. In three dimensions, this corresponds to two 2-surfaces intersecting in a line as depicted in Fig. 1.5b. Notice that on each of the two surfaces, the field can assume values in the vacuum. For example,



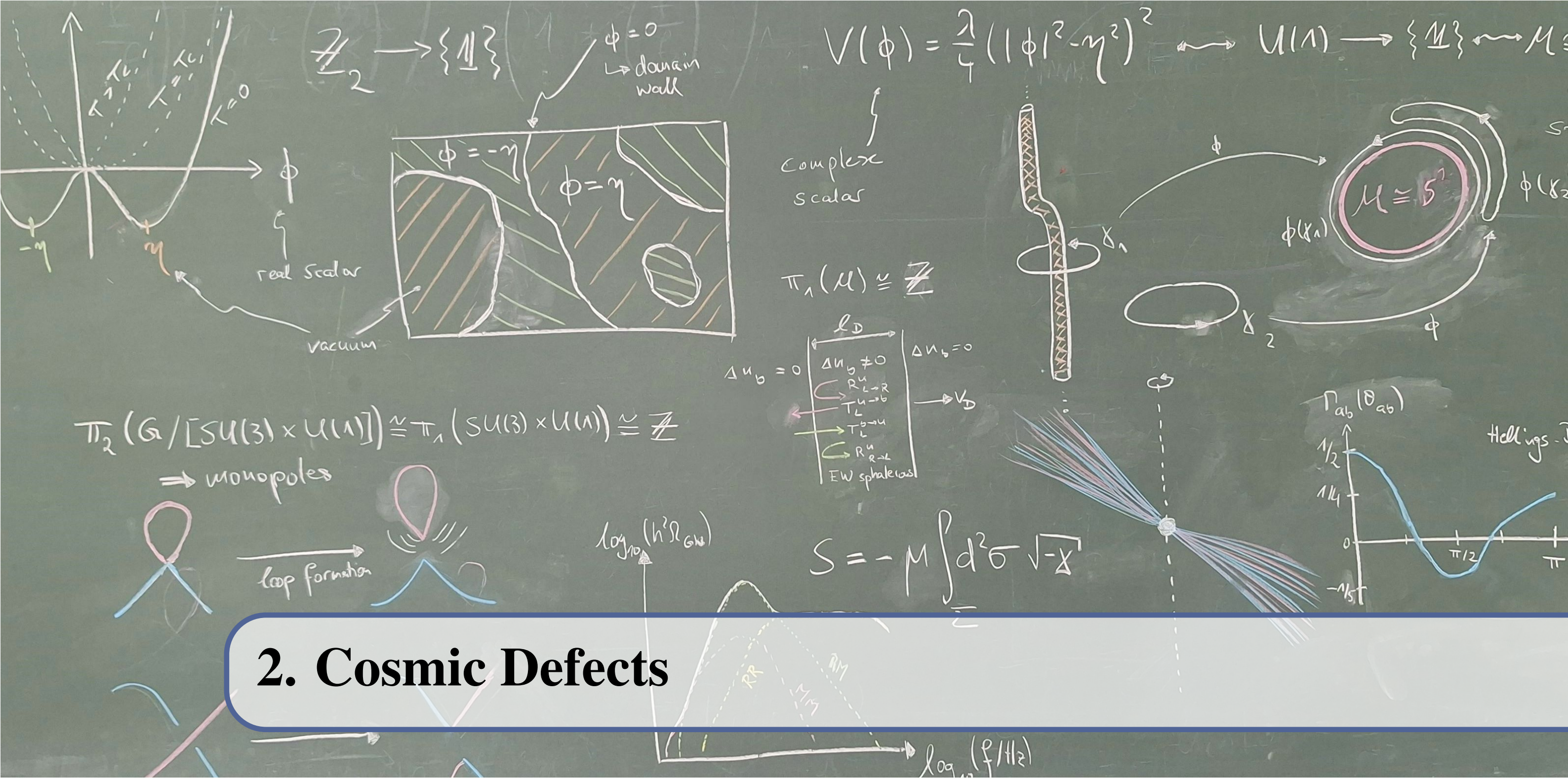
(a) **(b)**

Figure 1.5: Illustration of the formation of cosmic strings. **Left:** Zero-temperature potential of the Abelian Higgs model (Ex. 1.4). Starting from the symmetric phase $\phi = 0$ marked by an orange dot at high temperatures, the field will assume an expectation value in the vacuum manifold $\mathcal{M}$, the red-dotted circle at the bottom of the potential. In different uncorrelated regions, the Higgs expectation value will randomly be chosen in $\mathcal{M}$ and generally be different, as illustrated here for three uncorrelated points in space. **Right:** Uncorrelated regions A with $\phi_{1,2} > 0$ in the upper left and B with $\phi_{1,2} < 0$ in the lower right will be separated by two surfaces on which $\phi_1 = 0$ (blue) and $\phi_2 = 0$ (orchid). The surfaces intersect in a line, the cosmic string, along which $\phi = 0$, corresponding to the local maximum of the potential shown in the left panel.

on the $\phi_1 = 0$-surface, this can be realised by $\phi_2 = \pm\sqrt{2}\eta$ such that $|\phi_{\rm cl}| = \eta$, which corresponds to the bottom of the potential. Along the line that is the intersection of the two surfaces, the field is, however, forced to assume the value $\phi_1 = \phi_2 = 0$ and thus $\phi_{\rm cl} = 0$. This means that along the line, the field has to remain in the symmetric phase, the phase transition cannot complete, and the vacuum cannot be reached; we obtained another cosmic defect. Because the symmetry remains unbroken along a line, this defect is known as a *(cosmic) string*.

Finally, consider Ex. 1.2 in which an $SU(2)$ symmetry is spontaneously broken to a $U(1)$ symmetry with a corresponding vacuum manifold S^2. This was achieved by giving a non-vanishing VEV to a Higgs $SU(2)$-triplet $\Phi_{\rm cl}$ that can be written in terms of three real scalars $\phi_{1,2,3}$. As before, we can consider two regions A and B such that $\phi_{1,2,3}|_A < 0$ and $\phi_{1,2,3}|_B > 0$. The continuity of the fields implies the existence of three codimension-1 hypersurfaces $\phi_1(x) = 0$, $\phi_2(x) = 0$, and $\phi_3(x) = 0$ that generically intersect in a codimension-3 hypersurface. In three dimensions, this intersection is 0-dimensional, i.e., a point. In a point-like region of space, the unbroken phase $\Phi_{\rm cl} = 0$ remains and, since this corresponds to a local maximum of the zero-temperature potential, we find a non-vanishing energy density there. These defects are referred to as *monopoles* for reasons that will become clear in Sec. 2.2.4.

These simple examples already clarify the relevance of the correlation length L_ξ. It is the typical length over which the field values become uncorrelated, and different vacua can be chosen in the phase transition. Since defects form at the boundary of regions with different vacua, L_ξ is also the typical scale that separates cosmic defects. It is important to note that L_ξ continues to change dynamically after the end of the phase transition, typically growing over time as separated vacua can merge and erase the defects in between. The evolution of L_ξ will generally depend not only on the underlying microphysical model but also on defect interactions, interactions with the surrounding early-Universe plasma, and the expansion of the Universe.



2. Cosmic Defects

In the previous chapter, we saw that cosmological phase transitions can give rise to certain regions of space, where the phase transition cannot complete and the symmetry remains unbroken. These so-called defects are ubiquitous in extensions of the SM. In this chapter, we aim to provide a systematic overview of them. In Sec. 2.1, we begin with the simplest possible defect: a domain wall in one spatial dimension. This toy model serves as a foundation, introducing many concepts that generalise to more complex defect configurations. We then discuss Derrick's theorem, which constrains the existence of static, finite-energy defect solutions. In Sec. 2.2, we will then show how these basic ideas of defects are related to topological considerations and discuss the classification of stable defects in terms of homotopy groups. A classification in terms of topological considerations is, however, incomplete, and we will also study the field theoretical aspects of stable defects. Building on these ideas, we can consider the more complicated case of a series of phase transitions in which metastable defects can arise. We discuss their classification in Sec. 2.3. As an original contribution, we present an ansatz for field configurations describing strings with attached monopoles, where the monopoles carry unconfined flux. While stable and metastable defects are the main classes of topological defects, they contain important subclasses. Moreover, there are other important defect configurations beyond these two main categories. In Sec. 2.4, we will briefly discuss composite defects, textures, semilocal and embedded defects, superconducting strings, melting defects and time-dependent defects, each illustrated with a toy model and accompanied by a brief overview.

Since we will almost exclusively be concerned with classical background fields, we will henceforth write Φ instead of Φ_{cl}.

2.1 Some basics

Let us begin by studying the ϕ^4 kinks we have encountered at the end of Sec. 1.3 in more detail. These arise in one spatial dimension $d = 1$ from the simplest non-trivial symmetry breaking $\mathbb{Z}_2 \to \{\,\mathbb{1}\,\}$, which is discussed in Ex. 1.3. The Lagrangian of this model reads

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{\lambda}{4}\left(\phi^2 - \eta^2\right)^2 \,. \tag{2.1}$$

The equation of motion derived from this Lagrangian is

$$\left(\partial_t^2 - \partial_x^2 + V'(\phi)\right)\phi = 0 \tag{2.2}$$

and it has a non-trivial static solution of the form [156]

$$\phi(x) = \eta \tanh\left(\sqrt{\frac{\lambda}{2}}\eta x\right) . \tag{2.3}$$

This configuration, called a ϕ^4 kink, converges in the two asymptotic limits towards different vacua $\lim_{x\to\pm\infty}\phi(x) = \pm\eta$. Thus, in both limits, it minimises the energy density. This is exactly what we have plotted as a solid black line in Fig. 2.1a (and previously in Fig. 1.4a).

We can easily compute the energy density of this solution

$$\epsilon(x) = \frac{1}{2}\left(\partial_x\phi(x)\right)^2 + V(x) = \frac{\lambda\eta^4}{2}\text{sech}^4\left(\sqrt{\frac{\lambda}{2}}\eta x\right) , \tag{2.4}$$

and see that it has a profile of width $\delta = \sqrt{2/\lambda}\eta^{-1}$. This energy density peaks at $x = 0$, i.e., the location of unbroken symmetry, and falls off symmetrically. Hence, one can meaningfully regard the defect as being localised at $x = 0$. The function $\epsilon(x)$ is plotted as a solid black line in Fig. 2.1b. Integrating it, we find the finite energy of the defect

$$\sigma = \int_{-\infty}^{\infty} \mathrm{d}x\,\epsilon(x) = \frac{2\sqrt{2\lambda}\eta^3}{3} . \tag{2.5}$$

The ϕ^4 kink is undeniably not a trivial vacuum solution, and we would naturally expect it to decay over time into the vacuum. All the same, the ϕ^4 kink is static and does not decay.

To understand the underlying reason, let us examine how one could remove the ϕ^4 kink. The theory has two degenerate vacua $\phi = \pm\eta$. We have already seen that due to the continuity, as long as in one region of space $\phi \to -\eta$ and in another region $\phi \to \eta$, there must be a point at which $\phi = 0$, and this leads to the non-vanishing energy of the configuration. To remove the defect, it is hence necessary to bring ϕ to the same vacuum value, either $-\eta$ or η, at every point in space. We can, for example, try to reach the configuration $\phi(x) = \eta$. In order to do so, we need to lift all points $x < 0$ over the potential barrier, which has for $x \ll 0$ the height $\Delta V = V(0) - V(-\eta) = \lambda\eta^4/4$. If we lift the field only for a part of space over the potential barrier, as illustrated in Fig. 2.1a by means of grey dotted arrows, the position at which one crosses $\phi = 0$ just shifts. At the same time, the region with non-vanishing energy density shifts, as is visible in Fig. 2.1b, but the total energy of the configuration stays the same (or grows, but in this case, the resulting configuration will not be static). On the other hand, lifting the field everywhere in space over the potential barrier would erase the defect, but takes an energy $\sim \int_{-\infty}^{0} \Delta V$ which is infinite. Because of that, the defect is stable and will not decay.

The stability of the ϕ^4 kink can also be understood in terms of an associated conserved charge. For this, note that the current [157]

$$j^\mu = \frac{1}{2\eta}\epsilon^{\mu\nu}\partial_\nu\phi \tag{2.6}$$

is conserved $\partial_\mu j^\mu = 0$, with $\epsilon^{\mu\nu}$ the totally antisymmetric symbol and $\epsilon^{01} = 1$. Linked with the current is the charge

$$\mathcal{Q}_S = \int_{-\infty}^{\infty} \mathrm{d}x\, j^0 = \lim_{x\to\infty} \frac{\phi(t,x) - \phi(t,-x)}{2\eta} . \tag{2.7}$$

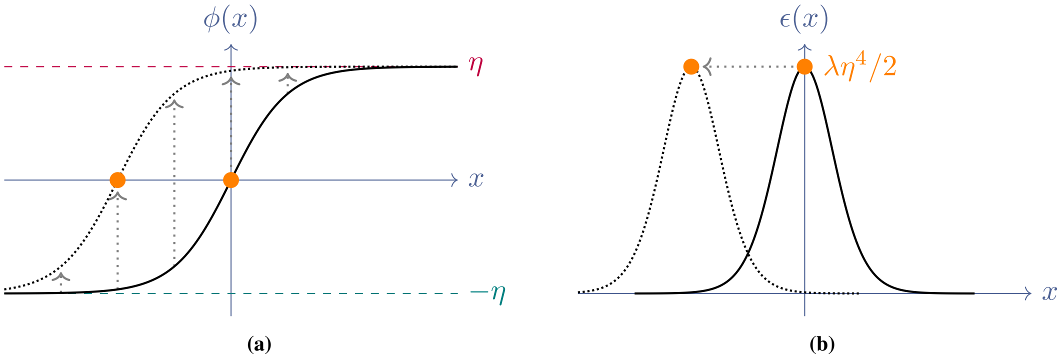


Figure 2.1: Left: Plot of the ϕ^4 kink field configuration given in Eq. (2.3) and shown as a solid black line. The configuration corresponds to a single $\mathbb{Z}_2$ domain wall in one spatial dimension located at $x = 0$, where $\phi = 0$ and the symmetry remains unbroken. The unbroken phase is indicated by an orange dot. The field ϕ asymptotically approaches the two degenerate vacua $\pm\eta$ highlighted by teal and purple dashed lines, respectively. Lifting the field over the potential barrier from one potential minimum $-\eta$ to the other η (illustrated by grey dotted arrows) leads to a topologically equivalent configuration (black dotted line), which also represents a domain wall, but is shifted to smaller values of x. The centre of the wall is again indicated by an orange dot. **Right:** Energy density given in Eq. (2.4) of the original ϕ^4 kink (black solid line) and of the shifted configuration (black dotted line). The orange dots mark again the location of the unbroken phase, which, visibly, corresponds to the maximum of the wall's energy density.

Any configuration ϕ should asymptotically approach a minimum of the potential. Transitioning between the two minima of the potential is impossible without leaving the vacuum manifold, and we must have $\partial_t \lim_{x\to\pm\infty} \phi(t,x) = 0$ such that the charge is conserved. For the ϕ^4 kink configuration of Eq. (2.3), we find $\mathcal{Q}_S = 1$, whereas a full vacuum configuration has $\mathcal{Q}_S = 0$. Charge conservation makes the transition impossible. As a consequence of the fact that there are only two minima of the potential, there is only one other possible value of the conserved charge, namely $\mathcal{Q}_S = -1$. Such a configuration can easily be obtained by multiplying the one in Eq. 2.3 by -1. It is usually referred to as an *antikink*. It is important that the charge $\mathcal{Q}_S$ is not a conserved charge in the sense of Noether's theorem: It does not derive from an underlying symmetry, nor does it generate one. The conservation is rather related to the fact that the vacuum manifold $\mathcal{M}$ is disconnected, and one cannot continuously transition from one minimum of the potential to the other without leaving $\mathcal{M}$. The conservation law follows therefore from the topology of $\mathcal{M}$, and $\mathcal{Q}_S$ is referred to as a topological charge.

The above discussion could tempt us to characterise cosmic defects as (topologically) stable, localised, static, finite-energy configurations. Unfortunately, this rules out most of the interesting cases. A useful theorem to realise this was proven by DERRICK in 1964 [158]. To understand it, consider a general model of n_Φ real scalar fields in d spatial dimensions with a potential $V(\Phi)$, as described by the Lagrangian in Eq. (1.10). Let us assume $\Phi_1(\boldsymbol{x})$ describes a stable, static, finite-energy solution to the equations of motion. The energy of any such field configuration is

$$E[\Phi_1] = I_1[\Phi_1] + I_2[\Phi_1]\,, \tag{2.8}$$

with

$$I_1[\Phi_1] = \int \mathrm{d}^d x\, \frac{1}{2}\,(\nabla\Phi_1(\boldsymbol{x}))^2 \qquad \text{and} \qquad I_2[\Phi_1] = \int \mathrm{d}^d x\, V(\Phi_1(\boldsymbol{x}))\,. \tag{2.9}$$

We can now examine the effects of a rescaling of Φ_1 and define $\Phi_\alpha(\boldsymbol{x}) = \Phi_1(\alpha\boldsymbol{x})$ for any $\alpha > 0$. Then we find for the energy

$$E_\alpha \equiv E[\Phi_\alpha] = \alpha^{2-d} I_1[\Phi_1] + \alpha^{-d} I_2[\Phi_1]\,, \tag{2.10}$$

and it follows that

$$\left.\frac{\mathrm{d}E_\alpha}{\mathrm{d}\alpha}\right|_{\alpha=1} = (2-d)\, I_1[\Phi_1] - dI_2[\Phi_1]\,. \tag{2.11}$$

Since the action becomes stationary for any solution to the equations of motion, we have $\delta S[\Phi_1] = 0$ and, due to the time-independence of the considered solution, this means $\delta E[\Phi_1] = 0$, which implies $(2-d)I_1[\Phi_1] = dI_2[\Phi_1]$. Upon using this relation, we can find that

$$\left.\frac{\mathrm{d}^2E_\alpha}{\mathrm{d}\alpha^2}\right|_{\alpha=1} = 2(2-d)I_1[\Phi_1]\,. \tag{2.12}$$

For the solution to be stable, it is required that $\delta^2 E[\Phi_1] > 0$. Since for any non-trivial configuration the inequality $I_1[\Phi_1] > 0$ holds, and the requirement of stability imposes via Eq. (2.12) that $d < 2$. We can summarise this result, known as *Derrick's theorem*, as follows: There exists no non-trivial, stable, static, finite-energy solution to the scalar field theory described by the Lagrangian of Eq. (1.10) in $d > 1$.

Although this theorem appears to place stringent constraints on the possibilities of allowed field configurations, most of the defects of interest do not satisfy the restrictive assumptions required for the theorem to apply.

This is easiest seen by returning to the ϕ^4 kink but now in three spatial dimensions $d = 3$, i.e., a domain wall. While $\phi(\boldsymbol{x})$ will depend on three spatial coordinates, we can choose it to be translationally invariant in the y and z directions. Under these conditions, the equations of motion reduce to the same form as for $d = 1$, which means that $\Phi(\boldsymbol{x}) = \eta \tanh\left(\sqrt{\lambda/2}\eta x\right)$ is a solution. We have clearly found a non-trivial, static solution to a scalar field theory that can be described by Eq. (1.10) in $d > 1$. This solution is very interesting since it can, at least locally, describe domain walls produced in cosmological phase transitions. Derrick's theorem is, however, not applicable, because this solution is not a finite-energy configuration: The form of the energy density, given in Eq. (2.4), remains unchanged compared to the one-dimensional case. Similarly, σ in Eq. (2.5) is now the constant surface energy density of the wall, named *domain wall tension*, and integrating it over the entire wall $\mathbb{R}^2$ yields a divergent energy. Other interesting solutions with infinite energy in three dimensions are infinitely long strings.

Apart from infinite energy solutions, we can also extend the simple model of Eq. (1.10). This can, for example, be done via the inclusion of higher derivative terms [157]. Much more relevant to us, though, is the extension to gauge fields, i.e., we should consider the Lagrangian in Eq. (1.25). If we include gauge fields A_μ, we add to the total energy another term $I_3[A_\mu] = \int \mathrm{d}^d x\, \frac{1}{2}\mathrm{tr}\left\{F_{\mu\nu}F^{\mu\nu}\right\}$. Denoting the rescaled field by $A^\alpha_\mu(\boldsymbol{x}) = \alpha A^1_\mu(\alpha\boldsymbol{x})$ (the gauge field components must scale as the partial derivative does, as can be seen, e.g., by considering the gauge-covariant derivative), we find $I_3[A^\alpha_\mu] \to \alpha^{4-d} I_3[A^1_\mu]$. The term I_1 will now also depend on A_μ because we need to replace partial with gauge-covariant derivatives, but this will not affect its scaling with α. From stationarity of the energy, we find then that $I_2 d = (4-d)I_3 + (2-d)I_2$ such that the solution is stable if $2(2-d)I_1 + 4(4-d)I_3 > 0$. Previously, we had the same equation with $I_3 = 0$, which led to the problem that for $d > 1$ the inequality could not be satisfied since $I_1 > 0$. But now, for $d = 2$ and 3, we have another positive contribution $I_3 > 0$ and stable, static, finite-energy configurations are allowed. These include, e.g., the monopoles we will encounter in Sec. 2.2.4.

Finally, we can also avoid Derrick's theorem by allowing time-dependent solutions. This can be either in a non-dissipative way, as is the case for so-called Q-balls, which we will briefly encounter in Sec. 2.4, or in a dissipative way, e.g., for defects of finite extent such as cosmic string loops, which will be the main subject of Pt. II of this thesis.

In summary, Derrick's theorem can be avoided in three sensible ways:

- Model extensions
- Time-dependent solutions
- Non-finite energy solutions

In this section, we have already discussed many characteristics of cosmic defects just based on the $\mathbb{Z}_2$-wall. The most important point was, however, that the stabilisation of the ϕ^4 kink is a consequence of the topological structure of the vacuum manifold. We will make this statement more precise in the next section, which allows us to classify different kinds of defects.

2.2 Topologically stable defects

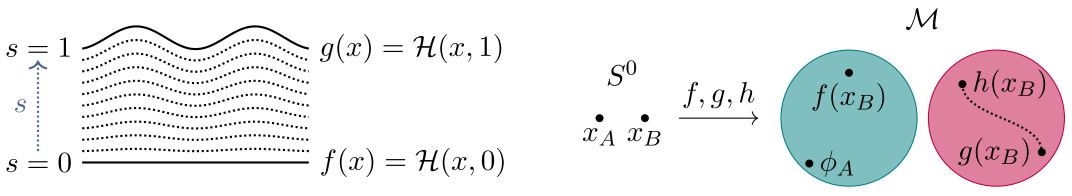
2.2.1 Domain walls and homotopy

We have seen in the previous section that the model of Ex. 1.3 giving rise to ϕ^4 kinks allows for solutions that cannot be continuously deformed into each other as they are separated by an infinite energy barrier. By contrast, a model without such a barrier would permit a continuous interpolation of ϕ from $-\eta$ to η while remaining entirely on the vacuum manifold $\mathcal{M}$. In other words, $\mathcal{M}$ would need to be path-connected.

Let us return to the more general setting in which we discuss what happens between two regions A and B. We could imagine a vacuum manifold $\mathcal{M}$ consisting of multiple path-connected components, for example, two components, as depicted in Fig. 2.2b. If the field takes in regions A and B a value from the same path-connected component, then there is no need for the field to leave $\mathcal{M}$ when transitioning from A to B. If the values in the same connected component are initially different, we would start out with a non-vacuum configuration due to the energy stored in the field's gradients. Nevertheless, there is no energy barrier to overcome, and one can continuously deform the field configuration such that the same vacuum field value is taken everywhere and the gradients vanish. On the other hand, if the field takes in regions A and B values corresponding to different path-connected components of $\mathcal{M}$, then the field must leave the vacuum when transitioning from A to B, and there is a domain wall. We can imagine yet another setup, where A and B have field values from the same path-connected component but are separated by a region C which has field values from a different component of $\mathcal{M}$. In this case, there will be two walls: One wall between A and C and another one between C and B. This time, we can, however, simply shrink region C continuously until it vanishes. In this way, we return to the first configuration for which there was no domain wall. The configuration with two walls is, therefore, not distinct from the trivial one, in the sense that they can continuously be deformed into each other.

We can now take the preceding intuitive picture and convert it into a formal statement. Consider two points $S^0 \cong \{ x_A, x_B \}$ as well as a map between this (pointed) 0-sphere and the (pointed) vacuum manifold $f : (S^0, x_A) \to (\mathcal{M}, \phi_A)$, such that $f(x_A) = \phi_A$. Considering all maps of this type, for some of them ϕ_A and $f(x_B)$ end up in the same connected component and for some of them they do not. If we imagine x_A and x_B to be two points in regions of physical space in which the vacuum is reached, e.g., regions A and B as before, each different map would describe a certain class of different field configurations. Each such configuration is fixed to the same vacuum value at x_B (the value at x_A is generally fixed anyway). This is, however, not yet what we are interested in to distinguish domain walls. If two field configurations can be continuously deformed into each other, they will give rise to an equivalent domain wall. To take this into account, we can identify maps f and g for which $f(x_B)$ and $g(x_B)$ are in the same connected component of $\mathcal{M}$. The set of all such maps after this identification has then for each different domain wall configuration (and the trivial vacuum configuration), one element.

Let us make the identification precise by using the notion of homotopy. Let X and Y be topological spaces and $f, g : X \to Y$ be continuous maps. Then we call a continuous map $\mathcal{H} : X \times [0, 1] \to Y$ such that $\forall x \in X : \mathcal{H}(x, 0) = f(x)$ and $\mathcal{H}(x, 1) = g(x)$ a homotopy between



(a) (b)

Figure 2.2: **Left**: Illustration of a homotopy $\mathcal{H} : [-1, 1] \times [0, 1] \to \mathbb{R}, (x, s) \mapsto \mathcal{H}(x, s)$ with $\mathcal{H}(x, 0) = f(x) = 0$ and $\mathcal{H}(x, 1) = g(x) = \sin^2(3x) + 5$. **Right**: The zero-sphere S^0 depicted by two dots is mapped to the vacuum manifold $\mathcal{M}$, consisting of two disconnected components (teal and purple). Three different maps $f, g, h : S^0 \to \mathcal{M}$ with $f(x_A) = g(x_A) = h(x_A) = \phi_A$ are shown. h and g are homotopic $h \sim g$, as one can define a homotopy that continuously moves $g(x_B)$ to $h(x_B)$. h and g describe, therefore, equivalent domain wall solutions; they correspond to the same element in $\pi_0(\mathcal{M})$. There is no such homotopy between f and g since $f(x_B)$ and $g(x_B)$ lie in different path-connected components of $\mathcal{M}$; they belong to different elements of $\pi_0(\mathcal{M})$. More explicitly, f is null-homotopic and describes a vacuum configuration; there is no domain wall.

f and g. Homotopy is an equivalence relation between continuous functions $f, g : X \to Y$, and if there exists a homotopy between f and g, these are called homotopic. We write $f \sim g$ and denote the corresponding equivalence classes $[f]$. Roughly speaking, this means $\mathcal{H}$ continuously interpolates between f and g. If a homotopy between two such maps f and g exists, this means they can be continuously deformed into each other as shown by means of an explicit example of a homotopy of two curves in Fig. 2.2a. Since we are working with pointed spaces (X, x_0), (Y, y_0) and pointed maps, we also consider our homotopy to be pointed in the following, i.e., $\forall s \in [0, 1] : \mathcal{H}(x_0, s) = y_0$.

With this, we can formalise the previous statement. To each element of

$$\pi_0\left(\mathcal{M}\right) = \left\{ \, [f] \;\middle|\; f : (S^0, x_A) \to (\mathcal{M}, \phi_A), \; f(x_A) = \phi_A \, \right\} \tag{2.13}$$

that does not contain the constant map, corresponds a class of topologically equivalent domain wall configurations.[1] The element that contains the constant map, which is the class of null-homotopic functions, corresponds to a configuration without any domain wall. Configurations associated with different elements of $\pi_0(\mathcal{M})$ cannot be continuously deformed into each other. Combining this with the knowledge of defect formation from Sec. 1.3, we can state that in a phase transition in which the symmetry $G \to H$ is broken, domain walls will arise if $\pi_0(G/H) \ncong \{ \, \mathbb{1} \, \}$, or equivalently, if the vacuum manifold is disconnected. The set $\pi_0(\mathcal{M})$ is called the zeroth-homotopy group. If $\mathcal{M}$ itself has a group structure, which is the case if H is a normal subgroup of G, and $\mathcal{M}_0$ is its identity component, then $\pi_0(\mathcal{M})$ can be equipped with a group structure and is isomorphic to the quotient group $\pi_0(\mathcal{M}) \cong \mathcal{M}/\mathcal{M}_0$. In general, $\pi_0(\mathcal{M})$ should, however, be thought of as a set. For the simple example of the symmetry breaking $\mathbb{Z}_2 \to \{ \, \mathbb{1} \, \}$, $\pi_0(\mathcal{M})$ actually inherits a group structure from the the vacuum manifold $\mathcal{M} \cong \mathbb{Z}_2$, namely $\pi_0(\mathcal{M}) \cong \mathbb{Z}_2$.

It should eventually be remarked that the non-triviality of these configurations relies on fixing the field values in regions A and B to lie within the vacuum. Although this need not hold for arbitrary regions, one typically takes these regions to be at spatial infinity, where one can sensibly assume a vacuum to be reached. Note that this picture is useful for the classification of defects, but not necessarily for practical purposes. There, one could, for example, consider a closed domain wall or multiple domain walls that are topologically equivalent to the trivial vacuum if A and B are considered to be at infinity.

---

[1] We have omitted the dependence of the set $\pi_0(\mathcal{M})$ on ϕ_A in our definition, since it is isomorphic to the set of path-connected components of $\mathcal{M}$ and, hence, independent of the basepoint ϕ_A.

2.2.2 Classification of stable defects

We have now seen that homotopy theory is the correct framework to describe domain walls. This is actually also true for other (stable) cosmic defects, as was realised by KIBBLE when he introduced a classification scheme of defects based on the homotopy groups of the vacuum manifold in Ref. [32].

Defects of different dimensions require different homotopy groups, and we can define them as

$$\pi_n\left(\mathcal{M}\right) = \left\{ \, [f] \mid (S^n, p) \to (\mathcal{M}, \phi_0) \text{ with } f \text{ continuous and } f(p) = \phi_0 \, \right\} . \tag{2.14}$$

This definition means that we map n-spheres to images of n-spheres in $\mathcal{M}$. We consider two maps to be equivalent if they are homotopic, i.e., if they can be continuously deformed into each other. We impose, however, the requirement that all maps are fixed at one point such that the basepoint $p \in S^n$ is mapped to the basepoint $\phi_0 \in \mathcal{M}$. These sets can be equipped with a group structure,[2] which justifies their name. Importantly, every homotopy group will contain a neutral element, which is the equivalence class that contains the constant map $f_{\text{const}} : S^n \to \{ \, \phi_0 \, \}$. Furthermore, homotopy groups are independent of their basepoint $\phi_0 \in \mathcal{M}$ if the vacuum space is path-connected $\pi_0\left(\mathcal{M}\right) \cong \{ \, 1 \, \}$.

This abstract definition directly relates to defects, as we want to discuss in the following. If the n-th homotopy group of the vacuum manifold is trivial $\pi_n(\mathcal{M}) \cong \{ \, \mathbb{1} \, \}$, this means that all (basepoint-preserving) continuous maps from the n-sphere to $\mathcal{M}$ are homotopic to the constant map. It is instructive to construct such a homotopy explicitly. Consider a map $\alpha_R : S^n \to \mathbb{R}^d$ from the abstract n-sphere that appears in the definition of the homotopy group to our physical space $\mathbb{R}^d$. This map should be basepoint-preserving $\alpha_R(p) = x_0$. We choose α_R in such a way that it maps S^n homeomorphically to an n-sphere $S^n_R \subset \mathbb{R}^d$ of radius R, i.e., if we restrict the codomain of α_R to S^n_R, then the restricted function is a homeomorphism. This is only possible if $n < d$, and we will confine ourselves to this case. We can now obtain a homotopy $\tilde{\mathcal{H}} : S^n \times [0,1] \to \mathbb{R}^d$ between α_R and the constant map $\alpha_0(S^n) = \{ \, x_0 \, \}$ as the continuous function defined in such a way that $\mathcal{H}(\cdot, s)$ sends S^n homeomorphically to the n-sphere $S^n_r \subset \mathcal{R}^d$ of radius $r = (1-s)R$; α_R is null-homotopic. By increasing s from 0 to 1, we shrink the sphere of radius R fixed at x_0 continuously down to the single point x_0. We can now make use of our Higgs field $\Phi : \mathbb{R}^d \to \mathcal{V}$ with $\mathcal{V}$ being the space of all Higgs field values. We are interested in the case where the Higgs field takes on S^n_R only vacuum values. As previously discussed, this makes sense if we consider $R \to \infty$. Moreover, we define $\phi_0 = \Phi(x_0)$ as the basepoint of our vacuum manifold. By composing Φ with α_R, we can define a function from the abstract n-sphere to the vacuum manifold $f = \Phi \circ \alpha_R : S^n \to \mathcal{M} \subseteq \mathcal{V}$ which is continuous and basepoint-preserving. Similarly, $f_0 = \Phi \circ \alpha_0$ is the constant map with image $\{ \, \phi_0 \, \}$. The map f describes all the different vacuum values that our field Φ takes at the different points on the sphere S^n_R. If $\pi_n(\mathcal{M})$ is trivial, this map must necessarily be null-homotopic. We can find a configuration Φ such that not only on S^n_R, but also in its interior, the vacuum is assumed everywhere, and thus, there is a homotopy $\mathcal{H} = \Phi \circ \tilde{\mathcal{H}} : S^n \times [0,1] \to \mathcal{M}$. While $\tilde{\mathcal{H}}$ shrinks the geometric sphere in physical space, $\mathcal{H}$ shrinks the image of that sphere under Φ.

If $\pi_n(\mathcal{M})$ is non-trivial, f can represent a non-trivial class in this group, and is thus not null-homotopic. In this case, the above construction is not possible. This means there is no configuration Φ for which the vacuum is assumed on a sphere S^n_R and also throughout its interior. If a vacuum configuration corresponding to such a non-trivial map f arises in a phase transition, the topology of the vacuum requires that there is at least one point inside the sphere S^n_R at which the vacuum cannot be reached, and the phase transition does not complete. The sphere must contain a defect. The argument so far requires a non-vacuum configuration at some point in the interior of an n-sphere, i.e., in an $n+1$-ball. We can now continuously move in physical space in any of the $d-(n+1)$ directions normal to the n-sphere and apply the same argument as long as our n-sphere in physical

---

[2] This holds only for $n \geq 1$. For a basic introduction to homotopy groups, see, e.g. [159].

space is mapped to the vacuum manifold. For each such n-sphere, we have then at least one point in its interior at which the vacuum is not reached. All these points along the normal directions combine to an object of codimension $n+1$.

We can summarise the preceding discussion: If $\pi_n(\mathcal{M}) \ncong \{ \mathbb{1} \}$, there are topologically stabilised non-vacuum configurations (defects) of codimension $n+1$. Conversely, if $\pi_n(\mathcal{M}) \cong \{ \mathbb{1} \}$, no such defects are topologically stable.

Let us restrict to the case of $d=3$. For $n=0$, we have already seen that $\pi_0(\mathcal{M}) \ncong \{ \mathbb{1} \}$ gives rise to domain walls. This agrees with our statement that defects associated with a disconnected vacuum manifold are 2-dimensional. Defects corresponding to a non-trivial first homotopy group (also referred to as the fundamental group) are 1-dimensional, namely strings. As an illustration, recall Ex. 1.2. The vacuum manifold of this model was $\mathcal{M} \cong S^1$, which has a fundamental group $\pi_1(S^1) \cong \mathbb{Z}$. This group is non-trivial, and the model contains topologically stabilised strings. Indeed, we have seen at the end of Sec. 1.3 that this model will give rise to cosmic strings in a cosmological phase transition. A non-trivial second homotopy group gives rise to 0-dimensional defects, i.e., monopoles. We also encountered them at the end of Sec. 1.3 as a result of a cosmological phase transition for the model given in Ex. 1.2. This model has $\mathcal{M} \cong S^2$ for which one finds the non-trivial second homotopy group $\pi_2(S^2) \cong \mathbb{Z}$, and the above formalism indeed predicts stable monopole configurations.

In summary, a phase transition $G \to H$ will produce topologically stabilised defects of codimension n if $\pi_n(G/H) \ncong \{ \mathbb{1} \}$. In three spatial dimensions, the correspondence between defect configurations and homotopy groups is:

- Domain walls: $\pi_0(G/H)$
- Cosmic strings: $\pi_1(G/H)$
- Monopoles: $\pi_2(G/H)$

There is one caveat to the above summary, which is related to the introduction of a basepoint in the definition of homotopy groups. Fixing a basepoint of $\mathcal{M}$ introduces the unphysical criterion that field configurations have to attain a certain value at the arbitrarily chosen basepoint x_0 of S^n_R. Physically different configurations are those that cannot be continuously deformed into each other after the maps are released from their basepoints. We should, therefore, actually consider another equivalence relation that identifies elements of the homotopy groups which are only different due to the fixed basepoints. We should consider free homotopy in contrast to based homotopy. In practice, only $\pi_1(\mathcal{M})$ requires transitioning from based to free homotopy (via conjugacy classes) if it is non-Abelian. For $\pi_2(\mathcal{M})$ (and all higher homotopy groups), the groups are inherently Abelian and, therefore, based and free homotopy coincide, provided $\mathcal{M}$ itself is simply connected. In the models studied in this thesis, there is no difference between based and free homotopy, so we shall not address this subtlety in more detail. The interested reader is referred to Ref. [36] for an explicit discussion.

Non-trivial homotopy groups provide sufficient conditions for defect stabilisation, but they are not necessary. Energetic or dynamical considerations can also prohibit decay. In addition, purely topological methods cannot predict physically important quantities such as defect tensions, masses, or detailed field profiles. We, therefore, now turn to the field-theoretical description of strings and monopoles.

2.2.3 Strings — field theoretical aspects

Local $U(1)$ strings: The archetypical model of strings is the Abelian Higgs model that was discussed in great detail in Sec. 1.1.1 and Ex. 1.4. Here, we reproduce its Lagrangian for convenience

$$\mathcal{L} = (\mathcal{D}_\mu \phi)^* (\mathcal{D}^\mu \phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{\lambda}{4} \left(|\phi|^2 - \eta^2 \right)^2 , \tag{2.15}$$

where $\mathcal{D}_\mu\phi = (\partial_\mu - ieA_\mu)\phi$. As we saw at the end of Sec. 1.3, a cosmological phase transition will in this model indeed lead to the production of cosmic strings. It can be illuminating to view the appearance of cosmic strings explicitly in the previously introduced formalism. Strings are labelled by the fundamental group of the vacuum manifold $\mathcal{M} \cong S^1$, which is non-trivial $\pi_1(S^1) = \mathbb{Z}$.

Consider now a circle S^1_R with very large radius $R \to \infty$ such that we can sensibly assume that the field has settled to the vacuum everywhere along the loop. A possible field configuration on this circle in cylindrical coordinates is

$$\phi(\rho = R, \varphi, z) = \eta e^{in\varphi}\,, \qquad \text{with} \qquad n \in \mathbb{Z}\,. \tag{2.16}$$

As required, the field values lie completely in the vacuum manifold since $|\phi| = \eta$. If $n = 0$, the configuration maps every point in space to the same value $\eta \in \mathcal{M}$ and is (trivially) null-homotopic, thus corresponding to the trivial element of $\pi_1(\mathcal{M})$. Consequently, the case $n = 0$ is not associated with a string. Everything else would have been a surprise since imposing $\phi = \eta$ on a circle does not hinder us in any way from choosing $\phi = \eta$ everywhere in space.

If $n \neq 0$, the situation is different. As an example, consider $n = 1$ for which the field winds once around the vacuum manifold. If we continuously shrink the circle in physical space and require that the field remains in the vacuum, then, due to continuity, on each smaller loop, the configuration still needs to wind once around $\mathcal{M}$. Eventually, when making the circle arbitrarily small, all different vacuum values must be realised in an arbitrarily small region of space, and the field's gradient energy will diverge. To put it differently, such a configuration must be multivalued and thus discontinuous. In conclusion, there is no physical and, therefore, continuous field configuration for which the field assumes the vacuum configuration $\phi = \eta e^{i\varphi}$ on S^n_R and at the same time takes values in $\mathcal{M}$ everywhere inside the circle. Instead, the field must leave the vacuum. This is depicted in Fig. 2.3. On the circle γ_1 in real space, the field behaves as in Eq. (2.16) and winds once around the vacuum manifold shown in purple on the right side. This map is not null-homotopic and is associated with a string. Inside the circle γ_1, the field must assume non-vacuum values. In fact, at least at one point in space, the field value must vanish to avoid multivaluedness. At this point, the local maximum of the potential is reached, and the symmetry remains unbroken. We indicate the region of unbroken symmetry as before in orange.

What we just argued is just a rephrasing of the situation we encountered already at the end of Sec. 1.3 and which is depicted in Fig. 1.5. In Fig. 2.3, we also show a second circle γ_2 in physical space that does not surround the string. Again, we fix the field along γ_2 to lie everywhere in the vacuum. The map from the circle γ_2 to the vacuum can, however, be continuously shrunk to a point without leaving the vacuum. The loop γ_2 does not contain a string.

Similar considerations apply for other winding numbers $n \neq 0$. For $n = 2, 3, \ldots$, the field values will just run $2, 3, \ldots$ times around the vacuum while traversing a closed loop around the string in physical space once. For $n < 0$, the winding of the vacuum manifold occurs in the opposite direction. In fact, each value of n gives rise to a string associated with a different element of $\pi_1(\mathcal{M}) \cong \mathbb{Z}$.

Let us now consider the case where the string configuration is invariant under translations along the z-axis. Its total energy is clearly divergent due to its infinite extent. Nevertheless, we can compute its finite energy per unit length, known as the *string tension* μ, as the integral of the energy density, which reads for a static configuration in temporal gauge [36]

$$\mu = \int_0^\infty \rho\, \mathrm{d}\rho \int_0^{2\pi} \mathrm{d}\varphi \left(-\left(\mathcal{D}_i\phi\right)^* \left(\mathcal{D}^i\phi\right) + \frac{1}{2}\left(\boldsymbol{E}^2 + \boldsymbol{B}^2\right) + \frac{\lambda}{4}\left(|\phi|^2 - \eta^2\right)^2 \right)\,. \tag{2.17}$$

The electric and magnetic fields are defined in the usual way $E_i = F_{0i}$ and $B_i = -1/2\epsilon_{ijk}F^{jk}$ (with $\epsilon_{xyz} = 1$), and we denote $\boldsymbol{E}^2 = -E_iE^i$ and $\boldsymbol{B}^2 = -B_iB^i$. Since we are considering a static configuration, finding solutions to the equations of motion reduces to the task of finding the

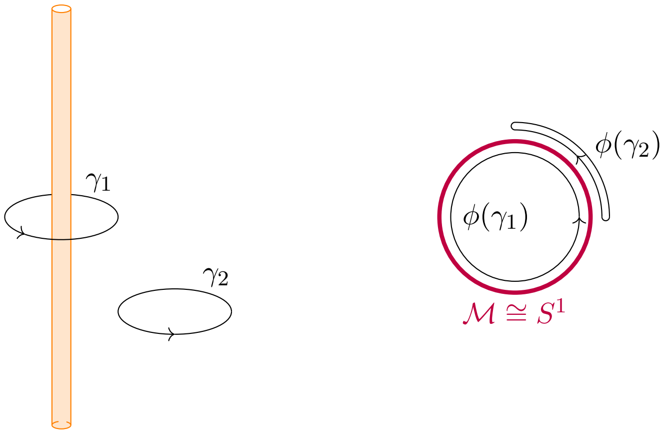


Figure 2.3: Illustration of the winding associated with a string. On the left, physical space with two loops γ_1 and γ_2 is shown. Both loops are mapped by ϕ to the vacuum manifold $\mathcal{M}$ in purple. The loop γ_1 winds $\mathcal{M}$ once, such that it must enclose a string, depicted as an orange cylinder. The loop γ_2 does not wind $\mathcal{M}$ and can be contracted to a point in the vacuum manifold. It does not contain a string.

configurations that minimise the string tension μ. All three terms contributing to the tension in Eq. (2.17) are non-negative.[3]

If our configuration at $\rho \to \infty$ is of the form given in Eq. (2.16), then the potential term already vanishes and the first term can be minimised if $\mathcal{D}_i\phi = 0$, or, equivalently, if

$$A_\varphi|_{\rho\to\infty} = \frac{n}{e} \tag{2.18}$$

and $A_\rho = A_z = 0$. The gauge field is pure gauge, and we have also minimised the energy of the electromagnetic field. Previously, we only considered the field ϕ and concluded that if it has the form given in Eq. (2.16) on a circle in space, it cannot be a vacuum configuration everywhere. Similarly, we can now see from the gauge field that, even though it is pure gauge on the circle we considered, it cannot be pure gauge everywhere. In fact, we can compute the magnetic flux along the string (parallel to $\hat{z}$) and find from Stokes' theorem

$$\mathbf{\Phi} = \int_0^\infty \rho\,\mathrm{d}\rho \int_0^{2\pi} \mathrm{d}\varphi\, B_z = -\int_0^{2\pi} \mathrm{d}\varphi\, A_\varphi|_{\rho\to\infty} = -\frac{2\pi n}{e}\,, \tag{2.19}$$

assuming regularity of the gauge field at $\rho = 0$. The string, therefore, carries a quantised magnetic flux and can be viewed as the analogue of quantised flux tubes in superconductors [24].

Let us now consider a possible field configuration that tells us what the field looks like everywhere in space, i.e., for every ρ, and restrict to solutions with cylindrical symmetry and translational invariance along the z-axis. In this case, one can find an ansatz that reads (in temporal and Lorenz gauge) [24, 28]

$$\phi(\rho,\varphi) = \eta e^{in\varphi}h(\rho)\,, \qquad A_0 = A_\rho = A_z = 0\,, \qquad A_\varphi(\rho) = \frac{n}{e}(1-k(\rho))\,, \tag{2.20}$$

and strings of this form are known as *Abrikosov–Nielsen–Olesen (ANO) strings*. We require that for $\rho \to \infty$, the configuration approaches a vacuum solution, and by imposing the conditions

$$h(\rho) \overset{\rho\to\infty}{\longrightarrow} 1 \qquad \text{and} \qquad k(\rho) \overset{\rho\to\infty}{\longrightarrow} 0\,, \tag{2.21}$$

---

[3] The negative sign in the first term cancels with the negative sign from the metric signature and makes the overall term positive.

it matches the asymptotic configuration of Eq. (2.16). Furthermore, the configuration is clearly multivalued if $h(0) \neq 0$. Similarly, the energy will diverge logarithmically if $k(0) \neq 1$, and we need to impose

$$h(0) = 0 \qquad \text{and} \qquad k(0) = 1\,. \tag{2.22}$$

To understand string solutions in detail and to fix the profile functions h and k, we need to study the equations of motion derived from the Lagrangian in Eq. (2.15). This can be complicated, and generally, one needs to resort to numerical methods. A lot can, however, already be learned if one restricts to the case of critical coupling $\lambda = 2e^2$.[4] Note that this is just the case in which Higgs and gauge field have the same masses (cf. Sec. 1.1.1). As shown originally in Ref. [161], it is possible to use an integration by parts to write the tension in Eq. (2.17) for the ANO string in the form

$$\mu = \int \mathrm{d}^2x \left(|\mathcal{D}_x \pm i\mathcal{D}_y\phi|^2 + \frac{1}{2}\left(B_z \pm \sqrt{\frac{\lambda}{2}}\left(|\phi|^2 - \eta^2\right)\right)^2 \right) \pm e\eta^2 \int \mathrm{d}^2x\, B_z\,. \tag{2.23}$$

The upper sign applies to $n < 0$, while the lower sign applies to $n \geq 0$. The second integral in this equation is, up to a prefactor, equivalent to the magnetic flux in Eq. (2.19). The first integral is clearly non-negative, such that the second integral determines a lower bound on the string tension

$$\mu \geq 2\pi\eta^2|n|\,, \tag{2.24}$$

which is referred to as the *Bogomol'nyi bound.* Saturation of this bound requires the first integral to vanish, or equivalently, that the *Bogomol'nyi equations* are satisfied:

$$|\mathcal{D}_x \pm i\mathcal{D}_y\phi| = 0 \qquad \text{and} \qquad B_z \pm \sqrt{\frac{\lambda}{2}}\left(|\phi|^2 - \eta^2\right) = 0\,. \tag{2.25}$$

For the critical coupling $\lambda = 2e^2$, the second-order equations of motion can be replaced by the first-order Bogomol'nyi equations, whose solutions automatically extremise the action. The classification of cosmic strings in terms of the fundamental group of the vacuum manifold allows us to see that having N strings with winding $n = 1$ is topologically the same as having one string with winding $n = N$. Let us refer to the tension of a string with winding n as μ_n. For critical coupling, we see that solutions to the Bogomol'nyi equations satisfy $\mu_N = N\mu_1$. Away from critical couplings, the behaviour was originally studied in Ref. [161, 162]. For type-II strings, i.e. $\lambda > 2e^2$, one finds $\mu_N > N\mu_1$. Accordingly, type-II strings repel each other to reduce their total energy. Type-II strings of larger-than-unit winding will not be stable and break up. Heuristically, this can be understood as a result of the fact that, in this case, the Compton wavelength of the gauge bosons is larger than that of the Higgs. Interactions mediated by the gauge field are, therefore, of longer range. These, however, are associated with the magnetic flux tubes, which repel each other. The opposite is true for $\lambda < 2e^2$, i.e., type-I strings. Here, one finds that $\mu_N < N\mu_1$ and accordingly, strings attract each other and form junctions. The heuristic interpretation is that now, since the Compton wavelength of the Higgs field is larger than that of the gauge field, it is more important to minimise the volume in which the Higgs field deviates from its vacuum.

To study general couplings, there is no way around the equations of motion, which are of the general form

$$\nabla^\mu F_{\mu\nu} - ie\left(\phi\partial_\nu\phi^* - \phi^*\partial_\nu\phi\right) + 2e^2A_\nu|\phi|^2 = 0\,, \tag{2.26}$$

$$\Box\phi - ie\left(\nabla_\mu A^\mu\phi + 2A^\mu\partial_\mu\phi\right) - e^2A^\mu A_\mu\phi - \frac{\lambda}{2}\left(\eta^2 - |\phi|^2\right)\phi = 0\,, \tag{2.27}$$

---

[4] This coupling is known as distinguishing between type-I superconductors for $\lambda < 2e^2$ and type-II superconductors for $\lambda > 2e^2$ [24, 160]; a nomenclature also widely encountered in the literature on cosmic strings.

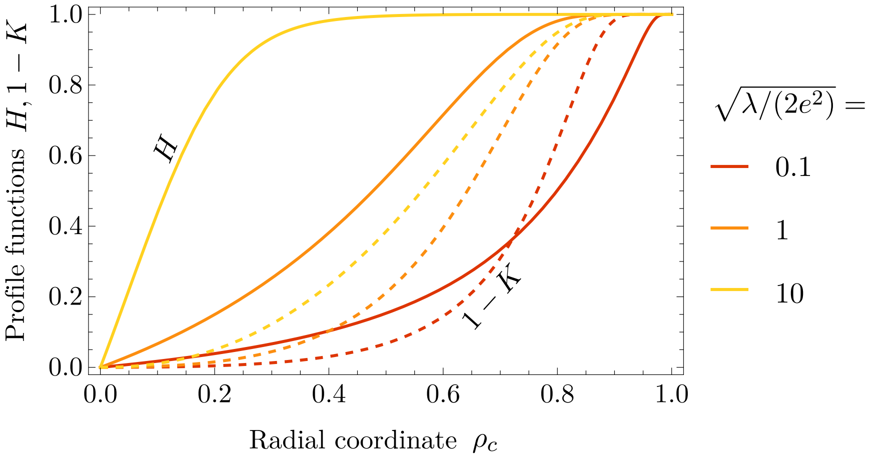


Figure 2.4: Higgs profile function H (solid) and gauge field profile function $1-K$ (dashed) of the Abrikosov–Nielsen–Olesen string with winding $n=1$ for three different values of the mass ratio $m_s/m_v = \sqrt{\lambda/(2e^2)}$. The functions H and K are plotted versus the compactified radial coordinate $\rho_c = \frac{\tilde{\rho}}{1+\tilde{\rho}} \in [0,1]$ and were obtained from numerically solving Eqs. (2.28).
The numerical data used in this figure were kindly provided by Jonah Fabian.

with $\Box = \nabla^\mu \nabla_\mu$. Upon using the ansatz in Eq. (2.16), rescaling the radial coordinate $\tilde{\rho} = \sqrt{2}e\eta\rho$, and introducing $K(\tilde{\rho}) = k(\rho)$ as well as $H(\tilde{\rho}) = h(\rho)$, the equations of motion reduce to the following second-order differential equations for the profile functions

$$K''(\tilde{\rho}) - \frac{K'(\tilde{\rho})}{\tilde{\rho}} - H^2(\tilde{\rho})K(\tilde{\rho}) = 0\,, \tag{2.28}$$

$$H''(\tilde{\rho}) + \frac{H'(\tilde{\rho})}{\tilde{\rho}} - \frac{n^2 H(\tilde{\rho})K^2(\tilde{\rho})}{\tilde{\rho}^2} + \frac{1}{2}\frac{\lambda}{2e^2}\left[1 - H^2(\tilde{\rho})\right] H(\tilde{\rho}) = 0\,. \tag{2.29}$$

The rescaling shows that for a fixed winding n, there is only one free parameter in the theory, namely $\lambda/(2e^2) = m_s^2/m_v^2$, the ratio of the Higgs and gauge field square masses.

Although we do not want to investigate the approximate profile functions of ANO strings in detail, they allow us to make statements about the string width δ — an important quantity for our discussion in Pt. II. For $m_s \lesssim 2m_v$, there exist two characteristic length scales on which the profile functions decay, namely the (reduced) Compton wavelengths m_s^{-1} for the Higgs field and m_v^{-1} for the gauge field [28, 163]. We see that the string consists of an inner magnetic flux tube and outer tube of unbroken symmetry. For $m_s > 2m_v$, the gauge field still decays over a characteristic length m_v^{-1}. For the scalar field, the characteristic length is, however, no longer independent but rather $(2m_v)^{-1}$ [164]. We find here an inner string core of unbroken symmetry surrounded by a magnetic flux tube. Of course, for coupling constants of order unity, $\delta \sim \eta^{-1}$ in both cases, as is already indicated by a dimensional analysis.

In Fig. 2.4, we show the profile functions of the Higgs field H and the gauge field $1-K$ for $n=1$ and three different choices of the mass ratio $m_s^2/m_v^2 = \lambda/(2e^2)$. We indeed see that for small values of this parameter, we obtain a region of unbroken symmetry which is large compared to the flux tube, whereas large mass ratios lead to a thin core of the symmetric Higgs phase surrounded by a wider flux tube. Given the numerical profile functions, one can insert them into the general form of the string tension in Eq. (2.17), and one obtains a result that is shown in Fig. 2.5. The numerical result for the tension is of the form

$$\mu = 2\pi\eta^2 \mathcal{F}\left(\frac{\lambda}{2e^2}\right)\,, \tag{2.30}$$

where $\mathcal{F}$ depends only on $\lambda/(2e^2)$ and can be shown to be monotonically increasing, although slowly [162, 165]. Notably, in the limit $\lambda/(2e^2) \to \infty$, one finds that the string tension diverges logarithmically as [24, 28]

$$\mu \simeq 2\pi\eta^2 \ln\left(\frac{\lambda}{2e^2}\right) . \tag{2.31}$$

This is not an unexpected result since we can view the above limit also as the limit $e \to 0$ in which we approach a string solution without gauge coupling. As we will see in a moment, the string tension is, in fact, divergent for strings obtained from a global symmetry breaking. Let us finally remark that from Fig. 2.5, it can be seen that $\mathcal{F}(1) = 1$, in agreement with a saturated Bogomol'nyi bound given in Eq. (2.24).

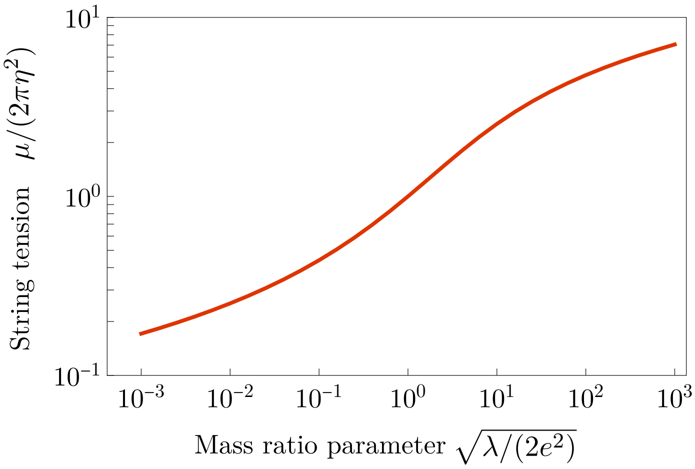


Figure 2.5: String tension normalised to the tension at critical coupling $2\pi\eta^2$ as a function of the parameter $m_s/m_v = \sqrt{\lambda/(2e^2)}$. The tension is obtained from using numerical solutions to Eqs. (2.28) in Eq. (2.17).
The numerical data used in this figure were kindly provided by Jonah Fabian.

Global $U(1)$ strings: After having considered strings from the Abelian Higgs model, it is natural to ask what changes if we do not couple the field ϕ to a gauge field and consider a global symmetry breaking. In this case, we can still use the ansatz given in Eq. (2.20) for the field ϕ. The equations of motion reduce drastically to

$$\Box\phi - \frac{\lambda}{2}\left(\eta^2 - |\phi|^2\right)\phi = 0 , \tag{2.32}$$

and, upon inserting the ansatz, one finds

$$h''(\rho) + \frac{h'(\rho)}{\rho} - \frac{n^2 h(\rho)}{\rho^2} + \frac{\lambda\eta^2}{2}\left(1 - h^2(\rho)\right) h(\rho) = 0 . \tag{2.33}$$

As before, one can derive asymptotic solutions to this equation, but it is sufficient for our purposes to observe that the only relevant length scale in the problem is $m_s^{-1} = 1/(\sqrt{\lambda}\eta)$. This scale must therefore set the characteristic length over which the string core decays.[5]

We can then roughly approximate the string as a configuration with unbroken symmetry inside $\delta = m_s^{-1}$ such that $h(\rho < \delta) \sim 0$, and otherwise described by the asymptotic form $h(\rho > \delta) \sim 1$.

---

[5] An important difference to local strings is that the profile functions solving the equations of motion do not decay exponentially for asymptotically large distances, but rather follow a power-law behaviour (cf. e.g. Ref. [36]).

Computing the string tension, for radii $\rho < \delta$, we only obtain contributions from the field's potential and for $\rho > \delta$, only the field's gradient contributes. In sum, we find

$$\mu \sim \eta^2 + \int_\delta^R \frac{\mathrm{d}\rho}{\rho} \int_0^{2\pi} \mathrm{d}\varphi \, |\partial_\varphi \phi|^2 \simeq 2\pi n^2 \eta^2 \ln\left(\frac{R}{\delta}\right) , \tag{2.34}$$

where we assumed that $R \gg \delta$. The global string tension is thus logarithmically divergent with R.

Considering both the above global string configuration and the local string in a plane orthogonal to the string direction, we observe two explicit ways to evade Derrick's theorem in two spatial dimensions. For the local string, the tension is finite due to the gauge fields cancelling the gradients of the Higgs field at large distances. The introduction of gauge fields makes Derrick's theorem inapplicable. For global strings, the field gradients cause the energy to diverge, and again, the theorem cannot be used.

Even though the above mild divergence of the string tension might be worrying, it is not problematic in a cosmological context. For example, integrating the finite and constant energy density in matter over a plane will give a result that is quadratically diverging with the length scale integrated over. In comparison, the logarithmic divergence of the global-string tension is relatively mild and stems from the long-range interactions mediated by the massless Goldstone modes present in the model. Moreover, the above estimate assumes a single isolated string. In realistic phase transitions, many strings form, and the cutoff length R is bounded from above by the typical inter-string distance. This naturally regulates the divergence and renders the tension physically meaningful.

Global strings will not be of further importance to this thesis. Nevertheless, we want to emphasise that they can be of phenomenological interest. They emerge, for example, as axion strings when a global Peccei–Quinn $U(1)$ symmetry is spontaneously broken [75, 166–170], but also in the case of a spontaneously broken global $U(1)$ associated with baryon-minus-lepton number $B-L$ [171, 172].

Non-Abelian local strings: Strings do not have to emerge from the spontaneous breaking of a $U(1)$ symmetry, but can also arise from the spontaneous breaking of non-Abelian groups. Since their study does not play a role in our further discussion, we will only briefly introduce them through a representative example and comment on their general features, without developing the formal framework.

■ **Example 2.1 — $SU(2) \to \mathbb{Z}_2$.** Let us start from the model in Ex. 1.2 and extend it by another scalar field Ψ in the adjoint representation of $SU(2)$ (in particular, we use the same conventions for Ψ as introduced in Ex. 1.2 for Φ). Importantly, we also introduce interactions between Φ and Ψ in the Lagrangian[6]

$$\mathcal{L} = -\frac{1}{2}\mathrm{tr}\left(F_{\mu\nu}F^{\mu\nu}\right) + \mathrm{tr}\left(\mathcal{D}_\mu\Phi\mathcal{D}^\mu\Phi\right) + \mathrm{tr}\left(\mathcal{D}_\mu\Psi\mathcal{D}^\mu\Psi\right) - V(\Phi,\Psi) , \tag{2.35}$$

with a potential

$$V(\Phi,\Psi) = \frac{\lambda_\Phi}{4}\left(\mathrm{tr}\left(\Phi^2\right) - \eta_\Phi^2\right)^2 + \frac{\lambda_\Psi}{4}\left(\mathrm{tr}\left(\Psi^2\right) - \eta_\Psi^2\right)^2 + \frac{\gamma}{2}\mathrm{tr}\left(\Phi\Psi\right)^2 . \tag{2.36}$$

We can choose the VEV of Φ to be aligned with the $SU(2)$-generator T^3 such that $\phi_1 = \phi_2 = 0$ as well as $\phi_3 = \sqrt{2}\eta_\Phi$, and Φ minimises the first term in the potential. The VEV of Ψ must satisfy $\mathrm{tr}(\Psi^2) = \eta_\Psi^2$ in order to minimise the second term of the potential, but simultaneously, it must also minimise the third term. The VEVs of Φ and Ψ are therefore not independent. More explicitly, Ψ

---

[6] That this model admits stable string solutions was in fact already found in the appendix of Ref. [28].

must be orthogonal to Φ and we can choose explicitly $\psi_1 = \psi_3 = 0$ and $\psi_2 = \sqrt{2}\eta_\Psi$. Note that the kernel of the triplet representation is the centre $\mathbb{Z}_2$ of $SU(2)$. With the two triplets acquiring a non-vanishing VEV as above, the symmetry $SO(3) \cong SU(2)/\mathbb{Z}_2$ is completely broken, and hence, the overall symmetry breaking is of the form $SU(2) \to \mathbb{Z}_2$. The particle spectrum consists of three massive Higgs fields and three massive gauge fields. Furthermore, the vacuum manifold $\mathcal{M} \cong SU(2)/\mathbb{Z}_2$ has one non-trivial homotopy group, namely $\pi_1(SU(2)/\mathbb{Z}_2) \cong \mathbb{Z}_2$ (in Eq. (2.55), we present a relation that allows us to easily compute this fundamental group). This tells us that the model contains only two topologically distinct sectors; one of them corresponding to strings and one of them corresponding to the vacuum.

If the Higgs fields Φ and Ψ acquire their VEV everywhere in space as described above, the configuration is pure vacuum and a representative of the trivial element of the fundamental group of $\mathcal{M}$. We can allow Ψ to wind once around the plane orthogonal to Φ, which leads to an ansatz of the form [173]

$$\phi_3 = \sqrt{2}\eta_\Phi, \quad \psi_1 = \sqrt{2}\eta_\Psi h(\rho)\cos(\varphi), \quad \psi_2 = \sqrt{2}\eta_\Psi h(\rho)\sin(\varphi), \quad A^3_\varphi = \frac{1}{c}(1 - k(\rho)), \tag{2.37}$$

and all other components vanish. The configuration formally resembles the ANO string configuration in Eq. (2.20) and the profile functions need to satisfy the same conditions as in Eqs. (2.21) and (2.22). The winding is associated with the field Ψ, which breaks the symmetry generated by T^3, which is why we find our winding again in the corresponding gauge field A^3. Defining our field strength as $\tilde{F}_{\mu\nu} = 2\partial_{[\mu}A^3_{\nu]}$, we can compute the magnetic flux through a plane with normal vector parallel to the z-direction as [173]

$$\mathbf{\Phi} = -\frac{2\pi}{c}\,. \tag{2.38}$$

The formal resemblance of the $\mathbb{Z}_2$-string to the ANO strings also seems to indicate that we should be able to find string solutions with arbitrary winding. However, the fundamental group does not support that statement. Let us reemphasise that topological criteria are sufficient for the stability of defects, but not necessary. In fact, the above model allows for strings of higher winding. To understand why strings with odd winding number are topologically equivalent to strings with unit winding and strings with even winding are topologically trivial, we have to observe that the strings we consider basically arise from the breaking $U(1) \to \mathbb{Z}_2$. This $U(1)$ is generated by T^3, the generator unbroken by the VEV of Φ, and indeed, the fundamental group $\pi_1\left(U(1)/\mathbb{Z}_2\right) \cong \mathbb{Z}$ would admit topologically stable strings for any non-trivial integer winding. In our model, the $U(1)$ group is, however, embedded in a larger $SU(2)$ and higher winding numbers can be unwound in the larger symmetry group. While this statement might seem abstract now, it will become clear when we have discussed the classification of metastable strings. We will comment on this model again in Sec. 2.4 when we introduce composite defects. ■

$\mathbb{Z}_2$-strings are not restricted to symmetry breakings involving $SU(2)$ groups. In fact, they will emerge in any symmetry breaking of the form $G \to H \times \mathbb{Z}_2$ where G is a connected and simply connected group [174].[7] A particular example is the GUT symmetry breaking $\mathrm{Spin}(10) \to SU(5) \times \mathbb{Z}_2$ [175].

More generally, the kernel of the adjoint representation of $SU(N)$ is $\mathbb{Z}_N$, which allows for symmetry breakings of the form $SU(N) \to \mathbb{Z}_N$ by means of N Higgs fields in the adjoint representation and a potential that fixes their expectation values in mutually orthogonal directions. This was originally investigated in Refs. [173, 176].

A comprehensive treatment of non-Abelian string solutions lies beyond the scope of this thesis. The interested reader is, in addition to the mentioned literature, referred to Refs. [177–182].

---

[7] More topologically distinct strings can emerge if H is not path-connected.

2.2.4 Monopoles — field theoretical aspects

’t Hooft–Polyakov monopoles: Monopoles are topologically stabilised if the second homotopy group of the vacuum is non-trivial $\pi_2(\mathcal{M}) \not\cong \{\, \mathbb{1}\, \}$. Monopole configurations were first introduced by ’t Hooft [29] and Polyakov [30, 183] in the context of the model in Ex. 1.2. The *’t Hooft–Polyakov monopole*, also known as *hedgehog* configuration, reads (in temporal and Lorenz gauge)

$$\phi_i = \sqrt{2}\eta f(r)\hat{\boldsymbol{x}}^i\,, \qquad A_0 = 0\,, \qquad A_i^a = \frac{1}{c}\epsilon_{aij}\hat{\boldsymbol{x}}^j\frac{1-g(r)}{r}\,. \tag{2.39}$$

For sufficiently large radii, we want this configuration to approach the vacuum, which requires $\mathrm{tr}(\Phi^2) \overset{r\to\infty}{\longrightarrow} \eta^2$. We can also see that it is only possible to avoid multivaluedness if $\Phi(r=0)=0$. For the profile function, this imposes the conditions

$$f(0) = 0 \qquad \text{and} \qquad f(r) \overset{r\to\infty}{\longrightarrow} 1\,. \tag{2.40}$$

Similarly, to make gradient energies at spatial infinity vanish, we require $\lim_{r\to\infty}\mathcal{D}_\mu\Phi = 0$ and together with regularity at the spatial origin, the gauge-field-profile function must satisfy

$$g(0) = 1 \qquad \text{and} \qquad g(r) \overset{r\to\infty}{\longrightarrow} 0\,. \tag{2.41}$$

Upon using this ansatz, rescaling $\tilde{r} = \sqrt{2}c\eta r$, and defining $F(\tilde{r}) = f(r)$ as well as $G(\tilde{r}) = g(r)$, the equations of motion derived from the Lagrangian in Eq. (1.39) yield the following differential equations for the profile functions:

$$\begin{aligned} F''(\tilde{r}) + \frac{2}{\tilde{r}}F'(\tilde{r}) - \frac{2}{\tilde{r}^2}F(\tilde{r})G^2(\tilde{r}) + \frac{1}{2}\frac{\lambda}{2c^2}\left[1 - F^2(\tilde{r})\right]F(\tilde{r}) &= 0\,, \\ G''(\tilde{r}) + \frac{G(\tilde{r})}{\tilde{r}^2}\left[1 - G^2(\tilde{r})\right] - F^2(\tilde{r})G(\tilde{r}) &= 0\,. \end{aligned} \tag{2.42}$$

In Fig. 2.6, we show the numerical solution of the above equations for different values of the mass ratio $\lambda/(2c^2) = m_s^2/m_v^2$. Due to the rescaling, the scale $\tilde{r}$ is in units of the inverse gauge boson mass m_v^{-1}. As one would expect, with increasing mass ratio, the monopole core of the unbroken Higgs phase becomes increasingly smaller compared to m_v^{-1}. For a detailed analysis of the asymptotic structure of the profile functions, see Ref. [184].

Let us compute the magnetic flux associated with the unbroken $U(1)$ subgroup through a sphere at infinity $S^2_{R\to\infty}$ around the monopole. Since Φ varies from point to point, the way the unbroken $U(1)$ is embedded in $SU(2)$ also changes continuously. Following Ref. [29], the $U(1)$ field strength can be defined as

$$\mathcal{F}_{\mu\nu} = \frac{2}{|\Phi|}\mathrm{tr}(\Phi A_{\mu\nu}) + \frac{2i}{c|\Phi|^3}\mathrm{tr}\left(\Phi\left[\mathcal{D}_\mu\Phi, \mathcal{D}_\nu\Phi\right]\right)\,, \tag{2.43}$$

with $|\Phi| = \sqrt{2\mathrm{tr}\,(\Phi^2)}$. This definition is both gauge invariant and ensures that if $\phi^a = \sqrt{2}\eta\delta^a_z$ everywhere in space, the field strength reduces to $\mathcal{F}_{\mu\nu} = \partial_\mu A^3_\nu - \partial_\nu A^3_\mu$. For the above monopole configuration, the only two non-vanishing components are

$$\mathcal{F}_{\theta\varphi} = -\mathcal{F}_{\varphi\theta} = -\frac{\sin(\theta)}{c}\,. \tag{2.44}$$

Correspondingly, we find that the only non-vanishing component of the associated magnetic field strength reads $B_r = 1/(cr^2)$. For the radially outgoing magnetic flux through a sphere centred at the monopole, we obtain then

$$\boldsymbol{\Phi} = \oint_{S^2_{R\to\infty}} \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi\,B_r r^2 = \frac{4\pi}{c}\,. \tag{2.45}$$

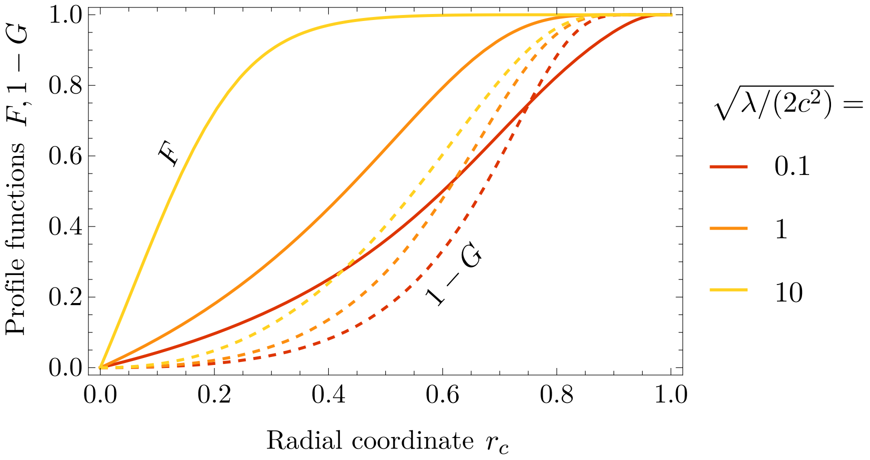


Figure 2.6: Higgs profile function F (solid) and gauge field profile function $1-G$ (dashed) of the 't Hooft–Polyakov monopole for three different values of the mass ratio $m_s/m_v = \sqrt{\lambda/(2c^2)}$. The functions are plotted versus the compactified radial coordinate $r_c = \frac{\tilde{r}}{1+\tilde{r}} \in [0,1]$ and were obtained from numerically solving Eqs. (2.42).
The numerical data used in this figure were kindly provided by Jonah Fabian.

Observe that both local cosmic strings and monopoles carry a magnetic flux. In the case of the string, the magnetic flux is associated with a broken generator and the flux is accordingly confined. In the case of the monopole, the magnetic field corresponds to an unbroken generator and is unconfined, which manifests as a long-range decay of the magnetic field $B_r \propto r^{-2}$.

When viewed purely in terms of their $U(1)$-magnetic field, and disregarding the embedding in a larger $SU(2)$ gauge theory, these monopoles simply appear as magnetically charged point particles. This possibility was already proposed by DIRAC in Ref. [185]. These magnetically charged point particles, which are absent in Maxwell's electrodynamics, are called Dirac monopoles and are the reason the above field configurations are referred to as monopoles.[8]

In contrast to the string and the domain wall, monopoles are localised in all three spatial dimensions. We can compute the total energy of the 't Hooft–Polyakov monopole as

$$m_{\mathrm{M}} = \int \mathrm{d}^3x \left(-\mathrm{tr}(B_i B^i) - \mathrm{tr}\left((\mathcal{D}_i\Phi)\left(\mathcal{D}^i\Phi\right)\right) + \frac{\lambda}{2}\left(\mathrm{tr}\left(\Phi^2\right) - \eta^2\right) \right) , \tag{2.46}$$

with $B^i = -\frac{1}{2}\epsilon^{ijk}F_{jk}$. This energy is referred to as the monopole mass. All three terms in the integrand are non-negative, and Bogomol'nyi's method can also be applied in this case. Focusing on the first two terms and completing the square, we obtain the inequality

$$m_{\mathrm{M}} \geq -\int \mathrm{d}^3x \,\mathrm{tr}\left((B_i \pm \mathcal{D}_i\Phi)\left(B^i \pm \mathcal{D}^i\Phi\right)\right) \pm 2\int \mathrm{d}^3x \,\mathrm{tr}\left(B_i \mathcal{D}^i\Phi\right) , \tag{2.47}$$

where the sign is chosen such that the last term, including the sign, is positive. Since the monopole satisfies $\mathcal{D}_i B^i = 0$ and $\mathrm{tr}(B_i\phi)$ is an $SU(2)$-singlet, we can perform an integration by parts on the last term. Since the first term is non-negative, we can compute

$$m_{\mathrm{M}} \geq \left| \oint_{S^2_{R\to\infty}} \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi\, B^a_r \phi^a \right| = \frac{4\pi}{c^2} m_v \equiv m_{\mathrm{BPS}} , \tag{2.48}$$

where m_v is the gauge boson mass. This lower bound on the monopole mass is again known as the Bogomol'nyi bound and due to Ref. [187].

---

[8] A non-singular description of Dirac's monopole in terms of principal bundles was found in Ref. [186] and is known by the name *Wu–Yang monopole*.

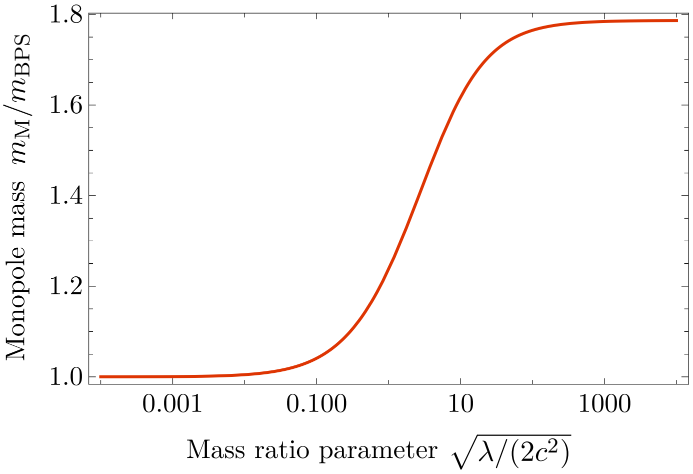


Figure 2.7: Monopole mass normalised to the monopole mass in the BPS limit $m_{\mathrm{BPS}} = 4\pi m_v/c^2$ as a function of the parameter $m_s/m_v = \sqrt{\lambda/(2c^2)}$. The tension is obtained from using numerical solutions to Eqs. (2.42) in Eq. (2.46). *The numerical data used in this figure were kindly provided by Jonah Fabian.*

We can saturate the bound in Eq. (2.47) by taking the limit $\lambda/(2c^2) \to 0$. In order for the Bogomol'nyi bound in Eq. (2.48) to be saturated, the first integrand in Eq. (2.47) must vanish too:

$$B_i \pm \mathcal{D}_i \Phi = 0\,. \tag{2.49}$$

This limit, now known as the *Bogomol'nyi–Prasad–Sommerfield (BPS) limit*, was first considered in Refs. [188, 189], where exact solutions to the above equations were found. In the BPS limit, there is no net force between equally charged monopoles and exact static multi-monopole solutions can be constructed [190].

The monopole mass obtained from using the numerical solutions to Eqs. (2.42) inserted into Eq. (2.46), is plotted in Fig. 2.7 as a function of the mass ratio m_s/m_v. As predicted, one finds that the monopole mass approaches $m_{\mathrm{M}} \to m_{\mathrm{BPS}}$ in the BPS limit. In the opposite limit $m_s/m_v \to \infty$, the monopole mass can also be found to approach a constant value of $m_{\mathrm{M}} \to 1.787\, m_{\mathrm{BPS}}$ [184].

Monopole problem: Monopoles do not only occur in simple symmetry breakings such as $SU(2) \to U(1)$, but are a generic prediction of GUTs [37, 38]. The reason for this is that for any symmetry breaking $G \to H$ with G connected and simply connected, the second homotopy group is

$$\pi_2(G/H) \cong \pi_1(H_0)\,, \tag{2.50}$$

where H_0 is the identity component of H. After electroweak symmetry breaking, the SM gauge group G_{SM} reduces to $SU(3)_C \times U(1)_{\mathrm{EM}}$. Any symmetry breaking (with possible intermediate steps) leading to this group $G \to SU(3)_C \times U(1)_{\mathrm{EM}}$ has, if the GUT group G is connected and simply connected, the second homotopy group

$$\pi_2\left(G/[SU(3)_C \times U(1)_{\mathrm{EM}}]\right) \cong \pi_1(SU(3) \times U(1)) \cong \mathbb{Z}\,. \tag{2.51}$$

Therefore, we would expect in this case the production of monopoles at some point in the early Universe. All the typical GUT groups are connected and simply connected. These include, for example, the special unitary groups $SU(5), SU(6)$, spin groups such as $\mathrm{Spin}(10)$, and the exceptional Lie groups such as E_6. These all predict the production of monopoles. Importantly, the electroweak symmetry breaking does not give rise to topologically stable magnetic monopoles since $SU(2)_L \times U(1)_Y$ is not simply connected and, in fact, $\pi_2([SU(2) \times U(1)]/U(1)) \cong \{\,\mathbb{1}\,\}$.

Let us give a rough order-of-magnitude estimate of the monopole energy density today. Consider monopoles of mass m_M of which we expect ~ 1 monopole per Hubble volume d_H^3, i.e., a number density at production of $n_M \sim d_H^{-3}$. The symmetry breaking leading to the production of monopoles will occur deep in the radiation-dominated era at a temperature T_G, such that we can approximate the horizon length scale as $d_H = H^{-1} \sim M_{\mathrm{Pl}}/T_G^2$. Assuming that the number of monopoles does not drastically change, we find for the monopole energy density today

$$\rho_{\mathrm{M}}^0 \sim m_{\mathrm{M}} \left(\frac{T_G}{M_{\mathrm{Pl}}} \right)^3 T_0^3 \sim 10^{14} \rho_{\mathrm{crit}} \left(\frac{m_{\mathrm{M}}}{10^{15}\,\mathrm{GeV}} \right) \left(\frac{T_G}{10^{15}\,\mathrm{GeV}} \right)^3 , \tag{2.52}$$

where we used the present-day values $\rho_{\text{crit}} \sim 10^{-47}\,(\mathrm{GeV})^4$ for the critical density of the Universe, $T_0 \sim 10^{-13}\,\mathrm{GeV}$ for the CMB temperature, and $M_{\mathrm{Pl}} \sim 10^{18}\,\mathrm{GeV}$ for the reduced Planck mass. This estimate shows that GUT-scale monopoles would overclose the Universe by many orders of magnitude.

Since almost all GUTs predict monopoles, this is a problem known as the *monopole problem*, and it was originally noticed in Refs. [37, 38].[9] The above estimate obviously neglects the fact that the monopole energy density can be reduced by monopole–antimonopole annihilations. This mechanism has, however, been shown to be inefficient to resolve the monopole problem [192]. Apart from the trivial solution, namely the absence of grand unification, many possible ways to avoid the problem have been proposed, for example, based on connecting monopoles and antimonopoles by flux tubes leading to highly efficient monopole annihilation [193], symmetry non-restoration at high temperatures [194, 195], unstable domain walls sweeping away monopoles [196], or primordial black holes eating up monopoles and subsequently evaporating into Hawking radiation [197]. The most popular mechanism, though, remains cosmic inflation, which dilutes the monopole number density to an unmeasurably small value by introducing a period of nearly exponential expansion of the Universe [39, 40]. As a solution to the monopole problem, inflation must occur after the GUT symmetry breaking that gives rise to monopoles. Furthermore, it is required that the reheating temperature remains below the critical temperature at which the GUT monopoles form.

Similar to the case of strings, not only local but also global monopoles can form in phase transitions. In the absence of gauge fields, the monopoles have very different properties from those of local monopoles. In particular, they do not carry magnetic charges and are, hence, not related to Dirac's magnetic monopoles. Moreover, the energy of a single monopole is divergent. Global monopoles are mainly of interest due to the gravitational field they exert [198], by which they leave imprints on the CMB and provide seeds for structure formation. For related studies, see, e.g., Refs. [199–201], in particular Ref. [202] for recent CMB bounds and Ref. [203] for supermassive black hole formation. We will not further discuss global monopoles in this thesis and refer the interested reader to the mentioned literature (see also Ref. [36]).

2.3 Metastable defects

Given our previous discussion of topologically stable defects, we can now introduce metastable defects. For this, we just need to consider the implications of having more than one spontaneous symmetry breaking. To start with, consider a symmetry breaking chain of the form

$$G \xrightarrow{\Phi} H \xrightarrow{\Psi} \{\, \mathbb{1} \,\} , \tag{2.53}$$

where G is path-connected and simply connected but H is disconnected. We would then expect domain walls from the second phase transition, but none from the overall phase transition. The

---

[9] The monopole abundance is not only constrained by energy density considerations. A summary of recent bounds can be found in Ref. [191].

naive expectation tells us that, if we have a huge scale separation between the two symmetry breakings, then a low-energy effective theory should give rise to domain walls that appear to us as stable; they should not just completely disappear because we embed our theory into a larger symmetry group.

To understand what actually happens, let us now be more detailed and start with the situation after the first phase transition. Consider a parametrised curve $\boldsymbol{x} : [0, 2\pi] \to \mathbb{R}^3$ in physical space and assume that everywhere along the path, the field has settled to a vacuum value. Furthermore, let us set $\Phi(\boldsymbol{x}(0)) = \Phi_0$ for some vacuum value Φ_0. Recall that we can view the vacuum as the G orbit of Φ_0:

$$\mathcal{M}_1 \cong \{ \, \Phi \mid \exists g \in G : \Phi = \rho_\Phi(g)\Phi_0 \, \} \cong G/H \, . \tag{2.54}$$

This means that we can find a curve in group space $g : [0, 2\pi] \to G$, which describes the values of the field's VEV along the path in physical space: $\Phi(\boldsymbol{x}(\theta)) = \rho_\Phi(g(\theta))\Phi_0$. Suppose that the path $\boldsymbol{x}$ is closed. It thus starts and ends in field space at the value Φ_0. This does, however, not yet mean that $g(0)$ and $g(2\pi)$ must be equal. In fact, any element of the stabiliser subgroup H will do as these act trivially on Φ_0, and the path described by g is only required to be a closed loop in G/H. Explicitly, this is achieved by fixing $g(0), g(2\pi) \in H$.[10] In this discussion, we define two such paths $g, \tilde{g}$ in G to be equivalent if there exists a homotopy between them that projects to a homotopy of closed loops in G/H. Anything else would correspond to a discontinuous deformation of the field configuration. There are now two qualitatively different options to consider. If the path g in G starts and ends in the identity component $H_0 \subset H$, it will be equivalent to another path $\tilde{g}$ that completely lies in H_0 because G is simply connected. Since $\tilde{g}$ projects to the trivial map in G/H, the closed loop corresponding to g in G/H is null-homotopic. If the path g starts, however, in H_0 but ends in a different component H_i, then g cannot be equivalent to $\tilde{g}$ as it must run in parts through a region of G outside of H. Any homotopy between them must necessarily move the endpoint of the path out of H such that it does not project to a homotopy of closed loops in G/H. The closed path associated with g in G/H is, therefore, not null-homotopic and describes a string configuration. With this argument, it is easy to see that for each non-trivial element of $\pi_0(H)$ there is a unique element in $\pi_1(G/H)$, which is not null-homotopic. We have thus demonstrated that

$$\pi_1(G/H) \cong \pi_0(H) \, , \tag{2.55}$$

and the first symmetry breaking must stabilise strings since $\pi_0(H) \ncong \{ \, \mathbb{1} \, \}$. This is an interesting result on its own as it allows us to find out whether a symmetry breaking leads to cosmic strings, by just considering the connectedness of the unbroken subgroup instead of considering the much more complicated question of whether the vacuum is simply connected or not.

So far, we have just considered a single symmetry breaking as before. Let us now turn to the second symmetry breaking. We see that H has not only the meaning of the stabiliser subgroup with respect to Φ_0, but is also the vacuum manifold of the second phase transition $\mathcal{M}_2 \cong H$. According to Eq. (2.55), the criterion that guarantees the stabilisation of strings in the first phase transition guarantees now the stabilisation of domain walls in the second one. The statement is actually even stronger in the sense that to each topological sector of strings, there is a corresponding sector of topologically equivalent wall configurations. Let us see in detail what happens after the first phase transition. In the second symmetry breaking, the Higgs field Ψ acquires a non-vanishing VEV. Suppose the field takes at the beginning of the path the VEV $\Psi(\boldsymbol{x} = 0) = \Psi_0$. The path around a string from the first phase transition is described by a path in G that starts and ends on

---

[10] Since the VEV of Φ changes along the path, the embedding of H in G will also change. Therefore, we actually obtain a continuous one-parameter family of subgroups $H(\theta)$, which are, however, isomorphic to H and we omit the label θ.

different connected components of H. Since Ψ_0 is not invariant under the action of any non-trivial element of H, we find that $\Psi(\boldsymbol{x}(2\pi)) = \rho_\Psi(g(2\pi))\Psi_0 \neq \Psi_0 = \Psi(\boldsymbol{x}(0))$: The expectation value of the field seems discontinuous. This is, of course, not possible. Instead, the field must assume a non-vacuum value in between the two disconnected vacua: A domain wall that ends on a string forms. Neither domain walls nor strings are stable since the total vacuum $\mathcal{M} \cong G$ is connected and simply connected. Metastable defects are formed, namely walls bounded by strings [204, 205].

Note that if we replace the symmetry breaking of this example with the more general case

$$G \longrightarrow H \longrightarrow K\,, \tag{2.56}$$

then the above argument is still true, with the only difference that H contains non-trivial elements that leave Ψ_0 invariant, exactly those contained in K. The elements that do not leave Ψ_0 invariant are, by definition, those of $\mathcal{M}_2 \cong H/K$. We can state that each string corresponding to a path in G whose endpoints lie in different connected components of H/K will become attached to a domain wall. But this is not necessarily every string. While in the example with $K = \{\, \mathbb{1} \,\}$, there were no stable strings since G was simply-connected, this time, the vacuum is $\mathcal{M} \cong G/K$ and stable strings are labelled by the fundamental group $\pi_1(G/K) \cong \pi_0(K)$. If K is not path-connected, then topologically stable strings will remain that do not become attached to domain walls.

2.3.1 Classification of metastable defects

For the classification of metastable defects, we consider the general symmetry-breaking scheme of Eq. (2.56). Our discussion is based on Ref. [33]. What we mean by a metastable defect is one that would be topologically stable if there were only the low-energy theory with the symmetry breaking $H \to K$, but which is no longer topologically stable if we embed our model in a theory with a larger symmetry group G. In other words, metastable defects are those stabilised by the vacuum $\mathcal{M}_2 \cong H/K$ but not by the full vacuum $\mathcal{M} \cong G/K$. In the following, we will also consistently denote the vacuum from the first phase transition as $\mathcal{M}_1 \cong G/H$.

Metastable domain walls: Due to the inclusion $\mathcal{M}_2 \cong H/K \hookrightarrow G/K \cong \mathcal{M}$, there is an induced map $p_0 : \pi_0(\mathcal{M}_2) \to \pi_0(\mathcal{M})$, which is a homomorphism if the two sets can be equipped with a group structure as discussed before. Since $\pi_0(\mathcal{M}_2)$ labels the domain walls arising from the second phase transition and $\pi_0(\mathcal{M})$ the ones that are stable in the overall symmetry breaking, the metastable domain walls — those which are stabilised by $\mathcal{M}_2$ but not by $\mathcal{M}$ — are the elements of the preimage under p_0 of the homotopy class $[c] \in \pi_0\,(\mathcal{M})$ which contains the constant map c, i.e., $p_0^{-1}\,(\{\,[c]\,\}) \subset \pi_0(\mathcal{M}_2)$. If p_0 is a homomorphism, then the preimage is just the kernel $\ker\{\, p_0 \,\} \subset \pi_0(\mathcal{M}_2)$. Elements of this preimage other than the homotopy class of the constant map (non-trivial elements of this kernel) correspond to maps $S^0 \to \mathcal{M}_2$ that are not null-homotopic in $\mathcal{M}_2$ but correspond to null-homotopic maps in $\mathcal{M}$. Therefore, there is a path in $\mathcal{M}$ that starts and ends in different connected components of $\mathcal{M}_2$. Every such map describes, up to homotopy, a non-null-homotopic map $S^1 \to \mathcal{M}/\mathcal{M}_2 \cong G/H \cong \mathcal{M}_1$ and, therefore, a representative of a non-trivial element of $\pi_1(\mathcal{M}_1)$. But these correspond just to the strings of the first symmetry breaking. We have recovered the result of the previous discussion: To each metastable domain wall produced in the second symmetry breaking, there is a string solution of the first symmetry breaking on which the domain wall can end.

Metastable strings: Similarly to the above discussion, there is an induced homomorphism $p_1 : \pi_1(\mathcal{M}_2) \to \pi_1(\mathcal{M})$. We know that $\pi_1(\mathcal{M}_2)$ classifies the cosmic strings stabilised by the second symmetry breaking, and to the different elements of $\pi_1(\mathcal{M})$ correspond topologically distinct strings of the full theory. Strings that are stabilised in the second symmetry breaking but not in the full theory correspond accordingly to the non-trivial elements of the kernel

$\ker\{p_1\} \subset \pi_1(\mathcal{M}_2)$. The non-trivial elements of this kernel are homotopy classes of maps $S^1 \to \mathcal{M}_2$ that cannot be continuously shrunk to a point in $\mathcal{M}_2$ but allow for homotopies that shrink them to the constant map in $\mathcal{M}$. Considering instead the quotient space $\mathcal{M}/\mathcal{M}_2$, these homotopies correspond to non null-homotopic maps $S^2 \to \mathcal{M}/\mathcal{M}_2 \cong \mathcal{M}_1$, thus representing non-trivial elements of $\pi_2(\mathcal{M}_1)$. Recall now that these non-trivial elements classify the monopoles of the first phase transition. The corresponding physical picture is that to each metastable string configuration arising from the second symmetry breaking, there exists a monopole configuration in the first symmetry breaking on which the string can end [206].

No metastable monopoles: In analogy to the above discussion, metastable monopoles correspond to the non-trivial elements of the kernel of the homomorphism $p_2 : \pi_2(\mathcal{M}_2) \to \pi_2(\mathcal{M})$. If H is a finite-dimensional Lie group, this map is injective, and its kernel is trivial. There are no metastable monopoles.

This concludes our discussion of the classification of metastable defects. For more details, we refer the reader to Refs. [33, 205].

2.3.2 Metastable strings — field configurations

Metastable strings are a particularly interesting class of metastable defects and have been studied in many particle physics models.[11] They can fragment into segments as a result of the nucleation of monopole–antimonopole pairs due to quantum tunnelling events. The nucleation rate per unit string length has been estimated using a semiclassical approximation, which assumes zero thickness for the string and monopole [33, 211]:

$$\Gamma \simeq \frac{\mu}{2\pi} e^{-\pi \frac{m_{\mathrm{M}}^2}{\mu}} , \tag{2.57}$$

with monopole mass m_{M} and string tension μ. This approximation works well for a large symmetry-breaking scale hierarchy. If this hierarchy is too large, $m_{\mathrm{M}} \gg \sqrt{\mu}$, the decay rate is exponentially suppressed, rendering the stings quasi-stable. For all practical purposes, they can then be treated as stable defects. In the absence of a large scale separation, however, the assumptions underlying the above estimate become less reliable. Nevertheless, a small separation will still lead to a high nucleation rate and rapid string decay. Consequently, the intermediate regime, where the scale separation is neither very large nor very small, is of particular interest. Considering a network of strings,[12] this decay process yields, over time, a population of short string segments, each terminating on a monopole at one end and on an antimonopole at the other. Two monopoles attached to one string will accelerate towards each other and annihilate, thereby erasing the string. Such metastable string networks can give rise to characteristic GWB signatures. These have been studied in Refs. [212–214] and recent PTA observations find good agreement with $m_{\mathrm{M}}/\sqrt{\mu} \sim 8$ [1] (see also Ch. 7). In this intermediate regime, the estimate of Eq. (2.57) may not hold, and a computation that takes the actual metastable string configurations into account is required. This highlights the importance of determining the actual field configurations of metastable strings, rather than approximating them as infinitely thin.

In Ref. [215] and more recently in Ref. [216], this was achieved for a simple model of metastable strings. There, one starts with the local symmetry breaking $SU(2) \xrightarrow{\Phi} U(1)$ of

---

[11] See, e.g., Refs. [35, 175, 207–209]. For a recent review of metastable strings with a particular focus on their GW signatures, see Ref. [210].

[12] This assumes that we start out with a network of strings without any monopoles. The monopoles from the first symmetry breaking must be diluted by an inflationary period between the two stages of symmetry breaking. Otherwise, strings would connect monopoles to antimonopoles and quickly erase the defects.

Ex. 1.2 by means of a Higgs triplet Φ. This symmetry breaking gives rise to 't Hooft–Polyakov monopoles, as discussed in Sec. 2.2.4. By extending the 't Hooft–Polyakov model by another Higgs field Ψ in the fundamental representation of $SU(2)$, the remaining $U(1)$ symmetry can be broken at a lower scale: $SU(2) \overset{\Phi}{\to} U(1) \overset{\Psi}{\to} \{\,\mathbb{1}\,\}$. The first symmetry breaking will lead to the formation of monopoles, the second symmetry breaking leads to strings. The overall vacuum $\mathcal{M} \cong SU(2) \cong S^3$, however, does not stabilise either of the defects. From the previously introduced formalism, we know that in this model, each monopole becomes attached to a string and vice versa. Since there is no remaining symmetry, there can also be no unconfined flux. The entire monopole flux becomes confined to the string. Due to an inflationary phase after the first phase transition, the second phase transition happens effectively in the absence of monopoles, and long strings can form. These strings can then unwind in the larger simply connected vacuum $\mathcal{M}$ — the process associated with the nucleation of a monopole–antimonopole pair on the string.

Model definition: We want to consider a slightly more complex model with the following symmetry-breaking pattern:[13,14]

$$SU(2)_P \times U(1)_K \xrightarrow{\Phi} U(1)_P \times U(1)_K \xrightarrow{\Psi} U(1)_Y\,. \tag{2.58}$$

The intermediate vacuum manifolds have non-trivial homotopy groups $\pi_2(\mathcal{M}_1) \cong \mathbb{Z}$ and $\pi_1(\mathcal{M}_2) \cong \mathbb{Z}$, so the first phase transition stabilises monopoles while the second stabilises strings. The overall vacuum $\mathcal{M}$ is of the same form as in the electroweak case and topologically stabilises neither strings nor monopoles. In contrast to the model discussed in Ref. [215], our model leaves an unbroken $U(1)$ subgroup such that the metastable monopoles can, in addition to the confined flux in the strings, also carry an unconfined flux after the second phase transition.

The symmetry breaking we want to consider was used in a different context in Ref. [2]. Following this reference, we can achieve the above symmetry breaking with a Higgs field Φ in the 3_0 representation that breaks the first symmetry and a Higgs field Ψ in the $2_{1/2}$ representation, which breaks the second symmetry. We consider Φ in matrix form as in Ex. 1.2.

Let us define the model more explicitly. We call the gauge fields associated with the $SU(2)_P$ and $U(1)_K$ symmetries $P_\mu = P_\mu^a T^a$ and K_μ, respectively. We denote their field strengths by

$$P_{\mu\nu} = 2\partial_{[\mu}P_{\nu]} - ig_P[P_\mu, P_\nu] \qquad \text{and} \qquad K_{\mu\nu} = 2\partial_{[\mu}K_{\nu]}\,, \tag{2.59}$$

and the gauge covariant derivatives by

$$\mathcal{D}_\mu\Phi = \partial_\mu\Phi - ig_P[P_\mu, \Phi] \qquad \text{and} \qquad \mathcal{D}_\mu\Psi = \partial_\mu\Psi - ig_P P_\mu\Psi - \frac{ig_K}{2}K_\mu\Psi\,. \tag{2.60}$$

The model is described by the Lagrangian density

$$\mathcal{L} = -\frac{1}{2}\mathrm{tr}\left(P_{\mu\nu}P^{\mu\nu}\right) - \frac{1}{4}K_{\mu\nu}K^{\mu\nu} + \mathrm{tr}\left[\left(\mathcal{D}_\mu\Phi\right)\left(\mathcal{D}^\mu\Phi\right)\right] + \left(\mathcal{D}_\mu\Psi\right)^\dagger\left(\mathcal{D}^\mu\Psi\right) - V\left(\Phi, \Psi\right)\,, \tag{2.61}$$

---

[13] After this chapter was completed, Ref. [217] appeared on arXiv. In that work, the authors consider a symmetry breaking chain similar to the one discussed here, with the difference that they extend it by an additional $SU(2)$ singlet that is charged under the original $U(1)$ group. This field is used in a further step to break the $U(1)$ group, which in our case remains unbroken. Their analysis is largely complementary to ours, as it focuses on the field profiles of the metastable strings from the common second symmetry breaking and of the stable strings from their third symmetry breaking. In contrast, our discussion focuses on the unwinding of the metastable string within the larger symmetry group, which is not studied in Ref. [217].

[14] Here, $U(1)_Y$ does not have to refer to the symmetry group generated by weak hypercharges, although it can, and we will make this identification in Ch. 3.

where the potential consists of three terms $V(\Phi,\Psi) = V_1(\Phi) + V_2(\Psi) + V_3(\Phi,\Psi)$. These read

$$V_1(\Phi) = \lambda_\Phi \left(\mathrm{tr}\left(\Phi^2\right) - \frac{v_\Phi^2}{2} \right)^2 , \tag{2.62}$$

$$V_2(\Psi) = \lambda_\Psi \left(\Psi^\dagger \Psi - \eta_\Psi^2 \right)^2 , \tag{2.63}$$

$$V_3(\Phi,\Psi) = M \Psi^\dagger \Phi \Psi . \tag{2.64}$$

To study the defects from the first and second phase transition separately, let us first restrict to a large scale separation $v_\Phi \gg \eta_\Psi$. This can be relaxed once we consider the full metastable defect configurations, although we require Φ to acquire its non-vanishing VEV in the first phase transition and Ψ in the second. For later purposes, we write down the gauge transformations explicitly. Denoting $U_P(x) = \exp(i\pi^a(x)T^a)$ and $U_K(x) = \exp\left(\frac{i}{2}\kappa(x)\right)$, we have

$$P_\mu \longmapsto U_P P_\mu U_P^{-1} + \frac{i}{g_P} U_P \partial_\mu U_P^{-1} , \qquad K_\mu \longmapsto K_\mu + \frac{1}{g_K}\partial_\mu \kappa , \tag{2.65}$$

$$\Psi \longmapsto U_P U_K \Psi , \qquad \Phi \longmapsto U_P \Phi U_P^{-1} . \tag{2.66}$$

Spontaneous symmetry breaking: Let us briefly go through the steps involved in the symmetry breaking. In the first stage, we consider the field Φ only and neglect effects due to Ψ because of the scale separation $v_\Phi \gg \eta_\Psi$. The field Φ acquires a non-vanishing VEV, and we may locally choose this VEV to align with the T^3-direction, i.e., $\Phi_0 = v_\Phi T^3$. This leaves one unbroken $SU(2)_P$ generator associated with rotations around the T^3-axis and leaves the $U(1)_K$ symmetry untouched as Φ is not charged under this group. The unbroken subgroup is therefore $U(1)_P \times U(1)_K$. Accordingly, the fields $P_\mu^{1,2}$ obtain an effective mass $m_P^{(1)} = g_P v_\Phi$ while P_μ^3 and K_μ remain the massless gauge bosons of the unbroken subgroup. The superscript index "(1)" on the mass indicates that this is the mass after the first stage of symmetry breaking.

Let us move on to the second symmetry-breaking stage in which the Higgs doublet Ψ acquires a non-vanishing VEV. Using again the large scale separation, we can treat Φ as a fixed background field, neglecting its dynamics. Inserting the VEV Φ_0 into the potential, the Higgs doublet Ψ will effectively experience a potential of the form

$$\tilde{V}(\Psi) = V_2(\Psi) + V_3(\Phi_0,\Psi) = \lambda_\Psi \left(|\psi_1|^2 + |\psi_2|^2 \right)^2 + \mu_1 |\psi_1|^2 - \mu_2 |\psi_2|^2 + \lambda_\Psi \eta_\Psi^4 . \tag{2.67}$$

For convenience, we introduced

$$\mu_1 = \frac{M v_\Phi}{2} - 2\lambda_\Psi \eta_\Psi^2 \qquad \text{and} \qquad \mu_2 = \frac{M v_\Phi}{2} + 2\lambda_\Psi \eta_\Psi^2 . \tag{2.68}$$

We consider the part of the parameter space for which $\mu_{1,2} > 0$.

The minimum of the effective potential that one needs to consider after the first symmetry breaking satisfies

$$|\psi_1|^2 = 0 \qquad \text{and} \qquad |\psi_2|^2 = \frac{\mu_2}{2\lambda_\Psi} \equiv v_\Psi^2 , \tag{2.69}$$

and we can choose the remaining gauge freedom to make ψ_2 real, i.e.,

$$\Psi_0 = \begin{pmatrix} 0 \\ v_\Psi \end{pmatrix} . \tag{2.70}$$

Under the subgroup $U(1)_P \times U(1)_K$, the field Ψ transforms as

$$\Psi = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix} \mapsto \begin{pmatrix} e^{\frac{i}{2}\left(\pi^3(x)+\kappa(x)\right)} \psi_1(x) \\ e^{\frac{i}{2}\left(-\pi^3(x)+\kappa(x)\right)} \psi_2(x) \end{pmatrix} . \tag{2.71}$$

Transformations affecting ψ_1, but leaving the ψ_2 component untouched, are still symmetries after the second spontaneous symmetry breaking. From Eq. (2.71), we can immediately see that transformations for which the relation $\pi^3(x) = \kappa(x)$ holds satisfy this criterion and lead to an unbroken $U(1)$ subgroup.

To express the gauge field associated with this remaining $U(1)_Y$ group in terms of the gauge fields of the full theory, it is useful to apply a rotation in the space of gauge fields:

$$\begin{pmatrix} X_\mu \\ Y_\mu \end{pmatrix} = \frac{1}{\tilde{g}} \begin{pmatrix} g_P & -g_K \\ g_K & g_P \end{pmatrix} \begin{pmatrix} P^3_\mu \\ K_\mu \end{pmatrix} . \tag{2.72}$$

Here, we have defined the gauge coupling $\tilde{g} = \sqrt{g_P^2 + g_K^2}$. Rewritten in this form, ψ_2 couples to X_μ only, and not to Y_μ. Inserting the VEV of Ψ into the kinetic term of this Higgs doublet Ψ, we find that the gauge field masses take the form

$$m_P^{(2)} = g_P \sqrt{v_\Phi^2 + \frac{v_\Psi^2}{2}} , \qquad\qquad m_X^{(2)} = \frac{\tilde{g}}{\sqrt{2}} v_\Psi . \tag{2.73}$$

Only the gauge field Y_μ remains massless and is associated with the remaining unbroken $U(1)_Y$ symmetry.

Defect configurations: We are now prepared to discuss the defect configurations. The first symmetry breaking is, due to the large scale separation, effectively the same as that considered in the case of ’t Hooft–Polyakov monopoles in Sec. 2.2.4. We can match these two models and find from Eq. (2.39) the configurations

$$\phi^a = v_\Phi f(r) \hat{\boldsymbol{x}}^a , \qquad P_0 = 0 , \qquad P_i^a = \frac{1}{g_P} \epsilon_{aij} \hat{\boldsymbol{x}}^j \frac{1 - g(r)}{r} , \tag{2.74}$$

with the same boundary conditions on the profile functions f and g as in Eqs. (2.40) and (2.41). The magnetic flux associated with the unbroken $SU(2)_P$ gauge field through a closed surface around the monopole can be read off from Eq. (2.45) as

$$\boldsymbol{\Phi}_P^{\text{monopole}} = \frac{4\pi}{g_P} . \tag{2.75}$$

To describe strings, it is useful to explicitly obtain an effective Lagrangian for the second symmetry breaking. This symmetry breaking is governed by the field ψ_2 and the gauge field X_μ with field strength $X_{\mu\nu}$. Introducing the covariant derivative

$$\tilde{\mathcal{D}}_\mu \psi_2 = \partial_\mu \psi_2 + i \frac{\tilde{g}}{2} X_\mu \psi_2 , \tag{2.76}$$

the symmetry breaking can be effectively described by the Abelian Higgs model of Sec. 1.1.1

$$\tilde{\mathcal{L}} = -\frac{1}{4} X_{\mu\nu} X^{\mu\nu} + \left(\tilde{\mathcal{D}}_\mu \psi_2\right)^* \left(\tilde{\mathcal{D}}^\mu \psi_2\right) - \lambda_\Psi \left|\psi_2\right|^4 + \mu_2 \left|\psi_2\right|^2 . \tag{2.77}$$

The $U(1)_Y$ gauge invariance requires that if $X_\mu \mapsto X_\mu + 1/\tilde{g}\, \partial_\mu \gamma$, then $\psi_2 \mapsto e^{-i\gamma/2} \psi_2$. In Sec. 2.2.3, we have seen that this model gives rise to ANO strings described by configurations of the form given in Eq. (2.20), and we can match the two models to find an analogous configuration:

$$\psi_1 = 0, \quad \psi_2(\rho, \varphi) = v_\Psi e^{in\varphi} h(\rho), \quad X_0 = X_\rho = X_z = 0, \quad X_\varphi(\rho) = -\frac{2n}{\tilde{g}} (1 - k(\rho)) . \tag{2.78}$$

The boundary conditions on the profile functions h and k are the same as in Eqs. (2.21) and (2.22). The string contains a confined magnetic flux associated with the gauge field X_μ and we can find it directly from Eq. (2.19) as

$$\mathbf{\Phi}_X^{\text{string}} = \frac{4\pi n}{\tilde{g}} , \tag{2.79}$$

with the magnetic field aligned with the unit vector $\hat{\boldsymbol{z}}$. The sign in front of the flux is reversed in comparison to Eq. (2.19) due to the matching $e = -\tilde{g}/2$.

Let us now see, how strings attach to monopoles after the second symmetry breaking. To this end, consider a sphere of radius R centred at the core of the monopole $r = 0$. The sphere can be covered by two coordinate charts, one covering the northern hemisphere, the other covering the southern hemisphere

$$\begin{aligned} \mathcal{U}_N &= \{ \, (\theta, \varphi) \mid \varphi \in [0, 2\pi); \, \theta \in [0, \pi/2 + \epsilon) \, \} \ , \\ \mathcal{U}_S &= \{ \, (\theta, \varphi) \mid \varphi \in [0, 2\pi); \, \theta \in (\pi/2 - \epsilon, \pi] \, \} \ . \end{aligned} \tag{2.80}$$

With $SU(2)_P$ gauge transformations of the form

$$U_N = \begin{pmatrix} \cos(\theta/2) & e^{-i\varphi}\sin(\theta/2) \\ -e^{i\varphi}\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} , \quad U_S = \begin{pmatrix} e^{i\varphi}\cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & e^{-i\varphi}\cos(\theta/2) \end{pmatrix} , \tag{2.81}$$

the monopole configuration in Eq. (2.74) can be brought into the trivial form

$$\Phi_{N,S} = U_{N,S} \Phi U_{N,S}^\dagger = v_\Phi f(r) T^3 , \tag{2.82}$$

on the respective hemisphere. Of course, it is impossible to write the field globally in this way. In the following, we will consider the limit $R \to \infty$. For the gauge fields, using differential form notation for brevity, we have then

$$P_N = \frac{1}{g_P} \left(\cos(\theta) - 1\right) T^3 \mathrm{d}\varphi \qquad \text{and} \qquad P_S = \frac{1}{g_P} \left(\cos(\theta) + 1\right) T^3 \mathrm{d}\varphi , \tag{2.83}$$

such that on the intersection of the two charts $\mathcal{U}_N \cap \mathcal{U}_S$, the gauge fields are related via

$$P_S^3 = P_N^3 + \frac{2}{g_P} \mathrm{d}\varphi . \tag{2.84}$$

Using the relation given in Eq. (2.72), this implies

$$X_S = X_N + \frac{2}{\tilde{g}} \mathrm{d}\varphi . \tag{2.85}$$

Moreover, we can utilise Eq. (2.72) to find that $X_{\mu\nu} = g_P/\tilde{g}\, \mathcal{P}_{\mu\nu} - g_K/\tilde{g}\, K_{\mu\nu}$ and $Y_{\mu\nu} = g_K/\tilde{g}\, \mathcal{P}_{\mu\nu} + g_P/\tilde{g}\, K_{\mu\nu}$, where $\mathcal{P}_{\mu\nu}$ is the $U(1)$ field strength corresponding to the gauge field P_μ after the first symmetry breaking as defined in Eq. (2.43). Since the fields associated with the monopole are in the trivial representation of $U(1)_K$, we can assume $K_{\mu\nu} = 0$, and the radially outgoing X-magnetic and Y-magnetic fluxes through a sphere centred around the monopole derive from Eq. (2.75) as

$$\mathbf{\Phi}_X^{\text{monopole}} = \frac{4\pi}{\tilde{g}} \qquad \text{and} \qquad \mathbf{\Phi}_Y^{\text{monopole}} = \frac{4\pi}{\tilde{g}} \frac{g_K}{g_P} . \tag{2.86}$$

Suppose now that the second symmetry breaking happens in the northern hemisphere in such a way that $\Psi_0 = (0, v_\Psi)^T$ everywhere. Then the gauge field will acquire a mass and decay exponentially over length scales of $\mathcal{O}(m_X^{-1})$. Therefore, in the northern hemisphere, $X_N = 0$. This implies that on the southern hemisphere, the configuration takes the form

$$X_S = \frac{2}{\tilde{g}}\mathrm{d}\varphi\,, \qquad \psi_2^S = v_\Psi e^{-i\varphi}\,. \tag{2.87}$$

This is clearly the asymptotic form of the string configuration given in Eq. (2.78) if we set $n = -1$. The presence of a monopole imposes the presence of a string attached to it after the second phase transition. Note that this is completely analogous to the situation of metastable domain walls, which we described in the introduction to Sec. 2.3 in terms of homotopy groups. There, the existence of strings requires that, after the second phase transition, domain walls attach to them.

Let us now see what happens to the magnetic flux. Using the result in Eq. (2.79), we find that the X-magnetic flux opposite to the $\hat{\boldsymbol{z}}$-direction which is enclosed in the string reads

$$\boldsymbol{\Phi}_X^{\mathrm{string}} = \frac{4\pi}{\tilde{g}}\,. \tag{2.88}$$

By comparison with Eq. (2.86), this is just the X-magnetic flux outgoing from the monopole, which becomes confined in the string. The Y-magnetic flux of the monopole associated with the unbroken $U(1)_Y$ symmetry remains, however, unconfined.[15]

We can now try to find an unwinding ansatz proceeding in a similar way as in Ref. [215]. The string we want to consider winds in the subgroup

$$U(1)_X \cong \left\{ \left(e^{i\gamma T^3}, e^{-i\gamma}\right) \,\middle|\, \gamma \in [0, 2\pi) \right\} \subset SU(2)_P \times U(1)_K\,. \tag{2.89}$$

Due to the associated winding in the multiply connected $U(1)_K$, the string cannot be unwound in the full group $SU(2)_P \times U(1)_K$. This is not a problem as we only require unwinding in the vacuum manifold $[SU(2)_P \times U(1)_K]/U(1)_Y$. The task reduces, therefore, to finding a homotopy in $SU(2)_P \times U(1)_K$ between a loop that winds once around $U(1)_X$ and a loop that only winds around $U(1)_Y$. A homotopy that achieves this aim is

$$\mathcal{H}(s,\varphi):\ [0,1]\times[0,2\pi) \to SU(2)_P \times U(1)_K,\ (s,\varphi) \mapsto \left(U_P(s,\varphi), e^{-i\varphi}\right)\,, \tag{2.90}$$

where

$$U_P(s,\varphi) = e^{i\varphi T^3}\cos\left(\frac{\pi}{2}s\right) + e^{i\pi T^2}e^{-i\varphi T^3}\sin\left(\frac{\pi}{2}s\right)\,. \tag{2.91}$$

One may verify that $U_P(s,\varphi)$ always assumes values in $SU(2)_P$. Using the respective representations of the fields, this allows us to write down the following unwinding ansatz:

$$\Phi_{(s)} = U_P(s,\varphi)v_\Phi T^3 U_P^\dagger(s,\varphi) + \tilde{\Phi}_{(s)}\,, \quad \Psi_{(s)} = U_P(s,\varphi)e^{-\frac{i\varphi}{2}}\begin{pmatrix} 0 \\ v_\Psi h_s(\rho)\end{pmatrix}\,, \tag{2.92}$$

$$P_{(s),\varphi} = \frac{i}{g_P}U_P(s,\varphi)\partial_\varphi U_P^\dagger(s,\varphi)(1-k_s(\rho))\,, \qquad K_{(s),\varphi} = -\frac{1}{g_K}(1-k_s(\rho))\,, \tag{2.93}$$

where all other field components vanish, and we introduced

$$\tilde{\Phi}_{(s)}(\rho,\varphi) = v_\Phi p_s(\rho)\sin(\pi s)\left(\cos(\varphi)T^1 - \sin(\varphi)T^2\right)\,. \tag{2.94}$$

Note that the negative sign appearing in K_φ accounts for the fact that a negative winding in K_μ corresponds to a positive winding in X_μ.

---

[15] In a similar fashion, we can also say that the flux $\boldsymbol{\Phi}_P$ linearly decomposes into a confined and an unconfined contribution $\boldsymbol{\Phi}_P^{\mathrm{unconf.}} = \frac{g_P^2}{\tilde{g}^2}\boldsymbol{\Phi}_P$ and $\boldsymbol{\Phi}_P^{\mathrm{conf.}} = \frac{g_K^2}{\tilde{g}^2}\boldsymbol{\Phi}_P$.

Observe that

$$\Psi_{(0)} = \begin{pmatrix} 0 \\ v_\Psi e^{-i\varphi} h_0(\rho) \end{pmatrix} \qquad \text{and} \qquad \Psi_{(1)} = \begin{pmatrix} v_\Psi h_1(\rho) \\ 0 \end{pmatrix}\,, \tag{2.95}$$

which indicates that, for $s = 0$, there is a string with winding $n = -1$ in $U(1)_X$ and, for $s = 1$, there is no more winding in $U(1)_X$. Furthermore, we have

$$\Phi_{(0)} = v_\Phi T^3 \qquad \text{and} \qquad \Phi_{(1)} = -v_\Phi T^3\,, \tag{2.96}$$

with both configurations corresponding to a constant vacuum. The change in the sign of Φ between $s = 0$ and $s = 1$ indicates the presence of a monopole. The above homotopy continuously interpolates between the string solution and the trivial vacuum. Therefore, it describes an unwinding of a metastable string.

Let us comment on $\Phi_{(s)}$, for which we introduced the additional term $\tilde{\Phi}_{(s)}$. This is necessary to avoid multivaluedness at $\rho = 0$ for $s \neq 0, 1$ (cf. Ref. [215]), which can be seen by explicitly writing $\Phi_{(s)}$ out:

$$\Phi_{(s)} = v_\Phi \left[\cos(\pi s) T^3 + \sin(\pi s) \left(\sin(\varphi) T^2 - \cos(\varphi) T^1\right)\right] + \tilde{\Phi}_{(s)}\,. \tag{2.97}$$

From Eq. (2.94) it is clear that regularity at $\rho = 0$ requires $p_s(0) = 1$. In particular, we cannot choose $\tilde{\Phi}_{(s)}$ to vanish. Naturally, $\Phi_{(s)}$ must also approach the vacuum infinitely far away from the string core and we need to impose the condition $\lim_{\rho\to\infty} p_s(\rho) = 0$.

We also need to fix the boundary conditions on the profile functions. For asymptotically large ρ, we require that the field settles to the vacuum and at $\rho = 0$, we require regularity, leading, as usual, to

$$h_s(0) = 0\,, \qquad \lim_{\rho\to\infty} h_s(\rho) = 1\,, \tag{2.98}$$

$$k_s(0) = 1\,, \qquad \lim_{\rho\to\infty} k_s(\rho) = 0\,. \tag{2.99}$$

The configuration should not only be viewed as describing the unwinding process via some abstract parameter s, but as an actual unwinding in space and time of a string that extends infinitely in one direction and ends on a monopole in the other direction. For $z \to -\infty$, we just have a string configuration of the ANO form corresponding to $s \to 0$, and for $z \to \infty$, we have the trivial vacuum corresponding to $s \to 1$. In between, the string ends on a monopole. The monopole at the end of the string does, however, not remain at rest but rather accelerates towards the negative z-direction due to the constant string tension. In conclusion, s is a function of z and t. The above ansatz is then inconsistent, as we fixed $P_z = P_t = 0$. To correct for this, one may use the more general ansatz for the gauge fields

$$P_{(s),\mu} = \frac{i}{g_P} U_P(s(t,z),\varphi)\partial_\mu U_P^\dagger(s(t,z),\varphi)(1 - k_{s(t,z)}(\rho))\,, \tag{2.100}$$

notwithstanding the option of other ansätze such as those used in Refs. [215, 216]. The same references also propose an "improved ansatz" with different profile functions for the different gauge fields. Similar adaptations could be used in the present model.

In Refs. [215, 216], the nucleation rate of monopole–antimonopole pairs on a string was computed in a model with fully confined flux. Future studies may extend this calculation to the case of unconfined flux, using the unwinding ansatz proposed above. This is especially interesting in light of GWB observations, as the spectra predicted for metastable strings with confined versus unconfined flux can be quite different [214] (see also Ch. 7).

2.4 Other defects

Before concluding our discussion of the field theory and topology of cosmic defects, we briefly highlight other types of defects that have not yet been covered. This overview is not intended to be exhaustive or serve as a comprehensive guide to the literature. Rather, it aims to convey the fundamental concepts underlying these additional defects. In this spirit, we limit our discussion to simple models and refrain from pursuing generalisations or detailed classifications. It should also be remarked that we do not review cosmic superstrings, as, although they might share some properties with other cosmic strings, they are fundamentally different. We will briefly comment on them in Ch. 7.

Composite defects: We begin with a brief comment on composite or hybrid defects, which are technically not different from the defects discussed in Sec. 2.3. A chain of symmetry breakings can give rise to defects of different dimensions, which can attach to each other. Walls can be bounded by strings and strings by monopoles. When speaking of metastable or transient defects, the emphasis is more often on the decay. Such composite configurations can, however, be effectively stable. A chain of phase transitions can also give rise to both metastable and topologically stable defects. In fact, we already saw a classical example of a composite defect in a model that contains both metastable and stable defects, namely Ex. 2.1. We talked about the symmetry breaking $SU(2) \to \mathbb{Z}_2$ as a one-step breaking involving two Higgs triplets and giving rise to $\mathbb{Z}_2$-strings. The two triplets can, however, acquire their non-vanishing VEVs at largely separated temperature scales, which amounts to a symmetry breaking chain $SU(2) \to U(1) \to \mathbb{Z}_2$. The first symmetry-breaking step is just the same we considered in Ex. 1.4, which gives rise to ’t Hooft–Polyakov monopoles with an unconfined magnetic flux $\mathbf{\Phi} = 4\pi/c$; see Eq. (2.45). In the second symmetry-breaking step, this flux is confined into strings that attach to the monopole. Note that we found in Eq. (2.38) that the $\mathbb{Z}_2$ strings have magnetic flux $\mathbf{\Phi} = 2\pi/c$, up to a sign that can be adjusted by the choice of orientation of the string. Hence, two strings have to attach to one monopole.

Imagine the two strings along a straight line with a monopole between them. On one side of the monopole, the string has winding $+1$; on the other side, it has winding -1. Both configurations are therefore topologically equivalent: Moving the monopole along the string will change the winding by ± 2. Similarly, we could imagine both strings pointing in almost the same direction. If we consider a closed path around this one side of the monopole, we enclose two strings of winding $+1$, giving a total winding $+2$, whereas on the other side of the monopole, there are no strings and the total winding is 0. We can simply eliminate the strings by moving the monopole in the direction of the strings. Strings with winding 2 are thus topologically trivial. Since the monopoles can carry winding numbers in $\mathbb{Z}$ and, correspondingly, higher magnetic fluxes, we can extend the argument to any odd and even winding number: All strings with odd winding are topologically equivalent to the string with winding 1, while all strings with even winding are topologically trivial. In this way, we recover the group structure $\pi_1(\mathcal{M}) \cong \mathbb{Z}_2$ of the full vacuum. This gives a more intuitive explanation as to why this model contains only one class of strings that are topologically distinct from the trivial vacuum. It does, however, not imply that defects consisting of strings and monopoles always decay to the simplest configuration. Instead, one can also find stable configurations on which many monopoles are connected by strings, a configuration dubbed necklace. If one continues this game, one can also break the remaining $\mathbb{Z}_2$ symmetry, which is a subcase of what we discussed at the beginning of Sec. 2.3. This makes all defects metastable and domain walls attach to the strings. A closed necklace could then turn into a domain wall bounded by a necklace.

In more complex models, there is also a greater variety of hybrid defects. A modern overview discussing different composite defects arising in various symmetry-breaking chains from $\mathrm{Spin}(10)$ to the SM gauge group G_{SM} can be found in Ref. [218].

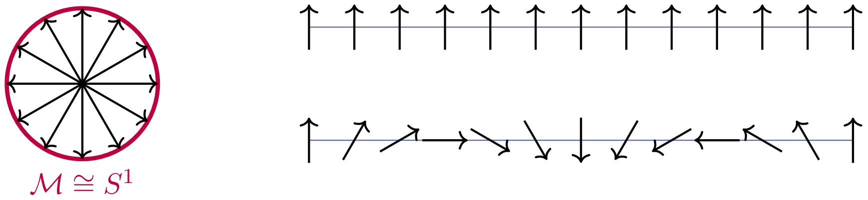


Figure 2.8: Illustration of a trivial configuration and a texture in one spatial dimension. **Left:** Multiply connected vacuum manifold $\mathcal{M} \cong S^1$ in purple, as obtained, e.g., from the model described by Eq. (1.1). Points in $\mathcal{M}$ can be represented by arrows. **Right:** To each point in the one spatial dimension (along the vertical lines) corresponds a field value, which lies, for the texture and the trivial configuration, in $\mathcal{M}$. The arrows associated with the vacuum values are shown along the lines. The configuration on the upper line does not wind and is trivial. The configuration depicted on the lower line winds once around the vacuum. The space can also be viewed as a circle by identifying the right and left ends of the straight lines, in which case the texture is stable.

Textures: In addition to domain walls, strings, and monopoles, there is another possible defect in three spatial dimensions. This remaining class of defects entails so-called *textures* which arise in phase transitions for which the third homotopy group of the vacuum manifold is non-trivial $\pi_3(\mathcal{M}) \ncong \{ \mathbb{1} \}$ [32]. However, they do not fit straightforwardly into the classification scheme laid out in Sec. 2.2.2, as this would mean that these defects have dimension -1. In reality, textures are very different from other defects, as the texture field configuration does not have to leave the vacuum manifold at any point in space. This does not mean, though, that they are trivial. Instead, the topological constraint hinders the fields from assuming the same vacuum value everywhere, such that a non-vanishing energy is stored in the field's gradients.

To understand textures, it is easiest to start in one spatial dimension. In this case, textures are associated with the fundamental group. We already know that the fundamental group of the $U(1)$ model of Sec. 1.1.1 has a circle as a vacuum manifold and is therefore multiply connected. Let us restrict in the following to the global case with the Lagrangian given in Eq. (1.1). As depicted on the left-hand side of Fig. 2.8, we can assign to each point of the vacuum a unique arrow. We can now imagine a configuration that, as traversing the entire space, runs once around the vacuum and approaches the same configuration in both asymptotic regions $\lim_{x\to-\infty} \phi(x) = \lim_{x\to\infty} \phi(x)$. If we imagine the one spatial dimension as compactified to a finite interval, this is the configuration we can see on the lower right-hand side of Fig. 2.8, where to each point is assigned the arrow corresponding to the vacuum value. Clearly, such a configuration is constructed in a way that the vacuum is assumed everywhere in space, but the spatially varying field values lead to non-vanishing gradients. If our space is really $\mathbb{R}$, there is no reason for the configuration to assume the same asymptotic values. If our space were S^1, though, the condition would automatically be satisfied for continuous fields. In Fig. 2.8, this amounts to nothing but identifying the left and the right end of the compactified space.

Let us now return to three spatial dimensions. The model described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \Phi^T \partial^\mu \Phi - \frac{\lambda}{4}\left(\Phi^T\Phi - \eta^2\right)^2 , \qquad \text{with} \qquad \Phi = (\phi_1, \phi_2, \phi_3, \phi_4) , \tag{2.101}$$

has a potential with an $O(4)$ symmetry which is spontaneously broken to $O(3)$ when the field acquires a VEV $\Phi^T\Phi = \eta^2$. The minima of the vacuum can be seen to describe a three-sphere, which can also be observed by considering $O(4)/O(3) \cong S^3$. Since $\pi_3(S^3) \cong \mathbb{Z}$, this model will give rise to textures. A texture field configuration can in this model be found as follows. Any texture or trivial vacuum configuration assumes everywhere in space values that lie in the vacuum manifold $\mathcal{M} \cong S^3$. A unit-vector $\hat{\boldsymbol{X}}$ on the three-sphere can be parametrised by three angular variables $\chi, \psi \in [0, \pi]$ and $\omega \in [0, 2\pi)$, analogous to spherical coordinates for S^2, and any vacuum field configuration can be written as $\Phi = \eta \hat{\boldsymbol{X}}$. The simplest non-trivial vacuum configuration one can consider is the spherically symmetric texture, obtained by setting $\psi = \theta$, $\omega = \varphi$, and letting χ be a function of t and r only. Explicitly, this means [219, 220]

$$\Phi(t, r, \theta, \varphi) = \eta \left(\sin(\chi(r,t)) \frac{x}{r}, \sin(\chi(r,t)) \frac{y}{r}, \sin(\chi(r,t)) \frac{z}{r}, \cos(\chi(r,t)) \right) . \tag{2.102}$$

This solution is not stable, as can be seen by a simple scaling argument. Following the discussion of Derrick's theorem in Eq. (2.8), the energy of a texture which is only due to gradients changes in three spatial dimensions under a rescaling $\boldsymbol{x} \to \alpha \boldsymbol{x}$ as $E \to \alpha^{-1} E$. The energy can, therefore, be reduced by increasing α, which corresponds to a shrinking of the texture. This computation assumes, however, that the space is $\mathbb{R}^3$. If the space were S^3, the texture could indeed be stable [219]. Moreover, a cosmological phase transition will, of course, produce textures also on super-Hubble scales. Even if our Universe is spatially flat, the texture collapse can then only happen once the texture has become sub-Hubble-sized.

If textures were to continue to shrink arbitrarily small, they would collapse into a black hole. That this is not the case can be seen following Ref. [203]. The energy density of the texture is computed to be [221]

$$\rho_T(t, r) = 2\eta^2 \frac{r^2 + 3t^2}{(r^2 + t^2)^2} , \tag{2.103}$$

with a collapse at $t = 0$ as indicated by the divergence of the density for $r \to 0$. However, this does not yet indicate the formation of a black hole as prior to this, the energy density below a certain radius r_f becomes comparable to the energy density in the unbroken phase $\rho \sim \lambda \eta^4$. When this happens, the texture can unwind and decay into Goldstone radiation. Since close to the collapse $\rho_T \sim \eta^2 / r^2$, we have $r_f \sim \lambda^{-1/2} \eta^{-1}$. At highest density, the texture has an energy $E_f \sim \rho_T(r_f) r_f^3 \sim \eta^2 r_f$ inside a sphere of radius r_f which has a corresponding Schwarzschild radius $r_S(E_f) = 2GE_f \sim G\eta^2 r_f \ll r_f$. The last inequality follows from the fact that CMB temperature anisotropies constrain $G\eta^2 \lesssim 10^{-6}$ in the case of global textures [222]. When the texture reaches a size still largely above the Schwarzschild radius, it will evaporate into Goldstone radiation and, therefore, not directly lead to the formation of a primordial black hole.

The shrinkage and decay of textures leads to their correlation length rapidly growing to the Hubble scale d_H and then continuing to grow with it $\xi \sim d_H$, thereby entering a scaling regime [220]. In this regime, textures produce an irreducible gravitational wave background [73, 88, 89, 223, 224] (see also Ref. [225]). While textures are no longer a viable option as the main seeds for structure formation, they could explain the overabundance of supermassive black holes [203]. Furthermore, textures are a possible source of cosmic rays and may lead to CMB spectral distortions [226].

So far, we have been concerned with textures from the spontaneous breaking of a global symmetry. In the case of a local symmetry, the scalar field gradient can be eliminated by the corresponding gauge field, and the configuration actually carries virtually no energy. While we do not want to consider gauged textures here, the previous statement does not mean that local textures can have non-trivial dynamics or a cosmologically interesting phenomenology. Indeed, local textures have, e.g., been considered as giving rise to non-Abelian magnetic fields [219] and in the context of electroweak baryogenesis [227].

Semilocal and embedded defects: From textures, it is clear that field gradients can actually become very important when studying cosmic defects. Other instances for which gradients are vital are *semilocal defects*, first realised in the case of semilocal strings in Ref. [228]. In this reference, the authors considered a model that exhibits, similarly to the electroweak model, an

$SU(2) \times U(1)$ symmetry. In contrast to the electroweak model, here, only the $U(1)$ symmetry is a gauge symmetry, whereas the $SU(2)$ symmetry is global. The Lagrangian of this model can be written as

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}^\mu \Phi) - \frac{\lambda}{4} \left(\Phi^\dagger \Phi - \eta^2 \right)^2 , \tag{2.104}$$

where Φ is in the fundamental representation of $SU(2)$ and $\mathcal{D}_\mu \Phi = (\partial_\mu - ieA_\mu) \Phi$ with A_μ denoting the $U(1)$ gauge field and $F_{\mu\nu}$ its associated field strength. When Φ acquires a non-vanishing VEV, the symmetry is spontaneously broken:

$$SU(2)_{\text{global}} \times U(1)_{\text{local}} \longrightarrow U(1)_{\text{global}} . \tag{2.105}$$

We can see from the potential in Eq. (2.104) as well as from the symmetry-breaking pattern that the vacuum manifold is a three-sphere $\mathcal{M} \cong [SU(2) \times U(1)]/U(1) \cong S^3$ and has a trivial fundamental group. Hence, the model does not exhibit topologically stable strings. These considerations are, however, incomplete.

To understand this, first note that the $U(1)_{\text{local}}$ symmetry generates an S^1 worth of gauge equivalent vacua. If the field changes in space in such a way that it moves along such an orbit, the gauge field can compensate for the Higgs gradient, and the gradient energy vanishes. If the field moves in a vacuum direction orthogonal to a $U(1)_{\text{local}}$ orbit, i.e., into a direction associated with a change in the global $SU(2)_{\text{global}}$ group, this compensation is not possible and the configuration will necessarily have non-vanishing gradient energy. We can now imagine moving along a loop in space such that when traversing it, the Φ-configuration winds around a $U(1)_{\text{local}}$ orbit. If we shrink the loop in space to a point, then the field either needs to move away from this orbit, which would mean overcoming a gradient-energy barrier, or the field would need to leave the vacuum and obtain a phase of unbroken symmetry somewhere inside the circle. By the same arguments as in Sec. 2.2.2, the latter option amounts to the presence of a cosmic string. Which option is chosen has no general answer but depends on the model parameters. The stability of semilocal strings was investigated, e.g., in Refs. [228–230] with the conclusion that for $\lambda < 2e^2$, the strings are stable and for $\lambda > 2e^2$ they are unstable. The same idea can, of course, be applied to defects of different dimensions as discussed in Ref. [231]. For a more detailed general overview, we refer the reader to Ref. [232].

This mentioned review goes beyond semilocal defects and introduces the more general concept of embedded defects, which have been classified in Refs. [233–236]. The rough idea is that in a symmetry breaking $G \to H$, we can consider subgroups embedded in the full symmetry group $G_{\text{emb}} \subset G$ that get broken in the process according to $G_{\text{emb}} \to G_{\text{emb}} \cap H$. This spontaneous symmetry breaking can give rise to defect configurations that are, in a sense, embedded in the full theory under conditions discussed in detail in the above references. Instead of reviewing this classification, we consider a simple example: the electroweak string. To this end, consider the scalar sector of the electroweak theory, which was shown in Refs. [237–239] to contain an embedded string solution of the form

$$H = (0, \phi_{\text{ANO}}) , \qquad Z_\mu = A_\mu^{\text{ANO}} , \qquad A_\mu = W_\mu^\pm = 0 , \tag{2.106}$$

where H denotes the electroweak Higgs doublet, Z_μ the Z boson, A_μ the photon, and $W_\mu^\pm$ the two W bosons. The functions ϕ_{ANO} and A_μ^{ANO} denote the ANO string configurations described by Eq. (2.20). These embedded strings are known as Z-strings and arise from breaking the embedded subgroup $U(1)_Z \to \{ \mathbb{1} \}$.

In Ref. [240], it was shown that for certain parts of parameter space (namely in the limit approaching the stable regime of semilocal strings), Z-strings can be stable. For the actual parameter values realised in Nature, they are, however, unstable. In the literature, different stabilisation mechanisms have been considered, such as the presence of bound states in Refs. [241, 242], external electric and magnetic fields in Ref. [243] and Ref. [244], as well as explicit symmetry breaking in Ref. [245]. Yet another way to stabilise embedded strings was found in Refs. [246, 247] using the fact that defects form in the early Universe surrounded by a thermal plasma. We will study this plasma stabilisation in more detail in Ch. 3 and see that it can be used to stabilise embedded domain walls in an extension of the SM of particle physics.

Superconducting strings: Another phenomenologically interesting possibility is that of superconducting strings, as first proposed in Ref. [248]. These can arise when, after a phase transition that gives rise to strings, there is still an unbroken Abelian gauge symmetry (electromagnetism), which gets broken inside the string. Inside the string, the otherwise massless gauge boson becomes massive, which results in the Meißner effect: External magnetic fields decay exponentially over a characteristic length given by the inverse of the photon mass inside the string. The screening happens by means of a superconducting current. We want to briefly sketch the idea behind such superconducting strings using a toy model introduced in Ref. [248]. For intermediate steps and an extended discussion, we refer the reader to this reference.

The toy model exhibits a gauge symmetry $U(1)_X \times U(1)_Y$ with respective gauge fields X_μ, Y_μ and field strengths $X_{\mu\nu}, Y_{\mu\nu}$. Moreover, it contains two complex scalar fields ϕ and ψ. The Lagrangian is of the form

$$\mathcal{L} = -\frac{1}{4} X_{\mu\nu} X^{\mu\nu} - \frac{1}{4} Y_{\mu\nu} Y^{\mu\nu} + \left(\mathcal{D}_\mu \phi\right)^* \left(\mathcal{D}_\mu \phi\right) + \left(\mathcal{D}_\mu \psi\right)^* \left(\mathcal{D}_\mu \psi\right) - V(\phi, \psi) \,, \tag{2.107}$$

where $\mathcal{D}_\mu \phi = \left(\partial_\mu - i g_X X_\mu\right) \phi$ and $\mathcal{D}_\mu \psi = \left(\partial_\mu - i g_Y Y_\mu\right) \psi$, and the potential reads

$$V(\phi, \psi) = \frac{\lambda_\phi}{4} \left(|\phi|^2 - \eta_\phi^2\right)^2 + \frac{\lambda_\psi}{4} \left(|\psi|^2 - \eta_\psi^2\right)^2 + \Lambda |\phi|^2 |\psi|^2 \,. \tag{2.108}$$

Consider the parameters chosen in a way such that the VEV of ϕ is non-vanishing and the $U(1)_X$ symmetry is spontaneously broken but $U(1)_Y$ remains unbroken, i.e., the VEVs are $|\phi_0| = \eta_\phi$ and $|\psi_0| = 0$. This can be ensured by choosing the parameters to satisfy $\lambda_\phi \eta_\phi^4 > \lambda_\psi \eta_\psi^4$ and $\Lambda \eta_\phi^2 \geq \lambda_\psi \eta_\psi^2 / 2$. The unbroken $U(1)_Y$ symmetry could, for example, be associated with the gauge group of electrodynamics $U(1)_{\mathrm{EM}}$.

While outside of the string, the field ϕ assumes a non-vanishing VEV, and the term $\Lambda |\phi|^2 |\psi|^2$ in the potential forces $\psi = 0$, inside the string, this interaction term vanishes. It can now be energetically favourable for ψ to assume a non-vanishing VEV and thereby minimise the second term in the potential given in Eq. (2.108). Thereby, the $U(1)_Y$ symmetry is unbroken in space except along ϕ-strings. Considering a straight ϕ-string along the z-axis, we can write for the ψ-field configuration $\psi(\rho, \theta, z) = \tilde{\psi}(\rho, \theta) e^{i\vartheta(z)}$. The function $\tilde{\psi}(\rho, \theta)$ will fall off over some width δ_ψ, which is typically very thin as the field ψ quickly approaches its unbroken vacuum value $\psi = 0$ outside of the string.

Associated with the unbroken $U(1)_Y$ symmetry is a conserved Noether current density

$$j_\mu = i g_Y \left(\psi \left(\mathcal{D}_\mu \psi\right)^* - \psi^* \mathcal{D}_\mu \psi\right) \,. \tag{2.109}$$

For the ψ-configuration described above, the integrated current can be found to read

$$I = \int \rho \, \mathrm{d}\rho \, \mathrm{d}\theta \, j_z = 2 g_Y K \left(\partial_z \vartheta - g_Y Y_z\right) \,, \tag{2.110}$$

where the integral runs over the cross-section of the string, and we assumed A_z to be constant inside the string. We also defined $K = \int \rho \, \mathrm{d}\rho \, \mathrm{d}\theta \, \tilde{\psi}^2(\rho, \theta)$.

Consider now a closed string loop of length $2\pi R$. Along the loop, ϑ can only change by an integer value $n \in \mathbb{Z}$ due to the single-valuedness of ψ. This integer can be computed as

$$n = \frac{1}{2\pi} \oint_{2\pi R} \mathrm{d}\sigma \, \frac{\mathrm{d}\vartheta}{\mathrm{d}\sigma} \,, \tag{2.111}$$

with σ running along the string (for our straight string, we can locally match $\sigma = z$). This number is conserved as it can only change discretely. From Maxwell's equations in Lorenz gauge, it follows that

$$Y_{\parallel} = \frac{1}{2\pi} \ln\left(\frac{R}{\delta}\right) I \,, \tag{2.112}$$

where $Y_{\parallel}$ is the gauge field projected along the string, and the string width δ appears as a natural cutoff that acts as a regulator. For a constant current along the string, we have $\mathrm{d}\vartheta/\mathrm{d}l = n/R$, and we can use Eqs. (2.110) and (2.112) to find

$$I = \frac{2 g_Y K}{1 + \frac{g_Y^2 K}{\pi} \ln\left(\frac{\delta}{R}\right)} \frac{n}{R} \,. \tag{2.113}$$

This demonstrates that along such a string, there is a persistent current flowing if $n \neq 0$. Superconducting strings are not constrained to the case of bosonic currents on them. The Higgs field making up the string can also couple to fermions via a Yukawa term and give rise to fermionic superconductivity, which is also discussed in Ref. [248]. For an extensive review, see also Ref. [36].

The currents carried by these strings can be large enough to be of cosmological relevance. They could play a role in primordial magnetogenesis [249–251], and are a possible source of fast radio bursts [252, 253] and high-energy cosmic rays [61, 254, 255]. Furthermore, they can lead to spectral distortions of the CMB [65–67].

Melting defects: While we typically consider symmetry restoration at high temperatures, it is also possible to restore symmetries at low temperatures that are broken at high temperatures [131]. The idea behind melting defects is to start out with a high-temperature symmetry breaking that gives rise to defects, but restore the symmetry at low temperatures. Defects in such models will generally have quite different properties, such as temperature-dependent tensions or masses. Most importantly, they will dissolve or "melt" as a consequence of the restoration process [167]. Melting defects have recently found more interest [44, 256–260] as they seem a fitting explanation of the recently found evidence for a nano-hertz GWB (cf. Ch. 7).

How melting defects can be implemented into a particle physics model is most easily understood in terms of a toy model. To this end, we will use the one considered in Refs. [44, 257, 258, 260], which gives rise to domain walls. Consider a model containing a real scalar field ϕ which is coupled to a multiplet Ψ of n_Ψ real scalar fields. The relevant part of the Lagrangian, including all ϕ-dependent terms, reads

$$\mathcal{L}_\phi = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \frac{\lambda_\phi}{4} \phi^4 + \Lambda \Psi^T \Psi \phi^2 - \frac{\lambda_\Psi}{4} \left(\Psi^T \Psi\right)^2 \,, \tag{2.114}$$

where $m, \lambda_\phi, \lambda_\Psi, \Lambda > 0$. The potential must be bounded from below, which requires $\lambda_\phi \lambda_\Psi \geq 4\Lambda$. This model has a $\mathbb{Z}_2$ symmetry as the Lagrangian remains invariant under transformations of the form $\phi \mapsto \pm\phi$.

Suppose that at early times, the field Ψ is in equilibrium with a thermal bath of temperature T and fluctuates about a classical background $\Phi = 0$ such that the finite-temperature effective potential of ϕ obtained from Eq. (1.51) reads (up to ϕ-independent terms)

$$V_{\text{eff}}^{(T)}(\phi) = \frac{1}{2} \left(m^2 - \frac{n_\Psi \Lambda}{6} T^2\right) \phi^2 + \frac{\lambda_\phi}{4} \phi^4 \,. \tag{2.115}$$

Observe that for sufficiently high temperatures $T > T_c = \sqrt{6/(n_\Psi \Lambda)}m$, the minimum of the potential is not at $\phi = 0$ but at

$$\phi = \pm\sqrt{\frac{n_\Psi \Lambda}{6\lambda_\phi}T^2 - \frac{m^2}{\lambda_\phi}}\,. \tag{2.116}$$

Assuming that at even earlier times, the field ϕ started in the symmetric phase at $\phi = 0$ (for details, see Ref. [257]), the $\mathbb{Z}_2$ symmetry is spontaneously broken at these high temperatures, and domain walls arise. Supposing furthermore that the changes in the temperature are adiabatic, the domain wall tension becomes according to Eq. (2.5) for $T \gg T_c$

$$\sigma(T) \simeq \frac{1}{3\lambda_\phi}\left(\frac{n_\Psi \Lambda}{3}\right)^{3/2} T^3\,. \tag{2.117}$$

The tension of these walls is therefore strongly temperature-dependent and decreases as the Universe cools down. At $T \sim T_c$, the domain walls will dissolve. Such walls are interesting as their temperature-dependent tension will drastically change their phenomenological implications in comparison to constant-tension walls, e.g., the GWB spectrum [258–260]. Additionally, entirely stable domain walls cannot exist as they overclose the Universe [31]. Melting domain walls decay and thereby avoid this *domain wall problem*.

The above idea can similarly be applied to other defects, such as strings (see Ref. [256]). In spite of the probably interesting phenomenology of melting defects, the literature on them is sparse, highlighting a promising direction for future studies.

Time-dependent defects: We have seen by now examples that avoid Derrick's theorem due to their infinite energy or due to the introduction of gauge fields. Another option is to consider localised, non-dissipative and time-periodic field configurations. These are known as *non-topological solitons* or *Q-balls* and were first introduced in Refs. [261–263]. These solitons are not topologically stabilised but rather by energetic minimisation and non-trivial conserved Noether charges (Q). Following Ref. [263], we can consider once again a complex scalar field ϕ with global $U(1)$ symmetry and Lagrangian

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi)\,. \tag{2.118}$$

We can choose the potential such that $V(0) = 0$, and we consider a potential for which this is the absolute minimum. A Q-ball is a time-periodic configuration

$$\phi(t, \boldsymbol{x}) = \hat{\phi}(\boldsymbol{x})e^{-i\omega t}\,, \tag{2.119}$$

which is spatially localised. We could, e.g., imagine a spherically symmetric configuration that is non-vanishing with a constant value $\hat{\phi}_0$ in some region of radius R with volume $V_R = 4\pi R^3/3$ and vanishing outside. In reality, the field will, of course, continuously transition to $\phi = 0$.

Associated with this configuration is, due to Noether's theorem, the conserved $U(1)$ charge

$$Q = i\int \mathrm{d}^3x \left(\phi^*\dot{\phi} - \dot{\phi}^*\phi\right) \simeq V_R \omega |\hat{\phi}_0|^2\,. \tag{2.120}$$

In the same approximation, we can compute the configuration's energy:

$$E \simeq V_R\left(\frac{1}{2}\omega^2|\hat{\phi}_0|^2 + V(\phi)\right) \simeq \frac{1}{2V_R}\frac{Q^2}{|\hat{\phi}_0|^2} + V_R V(\phi)\,, \tag{2.121}$$

where we neglected the surface contributions associated with the gradients, which is only problematic for too small Q. The energy is minimised for a volume $V_R^{\min}$:

$$E_{\min} = \sqrt{\frac{2V(\phi)Q^2}{|\hat{\phi}_0|^2}} \qquad \text{for} \qquad V_R^{\min} = \sqrt{\frac{Q^2}{2V(\phi)|\hat{\phi}_0|^2}} \,. \tag{2.122}$$

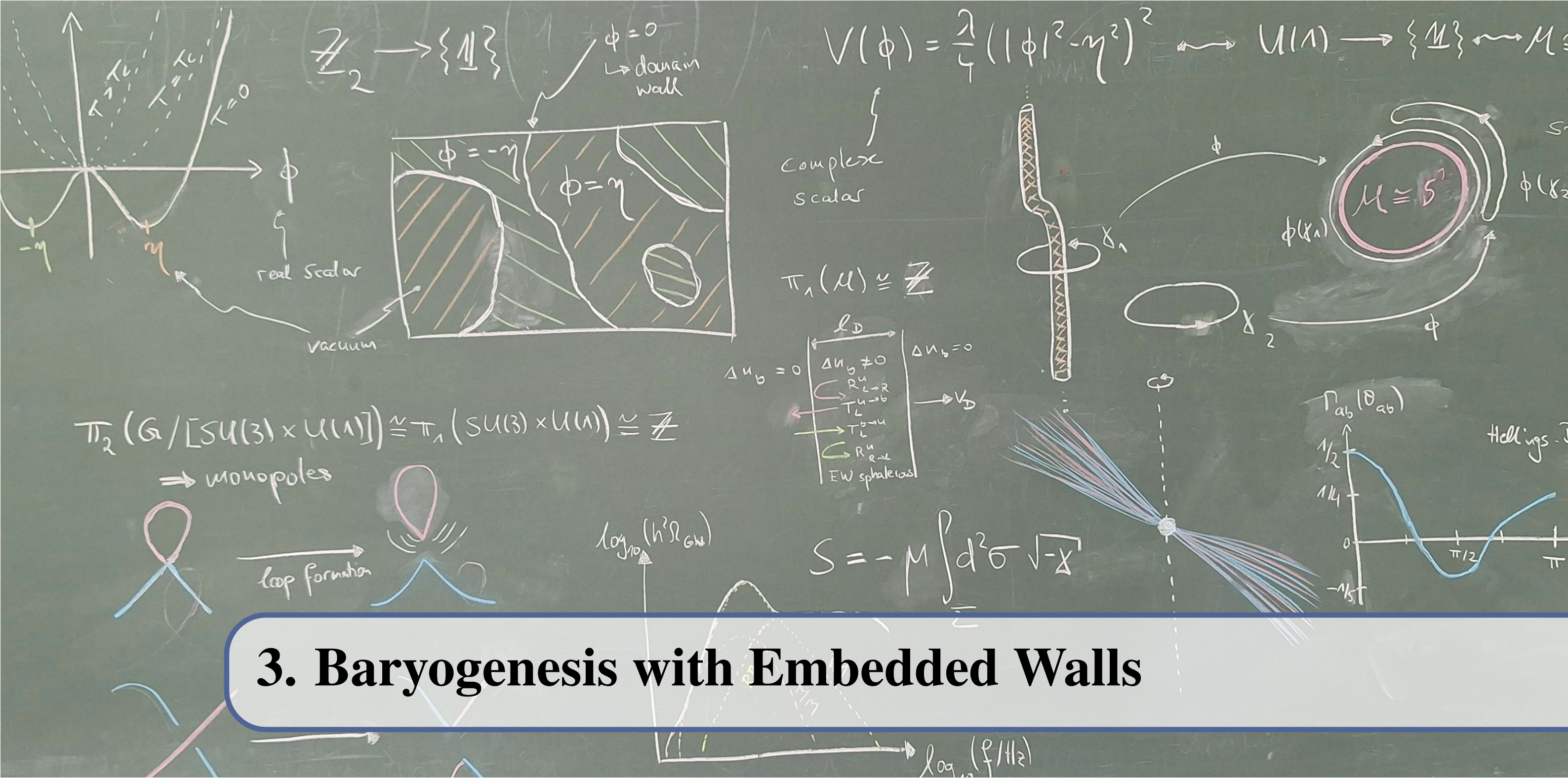
The Q-ball will be stable against the emission of ϕ-particles if the energy per unit charge $E_{\min}/Q$ is smaller than the minimum energy of a free ϕ-particle which is $m_\phi = \sqrt{V''(\phi)|_{\phi=0}}$. Therefore, stable Q-balls can exist if for some ϕ of the form in Eq. (2.119) the condition

$$V(\phi) < \frac{1}{2} m_\phi^2 |\phi|^2 \tag{2.123}$$

is satisfied [263]. Choosing the potential to satisfy this condition, we can obtain non-dissipative and spatially localised configurations, whose decay is prevented by Noether's theorem and the shape of the potential. No topological conservation laws are involved, and the Q-ball is only time-periodic, not static.

Apart from these simplest non-topological solitons, there are also solitons whose stability requires combinations of Noether conservation laws and topological conservation laws [264–268]. In addition to these stable solitons, there are similar configurations which are not protected by charge conservation and are dissipative. Nonetheless, they can be long-lived. These are referred to as quasi-solitons or oscillons [269–273].

Non-topological solitons can be interesting in the context of cosmology as they are possible dark matter candidates [274–276], can produce primordial GWs [277–281], and are used in some baryogenesis mechanisms [282, 283]. For recent reviews of non-topological (quasi) solitons, see Refs. [284, 285].



3. Baryogenesis with Embedded Walls

This chapter is based on Ref. [2]. The detailed contributions of the author to this article are outlined at the beginning of this thesis.

We have, by now, seen how cosmic defects arise in various particle physics models and mentioned their wide range of applications in cosmology. One phenomenologically appealing possibility is to utilise defects to generate the observed baryon asymmetry of the Universe. This asymmetry can be expressed in terms of the net baryon-to-entropy ratio with a measured value of [286, 287]

$$\mathcal{B}_{\text{obs}} \equiv \frac{n_b}{s} \simeq 9 \times 10^{-11} \,. \tag{3.1}$$

Here and in the following, we use n_b to denote the net baryon-number density, i.e., the difference between baryon-number density and antibaryon-number density.[1] As originally formulated by SAKHAROV in Ref. [288], a process generating a baryon-antibaryon asymmetry from symmetric initial conditions[2] requires three key ingredients. The process must

1. violate baryon number conservation,
2. violate C and CP symmetry,
3. take place out of thermal equilibrium.

The reason for the first condition is obvious. If the model under consideration does not allow for baryon-number-violating processes, an initially vanishing baryon number will remain zero. If the second condition were not fulfilled, then to each process that generates a baryon asymmetry, there would be an inverse process generating the equal and opposite baryon asymmetry at the same rate, such that no net baryon number B would be produced. For the third condition, one may use that, starting out from a symmetric initial condition, the expectation value of the baryon number satisfies

---

[1] For the observed ratio, n_b and the baryon–number density are, for all practical purposes, the same since the observed antibaryon-number density is negligibly small.

[2] This is a natural starting point if one assumes the Universe to undergo an inflationary period, which would wash out all asymmetries.

in thermal equilibrium

$$\begin{aligned}\langle B\rangle_T = \operatorname{tr}\left(e^{-\beta H}B\right) &= \operatorname{tr}\left(\mathcal{CPT}(\mathcal{CPT})^{-1}e^{-\beta H}B\right) = \\ &= \operatorname{tr}\left(e^{-\beta H}\,(\mathcal{CPT})^{-1}\,B\,\mathcal{CPT}\right) = -\left\langle B\right\rangle_T \,,\end{aligned}$$

from which it follows that $\langle B\rangle_T = 0$. Here, we used in the second-to-last step that the Hamiltonian H commutes with the $\mathcal{CPT}$ operator, and in the last step that B is even under $\mathcal{P}$ and $\mathcal{T}$, but odd under $\mathcal{C}$.

The SM fulfils the first two conditions: It includes C and CP violation in electroweak interactions and allows for baryon-number-violating sphaleron transitions, as pointed out in Ref. [289]. The third condition is also generally satisfied due to the expansion of our Universe. Unfortunately, the CP violation is too small, and the departure from thermal equilibrium is insufficient, such that for successful baryogenesis BSM physics is required.

In this chapter, we will propose a model of electroweak baryogenesis (EWBG) which utilises plasma-stabilised embedded domain walls to satisfy the out-of-thermal-equilibrium condition.

Let us start by summarising the original idea behind EWBG (see, e.g., Refs. [290–294] for reviews). Suppose the electroweak phase transition were strongly first-order. Then it would proceed through the nucleation and expansion of bubbles of the broken phase within the surrounding unbroken phase. On the bubble walls, which separate these two phases, thermal equilibrium could not be maintained. This departure from equilibrium would allow electroweak sphaleron processes inside the walls to generate a baryon asymmetry. This mechanism cannot be realised within the SM since, without any extensions, the electroweak phase transition is a smooth crossover [145]. Nonetheless, a variety of BSM models can, for the right parameter values, give rise to a first-order electroweak phase transition (see e.g. Ref. [295]); among them, the two-Higgs-doublet model (2HDM) (see, e.g., Refs. [296–298]). Note that these models usually also need a BSM source of CP violation at the boundary of the wall. This can again be achieved by involving a 2HDM, which can contain a CP-violating relative phase between the two Higgs doublets.

As discussed in Refs. [68, 69, 299–301], there is an alternative route to realise EWBG without a first-order phase transition, but rather by incorporating cosmic defects. Since defects represent field configurations which are out of thermal equilibrium, they offer a natural environment for baryogenesis, provided that the electroweak symmetry remains unbroken within the defect. Similar to EWBG via a first-order electroweak phase transition, an additional source of CP violation at the defect boundary is required. As before, this can be achieved involving a 2HDM.

Among the various defect types, cosmic strings have drawn particular interest due to the fact that a string network reaches a scaling regime in which it contributes a fixed fraction to the total energy density of the Universe (see Ch. 4). This prevents strings from any overclosure problems. It was, however, shown in Refs. [69, 300, 302] that baryogenesis mediated by cosmic strings is insufficient: As a result of the microscopic width of the strings, it proceeds only in a small fraction of the total volume of space and fails to account for the observed baryon-to-entropy ratio.

It was already noted in Refs. [69, 300, 301] that domain walls could, in principle, overcome this issue. Compared to walls, the volume available for baryogenesis inside strings is suppressed by a factor of the string width over the Hubble scale. However, models with stable domain walls are cosmologically ruled out, as they would overclose the Universe (domain wall problem) [31]. Plasma-stabilised embedded domain walls offer a promising alternative [301]: They can provide spatially extended environments for efficient EWBG without leading to the cosmological problems associated with the stability of domain walls.[3]

As an original contribution, we show in this chapter that plasma-stabilised embedded domain walls can actually be realised in an extension of the SM, and that such walls can indeed provide an

---

[3] A different proposal of defect-mediated baryogenesis was recently discussed in Ref. [283].

efficient mechanism of EWBG. We also discuss the effect of having a network of domain walls rather than a single wall and show that the simple enhancement factor found for domain-wall mediated baryogenesis in Refs. [69, 301] is only applicable under extremely restrictive conditions.

We begin this chapter with a short review to illustrate how interactions with the early-Universe plasma can stabilise embedded defects in Sec. 3.1. In Sec. 3.2, we estimate the baryon-to-entropy ratio produced by domain walls — first for a single wall, then for a network of walls. Importantly, this computation is independent of the decay mechanism of domain walls. It can also be used as a rough estimate if one chooses a model in which embedded walls are replaced by metastable, melting, or biased DWs. In Sec. 3.3, we provide an extension of the SM that gives rise to plasma-stabilised embedded walls which satisfy the properties required for EWBG. We conclude with a brief discussion in Sec. 3.4.

3.1 Plasma stabilisation of embedded defects

Let us first consider the mechanism of plasma stabilisation of embedded defects, which was originally studied in Refs. [246, 247] (for later work, see also Ref. [303]). We will illustrate the idea behind this mechanism in terms of the electroweak Z-string; an embedded string in the standard electroweak theory which we already encountered briefly in Sec. 2.4. The electroweak symmetry breaking $SU(2)_L \times U(1)_Y \to U(1)_{\mathrm{EM}}$ is governed by the Higgs doublet H, which acquires in the phase transition a non-vanishing VEV. Here, $SU(2)_L$ denotes the weak-isospin group, $U(1)_Y$ denotes the weak-hypercharge group, and $U(1)_{\mathrm{EM}}$ is the gauge group of electromagnetism. The symmetry breaking is achieved by a potential of the form

$$V_H(H) = \lambda_H \left(H^\dagger H - \frac{v_H}{2}\right)^2 = \frac{\lambda_H}{4}\left(\sum_{i=1}^{4} H_i^2 - v_H^2\right)^2 , \tag{3.2}$$

where λ_H is the Higgs coupling constant, v_H is the VEV of H in the broken phase, and we decomposed the Higgs field into four canonically normalised real scalar fields according to $H = (H_1 + iH_2, H_3 + iH_4)^T/\sqrt{2}$. We see that the vacuum of the theory is

$$\mathcal{M} = \left\{ (H_1, H_2, H_3, H_4)^T \in \mathbb{R}^4 \;\middle|\; \sum_{i=1}^{4} H_i^2 = v_H^2 \right\} \cong S^3 , \tag{3.3}$$

which does not admit topologically stable string solutions. As discussed previously in Sec. 1.2, as long as H is in thermal equilibrium, the full gauge symmetry is restored.[4] Thermal fluctuations of scalars, vectors, and spinors contribute a term to the finite-temperature effective potential of the Higgs field, which is of the form $\sim T^2H^2$, where T is the temperature of the thermal bath and the omitted prefactor consists of combinations of the different coupling constants. For sufficiently large temperatures, this shifts the minimum of the effective potential of the Higgs field to $H = 0$ and the symmetry is unbroken.

When studying the formation of electroweak Z-strings, one must consider the situation where H drops out of thermal equilibrium, but the massless photon field remains in thermal equilibrium. Choosing a gauge in which the Higgs doublet assumes a VEV of the form $H_0 = (0, v_H)^T$, we can write the complex doublet as $H = \left(H^+, H^0\right)$ with electric charges $Q_{\mathrm{EM}}(H^+) = 1$ and $Q_{\mathrm{EM}}(H^0) = 0$. Let us now consider the part of the SM Lagrangian that effectively describes the full Higgs field after the electroweak symmetry breaking as well as the photon field A_μ with an

---

[4] As we mentioned before, symmetry need not necessarily be restored at high temperatures [131] and there are models in which the electroweak symmetry is not restored, see e.g. Refs. [141, 304, 305], or restored only far above the electroweak scale, see e.g. Ref. [306].

associated field strength tensor $F_{\mu\nu}$. It is of the form

$$\mathcal{L}_H = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \left(\mathcal{D}_\mu H^+\right)^* \mathcal{D}^\mu H^+ + \partial_\mu H^0 \partial^\mu H^0 - V(H)\,, \tag{3.4}$$

where the covariant derivative reads $\mathcal{D}_\mu H^+ = \left(\partial_\mu - ieA_\mu\right) H^+$ and we used that H^0 is electrically neutral. The interaction between the Higgs and the thermal photon background leads at one-loop level to a contribution to the effective potential of the form [246, 247]

$$\sim e^2 T^2 \left(H_1^2 + H_2^2\right)\,. \tag{3.5}$$

The crucial point is that this contribution exists only for the electrically charged fields H_1, H_2, but not for the uncharged field directions H_3, H_4. The effective potential, which, in the absence of these contributions, has a vacuum manifold $\cong S^3$, is lifted in the charged field directions. This reduces the vacuum manifold effectively to a circle:

$$\mathcal{M}_{\text{eff}} = \left\{\, (H_1, H_2, H_3, H_4)^T \in \mathbb{R}^4 \;\middle|\; H_1 = H_2 = 0 \text{ and } H_3^2 + H_4^2 = v_H^2 \,\right\} \cong S^1\,. \tag{3.6}$$

The neutral Higgs component can have non-trivial winding in the effective vacuum manifold, which corresponds to strings stabilised by the effective vacuum. The strings stabilised in this way are nothing but the electroweak Z-strings we encountered in Eq. (2.106).

Plasma stabilisation of embedded defects is not restricted to the electroweak model. For example, consider the chiral limit of the linear sigma model of QCD. The model consists of the uncharged sigma field σ, and the three pions $\pi^\pm, \pi^0$, of which two are electrically charged and one is uncharged. The potential of this model is invariant under $O(4)$ rotations between these four fields. Spontaneous symmetry breaking reduces the symmetry group to $O(3)$, and the associated vacuum manifold is $O(4)/O(3) \cong S^3$. Due to a coupling to the electromagnetic field in thermal equilibrium, the symmetry-breaking potential gets lifted in the charged field directions and the vacuum manifold is reduced to the effective form $\mathcal{M}_{\text{eff}} \cong S^1$. This stabilises a string configuration known as *pion string* [307], which is of phenomenological interest as it can be used to generate primordial magnetic fields [308].

However, for our discussion, we are interested in embedded domain walls that are stabilised by interactions with the electromagnetic plasma, not strings. This requires a disconnected effective vacuum $\pi_0\left(\mathcal{M}_{\text{eff}}\right) \ncong \{\, \mathbb{1} \,\}$. As a toy model, we could imagine a real scalar triplet for which two scalar degrees of freedom are charged under $U(1)_{\text{EM}}$, while the third one is uncharged. Plasma interactions would then reduce a vacuum manifold $\mathcal{M} \cong S^2$ to the embedded vacuum manifold $\mathcal{M} \cong S^0$, and embedded walls would be stabilised. In Sec. 3.3, we will demonstrate that the SM can be extended in such a way that embedded walls arise, which are stabilised by the electromagnetic plasma.

Before we turn to this task of model building, we first review wall-mediated EWBG and investigate whether it can explain the observed baryon-to-entropy ratio.

3.2 Domain-wall-mediated electroweak baryogenesis

3.2.1 A single wall

Suppose we are in a phase of the early Universe just after electroweak symmetry breaking has occurred. An even earlier phase transition, governed by the Higgs field Φ, has given rise to embedded domain walls, which are stabilised due to interactions with the electromagnetic plasma. The Higgs field Φ associated with the embedded walls couples to the electroweak Higgs H in a way such that the electroweak symmetry is restored in a region of width l_D around the core of the embedded defect. In Sec. 3.3, we provide a microphysical model that fits this description. For the baryogenesis mechanism we are about to consider, the actual domain wall does not play a direct role. However, due to the above construction, the electroweak Higgs H will also form a wall-like configuration at the same position as the actual embedded defect. It is the latter wall configuration which is relevant for EWBG as at its core the Higgs expectation value vanishes $H = 0$ and the electroweak symmetry remains unbroken. When using the word "wall" in this section, it is this wall and not the actual domain wall we are referring to.

Let us start by describing defect-mediated EWBG in general. As in any EWBG scenario, the baryon-number-violating process is an electroweak sphaleron transition.[5] While these transitions are exponentially suppressed in the broken phase, they are not suppressed in the electroweak symmetric phase, which is, in our case, inside the wall. Moreover, the wall is a region which dropped out of thermal equilibrium, moving with relativistic velocities through the surrounding plasma. The remaining Sakharov condition is satisfied by the scattering of fermions from the advancing wall in a CP-violating manner, which can be implemented in different ways. For N_f particle families, the rate (per unit volume) at which the chiral asymmetry is converted into a baryon asymmetry can be computed as [311]

$$\dot{n}_b = -\frac{N_f \Gamma_s}{2T} \sum_i \left(3\mu^i_{u,L} + 3\mu^i_{d,L} + \mu^i_{l,L} + \mu^i_{\nu,L}\right) , \tag{3.7}$$

where the μ denote differences between particle and antiparticle chemical potentials, the index i runs over the families and where u, d, l, ν denote up-type, down-type quarks, charged leptons and neutrinos, respectively. The index L indicates left-handedness, and we use a dot to denote a derivative with respect to time. Furthermore, Γ_s is the weak sphaleron rate per unit volume [312]:

$$\Gamma_s = \kappa \alpha_W^4 T^4 , \tag{3.8}$$

with $\alpha_W = g^2/(4\pi)$, the $SU(2)_L$ coupling constant g, and a constant κ in the range $\kappa \sim 0.1 \ldots 1$.[6]

For massless fermions, net number densities n satisfy the relation

$$n = \frac{\mu T^2}{12} , \tag{3.9}$$

such that Eq. (3.7) may be rewritten at sufficiently large temperatures as

$$\dot{n}_b = -\frac{6 N_F \Gamma_s}{T^3} \left(3n_{b,L} + n_{l,L}\right) \equiv -\bar{\Gamma}_s \left(3n_{b,L} + n_{l,L}\right) , \tag{3.10}$$

where $n_{b,L}$ and $n_{l,L}$ denote the total number densities of left-handed baryons and leptons. In addition, we defined $\bar{\Gamma}_s = 6N_F \Gamma_s / T^3$.

There are now two different mechanisms by which defect-mediated baryogenesis can proceed — a local mechanism [227, 315, 316], in which CP violation and conversion of a chiral into a baryon asymmetry both have to happen at the wall's boundary and a non-local mechanism [311, 317–320], where one considers the diffusion of the chiral asymmetry from the wall's edges into the wall.

---

[5] Although baryon and lepton number are classically conserved in the SM, on a quantum level, the chiral anomaly gives rise to their non-conservation [309]. Baryon-minus-lepton number $B-L$ remains, however, conserved. At finite temperatures, a sphaleron process [310] allows to transition between electroweak vacua with different Chern-Simons numbers, which is associated with a change in baryon and lepton number such that $B-L$ is conserved: $\Delta B = \Delta L$. Furthermore, since these processes are linked to the $SU(2)_L$ gauge boson, they affect only left-handed fermions. Consequently, sphalerons can convert a chiral asymmetry into a baryon asymmetry.

[6] The above form can be found on dimensional grounds from the fact that the only relevant scale in the problem is the magnetic screening length $\sim (\alpha_W T)^{-1}$. Nonetheless, this oversimplifies the situation and in Refs. [313, 314], a different dependence on α_W was found $\Gamma_s \propto \ln(1/\alpha_W)\, \alpha_W^5 T^4$. For the numerical evaluation, this difference is, however, covered by the uncertainty in the prefactor κ.

The latter mechanism is strongly dominant in our scenario, and we will only briefly explain local baryogenesis in the following.

Consider a single wall moving through space. At the leading edge of the wall, fermions from the surrounding plasma are scattered off in a CP-violating manner. We restrict to the case where the necessary CP violation is obtained by introducing a second electroweak Higgs doublet, thus extending the Higgs sector to a 2HDM. The CP violation occurs from a relative phase between the two Higgs fields which changes over the wall's boundary by θ_{CP}. In this case, the baryon-to-entropy ratio produced locally at the edge of the wall was estimated in [69, 300] to be

$$\tilde{\mathcal{B}}_{\mathrm{l}} \simeq -\frac{4\Gamma_s}{g_s T^4}\left(\frac{m_F}{T}\right)^2 \theta_{CP}\,, \tag{3.11}$$

where m_F is the fermion mass. In principle, one would need to take the contributions of different fermion species into account, but for a rough order-of-magnitude estimate, we limit our computation to the most massive fermion species that contributes. Furthermore, g_s denotes the effective number of entropic degrees of freedom such that the entropy density reads

$$s = \frac{2\pi^2}{45} g_s T^3\,. \tag{3.12}$$

The name "local" refers, for this baryogenesis mechanism, to the fact that one only considers the baryon asymmetry produced locally at the edge of the wall and assumes that CP violation and baryon-number violation happen at the same point.

In principle, both quarks and leptons are capable of contributing to this process. However, any CP violation generated at the wall boundary that might be converted via weak sphalerons is, for quarks, effectively erased by the much faster strong sphaleron transitions. As a result, leptons are the only particles expected to play a significant role (see [292, 311, 321]). From the dependence of Eq. (3.11) on the fermion mass, it is evident that the heaviest lepton, i.e., the tau lepton, will dominate baryon production, and we therefore set $m_F = m_\tau$. For concreteness, let us assume that θ_{CP} is positive, which results in a net antibaryon number being generated at the leading edge of the wall. These antibaryons must still pass through the wall, which is a region of width l_D where electroweak symmetry is not broken. Within this region, weak sphalerons wash out the baryon asymmetry, leading to a suppression in their number density by a factor of $\exp\left(-\bar{\Gamma}_s l_D/v_D\right)$, where v_D is the wall velocity and l_D/v_D correspondingly the traversal time. $\bar{\Gamma}_s$ is the antibaryon decay rate defined in Eq. (3.10).

At the trailing edge of the wall, baryons are generated in an amount equal to the antibaryons produced at the leading edge. However, the baryons generated at the trailing edge do not pass through the wall and, therefore, are not diluted by sphalerons. As a result, the net baryon-to-entropy ratio after the wall has passed is given by [69, 300]

$$\mathcal{B} = \tilde{\mathcal{B}}_{\mathrm{l}}\left(e^{-\frac{\bar{\Gamma}_s l_D}{v_D}} - 1\right) \sim 10^{-16}\,. \tag{3.13}$$

For the numerical order-of-magnitude estimate, we assumed the width of the wall to be roughly determined by the Compton wavelength of the electroweak Higgs $l_D \sim m_H^{-1}$, and a temperature of $T \sim m_H$, as well as $g_s \sim 10^2$. We furthermore used the values of the free parameters that maximise the produced baryon-to-entropy ratio $\kappa, v_D, \theta_{\mathrm{CP}} \sim 1$. The value obtained in this way for $\mathcal{B}$ is at least 5 orders of magnitude too small to explain the observed baryon-to-entropy ratio given in Eq. (3.1).

Therefore, let us now turn to non-local EWBG originally discussed in Refs. [311, 317–320].[7]

---

[7] The computation of the CP-violating source is intricate and involves numerous potential contributions (for recent reviews, see, e.g., Refs. [293, 322]). Our analysis relies on computing the reflection and transmission coefficients at the wall boundary and provides only a rough order-of-magnitude estimate. For a recent discussion of some of the more subtle aspects, see [323].

As in local EWBG, fermions are reflected at the wall's boundaries in a CP-violating manner. Due to this, a net chiral asymmetry builds at the edges of the wall. The chiral current can then diffuse into the wall where electroweak sphaleron processes generate a baryon asymmetry from the chiral asymmetry. Let us now be more specific and begin by demonstrating, how the chiral current is computed. For this we follow Ref. [69] and focus on the trailing edge of the wall, which we assume to be straight. We fix our coordinates such that this edge lies in the xy-plane at $z = 0$. We denote the reflection coefficient for right-handed fermions incident on the edge from the unbroken phase by $R^u_{R\to L}$ and the transmission coefficient from the unbroken to the broken phase by $T^{u\to b}_R$. Additionally, we use the labels $\bar{L}$ and $\bar{R}$ to denote the CP-conjugates of left- and right-handed particles, respectively. In Ref. [311], it was found that the difference of reflection coefficients satisfies

$$\Delta R(p_z) = R^u_{R\to L} - R^u_{\bar{R}\to\bar{L}} = -4t(1-t^2)\theta_{CP}\frac{m_F}{m_H}\exp\left(-\frac{p_z^2}{m_H^2}\right)\Theta(|p_z| - m_F)\,, \tag{3.14}$$

where $t = \tanh(\vartheta)$, $\tanh(2\vartheta) \equiv m_F/|p_z|$, and, as before, θ_{CP} is the change in the CP-violating angle over the wall's boundary, which we assumed to be of width $\sim m_H^{-1}$. Furthermore, m_H is the mass of the electroweak Higgs, m_F is the mass of the fermion in the broken phase and p_z the z-component of its momentum in the wall frame. The Heaviside function Θ occurs due to the fact that particles with $|p_z| < m_F$ will be totally reflected from the wall's edge and consequently $\Delta R = 0$, while the suppression of large momenta $|p_z| > m_H$ is due to coherence effects across the boundary of the region in which the electroweak symmetry is restored [311]. Since $|p_z| > m_F$, we can write $t \simeq m_F/(2|p_z|)$. Keeping in mind that we are interested in a rough order-of-magnitude estimate, we approximate the exponential by a step function, thus obtaining

$$\Delta R(p_z) \simeq -\frac{2m_f^2}{|p_z|m_H}\theta_{CP}\Theta\left(m_H - |p_z|\right)\Theta\left(|p_z| - m_F\right). \tag{3.15}$$

Next, we want to consider the net flux of left-handed particles J_0 into the unbroken phase. For this, we need to calculate the difference between the flux densities of left-handed particles j_L and their CP conjugates $j_{\bar{L}}$ such that

$$J_0 = \int \frac{\mathrm{d}^3p}{(2\pi)^3}\left(j_L - j_{\bar{L}}\right)\,. \tag{3.16}$$

For both flux densities, we have to consider transmission from the broken into the unbroken phase and reflection back into the unbroken phase [317, 318]. For concreteness, let us assume that the wall moves with velocity v_D in positive z-direction. We want to calculate the influx of left-handed fermions from the broken phase on the left ($z < 0$) into the unbroken phase inside the wall on the right ($z > 0$). We can then write, for example, for the flux density of left-handed particles

$$j_L(p_z, p_\perp) = f^{\rightarrow}(p_z, p_\perp)T^{b\to u}_L + f^{\leftarrow}(p_z, p_\perp)R^u_{R\to L} - f^{\leftarrow}(p_z, p_\perp)\,. \tag{3.17}$$

Here, $f^{\leftarrow}$ denotes the flux density of particles moving in the unbroken phase to the left (towards the wall boundary) and $f^{\rightarrow}$ that of particles moving in the broken phase to the right (also towards the wall boundary). Moreover, p_z denotes the momentum along the z-direction, i.e., perpendicular to the wall, and $p_\perp = (p_x^2 + p_y^2)^{1/2}$ is the total momentum orthogonal to the z-axis. For the third term in Eq. (3.17), which actually consists of one contribution from reflection and one from transmission, we used that $R^u_{L\to R} + T^{u\to b}_L = 1$. The same term will occur in $j_{\bar{L}}$ and thus cancels out when considering the difference $j_L - j_{\bar{L}}$.

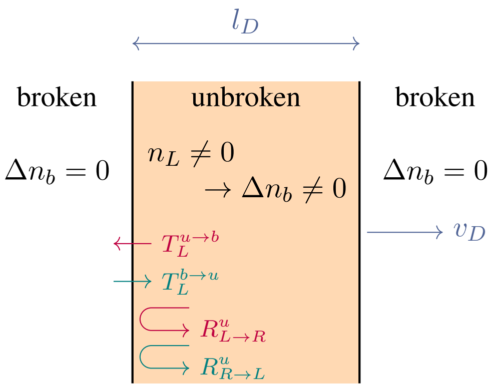


Figure 3.1: A wall of width l_D (shaded in orange), within which electroweak symmetry is restored, moves to the right with velocity v_D. Outside the wall, sphaleron transitions are exponentially suppressed, such that no net baryon number is produced ($\Delta n_b = 0$). CP-violating reflection and transmission of left-handed fermions at the trailing edge of the wall generate a chiral flux j_L (similarly for antifermions with $j_{\bar{L}}$), as defined in Eq. (3.17). Contributions to j_L entering with a positive sign are shown in teal; those entering with a negative sign are shown in purple. The chiral flux that builds at the edge of the wall diffuses in front of the trailing edge, i.e., into the symmetric phase, creating a net left-handed fermion density $n_{f,L} \neq 0$ inside the wall. This leads via electroweak sphalerons to a change in the net number density of baryons $\Delta n_b \neq 0$.

Assuming free field phase space densities and using the velocity $v_z = p_z/E$, we have in the wall frame

$$f^{\leftarrow} = \frac{|p_z|/E}{1+\exp\left(\frac{\gamma}{T}\left(E - v_D|p_z|\right)\right)}\,, \qquad f^{\rightarrow} = \frac{|p_z|/E}{1+\exp\left(\frac{\gamma}{T}\left(E + v_D\sqrt{p_z^2 - m_f^2}\right)\right)}\,, \tag{3.18}$$

where $E = (p_\perp^2 + p_z^2)^{1/2}$ is the energy in the wall frame and $\gamma = (1 - v_D^2)^{-1/2}$ is a Lorentz factor.

The expression for j_L in Eq. (3.17) simplifies tremendously if one observes that the appearing reflection and transmission coefficients are not independent of each other. In fact they satisfy

$$T_L^{b\to u} = T_{\bar{L}}^{u\to b} = 1 - R^u_{\bar{L}\to\bar{R}} = 1 - R^u_{R\to L}\,, \tag{3.19}$$

where we first used CPT invariance, then probability conservation, and finally again CPT invariance.

Doing the same for the coefficients appearing in the expression for $j_{\bar{L}}$, one obtains [311]

$$J_0 = \int_{p_z>0} \frac{d^3p}{(2\pi)^3}\left(f^{\leftarrow} - f^{\rightarrow}\right)\Delta R(p_z) \simeq -\frac{m_F^2\theta_{CP}}{2\pi^2 m_H}\int_{m_F}^{m_H} \mathrm{d}p_z \int_0^\infty p_\perp \mathrm{d}p_\perp\, \frac{f^{\leftarrow} - f^{\rightarrow}}{|p_z|}\,. \tag{3.20}$$

Suppose now that in the relevant temperature range $m_H/T \ll 1$ holds and use $m_F/m_H \ll 1$. Then the above integral evaluates to [311]

$$J_0 \simeq -\frac{v_D m_F^2 m_H \theta_{CP}}{4\pi^2}\,. \tag{3.21}$$

Similarly, we can calculate the average velocity of the chiral flux relative to the wall [311]

$$v_i \equiv \frac{\int_{p_z>0} \frac{\mathrm{d}^3p}{(2\pi)^3}\frac{|p_z|}{E}\left(f^{\leftarrow} - f^{\rightarrow}\right)\Delta R(p_z)}{\int_{p_z>0} \frac{\mathrm{d}^3p}{(2\pi)^3}\left(f^{\leftarrow} - f^{\rightarrow}\right)\Delta R(p_z)} \simeq \frac{1}{4\ln(2)}\frac{m_H}{T}\,, \tag{3.22}$$

where we expanded in leading order of v_D, m_F/m_H, $m_H/T \ll 1$. Note that we use the expansion of the above expressions just for illustrative purposes. For our final computation, we will use the full integral expressions.

In the next step, we want to compute the net chiral fermion number density inside the wall that arises due to diffusion of the chiral fermion current from the wall's edge into the wall. As before we focus on leptons, since quarks suffer, in addition to the already mentioned wash-out due to strong sphalerons, from a shorter diffusion length (for more details, see Refs. [311, 321]). The net chiral lepton number density $n_{l,L}$ will be stationary inside the wall frame, i.e., it is of the form $n_{l,L}(z,t) = n_{l,L}(z - v_D t)$ and hence satisfies $\dot{n}_{l,L} = -v_D n'_{l,L}$, where a prime denotes a derivative with respect to z. Modelling the injected current as localised at the edge of the wall $\xi_p J_0 \delta\left(z - v_D t\right)$, the diffusion equation takes the form

$$D n''_{l,L} + v_D n'_{l,L} = \xi_p J_0 \delta'(z - v_D t)\,, \tag{3.23}$$

where $D = 1/(8\alpha_W^2 T)$ is the diffusion constant, ξ_p is the persistence length. The persistence length can be estimated as $\xi_p \sim 6 D v_i$ (see Ref. [311]). This equation is solved by [69, 300]

$$n_{l,L}(z) = J_0 \frac{\xi_p}{D} e^{-\frac{v_D}{D} z} \tag{3.24}$$

for $z > 0$, and $n_{l,L}(z) = 0$ for $z < 0$. In addition to this diffusion tail inside the wall, there will be a similar diffusion tail in front of the leading edge. We do not need to consider it, though, as it is outside the wall where sphaleron processes are suppressed, so this chiral asymmetry will not contribute to the produced baryon asymmetry [69].

Making use of Eq. (3.10) and again of $\dot{n}_b = -v_D n'_b$, we can relate the chiral asymmetry to a baryon asymmetry

$$n'_b = \frac{\bar{\Gamma}_s}{v_D} n_{l,L}\,. \tag{3.25}$$

To obtain the produced net baryon density, we have to integrate the above expression from the front of the wall at l_D, where no baryogenesis has happened all the way to the trailing edge of the wall at 0, behind which no more baryons are produced. We obtain [69]

$$n_b = -\tilde{n}_{\rm nl}\left(1 - e^{-\frac{v_D}{D} l_D}\right)\,, \qquad \text{with} \qquad \tilde{n}_{\rm nl} = \frac{\bar{\Gamma}_s \xi_p}{v_D^2} J_0\,, \tag{3.26}$$

for the induced baryon number density.

For the associated baryon-to-entropy ratio, we have to divide by the entropy density given in Eq. (3.12) and find

$$\tilde{\mathcal{B}}_{\rm nl} \equiv \frac{\tilde{n}_{\rm nl}}{s} = \frac{45}{2\pi^2 g_s}\frac{\bar{\Gamma}_s \xi_p J_0}{v_D^2 T^3} \simeq -\frac{45}{8\pi^4 g_s}\frac{\bar{\Gamma}_s D}{v_D}\frac{\xi_p}{D}\theta_{CP}\left(\frac{m_F}{T}\right)^2\frac{m_H}{T}\,, \tag{3.27}$$

where the last approximate equality derives from the approximation in Eq. (3.21) and only holds as long as $m_H/T \ll 1$. Applying the above formula to lower temperatures would suggest a significant rise of $\mathcal{B}$ at late times. However, solving all the integrals for J_0 and v_i numerically shows that this behaviour is not realised and, instead, that (in the relevant temperature range with $m_\tau < T$) the function $\mathcal{B}(T)$ is slowly decreasing.

Based on Eq. (3.26), we can estimate the induced baryon-to-entropy ratio at a point in space which is passed once by a wall. For an order-of magnitude estimate at $T \sim m_H$, we employ the expansion in Eq. (3.27). Using $v_D^2/(\bar{\Gamma}_s D) \sim 10^2$, $v_D \sim 1$, $\theta_{CP} \sim 1$, $v_D l_D/D \sim 10^{-2}$, and $g_s \sim 10^2$, we find the value

$$\mathcal{B} \sim 10^{-11}\,, \tag{3.28}$$

marginally compatible with the measured baryon-to-entropy ratio in Eq. (3.1).

3.2.2 A network of walls

The above computation gives us the baryon-to-entropy ratio $\mathcal{B}$ one obtains when a wall sweeps over each point in space once. This neglects the fact that walls occur in a network, such that the entire space is populated with walls separated by some average distance L_D, and a given point in space can be passed by many walls. These walls have some curvature which will, due to the wall tension induce motion of the walls at speeds v_D of the order of the speed of light. For a network of walls, it is then possible to wash over each point in space multiple times before the temperature has decreased to T_f and the embedded walls unwind.

Assuming that embedded walls have, until they decay, the same dynamical behaviour as stable domain walls, they approach a scaling regime (see Refs. [324, 325] for simulations and Refs. [326, 327] for scaling solutions in the VOS model). In this scaling regime, the inter-wall distance changes proportionally to the horizon scale $L_D(t) \propto d_h(t) = 2t$, where in the last step, we assumed an evolution during the radiation-dominated era. We call the proportionality constant ξ the reduced correlation length and write

$$L_D(t) = 2\xi t \,. \tag{3.29}$$

To obtain an approximate result, we can take the horizon scale to change discretely after each time step that the wall requires to cross the length L_D. The time that it takes for each point in space to be passed once in the n-th time step is $\tau_n = L_D(t_n)/v_D = 2\xi t_n/v_D$. The time after the n-th passage over each point in space is accordingly

$$t_n = t_{n-1} + \tau_{n-1} = t_{n-1}\left(1 + \frac{2\xi}{v_D}\right) \,, \tag{3.30}$$

which implies

$$t_n = \left(1 + \frac{2\xi}{v_D}\right)^n t_0 \,. \tag{3.31}$$

It is important to note that t_0 is not the cosmic time today, but the time before the first passage of the walls, which is roughly the time of the electroweak phase transition.

In the radiation-dominated era, cosmic time and temperature are related via

$$\frac{1}{2t} = H(t) = \sqrt{\frac{8\pi^3 G}{90} g_\rho} T^2 \,, \tag{3.32}$$

where g_ρ is the effective number of energetic degrees of freedom. Under the approximation that $g_\rho = \text{const.}$, this translates a time t_n into a temperature

$$T_n = \frac{T_0}{\left(1 + \frac{2\xi}{v_D}\right)^{n/2}} \,. \tag{3.33}$$

We can identify T_0 as the temperature of the electroweak phase transition T_{EW} which we take for the numerical evaluation to be the $T_{\text{EW}} = 160\,\text{GeV}$ as predicted by the SM [145], which should be adequate for our order-of-magnitude estimates, even though, the value might differ in a 2HDM.

With this input, we can now compute the baryon-to-entropy ratio $\mathcal{B}_n$ after the walls swept over each point in space n times. For the case of local baryogenesis, we find

$$\mathcal{B}_{n+1} = \mathcal{B}_n \exp\left(-\frac{\bar{\Gamma}_s l_D}{v_D}\right) + \tilde{\mathcal{B}}_1\left(1 - \exp\left(-\frac{\bar{\Gamma}_s l_D}{v_D}\right)\right) \,, \qquad \text{with} \qquad \mathcal{B}_0 = 0 \,. \tag{3.34}$$

As before, l_D denotes the width of the region in which electroweak symmetry is restored around the domain wall, and $\tilde{\mathcal{B}}_{\rm l}$ is the baryon-to-entropy ratio that is produced locally at the edge of this region and given in Eq. (3.11). The previous expression describes the fact that in the n-th passage, a baryon-to-entropy ratio $\mathcal{B}_n$ was already produced. The exponential factor takes into account that during the $n+1$-st passage of a wall, this already produced baryon asymmetry is diluted due to sphaleron processes. The second term is the baryon number density newly produced in the $n+1$-st passage of a wall due to the difference between baryons produced at the trailing edge and antibaryons produced at the leading edge, which had time to decay during the wall passage.

In a similar fashion, we can find in the case of non-local baryogenesis the relation

$$\mathcal{B}_{n+1} = \mathcal{B}_n \exp\left(-\frac{\bar{\Gamma}_s l_D}{v_D}\right) + \tilde{\mathcal{B}}_{\rm nl}\left(1-\exp\left(-\frac{v_D l_D}{D}\right)\right) , \qquad \text{with} \qquad \mathcal{B}_0 = 0 , \tag{3.35}$$

where $\tilde{\mathcal{B}}_{\rm nl}$ was calculated in Eq. (3.27). While the first term in the previous equation remains unchanged compared to the first term in the local case, the second term, which describes the production mechanism, is modified. Baryogenesis now occurs, as a result of chiral leptons diffusing into the wall, throughout the entire wall and not just at its edge.

Under the very restrictive assumption that $\xi \ll 1$, which corresponds to a large number of domain walls per Hubble length, it is possible to sweep over each point in space many times without the temperature changing substantially. Let us assume that this is the case, such that $\bar{\Gamma}_s$, $\tilde{\mathcal{B}}_{\rm nl}$, $\tilde{\mathcal{B}}_{\rm l}$, D, l_D and v_D are roughly constant.[8] In Refs. [69, 301], this assumption of adiabaticity was made implicitly in order to arrive at an equilibrium expression for $\mathcal{B}$. We reemphasise that these assumptions are generally much too restrictive, and one needs to solve the previous equations numerically. They allow, however, for analytical solutions. Via a geometric series, Eqs. (3.34) and (3.35) lead to the explicit expressions

$$\mathcal{B}_n \sim \tilde{\mathcal{B}}_{\rm l}\left(1-\exp\left(-n\frac{\bar{\Gamma}_s l_D}{v_D}\right)\right) , \tag{3.36}$$

in the local case, and

$$\mathcal{B}_n \sim \tilde{\mathcal{B}}_{\rm nl}\left(1-\exp\left(-\frac{v_D l_D}{D}\right)\right)\frac{1-\exp\left(-n\frac{\bar{\Gamma}_s l_D}{v_D}\right)}{1-\exp\left(-\frac{\bar{\Gamma}_s l_D}{v_D}\right)} , \tag{3.37}$$

in the non-local case. The tilde indicates that this is, in any case, only a very rough approximation.

If $v_D l_D/D \ll 1$ and $\bar{\Gamma}_s l_D/v_D \ll 1$, we can get in the non-local case

$$\mathcal{B}_n \sim \frac{v_D^2}{\bar{\Gamma}_s D}\tilde{\mathcal{B}}_{\rm nl}\left(1-\exp\left(-n\frac{\bar{\Gamma}_s l_D}{v_D}\right)\right) , \tag{3.38}$$

which gives an enhancement factor $v_D^2/(\bar{\Gamma}_s D) \sim 10^2$ over the local case which was found in Refs. [69, 301].

After this approximate analytical estimate, let us come to the numerical evaluation of the integrals in Eqs. (3.20) and (3.22). We find that our mechanism can, for a reasonable part of the parameter space, comfortably explain the observed baryon-to-entropy ratio. The results of this numerical evaluation are shown in Fig. 3.2. The four plots depict the total baryon-to-entropy ratio $\mathcal{B}$ after the baryon production has ended as a function of the reduced average domain wall distance $\xi = L_D(t)/(2t)$ and the temperature T_f at which the embedded DWs decay and baryogenesis

---

[8] For the scaling network $v_D = \text{const.}$ is generally a good assumption while the temperature dependence of the wall width l_D is unknown.

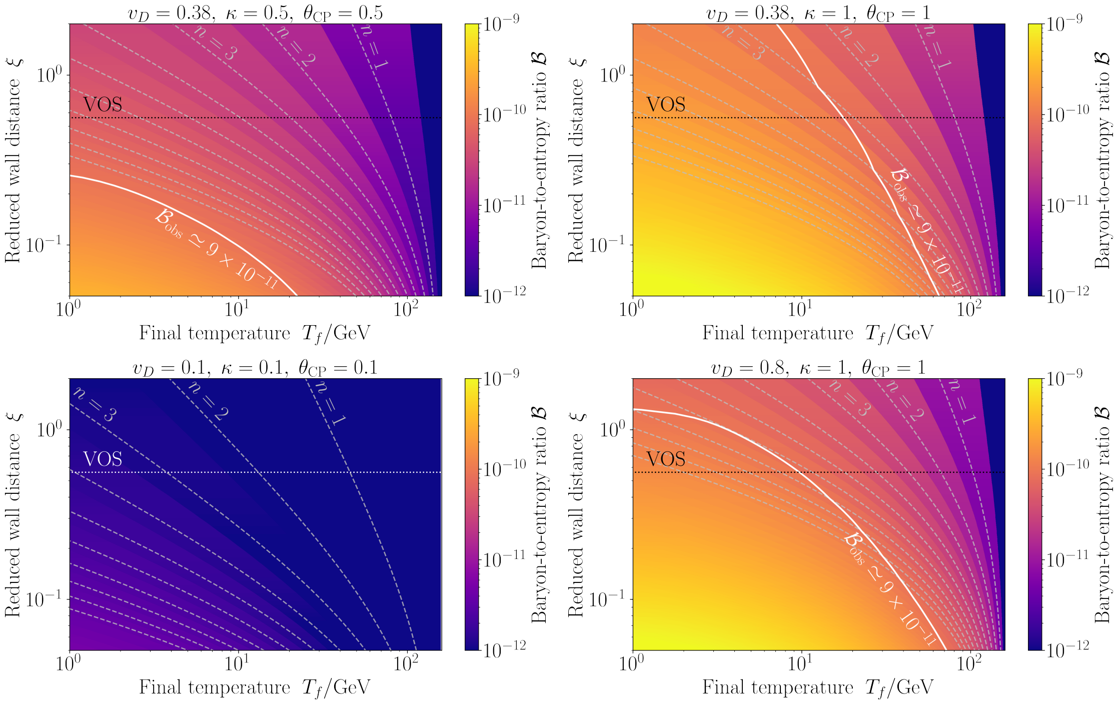


Figure 3.2: Plots showing the final baryon-to-entropy ratio obtained from Eq. (3.35) with numerically computed $\tilde{\mathcal{B}}_{\rm nl}$ given in Eq. (3.27) as a function of the reduced average domain wall distance ξ defined in Eq. (3.29) and the final temperature T_f at which the embedded walls dissolve such that baryogenesis terminates. The result is plotted for four different values of the parameters $(v_D, \kappa, \theta_{\rm CP})$. The wall velocity $v_D \simeq 0.38$ corresponds to the prediction of the velocity-dependent one-scale (VOS) model which we numerically computed on basis of Ref. [326]. The horizontal dashed line indicates the value $\xi \simeq 0.56$ was similarly computed on basis of the VOS model, and the solid white line shows the observed baryon-to-entropy ratio $\mathcal{B}_{\rm obs}$ given in Eq. (3.1). The light-grey dashed lines show how often the walls have swept over each point in space after $T_{\rm EW}$.
Figure adapted from Ref. [2].

ceases. In each plot, we consider a different set of parameters v_D, κ, and $\theta_{\rm CP}$. As a common feature of all four plots, lower ξ and lower T_f lead to a larger final baryon-to-entropy ratio $\mathcal{B}$. This is because both lead to an increase in the number of times each point in space is washed over by a wall before the embedded walls decay. The parameters κ and $\theta_{\rm CP}$ enter almost exclusively as a common prefactor $\kappa\theta_{\rm CP}$ in the equation for $\mathcal{B}$. The role of v_D is more complicated. For very low values, the domain walls wash over each point in space only a few times before they decay, leading to a small values of $\mathcal{B}$. On the other hand, too large values of v_D can reduce the baryon number produced in a single passage of a wall. The solid white line in the plots indicates the observed baryon-to-entropy ratio given in Eq. (3.1). As can be seen, for a large part of the displayed parameter space, the presented mechanism leads to a sufficient production of baryons to explain observations. It should be emphasised again that the parameter dependencies and numbers shown in Fig. 3.2 are based on order-of-magnitude estimations.

3.3 A model with embedded walls

Having seen that (embedded) domain walls can lead to an efficient mechanism of EWBG, we aim to construct a particle physics model that gives rise to the required wall solutions. More precisely, we are looking for an extension of the SM gauge symmetry group $G_{\rm SM}$ such that, when the enhanced symmetry group G breaks to $G_{\rm SM}$, embedded defects arise. Thus, we consider the symmetry breaking pattern

$$G \xrightarrow{\Phi} G_{\mathrm{SM}} = SU(3)_C \times SU(2)_L \times U(1)_Y \xrightarrow{H} SU(3)_C \times U(1)_{\mathrm{EM}}\,, \tag{3.39}$$

where $SU(3)_C$ is the gauge group of QCD and Φ is a new Higgs field that acquires a non-vanishing VEV in the first phase transition. While the vacuum manifold $\mathcal{M}$ should be path-connected to avoid topologically stable domain walls, it should allow for embedded domain wall solutions, which can be stabilised via interactions with the thermal plasma. We might, for example, have some Higgs field consisting of $n+1$ real scalars $\Phi = (\phi_1, \ldots, \phi_{n+1})$, such that its vacuum is described by

$$\mathcal{M} = \left\{ \Phi \in \mathbb{R}^{n+1} \,\middle|\, \sum_{i=1}^{n+1} \phi_i^2 = v_\Phi^2 \right\} \cong S^n\,. \tag{3.40}$$

If the fields $\phi_1, \ldots, \phi_n$ are charged under $U(1)_{\mathrm{EM}}$ while ϕ_{n+1} is uncharged, interactions of the fields with the thermal electromagnetic plasma in the early Universe can lift the potential in the direction of the charged fields, leading to the disconnected effective vacuum manifold

$$\mathcal{M}_{\mathrm{eff}} = \left\{ \Phi \in \mathbb{R}^{n+1} | \phi_1 = \cdots = \phi_n = 0 \text{ and } \phi_{n+1} = \pm v_\Phi \right\} \cong S^0\,. \tag{3.41}$$

Notice that this is only possible if n is even, since only an even number of real scalar fields can be charged under a $U(1)$ symmetry. The fact that only a single real scalar is allowed to be uncharged highlights again that the components ϕ_i must not be complex-valued.

As a simple realisation of this scenario, we propose studying a model in which the weak-hypercharge group $U(1)_Y$ is embedded in a $SU(2)_P \times U(1)_K$ gauge symmetry group

$$G = SU(3)_C \times SU(2)_L \times SU(2)_P \times U(1)_K\,. \tag{3.42}$$

Since $SU(3)_C$ remains unaffected in the chain given in Eq. (3.39), we do not have to consider it in detail, and will omit it from the following discussion. In fact, all fields we introduce in addition to the SM will be singlets under the QCD gauge group. It is noteworthy that the vacuum manifold after the entire symmetry breaking does not topologically stabilise monopoles, strings, or domain walls. Symmetry groups of a form similar to G appear in other phenomenological settings, for example, in left-right symmetric models (see, e.g., Refs. [328–330]).

Starting from a phase with the full internal symmetry, the breaking to the SM must involve the breaking $SU(2)_P \times U(1)_K \to U(1)_Y$. To obtain embedded domain walls, we realise this symmetry breaking in two steps:

$$SU(2)_P \times U(1)_K \xrightarrow{\Phi} U(1)_P \times U(1)_K \xrightarrow{\Psi} U(1)_Y\,. \tag{3.43}$$

The above pattern can be obtained by choosing Φ in the $(1,1,3)_0$ representation of G and Ψ in the $(1,1,2)_{1/2}$ representation. In fact, ignoring the $SU(3)_C$ group under which Φ and Ψ are singlets anyway, this is just the model of metastable strings we encountered in Sec. 2.3.2. For brevity, we will adapt the notations and conventions used in that section, although we focus on different aspects of the model.

First phase transition: Recall from Sec. 2.3.2 that the first symmetry breaking $SU(2)_P \times U(1)_K \to U(1)_P \times U(1)_K$ occurs as a result of Φ acquiring a non-vanishing VEV, which we chose to be of the form $\Phi_0 = v_\Phi T^3$, with T^3 being the single remaining generator of $U(1)_P$. For the plasma-stabilisation mechanism, it is essential to know the charges of the different Higgs-field components. Fortunately, they can be computed easily. The charge operator Q_P associated with the $U(1)_P$ group must be chosen such that the VEV of Φ is uncharged. It is easy to see that the (up to an arbitrary normalisation) unique charge operator which satisfies this is the unbroken generator T^3 itself: $Q_P = T^3$. To find the charges of the other components of Φ, we first have to determine the eigenstates corresponding to Q_P. For this purpose, we can define

$$T^{\pm} = \frac{1}{2}\left(T^1 \pm iT^2\right) . \tag{3.44}$$

From the commutation relation $[T^i, T^j] = \epsilon_{ijk}T^k$, it is easy to see that they satisfy $[Q_P, T^{\pm}] = \pm T^{\pm}$. Correspondingly, we find with the decomposition $\Phi = \phi_+ T^+ + \phi_- T^- + \phi_0 T^3$ that the transformation behaviour under the unbroken group $U(1)_P \subset SU(2)_P$ is

$$\phi_{\pm}(x) \longrightarrow e^{\pm i\pi^3(x)}\phi_{\pm}(x) \qquad \text{and} \qquad \phi_0(x) \longrightarrow \phi_0(x) . \tag{3.45}$$

This implies the charge assignment

$$Q_P(\phi_{\pm}) = \pm 1 \qquad \text{and} \qquad Q_P(\phi_0) = 0 . \tag{3.46}$$

At this stage, we have already obtained the plasma-stabilised embedded domain walls, which can be seen as follows. The kinetic term of the first Higgs field reads

$$\begin{aligned} \mathrm{tr}\left[(\mathcal{D}_\mu\Phi)(\mathcal{D}^\mu\Phi)\right] =& \mathrm{tr}\left(\partial_\mu\Phi\partial^\mu\Phi\right) - 2ig_P\mathrm{tr}\left(\partial_\mu\Phi[P^\mu,\Phi]\right) + \\ &+ 2g_P^2\mathrm{tr}\left(P_\mu P^\mu\Phi^2 - P_\mu\Phi P^\mu\Phi\right) . \end{aligned} \tag{3.47}$$

Substituting the VEV $v_\Phi T^3$ for Φ, we can read off that the gauge fields $P_\mu^{1,2}$ obtain effective masses $m_P^{(1)} = g_P v_\Phi$. They subsequently drop out of thermal equilibrium while the massless P_μ^3 does not, and, following Refs. [246, 247], we can apply thermal averages and approximate

$$\left\langle P_\mu^3 \right\rangle_T \simeq 0 \qquad \text{and} \qquad \left\langle P_\mu^3 P^{\mu,3} \right\rangle_T \simeq -\kappa_P T^2 , \tag{3.48}$$

with a prefactor κ_P of order one. Among the terms in Eq. (3.47), those quadratic in the gauge fields yield an effective contribution to the potential of the form

$$\frac{g_P^2\kappa_P}{2}T^2\left(\phi_1^2 + \phi_2^2\right) . \tag{3.49}$$

Let us now recall the potential of Φ from Eq. (2.62) which breaks $SU(2)_P \to U(1)_P$. Adding the new contribution from Eq. (3.49) lifts the potential in the directions of the charged fields $\phi_{1,2}$, but not in the direction of the uncharged component ϕ_3. The degeneracy between the minima is effectively lifted from an S^2 worth of minima to just two embedded points $S^0 \subset S^2$. A symmetry breaking that would be expected to give rise to monopoles instead leads, due to plasma stabilisation, to domain walls.

Second phase transition We now turn to the second stage of symmetry breaking. Recall from Sec. 2.3.2 that, for the second symmetry breaking due to Ψ, one can view Φ_0 as a fixed background field. This comes as a consequence of a large hierarchy between the scales at which Φ and Ψ acquire their non-vanishing VEVs. As we have seen, the resulting potential has minima satisfying $|\psi_1|^2 = 0$ and $|\psi_2|^2 = v_\Psi^2$. We chose the VEV such that $\psi_{2,0} = v_\Psi$. Transformations that leave ψ_2 invariant are still symmetries after the second symmetry breaking $U(1)_P \times U(1)_K \to U(1)_Y$. To make the connection to the SM, we identify $U(1)_Y$ with the group generated by weak hypercharges. The vacuum manifold of this second symmetry breaking is a circle S^1. Therefore, we expect (metastable) strings from the second phase transition.

Since ψ_2 has the charges $Q_K(\psi_2) = 1/2$ and $Q_P(\psi_2) = -1/2$, the generator of the unbroken $U(1)_Y$ subgroup must, up to an arbitrary normalisation, be of the form

$$Q_Y = Q_P + Q_K\,, \tag{3.50}$$

such that the requirement $Q_Y(\psi_2) = 0$ is satisfied. As one can easily check by using the rotation in gauge field space given in Eq. (2.72), the weak hypercharge coupling reads

$$g' = \frac{g_P g_K}{\sqrt{g_P^2 + g_K^2}}\,. \tag{3.51}$$

Similarly, we compute the weak hypercharge of the other component of the second Higgs field: $Q_Y(\psi_1) = 1$. Plasma interactions between ψ_1 and the thermalised B boson Y_μ lead to an additional mass term that lifts the potential in the ψ_1 direction. However, it does not affect the vacuum structure since, in the vacuum manifold, $\psi_1 = 0$ is already satisfied, whereas the field ψ_2 is uncharged and remains unaffected. Even when plasma effects are taken into account, the second symmetry breaking will give rise to strings.

Let us also examine how the embedded walls are affected. The field Φ is a singlet under $U(1)_K$. Hence, the weak hypercharges are identical to the Q_P charges of the components given in Eq. (3.46):

$$Q_Y\left(\phi_\pm\right) = \pm 1 \qquad \text{and} \qquad Q_Y\left(\phi_0\right) = 0\,. \tag{3.52}$$

As before, we can apply thermal averages to the B boson $\langle Y_\mu\rangle_T = 0$ and $\langle Y_\mu Y^\mu\rangle_T \simeq -\kappa_Y T^2$ with a constant κ_Y. Inserting these into Eq. (3.47) gives rise to a contribution to the potential of the form

$$\frac{g'^2\kappa_Y}{2} T^2 \left(\phi_1^2 + \phi_2^2\right)\,. \tag{3.53}$$

As in the first phase transition, plasma effects lift the potential of Φ in the directions of the charged components but not in the uncharged ones. The effective vacuum manifold remains S^0 and the embedded domain walls continue to be plasma-stabilised.

Recovering the Standard Model: To bring our model in contact with the real world, it is important to specify how the SM particles are charged under G so as to reproduce the correct charges under the SM group G_{SM}. Since the $SU(3)_C \times SU(2)_L$ part of the SM is not affected by the extension, all SM particles stay in their SM representations under these subgroups. Moreover, we can choose all SM particles to be singlets under $SU(2)_P$ and assign them $U(1)_K$ charges which are the same as their $U(1)_Y$ charges according to Eq. (3.52). Therefore, the particle content of the new theory stays the same as in the SM except for the introduction of the Higgs bosons Φ and Ψ, and the four gauge bosons of the $SU(2)_P \times U(1)_K$ symmetry, with one linear combination of these gauge fields identified as the electroweak B boson corresponding to $U(1)_Y$. The electroweak symmetry breaking $SU(2)_L \times U(1)_Y \xrightarrow{H} U(1)_{\text{EM}}$ is obtained with the usual electroweak Higgs H, which corresponds in the larger symmetry group G, according to the previous prescription, to the representation $(1, 2, 1)_{1/2}$.

For the baryogenesis mechanism described in Sec. 3.2 to be successful, we must ensure that the electroweak symmetry is restored inside the embedded domain walls. This is possible by introducing an additional term to the potential in Eq. (3.2) that couples H and Φ:

$$V_4(\Phi, H) = \sigma H^\dagger H \left(\frac{v_\Phi^2}{2} - \operatorname{tr}\left(\Phi^2\right)\right)\,, \tag{3.54}$$

where $\sigma > 0$ is a constant. On the one hand, outside the domain wall, the VEV of Φ assumes its broken-phase VEV $\mathrm{tr}(\Phi^2) = v_\Phi^2/2$ and the additional potential vanishes $V_4 = 0$. In this case, the usual electroweak theory is unaffected. On the other hand, inside the domain wall where $\Phi = 0$, one obtains $V_4(0, H) = \sigma v_\Phi^2 H^\dagger H/2$. The newly introduced potential contributes a term to the electroweak Higgs potential given in Eq. (3.2), resulting in a modified potential of the form

$$\tilde{V}_H(H) = V_4(0, H) + V_H(H) = \left(\frac{1}{2}\sigma v_\Phi^2 - 2\lambda_H v_H^2\right) H^\dagger H + \lambda_H (H^\dagger H)^2 + \lambda_H v_H^4 \,. \tag{3.55}$$

The potential has a unique minimum at $H = 0$ provided that $v_\Phi^2 \geq 4\lambda_H v_H^2/\sigma$, implying that the electroweak symmetry remains unbroken inside the embedded wall.

Third phase transition: We now have all the necessary ingredients to discuss the electroweak symmetry breaking, which leads us to the stage relevant for EWBG. Since the Higgs field H belongs to the $(1, 2, 1)_{1/2}$ representation of the full group given in Eq. (3.42), its covariant derivative takes the form

$$\mathcal{D}_\mu H \;=\; \left(\partial_\mu - igW_\mu^{\hat{a}} T_L^{\hat{a}} - \frac{i}{2} g_K K_\mu\right) H \,, \tag{3.56}$$

where $T_L^{\hat{a}} = \sigma^{\hat{a}}/2$ are the generators of the fundamental representation of $SU(2)_L$ and $W_\mu^{\hat{a}}$ denote the three gauge bosons of this group, namely, the W bosons. The Higgs field transforms according to its representation as $H(x) \to \exp(i\lambda^{\hat{a}}(x) T_L^{\hat{a}}) \exp(i\kappa(x)/2) H(x)$. From the potential in Eq. (3.2), we can read off that, outside the domain walls, the minimum of the potential is at $H^\dagger H = v_H^2/2$. We can choose the VEV correspondingly in the unitary representation as $H_0 = (0, v_H)^T$. This VEV is invariant under $SU(2)_L \times U(1)_Y$ transformations that satisfy $\lambda^{1,2}(x) = 0$ and $\lambda^3(x) = \kappa(x)$ corresponding to a $U(1)$ subgroup which we interpret as the gauge group of electrodynamics.

From the kinetic term of the electroweak Higgs, we can find that the linear combination

$$A_\mu = \frac{g' W_\mu^{\hat{3}} + g Y_\mu}{\sqrt{g^2 + g'^2}} \tag{3.57}$$

is the only gauge field that remains massless and can therefore be interpreted as the gauge boson of the unbroken group $U(1)_{\mathrm{EM}}$, i.e., the photon.

Considering the transformation behaviour of a field in an arbitrary representation of $SU(2)_L \times SU(2)_P \times U(1)_K$, one can check that the electromagnetic coupling constant is

$$e = \frac{g g'}{\sqrt{g^2 + g'^2}} \,, \tag{3.58}$$

with the electric charge operator

$$Q_{\mathrm{EM}} = T_L^{\hat{3}} + Q_Y = T_L^{\hat{3}} + Q_P + Q_K \,. \tag{3.59}$$

After electroweak symmetry breaking, the photon is in thermal equilibrium such that we can apply the thermal averages $\langle A_\mu \rangle_T \simeq 0$ and $\langle A_\mu A^\mu \rangle_T \simeq -\kappa_A T^2$ from which we find

$$\mathrm{tr}\left[(\mathcal{D}_\mu \Phi)(\mathcal{D}^\mu \Phi)\right] \supset -\frac{e^2}{2} \kappa_A T^2 \left(\phi_1^2 + \phi_2^2\right) \,. \tag{3.60}$$

Thus, even after electroweak symmetry breaking, plasma interactions continue to lift the potential of Φ in the charged $\phi_{1,2}$ directions, but not in the uncharged ϕ_3-direction. The effective vacuum manifold remains disconnected and embedded domain walls are stabilised.

It is also worth noting that the same mechanism stabilises the embedded Z-strings we discussed previously in Sec. 3.1.

In the above model, we have not yet specified how CP symmetry is violated at the domain wall boundary. To directly apply the discussion of Sec. 3.2, CP violation should be included via a 2HDM (see, e.g., [227, 315, 316]). The second Higgs doublet can be coupled to the Higgs field of the embedded walls Φ in complete analogy to the first Higgs doublet H in Eq. (3.54).

3.4 Discussion

We have investigated how the baryon asymmetry of the Universe can be generated by a network of plasma-stabilised embedded domain walls. In these walls, the electroweak symmetry is restored and fermions are reflected in a CP-violating manner at the wall edges. These walls form out-of-equilibrium configurations. As a result, the Sakharov conditions necessary for baryogenesis are fulfilled. Our results demonstrate that the observed baryon-to-entropy ratio can be reproduced within this framework. The main advantage of wall-mediated baryogenesis over string-mediated mechanisms is that the effective volume available for baryogenesis is significantly larger in the case of domain walls; for strings, it is suppressed by a factor proportional to the ratio of the string width to the Hubble length.

Since we assume that sufficient CP violation is provided by a two-Higgs-doublet model, it is worth noting that such models can give rise to a first-order electroweak phase transition. As discussed in the introduction to this chapter, electroweak baryogenesis can, in this case, also proceed via expanding walls of the broken electroweak phase into the unbroken phase. In this context, the bubble walls effectively play the role of the trailing edge of the walls in the non-local baryogenesis scenario, potentially contributing to the baryon asymmetry. However, we expect this contribution to be subleading, as a network of domain walls can sweep over a given region of space multiple times, whereas bubbles can do so only once.

While we have considered embedded domain walls, the same mechanism can also be applied to other types of walls, such as biased domain walls (see, e.g., Refs. [331–334]) or metastable domain walls (see, e.g., Refs. [33, 35]). It is only required that the walls decay before they can lead to a domain wall problem. Plasma-stabilised walls decay once plasma effects become inefficient. In this way, a domain wall problem is avoided as long as the decay temperature T_f remains above the temperature of Big Bang nucleosynthesis.

It remains to future work to determine the properties of the plasma-stabilised walls, mainly the wall width and tension. These can be of interest when considering phenomenological applications of embedded walls, e.g., the production of a stochastic GWB from wall collapse.

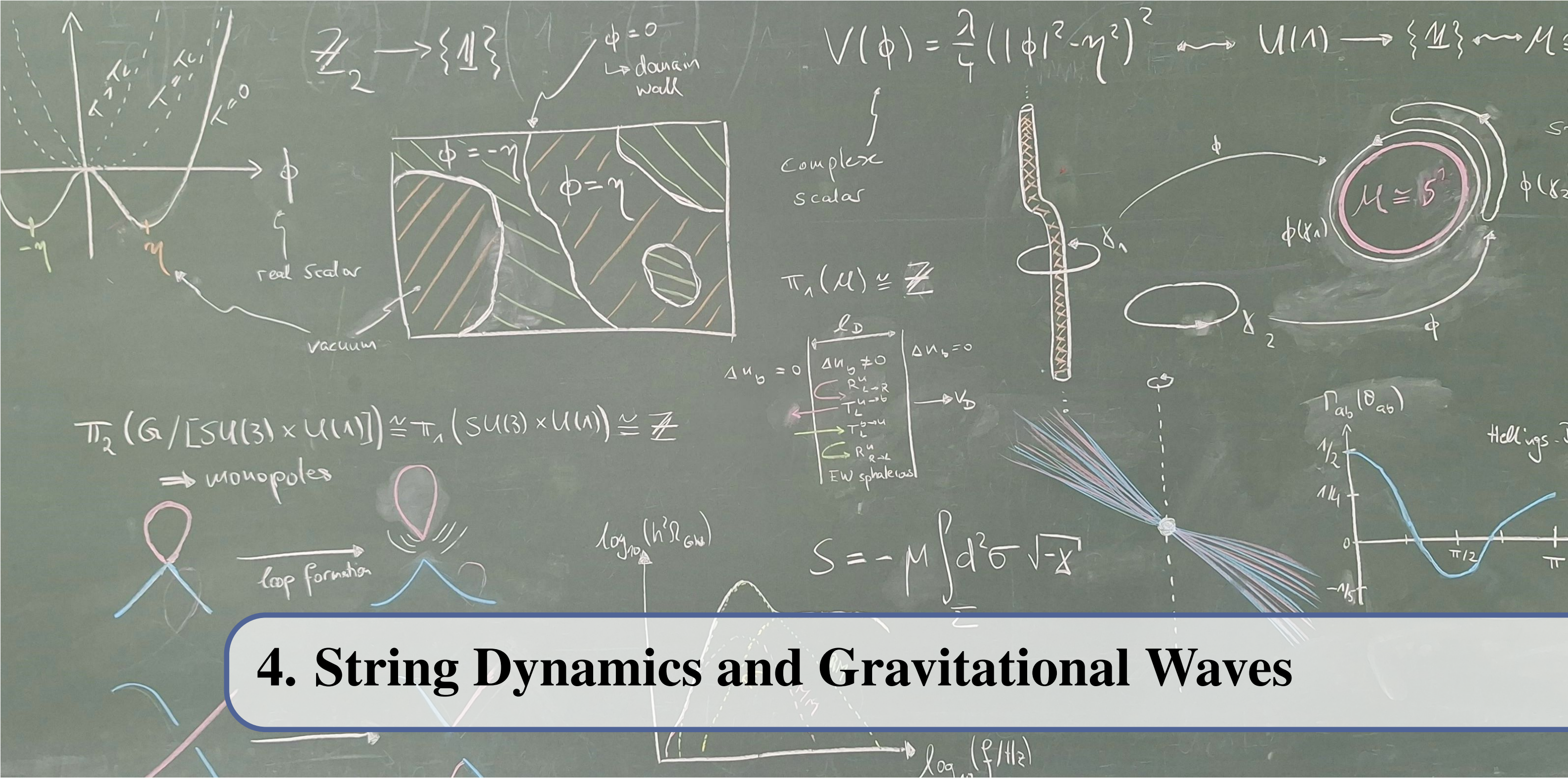
Beyond the specific case of embedded domain walls, there is a broader motivation to study plasma-stabilised defects. The fact that plasma interactions can change which defects are formed in a symmetry breaking could have far-reaching phenomenological implications, but has barely been studied. Future work should therefore investigate the plasma stabilisation more rigorously. In particular, backreaction of the defect on the plasma fields and the influence of vacuum fluctuations on the stability of embedded defects is neglected in present studies [2, 246, 247, 303]. These aspects will be necessary for any reliable determination of the temperature T_f at which the plasma stabilisation ceases. This temperature, which determines the lifetime of the defect will in turn be crucial for all observational consequences. We have, for example, already seen that the baryon-to-entropy ratio produced in the presented mechanism is very sensitive to T_f.

Similarly, understanding how defects behave as the plasma cools below this temperature is interesting, especially if remnant defects persist. For example, consider the symmetry-breaking chain $SU(2)_P \times U(1)_K \to U(1)_P \times U(1)_K \to U(1)_Y$. From this sequence, we would expect metastable strings: monopoles from the first transition and strings from the second, while the overall vacuum stabilises neither. Instead, plasma effects replace the monopoles from the first phase transition with domain walls. If, on the one hand, the temperature T_f lies below the second phase transition, one might wonder whether strings now form only as closed loops or whether they attach to domain walls. Below T_f, no energy barrier prevents the embedded walls from unwinding, and we would naively expect them to decay into monopoles, leading to the usual network of string segments with monopoles at their ends. If, on the other hand, the temperature T_f lies above the second phase transition, then monopoles may form directly at T_f.

This raises another important point. If a symmetry-breaking chain produces stable, sufficiently heavy monopoles, a monopole problem arises, as discussed around Eq. (2.50). This occurs in GUTs, where monopoles form at a temperature T_G. However, we have seen that plasma stabilisation could yield domain walls instead if the Higgs field responsible for the monopole carries the appropriate charges under the unbroken subgroup. The walls would then decay at temperature T_f back into monopoles. Although this does not affect the monopole mass, the correlation length between monopoles will have grown, most likely roughly in proportion to the Hubble scale. This dilutes the monopole density in Eq. (2.52) by a factor of $(T_f/T_G)^3$, which could be significant if plasma stabilisation lasts long enough such that T_f is sufficiently low. Nevertheless, it appears challenging for T_f to be actually low enough to fully resolve the monopole problem. Without a rigorous treatment, particularly a computation of T_f, all these plausible scenarios remain ultimately speculative.

II

Gravitational Waves from Cosmic Strings



4. String Dynamics and Gravitational Waves

This chapter, particularly Sec. 4.2 and Sec. 4.3, is based on Ref. [4]. The detailed contributions of the author to this article are outlined at the beginning of this thesis.

In this part, we are ultimately interested in the stochastic GWB produced by a network of cosmic strings. This thesis will not focus on global string loops, as they predominantly decay into Nambu–Goldstone radiation [74–78] and emit only a subleading fraction of their energy as GWs. Nevertheless, it is worth noting that global strings still produce potentially detectable GW signatures, and we refer the interested reader to Refs. [71–73] and references therein.

To investigate the gravitational radiation from local strings, we must first understand their dynamics. At first glance, this might appear feasible only through numerical simulations, and even then, it poses a significant challenge. In principle, we would need to solve the full non-linear equations of motion for a specific model, such as the Abelian Higgs model, starting from a string configuration more complicated than the ANO strings of Eq. (2.20). However, as we will see in Sec. 4.1, we can still gain a significant analytical understanding of string dynamics from an approximate action — the Nambu–Goto action. It is derived utilising the fact that the string width is microscopic while the other scales of interest for the dynamics are of cosmological size. The Nambu–Goto action leads to manageable equations of motion and we will explore properties of their solutions first in flat spacetime in Sec. 4.1.1 before considering the effects of an expanding spacetime on the string trajectories in Sec. 4.1.2. Building on this, we will introduce in Sec. 4.2 a semi-analytical model, called the velocity-dependent one-scale (VOS) model to describe the evolution of the stochastic properties of an entire network of strings.

This lays the groundwork for studying gravitational radiation from a network of strings in Sec. 4.3. First, in Sec. 4.3.1, we discuss how to compute the power emitted into GWs from a single loop and then describe key properties of a population of loops. Afterwards, in Sec. 4.3.2, we use the VOS model to derive an expression for the time dependence of the number density of string loops with a given length. Finally, in Sec. 4.3.3, we combine these results to compute the GWB spectrum of decaying string loops produced by a long-string network. As an original contribution to this section, based on Ref. [4], we also compute the effect of the finite string width which leads to corrections in the ultra-high-frequency regime of the spectrum.

4.1 The Nambu–Goto approximation

Recall from Sec. 2.2.3 that the width of cosmic strings is determined by the Compton wavelengths of the Higgs field and the gauge field that acquires a mass in the symmetry breaking. It therefore scales with the Higgs VEV η as

$$\delta \sim \eta^{-1} \simeq 2 \times 10^{-31}\,\mathrm{m} \left(\frac{\eta}{10^{15}\,\mathrm{GeV}}\right)^{-1} . \tag{4.1}$$

Strings with large widths arise at low symmetry-breaking scales and the lowest plausible scale for string production is the electroweak scale $\eta_{\mathrm{EW}} \sim 10^{2}\,\mathrm{GeV}$. Assuming coupling constants of order unity, one finds for this scale $\delta_{\mathrm{EW}} \sim 10^{-3}\, r_p$ with r_p the proton radius: Strings are microscopically thin objects. However, of cosmological interest is mostly the dynamics of strings on much larger scales, typically of the order of the Hubble scale. Correspondingly, it is not surprising that the string dynamics on cosmological scales can be well described by treating strings as effectively one-dimensional objects, only characterised in terms of their tension μ. To obtain a dimensionless quantity, this tension is often rescaled with Newton's constant $G = 1/(8\pi M_{\mathrm{Pl}}^2)$, where $M_{\mathrm{Pl}} \simeq 2.435 \times 10^{18}$ GeV is the reduced Planck mass. As discussed around Eq. (2.30), the tension roughly relates to the Higgs VEV via

$$G\mu \sim 4 \times 10^{-8} \left(\frac{\eta}{10^{15}\,\mathrm{GeV}}\right)^{2} . \tag{4.2}$$

In the following, we want to use the separation of length scales to derive an effective action on the basis of the Abelian Higgs model. Originally, this was studied in Ref. [335].

Let us recall the action of the Abelian Higgs model:

$$S_{\mathrm{AH}} = \int \sqrt{-g_y}\, \mathrm{d}^4 y \left\{ (\mathcal{D}_\mu \phi)^* \, (\mathcal{D}^\mu \phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{\lambda}{4} \left(|\phi|^2 - \eta^2 \right)^2 \right\} , \tag{4.3}$$

where $\mathcal{D}_\mu \phi = (\partial_\mu - ieA_\mu)\phi$ and g_y denotes the determinant of the metric $g_{\mu\nu}$ expressed in the y-coordinate basis. We are interested in studying how a string that appears in this model moves through space. This means we must consider string solutions that are, in contrast to the ANO solutions of Eq. (2.20), not restricted to be straight or static. At a given time, the centre of such a string, defined by the locus where $\phi = 0$, describes a one-dimensional line in space. In spacetime, this corresponds to a 2-surface called the *worldsheet*.

The two-dimensional worldsheet can be described by two coordinates σ^a with $a \in \{ 0, 1 \}$. It is embedded in spacetime and described by the coordinates $x^\mu(\sigma^a)$. We can define tangent vectors to the worldsheet as

$$e_a^\mu = \partial_a x^\mu , \tag{4.4}$$

where ∂_a is the derivative with respect to σ^a. We choose the two worldsheet coordinates such that σ^0 is timelike and σ^1 is spacelike with respect to the induced worldsheet metric

$$\gamma_{ab} = g_{\mu\nu} e_a^\mu e_b^\nu . \tag{4.5}$$

As usual, to raise indices, we use the inverse metric γ^{ab} defined by $\gamma^{ab}\gamma_{bc} = \delta^a_c$. Furthermore, for each point on the worldsheet, we can choose two spacelike vectors n_A^μ with $A \in \{ 2, 3 \}$ that are normal to the worldsheet and orthonormal to each other:

$$g_{\mu\nu} n_A^\mu e_b^\nu = 0 \qquad \text{and} \qquad g_{\mu\nu} n_A^\mu n_B^\nu = -\delta_{AB} . \tag{4.6}$$

By introducing two new coordinates ρ^A, any spacetime point near the worldsheet with coordinate y^μ can be described by the new set of coordinates $\xi^\alpha = (\sigma^a, \rho^A)$ as

$$y^\mu(\xi^\alpha) = x^\mu(\sigma^a) + \rho^A n^\mu_A(\sigma^a)\,. \tag{4.7}$$

Note that the coordinates ξ^α can become multivalued if the point described by y^μ has a larger spatial distance from the string than the string's curvature radius R.

Since the strings produced in a phase transition are typically of infinite or super-horizon length [82], it is sensible to assume that their curvature radius R is much larger than their width δ. In the plane locally orthogonal to the string, we can define polar coordinates: A radial coordinate $\rho = (\delta_{AB}\rho^A\rho^B)^{1/2}$ and an angular coordinate φ such that $\rho^A = \rho(\cos(\varphi), \sin(\varphi))^T$. We expect that for $\rho \ll R$, the string can, up to corrections of order ρ/R, still be approximately described by the ANO solutions given in Eq. (2.20) upon using the new definitions of ρ and φ.

To evaluate the action in Eq. (4.3), we not only need the field configuration in terms of our new coordinates ξ, but we also need to re-express the volume form $\sqrt{-g_y}\mathrm{d}^4y = \sqrt{-g_\xi}\mathrm{d}^2\sigma\mathrm{d}^2\rho$, where g_ξ is the determinant of the metric expressed in the ξ-coordinate basis. Using the relation between the two coordinate systems in Eq. (4.7) and the fact that in our approximation the normal vectors vary over a typical length of order R, we obtain

$$g_{ab}(\xi) = \gamma_{ab}(\sigma) + \mathcal{O}(\rho/R)\,, \qquad g_{aB}(\xi) = \mathcal{O}(\rho/R)\,, \qquad g_{AB} = -\delta_{AB}\,. \tag{4.8}$$

This, in turn, yields $\sqrt{-g_\xi} = \sqrt{-\gamma}$, with γ the determinant of γ_{ab}. We can then write to leading order in ρ/R

$$S_{\mathrm{AH}} \simeq \int \sqrt{-\gamma}\,\mathrm{d}^2\sigma \int \mathrm{d}^2\rho \left\{ (\mathcal{D}_\mu\phi)^* (\mathcal{D}^\mu\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\lambda}{4}\left(|\phi|^2 - \eta^2\right)^2 \right\}\,, \tag{4.9}$$

where the fields ϕ and A_μ need to be evaluated for solutions that are of the ANO string form discussed above. For such solutions, the expression for the Lagrangian coincides with the negative energy density. Integrating it over the plane normal to the string yields minus the string tension $-\mu$ (see Eq. (2.17)). Therefore, we obtain in total $S_{\mathrm{AH}} \simeq S_{\mathrm{NG}}$ with

$$S_{\mathrm{NG}} = -\mu \int \sqrt{-\gamma}\,\mathrm{d}^2\sigma\,. \tag{4.10}$$

This action is known as the *Nambu–Goto action* [336, 337].

The corrections of $\mathcal{O}(\rho/R)$ appearing in the integrand of the Abelian–Higgs action in Eq. (4.9) will not become large since the string width δ provides a natural truncation to the radial integration. Deviations from the Nambu–Goto action are therefore expected to be of $\mathcal{O}(\delta/R)$. Moreover, observe that this action is proportional to the area of the worldsheet in spacetime. It is the straightforward generalisation of the worldline action describing the motion of classical point particles, which is proportional to the proper length of the worldline. The Nambu–Goto action is well-known from bosonic string theory where it can be used to describe free relativistic strings. For the purpose of string quantisation, however, one usually considers the classically equivalent Polyakov action (for a recent review, see Ref. [338]).

For a more rigorous albeit less illustrative derivation of the effective action including higher order corrections, see, for example, Refs. [339–341]. It should also be emphasised that the above discussion cannot be applied to global strings as a consequence of the Goldstone modes, which lead to long-range interactions. In this case, the correct approximate action is the Kalb–Ramond action (see, e.g., Ref. [36]).

To analyse the dynamics of strings, we impose stationarity of the action in Eq. (4.10) under variations with respect to x^μ. From this, we obtain the equations of motion

$$\nabla^a e^\mu_a + \gamma^{ab}\Gamma^\mu_{\alpha\beta}e^\alpha_a e^\beta_b = 0\,, \tag{4.11}$$

where $\Gamma^\mu_{\alpha\beta}$ are the Christoffel symbols associated with the four-dimensional spacetime metric $g_{\mu\nu}$, and the covariant derivative ∇^a is taken with respect to the two-dimensional worldsheet, i.e., $\nabla^a e^\mu_a = \partial_a \left(\sqrt{-\gamma}\gamma^{ab} e^\mu_b\right)/\sqrt{-\gamma}$. We will refer to the above equations of motion as the *Nambu–Goto equations*.

4.1.1 Loop dynamics in flat spacetime

In the following, we will consider properties of solutions to the Nambu–Goto equations. Our discussion follows Ref. [36].

We are interested in the evolution of cosmic strings in an expanding Universe. However, there are two reasons why it is beneficial to start with the study of the string dynamics in flat spacetime. First, computations on a Minkowski background are simpler and therefore provide a good starting point. Second, string loops of sub-Hubble size will play an important role in our discussion. On these scales, effects of cosmic expansion are negligible, and the dynamics are effectively the same as in flat spacetime.

For the Minkowski background, we choose $x^\mu = (t, \boldsymbol{x})$ to be Cartesian coordinates such that the Christoffel symbols in Eq. (4.11) vanish. Observe that the Nambu–Goto action possesses a reparametrisation invariance $\sigma^a \to \tilde{\sigma}^a\left(\sigma^b\right)$. This invariance is not a proper symmetry but rather a gauge redundancy, namely local diffeomorphism invariance on the world-sheet.[1] We are free to pick the gauge in which the conditions

$$\sigma^0 = t\,, \qquad \dot{\boldsymbol{x}}\cdot\boldsymbol{x}' = 0\,, \qquad \dot{\boldsymbol{x}}^2 + \boldsymbol{x}'^2 = 1 \tag{4.12}$$

are satisfied. Here, $\dot{f}$ and f' denote derivatives of f with respect to σ^0 and σ^1 respectively. For brevity, we will use in the following $\sigma \equiv \sigma^1$. The second of these conditions fixes the local string velocity $\dot{\boldsymbol{x}}$ to be orthogonal to the string itself, while the third one amounts to the equation $\mathrm{d}\sigma = |\mathrm{d}\boldsymbol{x}|/(1-\dot{\boldsymbol{x}}^2)^{1/2}$. It thus relates the proper length $\mathrm{d}\sigma$ at a certain point on the string to the length $|\mathrm{d}\boldsymbol{x}|$ in the frame in which the considered point moves with velocity $\dot{\boldsymbol{x}}$.

With this gauge choice, the Nambu–Goto equation simplifies to a two-dimensional wave equation

$$\ddot{\boldsymbol{x}} - \boldsymbol{x}'' = 0\,. \tag{4.13}$$

Upon using lightcone coordinates $\sigma_\pm = \sigma \pm t$, one finds that the general solution of this wave equation reads

$$\boldsymbol{x}(t,\sigma) = \frac{1}{2}\left(\boldsymbol{a}(\sigma_-) + \boldsymbol{b}(\sigma_+)\right)\,, \qquad \text{with} \qquad \boldsymbol{a}'^2 = \boldsymbol{b}'^2 = 1\,. \tag{4.14}$$

Here, we introduced the functions $\boldsymbol{a}$ and $\boldsymbol{b}$ which are known as right- and left-movers respectively. These functions are, except for the above conditions, arbitrary.

In the chosen gauge, the stress–energy tensor of the string takes the simple form

$$T^{\mu\nu}(t,\boldsymbol{y}) = \mu \int \mathrm{d}\sigma\, \left(\dot{x}^\mu \dot{x}^\nu - x'^\mu x'^\nu\right) \delta^{(3)}\left(\boldsymbol{y} - \boldsymbol{x}(t,\sigma)\right)\,. \tag{4.15}$$

In particular, it allows to obtain the following expressions for the energy and momentum of the string:

$$E = \mu \int \mathrm{d}\sigma \qquad \text{and} \qquad \boldsymbol{P} = \mu \int \mathrm{d}\sigma\, \dot{\boldsymbol{x}}(t,\sigma)\,. \tag{4.16}$$

---

[1] Note that, in contrast to the Polyakov action, the Nambu–Goto action does not possess an additional gauge invariance under Weyl transformations.

For a string of finite length, i.e., a closed loop, we can now introduce the invariant length of the string $l = E/\mu$ such that $\sigma \in [0, l)$. Such a closed loop must satisfy at all times periodic boundary conditions $\boldsymbol{x}(t, \sigma + l) = \boldsymbol{x}(t, \sigma)$. From Eq. (4.14), it follows that

$$\boldsymbol{b}(\sigma_+ + l) - \boldsymbol{b}(\sigma_+) = -\boldsymbol{a}(\sigma_- + l) + \boldsymbol{a}(\sigma_-) = \frac{1}{\mu}\boldsymbol{P}\,, \tag{4.17}$$

where the last equality can be obtained by inserting Eq. (4.14) into Eq. (4.16).

When choosing the centre-of-mass frame of the string loop, which means that $\boldsymbol{P} = \boldsymbol{0}$, we find that $\boldsymbol{a}$ and $\boldsymbol{b}$ must be functions with period l. Combining these results implies for the string loop furthermore that

$$\boldsymbol{x}\left(t + \frac{l}{2}, \sigma + \frac{l}{2}\right) = \boldsymbol{x}\left(t, \sigma\right). \tag{4.18}$$

Accordingly, after a time of $l/2$, the loop takes on its original shape again; it is just shifted by $\sigma \to \sigma + l/2$. Therefore, the temporal period of the loop is actually not l but $l/2$.

We can now also compute the square velocity v^2 of the loop segment averaged over an oscillation period T and the length of the loop l. We find

$$v^2 = \frac{1}{Tl}\int_0^T \mathrm{d}t \int_0^l \mathrm{d}\sigma\, \dot{\boldsymbol{x}}^2 = \frac{1}{Tl}\int_0^T \mathrm{d}t \int_0^l \mathrm{d}\sigma\, \boldsymbol{x}'^2 = 1 - v^2\,, \tag{4.19}$$

where the first equality is a definition, the second equality follows from first integrating by parts over time, then using the wave equation $\ddot{\boldsymbol{x}} = \boldsymbol{x}''$ and integrating by parts over σ. The last equality uses the third condition in Eq. (4.12) and the definition of v^2. Therefore, we can see that in flat spacetime the root-mean-square (RMS) velocity of segments on the string loop reads

$$v = \frac{1}{\sqrt{2}}\,. \tag{4.20}$$

While this is a statement about an RMS velocity, parts of the string can momentarily become much faster and even reach the speed of light.

Cusps: Cusps are an important class of string features that move with luminal velocities. They appear generically on smooth strings as was shown in Ref. [342] and can be understood by a simple geometric argument. Using Eq. (4.14), we can compute the local speed of the string as

$$|\dot{\boldsymbol{x}}(t, \sigma)| = \frac{1}{2}\left|\boldsymbol{a}'(\sigma_-) - \boldsymbol{b}'(\sigma_+)\right|\,. \tag{4.21}$$

From Eq. (4.14), it follows that $\boldsymbol{a}'$ and $-\boldsymbol{b}'$ can be viewed as parametrised curves on the unit sphere S^2, referred to as the *Kibble–Turok sphere*. Due to the periodicity of these curves in the centre-of-mass frame derived from Eq. (4.17), it is clear that they are closed and satisfy

$$\int_0^l \mathrm{d}\sigma\, \boldsymbol{a}' = \int_0^l \mathrm{d}\sigma\, \boldsymbol{b}' = 0\,. \tag{4.22}$$

This means that none of the two curves described by $\boldsymbol{a}'$ and $-\boldsymbol{b}'$ can completely lie in a single hemisphere and, hence, they generically intersect. More precisely, there are values σ_a, σ_b such that $\boldsymbol{a}'(\sigma_a) = -\boldsymbol{b}'(\sigma_b)$. From Eq. (4.21), it immediately follows that at points which satisfy

$$(t, \sigma) = \left(\frac{\sigma_b - \sigma_a + nl}{2}, \frac{\sigma_a + \sigma_b + nl}{2}\right), \qquad \text{with} \qquad n \in \mathbb{Z}\,, \tag{4.23}$$

the velocity of the string becomes $|\dot{\boldsymbol{x}}| = 1$, the speed of light. Moreover, we can see that this formation of a cusp happens actually once per period $l/2$. These computations do not imply, though, that there cannot be multiple cusps per period. It is also important to emphasise that this argument breaks down if $\boldsymbol{a}'$ or $\boldsymbol{b}'$ are discontinuous. In this case, these curves are no longer closed on the Kibble sphere and intersections between them do not occur generically. Discontinuities correspond to a different string microstructure, known as a kink, and we will discuss it below.

However, before that, we want to investigate what happens in the local spatial neighbourhood of the point that reaches the speed of light. This will also clarify why it is called a cusp. Without loss of generality, we choose the point moving with the speed of light to develop at $\boldsymbol{x} = 0$ and for parameter values $(t, \sigma) = (0, 0)$. We can expand the right- and left-movers $\boldsymbol{a}$ and $\boldsymbol{b}$ around this point:

$$\boldsymbol{a}(\sigma_-) = \boldsymbol{a}'(0)\sigma_- + \frac{1}{2}\boldsymbol{a}''(0)\sigma_-^2 + \frac{1}{6}\boldsymbol{a}'''(0)\sigma_-^3 + \mathcal{O}\left(\sigma_-^4\right) , \tag{4.24}$$

$$\boldsymbol{b}(\sigma_+) = \boldsymbol{b}'(0)\sigma_+ + \frac{1}{2}\boldsymbol{b}''(0)\sigma_+^2 + \frac{1}{6}\boldsymbol{b}'''(0)\sigma_+^3 + \mathcal{O}\left(\sigma_+^4\right) . \tag{4.25}$$

For a cusp to form at $(t, \sigma) = (0, 0)$, we require $\boldsymbol{a}'(0) = -\boldsymbol{b}'(0)$, and we obtain

$$\boldsymbol{x}(0, \sigma) = \frac{\sigma^2}{4}\left(\boldsymbol{a}''(0) + \boldsymbol{b}''(0)\right) + \frac{\sigma^3}{12}\left(\boldsymbol{a}'''(0) + \boldsymbol{b}'''(0)\right) + \mathcal{O}\left(\sigma^4\right) . \tag{4.26}$$

If the term quadratic in σ does not vanish, we can choose our y axis to be parallel to $\boldsymbol{x}''(0, 0)$. When projecting onto the xy plane, this implies that $x(0, \sigma) \propto \sigma^3 + \mathcal{O}\left(\sigma^4\right)$ and that $y(0, \sigma) \propto \sigma^2 + \mathcal{O}(\sigma^3)$. The shape of this configuration is therefore to leading order $y(x) \propto x^{2/3}$ and we have plotted this function in Fig. 4.1.

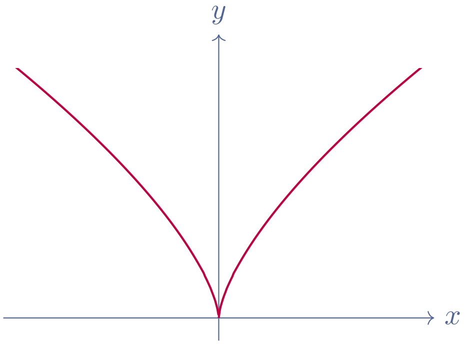


Figure 4.1: Projection of a string segment onto the xy plane with a cusp at $x = y = 0$. The vector $\boldsymbol{x}''(0, 0)$ points along the y axis, while the velocity of the string $\dot{\boldsymbol{x}}(0, 0)$ lies in the xz plane perpendicular to this direction.

Let us also figure out in which direction the velocity of the string points at the cusp. Eq. (4.14) forces the condition $\boldsymbol{a}'(\sigma_-)^2 = \boldsymbol{b}'(\sigma_+)^2 = 1$ to be satisfied in each order of σ_- and σ_+. Inserting the above expansion, this statement imposes the equations $\boldsymbol{a}'(0)^2 = \boldsymbol{b}'(0)^2$ as well as $\boldsymbol{a}'(0) \cdot \boldsymbol{a}''(0) = 0$ and $\boldsymbol{b}'(0) \cdot \boldsymbol{b}''(0) = 0$. Since at the cusp $\dot{\boldsymbol{x}}(0, 0) = \boldsymbol{b}'(0)$, the last condition implies $\boldsymbol{x}''(0, 0) \cdot \dot{\boldsymbol{x}}(0, 0) = 0$, which means that the cusp velocity is orthogonal to the direction of the cusp. For a more detailed study of cusps, we refer the reader to Ref. [81].

It should be emphasised that cusp-like features are also found in lattice field theory simulations. Although these cusps reach ultrarelativistic velocities, their speed remains below the speed of light. In fact, at the cusp, the string can overlap with itself, which locally lifts the stabilisation leading to evaporation of this part of the string (see Refs. [81, 343, 344]).

String interactions and kinks: We have studied some simple aspects of the evolution of a single string. The situation becomes more complicated if we consider self-intersecting trajectories. In this case, the finite width of the string becomes important around the string crossing and the Nambu–Goto equation breaks down. Then, one has to resort to the field-theoretical description of strings and largely relies on simulations.[2] The same is true if there are multiple strings that can intersect each other. Generally, there are three possible outcomes of an intersection event, which are depicted in Fig. 4.2. These are: (i) Nothing happens and the strings just move through each other, (ii) the strings "exchange legs" which is known as reconnection or intercommutation, or (iii) the strings entangle and junctions form.

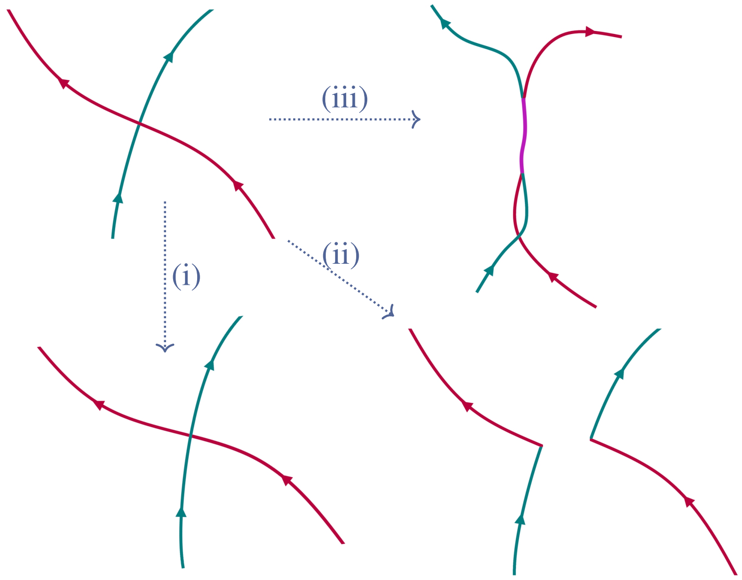


Figure 4.2: Sketch of the three possible outcomes of a string–string interaction starting from two strings in purple and green in the upper left corner. Arrows indicate the tangent vector field $\boldsymbol{x}'$ along the strings. In the string interaction the strings either (i) pass through each other, (ii) intercommute forming two new strings of the same type, or (iii) entangle and a new string type can form at their overlap (depicted in magenta). Note that the colours have no deeper meaning and solely serve to illustrate the origin of string segments after the interaction.

Let us begin with the last of these three cases, which is rarely encountered. It can occur, for example, if bound string configurations with higher winding numbers are stable, as can be the case for type-I Abelian Higgs strings (see below Eq. (2.25)). If two nearly parallel strings collide in this scenario at a low velocity, attractive forces can bind the strings together [348, 349]. Another possibility can be found if the strings are associated with a vacuum manifold $\mathcal{M}$ whose fundamental group $\pi_1(\mathcal{M})$ is non-Abelian. In this case, there exist strings that cannot pass through each other. Instead, a new third string that connects them forms at their intersection.

For the remaining two cases, numerical simulations have shown that, at least in the Abelian Higgs model, the second option, namely intercommutation, is almost always realised. This means, the intercommutation probability is $p \simeq 1$ [350, 351]. Exceptions with $p < 1$ include ultrarelativistic strings [352–354], cosmic superstrings [355], and colour-flux tubes in deconfinement–confinement phase transitions of pure Yang–Mills theories [356, 357]. Note that for colour-flux tubes and superstrings, there is indeed no interaction, while for ultrarelativistic strings, it was observed that strings always intercommute but may do so twice or more often. This can lead back to distorted versions of the initial strings and, therefore, to an outcome which is effectively the same as if they had passed through each other [358, 359].

---

[2] For sufficiently low string velocities, it is actually possible to obtain analytical results on string scattering [345] in the moduli space approximation [346, 347].

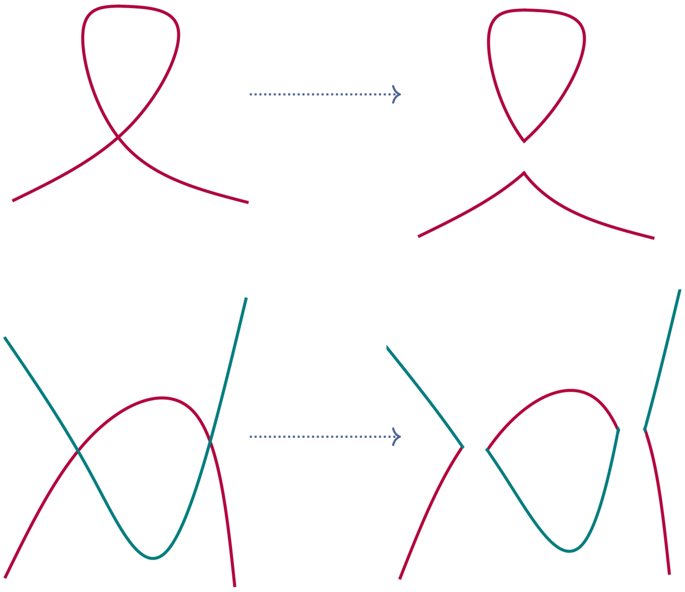


Figure 4.3: Formation of string loops from large, possibly infinite, cosmic strings via intercommutation. Loops can form when a single long string overlaps with itself, producing a small loop and a remaining long string (upper half), or when two long strings intersect, generating a small loop and again two long strings (lower half).

Let us now have a closer look at the typically realised case (ii) of intercommuting strings. In fact, based on the above discussion, this is the only case we will consider henceforth. It is easiest understood by considering Fig. 4.2. The two initial strings are illustrated with different colours: purple and teal. In the reconnection, the purple string provides two purple legs, one on each side of the interaction point. Similarly, the teal string provides two teal legs. In the intercommutation, two new strings form, each consisting of a purple leg and a teal leg.

We focus now on only one of the newly produced strings. At the moment of intersection, the initial strings will generically meet at an angle. More precisely, their tangent vectors at the interaction point will not be parallel. The tangent vector field $\boldsymbol{x}'$ along the newly formed string will, correspondingly, change discontinuously at this location.[3] Such a discontinuous feature is known as a *kink* [63, 360]. A discontinuity in $\boldsymbol{x}'$ immediately implies via Eq. (4.14) that at least one of the functions $\boldsymbol{a}'$ and $\boldsymbol{b}'$ is discontinuous, although, generically, both are. If the discontinuity in $\boldsymbol{x}'$ is located at (t_d, σ_d), the discontinuities in $\boldsymbol{a}'$ or $\boldsymbol{b}'$ will be at $\sigma_\mp^d = \sigma_d \mp t_d$ respectively. Since these functions depend only on $\sigma_\pm$, kinks will propagate along the string — for $\boldsymbol{a}$ along increasing values of σ, for $\boldsymbol{b}$ in the opposite direction. A kink formed in an intercommutation splits accordingly into two kinks which propagate in opposite directions along the loop. For a discontinuity in $\boldsymbol{a}'$, the location of the discontinuity evolves as $\sigma(t) = \sigma_d + t - t_d$ and, hence, we find that the kink propagates with a speed $|\dot{\boldsymbol{x}}| = \left|\boldsymbol{b}'\right| = 1$. Similarly, for a discontinuity in $\boldsymbol{b}'$, we have $|\dot{\boldsymbol{x}}| = |\boldsymbol{a}'| = 1$. Therefore, kinks propagate along the string with the speed of light.

We already mentioned that strings form in phase transitions with infinite or super-Hubble lengths. For our ultimate goal of computing the GWB, our focus will, however, be on string loops of sub-Hubble lengths. These loops can either form via interactions of two super-Hubble strings or via self-interactions of a single such string. Both processes are illustrated in Fig. 4.3. From our previous discussion it is clear that sub-Hubble loops are, therefore, generically formed with kinks on them [361]. In their further evolution, backreaction from particle and gravitational radiation on the loop will smoothen out the kinks. Subsequently, cusps can form on the loop. This, however,

---

[3] This is, of course, again an artefact of the Nambu–Goto approximation. For field strings, the tangent vector changes continuously, although over an extremely short distance of the order of the string's width.

is an oversimplification, and the formation and evolution of string microstructures such as kinks and cusps as well as backreaction effects are subject to ongoing research (see, e.g., Refs. [362–364]).

Beyond these generic features of strings, significant progress has been made by studying explicit analytical solutions to the Nambu–Goto equations. In particular, they have been used to study the gravitational decay of string loops. Before concluding this section, let us mention a few important examples. An explicit class of loops known as *Kibble–Turok loops* was found in Ref. [365] and subsequently generalised multiple times in Refs. [342, 366, 367]. These classes contain both self-intersecting as well as non-self-intersecting loops. To study loss of momentum due to GWs, Ref. [360] considered an asymmetric one-parameter class of loop solutions. Furthermore, a one-parameter family of typically self-intersecting loops known as *Burden loops* was constructed in Refs. [368, 369].

All the above string loops are kinkless. To provide a solution with kinks and to show that cusps are not ubiquitous, a one-parameter family of loop trajectories with four kinks but no cusps was constructed in Ref. [63]. These *Garfinkle–Vachaspati loops* are of a simple analytical form which we provide here as an example that we will encounter again later on:

$$\boldsymbol{x}(t,\sigma) = \frac{1}{2}\left(\hat{a}_0 f(\sigma_-) + \hat{b}_0 f(\sigma_+)\right) , \quad \text{with} \quad f(\sigma) = \begin{cases} \sigma - \frac{l}{4} & \text{for } \sigma \in \left[0, \frac{l}{2}\right) \\ \frac{3l}{4} - \sigma & \text{for } \sigma \in \left[\frac{l}{2}, l\right) \end{cases} , \tag{4.27}$$

where $\hat{a}_0$ and $\hat{b}_0$ are arbitrary unit vectors.

Another important class of kinky loops was described in Ref. [370] and is known as *Allen–Casper–Ottewill loops*. More general ways to construct solutions have been found in Ref. [371] and were applied in Refs. [372, 373]. For a more detailed overview of these different solutions, we refer the reader to Ref. [374].

4.1.2 String dynamics in an FLRW spacetime

Let us now examine how an expanding Universe affects the dynamics of cosmic strings. To this end, we move from Minkowski space to a spatially flat FLRW spacetime with the metric

$$\mathrm{d}s^2 = a^2(\tau)\left(\mathrm{d}\tau^2 - \mathrm{d}\boldsymbol{x}^2\right) , \tag{4.28}$$

where τ denotes conformal time and $\boldsymbol{x}$ comoving spatial coordinates, such that an observer comoving with the cosmic fluid has proper velocity $u^\mu = \left(a^{-1}, \boldsymbol{0}\right)$ in the corresponding coordinate basis.

In the following, we fix the gauge redundancy associated with the reparametrisation invariance of the worldsheet by imposing the transverse-gauge conditions

$$\sigma^0 = \tau \qquad \text{and} \qquad \dot{\boldsymbol{x}} \cdot \boldsymbol{x}' = 0 \,. \tag{4.29}$$

Note that, unlike in flat spacetime, we no longer have the freedom to choose the analogue of the third condition in Eq. (4.12). Upon fixing the above gauge conditions, the induced world-sheet metric has components

$$\gamma_{11} = a^2\left(1 - \dot{\boldsymbol{x}}^2\right) , \qquad \gamma_{12} = 0, \qquad \gamma_{22} = -a^2 \boldsymbol{x}'^2, \tag{4.30}$$

and the general Nambu–Goto Eq. (4.11) reduces to the following form[4] [375]

$$\ddot{\boldsymbol{x}} + 2\mathcal{H}\left(1 - \dot{\boldsymbol{x}}^2\right)\dot{\boldsymbol{x}} = \frac{1}{\epsilon}\left(\frac{\boldsymbol{x}'}{\epsilon}\right)' \qquad \text{and} \qquad \dot{\epsilon} + 2\mathcal{H}\dot{\boldsymbol{x}}^2\epsilon = 0 \,, \tag{4.31}$$

---

[4] The second equation follows, in the chosen gauge, immediately from the first one and is, therefore, not another independent equation of motion.

with $\dot{f}$ and f' denoting as before derivatives of f with respect to $\sigma^0 = \tau$ and $\sigma^1 \equiv \sigma$, respectively.[5] Moreover, we introduced the conformal Hubble parameter $\mathcal{H}$ and the quantity

$$\epsilon \equiv \sqrt{\frac{\boldsymbol{x}'^2}{1-\dot{\boldsymbol{x}}^2}}\,. \tag{4.32}$$

Energy, momentum, and RMS velocity: For the following discussion, three quantities are of great importance: the energy and momentum of the string as well as the root-mean-square (RMS) velocity of the different string segments. We begin by determining the energy. The stress–energy tensor of a Nambu–Goto string is directly obtained from the action in Eq. (4.10) as

$$T^{\mu\nu}(y) = \frac{\mu}{\sqrt{-g(y)}} \int \sqrt{-\gamma(x(\sigma^a))}\,\mathrm{d}^2\sigma\,\gamma^{ab}(x(\sigma^c))e^\mu_a e^\nu_b \delta^{(4)}(x(\sigma^a)-y)\,. \tag{4.33}$$

We define the energy of a single string as it is seen by an observer comoving with the cosmic fluid, namely $E = \int u_\mu u_\nu T^{\mu\nu}\,\mathrm{d}\Sigma = \int T^{\tau\tau} a^5\,\mathrm{d}^3x$, where $\mathrm{d}\Sigma = a^3\mathrm{d}^3x$ is the induced volume form on a spatial hypersurface normal to u^μ. With the explicit expression

$$T^{\tau\tau}(y) = \mu \int \mathrm{d}\sigma \sqrt{\frac{\boldsymbol{x}'^2(\tau_y,\sigma)}{1-\dot{\boldsymbol{x}}^2(\tau_y,\sigma)}}\frac{\delta^3(\boldsymbol{x}(\tau_y,\sigma)-\boldsymbol{y})}{\sqrt{-g(\tau_y,\boldsymbol{y})}}\,, \tag{4.34}$$

the energy in terms of ϵ takes the form

$$E = \mu a \int \epsilon\,\mathrm{d}\sigma\,. \tag{4.35}$$

This allows us to interpret $\mu a \epsilon$ as the energy per unit string-coordinate length. In a similar fashion, we can find the total momentum of the string as

$$\boldsymbol{P} = \mu a \int \epsilon \dot{\boldsymbol{x}}\,\mathrm{d}\sigma\,. \tag{4.36}$$

Finally, the RMS velocity of segments on the string is

$$v = \sqrt{\langle \dot{\boldsymbol{x}}^2 \rangle}\,, \tag{4.37}$$

where we introduced the average for a function f as

$$\langle f \rangle = \frac{\int f\epsilon\,\mathrm{d}\sigma}{\int \epsilon\,\mathrm{d}\sigma}\,. \tag{4.38}$$

Let us first discuss the time evolution of the string's energy and momentum. From the equations of motion in Eq. (4.31), we find directly

$$\dot{E} = \mathcal{H}\left(1-2v^2\right)E \qquad \text{and} \qquad \dot{\boldsymbol{P}} = -\mathcal{H}\boldsymbol{P}\,. \tag{4.39}$$

First, consider the evolution of the string's energy. For loops well inside the Hubble radius $l \equiv E/\mu \ll H^{-1}$, the evolution will not differ from that in flat spacetime. In Eq. (4.20), we have seen that the RMS velocity of the segments of a free loop averaged over one oscillation period is $v = 1/\sqrt{2}$. Eq. (4.39) tells us that this is consistent with a constant energy $\dot{E} = 0$ and,

---

[5] Note that, moving forward from this section, we do not use a dot to denote derivatives with respect to physical time t, but only with respect to conformal time τ.

correspondingly, a constant (physical) loop length l. This is indeed the behaviour one obtains in flat spacetime. On the other hand, for strings of super-Hubble length $l > H^{-1}$, Hubble friction will reduce the average velocity such that $v < 1/\sqrt{2}$. In this case, we find $\dot{E} > 0$ and the total energy of the string grows as a result of the string being stretched by the expanding spacetime. Let us also consider the total momentum of the loop. Eq. (4.39) tells us that the momentum redshifts as $\boldsymbol{P} \propto 1/a$ and, therefore, behaves like a point particle's momentum.

Perturbations around a straight string: To understand the impact of spacetime expansion on strings in a slightly more quantitative way, it is useful to follow Ref. [376] and consider the simplest non-trivial string solution: a comoving and straight string $\boldsymbol{x} = \boldsymbol{x}_0\sigma$. Without loss of generality, we can set $|\boldsymbol{x}_0| = 1$, as this can always be achieved by an appropriate reparametrisation of σ. We can study the dynamics of perturbations $\delta\boldsymbol{x}$ around this string

$$\boldsymbol{x}(\tau,\sigma) = \boldsymbol{x}_0\sigma + \delta\boldsymbol{x}\,(\tau,\sigma)\ , \tag{4.40}$$

which we assume to be small, i.e., $|\delta\dot{\boldsymbol{x}}| = |\delta\boldsymbol{x}'| \ll 1$. The equations of motion in Eq. (4.31) together with the gauge conditions of Eq. (4.29) give rise to the following equation in leading order of the perturbation:

$$\delta\ddot{\boldsymbol{x}} + 2\mathcal{H}\delta\dot{\boldsymbol{x}} = \delta\boldsymbol{x}'' \qquad \text{and} \qquad \boldsymbol{x}_0 \cdot \delta\dot{\boldsymbol{x}} = 0\,. \tag{4.41}$$

In a next step, we Fourier expand the perturbations,

$$\delta\boldsymbol{x}\,(\tau,\sigma) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \delta\tilde{\boldsymbol{x}}\,(\tau,k)\, e^{ik\sigma}\,, \tag{4.42}$$

where k denotes the comoving wavenumber. For the early Universe, we can consider a power-law expansion $a(\tau) \propto \tau^{\alpha}$ as is approximately realised with $\alpha = 1, 2$ during the radiation-dominated and matter-dominated eras, respectively. In this case the conformal Hubble parameter takes the form $\mathcal{H} = \alpha/\tau$, and the above equations are solved in Fourier space by

$$\delta\tilde{\boldsymbol{x}}\,(\tau,k) = \boldsymbol{x}_0\tau^{\alpha-\frac{1}{2}} J_{\alpha-\frac{1}{2}}\,(k\tau) \qquad \text{and} \qquad \boldsymbol{x}_0 \cdot \boldsymbol{c} = 0\,, \tag{4.43}$$

where $J_{\alpha-\frac{1}{2}}$ is the Bessel function of the first kind.[6]

Let us first consider super-Hubble modes $k\tau \ll 1$ and accordingly expand the Bessel function around small arguments. The resulting perturbation $\delta\tilde{\boldsymbol{x}}(\tau,k) \propto k^{\alpha-1/2}$ is time-independent. As this holds in comoving coordinates, we can see that both the physical wavelength as well as the amplitude of the perturbations grow linearly with the scale factor a. The ratio of amplitude to wavelength and correspondingly the curvature of the string remains constant for super-Hubble modes. Strings on these scales are frozen and just get conformally stretched. On the other hand, for sub-Hubble modes $k\tau \gg 1$, we find in leading order solutions of the form

$$\delta\tilde{\boldsymbol{x}}(\tau,k) \propto \tau^{-\alpha} \cos\left(\frac{\alpha\pi}{2} - k\tau\right)\,. \tag{4.44}$$

While the physical wavelength of the mode is still stretched in proportion to the scale factor, the physical amplitude of the perturbation remains, up to oscillations, constant. Hence, the ratio of amplitude to wavelength of the perturbation decreases over time and the string straightens out on sub-Hubble scales.

---

[6] The general solution contains also a term proportional to the Bessel function of the second kind. However, the term is unphysical since divergent for $k_n\tau \to 0$.

4.2 The VOS model

After this overview of the dynamical properties of individual strings and with the necessary definitions in place, we are equipped to turn to the evolution of a network of Nambu–Goto strings.

To describe the network, it is more suitable to work with the energy density of the strings $\rho \propto E/a^3$ instead of their energy given in Eq. (4.35). Eq. (4.39) can then be brought into the equivalent form

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2H\left(1+v^2\right)\rho = 0\,. \tag{4.45}$$

The network evolution is, in principle, very complex, allowing for a variety of string configurations, including different microstructures of the strings. Perhaps astonishingly, the statistical properties of the string network and their evolution can, nevertheless, be well described by a semi-analytic model called the velocity-dependent one-scale (VOS) model, which was developed primarily by KIBBLE, MARTINS, and SHELLARD in Refs. [45, 377, 378]. The apparent caveat is the need to introduce phenomenological parameters for the VOS model, which cannot directly be related to the underlying microphysics. These parameters need to be fixed by matching the VOS model to numerical simulations.

Generally, there are three different length scales that describe the stochastic properties of the strings in the network. The first scale is the correlation length, which is the distance beyond which string directions become uncorrelated. The other two scales are the curvature radius of the string and the average inter-string distance. The central assumption of the one-scale model introduced in Ref. [45] is that these three length scales are the same, and we denote them by L. To describe the network evolution, it is then useful to distinguish between long strings with $l > L$, denoted by an index "∞", and string loops with $l < L$. In this way, we can introduce the energy density of the network of long strings as

$$\rho_\infty = \frac{\mu}{L^2}\,. \tag{4.46}$$

According to Eq. (4.45), we can then write

$$\frac{\mathrm{d}\rho_\infty}{\mathrm{d}t} + 2H(1+v_\infty^2)\rho_\infty + \frac{\mathrm{d}\rho_\infty^{\mathrm{loss}}}{\mathrm{d}t} = 0\,, \tag{4.47}$$

where $\mathrm{d}\rho_\infty^{\mathrm{loss}}/\mathrm{d}t$ describes the exchange rate of energy density between the long-string network and the string loops. Since string loops are highly unlikely to recombine into long strings, but long strings will interact with each other and themselves to form string loops, this "exchange" rate is, in practice, a loss rate for the long-string network. Following Ref. [45] and using in addition to the one-scale assumption the numerically well-justified assumption that, for typical string velocities (not too close to 1), colliding strings intercommute with probability 1 (see Sec. 4.1.1), one finds

$$\frac{\mathrm{d}\rho_\infty^{\mathrm{loss}}}{\mathrm{d}t} = \frac{\tilde{c}\rho_\infty v_\infty}{L}\,. \tag{4.48}$$

Here, we implicitly defined the constant $\tilde{c}$, which is known as the "loop-chopping efficiency" and needs to be fixed by numerical simulations that find the value [378, 379]

$$\tilde{c} = 0.23 \pm 0.04\,. \tag{4.49}$$

Using Eq. (4.48) together with Eq. (4.46) in Eq. (4.47), one finds that the evolution of the universal length scale L is described by the equation

$$\frac{\mathrm{d}L}{\mathrm{d}t} - \left(1+v_\infty^2\right)HL - \frac{\tilde{c}v_\infty}{2} = 0\,. \tag{4.50}$$

Similarly, we can derive equations for the evolution of the network velocity. Differentiating Eq. (4.37) and using the Nambu–Goto equations in transverse gauge given in Eq. (4.31), one obtains

$$\dot{v}_\infty = \frac{1}{v_\infty}\left[-2\mathcal{H}v_\infty^2 + \mathcal{H}\left\langle \dot{\boldsymbol{x}}^4\right\rangle_\infty + \mathcal{H}v_\infty^4 + \left\langle \frac{\boldsymbol{x}'' \cdot \dot{\boldsymbol{x}}}{\epsilon^2}\right\rangle_\infty\right], \tag{4.51}$$

where we used the fact that $\boldsymbol{x}' \cdot \dot{\boldsymbol{x}} = 0$. Furthermore, we now make the approximation $\left\langle \dot{\boldsymbol{x}}^4\right\rangle_\infty = \left\langle \dot{\boldsymbol{x}}^2\right\rangle_\infty^2 = v_\infty^4$, which was numerically verified in Ref. [377]. To get rid of the second derivative of $\boldsymbol{x}$ in the last term of Eq. (4.51), let us first replace the derivatives with respect to σ by derivatives with respect to the physical length s along the string. They are related by $\mathrm{d}s/\mathrm{d}\sigma = |\boldsymbol{x}'|$. We can write $\mathrm{d}^2\boldsymbol{x}/\mathrm{d}s^2 = a\,\hat{\boldsymbol{u}}/R$ for some unit vector $\hat{\boldsymbol{u}}$ such that R is by definition the curvature radius. Recall now that by the one-scale assumption $R = L$. Defining furthermore $k \equiv \left\langle \left(1-\dot{\boldsymbol{x}}^2\right)\dot{\boldsymbol{x}}\cdot\hat{\boldsymbol{u}}\right\rangle_\infty / \left(v_\infty\left(1-v_\infty^2\right)\right)$, we can rewrite the differential equation for the network velocity as

$$\frac{\mathrm{d}v_\infty}{\mathrm{d}t} = \left(1-v_\infty^2\right)\left[\frac{k}{L} - 2Hv_\infty\right]. \tag{4.52}$$

The newly introduced k can be considered quantifying the string's microstructure [377]. In order to solve the coupled system of differential Eqs. (4.50) and (4.52), we need, in addition to the knowledge about the background evolution of the FLRW spacetime given by H, information about the phenomenological parameter k. This was extensively studied in Ref. [378], which resulted in an expression that interpolates between the expressions obtained from simulations and known analytical string solutions for the non-relativistic and ultra-relativistic regimes. This expression reads

$$k(v_\infty) = \frac{2\sqrt{2}}{\pi}\left(1-v_\infty^2\right)\left(1+2\sqrt{2}v_\infty^3\right)\frac{1-8v_\infty^6}{1+8v_\infty^6}. \tag{4.53}$$

The evolution of the string network is now solely described in terms of the variables L and v_∞.

The VOS equations have attractor solutions which are scaling solutions, i.e., $L \propto t$ and $v_\infty =$ const., if the scale factor follows a power-law evolution $a(t) \propto t^\beta$ with $\beta \in (0,1)$. Denoting $L = \xi t$,

$$\xi_\beta^2 = \frac{k(v_\beta)\left(k(v_\beta)+\tilde{c}\right)}{4\beta\left(1-\beta\right)} \qquad \text{and} \qquad v_\beta^2 = \frac{k(v_\beta)}{k(v_\beta)+\tilde{c}}\frac{1-\beta}{\beta} \tag{4.54}$$

give rise to stable fixed point solutions of the VOS equations [378]. For our analytical calculations, we require the values of the reduced correlation length ξ and the average velocity v_∞ during radiation domination ($\beta = 1/2$) and matter domination ($\beta = 2/3$). Solving the above scaling equations, one finds

$$\xi_r = 0.27\,, \qquad v_r = 0.66\,, \tag{4.55a}$$

$$\xi_m = 0.63\,, \qquad v_m = 0.58\,. \tag{4.55b}$$

The numerical solution to the VOS equations (4.50) and (4.52) is shown in Fig. 4.4. The necessary knowledge about the time dependence of the scale factor is obtained by solving the Friedmann equation numerically, using a model in which

$$H(a) = H_0\sqrt{\Omega_\Lambda + \Omega_m\left(a/a_0\right)^{-3} + \rho_r(a)/\rho_{\mathrm{crit}}}\,, \tag{4.56}$$

with Hubble constant $H_0 = 67.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and density parameters $\Omega_\Lambda = 0.685$ and $\Omega_m = 0.315$ [14]. We use in our numerical computation the time evolution of $\rho_r(a)$ based on tabulated

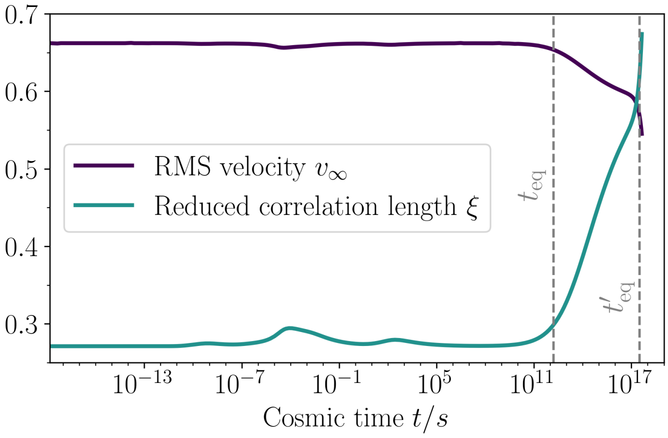


Figure 4.4: Time evolution of the reduced correlation length ξ (teal) and the RMS velocity v_∞ (deep purple) describing the long-string network. The evolution is obtained by numerically solving the coupled differential equations (4.2). The vertical dashed lines indicate matter–radiation equality $t_{\rm eq}$ and cosmological constant–matter equality $t'_{\rm eq}$.

data for temperature T, energetic DOFs g_ρ and entropic DOFs g_s from [380]. From Fig. 4.4, it is evident that replacing the numerical solution with the constant scaling solutions given in Eq. (4.55) is a good approximation during radiation domination, but not during the matter-dominated phase. This is because, after matter–radiation equality, the network does not settle fast enough in the new attractor solution for matter domination before the onset of dark-energy domination. These simplifications will be relevant to our analytical modelling of the GWB spectrum from decaying string loops in Ch. 6. Luckily, they only lead to small deviations in the GWB spectrum, as we will see in Sec. 6.1.5.

There is an important limitation to the above discussion. For all the computations, we neglected the fact that strings do not move in empty space but are surrounded by a thermal plasma with which they can interact. This leads to thermal friction. For local strings, the most relevant contribution to friction is a consequence of the magnetic flux $\mathbf{\Phi}$ along the string. A particle with charge q under the gauge field responsible for the string's magnetic flux tube experiences a phase shift of $q\mathbf{\Phi}$ as it traverses the string. Particles for which $\nu = q\mathbf{\Phi}/(2\pi)$ is not an integer will have a non-vanishing scattering cross section. This effect was computed in Refs. [381–383]. In Ref. [384], it was subsequently shown that the Hubble friction term in the string equations of motion, in particular for the long-string network in Eqs. (4.50) and (4.52), needs to be replaced via the substitution

$$H \longrightarrow H + \frac{\gamma T^3}{2\mu}\,, \qquad \text{where} \qquad \gamma = \frac{2\zeta(3)}{\pi^2}\sum_i b_i \sin^2(\pi\nu_i)\,. \tag{4.57}$$

Here, the index i runs over the particle species which are relativistic at the temperature T of the thermal bath. Moreover, $b_i = 1$ for bosons and $b_i = 3/4$ for fermions, and the value of the Riemann zeta function is $\zeta(3) \simeq 1.20$. At high temperatures, the string motion is, therefore, strongly suppressed by thermal friction, which severely reduces loop production. The network is frozen and merely undergoes conformal stretching [377, 385]. Only once the thermal friction drops below the Hubble friction, the above scaling solutions become applicable, although loop production can become significant earlier due to the large number density of strings (see Ref. [386] for a recent discussion on this point). We will revisit the discussion of thermal friction in Sec. 4.3.3 and in Ch. 5.

4.3 Gravitational radiation from strings

In this section, we are going to derive the expression for the GWB spectrum emitted from decaying string loops, which form in interactions of long strings. For this, we require information about (i) the gravitational radiation from individual loops and (ii) the ensemble properties of a network of cosmic strings.

4.3.1 Gravitational waves from individual loops

Let us first discuss how the power of gravitational radiation emitted by a single cosmic string loop can be computed. As we are considering loops of sub-Hubble size, it is appropriate to work with the flat spacetime results on loops. Due to the relativistic motion of the strings, applying the quadrupole formula is clearly inappropriate. Instead, we follow Ref. [387] and consider a general source which oscillates with period T and has stress–energy strength tensor

$$T_{\mu\nu}(t,\boldsymbol{x}) = \sum_{j=1}^{\infty} \int \mathrm{d}^3\boldsymbol{k}\, \tilde{T}_{\mu\nu}(\omega_j,\boldsymbol{k}) e^{-i\omega_j t + i\boldsymbol{k}\boldsymbol{x}} + \text{c.c.} \tag{4.58}$$

Here, c.c. denotes the complex conjugate of the preceding term and $\omega_j = 2\pi j/T$ is the angular frequency of the j-th harmonic mode. As we saw previously in Sec. 4.1.1 the string loop in its centre-of-mass frame is indeed a time-periodic source with $T = l/2$ and the above expansion is applicable. This assumes, of course, that the string length does not change significantly over the time T, which is an assumption we will confirm later. For general sources of this form, it was shown in Ref. [387] that the total emitted power $P = \sum_{j=1}^{\infty} P_j$ is of the form

$$P_j = \frac{G}{\pi}\omega_j^2 \oint_{S^2} \mathrm{d}^2\hat{x}\, \left(\tilde{T}^*_{\mu\nu}\left(\omega_j, \omega_j\hat{\boldsymbol{x}}\right) \tilde{T}^{\mu\nu}\left(\omega_j, \omega_j\hat{\boldsymbol{x}}\right) - \frac{1}{2}\left|\tilde{T}^{\mu}{}_{\mu}\left(\omega_j, \omega_j\hat{\boldsymbol{x}}\right)\right|^2 \right) . \tag{4.59}$$

Starting from this expression, it was shown in Ref. [369] that for an expansion into left- and right-movers as in Eq. (4.14), the total power takes the simpler form

$$P_j = 8\pi G\mu^2 j^2 \oint_{S^2} \mathrm{d}^2\hat{x}\, \left(|F_j(\hat{\boldsymbol{n}}_1,\hat{\boldsymbol{n}}_1) - F_j(\hat{\boldsymbol{n}}_2,\hat{\boldsymbol{n}}_2)|^2 + |F_j(\hat{\boldsymbol{n}}_1,\hat{\boldsymbol{n}}_2) + F_j(\hat{\boldsymbol{n}}_2,\hat{\boldsymbol{n}}_1)|^2 \right) , \tag{4.60}$$

where

$$F_j(\hat{\boldsymbol{n}}_a,\hat{\boldsymbol{n}}_b) = \frac{1}{l^2}\int_0^l \mathrm{d}\sigma_- \int_0^l \mathrm{d}\sigma_+ \,\left(\boldsymbol{a}'(\sigma_-)\cdot\hat{\boldsymbol{n}}_a\right)\left(\boldsymbol{b}'(\sigma_+)\cdot\hat{\boldsymbol{n}}_b\right) e^{i\frac{\omega_j(\sigma_+ - \sigma_-) - \hat{\boldsymbol{x}}\cdot(\boldsymbol{a}+\boldsymbol{b})}{2}} , \tag{4.61}$$

and $\hat{\boldsymbol{n}}_{1,2}$ are two unit vectors such that $\hat{\boldsymbol{n}}_1 \cdot \hat{\boldsymbol{n}}_2 = \hat{\boldsymbol{n}}_{1,2} \cdot \hat{\boldsymbol{x}} = 0$ (see also Ref. [388]).

Observe that the power in Eq. (4.60) is of the form $\Gamma_j G\mu^2$ where Γ_j is a numerical prefactor that depends on the loop configuration under consideration. Denoting the overall prefactor $\Gamma = \sum_{j=1}^{\infty}\Gamma_j$, such that Γ_j/Γ is the fraction of the total power in gravitational radiation that comes from the j-th oscillation mode of string loops, we can write

$$P = \Gamma G\mu^2 \,. \tag{4.62}$$

The first numerical and analytical studies using somewhat realistic loop configurations, namely Kibble–Turok and Burden loops, found $\Gamma \gtrsim 50$ and typical values of $\Gamma \sim 100$ [360, 369]. As mentioned previously, these loops are kinkless. Radiation from kinky strings, namely Garfinkle–Vachaspati loops, was investigated in Ref. [63]. Inserting the left- and right-movers of these loops given in Eq. (4.27) by means of Eq. (4.61) into Eq. (4.60), one finds the analytical expression

$$\Gamma = \frac{32}{\sin^2(\alpha)}\left((1+\cos(\alpha))\ln\left(\frac{2}{1+\cos(\alpha)}\right) + (1-\cos(\alpha))\ln\left(\frac{2}{1-\cos(\alpha)}\right) \right) , \tag{4.63}$$

where $\cos(\alpha) = \hat{\boldsymbol{a}}_0 \cdot \hat{\boldsymbol{b}}_0$. This formula does not apply to loops satisfying $\hat{\boldsymbol{a}}_0 \cdot \hat{\boldsymbol{b}}_0 = \pm 1$. In this case, the Garfinkle–Vachaspati loops describe an oscillating double line which leads to a divergent power when naively applied. Note that this is unphysical for two reasons. First, our assumption that gravitational backreaction on loops is negligible clearly breaks down if the emitted power becomes too large. Second, the Nambu–Goto approximation is not applicable to a string overlapping with itself. A line-like loop configuration would not be topologically stabilised anywhere, could simply unwind, and decay into massive particle radiation. Another class of solutions, namely Allen–Casper–Ottewill loops, was analysed in Ref. [370] and values of Γ comparable to those in Refs. [360, 369] were found. As shown in Ref. [389], this remains true if gravitational backreaction is taken into account.

For the purpose of studying the GWB of an entire network of strings, we are, however, not interested in the power emitted by a single loop. We are instead interested in the average power emitted by a realistic population of loops formed from a network of long strings. In Ref. [390], a large number of loops were obtained from a network simulation and time-evolved taking spacetime expansion effects into account. It was found that during all times the loop population radiates with a fairly peaked average of

$$\Gamma \simeq 50\,. \tag{4.64}$$

This value is independent of time and loop length. More recent simulations in Ref. [363, 364] account for gravitational backreaction on the loops and find that, over the lifetime of the loop, Γ actually decreases and assumes for most of this time values $\Gamma > 50$. Neglecting this effect can lead to deviations ranging from a few percent up to 30%. In this thesis, we are primarily interested in the identification of qualitatively new features of GWB spectra and more generally in the analytical modelling of GWB spectra from strings. For these purposes, the deviations are acceptably small. We will henceforth use $\Gamma = 50$ as a numerical reference value.

Apart from the total prefactor Γ, knowledge of the distribution of the power over the different oscillation modes is important, i.e., the dependence of Γ_j on the mode number j. Depending on the microstructure of the string loops, various scalings of Γ_j in the limit of large mode numbers have been found. If the radiation comes from cusps, kinks, or kink–kink collisions, one observes the scaling $\Gamma_j \propto j^{-4/3}$, $j^{-5/3}$, or j^{-2}, respectively [84, 360, 391]. Henceforth, we will parametrise

$$\Gamma_j = \frac{\Gamma}{H^{(q)}_{n_{\max}}} j^{-q}. \tag{4.65}$$

Here, $H^{(q)}_{n_{\max}} = \sum_{j=1}^{n_{\max}} j^{-q}$ denotes the $n_{\max}$-th generalised harmonic number of order q, ensuring proper normalisation. The reason for truncating the sum at $n_{\max}$ is that, in general, the sum to infinity cannot be evaluated analytically and numerical evaluation requires a finite $n_{\max}$. This can introduce artificial modifications of the spectra at extremely large frequencies, on which we will comment where relevant. The choice $q = 4/3$ matches the results of numerical simulations [390] very well, in particular for $j \gg 1$, and we adopt it as our benchmark value. The recent simulations including gravitational backreaction in Ref. [364] find that a constant value of q cannot completely reproduce their GWB spectra. This effect is included in the previously mentioned deviations if Γ is assumed to be constant.

4.3.2 Loop evolution

In the next step, we want to describe the evolution of the network properties of string loops produced from long-strings. Let us first consider individual loops. Over time, the loops lose a significant fraction of their energy due to the emission of GWs. We will neglect all other possible subdominant decay channels for now and return to this point in Ch. 5.

Since a string loop of length l has the energy $E = \mu l$ and the power of emission is $P = \Gamma G\mu^2$, one can directly conclude that a loop shrinks at a constant rate:

$$\frac{\mathrm{d}l}{\mathrm{d}t} = -\Gamma G\mu\,. \tag{4.66}$$

Hence, a loop which had at time t' a length l', has at time t the length

$$l = l' - \Gamma G\mu\left(t - t'\right)\,. \tag{4.67}$$

In particular, we can see that during one oscillation period $T = l/2$, the length of the loop is reduced by a relative amount $\Delta l/l = \Gamma G\mu/2 \ll 1$ where we used that $G\mu \ll 1$.[7] This actually validates the assumption we used in the previous section: The loop length stays approximately constant over the duration of an oscillation period.

The statistical information we need from the loops to describe the produced GWB is contained in the function $n(l,t)\,\mathrm{d}l$, which characterises the number density of string loops within the length interval $[l, l + \mathrm{d}l]$ at time t. To utilise the VOS model, it is useful to express this number density in terms of the loop production function $f(l,t)$ such that [394]

$$n(l,t) = \int_{t_{\mathrm{ini}}}^{t} f\left(l'(l,t,t'),t'\right)\left(\frac{a\left(t'\right)}{a\left(t\right)}\right)^3 \mathrm{d}t'\,. \tag{4.68}$$

Correspondingly, $f\left(l,t\right)\,\mathrm{d}l\,\mathrm{d}t$ is the number density of loops produced during the time interval $[t, t + \mathrm{d}t]$ within the length interval $[l, l + \mathrm{d}l]$. The lower integration boundary t_{ini} can be viewed as the time when loop production from the long string network becomes significant. We will discuss it in more detail at the end of Sec. 4.3.3. Since we previously assumed that the energy lost by the network of long strings is transferred into the production of string loops (see Eq. (4.47)), we have to demand that [64]

$$\frac{1}{\gamma}\frac{\mathrm{d}\rho_\infty^{\mathrm{loss}}}{\mathrm{d}t} = \mu\int_0^{L(t)} l f\left(l,t\right)\,\mathrm{d}l\,. \tag{4.69}$$

Here, we included a Lorentz factor γ accounting for the fact that loops are created with a non-vanishing centre-of-mass velocity. Therefore, a fraction of the energy of the long-string network goes into the kinetic energy of the loops and is subsequently lost by redshifting [377]. To fix the loop production function, we assume that all loops are created at all times with a length l_*, corresponding to a fixed fraction α of the one scale L:

$$l_*(t) = \alpha\,L(t)\,. \tag{4.70}$$

This assumption is well justified by the numerical network simulations in Ref. [47], which show that the distribution of $l_*(t)/t$ is sharply peaked around a value of 0.1, such that we can fix $\alpha\xi(t) \simeq 0.1$ to a good approximation. We can then write the loop production function as $f(l,t) = \mathcal{N}\delta\left(l - \alpha L(t)\right)$. Using Eq. (4.69) together with Eq. (4.48), we can, in principle, fix the normalisation $\mathcal{N}$. In this way, we would, however, introduce a large error. This is because, while the majority of string loops are produced with lengths close to $l_*(t)/t \simeq 0.1$, most of the energy of the long strings still goes into loops with $l_*(t)/t < 0.1$, namely in the form of kinetic energy. The error introduced by approximating the distribution of string lengths at production by a delta distribution was estimated in Ref. [47] and can be accounted for by an additional efficiency

---

[7] As we will see in Ch. 7, the currently strongest constraints on the string tension come from PTAs with $G\mu \lesssim 10^{-10}$ [1]. Weaker, but sufficiently strong bounds for the above statement can also be found without invoking GWs, but instead CMB temperature anisotropies, showing that $G\mu \lesssim 10^{-7}$ [51, 392, 393].

factor $\mathcal{F} \sim 0.1$ into the loop production function. The average initial velocity of the remaining largest loops in the network has been determined in simulations to be $\simeq 1/\sqrt{2}$ [395], such that we can set $\gamma \simeq \sqrt{2}$. Combining these results, one arrives at a loop production function of the form [394]

$$f(l,t) = \mathcal{F}\,\frac{\tilde{c}\,v_\infty(t)}{\sqrt{2}\,l\,L^3(t)}\,\delta\left(l - \alpha L(t)\right)\,. \tag{4.71}$$

We can now use this expression for the loop production function to calculate via Eq. (4.68) the loop number density and find[8]

$$n(l,t) = \Theta(t-t_*)\Theta(t_*-t_{\rm ini})\,\underbrace{\mathcal{F}\frac{\tilde{C}(t_*)}{\alpha\xi(t_*)t_*^4}\left[\Gamma G\mu + \alpha\xi(t_*) + \alpha t_*\frac{\mathrm{d}\xi}{\mathrm{d}t}(t_*)\right]^{-1}\left(\frac{a(t_*)}{a(t)}\right)^3}_{=\tilde{n}(l,t)}\,. \tag{4.72}$$

Here, t_* is implicitly defined as the solution to

$$l - \Gamma G\mu(t_* - t) - \alpha\xi(t_*)t_* = 0\,, \tag{4.73}$$

and since $\alpha\xi(t_*)t_*$ is the length of the loop at its birth, t_* is the corresponding time of birth. Note that t_* is a function of l and t. The two Heaviside step functions that occur in Eq. (4.72) arise automatically from the integration and ensure that the loops contributing to the loop number density at time t were already produced ($t > t_*$) and were produced after $t_{\rm ini}$. Moreover, it is often convenient to introduce the prefactor

$$\tilde{C}(t) = \frac{\tilde{c}v_\infty(t)}{\sqrt{2}\xi^3(t)}\,. \tag{4.74}$$

The general form of the number density is useful for obtaining numerical results, but it is intractable for analytical computations. However, it can be simplified by approximating the expansion history of the Universe following Ref. [4]. In particular, we will assume that the Universe is at early times radiation-dominated, and, after a time $t_{\rm eq}$, is purely matter-dominated until today t_0. While rough, this approximation allows a precise calculation of the GWB, as will be demonstrated later. In this spirit, we artificially introduce a factor

$$1 = \left[\Theta(t-t_{\rm eq}) + \Theta(t_{\rm eq}-t)\right]\left[\Theta(t_*-t_{\rm eq}) + \Theta(t_{\rm eq}-t_*)\right]\,, \tag{4.75}$$

which allows us to decompose the number density as

$$n(l,t) = n_{\rm rr}(l,t) + n_{\rm rm}(l,t) + n_{\rm mm}(l,t)\,, \tag{4.76}$$

where

$$n_{\rm rr}(l,t) = \Theta(t_{\rm eq}-t)\Theta(t-t_*)\Theta(t_*-t_{\rm ini})\tilde{n}(l,t)\,, \tag{4.77a}$$
$$n_{\rm rm}(l,t) = \Theta(t-t_{\rm eq})\Theta(t_{\rm eq}-t_*)\Theta(t_*-t_{\rm ini})\tilde{n}(l,t)\,, \tag{4.77b}$$
$$n_{\rm mm}(l,t) = \Theta(t-t_*)\Theta(t_*-t_{\rm eq})\tilde{n}(l,t)\,. \tag{4.77c}$$

---

[8] The loop number density was in this form derived in Ref. [396], however, without including the Heaviside functions. Their relevance will become clear in Ch. 6. The loop number densities we use are derived from the above loop production functions matched to those extracted from simulations in Ref. [47]. Another loop number distribution was developed in Ref. [397] based on simulations in Ref. [398]. Analytical results in this framework have, e.g., been developed in Refs. [399, 400]. A critical comparison of the two approaches can be found in Refs. [401, 402]. In our discussion, we stick to the former approach, which is also supported by more current network simulations [364] taking gravitational backreaction into account.

As we will see in a moment, the time t at which the loop number density is evaluated plays the role of the emission time of gravitational radiation from the respective loops. Hence, the above splitting distinguishes between loops that were born during radiation domination and emitted GWs during the radiation-dominated era (RR), loops that were produced during radiation domination and emitted GWs during matter domination (RM), and loops produced and emitting during the matter-dominated era (MM).

Consider the case that t_* is a time during which $a(t_*) \propto t_*^\beta$. In this case, one finds that

$$t_*(l,t) = \frac{l + \Gamma G\mu t}{\alpha\xi_\beta + \Gamma G\mu}\,, \tag{4.78}$$

and if, moreover, t is also a time during which $a(t) \propto t^\beta$, one obtains [403]

$$\tilde{n}_{\beta\beta}(l,t) = \mathcal{F}\frac{\tilde{C}_\beta}{\alpha\xi_\beta}\frac{(\alpha\xi_\beta + \Gamma G\mu)^{3(1-\beta)}}{t^{3\beta}(l + \Gamma G\mu t)^{4-3\beta}} \equiv \frac{C_\beta}{t^{3\beta}(l + \Gamma G\mu t)^{4-3\beta}}\,. \tag{4.79}$$

Here, we defined

$$C_\beta = \mathcal{F}\frac{\tilde{C}_\beta}{\alpha\xi_\beta}(\alpha\xi_\beta + \Gamma G\mu)^{3(1-\beta)}\,, \qquad \text{with} \qquad \tilde{C}_\beta = \frac{\tilde{c}}{\sqrt{2}}\frac{v_\beta}{\xi_\beta^3}\,, \tag{4.80}$$

where v_β and ξ_β given in Eq. (4.54). This covers the number densities of loops that are both produced and decay during radiation given in Eq. (4.77a) or matter domination in Eq. (4.77c) with $\beta = 1/2$ and $\beta = 2/3$, respectively. Finally, we have to consider the number density of loops that were produced during radiation domination but only decay during matter domination. Fortunately, we can, in this case, just evaluate the expression for $n_{\mathrm{rr}}(l,t)$ at t_{eq} and then redshift the expression to later times, i.e.,

$$\tilde{n}_{\mathrm{rm}}(l,t) = \left(\frac{a_{\mathrm{eq}}}{a(t)}\right)^3 \tilde{n}_{\mathrm{rr}}(l_{\mathrm{eq}}(l,t), t_{\mathrm{eq}})\,. \tag{4.81}$$

Here, $l_{\mathrm{eq}}(l,t) = l + \Gamma G\mu(t - t_{\mathrm{eq}})$ is the length of a loop at t_{eq} which has at time t length l. Inserting this into Eq. (4.79) yields [403]

$$\tilde{n}_{\mathrm{rm}}(l,t) = \frac{1}{t_{\mathrm{eq}}^{3/2}}\left(\frac{\Omega_r}{\Omega_m}\right)^3\left(\frac{a_0}{a(t)}\right)^3\frac{C_r}{(l + \Gamma G\mu t)^{5/2}}\,, \tag{4.82}$$

upon using that $a_{\mathrm{eq}}/a_0 = \Omega_r/\Omega_m$.

4.3.3 GWB spectrum

We can use the previous results to compute the GWB spectrum emitted by a population of decaying string loops (see, e.g., Refs. [36, 394, 404]). Let us start from a general expression for the power emitted in the form of gravitational radiation from string loops of length l, averaged over loop configurations, namely

$$\frac{\mathrm{d}P_l^{\mathrm{GW}}}{\mathrm{d}f'}(l,f') = G\mu^2 l\Delta\left(lf'\right)\,, \tag{4.83}$$

where f' is the frequency of the radiation at emission. Eq. (4.62) implies the normalisation condition $\int_0^\infty \Delta\left(x\right)\mathrm{d}x = \Gamma$. For the average energy density in gravitational radiation from a network of string loops, we can then write[9]

$$\frac{\mathrm{d}\rho_{\mathrm{GW}}}{\mathrm{d}t'\mathrm{d}f'}(t',f') = \int_0^{\alpha L(t')} \frac{\mathrm{d}P_l^{\mathrm{GW}}}{\mathrm{d}f'}(l,f')\, n(l,t')\,\mathrm{d}l\,, \tag{4.84}$$

where t' is the time of emission. To find the GW spectrum today at t_0 as a function of the observed GW frequency f, we need to take into account that the energy density in radiation redshifts as $\propto a^{-4}$ and the frequency satisfies $f = (a(t')/a_0)\, f'$, such that we obtain in total [36]

$$\frac{\mathrm{d}\rho_{\mathrm{GW}}}{\mathrm{d}f}(f) = G\mu^2 \int_{t_{\mathrm{ini}}}^{t_0} \left(\frac{a(t')}{a_0}\right)^3 \int_0^{\alpha L(t')} n(l,t')\, l\, \Delta\left(\frac{a_0}{a(t')} f l\right)\, \mathrm{d}l\, \mathrm{d}t'\,. \tag{4.85}$$

As usual in cosmology, we reexpress this spectrum in terms of the dimensionless density parameter

$$\Omega_{\mathrm{GW}}\,(f) = \frac{1}{\rho_{\mathrm{crit}}}\frac{\mathrm{d}\rho_{\mathrm{GW}}\,(f)}{\mathrm{d}\ln(f)}\,, \qquad \text{with} \qquad \rho_{\mathrm{crit}} = \frac{3H_0^2}{8\pi G}\,, \tag{4.86}$$

and find

$$\Omega_{\mathrm{GW}}\,(f) = \frac{8\pi G^2\mu^2 f}{3H_0^2} \int_{t_{\mathrm{ini}}}^{t_0} \left(\frac{a(t')}{a_0}\right)^3 \int_0^{\alpha L(t')} n(l,t')\, l\, \Delta\left(\frac{a_0}{a(t')} f l\right)\, \mathrm{d}l\, \mathrm{d}t'\,. \tag{4.87}$$

Recall from Sec. 4.1.1 that the string loops we have obtained from the Nambu–Goto equations oscillate with a period $T = l/2$ and we decomposed their stress–energy tensor and, correspondingly, the emitted power into a discrete Fourier series with modes of frequencies $f_j = \omega_j/(2\pi) = 2j/l$, $j \in \mathbb{N}$ (see, e.g., Eqs. (4.58) and (4.59)). Accordingly, we can set [405]

$$\Delta\,(x) = \sum_{j=1}^{\infty} \Gamma_j\, \delta\,(x - 2j)\,. \tag{4.88}$$

After carrying out the l-integration, the GWB spectrum takes the form

$$\Omega_{\mathrm{GW}}(f) = \frac{16\pi\Gamma G^2\mu^2}{3H_0^2 H_{n_{\max}}^{(q)}} \sum_{j=1}^{n_{\max}} \frac{1}{j^q}\frac{j}{f} \int_{t_{\mathrm{ini}}}^{t_0} \left(\frac{a(t')}{a_0}\right)^5 n\left(j l(t',f), t'\right)\, \mathrm{d}t'\,. \tag{4.89}$$

Here, $l(t',f) = 2\,(a(t')/a_0)\,/f$ is the length the loop had at the emission time t' of GWs if f is the frequency we observe today corresponding to those GWs from the fundamental string-loop excitation. It is now important to observe that, due to the functional dependence of the spectrum, we do not need to compute the contributions from different oscillation modes separately. Instead, the total spectrum can be expressed in terms of the fundamental-mode contribution via

$$\Omega_{\mathrm{GW}}\,(f) = \frac{1}{H_{n_{\max}}^{(q)}} \sum_{j=1}^{n_{\max}} \frac{1}{j^q}\,\Omega_{\mathrm{GW}}^{(j)}(f)\,, \qquad \text{with} \qquad \Omega_{\mathrm{GW}}^{(j)}(f) = \Omega_{\mathrm{GW}}^{(1)}\left(\frac{f}{j}\right)\,. \tag{4.90}$$

The contribution from the fundamental oscillation mode is

$$\begin{aligned} \Omega_{\mathrm{GW}}^{(1)}(f) &= \frac{16\pi}{3}\left(\frac{G\mu}{H_0}\right)^2 \frac{\Gamma}{f} \int_{t_{\mathrm{ini}}}^{t_0} n(l(t',f),t') \left(\frac{a(t')}{a_0}\right)^5 \mathrm{d}t' = \\ &= \frac{16\pi}{3}\left(\frac{G\mu}{H_0}\right)^2 \frac{\Gamma}{f} \int_{a_{\mathrm{ini}}/a_0}^{1} n(l(a,f),a) \left(\frac{a}{a_0}\right)^4 \frac{1}{H(a)}\, \mathrm{d}\left(\frac{a}{a_0}\right)\,. \end{aligned} \tag{4.91}$$

In Fig. 4.5, we show the numerically computed GWB spectrum for a benchmark scenario with $G\mu = 10^{-10}$ and $t_{\mathrm{ini}} = t_{\mathrm{Pl}}/(G\mu)^2$ as explained below in Eq. (4.96). For the scale factor and the

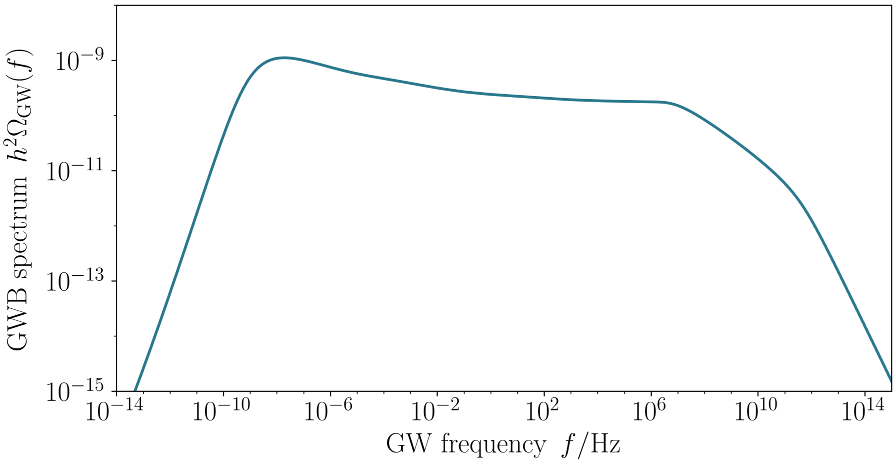


Figure 4.5: GWB spectrum obtained by numerical evaluation of Eqs. (4.90) and (4.91) for parameter values $G\mu = 10^{-10}$ $\alpha = 0.1/\xi_r \simeq 0.37$, $n_{\rm max} = 10^5$, $q = 4/3$, and $t_{\rm ini} = t_{\rm Pl}/(G\mu)^2$.

necessary input from the VOS model, we followed the details explained below Eq. (4.55). As described in the introduction of this thesis, the GWB spectrum of decaying cosmic string loops stretches over a vast range of frequencies. In particular, we point out the plateau region that extends to frequencies as high as 10^7 Hz. It is due to loops decaying in the radiation-dominated era (RR). Another outstanding feature is the peak of the spectrum at around 10^{-8} Hz, which is caused by string loops decaying during matter domination that were already produced in the radiation-dominated phase (RM). We will investigate the various features of this and other spectra in great detail in Ch. 6.

Up until now, we used the Nambu–Goto approximation and described strings as one-dimensional objects, albeit, in reality, strings have a finite width δ. Apart from possible effects on the particle decay of string loops, which we will discuss in the next chapter, the finite width will be irrelevant for most parts of the spectrum. Nevertheless, at extremely high frequencies that correspond to emission wavelengths which can resolve the string's microstructure, the non-vanishing string width will become relevant. Following our original work in Ref. [4], we will now show how this can be taken into account. Consider a string loop of length l and a corresponding oscillation mode j with wavelength $\lambda_j = l/(2j)$. If this wavelength becomes smaller than the string width δ, the Nambu–Goto approximation will break down, and we, furthermore, do not expect the emission of gravitational radiation. We account for this by modifying the function Δ introduced in Eq. (4.83) and expressed in Eq. (4.88). In practice, we implement an additional Heaviside function $\Theta\left(l - 2j\delta\right)$ and replace Δ in Eq. (4.87) by[10]

$$\tilde{\Delta}(x, f) = \sum_{j=1}^{\infty} \Gamma_j \, \delta\left(x - 2j\right) \Theta\left(x - 2j\delta \, \frac{a_0}{a(t')} f\right) . \tag{4.92}$$

Observe that this breaks the sole dependence of Δ on the variable x. The integration over l or, equivalently, over x in Eq. (4.87) leads to the same result as before, however, with an additional

---

[9] The upper integration boundary could also be set to infinity, but the largest loops that are present in the network at time t' are those formed at time t'. The loop production function in Eq. (4.71) sets the maximum loop length at time t' to $\alpha L(t')$. This fact is taken into account by the condition $t' > t_*(l, t')$, implemented via the first Heaviside function in the loop number density given in Eq. (4.72).

[10] In reality, the emission of radiation will not be unmodified above a length scale δ and completely die off below. We rather expect a smooth transition between the two regimes. Nonetheless, due to the absence of a better understanding of this transition, the modelling in terms of a Heaviside function should give a good first approximation.

Heaviside function of the form $\Theta\left(a(t')/a_0 - \delta f\right)$. This signifies that any frequency f observed nowadays in the spectrum is only allowed to correspond to an emission frequency at time t' associated with a wavelength larger than δ. This statement is, in particular, independent of the mode number. We can also view this as a constraint on the earliest possible time of emission depending on the frequency, which yields, upon using the Planck time $t_{\rm Pl} = G^{1/2}$,

$$\frac{a_{\rm min}(f)}{a_0} = \frac{t_{\rm Pl} f}{(G\mu)^{1/2}}\,, \tag{4.93}$$

where we used that the string width can be expressed in terms of the string tension as $\delta \sim \mu^{-1/2}$, up to prefactors which we assume to be of order one (see Sec. 2.2.3). This effect should be included by replacing $a_{\rm ini}$ (and correspondingly in $t_{\rm ini}$) with

$$a_{\rm ini} \longrightarrow \max\left\{a_{\rm min}(f), a_{\rm ini}\right\}\,. \tag{4.94}$$

It is important to pay attention to the fact that, while frequency-dependent, this minimum scale factor is independent of the mode number. Thus, the simple relation between the spectra from higher harmonics and the fundamental mode giving rise to Eq. (4.90) no longer holds. Fortunately, we can simply treat $a_{\rm ini}$ as an additional variable, carry out the sum over harmonics in Eq. (4.90), and only replace $a_{\rm ini}$ with the new frequency-dependent expression in the end. This is the route we will take in the following. As we will see at the end of Secs. 6.1.1 and 6.2.2, this change in the initial time can, indeed, have an impact on the GW spectrum at extremely high frequencies.

Let us also very briefly provide the standard estimate for $t_{\rm ini}$ — the time when GW emission from decaying string loops becomes significant. As we will see in more detail in the next chapter, $t_{\rm ini}$ is a parameter that can, in principle, be determined from other, microscopic model parameters. Nonetheless, there are substantial theoretical uncertainties in choosing the correct expression for $t_{\rm ini}$. The standard estimate follows from our discussion at the end of Sec. 4.2, where we saw that thermal friction suppresses the motion of strings, and thus GW emission. Setting all dimensionless prefactors to the value 1, the thermal friction and the Hubble friction in Eq. (4.57) are equal when

$$G^{1/2}T^2 \sim \frac{T^3}{\mu}\,, \tag{4.95}$$

where we used that during radiation domination $H \sim G^{1/2}T$. Solving this expression for T gives a rough estimate for the end of the so-called friction regime, which indeed occurs deep in the radiation-dominated era. We can find the corresponding time $t_{\rm fric}$ by using that $H = 1/(2t)$ and $t_{\rm Pl} = G^{1/2}$ which yields

$$t_{\rm fric} \sim \frac{t_{\rm Pl}}{(G\mu)^2}\,. \tag{4.96}$$

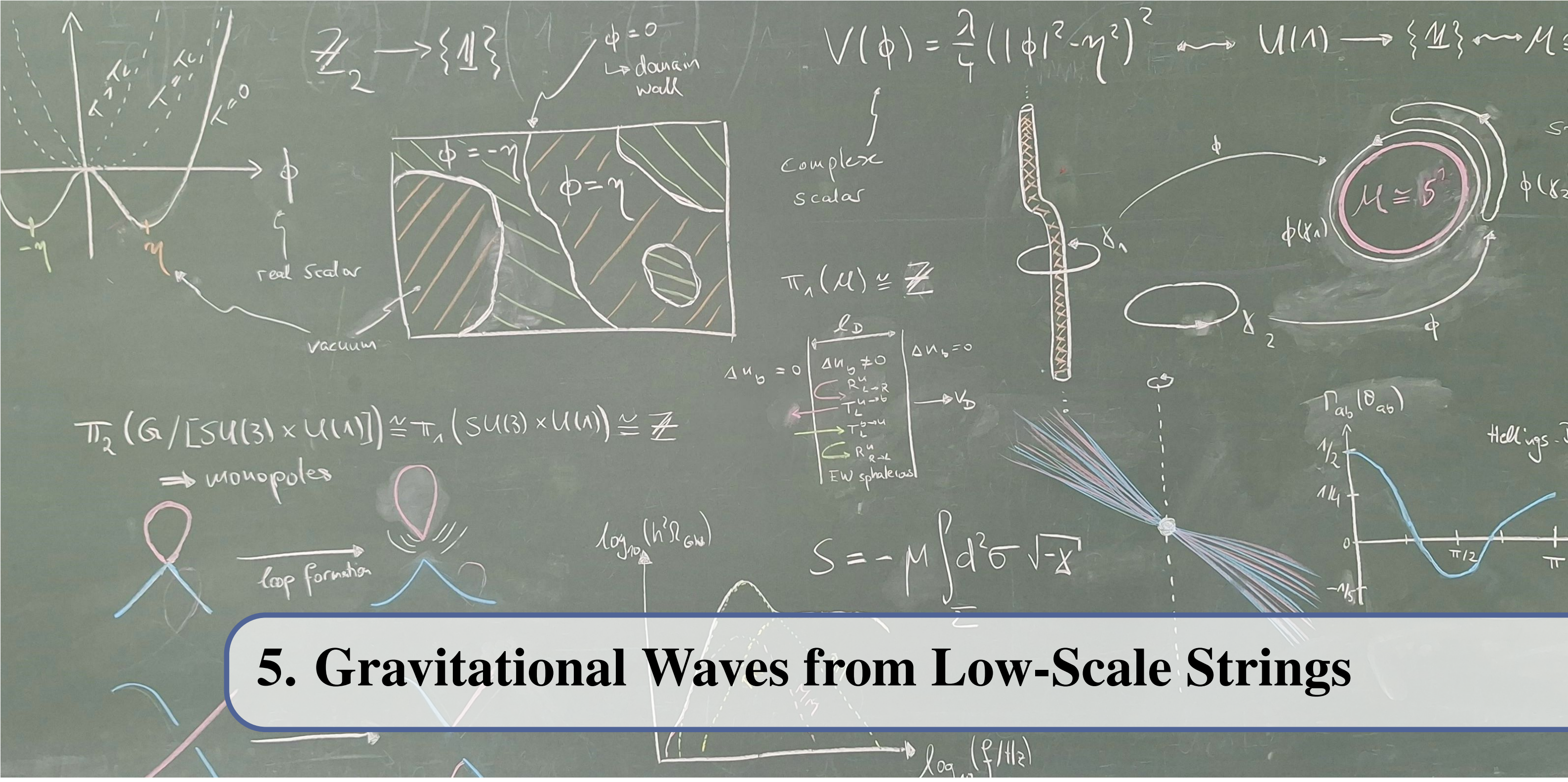
This provides a typical and well-motivated choice for the initial time $t_{\rm ini}$. It is also the earliest time when the network can reach the scaling regime, which we rely on in our calculations. Apart from that, particle decay might also affect the initial time since particle decay due to cusp formation or kink–kink collisions becomes the dominant decay channel of string loops below a critical loop length.[11] The initial time must, hence, be chosen late enough such that the loop production length has already become large enough to make the gravitational decay dominant. This point will be an essential ingredient for our discussion in the next chapter.

Lastly, it should also be pointed out that actually every defect network that adapts to remain in a scaling regime, in particular the long-string network, produces a GWB, as discussed, e.g., in

---

[11] In fact, this statement has been the subject of a long-lasting debate and no consensus has yet been reached. See, e.g., Refs. [344, 406–410] for recent articles in this regard.

Refs. [73, 88–90]. The GWs are produced with wavelength of around the Hubble scale at the time of emission. For radiation produced during the radiation-dominates era, this leads to an exactly scale invariant background. A similar plateau is also obtained from the decaying string loops, which is also visible in Fig. 4.5. The amplitude of the former scales like $(G\mu)^2$ while the amplitude of the latter scales like $(G\mu)^{1/2}$, as we will see in Secs. 6.1.1 and 6.2.2, and is therefore strongly suppressed [36, 91]. In the following, we will therefore only discuss GWs from loops and ignore the GW signal from long strings.



5. Gravitational Waves from Low-Scale Strings

This chapter is based on Refs. [3, 5]. The detailed contributions of the author to these articles are outlined at the beginning of this thesis.

As a possible source of GWs from the early Universe, cosmic strings have increasingly gained attention for current and prospective GW observations [85, 86, 394]. The GWB signal relevant for current and near-future observatories corresponds to string tensions $G\mu \gtrsim 10^{-17}$ or, equivalently, symmetry-breaking scales $\eta \gtrsim 10^{10}$GeV [394, 411]. While the GWB signal from such *high-scale strings* has been extensively investigated in the literature [364, 390, 402, 412], strings could also plausibly form at lower energies, $\eta \ll 10^{10}$ GeV. Cosmic strings with far too low tensions, which can be as low as $G\mu \sim 10^{-33}$ if they form in a symmetry breaking just above the electroweak scale, will hardly be observable via their GW signature. Nevertheless, for energy scales of the order $\eta \sim 10^9$ GeV, the GWB signal can be in the sensitivity reach of future space-borne GW interferometers such as BBO [98] and DECIGO [97]. The symmetry-breaking scales necessary for these scenarios can easily be realised in many SM extensions. Among the many natural possible $U(1)$ symmetries that could be spontaneously broken at $\eta \sim 10^9$ GeV is the one associated with the conservation of baryon-minus-lepton number $B-L$. This symmetry breaking is also of phenomenological interest beyond strings [91, 413, 414]. That said, in the following, we will not specify a microphysical model.

Instead, in this chapter, we demonstrate that the GWB spectrum produced by strings within a certain region of parameter space associated with low tensions is qualitatively distinct from the well-known spectra of high-scale strings that resemble those shown in Fig. 4.5. We will refer to strings with this new GW signature as *low-scale strings*.

The emergence of a qualitatively different GWB signature can be readily understood through a simple argument based on the discussion in the previous chapter: Cosmic string loops continuously shrink as they emit GWs. We have seen that they decay with a constant rate $\mathrm{d}l/\mathrm{d}t = -\Gamma G\mu$, which is directly proportional to the string tension (see Eq. (4.66)). The corresponding loop length evolution is provided in Eq. (4.67). The shortest loops are produced at the earliest times, let us say with an initial length l_{ini}. Considering increasingly lower string tensions $G\mu$, the gravitational

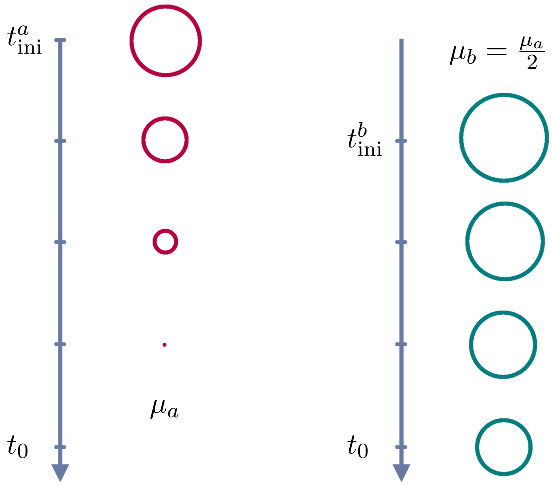


Figure 5.1: Illustration of the effect of string tension μ and initial time $t_{\rm ini}$ on loop decay. Loop sizes are shown at different time steps from their formation until today. **Left:** A loop (purple) with string tension μ_a produced at time $t^a_{\rm ini}$ decays at a constant rate $\Gamma G\mu_a$, sufficient for the loop to completely evaporate before today t_0. The depicted loop corresponds to high-scale strings. **Right:** Another loop (teal) with a smaller string tension $\mu_b = \mu_a/2 < \mu_a$, formed at a later time $t^b_{\rm ini} > t^a_{\rm ini}$, has a larger initial length $l^b_{\rm ini} > l^a_{\rm ini}$. Both the larger length and the lower constant decay rate $\Gamma G\mu_b < \Gamma G\mu_a$ increase the loop's decay time. The depicted loop has not fully evaporated by today. If $t^b_{\rm ini}$ is the formation time of the earliest loops that significantly contribute to the GWB, the depicted loop corresponds to low-scale strings. Note that, outside this illustration, we typically consider string tensions that deviate by orders of magnitude, not merely by a factor of 2.

decay of the loops will become less and less efficient. Below a critical string tension $G\mu$, the loop will not have enough time to fully evaporate between its time of production $t_{\rm ini}$ and today t_0. We have illustrated this in Fig. 5.1.

Since the loops produced with the shortest length $l_{\rm ini}$ are produced earliest, they have the longest available time to evaporate. If even these loops survive until today, then no loops formed in the early Universe will have completely decayed yet. Instead, there is an overall minimum length $l_{\rm min}$ in the distribution of loop lengths. Because the frequency of GWs emitted by loops is inversely proportional to the loop length at time of emission (see Sec. 4.1.1), the minimum loop length $l_{\rm min}$ corresponds a sharp ultraviolet cutoff frequency $f_{\rm cut}$ in the GWB spectrum sourced by the fundamental oscillation mode of string loops. In the following, we will refer to this spectrum often as the fundamental GWB spectrum. However, the actually observable GWB spectrum consists of contributions from all loop oscillation modes as described by Eq. (4.90) and we refer to is as the (total) GWB spectrum. The cutoff in the fundamental spectrum leads to a sequence of local peaks and dips in the total spectrum with each dip at an integer multiple of the cutoff frequency $f_{\rm cut}$. The GWB spectrum from low-scale strings, therefore, exhibits a distinct feature that does not appear in the GWB associated with high-scale strings. Importantly, if the string energy scale remains around $\eta \sim 10^9$ GeV, such spectral patterns could become observable in upcoming GW experiments. This presents an intriguing opportunity to explore the discrete spectrum of oscillation modes on string loops through future observations.

We begin this chapter by presenting in detail the main argument that explains the emergence of the cutoff frequency $f_{\rm cut}$ in Sec. 5.1. Afterwards, we discuss the uncertainty surrounding the shortest length $l_{\rm ini}$ at which loops contributing to the GWB are formed and review possible choices for its value. In Sec. 5.2, we compute the GWB from low-scale strings based on the VOS model, which was introduced in Sec. 4.2. This allows us to understand the origin of the oscillatory feature in the spectrum. The loop number densities used in this computation are built on the approximation that loops in the network are produced with a length that is a fixed fraction of the horizon scale. This assumption is also central for our argument leading to the sharp cutoff frequency and the respective oscillatory feature in the total GWB spectrum. However, in Ref. [47], BLANCO-PILLADO, OLUM, and SHLAER (BOS) numerically simulated a network of cosmic strings and found that the initial loop lengths follow a distribution of narrow but non-vanishing width. In Sec. 5.3, we will study the effects of such a finite-width distribution by replacing the Dirac distribution for the initial loop length with a Gaussian distribution which approximates the result found by BOS. As one might expect, this leads to a smoothing of the oscillatory feature in the GWB spectrum. Nevertheless, we find that the sequence of local peaks and dips in the spectrum remains at a detectable level.

What we refer to as a low-scale string clearly depends on the way string loops shrink, or similarly, the time evolution of the loop length. Until this point, we work under the assumption that the loop length evolution follows, due to the emission of gravitational radiation, the relation $\mathrm{d}l/\mathrm{d}t = -\Gamma G\mu$. This is the standard relation employed in models describing a scaling loop population. Note, in particular, that the right-hand side of this relation does not involve any explicit length scale, which would violate the assumption of scaling. The scaling of the loop population derives from the scaling of the long-string network via a loop production function that only depends on the scaling variable l/t but not on l and t separately. The scale independence of the loop's decay rate is only maintained as long as loops predominantly decay into gravitational radiation. For small loops, either already small at production or small due to advanced loop decay, evaporation processes other than gravitational decay can become dominant. In particular, in kink–kink collisions (k) and in the formation of cusps (c), the topological criterion stabilising the strings is lifted in a small region of the loop, and the string can decay into particle radiation. This breaks the scaling of loops as it introduces additional scales for particle radiation into the problem. In Secs. 5.4 and 5.5, we incorporate these effects by employing the nonscaling models presented in Ref. [64] and investigate their influence on the GWB spectra from low-scale cosmic strings.

In Sec. 5.4, we first review the modifications of the loop length evolution due to particle decay. In the framework of the associated nonscaling models, we will again provide analytical expressions for the cutoff frequency f_{cut}. We will derive criteria for the existence of f_{cut} and show how these constrain the parameter space of strings. We will also discuss how the results obtained in nonscaling models relate to the results in the scaling model that we considered in the preceding sections. Following the discussion of the parameter space, we will compute the GWB spectra in Sec. 5.5. Once again, the computation will be based on the VOS equations for the network of long strings, which continue to hold also in the nonscaling models. We provide the GWB spectra for nine different benchmark points that cover the qualitatively different regimes of our parameter space. These spectra are obtained by a fully numerical evaluation of the general expressions for the GWB. Afterwards, we turn to an approximate treatment and derive analytical estimates for the GWB spectrum of low-scale strings in the nonscaling models. We derive expressions for the peak frequency and amplitude when transitioning from the low-scale to the high-scale regime and find very good agreement between our analytical and numerical approaches.

Finally, in Sec. 5.6, we conclude with a brief outlook. We will summarise our findings and discuss their robustness as well as their relevance. We also point towards open issues that need to be addressed in future studies.

5.1 Sharp cutoff frequency

5.1.1 Principal argument

Recall from Sec. 4.3 that string loops oscillate for a given loop length l at discrete frequencies $f = 2j/l$ with $j \in \mathbb{N}$, which match the frequencies at which gravitational radiation is emitted. Since the Universe expands between the time of emission and the present time, the GWs are observed today at the redshifted frequencies

$$f = \frac{a\,(t)}{a_0}\frac{2j}{l\,(t)}\,, \qquad \text{with} \qquad j \in \mathbb{N}\,. \tag{5.1}$$

Focusing on the fundamental mode $j = 1$, we can observe that the largest possible frequency $f_{\rm cut}$ in the spectrum is set by the smallest loops with length $l_{\rm min}$. Since loops only shrink over time, $l_{\rm min}$ is only reached today and we do not need to account for redshifting:

$$f_{\rm cut} = \frac{2}{l_{\rm min}}\,. \tag{5.2}$$

Let us now compute the minimum length. Since our argument depends crucially on the evolution of loop lengths, we reproduce here, for convenience, the result from Eq. (4.67):

$$l(t,t_*) = l_*(t_*) - \Gamma G\mu(t - t_*)\,. \tag{5.3}$$

This equation tells us that a loop produced at time t_* with length $l_*(t_*)$ has at time t a length $l(t,t_*)$. Recall now also from Eq. (4.70) that numerical simulations by BOS show the initial loop lengths to be narrowly distributed around a fixed fraction of the horizon scale d_h during the radiation-dominated era:

$$l_* = \alpha_h\, d_h\,(t_*)\ , \qquad \alpha_h \simeq 0.05\,, \qquad d_h\,(t_*) = 2t_*\,. \tag{5.4}$$

The size of the loops produced in the network increases monotonically with the growing horizon scale d_h. Here, we expressed l_* in terms of d_h, whereas in Eq. (4.70), it was expressed in relation to the VOS model's universal length scale L. We introduced an alternative reference scale because, in this chapter, we will not only work in the VOS model but also in the BOS model. To account for the change in reference scale, we introduced a new proportionality factor $\alpha_h = \alpha\xi/2$ with ξ and α as previously defined in Secs. 4.2 and 4.3.2.

From Eq. (5.4), it is evident that the shortest lengths at birth are those produced at the earliest times $t_{\rm ini}$. Then the shortest loop production length is $l_{\rm ini} = l_*(t_{\rm ini}) = 2\alpha_h t_{\rm ini}$. At the same time, these earliest loops are the ones that have had the longest time to decay until today. This implies that their current length sets the minimum length of all string loops, and we will denote this overall minimum length by

$$l_{\rm min} = l_{\rm ini} - \Gamma G\mu\,(t_0 - t_{\rm ini})\ . \tag{5.5}$$

The minimum length $l_{\rm min}$ sets, via Eq. (5.2), the ultraviolet cutoff in the fundamental spectrum:

$$f_{\rm cut} = \frac{2}{l_{\rm min}} = \frac{2}{2\alpha_h t_{\rm ini} - \Gamma G\mu\,(t_0 - t_{\rm ini})}\,. \tag{5.6}$$

Suppose that none of the loops has fully evaporated yet, which is equivalent to demanding that $l_{\rm min} > 0$. In this case, the frequency $f_{\rm cut}$ is finite as well as positive and a sharp high-frequency cutoff appears in the GWB spectrum sourced by the fundamental mode. The condition $l_{\rm min} > 0$ is satisfied if the string tension is sufficiently small that

$$\Gamma G\mu\,(t_0 - t_{\rm ini}) < 2\alpha_h t_{\rm ini}\,. \tag{5.7}$$

Since loop production begins still in the early Universe, we can use the hierarchy $t_0 \gg t_{\rm ini}$ to rewrite this condition as

$$t_{\rm ini} > t_{\rm cut} = \frac{\Gamma G\mu}{2\alpha_h}\,t_0 \simeq 2.2\,{\rm s}\left(\frac{G\mu}{10^{-20}}\right)\,. \tag{5.8}$$

If the onset of loop production happens after $t_{\rm cut}$, we identify the strings as low-scale strings: Their fundamental GWB spectrum exhibits a sharp cutoff frequency. For the values of $G\mu$ of interest in this chapter, the threshold time $t_{\rm cut}$ always falls within the radiation-dominated epoch. Therefore, it is natural to restate this condition in terms of the temperature of the thermal bath at that time. Explicitly, we have during radiation domination the relation

$$H = \frac{1}{2t} = \frac{T^2}{M_*(T)}\,, \qquad \text{where} \qquad M_*(T) = \sqrt{\frac{90}{\pi^2 g_\rho(T)}}\,M_{\rm Pl}\,. \tag{5.9}$$

M_* is related to the reduced Planck mass $M_{\rm Pl} = (8\pi G)^{-1/2} \simeq 2.435 \times 10^{18}$ GeV and accounts for changes in the effective number of relativistic degrees of freedom g_ρ. We can, therefore, write

$$T_{\rm ini} = \sqrt{\frac{M_*(T_{\rm ini})}{2t_{\rm ini}}} = \sqrt{\frac{\alpha_h M_*(T_{\rm ini})}{l_{\rm ini}}}\,. \tag{5.10}$$

Expressed in terms of temperature, the criterion for the existence of a maximum frequency becomes

$$T_{\rm ini} < T_{\rm cut} = \sqrt{\frac{M_*}{2t_{\rm cut}}} \simeq 330\,{\rm keV}\left(\frac{10^{-20}}{G\mu}\right)^{1/2}\,. \tag{5.11}$$

For the numerical evaluation, we assumed $g_\rho = 100$. This value will be retained in all analytical calculations for simplicity, while a temperature-dependent g_ρ is used in our numerical analysis and in all plots.

The resulting cutoff temperature $T_{\rm cut}$ is relatively low. For a string tension of $G\mu \sim 10^{-20}$, it is comparable to the Big-Bang nucleosynthesis scale $T_{\rm BBN} \sim 1\,{\rm MeV}$. This is well below the energy scales usually associated with the generation of cosmic strings. Consequently, it may not be obvious that the initial temperature $T_{\rm ini}$ can indeed lie below $T_{\rm cut}$ — the criterion that must be satisfied for the presence of an ultraviolet cutoff frequency $f_{\rm cut}$ in the GWB signal from the fundamental oscillation mode. In what follows, we will explore various plausible choices for $T_{\rm ini}$ and argue that the inequality $T_{\rm ini} < T_{\rm cut}$ can indeed be fulfilled over a wide range of string tensions $G\mu$.

5.1.2 Initial loop length

Let us now have a closer look at the initial time $t_{\rm ini}$. Recall from the discussion in Sec. (4.3.3) that, to obtain the GWB from a cosmic string network, we need to evaluate the integral in Eq. (4.91), which runs over all possible GW emission times. The lower integration boundary, set by the earliest emission time, is $t_{\rm ini}$. Therefore, the value of $t_{\rm ini}$ can have a considerable impact on the GWB spectrum. This is already evident from the fact that the minimum lengths in Eq. (5.90), and hence the cutoff frequency in the fundamental spectrum, mainly depend on the value of $t_{\rm ini}$.

Several possible choices for $t_{\rm ini}$ have been proposed in the literature (see, e.g., Refs. [412, 415]). In the following, we briefly review and critically assess these proposals.

Network formation: The very first time loops can start to form and produce gravitational radiation is when the string network forms. The formation occurs during a phase transition that takes place at a temperature $T_{\rm form} \sim \eta$ or, equivalently, at the temperature at which the critical density satisfies $\rho_{\rm crit}(T_{\rm form}) = 3H^2(T_{\rm form})M_{\rm Pl}^2 \sim \mu^2$. For definiteness, we will use the latter form of this condition. Since the network formation occurs deep in the radiation-dominated era, the corresponding temperature of network formation is

$$T_{\rm form} = \left(\frac{30}{\pi^2 g_*}\right)^{1/4} \mu^{1/2} \simeq 5.1 \times 10^8\,{\rm GeV}\left(\frac{G\mu}{10^{-20}}\right)^{1/2}\,, \tag{5.12}$$

and the associated time reads

$$t_{\rm form} \simeq 9.3 \times 10^{-25}\,{\rm s} \left(\frac{G\mu}{10^{-20}}\right)^{-1} . \tag{5.13}$$

The time $t_{\rm form}$ is a strict lower bound on $t_{\rm ini}$. However, recall that strings formed in the phase transition are typically infinitely long or of super-Hubble size [82], so they will not directly contribute to the GWB from decaying string loops. Only once the long strings start to move around freely will the production of shorter loops become significant.

Thermal friction: As we know from the discussion at the end of Sec. 4.2, in realistic microscopic models, the dynamics of strings shortly after the phase transition is dominated by friction due to interactions between the strings and the surrounding thermal plasma. The production of string loops that decay into GWs thus takes place at times later than the network formation and increases $t_{\rm ini}$.[1] The total friction experienced by strings consists of a contribution from the Hubble friction and a temperature-dependent friction term. As discussed in Sec. 4.2, the total friction term that enters the Nambu–Goto equations of motion is of the form

$$2H + \frac{\gamma T^3}{\mu} , \tag{5.14}$$

where the parameter γ needs to be determined from the particle physics model under consideration. Here, we will work with $\gamma \sim 1$ for temperatures above $T_e \sim 0.5\,{\rm MeV}$, the temperature when electrons and positrons annihilate. Below T_e, there are no more charged particles contributing to the thermal bath, and we set $\gamma = 0$. To determine at which temperature the motion of strings becomes effectively free, we need to compare the thermal friction to the Hubble friction. The temperature and time at which the thermal friction becomes subdominant are

$$T_{\rm fric} = \frac{2\mu}{\gamma M_*} \simeq 4.1\,{\rm GeV} \left(\frac{1}{\gamma}\right) \left(\frac{G\mu}{10^{-20}}\right) , \tag{5.15}$$

$$t_{\rm fric} = \frac{M_*}{2T_{\rm fric}^2} \simeq 15\,{\rm ns} \left(\frac{\gamma}{1}\right)^2 \left(\frac{G\mu}{10^{-20}}\right)^{-2} . \tag{5.16}$$

Observe that the above relation holds only as long as $T_{\rm fric} > T_e$, or equivalently, for string tensions $G\mu > 1.2 \times 10^{-24}\gamma$. Below that, $T_{\rm fric}$ assumes the constant value T_e.

The time $t_{\rm fric}$ is, for the string tensions of interest to us, much later than $t_{\rm form}$ and can, depending on the underlying microphysical model, provide a reasonable estimate of $t_{\rm ini}$. With this choice of initial temperature, we can see from Eqs. (5.8) and (5.16) that a sharp cutoff frequency appears in the fundamental contribution to the GWB spectrum if $G\mu \lesssim 1.9 \times 10^{-23}$. Substituting $t_{\rm fric}$ for the initial time, the initial loop length evaluates to $l_{\rm fric} = l_*(t_{\rm fric})$ or, more explicitly, to

$$l_{\rm fric} = \frac{\alpha_h \gamma^2 M_*^3}{4\mu^2} \simeq 440\,{\rm km} \left(\frac{\gamma}{1}\right)^2 \left(\frac{10^{-23}}{G\mu}\right)^2 . \tag{5.17}$$

For string tensions $G\mu$ far below the threshold value for a cutoff frequency, i.e., $t_{\rm fric} \gg t_{\rm cut}$, loops will only lose a small fraction of their initial length until today. Consequently, their current length roughly equals their initial length $l_{\rm min} \simeq l_{\rm fric}$ and we see from Eq. (5.6) that $\tilde{f}_{\rm cut}^{\rm fric} \simeq 2/l_{\rm fric}$, which

---

[1] In fact, string loop production can become significant at earlier times, although the resulting GW emission remains suppressed. For a possible high-frequency GW signal from loops produced during the friction era in certain models, see the recent discussion in Ref. [386].

explicitly yields[2]

$$\tilde{f}_{\rm cut}^{\rm fric} \simeq \frac{8\mu^2}{\alpha_h \gamma^2 M_*^3} \simeq 1.4\,{\rm kHz} \left(\frac{1}{\gamma}\right)^2 \left(\frac{G\mu}{10^{-23}}\right)^2 . \tag{5.18}$$

This analytical estimate provides a good approximation to the actual cutoff frequency as we will see in Sec. 5.2, where we compute the GWB fully numerically on the basis of the VOS model.

Particle radiation from kink–kink collisions: The formation and friction cutoffs provide well-motivated choices for the initial time. However, this is only true if string loops decay during their lifetime until today mainly into gravitational radiation, which is not always the case. In fact, in many models, loops below a certain length evaporate predominantly via the emission of particle radiation [64, 81, 343, 344, 416]. This second decay channel is a consequence of the small scale structure of loops — the cusps and kinks discussed in Sec. 4.1.1. These substructures can lead to the string overlapping with itself, which lifts the topological stabilisation in small regions of the loop and the Nambu–Goto approximation breaks down. These small regions subsequently evaporate in the form of bursts of particle radiation. The emitted particles have typical masses of order $\sim \eta$.

One of two possible origins of particle emission is from collisions of two kinks which propagate with luminal velocities in opposite directions along the string. In Ref. [416], it was argued and numerically verified that the power emitted into particle radiation originating from kink–kink collisions can be estimated as

$$P_k \sim \frac{N_k \epsilon}{l} . \tag{5.19}$$

Here, $\epsilon \sim \mu^{1/2}$ is the l-independent energy emitted in a single radiation burst and N_k is the number of such bursts per oscillation period. In extreme cases, N_k can assume values up to $N_k \sim 10^3 \cdots 10^6$ [399, 417, 418].

Note that the power emitted due to kink–kink collisions in Eq. (5.19) depends on the loop length l, while the power emitted into gravitational radiation $P_{\rm NG} = \Gamma G\mu^2$ does not (see Eq. (4.62)). The subscript “NG” indicates that only true Nambu–Goto loops decay purely gravitationally. From this observation, it is evident that for loop lengths shorter than a critical length l_k, particle emission from kink–kink collisions dominates over the gravitational decay. This critical length can be found as

$$l_{\rm k} \sim \frac{N_k}{\Gamma G\mu^{3/2}} \simeq 320\,{\rm nm} \left(\frac{N_k}{1}\right) \left(\frac{10^{-20}}{G\mu}\right)^{3/2} . \tag{5.20}$$

Loops formed before the time $t_k = l_k/(2\alpha_h)$, which evaluates to

$$t_k \sim \frac{N_k}{2\alpha_h\,\Gamma G\mu^{3/2}} \simeq 11\,{\rm fs} \left(\frac{N_k}{1}\right) \left(\frac{10^{-20}}{G\mu}\right)^{3/2} , \tag{5.21}$$

will be shorter than l_k for their entire lifetime. The temperature corresponding to the time t_k reads

$$T_k = \sqrt{\frac{M_*}{2t_k}} \simeq 4.7\,{\rm TeV} \left(\frac{1}{N_k}\right)^{1/2} \left(\frac{G\mu}{10^{-20}}\right)^{3/4} . \tag{5.22}$$

Comparing Eqs. (5.16) and (5.21), we can see that, for the string tensions of interest for low-scale strings, there is hierarchy $t_k \ll t_{\rm fric}$, and t_k does not play a role as an initial time. For

---

[2] The tilde over the frequency carries no deeper meaning and is introduced solely to avoid confusion with quantities defined later.

an extreme number of kinks on the loop, $N_k \sim 10^6$, t_k can become as large as $t_{\rm fric}$, but still not larger. Only for a combination of such an extreme number of kinks and very low string tensions for which $T_k < T_e \simeq 0.5\,{\rm MeV}$, i.e., $G\mu \ll 5 \times 10^{-26}$, the time t_k becomes a relevant choice for $t_{\rm ini}$. However, such low tensions are of little interest, as they do not lead to GWB signals that can be observed in any near future experiment. Hence, in the following we will not identify $t_{\rm ini}$ with t_k.

Particle radiation from cusps: The other possible origin of bursts of particle radiation is the formation of cusps [81, 343]. Around a cusp, the string can, as a consequence of its finite width, overlap with itself. The Nambu–Goto approximation breaks down for the overlapping parts and the topological restriction preventing the string from decaying is locally lifted. Taking the relativistic length contraction into account, the size of this overlap was determined to be $\sqrt{\delta l}$ [81, 343] for a cusp forming on a loop of length l with string width $\delta \sim 1/\sqrt{\mu}$. After the formation of the cusp, the overlapping region of size $\sqrt{\delta l}$ will evaporate into particle radiation and thereby shrink to a region that is roughly of size δ. This remaining region can be neglected since we are interested in loops that satisfy $l \gg \delta$ and the particle evaporation of the cusp approximately reduces the length of the string from l to $l - \sqrt{\delta l}$. The energy lost by the string loop in this process is $\mu\sqrt{\delta l} \sim l^{1/2}\mu^{3/4}$. If a number of N_c cusps forms per oscillation period $T \sim l$, then the emitted power can be estimated as

$$P_c \sim \frac{N_c \mu^{3/4}}{l^{1/2}} \,. \tag{5.23}$$

In Ref. [361], it was found in a toy model that almost all loops develop $N_c \sim 1$ cusps per oscillation period in agreement with the geometric argument for cusp formation discussed in Sec. 4.1.1. As for the kink–kink collisions, we can observe that P_c depends on the loop length l, while $P_{\rm NG} = \Gamma G\mu^2$ does not. Equating P_c with $P_{\rm NG}$ leads to the critical length

$$l_c \sim \frac{N_c^2}{\Gamma^2 G^2 \mu^{5/2}} \simeq 4.3\,{\rm AU}\left(\frac{N_c}{1}\right)^2 \left(\frac{10^{-20}}{G\mu}\right)^{5/2} \,, \tag{5.24}$$

below which evaporation into particles due to cusp formation is the dominant decay mechanism and gravitational decay is subdominant. Loops for which GW emission is the dominant decay mechanism can form only after the time $t_c = l_c/(2\alpha_h)$. This time evaluates to

$$t_c \sim \frac{N_c^2}{2\alpha_h \Gamma^2 G^2 \mu^{5/2}} \simeq 6.0\,{\rm h}\left(\frac{N_c}{1}\right)^2 \left(\frac{10^{-20}}{G\mu}\right)^{5/2} , \tag{5.25}$$

with a corresponding temperature of

$$T_c = \sqrt{\frac{M_*}{2t_c}} \simeq 3.4\,{\rm keV}\left(\frac{1}{N_c}\right) \left(\frac{G\mu}{10^{-20}}\right)^{5/4} \,. \tag{5.26}$$

Note that for any realistic choice of parameter values, we find $T_c \ll T_{\rm fric}$. Although the network leaves the friction regime around the time $t_{\rm fric}$, GW emission will not become the dominant decay mechanism of loops until the much later time t_c. Assuming that loops with production lengths shorter than l_c do not significantly contribute to the GWB, we should identify $t_{\rm ini}$ with t_c. We will discuss the range of validity of this assumption and an alternative scenario in Secs. 5.4 and 5.5.

Having identified t_c as the initial time, we can also compare T_c to $T_{\rm cut}$ in Eq. (5.11) and find for $N_c = 1$ that $T_c < T_{\rm cut}$ is satisfied for string tensions $G\mu < 1.3 \times 10^{-19}$. Consequently, for these parameters, the fundamental GWB spectrum will exhibit a sharp cutoff frequency.

For string tensions $G\mu$ sufficiently far below the critical value for a cutoff frequency, i.e., $t_c \gg t_{\rm cut}$, loops will barely decay at all. The current minimum loop length in the network is, therefore, roughly equal to the production length of the earliest loops $l_{\min} \simeq l_c$, and we see from Eq. (5.6) that $\tilde{f}^c_{\text{cut}} \simeq 2/l_c$, which evaluates to

$$\tilde{f}^c_{\text{cut}} \simeq \frac{2\,\Gamma^2 G^2 \mu^{5/2}}{N_c^2} \simeq 0.93\,\text{mHz} \left(\frac{1}{N_c}\right)^2 \left(\frac{G\mu}{10^{-20}}\right)^{5/2} . \tag{5.27}$$

For typical parameter values, this GW frequency lies in the millihertz range and may be detectable by future space-borne interferometers. We will later show that for tensions of $G\mu \sim 10^{-19}$, not only the frequency band is appropriate but also the amplitude of the signal falls into the sensitivity reach of BBO and DECIGO.

Identifying t_{ini} with t_c instead of t_{fric} also has an advantage when it comes to the computation of the GWB spectrum. The formalism that allows for this computation, outlined in Sec. 4.3.3, relies on the long-string network having settled in the scaling regime for all times of GW emission. The scaling regime can at the earliest be reached when the thermal friction has become subdominant to the Hubble friction, which happens at t_{fric}. However, with t_{fric} as an initial time, the network may not yet have entered the scaling regime. This introduces uncertainties in the contributions to the GWB spectrum from early emission times. By contrast, with t_c as an initial time, the network will have sufficient time to reach the scaling regime as a result of the large hierarchy $t_c \gg t_{\text{fric}}$.

For the following discussion, we will treat the time scales t_{form}, t_{fric}, t_k, and t_c as hard cutoffs as done previously, e.g., in Refs. [412, 415]. In reality, the onset of GW emission from loops will, of course, be smooth. In Sec. 5.4, we will discuss the transition from particle decay to gravitational decay in more detail and argue that the hard cutoff is one of two extreme cases that might be realised.

In Fig. 5.2, we show the temperature scales T_{form} (black), T_{fric} (orange), T_k (dark yellow), T_c (red-orange) as well as T_{cut} (blue) as functions of the string tension. The parameter region above T_{form} is shaded black, since the absence of cosmic strings excludes it from parameter space. The remaining features of the plot will become clear in a moment.

5.2 GWB spectrum in the VOS model

Our next goal is to compute the GWB signal from low-scale cosmic strings. In this section, we follow the procedure outlined in Sec. 4.3. In particular, our computations are based on the VOS model.

To streamline the subsequent calculations, we rewrite Eq. (4.90) and separate out all constant prefactors:

$$\Omega_{\text{GW}}\left(f\right) = \sum_{j=1}^{n_{\max}} \frac{\Gamma_j}{\Gamma} \Omega_{\text{GW}}^{(j)}(f) = \frac{8\pi}{3H_0^2}\left(G\mu\right)^2 \sum_{j=1}^{n_{\max}} \Gamma_j\, \mathcal{I}_j\left(f\right) , \tag{5.28}$$

where $\Gamma_j/\Gamma = j^{-q}/H_{n_{\max}}^{(q)}$ as before. The sole frequency and mode number dependence is then contained in $\mathcal{I}_j(f)$. Importantly, the function $\mathcal{I}_j$ inherits its properties from $\Omega_{\text{GW}}^{(j)}$ given in Eqs. (4.90) and (4.91):

$$\mathcal{I}_j(f) = \mathcal{I}_1\left(\frac{f}{j}\right) , \qquad \text{with} \qquad \mathcal{I}_1(f) = \frac{2}{f}\int_{t_{\text{ini}}}^{t_0} \mathrm{d}t \left(\frac{a\left(t\right)}{a_0}\right)^5 n\left(\frac{2}{f}\frac{a\left(t\right)}{a_0}, t\right) . \tag{5.29}$$

The loop number density $n(l,t)$ is given in Eq. (4.72).

As described in Sec. 4.3, we obtain the GWB spectrum by numerically solving the VOS equations in Eqs. (4.50) and (4.52) with a standard expansion history consisting of an early radiation-dominated, a late matter-dominated, and a very late cosmological-constant-dominated phase as

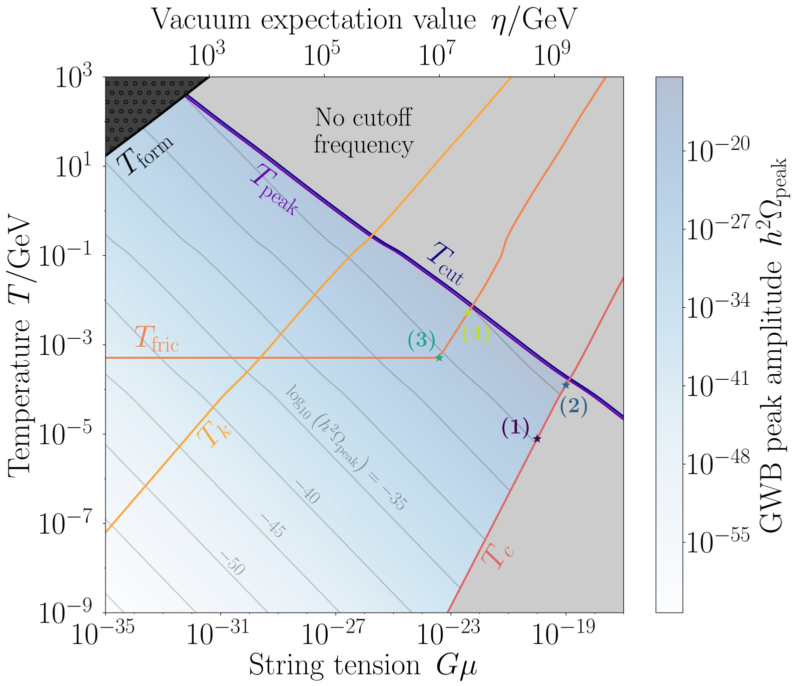


Figure 5.2: Relevant temperature scales as functions of the string tension $G\mu$; see T_cut in Eq. (5.11), T_form in Eq. (5.12), T_fric in Eq. (5.15), T_k in Eq. (5.22), T_c in Eq. (5.26), and T_peak in Eq. (5.41). The black region above T_form is entirely excluded as the string network has not yet formed for these parameter values. The blue shading and grey contours indicate the value of $h^2\Omega_\mathrm{peak}$ in Eq. (5.47) in dependence of $G\mu$ and T_ini, where T_ini is identified with T on the vertical axis. In realistic models, T_ini is not a free parameter, though, but must be identified with T_form, T_fric, T_k, or T_c. As explained in the text, the identifying T_ini with T_fric or T_c is particularly well motivated. In Sec. 5.2, we discuss four benchmark points for these two scenarios, whose $G\mu$ and T_ini values are summarised in Tab. 5.2 and shown above. *Figure reproduced from Ref. [3].*

described in detail around Eq. (4.56). Equipped with the explicit numerical time dependence of the scale factor $a(t)$ and the VOS variables $\xi(t)$ and $v_\infty(t)$, we can compute the total GWB spectrum given in Eq. (5.28) for any string tension $G\mu$ and any initial time t_ini, or equivalently initial temperature T_ini, without further approximations.

Based on the discussion in the previous section, our focus lies on parameter values for which T_ini is identified with T_fric or T_c. We consider two benchmark points for each of the two choices of the initial temperature. These points are specified in Tab. 5.2 and depicted as stars in Fig. 5.2. It should be emphasised that T_ini is not an actual free parameter of the theory. Once we identify T_ini with either T_fric or T_c, it is determined by the choice of $G\mu$. This is the way, in which the parameter values in Tab. 5.1 are obtained: $G\mu$ is freely chosen and T_ini is computed from this value of the string tension for points 1 and 2 by evaluating T_c in Eq. (5.26) and for points 3 and 4 by evaluating T_fric in Eq. (5.15). These benchmark points roughly cover the part of parameter space that is of phenomenological interest. Parameter points within this region could, for example, be realised if particle decay due to cusp formation is less efficient than described in the previous section. In this case, the curve describing T_c in Fig. 5.2 would be shifted towards larger temperatures, i.e., towards T_fric.

In Fig. 5.3, we show the total GWB spectra for the four benchmark points specified in Tab. 5.1 for $q = 4/3$, corresponding to GW emission due to gravitational radiation bursts from cusps. For the evaluation of the sum in Eq. (5.28), we set an upper cutoff on the mode number at $n_\mathrm{max} = 10^5$. This cutoff is large enough to understand the high-frequency behaviour of the spectrum while still allowing for numerical evaluation, even though the physical cutoff in the sum is much higher (for a

Table 5.1: Choice of the string tension $G\mu$ and initial temperature $T_{\rm ini}$ for the four benchmark points in Figs. 5.2 and 5.3.

Benchmark point	$\log_{10}(G\mu)$	$\log_{10}(T_{\rm ini}/\mathrm{GeV})$
1	-20.0	-5.107
2	-19.0	-3.903
3	-23.4	-3.292
4	-22.4	-2.277

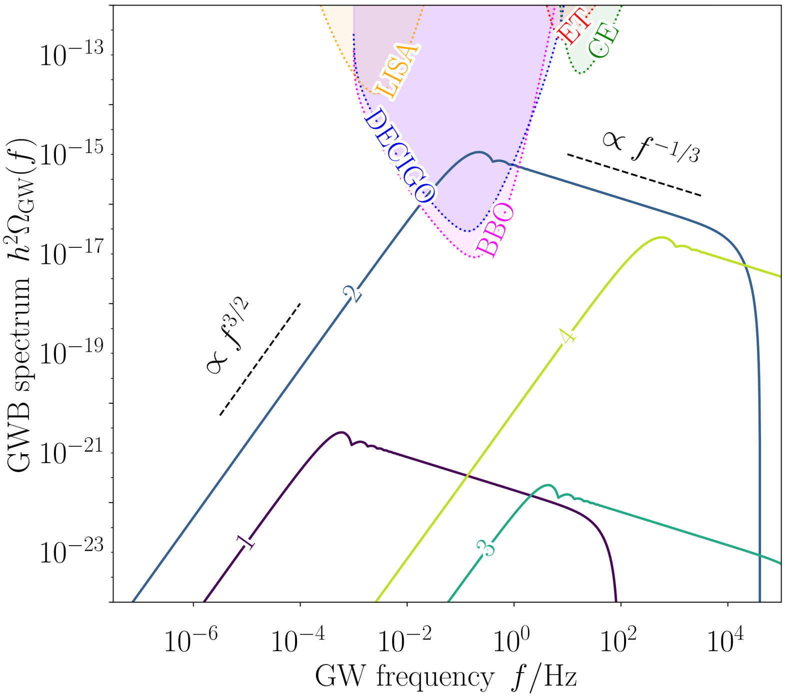


Figure 5.3: Numerical GWB spectra based on the VOS model for the four benchmark points indicated in Fig. 5.2 and specified in Tab. 5.1. Benchmark points 1 and 2 identify the initial temperature $T_{\rm ini}$ with $T_{\rm cusp}$, while benchmark points 3 and 4 identify $T_{\rm ini}$ with $T_{\rm fric}$. The dotted lines and shaded regions show the power-law-integrated sensitivity curves for future experiments and are taken from Ref. [419].
Figure adapted from Ref. [3].

detailed discussion of high-frequency effects of the mode summation, see Secs. 6.1.3 and 6.2.2).

As discussed in the beginning of this chapter, the GWB spectra exhibit indeed a characteristic oscillatory feature close to their peak. We emphasise that the spectra depicted in Fig. 5.3 are derived from the standard VOS framework without further ingredients for the parameter values in Tab. 5.1. These spectra, which are qualitatively very different from the well-known GWB spectra from high-scale strings (cf. Fig. 4.5) are the main result of this chapter. As is evident from the sensitivity curves in Fig. 5.3, the peak region of the spectra, including the characteristic feature of alternating local minima and maxima, may be observed by future space-borne interferometers such as BBO and DECIGO.

5.2.1 Analytical results

To understand the main features of the GWB spectra in Fig. 5.3, it is useful to follow an analytical approach. Let us begin by investigating the GWB spectrum from the fundamental oscillation mode

$$\Omega^{(1)}_{\rm GW}\left(f\right) = \frac{8\pi}{3H_0^2}\left(G\mu\right)^2\Gamma\,\mathcal{I}_1\left(f\right)\,. \tag{5.30}$$

As explained in Sec. 4.3.2, for analytical computations, it is useful to decompose the spectrum into three different contributions, namely from loops decaying during radiation domination (RR), loops produced during radiation domination and decaying during matter domination (RM), and loops produced as well as decaying during the matter-dominated era (MM). We will focus here on the RM spectrum, which dominates over the RR and the MM spectra. Since loops in the low-scale regime decay barely at all in a time span $\sim t_0$, they will lose only a fraction $\sim t_{\rm eq}/t_0$ of energy during the radiation-dominated era, which strongly suppresses the RR contribution. The MM spectrum is suppressed due to the comparatively small number density of loops born during matter domination.[3] This understanding is confirmed by the numerical computation of the complete GWB spectrum and in the following, we will only consider the RM contribution.

To compute the contribution of RM loops to $\Omega^{(1)}_{\rm GW}$, we need their loop number density. We have already computed it in Eq. (4.82) and reproduce the result here in slightly rewritten form:

$$n_{\rm rm}\left(l,t\right) = \Theta_{\rm rm}\left(l,t\right)\frac{\mathcal{F}\tilde{C}_r}{\alpha\xi_r}\frac{\left(\alpha\xi_r+\Gamma G\mu\right)^{3/2}}{t_{\rm eq}^{3/2}\left(l+\Gamma G\mu\,t\right)^{5/2}}\left(\frac{a_{\rm eq}}{a\left(t\right)}\right)^3\,. \tag{5.31}$$

This equation assumes pure radiation domination with a fixed number of effective degress of freedom g_ρ until matter–radiation equality at $t_{\rm eq}\simeq 1.59\times10^{12}\,$s with a corresponding scale factor $a_{\rm eq}\simeq a_0/3400$, immediately followed by a purely matter-dominated phase. We used the shorthand $\Theta_{\rm rm}(l,t)=\Theta\left(t-t_{\rm eq}\right)\Theta\left(t_{\rm eq}-t_*\right)\Theta\left(t_*-t_{\rm ini}\right)$ for the Heaviside functions introduced in Eq. (4.77b). The values for the VOS parameters are $\xi_r\simeq0.27$ and $v_r\simeq0.66$ from Eq. (4.55), and furthermore $\tilde{C}_r\simeq5.5$. Note that, as before, $\alpha=0.1/\xi_r\simeq0.37$, in agreement with $\alpha_h=\alpha\xi_r/2=0.05$.

Since we are only interested in string tensions that satisfy $\Gamma G\mu\ll\alpha\xi_r$, the loop number density can be further simplified to

$$n_{\rm rm}\left(l,t\right)\simeq\Theta_{\rm rm}(l,t)\frac{C_r}{t_{\rm eq}^{3/2}\left(l+\Gamma G\mu\,t\right)^{5/2}}\left(\frac{a_{\rm eq}}{a\left(t\right)}\right)^3\,, \tag{5.32}$$

where the normalisation $C_r\simeq\mathcal{F}\tilde{C}_r(\alpha\xi_r)^{1/2}\simeq0.17$ (cf. Eq. (4.80)) is in agreement with the results provided in Eq. (3.20) of Ref. [394].

Using this loop number density and that the scale factor during pure matter domination is of the form

$$a\left(t\right)=a_0\left(\frac{3}{2}H_m^0t\right)^{2/3}\,,\qquad\text{with}\qquad H_m^0=\Omega_m^{1/2}H_0\qquad\text{and}\qquad h^2\Omega_{\rm m}\simeq0.14\,, \tag{5.33}$$

we can analytically evaluate the integral $\mathcal{I}_1$ in Eq. (5.29) and obtain the fundamental GWB spectrum via Eq. (5.30). Let us first consider a low-frequency expansion. This results in an $f^{3/2}$ power law:

$$h^2\Omega^{(1)}_{\rm low}\left(f\right)\simeq A_{\rm low}f^{3/2}\simeq7.7\times10^{-23}\left(\frac{G\mu}{10^{-20}}\right)^2\left(\frac{f}{10^{-5}\,\rm Hz}\right)^{3/2}\,, \tag{5.34}$$

---

[3] In Ch. 6, the effect of this behaviour can be seen in the dependence of the amplitudes of the spectra on the string tension. The ratio of the amplitudes of the RM and the MM spectrum scales as $\mathcal{A}_{\rm mm}/\mathcal{A}_{\rm rm}\propto(G\mu)^{1/2}$ such that the MM contribution becomes increasingly less important for lower $G\mu$.

where we defined

$$A_{\mathrm{low}} = \frac{A}{2^{5/2} H_m^0} \left(1 - \frac{a_{\mathrm{eq}}}{a_0}\right) , \qquad \text{with} \qquad A = \frac{16\pi \, (G\mu)^2 \, \Gamma}{3 \, (H_0/h)^2} \frac{C_r}{t_{\mathrm{eq}}^{3/2}} \left(\frac{a_{\mathrm{eq}}}{a_0}\right)^3 . \tag{5.35}$$

Similarly, we can consider the high-frequency behaviour. For this, let us first assume that we are considering high-scale strings, i.e., string tensions $G\mu$ sufficiently large such that there is no high-frequency cutoff. We can then perform a high-frequency expansion of the GWB spectrum in Eq. (5.30) and find an f^{-1} power-law behaviour given by

$$h^2 \Omega_{\mathrm{high}}^{(1)} (f) \simeq A_{\mathrm{high}} f^{-1} \simeq 9.5 \times 10^{-17} \left(\frac{10^{-20}}{G\mu}\right)^{1/2} \left(\frac{10^5 \, \mathrm{Hz}}{f}\right)^{-1} , \tag{5.36}$$

where

$$A_{\mathrm{high}} = 4A \, \left(H_{\mathrm{m}}^0\right)^{3/2} \left(\frac{3}{2 \, \Gamma G\mu}\right)^{5/2} \left[\left(\frac{a_0}{a_{\mathrm{eq}}}\right)^{1/4} - 1\right] . \tag{5.37}$$

For frequencies between the low- and high-frequency regimes, the GWB spectrum is described by a broken power law that interpolates between the two asymptotic behaviours. We can estimate the peak frequency f_{peak} by equating the power-law expressions in Eqs. (5.34) and (5.36), which yields

$$h^2 \Omega_{\mathrm{low}}^{(1)} \left(f_{\mathrm{peak}}\right) = h^2 \Omega_{\mathrm{high}}^{(1)} \left(f_{\mathrm{peak}}\right) . \tag{5.38}$$

From this relation, we obtain

$$f_{\mathrm{peak}} = \frac{3 H_{\mathrm{m}}^0}{\Gamma G\mu} \left(4 \, \frac{\left(a_0/a_{\mathrm{eq}}\right)^{1/4} - 1}{1 - \left(a_{\mathrm{eq}}/a_0\right)}\right)^{2/5} \simeq 27 \, \mathrm{Hz} \left(\frac{G\mu}{10^{-20}}\right)^{-1} . \tag{5.39}$$

We can now compare this peak frequency to the cutoff frequency given in Eq. (5.6). If $f_{\mathrm{cut}} < f_{\mathrm{peak}}$, then the GWB spectrum does not show a broken power-law behaviour. Instead, it rises at low frequencies like $f^{3/2}$ and then sharply drops to zero at f_{cut}. We can rewrite $f_{\mathrm{cut}} < f_{\mathrm{peak}}$ as a condition on the initial time:

$$t_{\mathrm{ini}} > t_{\mathrm{peak}} = \frac{1}{\alpha_h f_{\mathrm{peak}}} + t_{\mathrm{cut}} , \tag{5.40}$$

where we used that $t_0 \gg t_{\mathrm{ini}}$. As before, this lower bound on the initial time corresponds via Eq. (5.9) to an upper bound on the initial temperature

$$T_{\mathrm{ini}} < T_{\mathrm{peak}} = \sqrt{\frac{M_*}{2 t_{\mathrm{peak}}}} . \tag{5.41}$$

We show this temperature as a function of the string tension $G\mu$ in violet in Fig. 5.2. As one can see, this temperature is always lower than T_{cut}, although the difference is only minor. This is a direct consequence of the fact that the requirement $t_{\mathrm{ini}} > t_{\mathrm{peak}}$ given in Eq. (5.40) is minimally stronger than the condition $t_{\mathrm{ini}} > t_{\mathrm{cut}}$, with t_{cut} given in Eq. (5.8). Technically, to obtain the well-pronounced ultraviolet cutoff in the fundamental GWB spectrum, we require the stronger condition $t_{\mathrm{ini}} > t_{\mathrm{peak}}$, which allows us to roughly approximate the spectrum as

$$h^2 \Omega_{\mathrm{GW}}^{(1)} (f) \simeq \Theta \left(f_{\mathrm{cut}} - f\right) \, h^2 \Omega_{\mathrm{low}}^{(1)} (f) . \tag{5.42}$$

However, in practice, it is sufficient to work with the condition $T_{\rm ini} < T_{\rm cut}$, since a strong fine-tuning of $G\mu$ is required to obtain a spectrum in the region where $T_{\rm peak} < T_{\rm ini} < T_{\rm cut}$. Hence, we will not investigate this scenario further.

With an analytical approximation for the GWB spectrum from the fundamental oscillation mode in Eq. (5.42) in place, we can continue and compute the total GWB spectrum from all modes. For this, we need to insert $\Omega_{\rm GW}^{(1)}$ into Eq. (5.28). We find

$$h^2\Omega_{\rm GW}\left(f\right) \simeq \frac{A_{\rm low}}{H_{n_{\rm max}}^{(q)}} \sum_{j=j_{\rm min}}^{n_{\rm max}} \frac{1}{j^q}\left(\frac{f}{j}\right)^{3/2} , \tag{5.43}$$

where the sum starts from the first number $j_{\rm min} \in \mathbb{N}$ that is larger than $f/f_{\rm cut}$, i.e., $j_{\rm min} = \lceil f/f_{\rm cut} \rceil$. The sum can be evaluated as

$$h^2\Omega_{\rm GW}\left(f\right) \simeq \frac{A_{\rm low}}{H_{n_{\rm max}}^{(q)}} \left[\zeta\left(q+3/2, j_{\rm min}\right) - \zeta\left(q+3/2, n_{\rm max}+1\right)\right] f^{3/2} , \tag{5.44}$$

where we used the definition of the Hurwitz zeta function $\zeta\left(p,a\right) = \sum_{j=0}^{\infty} 1/\left(j+a\right)^p$.

For $f \leq f_{\rm cut}$, the minimum mode number is $j_{\rm min} = 1$ and the factor containing the Hurwitz zeta functions takes on the constant value $H_{n_{\rm max}}^{(q+3/2)}$. Therefore, at low frequencies, the total spectrum is proportional to the fundamental GWB spectrum:

$$h^2\Omega_{\rm low}\left(f\right) \simeq \frac{H_{n_{\rm max}}^{(q+3/2)}}{H_{n_{\rm max}}^{(q)}} A_{\rm low}\, f^{3/2} . \tag{5.45}$$

Evaluation for our reference choice of parameters, $q = 4/3$ and $n_{\rm max} = 10^5$, yields

$$h^2\Omega_{\rm low}\left(f\right) \simeq 0.35 A_{\rm low} f^{3/2} . \tag{5.46}$$

This is indeed the $f^{3/2}$ power-law behaviour that can be observed at low frequencies for the total GWB spectra depicted in Fig. 5.3.

This result for the low-frequency behaviour of the total GWB spectrum can be used to obtain an estimate for the peak amplitude. For this, we evaluate the low-frequency expression at the cutoff frequency $f_{\rm cut}$, and obtain

$$h^2\Omega_{\rm peak} \simeq 0.35 A_{\rm low}\, f_{\rm cut}^{3/2} . \tag{5.47}$$

Observe that this expression depends via $A_{\rm low}$ on $G\mu$ and via $f_{\rm cut}$ on $t_{\rm ini}$ or, equivalently, on $T_{\rm ini}$ (see Eqs. 5.35 and 5.6).

Due to the large uncertainties in the correct choice of $T_{\rm ini}$ as a function of $G\mu$ as described in Sec. 5.1.2, it is often useful to treat $G\mu$ and $T_{\rm ini}$ as independent parameters. Following this approach, we can compute the peak amplitude $h^2\Omega_{\rm peak}$ for any pair of values $\left(G\mu, T_{\rm ini}\right)$. Identifying the temperature axis in Fig. 5.2 with the initial temperature $T_{\rm ini}$, the result is shown as a blue scale. Furthermore, Fig. 5.2 shows also the corresponding contour lines in grey. This demonstrates that the four benchmark points we chose in Tab. 5.1 represent the region of parameter space that is relevant for future observations. At smaller string tensions or lower initial temperatures, the GWB signal will be below the sensitivity reach of future GW observatories. As we have already seen, to benchmark point 2 corresponds a GWB spectrum from low-scale strings that can be detected by future experiments such as DECIGO or BBO. From the estimate of the peak amplitude, we also find that choices of parameter values around this benchmark point still lead to observable GWB signals: Although the GWB spectra from low-scale strings can only be detected for a small region

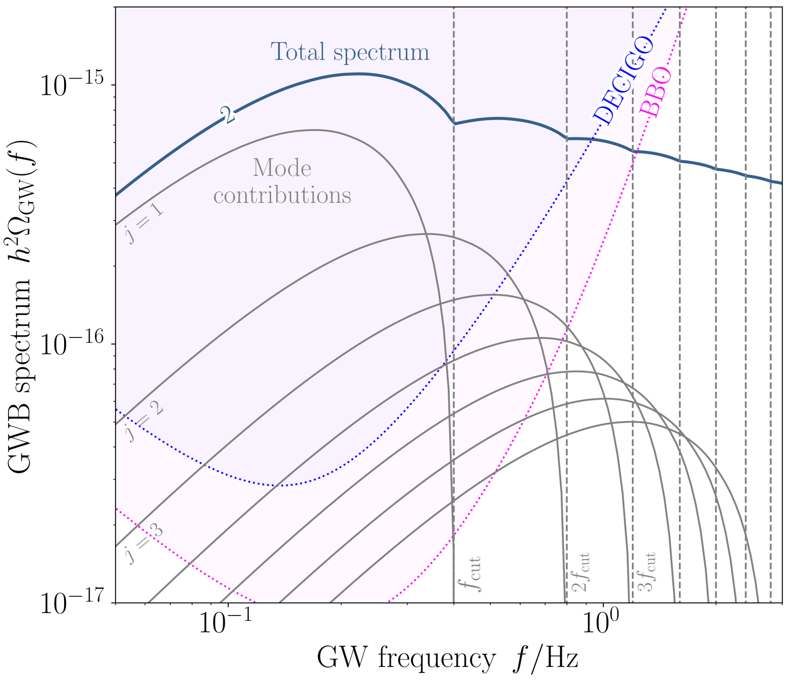


Figure 5.4: GWB spectrum for benchmark point 2; see Tab. 5.1 and Fig. 5.3. The solid blue line shows the total spectrum, while the thin grey lines indicate the respective contributions from harmonic string oscillations with $j = 1, 2, 3, \ldots$. The grey dashed vertical lines represent integer multiples of f_{cut} in Eq. (5.6).
Figure adapted from Ref. [3].

of parameter space by near future experiments, obtaining an observable peak amplitude does not require fine tuning.

So far, we have only considered the low-frequency limit. For intermediate frequencies $f_{\mathrm{cut}} \ll f \ll n_{\mathrm{max}} f_{\mathrm{cut}}$, we have the hierarchy $1 \ll j_{\mathrm{min}} \ll n_{\mathrm{max}}$, for which we can make the approximation

$$\zeta\left(q + 3/2, j_{\mathrm{min}}\right) - \zeta\left(q + 3/2, n_{\mathrm{max}} + 1\right) \simeq \frac{1}{q + 1/2}\left(\frac{f_{\mathrm{cut}}}{f}\right)^{q+1/2}, \tag{5.48}$$

to leading order in f_{cut}/f. For frequencies sufficiently above f_{cut}, but below $n_{\mathrm{max}} f_{\mathrm{cut}}$, the total GWB spectrum exhibits the following power-law behaviour:

$$h^2\Omega_{\mathrm{tail}}\left(f\right) \simeq \frac{A_{\mathrm{low}}\, f_{\mathrm{cut}}^{3/2}}{H_{k_{\mathrm{max}}}^{q}\left(q + 1/2\right)}\left(\frac{f_{\mathrm{cut}}}{f}\right)^{q-1}. \tag{5.49}$$

For the reference values $q = 4/3$ and $n_{\mathrm{max}} = 10^5$, this yields

$$h^2\Omega_{\mathrm{tail}}\left(f\right) \simeq 0.15 A_{\mathrm{low}}\, f_{\mathrm{cut}}^{3/2}\left(\frac{f_{\mathrm{cut}}}{f}\right)^{1/3}. \tag{5.50}$$

The spectral index $-1/3$ is indeed what we observe at frequencies above f_{cut} in the GWB spectra depicted in Fig. 5.3.

Lastly, we have to consider the frequency range $f > n_{\mathrm{max}} f_{\mathrm{cut}}$ for which $j_{\mathrm{min}} > n_{\mathrm{max}}$. In this case, the GWB spectrum vanishes and we can see from Fig. 5.3 that the $f^{-1/3}$ power-law abruptly drops to zero as f grows larger than $n_{\mathrm{max}} f_{\mathrm{cut}}$. However, note that the value $n_{\mathrm{max}} = 10^5$, which we have used for our plots, is an arbitrary choice. For strings of zero widths, we would have to send $n_{\mathrm{max}} \to \infty$. Nonetheless, realistic strings have a finite width and the spectrum indeed strictly vanishes above a certain frequency. As we show in more detail in Sec. 6.1.1, this frequency is however, far higher than the reach of any current or near future experiment, and for all practical purposes, we can use Eq. (5.49) to describe the GWB spectrum for frequencies $f \gg f_{\mathrm{cut}}$.

Our treatment also allows to understand the characteristic oscillatory feature of the total spectrum. This is best illustrated in Fig. 5.4, where we zoom in on the near-peak region of the total GWB spectrum that corresponds to benchmark point 2. In addition to the total spectrum given by a solid blue line, we show the separate mode contributions $h^2\Omega_{\mathrm{GW}}^{(j)}$ for the first seven modes as solid grey lines. The fundamental mode drops to zero at the frequency f_{cut} and strictly vanishes above. According to Eq. (5.28), the j-th mode similarly drops to zero as it approaches jf_{cut} and strictly vanishes above. As all these contributions add up to the total spectrum, the latter displays a series of dips and peaks for frequencies $f \gtrsim f_{\mathrm{cut}}$. The dips coincide with the frequencies $f_{\mathrm{cut}}, 2f_{\mathrm{cut}}, 3f_{\mathrm{cut}}, \ldots$ at which the first, second, third, ... modes drop to zero and no longer contribute to the total spectrum. The GWB spectrum therefore exhibits an extremely characteristic feature consisting of a sequence of dips located at integer multiples of the frequency of the first dip. Together with the low-frequency and high-frequency power laws with respective spectral indices $3/2$ and $1-q$, the GWB spectra of low-scale strings constitute a compelling and well-defined target for future GW experiments in the $\mathcal{O}\left(0.1\ldots1\right)$ Hz frequency band such as BBO and DECIGO.

5.2.2 Loop number and energy densities

The central assumption of our analysis is that we can consider string tensions so small that no loop formed in the early Universe has yet fully decayed via GW emission. As a result, all loops generated in the history of the string network are still present today. This naturally leads to the question of whether their continued existence could pose any cosmological issues. To investigate this, we will compute the current number and energy densities of string loops, demonstrating that their present-day abundance seems cosmologically unproblematic.

The total loop number density at a given time t is easily obtained as the integral

$$N\left(t\right) = \int_0^\infty \mathrm{d}l\, n\left(l,t\right) \,. \tag{5.51}$$

Recall that $n(l,t)\mathrm{d}l$ was the number density of loops with lengths in the interval $[l, l+\mathrm{d}l]$. We are ultimately interested in the present-day value of the total loop number density $N(t_0)$ and, therefore, in particular in $N(t)$ at times after matter–radiation equality. When evaluated for $t > t_{\mathrm{eq}}$, the RR contribution to $n(l,t)$ is zero and only n_{rm} and n_{mm} contribute (cf. Sec. 4.3.2). The RM contribution corresponds to loops produced before t_{eq}, the MM contribution to loops born after t_{eq}. As discussed previously, n_{mm} is strongly subdominant and we can, again, focus on n_{rm} only:

$$N\left(t\right) \simeq \int_0^\infty \mathrm{d}l\, n_{\mathrm{rm}}\left(l,t\right) \,. \tag{5.52}$$

We can use the explicit expression for n_{rm}, derived in the VOS model and provided in Eq. (5.31), to obtain

$$N\left(t\right) \simeq \frac{2C_r}{3\left(\alpha\xi_r t_{\mathrm{ini}} t_{\mathrm{eq}}\right)^{3/2}} \left(\frac{a_{\mathrm{eq}}}{a\left(t\right)}\right)^3 \,. \tag{5.53}$$

Evaluating this total loop number density at the present time t_0, we find roughly

$$N\left(t_0\right) \simeq \frac{49}{\mathrm{kpc}^3} \left(\frac{10^2\,\mathrm{s}}{t_{\mathrm{ini}}}\right)^{3/2} \,. \tag{5.54}$$

This total number density contains loops of any size, not only loops of astrophysical extent. As an illustration of this point, we can also compute the current density of loops with length equal or larger than 1 AU:

$$N_{>1\,\mathrm{AU}}\left(t_0\right) \simeq \int_{1\,\mathrm{AU}}^{\infty} \mathrm{d}l\, n_{\mathrm{rm}}\left(l, t_0\right) \simeq \frac{2 C_r}{3 t_{\mathrm{eq}}^{3/2}}\left(\frac{a_{\mathrm{eq}}}{a_0}\right)^3 \frac{1}{\left(1\,\mathrm{AU}+\Gamma G\mu\, t_0\right)^{3/2}}\,. \tag{5.55}$$

For all string tensions of interest, namely $G\mu \lesssim 10^{-19}$, this number density assumes a value of about 0.14 string loops per kpc^3, which appears to be harmless.

One might also ask whether string loops in the immediate neighbourhood of the solar system could give rise to observable effects, such as microlensing events [55–58] or GW bursts [83, 84, 420, 421]. However, we expect any such signals to be strongly suppressed by the low string tensions considered in this chapter. Still, exploring potential signatures of nearby low-scale strings may prove worthwhile in future studies.

In cosmology, the energy density of string loops is the more relevant quantity compared to their number density. Taking again only RM loops into account, the density parameter $\Omega_{\mathrm{loops}} = \rho_{\mathrm{loops}}/\rho_{\mathrm{crit}}$ is found to be

$$\Omega_{\mathrm{loops}}\left(t\right) \simeq \frac{\mu}{3H^2 M_{\mathrm{Pl}}^2}\int_0^{\infty} \mathrm{d}l\, l\, n_{\mathrm{rm}}\left(l, t\right) \simeq \frac{\mu}{3H^2 M_{\mathrm{Pl}}^2}\frac{2}{\left(\alpha \xi_r t_{\mathrm{ini}}\right)^{1/2}}\frac{C_r}{t_{\mathrm{eq}}^{3/2}}\left(\frac{a_{\mathrm{eq}}}{a\left(t\right)}\right)^3 . \tag{5.56}$$

Evaluated at the present time t_0, this yields

$$h^2 \Omega_{\mathrm{loops}}\left(t_0\right) \simeq 1.1 \times 10^{-13}\left(\frac{G\mu}{10^{-19}}\right)\left(\frac{10^2\,\mathrm{s}}{t_{\mathrm{ini}}}\right)^{1/2} . \tag{5.57}$$

In conclusion, the current value of the energy density in cosmic string loops is many orders of magnitude too small to cause an overclosure problem. Hence, as expected, low-scale cosmic strings, formed in the early Universe at symmetry-breaking scales $\eta \lesssim 10^9$ GeV and still present today, constitute a viable scenario.

5.3 GWB spectrum in the BOS model

The previous sections considered in detail a new feature of the GWB spectrum from cosmic strings. The spectra were computed in the standard way based on the VOS model, merely for parameters chosen from an unusual region of parameter space — the regime of low-scale strings. While the assumptions of the standard treatment might be appropriate for high-scale strings, they could be invalid for low-scale strings. In this and the following sections, we scrutinise those assumptions whose relaxation is expected to have the greatest impact on the GWB spectrum from low-scale strings. This includes the study of a modified loop evolution due to particle decay in nonscaling models and production of loops with lengths following a finite-width distribution. In this section, we explore the latter.

The sharp cutoff in the fundamental GWB spectrum at the frequency f_{cut} originates from the assumption that, once the string network has entered the scaling regime, loops are produced with a length l, which is a fixed fraction $\alpha_h \simeq 0.05$ of the particle horizon d_h. This assumption is commonly employed in the literature, as it largely simplifies numerical and analytical computations, without significant effects on the computed GWB spectra. However, for low-scale strings, this assumption plays a more critical role. The oscillatory feature in the total GWB spectrum from low-scale strings is a direct consequence of the sharpness of the cutoff in the fundamental GWB spectrum. Accordingly, it is also a consequence of working with a single production length $\alpha_h d_h$ (see Eq. (5.4)), as opposed to a broader distribution. For this reason, we now proceed to relax this assumption and generalise our calculation to incorporate a spectrum of initial lengths at a given time. For this, we use the distribution derived from the numerical simulations conducted by BOS in Ref. [47]. This will enable us to demonstrate that our findings based on the VOS model remain reliable, provided that the distribution of initial loop lengths remains sufficiently narrow, which is indeed the case in the BOS scenario.

Recall from Sec. 4.3.2 that the assumption about the initial loop length distribution enters the computation in the form of the loop production function $f(l,t)$. For a general loop production function, the loop number density is of the form given in Eq. (4.68), which we reproduce here for convenience:

$$n\left(l,t\right)=\int_{t_{\text{ini}}}^{t}\mathrm{d}t'\,f\left(l',t'\right)\left(\frac{a\left(t'\right)}{a\left(t\right)}\right)^{3}, \qquad \text{where} \qquad l'=l+\Gamma G\mu\left(t-t'\right). \tag{5.58}$$

When the string network has reached the scaling regime, the loop production function factorises,

$$f\left(l,t\right)=\frac{\tilde{f}\left(x\right)}{d_h^5}, \qquad \text{with} \qquad x=\frac{l}{d_h}, \tag{5.59}$$

where $d_h(t)$ carries the time dependence and $\tilde{f}(x)$ depends only on the scaling variable $x=l/d_h(t)$. The factor $1/d_h^5$ can directly be obtained from dimensional analysis, while the actual information on the initial distribution of the scaling variable x is contained in $\tilde{f}$.

Let us focus on loops that are produced during the radiation-dominated epoch and neglect changes in the effective number of degrees of freedom. Furthermore, let us begin with the loop number density evaluated during the same era, i.e., n_{rr}. We then have $d_h=2t$ and $a(t)\propto t^{1/2}$, such that Eq. (5.58) reduces to

$$n_{\text{rr}}\left(l,t\right)=\int_{t_{\text{ini}}}^{t}\mathrm{d}t'\,\frac{\tilde{f}\left(x'\right)}{\left(2t'\right)^5}\left(\frac{t'}{t}\right)^{3/2}, \qquad \text{where} \qquad x'=\frac{l'}{2t'}. \tag{5.60}$$

Using the relation for l' in Eq. (5.58) we find

$$t'=\frac{l+\Gamma G\mu t}{2x'+\Gamma G\mu}, \tag{5.61}$$

which allows us to rewrite the above integral as

$$n_{\text{rr}}\left(l,t\right)=\frac{\int_{x'(t)}^{x'(t_{\text{ini}})}\mathrm{d}x'\,\left(2x'+\Gamma G\mu\right)^{3/2}\tilde{f}\left(x'\right)}{2^4\,t^{3/2}\left(l+\Gamma G\mu t\right)^{5/2}}. \tag{5.62}$$

Here, we introduced the shorthand notation $x'\left(t_{\text{ini}}\right)$ and $x'\left(t\right)$ for

$$x'\left(t'\right)=\frac{l+\Gamma G\mu\left(t-t'\right)}{2t'}, \tag{5.63}$$

evaluated at $t'=t$ and $t'=t_{\text{ini}}$, respectively. Although we allow for a broader distribution of x values than previously, we still assume it to peak around α_h. Moreover, we are interested in small string tensions $\Gamma G\mu\ll 2\alpha_h$, such that we can comfortably neglect the $\Gamma G\mu$ term in the numerator of the integral over x'. Consequently, the loop number density simplifies to

$$n_{\text{rr}}\left(l,t\right)\simeq\frac{\mathcal{N}\left(t,t_{\text{ini}}\right)}{t^{3/2}\left(l+\Gamma G\mu t\right)^{5/2}}, \quad \text{with} \quad \mathcal{N}\left(t,t_{\text{ini}}\right)=\frac{1}{2^{5/2}}\int_{x'(t)}^{x'(t_{\text{ini}})}\mathrm{d}x'\,x'^{3/2}\tilde{f}\left(x'\right). \tag{5.64}$$

To continue, we must now specify the function $\tilde{f}(x)$. Our goal is to match it to the BOS distribution obtained from numerical simulations in Ref. [47]. BOS describe their distribution of initial values of the scaling variable x in terms of another density, $\rho\left(\ln x\right)$, which relates to $\tilde{f}(x)$ via[4]

$$\rho\left(\ln x\right)=\left[\frac{1}{2^{5/2}}\,x^{5/2}\tilde{f}\left(x\right)\right]_{x=e^{\ln x}}\,. \tag{5.65}$$

The normalisation $\mathcal{N}$ can then be rewritten in the equivalent form

$$\mathcal{N}\left(t,t_{\rm ini}\right)=\int_{\ln x'(t)}^{\ln x'(t_{\rm ini})}{\rm d}\left(\ln x'\right)\,\rho\left(\ln x'\right)\,. \tag{5.66}$$

In previous sections, we worked with the loop production function in Eq. (4.71), which is associated with the loops forming at a fixed fraction of the horizon scale. We can also express it by means of ρ and find

$$\rho\left(\ln x\right)\simeq C_r\,\delta\left(\ln\left(x/\alpha\right)\right)\,, \tag{5.67}$$

where, as before, $C_r\simeq\mathcal{F}\tilde{C}_r\left(\alpha\xi_r\right)^{1/2}\simeq 0.17$. With this standard loop production function of the VOS model, just rewritten in a different form, we can easily carry out the integration in Eq. (5.66) and find

$$\mathcal{N}\left(t,t_{\rm ini}\right)=C_r\Theta\left(t-t_*\right)\Theta\left(t_*-t_{\rm ini}\right)\,, \tag{5.68}$$

where t_* is, as usual, the loop formation time given in Eq. (4.78) with the identification $2\alpha_h=\alpha\xi_r$. This result, leads via Eq. (5.64) to the loop number density

$$n_{\rm rr}\left(l,t\right)\simeq\frac{C_r\,\Theta\left(t-t_*\right)\Theta\left(t_*-t_{\rm ini}\right)}{t^{3/2}\left(l+\Gamma G\mu t\right)^{5/2}}\,, \tag{5.69}$$

which, by construction, matches the density we previously found in Eq. (4.79).

We can now turn to our actual task of generalising our result to a wider distribution of initial loop lengths. For this, observe that the function ρ obtained from the simulations performed by BOS (see left panel of Fig. 2 of Ref. [47]), can be approximated by a Gaussian function with mean $\nu=\ln\alpha\simeq-3.0$ and standard deviation $\sigma\simeq 0.14$:

$$\rho\left(\ln x\right)\simeq\frac{C_r}{\sqrt{2\pi}\,\sigma}\,\exp\left(-\frac{\left(\ln x-\nu\right)^2}{2\sigma^2}\right)\,. \tag{5.70}$$

The corresponding loop production function is of the form

$$f\left(l,t\right)=\frac{C_r}{\sqrt{2\pi}\,\sigma\,l^{5/2}t^{5/2}}\,\exp\left(-\frac{\left(\ln\left(\frac{l}{2t}\right)-\nu\right)^2}{2\sigma^2}\right)\,. \tag{5.71}$$

The density function ρ in Eq. (5.70) offers a useful compromise. Although it is not the exact numerical BOS distribution, it accurately reflects the distribution's finite width near its maximum. Moreover, it remains sufficiently simple to permit an entirely analytical evaluation of the RR loop

---

[4] In Ref. [47], $x^{5/2}\tilde{f}\left(x\right)$ is denoted by $\alpha^{5/2}f_r\left(\alpha\right)$.

number density. Let us first perform the integral in Eq. (5.66) with the density function given in Eq. (5.70):

$$\mathcal{N}\left(t, t_{\text{ini}}\right) = \frac{C_r}{2} \operatorname{erf}\left(\frac{\ln x' - \nu}{\sqrt{2}\sigma}\right)\Bigg|_{\ln x'(t)}^{\ln x'(t_{\text{ini}})} . \tag{5.72}$$

Here, $\operatorname{erf}(z)$ is the Gauss error function. As a cross-check, it is easy to verify that the above expression for $\mathcal{N}$ reduces to the standard result in Eq. (5.68) upon taking the limit $\sigma \to 0$.

Using this normalisation in Eq. (5.64), we find the generalised RR loop number density

$$n_{\text{rr}}\left(l, t\right) \simeq \frac{C_r\left(\text{erf}_{t_{\text{ini}}} - \text{erf}_t\right)}{2t^{3/2}\left(l + \Gamma G\mu t\right)^{5/2}}, \qquad \text{where} \qquad \text{erf}_{t'} = \operatorname{erf}\left(\frac{\ln x'\left(t'\right) - \nu}{\sqrt{2}\sigma}\right) . \tag{5.73}$$

This loop number density accounts for the finite width of distribution of initial x values and can be matched to the numerical simulations performed by BOS, upon using the above values for ν and σ.

This result motivates two remarks. First, we emphasise that our analysis relies on approximating the actual distribution by a log-normal distribution given in Eq. (5.70). This choice is motivated by our interest in obtaining analytical expressions that can, in a certain limit, be reduced to the standard scenario of a single initial loop length, here in the limit $\sigma \to 0$. Nevertheless, the framework developed in this section is completely general and could, in principle, accommodate a fully numerical evaluation of the normalisation factor $\mathcal{N}$ in Eq. (5.66). For instance, one may use the exact numerical form of ρ provided by BOS in Ref. [47], or any distribution obtained from future numerical simulations. Instead of pursuing a numerical approach, we will, however, proceed with the analytical expressions derived above. In fact, we will employ these to explore a range of density functions associated with different widths σ. This enables us to assess the uncertainty inherent in our predictions of the GWB spectrum from low-scale strings.

Second, we note that the expression for the RR loop number density in Eq. (5.73) has applications beyond low-scale strings. In principle, any study making use of BOS-based loop number densities, should favour the more accurate number densities in Eq. (5.73) over those in Eq. (5.69). In practice, however, the numerical differences become relevant only for low-scale strings. For high-scale strings, we have explicitly verified that all corrections to the GWB spectrum introduced by the improved loop number density remain below the one-percent level.

Let us now return to our primary goal of assessing the impact of the finite-width distribution on the GWB spectra from low-scale strings. For their computation, we need the RM loop number density, rather than the RR loop number density. The former is easily obtained from the latter, following the procedure outlined around Eq. (4.81). We find

$$n_{\text{rm}}\left(l, t\right) = n_{\text{rr}}\left(l_{\text{eq}}, t_{\text{eq}}\right)\left(\frac{a_{\text{eq}}}{a\left(t\right)}\right)^3, \qquad \text{where} \qquad l_{\text{eq}} = l + \Gamma G\mu\left(t - t_{\text{eq}}\right) . \tag{5.74}$$

More explicitly, with n_{rr} from Eq. (5.73), this evaluates to

$$n_{\text{rm}}\left(l, t\right) \simeq \frac{C_r\left(\text{erf}_{t_{\text{ini}}} - \text{erf}_{t_{\text{eq}}}\right)}{2t_{\text{eq}}^{3/2}\left(l + \Gamma G\mu t\right)^{5/2}}\left(\frac{a_{\text{eq}}}{a\left(t\right)}\right)^3 , \tag{5.75}$$

where $\text{erf}_{t_{\text{eq}}}$ is erf_t evaluated at $t = t_{\text{eq}}$. As for the RR case, this RM loop number density captures results of the numerical simulations by BOS more accurately than the standard loop number density given in Eq. (5.32).

The above expression can further be simplified if we consider the Universe to be purely radiation dominated for $t \leq t_{\text{eq}}$ such that the scale factor takes the form

$$a(t) = a_0\left(2H_r^0 t\right)^{1/2}, \qquad \text{with} \qquad H_r^0 = H_0\Omega_r^{1/2} \qquad \text{and} \qquad h^2\Omega_{\text{r}} \simeq 4.2 \times 10^{-5}, \tag{5.76}$$

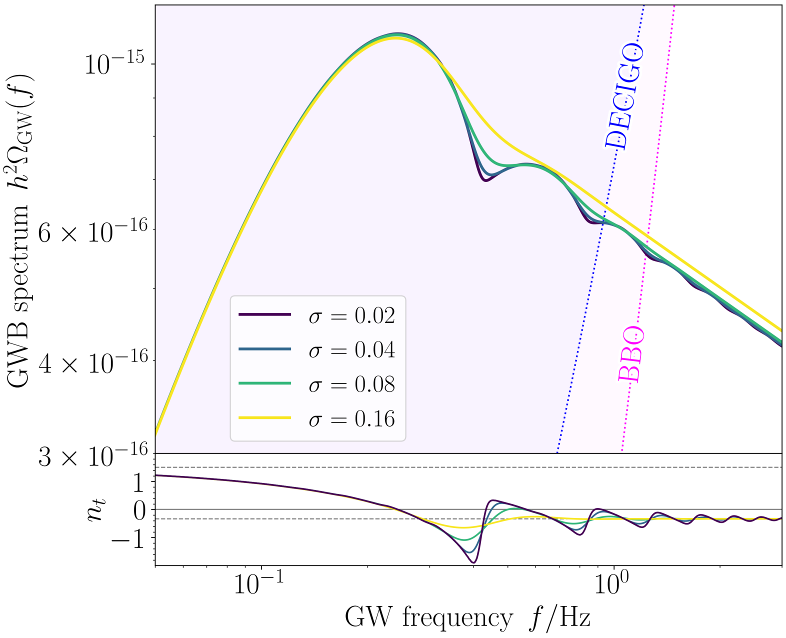


Figure 5.5: GWB spectrum for benchmark point 2 in Tab. 5.1 based on the generalised BOS loop number density n_{rm} in Eq. (5.78), which assumes a log-normal distribution of the loop lengths at the time of formation. The mean of the distribution is fixed at $\nu = -3.0$, while its standard deviation is varied, $\sigma \in \{0.02, 0.04, 0.08, 0.16\}$. The lower panel shows the spectral index n_t as a function of frequency as defined in Eq. (5.79). The horizontal dashed lines in this panel indicate the asymptotic values of the spectral index: $n_t \to 3/2$ at low frequencies and $n_t \to -1/3$ at high frequencies, as seen in the GWB spectra in Fig. 5.3.
Figure adapted from Ref. [3].

and evaluates at matter–radiation equality to $a_{\mathrm{eq}} = a_0 \left(2H_r^0 t_{\mathrm{eq}}\right)^{1/2}$. Assuming also a purely matter-dominated Universe for $t > t_{\mathrm{eq}}$ and the corresponding scale-factor evolution in Eq. (5.33), we find that

$$n_{\mathrm{rm}}(l,t) \simeq R\,\frac{C_r\left(\mathrm{erf}_{t_{\mathrm{ini}}} - \mathrm{erf}_{t_{\mathrm{eq}}}\right)}{2t^2\left(l + \Gamma G\mu t\right)^{5/2}}\,, \qquad \text{with} \qquad R = \frac{\left(2H_r^0\right)^{3/2}}{\left(3/2H_m^0\right)^2}\,. \tag{5.77}$$

The approximation of $a_{\mathrm{eq}}/a(t)$ made to obtain Eq. (5.77) from Eq. (5.75) does not properly account for the smooth transition from the radiation-dominated epoch to matter domination. Hence, the prefactor R is slightly inaccurate which, in turn, leads to an inaccuracy in the amplitude of the GWB spectrum. The magnitude of this effect can be determined by considering Eq. (5.77) in the limit $\sigma \to 0$ and computing its corresponding GWB spectrum. We can then compare it to the GWB spectrum obtained with the standard loop number density in Eq. (4.72) and a fully numerical evaluation of the scale factor evolution. From this, we find that R overestimates the actual prefactor by a factor of 1.5. We can compensate for this by substituting $R' = R/1.5$ for R in Eq. (5.77), such that

$$n_{\mathrm{rm}}(l,t) \simeq R'\,\frac{C_r\left(\mathrm{erf}_{t_{\mathrm{ini}}} - \mathrm{erf}_{t_{\mathrm{eq}}}\right)}{2t^2\left(l + \Gamma G\mu t\right)^{5/2}}\,. \tag{5.78}$$

This constitutes our final result for the RM loop number density that incorporates the log-normal distribution of initial x values based on the BOS model.

Finally, we employ the loop number density from Eq. (5.78) to evaluate the resulting GWB spectrum for a log-normal distribution of initial x values, characterised by a fixed mean $\nu = -3.0$

and four different standard deviations $\sigma \in \{\, 0.02, 0.04, 0.08, 0.16 \,\}$. The result is depicted in Fig. 5.5. The range of widths of the distributions includes the value $\sigma \simeq 0.14$, which we obtained from Fig. 2 in Ref. [47], while also spanning the range of expected variation across different models. As illustrated in Fig. 5.5, increasing the width of the initial-loop-length distribution tends to smooth out the characteristic pattern of peaks and dips in the spectrum. The lower panel of Fig. 5.5 displays the spectral index of the GWB as a function of frequency,

$$n_t = \frac{\mathrm{d}\ln h^2\Omega_{\mathrm{GW}}\,(f)}{\mathrm{d}\ln f}\,, \tag{5.79}$$

which allows us to highlight the oscillations of the spectrum around a spectral index of $n_t = -1/3$ for frequencies $f \gtrsim f_{\mathrm{cut}}$. This behaviour persists even for $\sigma = 0.16$, suggesting that experiments such as BBO and DECIGO could potentially observe at least one local maximum and one local minimum in n_t. As long as the distribution of initial loop lengths remains sufficiently narrow, the distinctive features predicted in the VOS model for a fixed initial loop size are preserved in the more realistic BOS scenario.

5.4 Particle decay and scale dependence

Let us return to the case where, at a given time, all loops are produced with the same length:

$$l_*(t) = \alpha_h d_h(t) = 2\alpha_h t\,. \tag{5.80}$$

So far, we described the loop evolution via the scale-independent rate $\mathrm{d}l/\mathrm{d}t = -\Gamma G\mu$ resulting from GW emission. As we discussed in Sec. 5.1.2, for loops shorter than a critical length, this dynamics can drastically change as the emission of particle radiation due to kink–kink collisions or cusp formation becomes the dominant decay channel. The aim of the remaining sections of this chapter is to compute the impact of this explicit scale dependence in the loop decay rate on the GWB spectra from low-scale strings. We will work with the nonscaling description of the loop length evolution presented in Ref. [64].

Since, in the following, particle decay of loops plays an important role, we re-emphasise that our understanding of low-scale cosmic strings relies on the assumption that GW emission from string loops can dominate over energy loss via particle emission for long periods of time. In the literature, there is actually a long-lasting debate on this point, in particular, on the question of whether loops mainly decay gravitationally, which is realised in the Nambu–Goto limit, or whether particle decay is dominant on all scales. For recent articles on this issue see, e.g., Refs. [344, 406–410]. In this sense, our analysis is more closely aligned with the usual discussion of cosmic strings in the Nambu–Goto approximation.

5.4.1 Modified loop length evolution

To describe the possible effects of particle radiation on the string length evolution, recall that the string tension is constant and the energy of a string of length l is given as $E = \mu l$, such that we can immediately conclude that a string that loses energy with total power $P(l)$ satisfies

$$\frac{\mathrm{d}l}{\mathrm{d}t} = -\frac{P(l)}{\mu}\,. \tag{5.81}$$

Sufficiently large string loops decay predominantly into gravitational radiation:

$$P_{\mathrm{NG}} = \Gamma G\mu^2\,. \tag{5.82}$$

As discussed in Sec. 5.1.2, specifically Eqs. (5.19) and (5.23), the power emitted into particle radiation due to kink–kink collisions and cusp formation can be estimated as

$$P_k \sim \frac{N_k \mu^{1/2}}{l}\,, \qquad\qquad P_c \sim \frac{N_c \mu^{3/4}}{l^{1/2}}\,. \tag{5.83}$$

The different l-dependencies of P_k and $P_{\rm NG}$ as well as P_c and $P_{\rm NG}$, lead to particle radiation becoming the dominant decay channel for short loops of length $l < l_{k(c)}$ given in Eqs. (5.20) and (5.24), respectively.

The total emitted power will be a function that interpolates between the asymptotic expressions in Eq. (5.82) for $l \gg l_{k(c)}$ and in Eq. (5.83) for $l \ll l_{k(c)}$. In Ref. [64], this fact was incorporated by a modification of the string loop's length evolution:

$$\frac{\mathrm{d}l}{\mathrm{d}t} = -\Gamma G\mu\, \mathcal{J}(l)\,, \tag{5.84}$$

where

$$\mathcal{J}(l) = \begin{cases} 1 & \text{for NG loops} \\ \left(1 + \left(\frac{l_k}{l}\right)^2\right)^{1/2} & \text{for kinky loops} \\ \left(1 + \left(\frac{l_c}{l}\right)^{3/2}\right)^{1/3} & \text{for cuspy loops} \end{cases}\,. \tag{5.85}$$

This expression explicitly depends on the loop length l and the two length scales l_k and l_c, signalling the breakdown of the usual scaling regime. Until now, we extensively discussed the first of these three cases (i.e., loops that behave as NG loops in the GW-dominated decay regime). In the following, our goal is to supplement this with a discussion of "kinky" and "cuspy" loops.

We continue by reproducing some results found in Ref. [64] that will be needed afterwards. We can straightforwardly integrate Eq. (5.84) and obtain

$$\zeta(l(t)) + \Gamma G\mu t = \zeta(l(t')) + \Gamma G\mu t'\,, \tag{5.86}$$

where $\zeta(l)$ is an auxiliary function defined as

$$\zeta(l) = \int^l \frac{\mathrm{d}l'}{\mathcal{J}(l')}\,. \tag{5.87}$$

Upon inserting the different forms of $\mathcal{J}$ in Eq. (5.85) into the integral, one promptly obtains ζ for the different cases:

$$\zeta(l) = \begin{cases} l & \text{for NG loops} \\ \left(l^2 + l_k^2\right)^{1/2} & \text{for kinky loops} \\ \left(l^{3/2} + l_c^{3/2}\right)^{2/3} & \text{for cuspy loops} \end{cases}\,. \tag{5.88}$$

Observe that for the NG case, we obtain from Eq. (5.86) with $\zeta(l) = l$ the standard loop length evolution in scaling, given, e.g., on the right-hand side of Eq. (5.58).

5.4.2 Criteria for a cutoff frequency

Next, we want to repeat the discussion of Sec. 5.1.1, but this time with the modified loop length evolution. In particular, we determine the criteria which allow us to identify the regions of parameter space in which the fundamental GWB spectrum is sharply cutoff at a frequency $f_{\rm cut}$ and vanishes for higher frequencies. Recall that a cutoff will occur if none of the string loops has fully decayed until today. The shortest loops are born with a length $l_{\text{ini}} = l_*(t_{\text{ini}}) = 2\alpha_h t_{\text{ini}}$, and since they also have the longest time to decay, their current length sets the minimum length of all string loops l_{min}. From Eq. (5.86), this minimum length is implicitly given as the solution to the equation

$$\zeta(l_{\text{min}}) = \zeta(l_{\text{ini}}) - \Gamma G\mu\,(t_0 - t_{\text{ini}}) \,. \tag{5.89}$$

Note that this is only true if $l_{\text{min}} > 0$, i.e., only for low-scale strings. If $l_{\text{min}} = 0$ then t_0 needs to be replaced by $t_{\text{max}} = \left[\zeta(l_{\text{ini}}) - l_{k(c)}\right]/(\Gamma G\mu) - t_{\text{ini}}$. Generally, the equation is only meaningful if $\zeta(l_{\text{ini}}) > \Gamma G\mu(t_{\text{max}} - t_{\text{ini}})$. Using the explicit forms of ζ given in Eq. (5.88), one readily finds for the specific cases

$$l_{\text{min}} = \begin{cases} l_{\text{ini}} - \Gamma G\mu\,(t_0 - t_{\text{ini}}) & \text{for NG loops} \\ \left(\left(\zeta_k(l_{\text{ini}}) - \Gamma G\mu\,(t_0 - t_{\text{ini}})\right)^2 - l_k^2\right)^{1/2} & \text{for kinky loops} \\ \left(\left(\zeta_c(l_{\text{ini}}) - \Gamma G\mu\,(t_0 - t_{\text{ini}})\right)^{3/2} - l_c^{3/2}\right)^{2/3} & \text{for cuspy loops} \end{cases} . \tag{5.90}$$

The expression for the NG case coincides, of course, with our earlier result in Eq. (5.5). Due to the large hierarchy between the initial time and the present time, we will henceforth use the approximation $t_0 - t_{\text{ini}} \simeq t_0$ for the minimum length in all three expressions. Observe that for $\Gamma G\mu t_0 \gg l_{k(c)}$, we return approximately to the NG case. This is expected since, in this case, most of the string's length and, thus, energy is lost into gravitational radiation and not particle radiation. Reduction of the initial length due to the latter is correspondingly negligible.

Let us elaborate a bit more on this point to avoid confusion. One might naively expect that loops with lengths $l < l_{k(c)}$ quickly evaporate into particle radiation, and gravitational radiation becomes negligible. Since $l_{k(c)}$ grows strongly with decreasing $G\mu$, the particle-decay-dominated regime is usually entered for low string tensions, which can strongly suppress any kind of emitted radiation. Thus, even though strings might have lengths $l < l_{k(c)}$, they may hardly decay at all. To make this statement more quantitative, we provide here expressions for the energy a string produced with length l_* loses into gravitational and particle radiation, respectively. To condense the notation, let us denote the typical length associated with the emission of GWs by

$$l_0^{\text{GW}} = \Gamma G\mu t_0 \,. \tag{5.91}$$

With this, we find

$$E_{\text{GW}} = -\mu \int_{l_*}^{l_{\text{min}}} \frac{\mathrm{d}l}{\mathcal{J}(l)} = \mu\left[\zeta(l_*) - \zeta(l_{\text{min}})\right] = \mu \begin{cases} l_0^{\text{GW}} & \text{if } l_{\text{min}} > 0 \\ \zeta(l_*) - l_{k(c)} & \text{if } l_{\text{min}} = 0 \end{cases} , \tag{5.92}$$

$$E_{\text{part}} = \int_{l_*}^{l_{\text{min}}} \Gamma G\mu^2 \left[\mathcal{J}(l) - 1\right] \left(\frac{\mathrm{d}l}{\mathrm{d}t}\right)^{-1} \mathrm{d}l = \tag{5.93}$$

$$= \mu\left[l_* - l_{\text{min}}\right] - E_{\text{GW}} = \mu \begin{cases} l_* - l_0^{\text{GW}} - l_{\text{min}} & \text{if } l_{\text{min}} > 0 \\ l_* + l_{k(c)} - \zeta(l_*) & \text{if } l_{\text{min}} = 0 \end{cases} .$$

Here, l_* is generally the length of loop birth and our focus is on the shortest loops with $l_* = l_{\text{ini}}$. Note that E_{GW} remains unaffected by particle decay as long as $l_{\text{min}} > 0$, i.e., as long as strings do not completely decay until today. This is expected, as the power emitted into GWs only depends on the string tension.

With an expression for the minimal length at hand, the criterion for the existence of a maximum frequency is $l_{\text{min}} > 0$. We can generally rephrase this as a condition on the initial length and find that a maximum frequency exists if $l_{\rm ini} > l_{\rm cut}$ and find from Eq. (5.90) that

$$l_{\rm cut} = \begin{cases} l_0^{\rm GW} & \text{for NG loops} \\ l_0^{\rm GW}\left(1+\frac{2l_k}{l_0^{\rm GW}}\right)^{1/2} & \text{for kinky loops} \\ l_0^{\rm GW}\left(\left(1+\frac{l_c}{l_0^{\rm GW}}\right)^{3/2}-\left(\frac{l_c}{l_0^{\rm GW}}\right)^{3/2}\right)^{2/3} & \text{for cuspy loops} \end{cases} . \tag{5.94}$$

The lower bound on $l_{\rm ini}$ in Eq. (5.94) interpolates for kinky and cuspy strings between two behaviours: For $l_{k(c)} \ll \Gamma G\mu t_0$, we recover the NG limit, as discussed before. For $l_{k(c)} \gg \Gamma G\mu t_0$, loop decay is dominated by particle radiation, and gravitational decay is negligible. To see this, note that the bound $l_{\rm ini} > l_{\rm cut}$ follows from setting $l_{\rm min} = 0$ while $t_{\rm max} = t_0$. Using the expressions for the energy emitted due to gravitational and particle decay in Eqs. (5.92) and (5.93), we find $E_{\rm part}/E_{\rm GW} = l_{\rm ini}/(\Gamma G\mu t_0) - 1 \simeq (2l_k/(\Gamma G\mu t_0))^{1/2} \gg 1$ and $E_{\rm part}/E_{\rm GW} \simeq (9l_c/(4\Gamma G\mu t_0))^{1/3} \gg 1$ for kinky and cuspy loops, respectively.

Identifying the string tension as our remaining free parameter after having fixed $t_{\rm ini} = t_{\rm cut}$, we can determine the critical string tension $G\mu_{\rm crit}$, at which we transition from the regime dominated by gravitational decay $G\mu \gg G\mu_{\rm crit}$ to the particle-decay-dominated regime $G\mu \ll G\mu_{\rm crit}$. These critical string tensions are for kink–kink-collision and cusp-induced particle decay, respectively, given by

$$G\mu^k_{\rm crit} = \left(\frac{N_k}{\Gamma^2}\frac{t_{\rm Pl}}{t_0}\right)^{2/5} \simeq 1.90\times 10^{-26}\left(\frac{N_k}{1}\right)^{2/5}\left(\frac{\Gamma}{50}\right)^{-4/5} , \tag{5.95}$$

$$G\mu^c_{\rm crit} = \left(\frac{N_c^2}{\Gamma^3}\frac{t_{\rm Pl}}{t_0}\right)^{2/7} \simeq 1.39\times 10^{-19}\left(\frac{N_c}{1}\right)^{4/7}\left(\frac{\Gamma}{50}\right)^{-6/7} . \tag{5.96}$$

All initial lengths $l_{\rm ini}$ can, following Sec. 5.1.1 and specifically Eq. (5.10), be related to initial times $t_{\rm ini} = l_{\rm ini}/(2\alpha_h)$ and initial temperatures $T_{\rm ini}$. Expressed in terms of these, the criterion for the existence of a maximum frequency are the equivalent conditions

$$t_{\rm ini} > t_{\rm cut} \qquad \text{and} \qquad T_{\rm ini} < T_{\rm cut} . \tag{5.97}$$

For the case of NG loops, we obtained the values for initial times and temperatures that set the boundary between the regime of low-scale and high-scale strings in Eqs. (5.8) and (5.11). For $G\mu \gg G\mu^{k(c)}_{\rm crit}$, these expressions also hold true for kinky and cuspy loops. Far below the critical string tension $G\mu \ll G\mu_{\rm crit}$, we obtain a very different parameter dependence of the cutoff times and temperatures. For kinky loops, we find

$$t^k_{\rm cut} \simeq \frac{(N_k t_{\rm Pl} t_0/2)^{1/2}}{\alpha_h\,(G\mu)^{1/4}} \simeq 6.9\times 10^{-5}\,{\rm s}\left(\frac{G\mu}{10^{-30}}\right)^{-1/4} , \tag{5.98}$$

$$T^k_{\rm cut} \simeq 94\,{\rm MeV}\left(\frac{G\mu}{10^{-30}}\right)^{1/8} , \tag{5.99}$$

whereas for cuspy loops, we have

$$t^c_{\rm cut} \simeq \frac{\left(3N_c t_0 t_{\rm Pl}^{1/2}\right)^{1/3}}{2\alpha_h\,(G\mu)^{1/6}} \simeq 2.8\times 10^{3}\,{\rm s}\left(\frac{G\mu}{10^{-30}}\right)^{-1/6} , \tag{5.100}$$

$$T^c_{\rm cut} \simeq 21\,{\rm keV}\left(\frac{G\mu}{10^{-30}}\right)^{1/12} . \tag{5.101}$$

Table 5.2: Choice of the string tension $G\mu$ and initial temperature $T_{\rm ini}$ for nine benchmark points (BMPs), depicted by stars in the parameter space in Fig. 5.6. In Fig. 5.7, we show the corresponding GWB spectra. Additionally, we list the cutoff frequencies $f_{\rm cut}^{{\rm NG}(k,c)} = 2/l_{\rm min}^{{\rm NG}(k,c)}$ according to Eq. (5.90) for the cases of NG, kinky, and cuspy loops, respectively.

BMP	$\log_{10}(G\mu)$	$\log_{10}\left(\frac{T_i}{\rm GeV}\right)$	$\log_{10}\left(\frac{f_{\rm cut}^{\rm NG}}{\rm Hz}\right)$	$\log_{10}\left(\frac{f_{\rm cut}^{k}}{\rm Hz}\right)$	$\log_{10}\left(\frac{f_{\rm cut}^{c}}{\rm Hz}\right)$
1	−20.0	−5.107	−3.04	−3.04	−3.04
1_A	−20.0	−4.607	−2.04	−2.04	−2.04
1_B	−20.0	−4.107	−1.03	−1.03	−0.98
1_C	−20.0	−3.607	0.22	0.22	—
1_D	−20.0	−3.107	—	—	—
2	−19.0	−3.903	−0.39	−0.39	−0.33
3^*	−22.7	−2.577	2.29	2.29	—
4	−22.4	−2.277	3.03	3.03	—
5	−22.1	−1.977	—	—	—

Treating $T_{\rm ini}$ and $G\mu$ as before as free parameters, the above expressions yield lines in a two-dimensional parameter space, separating the region in which a cutoff frequency occurs and the region in which the fundamental GWB spectrum extends to arbitrarily large frequencies (i.e., up to an ultra-high frequency cutoff mentioned previously).

In Fig. 5.6, we plot contour lines for $T_{\rm cut}^{k(c)}$ in blue as obtained from combining Eqs. (5.94) and (5.10) for kinky and cuspy loops, respectively. We also plot $G\mu_{\rm crit}^{k(c)}$ as magenta dotted lines, which illustrate the transition in the behaviour of $T_{\rm cut}^{k(c)}$ from string tensions below the critical tension described by Eqs. (5.99) and (5.101), respectively, to the NG regime described by Eq. (5.11). On the one hand, the relevant parameter space of string loops affected by kink–kink collisions remains in large parts unchanged in comparison to NG loops, which is a consequence of the low value of $G\mu_{\rm crit}^{k}$. On the other hand, we find for cuspy loops a value of $G\mu_{\rm crit}^{c}$ that is relatively large for low-scale strings, and the low-scale-string parameter space is severely reduced (cf. Fig. 5.2).

In the parameter space spanned by $G\mu$ and $T_{\rm ini}$, we choose five benchmark points 1 to 5 for which we shall compute the GWB spectra numerically. Note that points $1, 2$ and 4 coincide with the points given in Tab. 5.1, while we have replaced point 3 with point 3^*.[5] The points are selected in such a way that they cover a broad range of qualitatively different regions of the parameter space. To show the transition from the low-scale to the high-scale regime for a fixed string tension $G\mu$, we also choose four benchmark points 1_A to 1_D whose initial temperature increases by factors of $10^{0.5} \simeq 3.16$ starting from benchmark point 1. These nine benchmark points are summarised in Tab. 5.2 and indicated by stars in Fig. 5.6.

5.4.3 Initial loop length in the nonscaling model

In Sec. 5.1.2, we have discussed different choices for the initial time $t_{\rm ini}$. The network formation time $t_{\rm form}$ and the time of the end of thermal friction domination $t_{\rm fric}$ provide well-motivated choices for the initial time. However, if particle radiation is not directly accounted for at the level of the function $\mathcal{J}$ in the differential equation for $\mathrm{d}l/\mathrm{d}t$, these cutoffs neglect the effect that particle decay can become more dominant than gravitational decay for $l < l_{k(c)}$. In Sec. 5.1.2, we addressed this point by introducing two additional choices for the initial time $t_{\rm ini}$, namely t_k

---

[5] This has the simple reason that we will numerically compute the GWB spectra associated with these points later. The numerical evaluation taking the modified loop evolution into account is very unstable for point 3, but offers no additional insights compared to computing the spectra for point 3^*.

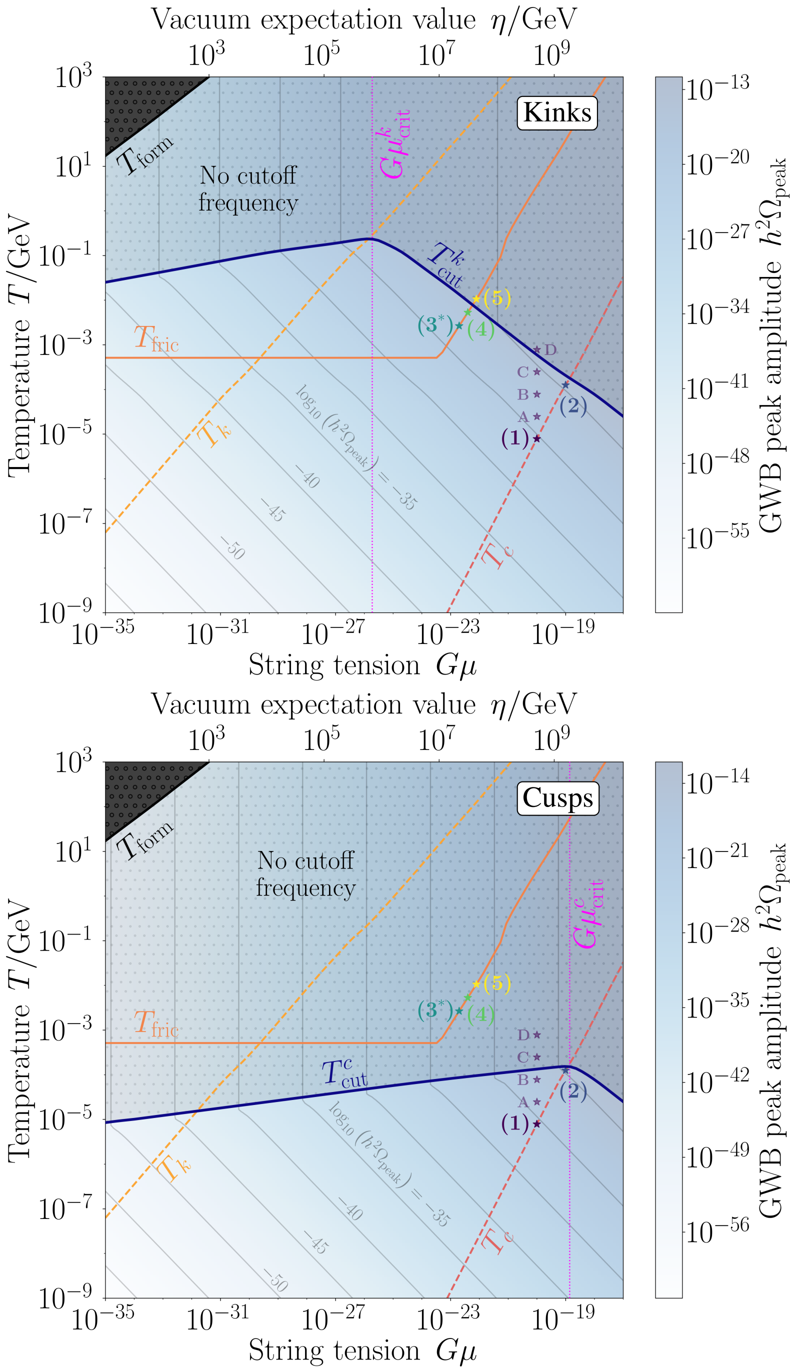


Figure 5.6: Plot of the two-dimensional parameter space spanned by string tension $G\mu$ and (initial) temperature T, i.e., the two parameters needed to describe the qualitatively different GWB spectra. Temperatures above T_{form} in Eq. (5.12) are excluded as in this region, strings have not formed yet, and we depict it as a black shaded region. Furthermore, we show T_{fric} in Eq. (5.15) as a solid orange line and $T_{k(c)}$ in Eqs. (5.22) and (5.26) as dashed dark-yellow and red-orange lines. In the upper panel, we assume particle decay due to kink–kink collisions, in the lower panel due to cusp formation with the respective temperatures $T^{k(c)}_{\text{cut}}$ as obtained from Eq. (5.94) as blue lines. The parameter space of low-scale strings is the entire region that lies below these lines. We show the region above $T^{k(c)}_{\text{cut}}$ with a grey shading to point at the absence of a cutoff frequency in the fundamental GWB spectrum. Moreover, we indicate the critical tensions $G\mu^{k(c)}_{\text{crit}}$ as magenta dotted lines above and below which $T^{k(c)}_{\text{cut}}$ approach the gravitational and particle decay dominated power-law behaviours given in Eqs. (5.11), (5.99), and (5.101). The benchmark points specified in Tab. 5.2 are depicted by stars. Finally, we visualise the GWB peak amplitude determined from Eq. (5.119) for each point in the parameter space in terms of a blue shading and corresponding contour lines in grey.
Figure reproduced from Ref. [5].

and t_c. This is based on the assumption that, for initial loop lengths $l_{\text{ini}} < l_{k(c)}$, loops quickly evaporate because of particle emission or gravitational decay is otherwise suppressed, resulting in no appreciable GW signal. Under this assumption, it is justified to replace $t_{\text{ini}} \to t_{\text{ini}}^{\text{eff}} = t_{k(c)}$ for any $t_{\text{ini}} < t_{k(c)}$ — simply because efficient particle emission prevents the emission of any relevant GW signal at all times $t_{\text{ini}} < t_{k(c)}$ — which renders t_k and t_c two valid choices for the initial time. For completeness, we show the corresponding temperatures $T_{k(c)}$ provided in Eqs. (5.22) and (5.26) as dashed dark-yellow lines for kink–kink collisions and red-orange lines for cusps in Fig. 5.6.

Here and afterwards, we shall take the influence of particle decay on the loop length directly into account (i.e., at the level of the function $\mathcal{J}$), such that, in the considered nonscaling models, T_k and T_c lose their meaning as possible choices for the initial temperature. Nonetheless, it is interesting to discuss how the results for the scaling model discussed previously are related to the formalism described in this section. In fact, we can conveniently summarise the assumptions going into the construction of the different models in terms of P_{GW}.

For low-scale strings with initial temperatures not too close to the cutoff temperature $T_{\text{ini}} \ll T_{\text{cut}}^{k(c)}$, string loops will only lose negligible amounts of their total energy into any form of radiation, be it particles or GWs. In this regime, ignoring loops produced with lengths $l_{\text{ini}} < l_{k(c)}$ is, hence, approximately equivalent to ignoring any loop with length $l < l_{k(c)}$ — loops produced after t_{ini} will not significantly shrink below their initial length l_{ini} until today. For this part of parameter space, the previous treatment in the scaling model, amounts to setting $P_{\text{GW}} = \Gamma G\mu^2\Theta(l - l_{k(c)})$. This can be seen as one of two extreme cases: Either there is no gravitational radiation below $l_{k(c)}$, or the emission of gravitation radiation is completely independent of the loop length, and we can summarise this with

$$P_{\text{GW}} = \Gamma G\mu^2\,\mathcal{K}(l)\,, \tag{5.102}$$

where the two extreme cases correspond to $\mathcal{K}(l) = \Theta(l - l_{k(c)})$ and $\mathcal{K}(l) = 1$. Most plausibly, realistic physical scenarios will correspond to solutions in between these two extreme cases. The scaling model used in the previous sections is based on the choice $\mathcal{K}(l) = \Theta(l - l_{k(c)})$, while here, we work with $\mathcal{K}(l) = 1$, or equivalently, $P_{\text{GW}} = P_{\text{NG}}$. At the same time, we supplement P_{GW} by P_k or P_c, which ultimately results in Eqs. (5.84) and (5.85).

As remarked before, due to the large model uncertainties, T_{ini} should be treated as a free phenomenological parameter that needs to be fixed in order to compute GWB spectra. In line with this sentiment, we specified our benchmark points to vary along different lines as choices for the initial temperature (T_c for points 1 and 2, and T_{fric} for points 3^*, 4, and 5) and, furthermore, picked different initial temperatures for a given value of $G\mu$ (points 1_A to 1_D) as can be seen in Fig. 5.6.

5.5 GWB spectrum from low-scale strings without scaling

Let us now turn to the actual computation of the GWB signal. If the effects of particle radiation are incorporated into the loop length evolution as discussed in Sec. 5.4.1, the loop number density takes the form

$$n(t,l) = \frac{\Theta(t-t_*)\Theta(t_*-t_i)\mathcal{F}\tilde{C}(t_*)\mathcal{J}(\alpha\xi(t_*)t_*)}{\mathcal{J}(l)\alpha\xi(t_*)\left(\alpha\xi(t_*) + \alpha\xi'(t_*)t_* + \Gamma G\mu\mathcal{J}(\alpha\xi(t_*)t_*)\right)}\left(\frac{1}{t_*}\right)^4\left(\frac{a(t_*)}{a(t)}\right)^3\,, \tag{5.103}$$

as derived in detail in Ref. [64]. This result represents a nonscaling solution for the loop number density, as evident from its explicit dependence on the length scale $l_{k(c)}$. At the same time, it relies on the standard VOS dynamics of the long-string network described in Sec. 4.2. In other words, for the nonscaling models considered here, the long strings in the string network display the same behaviour as long NG strings in the scaling regime. This means that the VOS equations for the long strings still apply, while the behaviour of the loop population in the string network differs from the usual NG picture. In particular, the functions $\tilde{C}(t)$ and $\xi(t)$ are as defined in Sec. 4.2. For the numerical computation of the GWB spectra, we obtain these functions, as before, by solving the VOS equations fully numerically, using numerical solutions to the Friedmann equations for the scale factor evolution, including changes in the effective number of radiative degrees of freedom.

To arrive at the expression in Eq. (5.103), we again assumed the single-length loop production function of Eq. (4.71) we also employed for the computation of the GWB spectrum in the scaling model in Sec. 5.2. This again assumes loop formation at time t_* with length $l_* = \alpha\xi(t_*)t_*$. Since we know the loop length evolution, given the length of a loop l at a time t, we can implicitly determine its production time $t_*(l,t)$ via the equation

$$\Gamma G\mu t_* + \zeta(\alpha\xi(t_*)t_*) = \Gamma G\mu t + \zeta(l)\,. \tag{5.104}$$

For the numerical computation of GWB spectra, we will solve this equation numerically for t_*. As initial values, we provide our solver with the analytical estimates for t_* found in Ref. [64]. With the numerical data for the loop number density and the scale factor evolution, the GWB is computed by inserting these results into Eq. (5.29) and then into Eq. (5.28).

In Fig. 5.7, we show the total GWB spectra for the nine benchmark points 1 to 5 and 1_A to 1_D from Tab. 5.2, including the effects of particle radiation due to kink–kink collisions in the left panel and due to cusps in the right panel. Correspondingly, we set $q = 5/3$ in the former case and $q = 4/3$ in the latter case.

Let us now compare the GW spectra belonging to benchmark points 1 and 1_A to 1_D in Fig. 5.7 to the spectra we obtained for the scaling model in Sec. 5.2. To obtain the latter, we had assumed that loops shorter than $l_{k(c)}$ will not contribute to the GWB at all, so that increasing T_{ini} starting from $T_{k(c)}$ while keeping $G\mu$ fixed left the spectrum almost unchanged in the case of particle radiation from kink–kink collisions (cusps). This behaviour is a consequence of the assumption that the power emitted into GWs is strongly suppressed for lengths $l < l_{k(c)}$ as discussed around Eq. (5.102). If we assume, instead, that $P_{\text{GW}} = P_{\text{NG}} = \Gamma G\mu^2$ for all loop lengths, we find a different outcome. Recall that for $T_{\text{ini}} \ll T_{\text{cut}}^{k(c)}$, loops typically do not significantly change their length during their evolution until today. This means that, for the description of the loop evolution in this regime, we can set $\mathrm{d}l/\mathrm{d}t \simeq 0$ or equivalently $\Gamma G\mu \simeq 0$. The time of production of a loop that has at time t a length l becomes

$$t_*(l) = \frac{l}{\alpha\xi(t_*)}\,, \tag{5.105}$$

and the loop number density in Eq. (5.103) simplifies tremendously:

$$n(t,l) = \frac{\mathcal{F}\tilde{C}(t_*)t_*^{-4}}{(\alpha\xi(t_*))^2}\left(\frac{a(t_*)}{a(t)}\right)^3\,. \tag{5.106}$$

For our analytical approximations, let us, as before, simplify the scale factor evolution by neglecting changes in g_ρ and by assuming an instantaneous transition from radiation to matter domination at $t_{\text{eq}} \simeq 1.59 \times 10^{12}$ s with the scale factor evolution during the radiation-dominated era given in Eq. (5.76) and during matter domination given in Eq. (5.33). For the same reasons as described in Sec. 5.2.1, we will focus here on the RM spectrum as the predominant contribution to the spectrum. In this case t_* lies in the radiation-dominated era and we set $\xi(t_*) = \xi_r \simeq 0.27$ and similarly $\tilde{C}(t_*) = \tilde{C}_r = 5.5$.

To obtain the number density of RM loops, we must compute the number density of RR loops first. One finds from Eq. (5.106)

$$n_{\text{rr}}(t,l) = \Theta_{\text{rr}}(t,l)\frac{\mathcal{F}\tilde{C}_r\,(\alpha\xi_r)^{1/2}}{l^{5/2}t^{3/2}}\,, \tag{5.107}$$

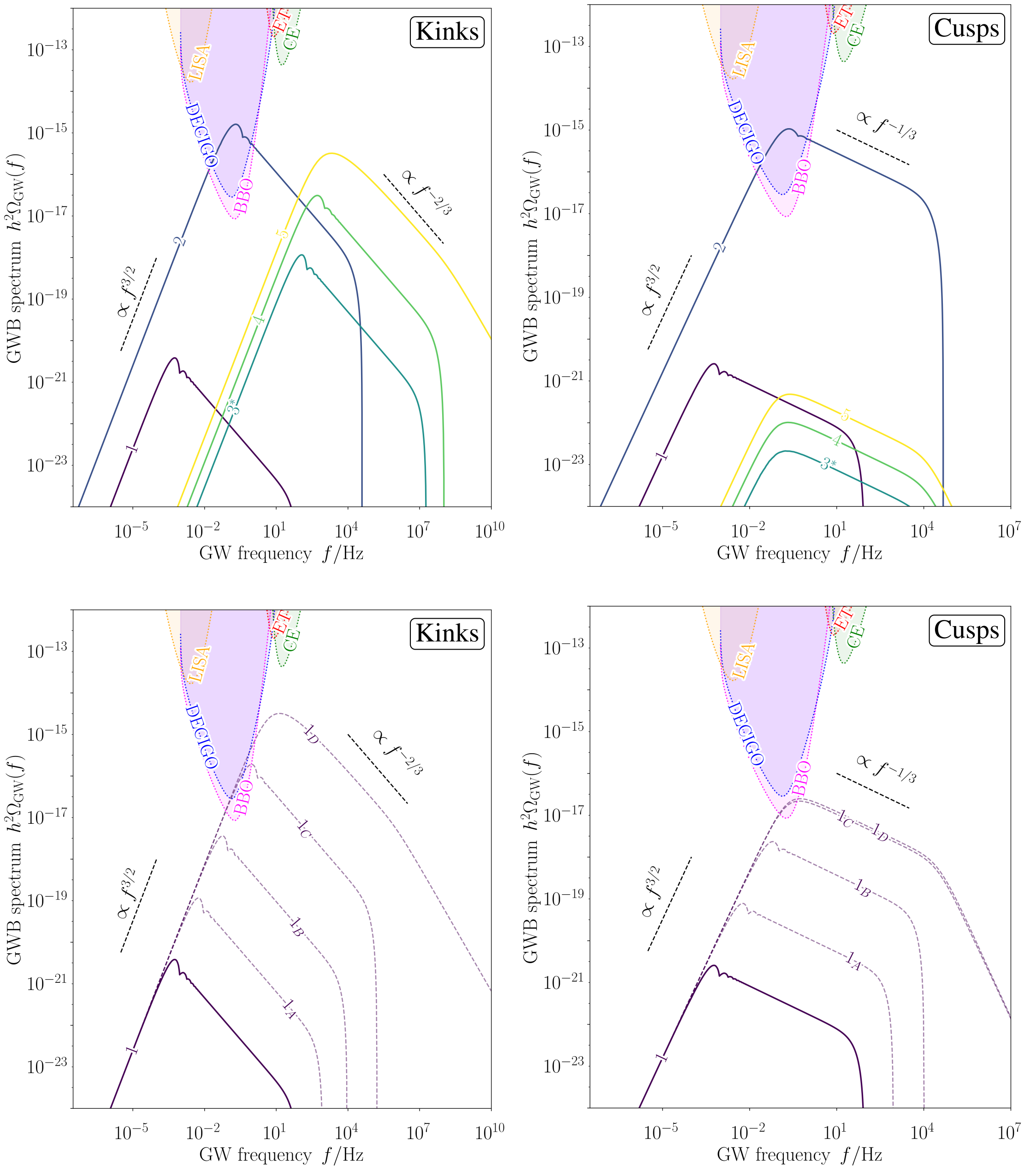


Figure 5.7: Numerical GWB spectra based on the VOS equations for the long-string network for benchmark scenarios 1 to 5 (upper panels) and 1, 1_A to 1_D (lower panels) as indicated in Fig. 5.6 and specified in Tab. 5.2. The left and right columns assume particle radiation from kink–kink collisions and cusps, respectively. The dotted lines and shaded regions show the power-law-integrated sensitivity curves of future experiments and are taken from Ref. [419].
Figure adapted from Ref. [5].

with $\Theta_{\rm rr}(t,l) = \Theta(t_{\rm eq} - t)\Theta(t - t_*)\Theta(t_* - t_i)$ a shorthand for the Heaviside functions introduced in Eq. (4.77a).

We can then find the RM loop number density via $n_{\rm rm}(t,l) = \Theta_{\rm rm}(t,l)\,(a_{\rm eq}/a(t))^3\, n_{\rm rr}(t_{\rm eq}, l)$ which, when explicitly evaluated, yields

$$n_{\rm rm}(t,l) = \Theta_{\rm rm}(t,l)\frac{8\sqrt{2}}{9}\frac{\Omega_r^{3/4}}{H_0^{1/2}\Omega_m}\frac{\mathcal{F}\tilde{C}_r\,(\alpha\xi_r)^{1/2}}{l^{5/2}t^2}\,,$$

where $\Theta_{\rm rm}(t,l) = \Theta(t - t_{\rm eq})\Theta(t_{\rm eq} - t_*)\Theta(t_* - t_{\rm ini})$, as introduced in Eq. (4.77b) and as used above.[6] We can use these loop number densities together with Eq. (5.29), to obtain an expression for the spectral shape function $\mathcal{I}_1(f)$. While we are eventually only interested in the RM spectrum, we also provide an expression for the RR spectrum since we have already computed the loop number density. For the RR case, we find

$$\mathcal{I}_{\rm rr,1}(f) = \Theta(f_{\rm rr}^{\rm max} - f)\Theta(f - f_{\rm rr}^{\rm min})\frac{2}{3}\mathcal{F}\tilde{C}_r\left(H_r^0\alpha\xi_r\right)^{1/2} f^{3/2}\left(\frac{a}{a_0}\right)^{3/2}\Bigg|_{a_{\rm rr}^{\rm min}(f)}^{a_{\rm eq}}. \tag{5.108}$$

The Heaviside functions present in the loop number density force the lower integration boundary to be frequency dependent:

$$\frac{a_{\rm rr}^{\rm min}}{a_0} = \max\left\{\frac{\alpha\xi_r t_{\rm ini} f}{2}, \frac{4H_r^0}{\alpha\xi_r f}\right\}, \qquad f_{\rm rr}^{\rm max} = \frac{2}{\alpha\xi_r t_{\rm ini}}\frac{a_{\rm eq}}{a_0}, \qquad f_{\rm rr}^{\rm min} = \frac{4H_r^0}{\alpha\xi_r}\frac{a_0}{a_{\rm eq}}. \tag{5.109}$$

The spectral shape function for the RM spectrum reads

$$\mathcal{I}_{\rm rm,1}(f) = \Theta(f_{\rm rm}^{\rm max} - f)\frac{H_0^{1/2}\Omega_r^{3/4}}{\Omega_m^{1/2}}\mathcal{F}\tilde{C}_r(\alpha\xi_r)^{1/2} f^{3/2}\left(\frac{a}{a_0}\right)\Bigg|_{a_{\rm rm}^{\rm min}(f)}^{a_{\rm rm}^{\rm max}(f)}, \tag{5.110}$$

and the Heaviside functions in the loop number density again imply frequency-dependent integration boundaries:

$$\frac{a_{\rm rm}^{\rm max}}{a_0} = \min\left\{1, \frac{\alpha\xi_r t_{\rm eq} f}{2}\right\}, \quad \frac{a_{\rm rm}^{\rm min}}{a_0} = \max\left\{\frac{a_{\rm eq}}{a_0}, \frac{\alpha\xi_r t_{\rm ini} f}{2}\right\}, \quad f_{\rm rm}^{\rm max} = \frac{2}{\alpha\xi_r t_{\rm ini}}. \tag{5.111}$$

Above the frequency $f_* = 2/(\alpha\xi_r t_{\rm eq}) = f_{\rm rm}^{\rm max} t_{\rm ini}/t_{\rm eq} \ll f_{\rm rm}^{\rm max}$, the maximum scale factor takes the value $a_{\rm rm}^{\rm max}/a_0 = 1$. In our analytical estimates, we are not interested in values of the spectrum at very low frequencies but rather frequencies close to the peak around $f_{\rm rm}^{\rm max}$, which is why we will restrict ourselves to $a_{\rm rm}^{\rm max}/a_0 = 1$ in the following.

Using this expression for the shape function $\mathcal{I}_{\rm rm,1}$, the GWB spectrum from the fundamental oscillation mode reads

$$\Omega_{\rm GW}^{(1)}(f) = \Theta(f_{\rm cut} - f)A_{\rm rm} f^{3/2}\left(1 - \frac{a_{\rm rm}^{\rm min}(f)}{a_0}\right), \tag{5.112}$$

with amplitude

$$A_{\rm rm} = \frac{8\pi}{3}\frac{\Omega_r^{3/4}}{H_0^{3/2}\Omega_m^{1/2}}\mathcal{F}\tilde{C}_r\left(\alpha\xi_r\right)^{1/2}\Gamma(G\mu)^2\,, \tag{5.113}$$

as well as cutoff frequency and minimum scale factor

$$f_{\rm cut} = \frac{2}{\alpha\xi_r t_{\rm ini}} \qquad \text{and} \qquad \frac{a_{\rm rm}^{\rm min}(f)}{a_0} = \max\left\{\frac{a_{\rm eq}}{a_0}, \frac{f}{f_{\rm cut}}\right\}. \tag{5.114}$$

---

[6] For $n_{\rm rm}$, see Eq. (4.81) and note that under the assumption of non-decaying loops, one finds $l_{\rm eq} = l$.

As one approaches the maximum frequency from below, the second term in $a_{\rm rm}^{\rm min}$ will become dominant at some point. Correspondingly, the spectrum will have the spectral shape $\Omega_{\rm GW}^{(1)}(f) \propto f^{3/2}\,(1-f/f_{\rm cut})$, which allows us to determine the peak frequency as $f_{\rm peak} = 3/5 f_{\rm cut}$.[7] In this approximation, the peak amplitude of the fundamental spectrum is thus given by

$$h^2\Omega_{\rm peak}^{(1)} \simeq \frac{2}{5} h^2 A_{\rm rm} f_{\rm peak}^{3/2} \simeq 1.51 \times 10^{-23} \left(\frac{t_{\rm ini}}{1\,{\rm s}}\right)^{-3/2} \left(\frac{G\mu}{10^{-25}}\right)^2 . \tag{5.115}$$

Generally, the above result will slightly overestimate the peak amplitude of the fundamental spectrum. The reason is that, when approaching $T_{\rm ini} \to T_{\rm cut}^{k(c)}$ from the parameter region of low-scale strings, i.e., from below, the effect of the decay processes will slowly start to suppress the amplitude of the spectrum. In this limit and when transitioning to the high-scale-string regime, this result must clearly break down as it would otherwise predict arbitrarily large GWB amplitudes as $t_{\rm ini} \to 0$. To understand the behaviour in the absence of a cutoff frequency, we can artificially split our integral into contributions for which the integrand is evaluated for $l > l_{\rm cut}$ and for which $l < l_{\rm cut}$, denoted $\Omega_{>}^{(1)}$ and $\Omega_{<}^{(1)}$ respectively. We then have

$$\Omega_{\rm GW}^{(1)}(f) = \Omega_{>}^{(1)}(f) + \Omega_{<}^{(1)}(f) . \tag{5.116}$$

The first term is, by definition, only evaluated in the regime in which our previous treatment approximately applies. Since the dependence on l in the integrand is evaluated via $l = 2a(t)/(a_0 f)$, the conditions on the allowed lengths translate into conditions on the allowed frequency ranges for a given emission time t. In particular, the condition $l < l_{\rm cut}$ implies $f < 2a(t)/(a_0 l_{\rm cut})$. Since $\Omega_{>}^{(1)}$ has the same spectral shape as in Eq. (5.112), it will peak at $f_{\rm peak}^{>} = 3/5 f_{\rm cut}^{>}$, where $f_{\rm cut}^{>}$ is the highest possible frequency contributing to $\Omega_{>}^{(1)}$. This frequency is obtained from the previous time-dependent upper bound on the frequencies when evaluating it at the largest scale factor and thus the latest time. Since the integral runs until t_0, the largest frequency that contributes overall is $f_{\rm cut}^{>} = 2/l_{\rm cut}$. For the second part of the spectrum $\Omega_{<}^{(1)}$, we expect the effects of particle decay to become important and therefore to suppress the spectrum. Note that $\Omega_{<}^{(1)}$ has no cutoff frequency for $T_{\rm ini} > T_{\rm cut}$ and extends to arbitrarily large frequencies. While it is, therefore, important for the high-frequency spectrum, we do not expect a big effect on the peak position or amplitude. The peak amplitude of the fundamental spectrum is thus

$$\Omega_{\rm peak}^{(1)}(f) \simeq \frac{2}{5} A_{\rm rm} \tilde{f}_{\rm peak}^{3/2} . \tag{5.117}$$

By introducing

$$\tilde{f}_{\rm peak} = \frac{3}{5}\,\min\left\{f_{\rm cut}^{>}, f_{\rm cut}\right\} , \tag{5.118}$$

we include both the case of low-scale strings as well as the case of high-scale strings, for which $f_{\rm cut}$ goes to infinity. To account for the fact that particle and gravitational radiation may have a non-negligible effect on the position of $f_{\rm cut}$, we can sensibly use the exact expressions obtained from Eqs. (5.2) and (5.90) in Eq. (5.118). With this treatment, we indeed find the peak position and amplitude to be in good agreement with our numerical results.

We can, of course, not measure the GWB spectrum from separate oscillation modes, but only the combined spectrum from all harmonics, as in Eq. (5.28). As shown in Sec. 5.2.1, the peak

---

[7] Note that this peak amplitude is an estimate of the peak amplitude of low-scale strings and cannot be applied to high-scale strings. It also becomes a bad approximation if $T_{\rm ini} \ll T_{\rm cut}$ is not satisfied and loop decay cannot be completely neglected. It should, in particular, not be confused with the would-be peak frequency defined in Eq. (5.38) that is an estimate for the peak frequency of a high-scale-string spectrum.

amplitude of the total spectrum follows from the peak amplitude of the fundamental spectrum by multiplying with a constant prefactor

$$\Omega_{\mathrm{peak}} \simeq \frac{2}{5} \frac{H_{n_{\max}}^{(q+3/2)}}{H_{n_{\max}}^{(q)}} A_{\mathrm{rm}} \tilde{f}_{\mathrm{peak}}^{3/2} \,. \tag{5.119}$$

For $n_{\max} = 10^5$ and $q = 5/3, 4/3$, we have $H_{n_{\max}}^{(q+3/2)}/H_{n_{\max}}^{(q)} \simeq 0.55,\ 0.35$. We show the resulting amplitude as a blue-shading with corresponding grey contour lines in our parameter space in Fig. 5.6. The result is what one would physically expect. For $T_{\mathrm{ini}} < T_{\mathrm{cut}}^{k(c)}$, the contour lines behave in the same way as they did for NG strings, because string decay, and therefore the form of string decay, change the peak amplitude only imperceptibly. Once $T_{\mathrm{ini}} > T_{\mathrm{cut}}^{k(c)}$, loop decay becomes very important, and strings produced at even earlier times will immediately decay and only contribute to the very high frequency tail of the spectrum, leaving the peak amplitude unchanged. The contour lines become, therefore, independent of T_{ini} and vary only when changing the value of $G\mu$. This behaviour can also be observed when considering the numerical spectra corresponding to benchmark points 1_C and 1_D in the case of cusps in Fig. 5.7.

As shown in Sec. 5.2.1, the $f^{3/2}$ power-law behaviour of the fundamental mode of low-scale strings also implies an $f^{3/2}$ power law for the total spectrum below f_{cut}. Above f_{cut} but below $n_{\max} f_{\mathrm{cut}}$, the total spectrum will develop an f^{q-1} power law. Both behaviours can be seen in all spectra shown in Fig. 5.7. The behaviour of the spectrum at frequencies $f > n_{\max} f_{\mathrm{cut}}$ is unphysical since it depends on our arbitrarily chosen value of $n_{\max}$.

In Fig. 5.7, we can also see the oscillatory feature characteristic of GWB spectra from low-scale strings, as discussed in detail in Sec. 5.2. While the parameter space in which low-scale strings arise is altered when including the effects of particle radiation in the form of nonscaling loop number densities into our computations of the GWB spectra, this oscillatory feature remains unchanged as long as one considers low-scale strings. When transitioning to high-scale strings, we can see that this feature disappears since modes do not drop to zero and, therefore, still contribute to the total spectrum at high frequencies.

5.6 Discussion

In this chapter, we investigated the GWB spectrum sourced by local cosmic strings emerging at symmetry-breaking scales in the range $\eta \sim 10^2$ GeV to $\eta \sim 10^9$ GeV. We refer to such strings as "low-scale strings" due to their small tension, spanning values from $G\mu \sim 10^{-33}$ to $G\mu \sim 10^{-19}$. Because of their low tension, these strings radiate GWs with a very low power. As a result, string loops formed in the early Universe with sufficiently large initial sizes have not had enough time to completely decay until today and are still present today.

The loops of shortest length that are present today are the remnants of loops formed at t_{ini}, the time at which production of loops contributing to the GWB signal becomes efficient for the first time, because it is no longer suppressed by thermal friction and GW emission from loops overtakes particle emission as the dominant energy loss mechanism. Defined in this way, t_{ini} marks the onset of the regime where the standard computation of the GWB signal from string loops becomes applicable. We have explored several well-motivated scenarios for setting t_{ini} and discussed them in detail. Notably, we find that t_{ini} can be significantly delayed for small string tensions. This leads to a large hierarchy between the temperature T_{form} at which the string network forms and the temperature at which the conventional picture of a scaling network emitting gravitational radiation becomes valid (T_{fric} or T_{cusp}), as illustrated in Fig. 5.2.

Low-scale strings arise from a combination of small string tensions $G\mu$ and a large initial loop length, i.e., large values of t_{ini}. This prevents string loops from fully evaporating and, therefore, their length distribution has an overall minimum length l_{min} and a corresponding sharp cutoff frequency $f_{\mathrm{cut}} = 2/l_{\mathrm{min}}$ in the GWB spectrum emitted from the fundamental oscillation mode. This cutoff induces a characteristic pattern of peaks and dips, where the dips are located at integer multiples of f_{cut} in the total GWB spectrum, which is the superposition of all harmonic mode contributions. The resulting oscillatory spectral feature is very distinct and, potentially, observable by upcoming GW experiments. In particular, around $G\mu \sim 10^{-19}$ and $T_{\mathrm{ini}} \sim 100\,\mathrm{keV}$ exists a favourable region of parameter space that falls within the projected sensitivity reach of BBO and DECIGO.

We computed the GWB spectrum generated by low-scale cosmic strings using both the VOS and BOS models. Apart from our numerical computations, we also provide analytical estimates in Eqs. (5.45) and (5.49) that allow us to describe the spectral shape of the GWB. One result of our study has relevance beyond low-scale strings: We found a generalisation of the standard loop number densities, n_{rr} and n_{rm}, given in Eqs. (5.73) and (5.75) obtained by incorporating the finite-width distribution of initial loop lengths observed in simulations by BOS. Our generalised expressions for n_{rr} and n_{rm} provide a useful foundation for future studies of cosmic strings that aim to incorporate the BOS loop number densities in a more realistic manner. Applied to low-scale strings, the improved loop number densities revealed that the presence of a non-zero width in the initial loop length distribution tends to smooth out the characteristic oscillatory feature in the total GWB spectrum (see Fig. 5.5). Nonetheless, when the distribution remains sufficiently peaked, these spectral features can still partially persist and continue to offer promising observational signatures.

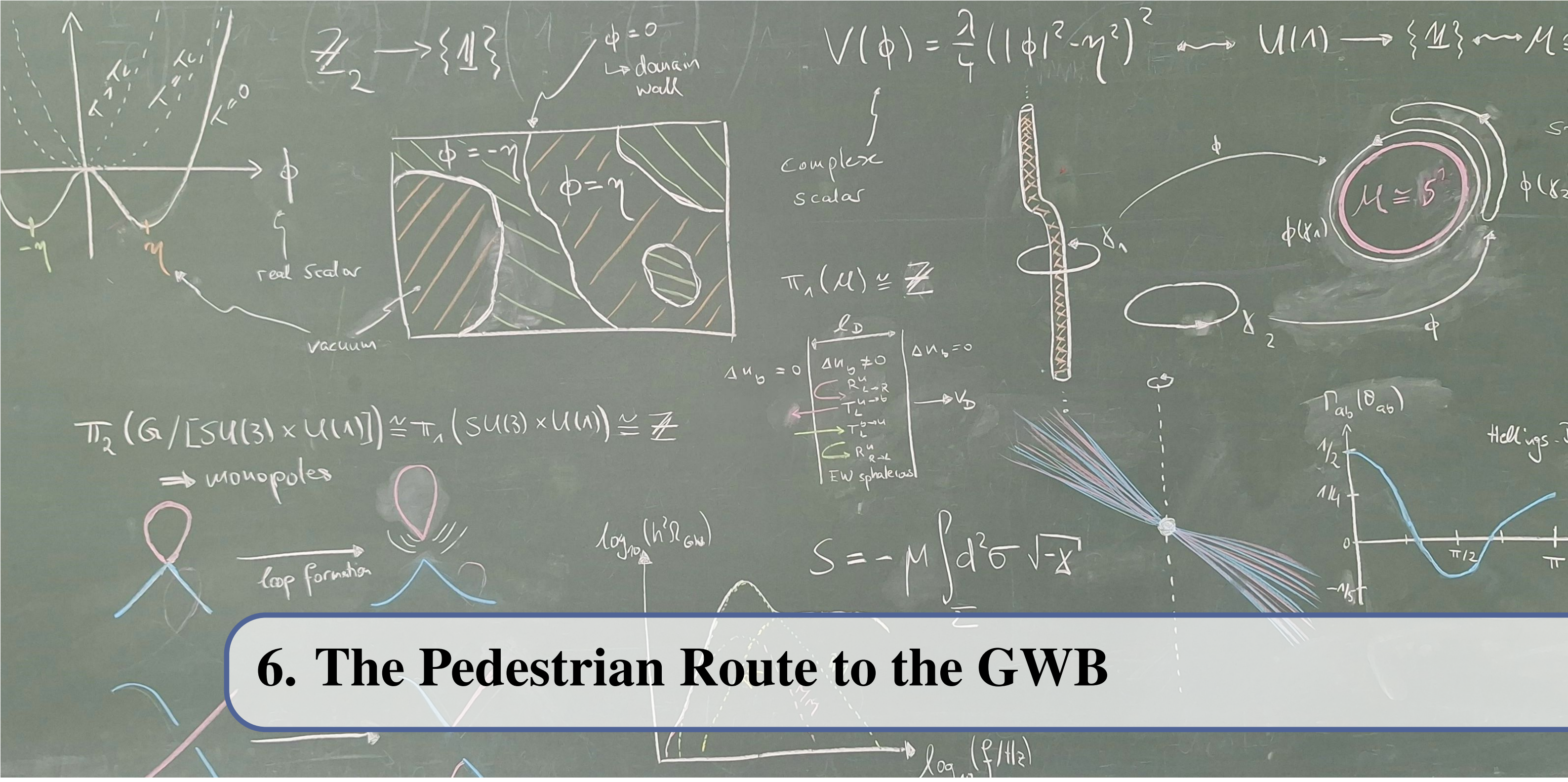
The computations discussed so far were all based on scaling models, in which loops shrink only due to the emission of gravitational radiation whose power is scale-independent. However, the scaling behaviour of the loop population is violated in models where particle emission dominates over GW emission once a loop has shrunk below certain length. Correspondingly, we also studied the impact of modelling the string loop decay due to particle emission from kink–kink collisions and cusps on the GWB from low-scale strings in nonscaling scenarios.

We found that, for a large part of parameter space, the spectra are quite independent of the decay mode. This is simply due to the fact that low-scale strings barely decay until today. Loop decay and, hence, the kind of loop decay, will only slightly affect the loop number densities and, correspondingly, the GWB spectrum. Indeed, we demonstrated that the inclusion of particle decay neither alters the power-law behaviours describing low-scale-string GWB spectra, nor does it change the characteristic oscillatory features of the spectra. This, of course, holds true only as long as we are actually considering low-scale strings. In fact, the parameter space from which low-scale strings are obtained can change quite strongly when one allows for the particle decay of string loops. In Eq. (5.94), we obtained criteria that allow us to quantify for which regions of parameter space a cutoff frequency in the fundamental spectrum arises. We found that for the particle decay via kink–kink collisions as well as via cusp formation, there exist critical string tensions $G\mu_{\mathrm{crit}}^{k(c)}$. Above the respective tension, there is virtually no difference whether particle decay is taken into account or not: The parameter space and spectra are just the same as if one allowed for gravitational evaporation only. Below the critical string tension, particle decay can become important, and we find that it moderately reduces the parameter space of low-scale strings in the case of kink–kink collisions and rather drastically in the case of cusp formation.

Generally, additional decay channels will always lead to a reduction of the low-scale-string parameter space as they decrease the decay time. In some parts of parameter space, string loops that would otherwise not have fully evaporated until today will completely decay because of the extra decay mode. These parts of parameter space will then no longer correspond to low-scale strings. As described, in the bulk of low-scale-string parameter space, the overall shortest loop length scale l_{min} will not depend notably on loop decay and roughly be set by the initial length. When approaching the parameter space boundary towards the high-scale string regime, loop evaporation and the dominant kind of decay channel become increasingly important to determine $l_{\rm min}$ and thus $f_{\rm cut}$. In Eq. (5.90), we provide explicit expressions for $l_{\rm min}$, covering the different possible decays. Particle decay cannot only have a dramatic effect on the peak position and amplitude of low-scale-string spectra close to the transition to high-scale strings, but also on high-scale-string spectra themselves. In Eqs. (5.118) and (5.119), we analytically estimated the peak frequency and amplitude of the GWB spectrum that apply to both low and high-scale strings, and they agree well with our numerical results.

Many open questions remain that need to be considered in the future. In the current treatment, we use the temperature $T_{\rm ini}$ (or, equivalently, the time $t_{\rm ini}$) when loop production becomes significant as a free parameter. This is reasonable as long as one does not fix a specific particle physics model and views $T_{\rm ini}$ as a phenomenological parameter. Nevertheless, $T_{\rm ini}$ is, at least in principle, fixed for a given microscopic model. It is an important task to figure out whether all values of $(G\mu, T_{\rm ini})$ that lie in the low-scale-string region of parameter space can actually be realistically obtained from particle physics models. For a better modelling of low-scale strings, it will also be of great importance to improve upon the understanding of the total power P emitted by string loops and of the distribution of this power between particle $P_{k(c)}$ and gravitational radiation $P_{\rm GW}$. In this chapter, we discussed two extreme cases summarised in Eq. (5.102): In the first case, we worked with a scaling model and assumed that the emission of gravitational radiation ceases completely once loops have shrunken below a length at which particle emission becomes dominant. In the second case, we employed the nonscaling model of Ref. [64] and described the total radiated power as an interpolation between the asymptotic expressions found for particle decay and gravitational decay. This also means that we assumed gravitational decay to be completely independent of the loop length. However, for very small loops, gravitational decay might be substantially different and is most likely not captured by either of the two extreme cases but rather by an intermediate scenario.

In spite of these open questions, we demonstrated the robustness of the central result of this chapter — a GWB with a characteristic oscillatory feature produced from low-scale cosmic strings. The feature remains if one accounts for the finite widths in the distribution of lengths at which loops are produced and can be found in both scaling and nonscaling models. We therefore conclude that strings with relatively small tensions should receive more attention than is currently the case in the literature. In particular strings that are formed in phase transitions at energy scales around $\eta \sim 10^9$ GeV could lead to the GWB spectra characteristic of low-scale strings while at the same time being potentially observable. Cosmic string formation at these scales is not uncommon in BSM scenarios. In future studies, it is worth exploring which particle physics models can give rise to the distinct GWB spectra we discussed in this chapter.



6. The Pedestrian Route to the GWB

This chapter is based on Ref. [4]. The detailed contributions of the author to this article are outlined at the beginning of this thesis.

The computation of the GWB from cosmic strings serves ultimately the purpose of confronting theoretical models with observational evidence. To draw any conclusions from the observation of a GWB about the underlying physics, a proper analytical understanding of all features of the GWB from cosmic strings for different parts of parameter space is vital. In this chapter, we provide detailed derivations and analytical expressions of such features as well as analytical approximations of the fundamental and the total GWB spectrum. Previous work in this regard was mainly conducted by SOUSA, AVELINO and GUEDES in Ref. [403] with a focus on the fundamental-mode contribution of the GWB and by BLANCO-PILLADO et al. in Ref. [411] with a focus on the total GWB spectrum. The expressions we derive in the following extend over a wider range of parameters and cover qualitatively different spectra that cannot be described by means of the expressions derived in earlier works [403, 411, 415]. In addition to improving the understanding of all the parameter dependencies from the analytical expressions in detail, the typically long computation times to obtain the GWB signal numerically are drastically reduced. In this sense, this chapter can be regarded as the "pedestrian" route to the GW signal from cosmic strings. While the accuracy of our results will not always match the accuracy of exact numerical computations, we achieve a purely analytical derivation that does not require the numerical evaluation of complicated integral and sum expressions.

Our results are thus suited for fast parameter scans at the cost of a minimally reduced accuracy. For other sources of GWs from the early Universe, GWB templates of this form have already been available for a long time. In simple models, the GW signal from cosmic inflation can, e.g., be modelled in terms of a constant or running power law [422]. Similarly, the GW signal from a cosmological phase transition can in many models be approximated by a broken or doubly broken power law [423]. For cosmic strings, simple analytical templates were, however, unavailable until now, at least no templates valid at all frequencies and capable of covering all possible types of GW spectra that can be realised in cosmic-string models. An important motivation behind this chapter is to close this gap and work out such analytical templates that no longer require further numerical steps.

Recall from the introduction to this thesis that PTA experiments recently found strong evidence for the existence of a GWB [102–105]. As we will see in Ch. 7, the spectral shape of the measurements disfavours stable strings as a sole explanation. Therefore, the current PTA measurements give, most likely, only an upper bound on the string tension. Meanwhile, cosmic superstrings, which are fundamental strings or one-dimensional D-branes of cosmologically relevant size that arise at the end of brane inflation [424, 425], yield spectra that give an excellent fit to the NANOGrav 15-year data [1, 426]. Current modelling predicts that the GW signal from colour-flux tubes can simply be obtained from the one of field strings by rescaling the amplitude of the GW spectrum with a constant prefactor [356, 357]. The same is, at least in the range of nanohertz frequencies, true for cosmic superstrings, although more improved modelling shows this to be an oversimplification for a wider range of frequencies [427, 428]. The analytical results presented in this chapter are thus, in specific frequency ranges, directly applicable to GWs from colour-flux tubes and cosmic superstrings.

This chapter is structured as follows: Assuming the evolution of the long-string network to be described by the VOS model presented in Sec. 4.2, we will rigorously derive analytical expressions for the GW spectrum obtained from the fundamental excitation of string loops in Sec. 6.1. We will discuss all qualitatively different spectra obtained in different regions of parameter space and analytically investigate their features in detail. Additionally, we will show how the effect of variations in the effective number of relativistic degrees of freedom (DOFs) in the early Universe can be included in the spectrum. Equipped with these results, we will continue in Sec. 6.2 by providing an accurate analytic estimate of the total spectrum. We will briefly disuss our results in Sec. 6.3. A handy summary of all expressions necessary to calculate the total spectrum is provided in Sec. 6.4.

Numerical spectra: In order to assess the quality of our analytical results later on, we need a comparison. Therefore, in all of our plots, we will show numerical spectra in addition to the analytical spectra. Here, we summarise the assumptions we make for our numerical computations. To obtain the fundamental spectrum numerically, we evaluate Eq. (4.91) with the loop number density in Eq. (4.72) and Heaviside functions adapted to the time interval (RR, RM, or MM) under consideration. The time dependence of the scale factor is obtained by solving the Friedmann equation numerically, using a model, in which

$$H(a) = H_0\sqrt{\Omega_\Lambda + \Omega_m\,(a/a_0)^{-3} + \rho_r(a)/\rho_{\mathrm{crit}}}\,, \tag{6.1}$$

with Hubble constant $H_0 = 67.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and density parameters $\Omega_\Lambda = 0.685$ and $\Omega_m = 0.315$ [14]. For the comparison to analytical results in which we take changes in the effective number of relativistic DOFs into account, we use in our numerical computation the time evolution of $\rho_r(a)$ based on tabulated data for temperature T, energetic DOFs g_ρ and entropic DOFs g_s from Ref. [380]. If we assume instead a constant number of effective DOFs, we only set $\rho_r^0 = \pi^2/30 g_\rho(T_0)T_0^4$ for a temperature of $T_0 = 2.73\,\mathrm{K}$ today and use the scale factor dependence $\rho_r(a) \propto a^{-4}$. The necessary input from the VOS model is obtained by numerically solving the VOS equations (4.50) and (4.52) based on the evolution of the scale factor obtained from the previously described calculation. Having thus found the fundamental-mode contribution to the spectrum, we obtain the total spectrum by explicitly carrying out the sum from mode 1 to $n_{\max}$ using the relation in Eq. (4.90).

6.1 Fundamental-mode contributions

In this section, we shall derive analytical expressions for the fundamental-mode contribution to the GW spectra from decaying string loops (see Eq. (4.91)). Similar analytical expressions have been found in Ref. [403], but were derived less rigorously leading to a smaller range of applicability. We will discuss where our GWB spectra lead to improvements. For our spectra, we will also study all qualitatively different parts of parameter space and, moreover, provide analytical expressions for all relevant features of these spectra. Following the splitting in Eq. (4.76), we decompose the GW spectrum according to the production time of loops and the emission time of GWs,

$$h^2\Omega_{\mathrm{GW}} = h^2\Omega_{\mathrm{rr}} + h^2\Omega_{\mathrm{rm}} + h^2\Omega_{\mathrm{mm}}\,. \tag{6.2}$$

6.1.1 RR loops

Let us begin with the GWs produced during radiation domination stemming from the fundamental oscillation mode of string loops. Using the loop number density in Eq. (4.77a) together with Eq. (4.79) in the general expression for the spectrum given in Eq. (4.91), we obtain

$$\begin{aligned} h^2\Omega_{\mathrm{rr}}^{(1)} = h^2\frac{16\pi}{3}\left(\frac{G\mu}{H_0}\right)^2\frac{\Gamma}{f}C_r\Theta\left(a_{\mathrm{end}}-a_{\mathrm{start}}\right)\times \\ \times\int_{a_{\mathrm{start}}/a_0}^{a_{\mathrm{end}}/a_0}\frac{\left(\frac{a}{a_0}\right)^4\mathrm{d}\left(\frac{a}{a_0}\right)}{t_r^{3/2}(a)\left(\frac{2}{f}\frac{a}{a_0}+\Gamma G\mu t(a)\right)^{5/2}H_r(a)}\,. \end{aligned} \tag{6.3}$$

The integration boundaries $a_{\mathrm{start/end}}$ are set by the Heaviside functions occurring in the loop number density, and we will discuss their explicit values in a moment. Since we only integrate over times, or equivalently scale factors, during radiation domination, we were able to simplify the expressions for the Hubble parameter and the cosmic time as a function of a occurring in the integrand,

$$H_r\left(a\right) = H_r^0\left(\frac{a_0}{a}\right)^2 \qquad \text{and} \qquad t_r\left(a\right) = \frac{1}{2H_r^0}\left(\frac{a}{a_0}\right)^2, \tag{6.4}$$

where we defined $H_r^0 = H_0\Omega_r^{1/2}$. For the moment, we keep the effective number of relativistic DOFs constant and consider the impact of changes in it later.

For the further discussion, it will be beneficial not to work directly with the scale factor but with the variable

$$x_r = \frac{f}{f_r}\,, \qquad \text{with} \qquad f_r = \frac{a_0}{a}\,h_r^0\,, \tag{6.5}$$

instead. Note that the reference frequency f_r carries the dependence on the scale factor. Furthermore, we introduced

$$h_r^0 = \frac{4H_r^0}{\Gamma G\mu} \simeq 1.68\times10^{-11}\,\mathrm{Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right). \tag{6.6}$$

Using the new variable x_r, the integral in Eq. (6.3) takes the form

$$h^2\Omega_{\mathrm{rr}}^{(1)} = \frac{64\pi}{3}C_r h^2\Omega_r\left(\frac{G\mu}{\Gamma}\right)^{1/2}\Theta\left(x_{\mathrm{rr}}^{\mathrm{end}}-x_{\mathrm{rr}}^{\mathrm{start}}\right)\int_{x_{\mathrm{rr}}^{\mathrm{start}}}^{x_{\mathrm{rr}}^{\mathrm{end}}}\frac{x_r^{1/2}}{\left(1+x_r\right)^{5/2}}\,\mathrm{d}x_r \tag{6.7}$$

and evaluates to the simple expression

$$h^2\Omega_{\mathrm{rr}}^{(1)} = \mathcal{A}_{\mathrm{rr}}\,\Theta(x_{\mathrm{rr}}^{\mathrm{end}}-x_{\mathrm{rr}}^{\mathrm{start}})\left[S_{\mathrm{rr}}^{(1)}(x_{\mathrm{rr}}^{\mathrm{end}})-S_{\mathrm{rr}}^{(1)}(x_{\mathrm{rr}}^{\mathrm{start}})\right], \tag{6.8}$$

with amplitude

$$\begin{aligned}\mathcal{A}_{\rm rr} &= \frac{128\pi}{9}\, C_r\, h^2\Omega_r \left(\frac{G\mu}{\Gamma}\right)^{1/2} \\ &\simeq 4.52\times 10^{-10} \left(\frac{C_r}{0.171}\right)\left(\frac{50}{\Gamma}\right)^{1/2}\left(\frac{G\mu}{10^{-10}}\right)^{1/2} ,\end{aligned} \tag{6.9}$$

and spectral shape function

$$S_{\rm rr}^{(1)}\left(x\right) = \left(\frac{x}{1+x}\right)^{3/2} . \tag{6.10}$$

In $h^2\Omega_{\rm rr}^{(1)}$, we are interested in two values of x_r given by the integration boundaries $x_{\rm rr}^{\rm start}$ and $x_{\rm rr}^{\rm end}$, which are in turn set by the Heaviside functions in the loop number density given in Eq. (4.77a). First, the condition $t_{\rm eq} > t$ gives rise to an upper integration boundary

$$x_{\rm rr}^{\rm end} = x_r^{\rm eq} = \frac{f}{f_r^{\rm eq}}\,, \qquad \text{with} \qquad f_r^{\rm eq} = \frac{a_0}{a_{\rm eq}}\, h_r^0 \simeq 5.74\times 10^{-8}\,\text{Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right) , \tag{6.11}$$

where $a_{\rm eq}/a_0 \simeq 2.92\times 10^{-4}$ is the scale factor at matter–radiation equality. Meanwhile, the two conditions $t > t_* > t_{\rm ini}$ set competing lower integration boundaries

$$x_{\rm rr}^{\rm start} = \max\left\{\chi_r^{-1}, \varphi_2\left(x_r^{\rm ini}, \chi_r\right)\right\} . \tag{6.12}$$

Here, we introduced

$$\chi_r = \frac{\alpha\xi_r}{\Gamma G\mu}\,, \qquad x_r^{\rm ini} = \frac{f}{f_r^{\rm ini}}\,, \qquad f_r^{\rm ini} = \frac{a_0}{a_{\rm ini}}\, h_r^0\,, \tag{6.13}$$

and φ_2 is the positive solution of the quadratic equation

$$\varphi^2 + \varphi - \left(1+\chi_r\right)\left(x_r^{\rm ini}\right)^2 = 0\,. \tag{6.14}$$

Explicitly, we have

$$\varphi_2\left(x,\chi\right) = \left[\frac{1}{4} + \left(1+\chi\right)x^2\right]^{1/2} - \frac{1}{2}\,. \tag{6.15}$$

For our analytical results, we will keep $t_{\rm ini}$ as a free parameter due to large uncertainties in its choice as discussed extensively in the previous chapter in Sec. 5.1.2. Nevertheless, we can safely assume it to be a time deep in radiation domination such that we can write to a good approximation $a_{\rm ini}/a_0 = \left(2H_r^0 t_{\rm ini}\right)^{1/2}$ and find

$$f_r^{\rm ini} = \frac{2\sqrt{2}}{\Gamma G\mu}\left(\frac{H_r^0}{t_{\rm ini}}\right)^{1/2} \simeq 8.19\times 10^{10}\,\text{Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right)\left(\frac{10^{-24}\,\text{s}}{t_{\rm ini}}\right)^{1/2} . \tag{6.16}$$

In most of our plots, we will assume that $t_{\rm ini} = t_{\rm fric} \propto t_{\rm Pl}/(G\mu)^2$ and we will set proportionality factors to be 1 for the purpose of plotting (see Eq. (4.96)). With this, we find

$$f_r^{\rm fric} = \frac{a_0}{a_{\rm fric}}\, h_r^0 \simeq 3.53\times 10^{10}\,\text{Hz}\left(\frac{50}{\Gamma}\right) . \tag{6.17}$$

The GW spectra obtained with the above formulae are shown in Figs. 6.1 and 6.3 for different parts of parameter space. The different characteristic frequencies and features of the spectra are discussed below. Before turning to this discussion, we would like to point out that our analytical expressions lead to an improvement over those found in Ref. [403] as can be seen in the plots when comparing the analytical results to our fully numerical calculation described at the end of the introduction to this chapter. We quantify the deviation of the analytical from the numerical spectrum in our plots in terms of δ_Ω, defined as

$$\delta_\Omega(f) = \frac{\Omega_{\rm ana}(f) - \Omega_{\rm num}(f)}{\Omega_{\rm num}(f)} \,. \tag{6.18}$$

In particular, our spectra capture the transition from the plateau to the high-frequency part of the spectrum better, as can be seen in the right panel of Fig. 6.2. In the left panel, one can also observe that our analytical expression for the spectrum goes to zero at a frequency $f_{\rm rr}^{\rm min}$ as does the numerically calculated spectrum. The result found in Ref. [403] extends, in principle, to arbitrarily low frequencies and needs to be cut off by hand. In Ref. [403], this cutoff was implicitly given and can be shown to evaluate to the same value of $f_{\rm rr}^{\rm min}$ as derived here in the following. Nonetheless, this still leads to large deviations between the numerical spectra and those computed in Ref. [403]. Finally, our expressions are applicable in the case of very low string tensions $G\mu \lesssim 10^{-20}$ and reproduce the numerical spectra as can be seen in Fig. 6.3. The analytical expressions of Ref. [403], on the other hand, turn negative and, therefore, unphysical in this part of parameter space.

Discussion of the spectrum: Next, let us discuss the characteristics of the spectrum, first the minimum and maximum frequencies $f_{\rm rr}^{\rm min/max}$, between which the spectrum is non-vanishing. From Eq. (6.8), it is obvious that these frequencies are reached when

$$x_{\rm rr}^{\rm start} = x_{\rm rr}^{\rm end} \,. \tag{6.19}$$

While $x_{\rm rr}^{\rm end} = x_r^{\rm eq}$ is always true, for $x_{\rm rr}^{\rm start}$, we have to distinguish between two cases as can be seen from Eq. (6.12). We denote the frequency at which we have to switch from one case to the other by $f_{\rm rr}^{\rm start,*}$. Using Eq. (6.12) together with Eq. (6.14), one finds

$$f_{\rm rr}^{\rm start,*} = \frac{f_r^{\rm ini}}{\chi_r}. \tag{6.20}$$

Let us begin with the minimum frequency $f_{\rm rr}^{\rm min}$ and, thus, low frequencies. Considering Eq. (6.12), we note that $\chi_r^{-1} > 0$ and that $\varphi_2(x_r^{\rm ini}, \chi_r)$ is monotonically increasing as a function of f, starting from $\varphi_2 = 0$ at $f = 0$ as can be recognised from the explicit form given in Eq. (6.15). Therefore, we find at sufficiently low frequencies that $x_{\rm rr}^{\rm start} = \chi_r^{-1}$. Here, sufficiently low means $f \leq f_{\rm rr}^{\rm start,*}$. Assuming that the lowest possible frequency satisfies $f_{\rm rr}^{\rm min} \leq f_{\rm rr}^{\rm start,*}$ allows us to replace $x_{\rm rr}^{\rm start}$ by χ_r^{-1} and we find from Eq. (6.19) that

$$f_{\rm rr}^{\rm min} = \frac{f_r^{\rm eq}}{\chi_r}. \tag{6.21}$$

Since $a_{\rm ini} < a_{\rm eq}$ and accordingly $f_r^{\rm eq} < f_r^{\rm ini}$, it follows that $f_{\rm rr}^{\rm min} < f_{\rm rr}^{\rm start,*}$ and hence, our assumption is always satisfied.

The physics behind this result can easily be understood. The lowest frequencies correspond to the largest loop sizes. The largest loops are those with both the longest initial length and the least time to decay, and hence, those produced at the latest times. For the RR spectrum, these are the string loops produced at $t_{\rm eq}$ with length $l_{\rm rr}^{\rm max} = \alpha\xi_r t_{\rm eq} = \Gamma G\mu\chi_r t_{\rm eq}$. The corresponding observed frequency today is

$$f_{\rm rr}^{\rm min} = \frac{a_{\rm eq}}{a_0}\frac{2}{l_{\rm rr}^{\rm max}} = \frac{a_0}{a_{\rm eq}}\frac{4H_r^0}{\Gamma G\mu}\frac{1}{\chi_r} = \frac{f_r^{\rm eq}}{\chi_r} \,, \tag{6.22}$$

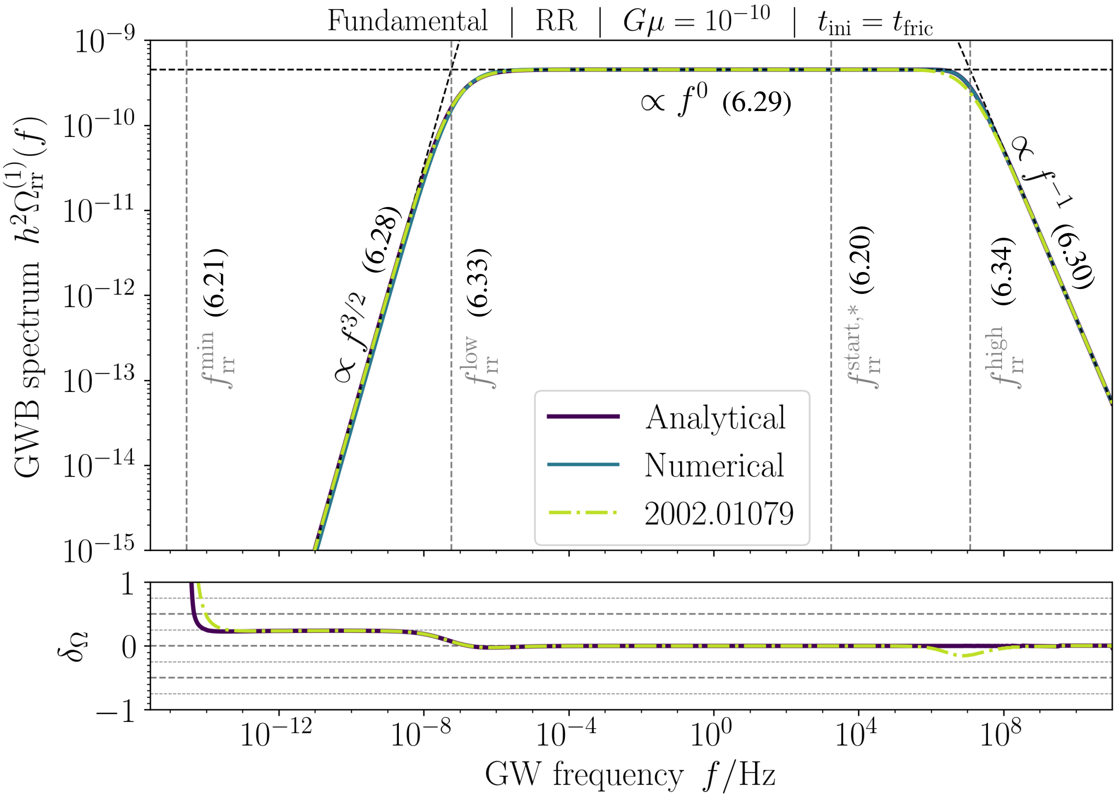


Figure 6.1: Fundamental GW spectrum for the RR case, a string tension of $G\mu = 10^{-10}$ and an initial time $t_{\rm ini} = t_{\rm fric}$. Upper panel: Our exact numerical result based on the VOS model accounting for the full time dependence of all relevant quantities and for a fixed effective number of DOFs as explained at the end of the introduction to this chapter (teal), our analytical result derived in this chapter (deep purple), and the analytical result derived in Ref. [403] (light green, dash-dotted). Grey dashed lines show characteristic frequencies of the spectrum, and black dashed lines indicate different power-law behaviours of the spectrum. Lower panel: Relative deviation δ_Ω (cf. Eq. (6.18)) of the two analytical results from our numerical result.
Figure reproduced from Ref. [4].

where we assumed perfect radiation domination to express $a_{\rm eq} = a(t_{\rm eq})$. This sets the lowest possible frequency for which the spectrum does not yet vanish and, thus, reproduces Eq. (6.21).

Let us now turn to $f_{\rm rr}^{\rm max}$. In the frequency range in which $f \geq f_{\rm rr}^{\rm start,*}$, Eq. (6.19) reads $x_r^{\rm eq} = \varphi_2(x_r^{\rm ini}, \chi_r)$. Using this equality in Eq. (6.14), it can be brought into the form

$$f_{\rm rr}^{\rm max} = \frac{f_r^{\rm eq}}{(1+\chi_r)\left(\frac{a_{\rm ini}}{a_{\rm eq}}\right)^2 - 1}\,. \tag{6.23}$$

This maximum has to be a non-negative number and, therefore, only exists if the condition $(1+\chi_r)\,(a_{\rm ini}/a_{\rm eq})^2 > 1$ holds true. If it exists, our initial assumption $f_{\rm rr}^{\rm max} \geq f_{\rm rr}^{\rm start,*}$ is, upon using again that $a_{\rm eq} > a_{\rm ini}$, equivalent to $\chi_r > -\left(a_{\rm ini} + a_{\rm eq}\right)/a_{\rm ini}$ and thus always satisfied.

As before, the condition for the existence of the maximum can be understood intuitively. The ratio $(1+\chi_r)\,(a_{\rm ini}/a_{\rm eq})^2$ is, in perfect radiation domination, the ratio between the time of complete decay of the loops born at $t_{\rm ini}$ and matter–radiation equality $t_{\rm eq}$. The condition that this ratio be larger than 1 amounts to the statement that the earliest loops have not completely decayed before $t_{\rm eq}$. These loops are, however, the ones with the smallest initial loop size and have, in addition, the longest time to decay. Correspondingly, they are the smallest loops that contribute to the RR spectrum and, therefore, lead to the highest frequencies. In case they decay completely, the frequencies in the RR spectrum can (at least in the framework considered here) become infinitely

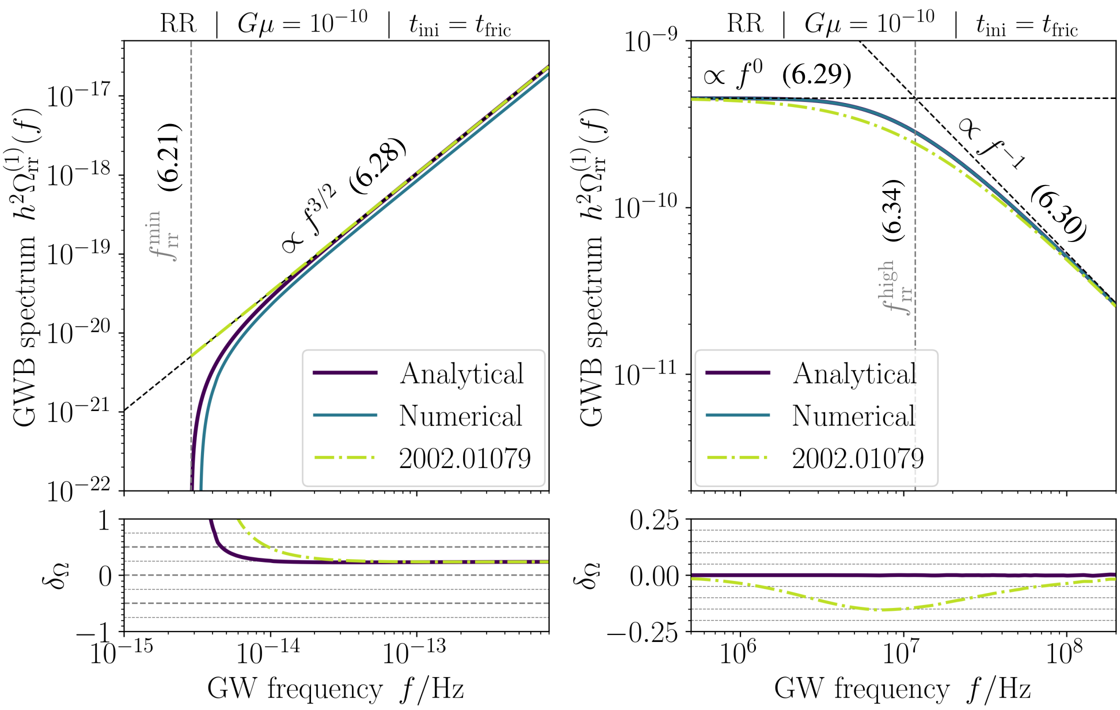


Figure 6.2: Left: Very-low-frequency region of the spectra shown in Fig. 6.1. Our numerical spectrum and the analytical spectrum of this work drop to zero as they approach $f_{\rm rr}^{\rm min}$, while the result of Ref. [403] needs to be cut off by hand. **Right:** Region of the spectra shown in Fig. 6.1 in which the transition from the plateau to the high-frequency regime occurs. The spectrum determined in Ref. [403] deviates in this regime by $\sim 20\%$ from our numerically calculated one, while our analytical result shows excellent agreement.
Figure adapted from Ref. [4].

large. If there is not enough time for the decay to complete, the highest frequency will correspond to the smallest loop size reached before matter–radiation equality, thus explaining why there is a maximum frequency. We can calculate this highest allowed frequency. The loops born at $t_{\rm ini}$ with length $l_{\rm ini} = \alpha\xi_r t_{\rm ini}$ have, according to Eq. (4.67), at $t_{\rm eq}$ a length

$$l_{\rm rr}^{\rm min} = \alpha\xi_r t_{\rm ini} - \Gamma G\mu\left(t_{\rm eq} - t_{\rm ini}\right) = \Gamma G\mu\left[(1+\chi_r)t_{\rm ini} - t_{\rm eq}\right] . \tag{6.24}$$

Under the assumption of perfect radiation domination, i.e., $t_r(a) = (a/a_0)^2 \,/\, \left(2H_r^0\right)$, this corresponds to a maximum observed frequency of

$$f_{\rm rr}^{\rm max} = \frac{a_{\rm eq}}{a_0}\frac{2}{l_{\rm rr}^{\rm min}} = \underbrace{\frac{a_0}{a_{\rm eq}}\frac{4H_r^0}{\Gamma G\mu}}_{=f_r^{\rm eq}}\frac{1}{(1+\chi_r)\left(\frac{a_{\rm ini}}{a_{\rm eq}}\right)^2 - 1} , \tag{6.25}$$

and, hence, reproduces our previous result given in Eq. (6.23). Observe that this is analogous to the argument that led us to the cutoff frequency of low-scale strings in Ch. 5. However, here we consider only RR loops and, therefore, a cutoff frequency occurs already if none of the loops decayed until $t_{\rm eq}$, not only if they survived until the present time t_0.

Having found the minimum and maximum frequency for the spectrum, the Heaviside function in Eq. (6.8) can be replaced by the function

$$\left[1 - \Theta(f - f_{\rm rr}^{\rm max})\Theta(f_{\rm rr}^{\rm max})\right]\Theta(f - f_{\rm rr}^{\rm min}) . \tag{6.26}$$

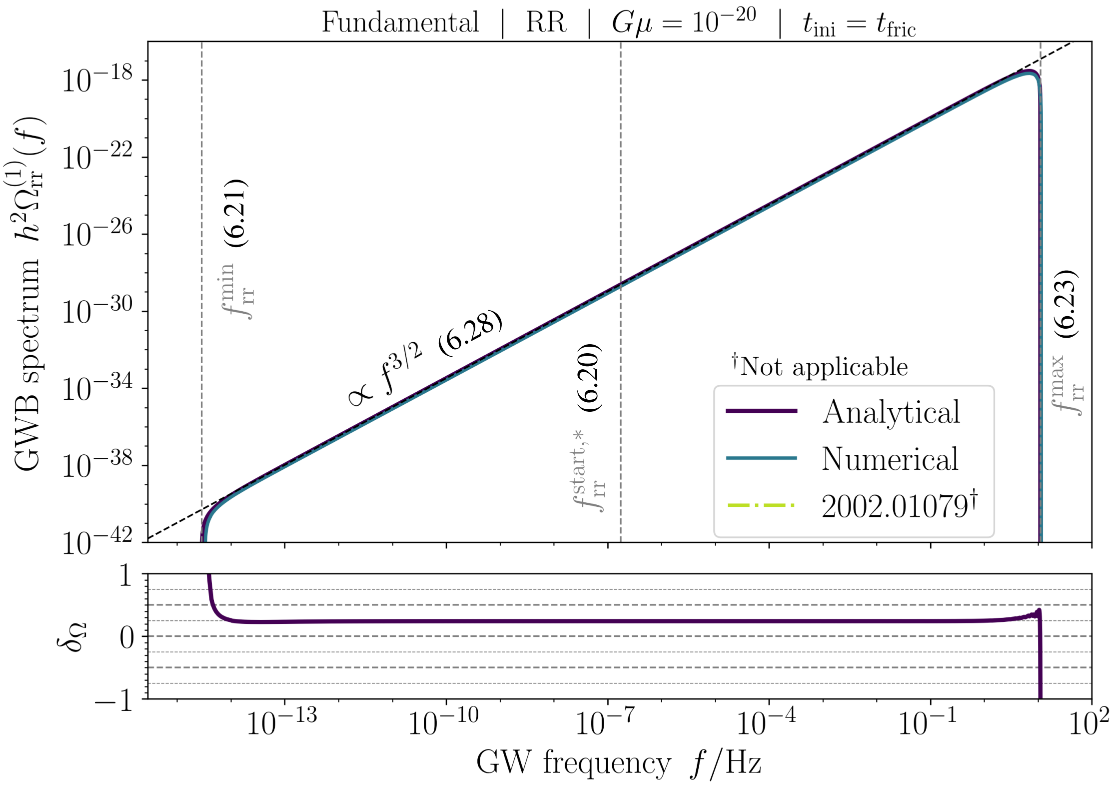


Figure 6.3: Fundamental GW spectrum for the RR case, a string tension of $G\mu = 10^{-20}$ and an initial time $t_{\rm ini} = t_{\rm fric}$. Panels and colour codes are the same as in Fig. 6.1. The spectrum derived in Ref. [403] does not appear in the plot since it turns negative.
Figure reproduced from Ref. [4].

Let us continue by discussing the shape of the spectrum and start with its low-frequency behaviour. For $f \ll f^{\rm eq}_r$ and $f \leq f^{\rm start,*}_{\rm rr}$, the spectrum reads

$$h^2\Omega^{(1)}_{\rm rr,low} = \mathcal{A}_{\rm rr}\left(\left(x^{\rm end}_{\rm rr}\right)^{3/2} - \left(\frac{1}{1+\chi_r}\right)^{3/2}\right) . \tag{6.27}$$

If furthermore $f \gg f^{\rm eq}_r/\left(1+\chi_r\right)$, which can with Eq. (6.21) be understood as the condition that we are at low frequencies but still sufficiently far above the minimum frequency, the spectrum is dominated by the first term. Correspondingly, the spectrum goes like

$$\begin{aligned} h^2\Omega^{(1)}_{\rm rr,low} &= \mathcal{A}_{\rm rr}\left(\frac{f}{f^{\rm eq}_r}\right)^{3/2} \\ &\simeq 3.29\times 10^{-14}\left(\frac{C_r}{0.171}\right)\left(\frac{\Gamma}{50}\right)\left(\frac{G\mu}{10^{-10}}\right)^2\left(\frac{f}{10^{-10}\,\mathrm{Hz}}\right)^{3/2} . \end{aligned} \tag{6.28}$$

For intermediate frequencies $f^{\rm end}_{\rm rr} \ll f \ll f^{\rm start}_{\rm rr}$, the term $S^{(1)}_{\rm rr}(x^{\rm end}_{\rm rr})$ in Eq. (6.8) is roughly one, while the term $S^{(1)}_{\rm rr}(x^{\rm start}_{\rm rr})$ is still negligible. This gives rise to a plateau of height

$$h^2\Omega^{(1)}_{\rm rr,plateau} = \mathcal{A}_{\rm rr} . \tag{6.29}$$

Finally, at large frequencies $f^{\rm start}_{\rm rr} \ll f$, both $S^{(1)}_{\rm rr}(x^{\rm start}_{\rm rr})$ and $S^{(1)}_{\rm rr}(x^{\rm end}_{\rm rr})$ are relevant and the value of the spectrum follows from their impartial cancellation. Asymptotically, it approaches the form

$$h^2\Omega^{(1)}_{\mathrm{rr,high}} = \frac{3\mathcal{A}_{\mathrm{rr}}}{2(1+\chi_r)^{1/2}x_r^{\mathrm{ini}}}\left(1-(1+\chi_r)^{1/2}\,\frac{a_{\mathrm{ini}}}{a_{\mathrm{eq}}}\right)\,. \tag{6.30}$$

This can, of course, only hold if the high-frequency region is not cut off by the appearance of a maximum frequency. If, moreover, $(1+\chi_r)^{1/2}a_{\mathrm{ini}}/a_{\mathrm{eq}} \ll 1$, which is always the case in the absence of a maximum frequency, this asymptotic expression can be further simplified to

$$h^2\Omega^{(1)}_{\mathrm{rr,high}} \simeq \frac{3\mathcal{A}_{\mathrm{rr}}}{2\,(1+\chi_r)^{1/2}\,x_r^{\mathrm{ini}}}\,. \tag{6.31}$$

For all physically relevant cases, $\chi_r \gg 1$ holds. In this regime, we find

$$\begin{aligned} h^2\Omega^{(1)}_{\mathrm{rr,high}} &\to \frac{3\,\mathcal{A}_{\mathrm{rr}}}{2\,\chi_r^{1/2}}\left(\frac{f_r^{\mathrm{ini}}}{f}\right) \\ &\simeq 5.34\times10^{-18}\left(\frac{C_r}{0.171}\right)\left(\frac{50}{\Gamma}\right)\left(\frac{G\mu}{10^{-10}}\right)\left(\frac{10^{15}\,\mathrm{Hz}}{f}\right)\,. \end{aligned} \tag{6.32}$$

Having derived expressions for the three regions of the spectrum, we can also find the frequencies at which we transition from one region to another. The low-frequency regime ends, and the plateau begins at $f^{\mathrm{low}}_{\mathrm{rr}}$ where $\Omega^{(1)}_{\mathrm{rr,low}}(f^{\mathrm{low}}_{\mathrm{rr}}) = \Omega^{(1)}_{\mathrm{rr,plateau}}(f^{\mathrm{low}}_{\mathrm{rr}})$, which means due to Eqs. (6.28) and (6.29) that

$$f^{\mathrm{low}}_{\mathrm{rr}} = f^{\mathrm{eq}}_r. \tag{6.33}$$

Similarly, the plateau region ends and the high-frequency regime begins at the frequency $f^{\mathrm{high}}_{\mathrm{rr}}$ which satisfies $\Omega^{(1)}_{\mathrm{rr,high}}(f^{\mathrm{high}}_{\mathrm{rr}}) = \Omega^{(1)}_{\mathrm{rr,plateau}}(f^{\mathrm{high}}_{\mathrm{rr}})$. This translates via Eqs. (6.29) and (6.30) to

$$f^{\mathrm{high}}_{\mathrm{rr}} = \frac{3f_r^{\mathrm{ini}}}{2(1+\chi_r)^{1/2}}\left(1-(1+\chi_r)^{1/2}\,\frac{a_{\mathrm{ini}}}{a_{\mathrm{eq}}}\right) \simeq \frac{3}{2\,(1+\chi_r)^{1/2}}\,f_r^{\mathrm{ini}}\,. \tag{6.34}$$

In summary, we see that, typically, $h^2\Omega^{(1)}_{\mathrm{rr}}$ first rises like $f^{3/2}$ until it reaches a flat plateau and then decays like f^{-1}. This shape of the spectrum can be seen in Fig. 6.1. If a maximum frequency occurs in the spectrum, the high-frequency part and the plateau might disappear completely, and the spectrum is described by the low-frequency $f^{3/2}$-behaviour only, dropping rapidly to zero as it approaches the minimum and maximum frequencies; see Fig. 6.3.

Ultrahigh-frequency (UHF) regime: Before moving on to the next type of loops, we want to return to the discussion at the end of Sec. 4.3.3. There, we included the fact that we do not expect the emission of GWs with wavelengths smaller than the string width. We found that this leads to a frequency-dependent earliest emission time, which is, in terms of the scale factor, expressed in Eq. (4.93). While the general way to include this in the total spectrum is clear from the discussion in the mentioned section, we want to see how this affects the spectrum explicitly. Since the effect concerns the initial time, it will influence the spectrum at extremely large frequencies. As discussed before (see e.g. Sec. 4.3.3), a typical choice for the initial time is the moment when the friction regime ends and the scaling regime begins, which is roughly at $t_{\mathrm{fric}} = t_{\mathrm{Pl}}/\left(G\mu\right)^2$. The initial time would then be

$$t_{\mathrm{ini}} = \max\left\{t_{\mathrm{min}}(f), t_{\mathrm{fric}}\right\}\,, \tag{6.35}$$

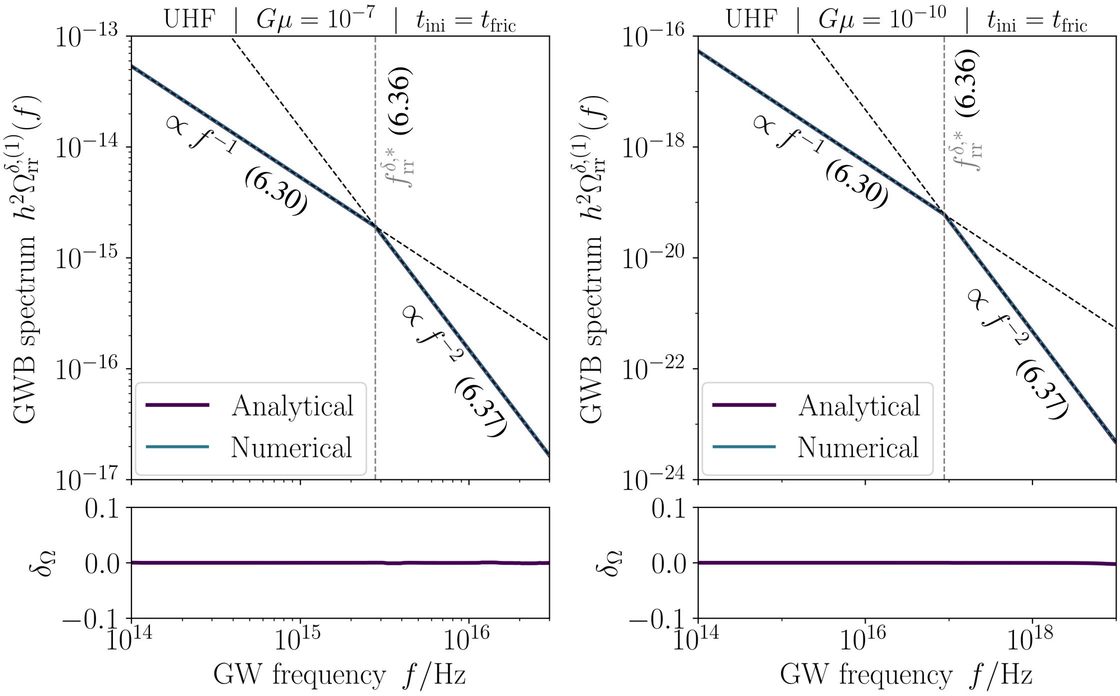


Figure 6.4: UHF regime of the fundamental GW spectrum for the RR case, string tensions of $G\mu = 10^{-7}$ (left) and $G\mu = 10^{-10}$ (right) and initial times given by $t_{\rm ini} = t_{\rm fric}$. Panels and colour codes are the same as in Fig. 6.1. The grey dashed line shows the frequency at which the UHF effects become relevant.
Figure reproduced from Ref. [4].

such that our consideration of the finite string width becomes only relevant at frequencies larger than

$$f_{\rm rr}^{\delta,*} = \left(\frac{2H_r^0}{G\mu\, t_{\rm Pl}}\right)^{1/2} = 8.82\times 10^{16}{\rm Hz}\left(\frac{10^{-10}}{G\mu}\right)^{1/2}. \tag{6.36}$$

Above this frequency, the power law will also change since $a_{\rm ini}/a_0$ in the high-frequency power-law of Eq. (6.32) is now frequency-dependent as well. Just replacing $a_{\rm min}$ with the expression in Eq. (4.93) yields

$$\begin{aligned} h^2\Omega_{\rm rr,high}^{\delta,(1)} &\simeq \frac{3\mathcal{A}_{\rm rr}}{2}\left(\frac{G\mu}{\chi_r}\right)^{1/2}\frac{h_r^0}{t_{\rm Pl}f^2} = \\ &= 4.71\times 10^{-26}\left(\frac{C_r}{0.171}\right)\left(\frac{50}{\Gamma}\right)\left(\frac{G\mu}{10^{-10}}\right)^{1/2}\left(\frac{10^{20}{\rm Hz}}{f}\right)^2. \end{aligned} \tag{6.37}$$

Both the f^{-1} power law derived before, as well as the f^{-2} decay of the spectrum at ultrahigh frequencies, can be seen in the plots in Fig. 6.4 for two different string tensions. Looking at the relative deviation between the analytical expression and the numerically computed spectrum, we find that the agreement is excellent (up to numerical noise).

Similar considerations can be found in Ref. [415], which reaches the conclusion that only higher harmonic contributions are affected. This is in contrast to the results that we arrive at. We will compare our results to those of Ref. [415] when turning to the total spectrum.

Finally, it is clear that the spectrum will drop to zero once $a_{\rm min}(f) = a_{\rm eq}$. That is, there is always a maximum frequency. Solving this condition for f, one finds this to be at

$$f_{\rm rr}^{\delta,\max} = \frac{a_{\rm eq}}{a_0}\frac{(G\mu)^{1/2}}{t_{\rm Pl}} = 5.42\times 10^{34}{\rm Hz}\left(\frac{G\mu}{10^{-10}}\right)^{1/2}, \tag{6.38}$$

which is, for any reasonable string tension, far out of reach of any current experiment.

6.1.2 Varying degrees of freedom

From the previous discussion, it is clear that there is no bijection between the GW frequency occurring in the spectrum and the time the radiation was emitted. At each time, there is a large number of loops emitting at different frequencies and, furthermore, the frequency of emission is redshifted from the time of emission until today. Nevertheless, it is also clear that high frequencies in the spectrum are predominantly associated with gravitational radiation produced at early times and low frequencies with gravitational radiation produced at late times. In particular, changes in the effective number of DOFs during radiation domination will lead to modifications of the plateau region of the RR spectrum: A larger number of DOFs at early times suppresses the plateau at larger frequencies. As pointed out in many articles (cf. e.g. Refs. [390, 429–432]), modelling these modifications properly is of great importance since a measurement of the spectrum allows then to draw conclusions on the thermal history of the Universe, including phase transitions after the formation of the strings.

Let us now try to approximate the changes due to variations in the effective number of relativistic DOFs. These enter the GW spectrum through the scale factor and, hence, it is advisable to first reflect on the Friedmann equation, which during radiation domination reads

$$H = \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t} = H_r^0\,\mathcal{G}^{1/2}(t)\left(\frac{a_0}{a}\right)^2 , \qquad \text{with} \qquad \mathcal{G}(t) = \frac{g_\rho(t)}{g_\rho^0}\left(\frac{g_s^0}{g_s(t)}\right)^{4/3} \tag{6.39}$$

where $g_{\rho,s}$ are the effective number of energetic and entropic DOFs at a given time. We define

$$\bar{H}_r(t) \equiv \frac{H_r^0}{t}\int_0^t \sqrt{\mathcal{G}(t')}\,\mathrm{d}t' , \tag{6.40}$$

such that we can rewrite the Friedmann equation as

$$\frac{a(t)}{a_0} = \left(2\bar{H}_r(t)\,t\right)^{1/2} . \tag{6.41}$$

If we were to consider no changes in the effective DOFs, we would simply have $\bar{H}_r(t) \to H_r^0$.

This modification of the scale factor evolution enters the GW spectrum now in three places. First, it alters the loop number density in Eq. (4.72), such that we obtain

$$\tilde{n}_{\mathrm{rr,DOF}}(l,t) = \left(\frac{\bar{H}_r(t_*(l,t))}{\bar{H}_r(t)}\right)^{3/2} \tilde{n}_{\mathrm{rr}}(l,t) , \tag{6.42}$$

where the time of loop production t_* is given in Eq. (4.78). Second, the integrand of the fundamental spectrum given in Eq. (4.91) contains, in addition to the loop number density, scale factor ratios accounting for redshifting between the time of GW emission and observation. Finally, we need to express the length l in the loop number density in the integrand in terms of frequency and scale factor, $l(f,t) = 2\sqrt{2\bar{H}_r(t)t}/f$. Putting all of this together in Eq. (4.91), we find

$$\Omega_{\mathrm{GW}}^{(1)}(f) = \frac{16\pi}{3}\left(\frac{G\mu}{H_0}\right)^2 \frac{\Gamma}{f}\int_{t_{\min}}^{t_{\max}} \left(\bar{H}_r\left(t_*\right)\right)^{3/2} \bar{H}_r(t)\,\frac{2^{5/2}C_r t\,\mathrm{d}t}{\left(\frac{2\sqrt{2\bar{H}_r(t)t}}{f} + \Gamma G\mu t\right)^{5/2}} . \tag{6.43}$$

In order to carry out this integral as well as for the later purpose of mode summation, we simplify the expansion history in the already familiar manner. We assume that we can divide the radiation-dominated era into N subintervals during which $\mathcal{G}$ is constant, i.e., we further divide the product of

Heaviside functions occurring in Eq. (4.77a):

$$\Theta\left(t_{\rm eq}-t\right)\Theta\left(t-t_*\right)\Theta\left(t_*-t_{\rm ini}\right)=\sum_{i=0}^{N-1}\left(\Theta\left(t_{(i+1)}-t\right)\Theta\left(t_*-t_{(i)}\right)\Theta\left(t-t_*\right)+\right. \tag{6.44}$$
$$\left.+\sum_{j=0}^{i-1}\Theta\left(t_{(i+1)}-t\right)\Theta\left(t-t_{(i)}\right)\Theta\left(t_{(j+1)}-t_*\right)\Theta\left(t_*-t_{(j)}\right)\right),$$

with $t_0=t_{\rm ini}$ and $t_N=t_{\rm eq}$. The above sum contains $N(N+1)/2$ terms, namely one term for each combination of birth and emission time interval, with the restriction that the birth time cannot be in an interval later than the one the emission time is in. We denoted the emission time interval with an index i and the birth time interval with an index j. Restricting to this simplification, we are able to carry out the integral in Eq. (6.43) and find[1]

$$h^2\Omega^{(1)}_{ij}(f)=\Theta\left(x_{ij}^{\rm end}-x_{ij}^{\rm start}\right)\mathcal{A}_{\rm rr}\left(\frac{\bar{H}_r^j}{H_r^0}\right)^{3/2}\left(\frac{\bar{H}_r^i}{H_r^0}\right)^{1/2}S^{(1)}_{\rm rr}(x)\Big|_{x_{ij}^{\rm start}}^{x_{ij}^{\rm end}}. \tag{6.45}$$

Let us now specify the form of the integration boundaries. These derive, just as before, directly from the Heaviside functions in Eq. (6.44). There are clearly two distinct cases, namely, either $i=j$ or $i>j$, which we label A and B, respectively. Case A corresponds to GW emission occurring within the same constant-DOF interval in which the emitting loop was produced — analogous to the RR and MM cases. Case B, on the other hand, corresponds to the emission taking place during a later constant-DOF interval than the one in which the loop was formed — analogous to the RM case, although the intervals corresponding to i and j need not be adjacent. In case A, the structure is exactly the same as before for the RR spectrum upon replacing $t_{\rm eq}\to t_{(i+1)}$ and $t_{\rm ini}\to t_{(i)}$, yielding

$$x_{A,i}^{\rm start}=\max\left\{\chi_r^{-1},\varphi_2(x_i^{(i)},\chi_r)\right\}\qquad\text{and}\qquad x_{A,i}^{\rm end}=x_i^{(i+1)}, \tag{6.46}$$

with $x_k^{(i)}=f/f_k^{(i)}$ and $f_k^{(i)}=\frac{4\mathcal{G}_k^{1/2}H_r^0}{\Gamma G\mu}\frac{a_0}{a_{(i)}}$. In case B, we find

$$x_{B,ij}^{\rm start}=\max\left\{x_i^{(i)},\varphi_2(x_i^{(j)},\chi_r)\right\}\qquad\text{and}\qquad x_{B,ij}^{\rm end}=\min\left\{x_i^{(i+1)},\varphi_2(x_i^{(j+1)},\chi_r)\right\}. \tag{6.47}$$

We can then replace the previous expression for the fundamental RR spectrum by

$$h^2\Omega^{(1)}_{\rm rr}=\mathcal{A}_{\rm rr}\sum_{i=0}^{N-1}\left(\mathcal{G}_i\,S^{(1)}_{\rm rr}(x)\Big|_{x_{A,i}^{\rm start}}^{x_{A,i}^{\rm end}}\Theta(x_{A,i}^{\rm end}-x_{A,i}^{\rm start})+\right. \tag{6.48}$$
$$\left.+\sum_{j=0}^{i-1}\mathcal{G}_j^{3/4}\mathcal{G}_i^{1/4}\,S^{(1)}_{\rm rr}(x)\Big|_{x_{B,ij}^{\rm start}}^{x_{B,ij}^{\rm end}}\Theta(x_{B,ij}^{\rm end}-x_{B,ij}^{\rm start})\right).$$

Approximating the numerical evolution of the SM effective DOFs taken from Ref. [380], we will use $N=4$ with $a_{(0)}=a_{\rm ini}$ and $a_{(4)}=a_{\rm eq}$ as well as

$$a_{(1)}/a_0=1.6\times10^{-15},\qquad a_{(2)}/a_0=3.4\times10^{-12},\qquad a_{(3)}/a_0=2.0\times10^{-9}, \tag{6.49}$$

---

[1] We set $\bar{H}_r(t)$ constant during an interval with constant number of DOFs, which is not exact but a very good approximation since these intervals have, on order of magnitude, the same length in logarithmic time. For the linear time integral determining $\bar{H}_r(t)$, earlier values of the DOFs are, therefore, negligible.

and

$$\mathcal{G}_0 = 0.39\,, \qquad \mathcal{G}_1 = 0.43\,, \qquad \mathcal{G}_2 = 0.83\,, \qquad \mathcal{G}_3 = 1\,.$$

One needs to take care of the fact that $a_{\rm ini}$ can become large for large values of $t_{\rm ini}$. Keeping the labelling $a_{(0)} = a_{\rm ini}$ fixed, this means that N can decrease due to this. The earliest value of $\mathcal{G}$, still labelled $\mathcal{G}_0$, will then be the one that is present in the time interval starting with $a_{\rm ini}$.

A final ingredient that will be useful when turning to the total spectrum are the switch, minimum and maximum frequencies. For the former, we have

$$f_{A,i}^{\rm start,*} = \frac{f_i^{(i)}}{\chi_r}, \quad f_{B,ij}^{\rm start,*} = \frac{f_i^{(i)}}{(1+\chi_r)\left(\frac{a_{(j)}}{a_{(i)}}\right)^2 - 1}, \quad f_{B,ij}^{\rm end,*} = \frac{f_i^{(i+1)}}{(1+\chi_r)\left(\frac{a_{(j+1)}}{a_{(i+1)}}\right)^2 - 1}. \tag{6.50}$$

With these (and knowledge about the low frequency-behaviour of φ_2), we can write

$$x_{A,i}^{\rm start} = \begin{cases} \chi_r^{-1} & \text{if } f < f_{A,i}^{\rm start,*} \\ \varphi_2(x_i^{(i)}, \chi_r) & \text{else} \end{cases}, \tag{6.51}$$

$$x_{B,ij}^{\rm end} = \begin{cases} \varphi_2(x_i^{(j+1)}, \chi_r) & \text{if } f < f_{B,ij}^{\rm end,*} \\ x_r^{i+1} & \text{else} \end{cases}, \tag{6.52}$$

$$x_{B,ij}^{\rm start} = \begin{cases} x_i^{(i)} & \text{if } f < f_{B,ij}^{\rm start,*} \\ \varphi_2(x_i^{(j)}, \chi_r) & \text{else} \end{cases}. \tag{6.53}$$

For the latter, we find

$$f_{A,i}^{\rm min} = \frac{f_i^{(i+1)}}{\chi_r}, \qquad f_{A,i}^{\rm max} = \frac{f_i^{(i+1)}}{(1+\chi_r)\left(\frac{a_{(i)}}{a_{(i+1)}}\right)^2 - 1}, \tag{6.54}$$

$$f_{B,ij}^{\rm min} = \frac{f_i^{(i)}}{(1+\chi_r)\left(\frac{a_{(j+1)}}{a_{(i)}}\right)^2 - 1}, \qquad f_{B,ij}^{\rm max} = \frac{f_i^{(i+1)}}{(1+\chi_r)\left(\frac{a_{(j)}}{a_{(i+1)}}\right)^2 - 1}. \tag{6.55}$$

Finally, let us also mention, that since a larger effective number of DOFs at high energies will suppress the amplitude of the spectrum at high frequencies, the high-frequency power law will be modified. Instead of Eq. (6.31), we find

$$h^2\Omega_{\rm rr,high}^{(1)} \simeq \frac{3\mathcal{A}_{\rm rr}\mathcal{G}_0}{2\,(1+\chi)^{1/2}\, x_0^{(0)}}\,. \tag{6.56}$$

6.1.3 RM loops

We continue with the consideration of string loops produced during radiation domination, but investigate now the gravitational radiation that they emit during the matter-dominated era. Correspondingly, this time, we use in the expression for the GW spectrum in Eq. (4.91) the loop number density given in Eqs. (4.82) and (4.77b), which yields

$$\begin{aligned} h^2\Omega_{\rm rm}^{(1)} = h^2\frac{16\pi}{3}\left(\frac{G\mu}{H_0}\right)^2 \frac{\Gamma}{f}\frac{1}{t_{\rm eq}^{3/2}}\left(\frac{\Omega_r}{\Omega_m}\right)^3 C_r\Theta\left(a_{\rm end} - a_{\rm start}\right) \times \\ \times \int_{a_{\rm start}/a_0}^{a_{\rm end}/a_0} \frac{\frac{a}{a_0}{\rm d}\left(\frac{a}{a_0}\right)}{\left(\frac{2}{f}\frac{a}{a_0} + \Gamma G\mu t_m(a)\right)^{5/2} H_m(a)}\,. \end{aligned} \tag{6.57}$$

The above integral runs only over scale factors in the matter-dominated era, and due to our assumption of pure matter domination during this time, we were able to simplify the expressions for the Hubble parameter and the cosmic time as a function of the scale factor:

$$H_m\left(a\right) = H_m^0 \left(\frac{a_0}{a}\right)^{3/2} \qquad \text{and} \qquad t_m\left(a\right) = \frac{2}{3H_m^0}\left(\frac{a}{a_0}\right)^{3/2}\,, \tag{6.58}$$

where we defined $H_m^0 = H_0\left(\Omega_m\right)^{1/2}$.

As before, it is advantageous to change integration variables to

$$x_m = \frac{f}{f_m}\,, \qquad \text{with} \qquad f_m = \left(\frac{a_0}{a}\right)^{1/2} h_m^0, \tag{6.59}$$

where f_m carries the dependence on the scale factor and we introduced the frequency

$$h_m^0 = \frac{3H_m^0}{\Gamma G\mu} \simeq 7.36\times 10^{-10}\,\text{Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right)\,. \tag{6.60}$$

With this substitution, the above integral takes the form

$$h^2\Omega_{\rm rm}^{(1)} = \sqrt{96}\pi C_r \frac{h^2\Omega_m\left(\frac{a_{\rm eq}}{a_0}\right)^3}{\left(H_m^0 t_{\rm eq}\right)^{3/2}}\left(\frac{G\mu}{\Gamma}\frac{h_m^0}{f}\right)^{1/2}\Theta\left(x_{\rm rm}^{\rm end}-x_{\rm rm}^{\rm start}\right)\int_{x_{\rm rm}^{\rm start}}^{x_{\rm rm}^{\rm end}}\frac{x_m\,\mathrm{d}x_m}{\left(1+x_m\right)^{5/2}} \tag{6.61}$$

and evaluates to the following expression for the spectrum

$$h^2\Omega_{\rm rm}^{(1)} = \mathcal{A}_{\rm rm}\Theta\left(x_{\rm rm}^{\rm end}-x_{\rm rm}^{\rm start}\right)\left[S_{\rm rm}^{(1)}(x_{\rm rm}^{\rm start}) - S_{\rm rm}^{(1)}(x_{\rm rm}^{\rm end})\right]\,. \tag{6.62}$$

The amplitude introduced above reads

$$\begin{aligned}\mathcal{A}_{\rm rm} &= \sqrt{\frac{128}{3}}\,\pi\, C_r\,\frac{h^2\Omega_m}{\left(H_m^0 t_{\rm eq}\right)^{3/2}}\left(\frac{a_{\rm eq}}{a_0}\right)^3\left(\frac{G\mu}{\Gamma}\right)^{1/2} \\ &\simeq 6.51\times 10^{-9}\left(\frac{C_r}{0.171}\right)\left(\frac{50}{\Gamma}\right)^{1/2}\left(\frac{G\mu}{10^{-10}}\right)^{1/2}\,,\end{aligned} \tag{6.63}$$

where $t_{\rm eq} \simeq 5.05\times 10^4\,$yr denotes the time of matter–radiation equality. The spectral shape function for RM loops takes the form

$$S_{\rm rm}^{(1)}\left(x\right) = \frac{2+3\,x}{\left(x_m^0\right)^{1/2}\left(1+x\right)^{3/2}} \tag{6.64}$$

and needs to be evaluated at the integration boundaries $x_{\rm rm}^{\rm start}$ and $x_{\rm rm}^{\rm end}$. These boundaries occur, as before, due to the Heaviside function in the loop number density in Eq. (4.77b). The conditions $t_* > t_{\rm ini}$ and $t > t_{\rm eq}$ contribute to the lower integration boundary $x_{\rm rm}^{\rm start}$, while the conditions $t_* < t_{\rm eq}$ and $t < t_0$ are taken into account by $x_{\rm rm}^{\rm end}$. To write down explicit expressions for these boundaries, it is helpful to introduce another variable $\tilde{x}_m$ that does not involve the scale factor a, but instead the (would-be) scale factor

$$\frac{\tilde{a}_m\left(t\right)}{a_0} = \left(\frac{3}{2}H_m^0 t\right)^{2/3}\,, \tag{6.65}$$

which describes the evolution of a purely matter-dominated Universe. We define in this way

$$\tilde{x}_m = \frac{f}{\tilde{f}_m}\,, \qquad \text{with} \qquad \tilde{f}_m = \left(\frac{a_0}{\tilde{a}_m}\right)^{1/2} h_m^0\,. \tag{6.66}$$

During the matter dominated era, the function $\tilde{a}_m\left(t\right)$ and the actual scale factor $a\left(t\right)$ in the standard Λ-cold-dark-matter (ΛCDM) model are (nearly) identical. However, in the following, we will also evaluate $\tilde{a}_m\left(t\right)$ at times outside the matter dominated era, where $\tilde{a}_m\left(t\right)$ and $a\left(t\right)$ no longer coincide with each other. Utilising this, the integration boundaries for x_m can be expressed as

$$x_{\rm rm}^{\rm start} = \max\left\{x_m^{\rm eq}, \varphi_3\left(\tilde{x}_m^{\rm ini}, \chi_r\right)\right\} \,, \qquad x_{\rm rm}^{\rm end} = \min\left\{x_m^{0}, \varphi_3\left(\tilde{x}_m^{\rm eq}, \chi_r\right)\right\} \,, \tag{6.67}$$

where $\varphi_3(x,\chi)$ is the real and positive solution of the cubic equation

$$\varphi^3 + \varphi^2 = \left(1+\chi\right)x^3 \,. \tag{6.68}$$

Explicitly, one finds[2]

$$\varphi_3\left(x,\chi\right) = \frac{1}{3}\left[\phi^{1/3}\left((1+\chi)x^3\right) + \frac{1}{\phi^{1/3}\left((1+\chi)x^3\right)} - 1\right] \,, \tag{6.69}$$

with

$$\phi\left(y\right) = \frac{27}{2}\,y - 1 + 3^{3/2}\left(\frac{27}{4}\,y^2 - y\right)^{1/2} \,. \tag{6.70}$$

In Eq. (6.67), the function φ_3 is evaluated for $\chi = \chi_r$ as well as

$$\tilde{x}_m^{\rm ini} = \frac{f}{\tilde{f}_m^{\rm ini}} \qquad \text{and} \qquad \tilde{x}_m^{\rm eq} = \frac{f}{\tilde{f}_m^{\rm eq}} \,. \tag{6.71}$$

Here, we defined the frequencies

$$\tilde{f}_m^{\rm ini} = \left(\frac{a_0}{\tilde{a}_m^{\rm ini}}\right)^{1/2} h_m^0 \simeq 6.00\times 10^4\,{\rm Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right)\left(\frac{10^{-24}\,{\rm s}}{t_{\rm ini}}\right)^{1/3} \,, \tag{6.72}$$

$$\tilde{f}_m^{\rm eq} = \left(\frac{a_0}{\tilde{a}_m^{\rm eq}}\right)^{1/2} h_m^0 \simeq 5.14\times 10^{-8}\,{\rm Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right) \,, \tag{6.73}$$

where we denoted $\tilde{a}_m^{\rm ini} = \tilde{a}_m\left(t_{\rm ini}\right)$ and $\tilde{a}_m^{\rm eq} = \tilde{a}_m\left(t_{\rm eq}\right)$. In Eq. (6.67), we also need

$$x_m^{\rm eq} = \frac{f}{f_m^{\rm eq}} \qquad \text{and} \qquad x_m^{0} = \frac{f}{f_m^{0}} \tag{6.74}$$

with the additional frequencies

$$f_m^{\rm eq} = \left(\frac{a_0}{a_{\rm eq}}\right)^{1/2} h_m^0 \simeq 4.30\times 10^{-8}\,{\rm Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right) , \tag{6.75}$$

and $f_m^0 = h_m^0$. Returning to $t_{\rm fric} = t_{\rm Pl}/\left(G\mu\right)^2$ as the standard choice for $t_{\rm ini}$, $\tilde{f}_m^{\rm ini}$ becomes

$$\tilde{f}_m^{\rm fric} = 3.42\times 10^4\,{\rm Hz}\left(\frac{50}{\Gamma}\right)\left(\frac{10^{-10}}{G\mu}\right)^{1/3} \,. \tag{6.76}$$

The RM spectra obtained as a result of the above computation are depicted in Figs. 6.5a to 6.5d, and we will turn to a discussion of their features and qualitative differences shortly. Before

---

[2] That this solution is always real can be seen as follows. First, consider Eq. (6.70). As long as $y \geq \frac{4}{27}$, $\phi(y)$ is real and positive and φ_3 is (upon picking the real solution of $\phi^{1/3}$) real and positive as well. In the case $0 < y < \frac{4}{27}$, $\phi(y)$ acquires a non-vanishing imaginary part. At the same time, we find $|\phi(y)|^2 = 1$. This implies that $\phi^*(y) = 1/\phi(y)$ and hence $\varphi_3(x,\chi) = \frac{1}{3}[2\,{\rm Re}(\phi^{1/3}((1+\chi)\,x^3)) - 1]$, which is manifestly real.

that, let us briefly comment on improvements over the RM spectra derived in Ref. [403] that are visible in those plots. A difference common to all four figures is that our spectrum exhibits a minimum frequency $f_{\rm rm}^{\rm min}$ below which it strictly drops to zero. This is the same behaviour one finds in the numerical calculation. In contrast, the analytical spectrum predicted in earlier works extends, in principle, to arbitrarily low frequencies, which would physically correspond to GWs from arbitrarily large string loops. As explained in Ref. [403], their spectrum must be cut off by hand. In later work by some of the same authors in Ref. [411], the cutoff was given explicitly and coincides with our expression for $f_{\rm rr}^{\rm min}$, which is slightly different from the actual cutoff $f_{\rm rm}^{\rm min}$, as can be seen in Figs. 6.5a to 6.5d. However, more importantly, since the cutoff is not contained in the spectral shape of the expressions given in Ref. [403], these largely deviate at low frequencies from the numerically computed spectra. This deviation is clearly visible in all four plots. For certain regions of parameter space, namely low string tensions and late initial times, our spectrum also contains a maximum frequency $f_{\rm rm}^{\rm max}$ and drops to zero above, as illustrated in Fig. 6.5d. This high-frequency cutoff is nothing but the cutoff of low-scale strings that we discussed in Ch. 5. This result could not be predicted by earlier analytical studies such as Ref. [403]. The remaining differences come from a change in the power law describing the spectrum at low frequencies, as we will discuss below.

Discussion of the spectrum: We are now going to examine the details of the RM spectrum. Two of the most important characteristics are the minimum and maximum frequency, which are reached when the spectrum in Eq. (6.62) vanishes. This clearly is the case if

$$x_{\rm rm}^{\rm start} = x_{\rm rm}^{\rm end} . \tag{6.77}$$

From the explicit expressions for the integration boundaries in Eq. (6.67), we see that it is necessary to distinguish between two cases for each of them. Let us begin by determining the frequencies $f_{\rm rm}^{\rm start,*}$ and $f_{\rm rm}^{\rm end,*}$ at which we have to switch from one behaviour in $x_{\rm rm}^{\rm start}$ or $x_{\rm rm}^{\rm end}$ respectively to another one. In the former case, the switch occurs when $x_m^{\rm eq} = \varphi_3(\tilde{x}_m^{\rm ini}, \chi_r)$, in the latter case when $x_m^0 = \varphi_3(\tilde{x}_m^{\rm eq}, \chi_r)$. Using these conditions to substitute φ_3 in Eq. (6.68) yields

$$f_{\rm rm}^{\rm start,*} = \frac{f_m^{\rm eq}}{\left(\frac{\tilde{a}_m^{\rm ini}}{a_{\rm eq}}\right)^{3/2}(1+\chi_r) - 1} \qquad \text{and} \qquad f_{\rm rm}^{\rm end,*} = \frac{f_m^0}{\left(\frac{\tilde{a}_m^{\rm eq}}{a_0}\right)^{3/2}(1+\chi_r) - 1} . \tag{6.78}$$

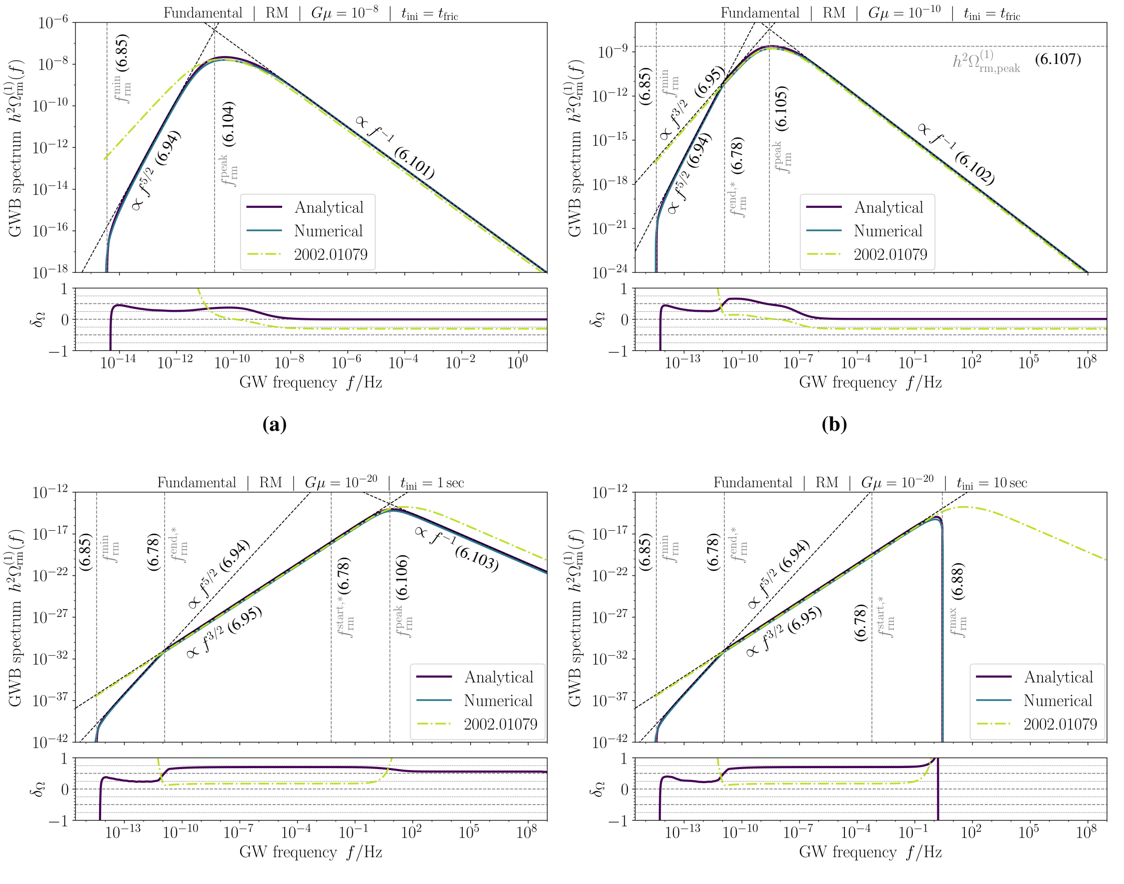
Now, consider first the switch at $f_{\rm rm}^{\rm start,*}$, which only occurs if this frequency is positive or, equivalently, if $(1+\chi_r)^{2/3}\tilde{a}_m^{\rm ini} > a_{\rm eq}$. In typical scenarios, this is not fulfilled since a combination of very large values of $t_{\rm ini}$ and very low string tensions $G\mu$ is necessary. To see which solution is realised below and above the switch frequency, consider the behaviour at high frequencies first. At those, we have $\varphi_3(x,\chi) \longrightarrow (1+\chi)^{1/3}x$ and, accordingly, find the general asymptotic behaviour

$$x_{\rm rm}^{\rm start} \longrightarrow \begin{cases} (1+\chi_r)^{1/3}\,\tilde{x}_{\rm m}^{\rm ini} & \text{if } (1+\chi_r)^{2/3}\,\tilde{a}_{\rm m}^{\rm ini} > a_{\rm eq} \\ x_{\rm m}^{\rm eq} & \text{else} \end{cases} . \tag{6.79}$$

In the upper case, a switch from one solution to the other must occur since $f_{\rm rm}^{\rm start,*}$ exists. Since we just found that the high-frequency behaviour corresponds to $\varphi_3(\tilde{x}_m^{\rm ini}, \chi_r)$, this is the solution relevant above $f_{\rm rm}^{\rm start,*}$. Accordingly, $x_m^{\rm eq}$ is the relevant solution below the switch frequency. In the lower case, there is no switch frequency and since $x_m^{\rm eq}$ is relevant at high frequencies, it is relevant everywhere. Note that, even though we have used a large-frequency expansion, the result of this argument holds exactly. In total, we find:

$$\text{If } (1+\chi_r)^{2/3}\,\tilde{a}_m^{\rm ini} > a_{\rm eq}, \text{ then} \qquad x_{\rm rm}^{\rm start} = \begin{cases} x_m^{\rm eq} & \text{if } f < f_{\rm rm}^{\rm start,*} \\ \varphi_3(\tilde{x}_m^{\rm ini}, \chi_r) & \text{else} \end{cases} .$$

$$\text{Otherwise} \qquad x_{\rm rm}^{\rm start} = x_m^{\rm eq} . \tag{6.80}$$



(a) **(b)** **(c)** **(d)**

Figure 6.5: Four qualitatively different fundamental GW spectra in the RM case. While Figs. 6.5b and 6.5c exhibit the same power laws, they are qualitatively different due to the occurrence of $f_{\rm rm}^{\rm start,*}$ as discussed at the very end of Sec. 6.1.3. Panels and colour codes are the same as in Fig. 6.1. Recall from our discussion in Sec. 5.1.2 that large initial times such as $t_{\rm ini} = 1\,$sec or $10\,$sec chosen in the lower panels are well motivated, especially for very low string tensions. Note that the RM spectra provided in Ref. [403] are independent of $t_{\rm ini}$. Following the discussion in Refs. [403, 411], we cut their RM spectra off at the frequency $f_{\rm rr}^{\rm min}$.
Figure adapted from Ref. [4].

Next, consider the switch at $f_{\rm rm}^{\rm end,*}$, which requires $(1+\chi_r)^{2/3}\tilde{a}_m^{\rm eq} > a_0$. Unlike the condition for the first case, this is satisfied for most of the physically relevant parameter choices.[3] To identify which solution is realised at which frequency, we progress in a similar manner as before, resulting in the statement:

$$\text{If } (1+\chi_r)^{2/3}\tilde{a}_m^{\rm eq} > a_0, \text{ then} \qquad x_{\rm rm}^{\rm end} = \begin{cases} \varphi_3(\tilde{x}_m^{\rm eq}, \chi_r) & \text{if } f < f_{\rm rm}^{\rm end,*} \\ x_m^0 & \text{else} \end{cases} .$$
$$\text{Otherwise} \qquad x_{\rm rm}^{\rm end} = \varphi_3(\tilde{x}_m^{\rm eq}, \chi_r)\,. \tag{6.81}$$

If both switch frequencies exist, which is automatically the case if $f_{\rm rm}^{\rm start,*}$ does, then we can also show that usually

$$f_{\rm rm}^{\rm end,*} < f_{\rm rm}^{\rm start,*}. \tag{6.82}$$

To this end, first note that $f_m^0/f_m^{\rm eq} = (a_{\rm eq}/a_0)^{1/2} < 1$. Besides, since we typically[4] have $\tilde{a}_m^{\rm ini}/a_{\rm eq} <$

---

[3] In particular, for our standard parameter values, this is satisfied as long as $G\mu < 5.85 \times 10^{-9}$.

[4] This is true for initial times as late as $t_{\rm ini} \sim 5 \times 10^6$s. For much larger $t_{\rm ini}$ the order can be inverted.

Table 6.1: Case distinction for $x_{\rm rm}^{\rm start}$ and $x_{\rm rm}^{\rm end}$ relevant for the RM spectrum.

Case	**Condition**	$\boldsymbol{x_{\rm rm}^{\rm start}}$	$\boldsymbol{x_{\rm rm}^{\rm end}}$
1	Either $f_{\rm rm}^{{\rm end},*}$ does not exist or $f < f_{\rm rm}^{{\rm end},*}$	$x_m^{\rm eq}$	$\varphi_3(\tilde{x}_m^{\rm eq}, \chi_r)$
2	$f_{\rm rm}^{{\rm end},*} < f$ and either $f_{\rm rm}^{{\rm start},*}$ does not exist or $f < f_{\rm rm}^{{\rm start},*}$	$x_m^{\rm eq}$	x_m^0
3	$f_{\rm rm}^{{\rm start},*}$ exists and $f_{\rm rm}^{{\rm start},*} < f$	$\varphi_3(\tilde{x}_m^{\rm ini}, \chi_r)$	x_m^0

$\tilde{a}_m^{\rm eq}/a_0$, we know that

$$1 < \frac{\left(\frac{\tilde{a}_m^{\rm eq}}{a_0}\right)^{3/2}(1+\chi_r) - 1}{\left(\frac{\tilde{a}_m^{\rm ini}}{a_{\rm eq}}\right)^{3/2}(1+\chi_r) - 1} \,. \tag{6.83}$$

Putting these two results together, we immediately obtain Eq. (6.82).

Given a frequency f at which we want to investigate the spectrum, we have to distinguish in total three possible cases, which are summarised in Tab. 6.1. With the distinction between these three cases at hand, we can return to the initial problem of determining the maximum and minimum frequencies of the spectrum. The criterion for them to occur is given in Eq. (6.77). Let us begin with case 1 from Tab. 6.1 in which we assume that either $f_{\rm rm}^{{\rm end},*}$ exists and $f_{\rm rm}^{\rm min} < f_{\rm rm}^{{\rm end},*}$ or $f_{\rm rm}^{{\rm end},*}$ does not exist. Then, Eq. (6.77) becomes

$$x_m^{\rm eq} = \varphi_3(\tilde{x}_m^{\rm eq}, \chi_r)\,. \tag{6.84}$$

With the same argument as in the derivation of $f_{\rm rm}^{{\rm start},*}$, upon replacing $\tilde{x}_m^{\rm eq} \leftrightarrow \tilde{x}_m^{\rm ini}$, one finds

$$f_{\rm rm}^{\rm min} = \frac{f_m^{\rm eq}}{\left(\frac{\tilde{a}_m^{\rm eq}}{a_{\rm eq}}\right)^{3/2}(1+\chi_r) - 1}\,. \tag{6.85}$$

Hence, a minimum can only exist if $\tilde{a}_m^{\rm eq}(1+\chi_r)^{2/3} > a_{\rm eq}$ which is always realised for physically reasonable string tensions.[5] Before we continue with the next case, we have to make sure that, given the existence of $f_{\rm rm}^{{\rm end},*}$, our assumption $f_{\rm rm}^{\rm min} < f_{\rm rm}^{{\rm end},*}$ is actually satisfied. Let us start from this assumption and bring it into the form

$$\frac{f_m^{\rm eq}}{f_m^0} = \left(\frac{a_0}{a_{\rm eq}}\right)^{1/2} < \frac{\left(\frac{\tilde{a}_m^{\rm eq}}{a_{\rm eq}}\right)^{3/2}(1+\chi_r) - 1}{\left(\frac{\tilde{a}_m^{\rm eq}}{a_0}\right)^{3/2}(1+\chi_r) - 1}\,. \tag{6.86}$$

First, note that the derivative of the right side with respect to χ_r is always negative and hence monotonically decreasing in χ_r. This function reaches a minimum of $(a_0/a_{\rm eq})^{3/2}$ for $\chi_r \to \infty$. In

---

[5] The non-existence of a minimum, however, does not mean that the spectrum extends to $f = 0$. This would not be reasonable either since the radiation would need to stem from arbitrarily large string loops. It is rather the case that non-satisfaction of the previous criterion implies $x_m^{\rm eq} \geq \varphi_3\left(\tilde{x}_m^{\rm eq}, \chi_r\right)$ and, thus, the Heaviside function in Eq. (6.62) vanishes and there is no RM contribution. From a physical viewpoint, this is only possible if loops produced during radiation domination have decayed fast enough so that none reaches matter domination. This, in turn, is only possible if there is a gap between the epoch of radiation and matter domination, which in reality does not occur. Nonetheless, it appears as an artefact of our treatment because of the splitting of the history of the Universe into a pure matter and a pure radiation part due to which $\tilde{a}_m^{\rm eq}/a_{\rm eq} < 1$, while at the same time we require loop production before $\tilde{a}_m^{\rm eq}$ and loop decay after $a_{\rm eq}$. Since $\tilde{a}_m^{\rm eq}/a_{\rm eq} = (2-\sqrt{2})^{2/3} \simeq 0.7$, this happens only if $\chi_r < (2-\sqrt{2})^{-1} - 1 \simeq 0.7$ which corresponds, under the assumption that $\alpha\xi_r = 0.1$, to huge string tensions, $\log_{10}\left(G\mu\right) \gtrsim -2.6$. However, if one allows $\alpha\xi_r$ to be several order of magnitude smaller than 1, it can actually become relevant for more realistic string tensions.

this case, the inequality (6.86) is indeed satisfied, and due to the monotonicity of the right-hand side, it also holds true for any other allowed value of χ_r implying that our initial assumption was justified. Correspondingly, we do not need to consider cases 2 or 3 since they would require $f_{\rm rm}^{\rm min}$ to be larger than the existing switch frequency $f_{\rm rm}^{{\rm end},*}$.

On the other hand, to find the maximum frequency, we need to examine cases 2 and 3 in Tab. 6.1. Let us begin with case 2, i.e., $f_{\rm rm}^{{\rm end},*}$ exists and fulfils $f_{\rm rm}^{{\rm end},*} < f_{\rm rm}^{\rm max}$. Furthermore, either $f_{\rm rm}^{{\rm start},*}$ exists and meets the condition $f_{\rm rm}^{\rm max} < f_{\rm rm}^{{\rm start},*}$ or $f_{\rm rm}^{{\rm start},*}$ does not exist. The criterion in Eq. (6.86) turns into $x_m^{\rm eq} = x_m^0$ or equivalently $a_{\rm eq}/a_0 = 1$, which is never true. This means that either $f_{\rm rm}^{{\rm start},*}$ exists, which implies that a possible maximum frequency must be larger than this frequency, or $f_{\rm rm}^{{\rm start},*}$ does not exist, and there is no maximum frequency such that the spectrum extends to arbitrarily high frequencies.

The last case is that $f_{\rm rm}^{{\rm start},*}$ exists and $f_{\rm rm}^{\rm max} > f_{\rm rm}^{{\rm start},*}$. The criterion in Eq. (6.77) becomes

$$\varphi_3(\tilde{x}_m^{\rm ini}, \chi_r) = x_m^0 \,. \tag{6.87}$$

Using the same argument as in the derivation of $f_{\rm rm}^{{\rm end},*}$, we obtain, upon exchanging $\tilde{x}_m^{\rm ini} \leftrightarrow \tilde{x}_m^{\rm eq}$,

$$f_{\rm rm}^{\rm max} = \frac{f_m^0}{\left(\frac{\tilde{a}_m^{\rm ini}}{a_0}\right)^{3/2}(1+\chi_r) - 1} \,. \tag{6.88}$$

Therefore, a maximum can only exist if $\tilde{a}_m^{\rm ini}(1+\chi_r)^{2/3} > a_0$. The validity of our assumption $f_{\rm rm}^{\rm max} > f_{\rm rm}^{{\rm start},*}$ can easily be proven by following an argument analogous to the one below Eq. (6.85). As for the RR spectrum, one can heuristically argue that there is the possibility for a maximum frequency below which the spectrum strictly vanishes. Such a high-frequency cutoff must arise if none of the string loops has fully decayed until today. The ones that would decay first are those produced with the smallest length and the longest time to decay. These loops are those produced at $t_{\rm ini}$ and their length today sets the minimum length $l_{\rm rm}^{\rm min}$ of all loops, in particular all RM loops. Via $f_{\rm rm}^{\rm max} = 2/l_{\rm rm}^{\rm min}$, this minimum length sets the maximum frequency of the spectrum, which, if one follows the presented arguments, leads again to the frequency in Eq. (6.88) we obtained by evaluating the Heaviside function in the spectrum.

The relevant frequencies, their order, and the cases in which they occur can be schematically summarised as follows:

$$f_{\rm rm}^{\rm min} \mathrel{\substack{\Leftarrow \\ <}} f_{\rm rm}^{{\rm end},*} \mathrel{\substack{\Leftarrow \\ <}} f_{\rm rm}^{{\rm start},*} \mathrel{\substack{\Leftarrow \\ <}} f_{\rm rm}^{\rm max} \,. \tag{6.89}$$

The implication arrow indicates that, e.g., if $f_{\rm rm}^{\rm max}$ exists, then so does $f_{\rm rm}^{{\rm start},*}$ and so on. The minimum frequency always exists, while the first switch frequency almost always exists, except for very large string tensions. The second switch frequency and the maximum frequency exist only in the case of very low $G\mu$ and very large $t_{\rm ini}$. The physical reason for the emergence of a maximum frequency in this parameter regime was one of the central aspects discussed in Ch. 5.

Equipped with the minimum and maximum frequencies, we can replace the Heaviside function in Eq. (6.62) with the more explicit expression

$$\left[1 - \Theta(f - f_{\rm rm}^{\rm max})\Theta(f_{\rm rm}^{\rm max})\right]\left[1 - \Theta(f_{\rm rm}^{\rm min} - f)\Theta(f_{\rm rm}^{\rm min})\right] \,. \tag{6.90}$$

Having determined the most important frequencies, except for the peak frequency to which we will come later, we shall now discuss the shape of the spectrum. Let us begin with asymptotically low frequencies, at which we find from a small x expansion of the spectral shape function in Eq. (6.64) that the spectrum can in first non-vanishing order be written as

$$h^2\Omega_{\rm rm,low}^{(1)} = \frac{3\mathcal{A}_{\rm rm}}{4\sqrt{x_m^0}}\left((x_{\rm rm}^{\rm end})^2 - (x_{\rm rm}^{\rm start})^2\right) \,. \tag{6.91}$$

Moreover, it is possible to use the low-frequency expansion

$$\varphi_3(x,\chi) \overset{x\to 0}{\longrightarrow} \sqrt{1+\chi}\, x^{3/2}\,. \tag{6.92}$$

Therefore, we obtain in case 1 of Tab. 6.1

$$h^2\Omega^{(1)}_{\mathrm{rm,low}} = \frac{3\mathcal{A}_{\mathrm{rm}}}{4(x_m^0)^{1/2}}\left((1+\chi_r)\,(\tilde{x}_m^{\mathrm{eq}})^3 - (x_m^{\mathrm{eq}})^2\right). \tag{6.93}$$

If we are considering low frequencies but still $f \gg f_{\mathrm{rm}}^{\mathrm{min}}$, then the second term is negligible. Upon noting that furthermore $\chi_r \gg 1$, we can approximate

$$\begin{aligned} h^2\Omega^{(1)}_{\mathrm{rm,low}} &= \frac{3\mathcal{A}_{\mathrm{rm}}}{4}\chi_r\left(\frac{\tilde{a}_m^{\mathrm{eq}}}{a_0}\right)^{1/4}\left(\frac{f}{\tilde{f}_m^{\mathrm{eq}}}\right)^{5/2} \simeq \\ &\simeq 1.95\times 10^{-14}\left(\frac{C_r}{0.171}\right)\left(\frac{\Gamma}{50}\right)\left(\frac{G\mu}{10^{-10}}\right)^2\left(\frac{f}{10^{-12}\,\mathrm{Hz}}\right)^{5/2}. \end{aligned} \tag{6.94}$$

In case 2, we find straightforwardly

$$\begin{aligned} h^2\Omega^{(1)}_{\mathrm{rm,low}} &= \frac{3\mathcal{A}_{\mathrm{rm}}}{4(x_m^0)^{1/2}}\left((x_m^0)^2 - (x_m^{\mathrm{eq}})^2\right) = \frac{3\mathcal{A}_{\mathrm{rm}}}{4}\left(\frac{f}{f_m^0}\right)^{3/2}\left(1-\frac{a_{\mathrm{eq}}}{a_0}\right) \simeq \\ &\simeq 2.45\times 10^{-10}\left(\frac{C_r}{0.171}\right)\left(\frac{\Gamma}{50}\right)\left(\frac{G\mu}{10^{-10}}\right)^2\left(\frac{f}{10^{-10}\,\mathrm{Hz}}\right)^{3/2}. \end{aligned} \tag{6.95}$$

In case 3, we again have to use Eq. (6.92) and find

$$h^2\Omega^{(1)}_{\mathrm{rm,low}}(f) = \frac{3\mathcal{A}_{\mathrm{rm}}}{4(x_m^0)^{1/2}}\left((x_m^0)^2 - (1+\chi_r)(\tilde{x}_m^{\mathrm{ini}})^3\right). \tag{6.96}$$

If we furthermore assume that $f \ll f_{\mathrm{rm}}^{\mathrm{max}}$, then the second term becomes negligible, and this expression simplifies to

$$\begin{aligned} h^2\Omega^{(1)}_{\mathrm{rm,low}}(f) &= \frac{3\mathcal{A}_{\mathrm{rm}}}{4}\left(\frac{f}{f_m^0}\right)^{3/2} \simeq \\ &\simeq 2.45\times 10^{-18}\left(\frac{C_r}{0.171}\right)\left(\frac{\Gamma}{50}\right)\left(\frac{G\mu}{10^{-20}}\right)^2\left(\frac{f}{10^{-2}\,\mathrm{Hz}}\right)^{3/2}, \end{aligned} \tag{6.97}$$

which is the same as the expression derived for case 2, up to a relative deviation in the prefactor of the order of 3×10^{-4}.

For asymptotically large frequencies, the spectrum takes, at leading order, generally the form

$$h^2\Omega^{(1)}_{\mathrm{rm,high}} = \frac{3\mathcal{A}_{\mathrm{rm}}}{(x_m^0)^{1/2}}\left(\frac{1}{(x_{\mathrm{rm}}^{\mathrm{start}})^{1/2}} - \frac{1}{(x_{\mathrm{rm}}^{\mathrm{end}})^{1/2}}\right). \tag{6.98}$$

We, furthermore, can use the behaviour

$$\varphi_3(x,\chi) \overset{x\to\infty}{\longrightarrow} (1+\chi)^{1/3}\, x \tag{6.99}$$

and immediately see that $h^2\Omega^{(1)}_{\mathrm{rm,high}}(f) \propto f^{-1}$ in all cases of Tab. 6.1. In more detail, we obtain in the first case

$$h^2\Omega^{(1)}_{\mathrm{rm,high}} = \frac{3\mathcal{A}_{\mathrm{rm}}}{f}\left(\left(f_m^0 f_m^{\mathrm{eq}}\right)^{1/2} - \frac{\left(f_m^0\tilde{f}_m^{\mathrm{eq}}\right)^{1/2}}{(1+\chi_r)^{1/6}}\right). \tag{6.100}$$

As an approximation, which becomes slightly rough for large string tensions, we can neglect the second term since $\chi_r \gg 1$ and find the parameter dependence

$$h^2\Omega^{(1)}_{\rm rm,high} \simeq 1.10 \times 10^{-16} \left(\frac{C_r}{0.171}\right)\left(\frac{50}{\Gamma}\right)^{3/2}\left(\frac{10^{-10}}{G\mu}\right)^{1/2}\left(\frac{1\,{\rm Hz}}{f}\right). \tag{6.101}$$

For case 2, we have

$$\begin{aligned} h^2\Omega^{(1)}_{\rm rm,high} &= \frac{3\mathcal{A}_{\rm rm}}{(x_m^0)^{1/2}}\left(\frac{1}{(x_m^{\rm eq})^{1/2}} - \frac{1}{(x_m^0)^{1/2}}\right) = 3\mathcal{A}_{\rm rm}\frac{f_m^0}{f}\left(\left(\frac{a_0}{a_{\rm eq}}\right)^{1/4} - 1\right) \simeq \\ &\simeq 9.54 \times 10^{-17}\left(\frac{C_r}{0.171}\right)\left(\frac{50}{\Gamma}\right)^{3/2}\left(\frac{10^{-10}}{G\mu}\right)^{1/2}\left(\frac{1\,{\rm Hz}}{f}\right), \end{aligned} \tag{6.102}$$

while in the third case, we find

$$h^2\Omega^{(1)}_{\rm rm,high} = 3\mathcal{A}_{\rm rm}\left(\frac{1}{(1+\chi_r)^{1/6}}\left(\frac{a_0}{\tilde{a}_m^{\rm ini}}\right)^{1/4} - 1\right)\frac{f_m^0}{f}. \tag{6.103}$$

For the reasonable part of parameter space, there is no large hierarchy between the two occurring terms, and the expression cannot be further simplified in a sensible way, so one cannot extract a simple power-law dependence on the model parameters.

From the power laws we derived for the high-frequency and the low-frequency behaviour, we can acquire comparatively simple approximations for the peak frequency by determining the intersection between these asymptotic expressions: $h^2\Omega^{(1)}_{\rm rm,low}(f_{\rm rm}^{\rm peak}) = h^2\Omega^{(1)}_{\rm rm,high}(f_{\rm rm}^{\rm peak})$. If the peak lies in a region where case 1 of Tab. 6.1 applies, one obtains

$$f_{\rm rm}^{\rm peak} = \frac{\tilde{f}_{\rm m}^{\rm eq}}{(1+\chi_r)^{1/3}}\left(4(1+\chi_r)^{1/6}\left(\frac{\tilde{a}_m^{\rm eq}}{a_{\rm eq}}\right)^{1/4} - 4\right)^{2/7}. \tag{6.104}$$

If the peak lies in a region where the second case is valid, one finds

$$f_{\rm rm}^{\rm peak} = \gamma f_m^0\,, \qquad \text{with} \qquad \gamma = \left(4\frac{\left(\frac{a_0}{a_{\rm eq}}\right)^{1/4} - 1}{1 - \frac{a_{\rm eq}}{a_0}}\right)^{2/5}. \tag{6.105}$$

Finally, in the third case, the peak frequency lies at[6]

$$f_{\rm rm}^{\rm peak} = f_m^0\left(\frac{4}{(1+\chi_r)^{1/6}}\left(\frac{a_0}{\tilde{a}_m^{\rm ini}}\right)^{1/4} - 4\right)^{2/5}. \tag{6.106}$$

In the usually realised case at which the peak frequency satisfies the criteria of case 2 in Tab. 6.1, the full RM spectrum can be evaluated at this peak frequency and one finds a relatively simple expression for the peak amplitude. This amplitude is proportional to $\mathcal{A}_{\rm rm}$ and the proportionality factor depends on $a_{\rm eq}/a_0$ only, i.e., it is independent of string-related parameters:

$$h^2\Omega^{(1)}_{\rm rm,peak} = \frac{\mathcal{A}_{\rm rm}}{\gamma^{1/2}}\left(\frac{2 + 3\left(\frac{a_{\rm eq}}{a_0}\right)^{1/2}\gamma}{\left(1 + \left(\frac{a_{\rm eq}}{a_0}\right)^{1/2}\gamma\right)^{3/2}} - \frac{2+3\gamma}{(1+\gamma)^{3/2}}\right) \simeq 0.37\mathcal{A}_{\rm rm}\,. \tag{6.107}$$

---

[6] This case can only be realised for a small range of $G\mu$ values if one chooses $t_{\rm ini} = t_{\rm fric}$, since in this case, the existence of this peak frequency also requires $f_{\rm rm}^{\rm start,*}$ to exist, which translates to the criterion $3H_m^0\alpha\xi_r t_{\rm Pl}/(2\Gamma) < (G\mu)^3 < (a_0/a_{\rm eq})^{3/2}\, 3H_m^0\alpha\xi_r t_{\rm Pl}/(2\Gamma)$. For $\alpha\xi_r = 0.1$ and $\Gamma = 50$, this means $5\times 10^{-22} \lesssim G\mu \lesssim 4\times 10^{-20}$.

Let us briefly summarise the features that can be found in the RM spectrum. Starting at low frequencies, it generally contains a minimum frequency $f_{\mathrm{rm}}^{\mathrm{min}}$ below which it vanishes. Still at low frequencies, but sufficiently far above this minimum cutoff, we also find a generically occurring $f^{5/2}$ power-law behaviour of the spectrum, as can be seen in all plots 6.5a to 6.5d. If, on the other hand, $(1+\chi_r)^{2/3}\,\tilde{a}_m^{\mathrm{eq}} > a_0$, as is the case for sufficiently low string tensions $G\mu \lesssim 10^{-9}$, this low-frequency behaviour turns first into an $f^{3/2}$ power law above a frequency $f_{\mathrm{rm}}^{\mathrm{end},*}$. Such a transition is present for the parameter values chosen for Figs. 6.5b to 6.5d. At even lower string tensions or later initial times t_{ini}, another switch frequency $f_{\mathrm{rm}}^{\mathrm{start},*}$ occurs. At frequencies below this switch frequency, the lower integration boundary in Eq. (6.57) is determined by the condition $t > t_{\mathrm{eq}}$. Above the switch frequency, the condition $t_* > t_{\mathrm{ini}}$ determines the lower integration boundary. While qualitatively different, the $f^{3/2}$ power law remains, up to a prefactor, the same, and can be observed in Figs. 6.5c and 6.5d. Leaving the low-frequency range, the spectrum has two options. In the first case, it has no high-frequency cutoff, implying that it extends to arbitrarily large frequencies. Here, one finds an f^{-1} decay at asymptotically large frequencies, as illustrated in Figs. 6.5a to 6.5c. The spectrum peaks then at one of the frequencies given in Eqs. (6.104), (6.105), and (6.106). The respective behaviour is realised depending on whether the region below the peak is described by the $f^{5/2}$ power law, the first, or the second $f^{3/2}$ power law. In the second case, realised at very low string tensions and very late initial times t_{ini}, or, more quantitatively, if $\tilde{a}_m^{\mathrm{ini}}\,(1+\chi_r)^{2/3} > a_0$, a maximum frequency $f_{\mathrm{rm}}^{\mathrm{max}}$ arises, above which the spectrum vanishes and which cuts off the high-frequency region. This behaviour is shown in Fig. 6.5d and we discussed it in detail in Ch. 5. Note that the analytic expressions in Ref. [403] do not capture the $f^{5/2}$ power law at low frequencies and, furthermore, a minimum frequency cutoff has to be imposed by hand.

6.1.4 MM loops

The last contribution to the fundamental spectrum is the gravitational radiation emitted from string loops produced and decaying during the matter-dominated era. For this case, we have to use the loop number density of Eq. (4.77c) with Eq. (4.79) in the general formula for the GW spectrum given in Eq. (4.91), which yields

$$\begin{aligned} h^2\Omega_{\mathrm{mm}}^{(1)} = h^2\,\frac{16\pi}{3}\left(\frac{G\mu}{H_0}\right)^2\frac{\Gamma}{f}C_m\Theta(a_{\mathrm{end}}-a_{\mathrm{start}})\times \\ \times\int_{a_{\mathrm{start}}/a_0}^{a_{\mathrm{end}}/a_0}\frac{\left(\frac{a}{a_0}\right)^4\mathrm{d}\left(\frac{a}{a_0}\right)}{t_m^2(a)\left(\frac{2}{f}\frac{a}{a_0}+\Gamma G\mu t_m(a)\right)^2 H_m(a)}\,, \end{aligned} \tag{6.108}$$

where the expressions for the Hubble parameter H_m and cosmic time t_m are given in Eq. (6.58). As in the RM case, we change the integration variable to x_m defined in Eq. (6.66), such that

$$h^2\Omega_{\mathrm{mm}}^{(1)} = 18\pi h^2\Omega_m G\mu C_m\left(\frac{h_m^0}{f}\right)^2\Theta(x_{\mathrm{mm}}^{\mathrm{end}}-x_{\mathrm{mm}}^{\mathrm{start}})\int_{x_{\mathrm{mm}}^{\mathrm{start}}}^{x_{\mathrm{mm}}^{\mathrm{end}}}\frac{x_m^2}{(1+x_m)^2}\,\mathrm{d}x_m\,, \tag{6.109}$$

which can easily be evaluated as

$$h^2\Omega_{\mathrm{mm}}^{(1)} = \mathcal{A}_{\mathrm{mm}}\Theta\left(x_{\mathrm{mm}}^{\mathrm{end}}-x_{\mathrm{mm}}^{\mathrm{start}}\right)\left[S_{\mathrm{mm}}^{(1)}(x_{\mathrm{mm}}^{\mathrm{end}})-S_{\mathrm{mm}}^{(1)}(x_{\mathrm{mm}}^{\mathrm{start}})\right]\,. \tag{6.110}$$

The resulting amplitude reads

$$\mathcal{A}_{\mathrm{mm}} = 18\pi\,C_m\,h^2\Omega_m\,G\mu \simeq 3.13\times 10^{-11}\left(\frac{C_m}{0.0387}\right)\left(\frac{G\mu}{10^{-10}}\right)\,, \tag{6.111}$$

and the spectral shape function is of the form

$$S_{\rm mm}^{(1)}(x) = \frac{1}{(x_m^0)^2}\left(\frac{x^2+x-1}{1+x} - 2\,\ln\left(1+x\right)\right)\,. \tag{6.112}$$

Using again the notation defined in Eqs. (6.71) and (6.74), one finds for the integration boundaries in the MM case

$$x_{\rm mm}^{\rm start} = \max\left\{\chi_m^{-1}, \varphi_3\left(\tilde{x}_m^{\rm eq}, \chi_m\right)\right\} \qquad \text{and} \qquad x_{\rm mm}^{\rm end} = x_m^0\,, \tag{6.113}$$

where we further introduced

$$\chi_m = \frac{\alpha\xi_m}{\Gamma G\mu}\,. \tag{6.114}$$

The boundary $x_{\rm mm}^{\rm start}$ derives directly from the Heaviside functions in Eq. (4.77c), i.e., it accounts for the conditions $t > t_* > t_{\rm eq}$, while the boundary $x_{\rm mm}^{\rm end}$ corresponds to the upper integration boundary in the general expression for the GW spectrum today given in Eq. (4.91), which is merely $t < t_0$.

The spectrum obtained from these analytical considerations is depicted in Figs. 6.6a and 6.6b, which we briefly want to discuss before turning to studying the features of this spectrum in detail. As can easily be seen from the plots, the agreement between the analytical prediction and the numerical result is much worse than it is for the RM and the RR cases. This is, however, not unexpected and can be explained by the rough approximation we used for the cosmic expansion history. In particular, our analytical treatment assumes the reduced correlation length ξ to be constant during the radiation-dominated and the matter-dominated era. Recalling the numerical solution to the VOS model shown in Fig. 4.4, it is clear that, while this is a good approximation during radiation domination, it becomes poor for the matter-dominated era. In spite of this, our analytical approximation correctly reproduces the qualitative behaviour of the spectrum and, in contrast to Ref. [403], predicts the existence of a maximum frequency above which the spectrum vanishes. In absence of a maximum frequency, the results from Ref. [403] reproduce the numerical spectra slightly better. Overall, the MM spectrum, however, is less important for the complete spectrum than the RM and the RR spectrum since it typically yields a subdominant contribution. In Sec. 6.1.5, we will see how a simple modification of our analytical expressions can restore very good agreement between the analytical and the numerical complete fundamental spectrum.

Discussion of the spectrum: Let us now discuss the characteristics of the MM spectrum and begin with the determination of the minimum and maximum frequencies. These are reached when the spectrum given in Eq. (6.110) vanishes, which happens if

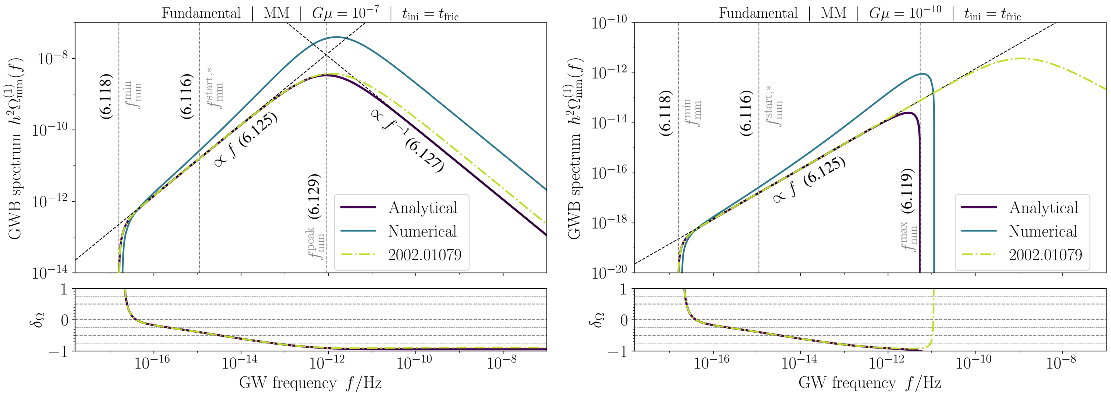
$$x_{\rm mm}^{\rm start} = x_{\rm mm}^{\rm end}\,. \tag{6.115}$$

Depending on the values of $f_{\rm mm}^{\rm min}$ and $f_{\rm mm}^{\rm max}$, we have to distinguish between two cases for $x_{\rm mm}^{\rm start}$. The switch from one behaviour to the other is located at the frequency which satisfies $\chi_m^{-1} = \varphi_3(\tilde{x}_m^{\rm eq}, \chi_m)$. Upon using Eq. (6.68), this can be seen to be equivalent to

$$f_{\rm mm}^{\rm start,*} = \frac{\tilde{f}_m^{\rm eq}}{\chi_m}\,. \tag{6.116}$$

We recall from the discussion of the RM spectrum that at very large frequencies, $\varphi_3(x,\chi) \propto x$. Hence, if f is sufficiently large, then $\varphi_3(\tilde{x}_m^{\rm eq}, \chi_m) > \chi_m^{-1}$ and, therefore,

$$x_{\rm mm}^{\rm start} = \begin{cases} \chi_m^{-1} & \text{if } f < f_{\rm mm}^{\rm start,*} \\ \varphi_3(\tilde{x}_m^{\rm eq}, \chi_m) & \text{else} \end{cases}\,. \tag{6.117}$$



(a) **(b)**

Figure 6.6: Fundamental GW spectra for the two qualitatively different cases of MM spectra. Panels and colour codes are the same as in Fig. 6.1.
Figure reproduced from Ref. [4].

Let us now come back to the minimum frequency and start from the assumption that $f^{\rm min}_{\rm mm} < f^{\rm start,*}_{\rm mm}$. Then Eq. (6.115) is equivalent to $x^0_m = \chi_m^{-1}$. This results in

$$f^{\rm min}_{\rm mm} = \frac{f^0_m}{\chi_m}\,. \tag{6.118}$$

To verify our initial assumption, observe that $a_0 > \tilde{a}^{\rm eq}_m$. Thus $f^0_m < \tilde{f}^{\rm eq}_m$, and correspondingly $f^{\rm min}_{\rm mm} = f^0_m/\chi_m < \tilde{f}^{\rm eq}_m/\chi_m = f^{\rm start,*}_{\rm mm}$. Observe that the minimum frequency always occurs in the MM spectrum. It is associated with the radiation from the largest loops ever produced, which are those created today with length $l^{\rm max}_{\rm mm} = \alpha\xi_m t_0$. This maximum length gives rise to the minimum frequency $f^{\rm min}_{\rm mm} = 2/l^{\rm max}_{\rm mm}$, which is, moreover, the overall minimum frequency.

Next, we turn to the maximum frequency and set out from the assumption that $f^{\rm max}_{\rm mm} > f^{\rm start,*}_{\rm mm}$. This means, the spectrum vanishes if $x^0_m = \varphi_3(\tilde{x}^{\rm eq}_m, \chi_m)$. The equation is the same as the one we had to solve for $f^{\rm start,*}_{\rm rm}$ upon making the replacements $\tilde{x}^{\rm ini}_m \leftrightarrow \tilde{x}^{\rm eq}_m$, $x^{\rm eq}_m \leftrightarrow x^0_m$, and $\chi_r \leftrightarrow \chi_m$. Hence, we find

$$f^{\rm max}_{\rm mm} = \frac{f^0_m}{(1+\chi_m)\left(\frac{\tilde{a}^{\rm eq}_m}{a_0}\right)^{3/2} - 1}\,. \tag{6.119}$$

While the switch and minimum frequencies always exist, a maximum is only present if $(1+\chi_m)^{2/3}\tilde{a}^{\rm eq}_m > a_0$. Heuristically, this is satisfied if and only if the earliest loops in matter domination born at time $t_{\rm eq}$ have not fully decayed until today. The maximum frequency then corresponds to the length $l^{\rm min}_{\rm mm}$ these shortest loops have nowadays due to the relation $f^{\rm max}_{\rm mm} = 2/l^{\rm min}_{\rm mm}$. We still have to check whether our supposition that $f^{\rm max}_{\rm mm} > f^{\rm start,*}_{\rm mm}$ was actually legitimate. This assumption can be brought into the equivalent form

$$1 < \frac{\left(\frac{\tilde{a}^{\rm eq}_m}{a_0}\right)^{1/2}\chi_m}{(1+\chi_m)\left(\frac{\tilde{a}^{\rm eq}_m}{a_0}\right)^{3/2} - 1}\,. \tag{6.120}$$

Since $(\tilde{a}^{\rm eq}_m/a_0)^{1/2} < 1$, it follows that $(\tilde{a}^{\rm eq}_m/a_0)^{1/2}\chi_m > (\tilde{a}^{\rm eq}_m/a_0)^{1/2}(1+\chi_m) - 1$ and, besides,

$$1 < \frac{(1+\chi_m)\left(\frac{\tilde{a}^{\rm eq}_m}{a_0}\right)^{1/2} - 1}{(1+\chi_m)\left(\frac{\tilde{a}^{\rm eq}_m}{a_0}\right)^{3/2} - 1}\,. \tag{6.121}$$

Combining these two results, we find that Eq. (6.120) is indeed satisfied.

The determination of the extremal frequencies allows us to replace the Heaviside function occurring in Eq. (6.110) with the expression

$$\left[1 - \Theta(f - f_{\mathrm{mm}}^{\mathrm{max}})\Theta(f_{\mathrm{mm}}^{\mathrm{max}})\right]\Theta(f - f_{\mathrm{mm}}^{\mathrm{min}})\,. \tag{6.122}$$

After discussing the most important frequencies of the spectrum, we want to consider its asymptotic behaviour. Let us begin with the low-frequency regime. The shape function in Eq. (6.112) allows for the expansion

$$S_{\mathrm{mm}}^{(1)}(x) \xrightarrow{x\to 0} \frac{1}{(x_m^0)^2}\left(-1 + \frac{x^3}{3}\right). \tag{6.123}$$

We first look at the case in which $f < f_{\mathrm{mm}}^{\mathrm{start},*}$ and hence $x_{\mathrm{mm}}^{\mathrm{start}} = \chi_m^{-1}$. This means we cannot use the above expansion of the shape function for $S_{\mathrm{mm}}^{(1)}(x_{\mathrm{mm}}^{\mathrm{start}})$ but need to continue with the full form instead. We then find

$$h^2\Omega_{\mathrm{mm,low}}^{(1)} = \mathcal{A}_{\mathrm{mm}}\left(-\frac{1}{(x_m^0)^2} + \frac{x_m^0}{3} - S_{\mathrm{mm}}^{(1)}(\chi_m^{-1})\right). \tag{6.124}$$

Still restricting to low frequencies, but assuming that we are sufficiently far above the minimum frequency $f \gg f_{\mathrm{mm}}^{\mathrm{min}}$, the spectrum further simplifies to

$$h^2\Omega_{\mathrm{mm,low}}^{(1)} = \frac{\mathcal{A}_{\mathrm{mm}}}{3}\,x_m^0 = 1.42\times 10^{-18}\left(\frac{C_m}{0.0387}\right)\left(\frac{\Gamma}{50}\right)\left(\frac{G\mu}{10^{-10}}\right)^2\left(\frac{f}{10^{-16}\,\mathrm{Hz}}\right). \tag{6.125}$$

For the second case, in which $f > f_{\mathrm{mm}}^{\mathrm{start},*}$, we have $x_{\mathrm{mm}}^{\mathrm{start}} = \varphi_3(\tilde{x}_m^{\mathrm{eq}}, \chi_m)$. Expanding at low frequencies to leading order, we obtain $x_{\mathrm{mm}}^{\mathrm{start}} \to (1+\chi_m)^{1/2}(\tilde{x}_m^{\mathrm{eq}})^{3/2}$, and hence find the same expression as before:

$$h^2\Omega_{\mathrm{mm,low}}^{(1)} = \frac{\mathcal{A}_{\mathrm{mm}}}{3}\,x_m^0\,. \tag{6.126}$$

For the high-frequency regime, we can, in principle, also distinguish the cases $f < f_{\mathrm{mm}}^{\mathrm{start},*}$ and $f > f_{\mathrm{mm}}^{\mathrm{start},*}$. The former can, however, never be realised since $f_{\mathrm{mm}}^{\mathrm{start},*}/f_{\mathrm{mm}}^{\mathrm{min}} = (a_0/\tilde{a}_m^{\mathrm{eq}})^{1/2}$, which means that the minimum frequency $f_{\mathrm{mm}}^{\mathrm{min}}$ is always only about two orders of magnitude smaller than $f_{\mathrm{mm}}^{\mathrm{start},*}$. For the remaining case that $f > f_{\mathrm{mm}}^{\mathrm{start},*}$, we find, upon using the expansion $x_{\mathrm{mm}}^{\mathrm{start}} = \varphi_3(\tilde{x}_m^{\mathrm{eq}}, \chi_m) \longrightarrow (1+\chi_m)^{1/3}\tilde{x}_m^{\mathrm{eq}}$ that

$$h^2\Omega_{\mathrm{mm,high}}^{(1)} = \frac{\mathcal{A}_{\mathrm{mm}}}{x_m^0}\left(1 - (1+\chi_m)^{1/3}\left(\frac{\tilde{a}_m^{\mathrm{eq}}}{a_0}\right)^{1/2}\right). \tag{6.127}$$

Note that, sensibly, this expansion breaks down at the latest when the condition for the existence of the maximum frequency $f_{\mathrm{mm}}^{\mathrm{max}}$ in Eq. (6.119) is satisfied. For sufficiently small χ_m and thus large enough string tensions, the first term on the right-hand side of Eq. (6.127) dominates,

$$h^2\Omega_{\mathrm{mm,high}}^{(1)} = \frac{\mathcal{A}_{\mathrm{mm}}}{x_m^0} \simeq 2.31\times 10^{-10}\left(\frac{C_m}{0.0387}\right)\left(\frac{\Gamma}{50}\right)\left(\frac{G\mu}{10^{-7}}\right)^2\left(\frac{10^{-10}\,\mathrm{Hz}}{f}\right). \tag{6.128}$$

With the results for the behaviour of the spectrum at low and high frequencies given in Eqs. (6.125) and (6.127) at hand, we can determine the peak frequency by requiring that both expressions yield the same value at this frequency. This yields

$$f_{\mathrm{mm}}^{\mathrm{peak}} = f_m^0\left(3 - 3\,(1+\chi_m)^{1/3}\left(\frac{\tilde{a}_m^{\mathrm{eq}}}{a_0}\right)^{1/2}\right)^{1/2} \tag{6.129}$$

and will, of course, break down if $f_{\rm mm}^{\rm peak} \sim f_{\rm mm}^{\rm min}$ or if $f_{\rm mm}^{\rm max}$ exists and thus the high-frequency part of the spectrum does not. The latter is actually the typical case for reasonably small values of the string tension. This being said, the MM contribution is normally, at frequencies around its peak, strongly subdominant compared to the RM contribution, which makes the peak position and height less relevant from a phenomenological point of view.[7]

We conclude with a short summary of the features of the MM spectrum. It always exhibits a minimum frequency below which it is zero. Above, it rises according to a power law $\propto f$. The spectrum invariably contains a switch frequency, which leaves the power-law behaviour unaffected, though. All these characteristics can be observed in both Figs. 6.6a and 6.6b. For very large string tensions, the spectrum then reaches its peak and turns into an f^{-1} power law, as one can see in Fig. 6.6a. For lower string tensions, the spectrum exhibits a maximum frequency, is cut off, and a high-frequency power-law behaviour is not present (cf. Fig. 6.6b). As mentioned previously, the difference between our analytical calculation and the numerical results can be understood as a result of an improper treatment of the reduced correlation length ξ as a constant ξ_m during matter domination. This explains, in particular, two features, best illustrated in Fig. 6.6b. Recall that the maximum frequency is associated with loops produced at $t_{\rm eq}$. At this time, we have $\xi(t_{\rm eq}) \simeq 0.30$ in comparison to $\xi_m = 0.63$. This will typically shift the analytical maximum frequency to about half the value that is found numerically. Indeed, replacing ξ_m in Eq. (6.119) with $\xi(t_{\rm eq})$ reproduces the numerically determined maximum frequency correctly. Moreover, an increase in ξ over time, as is found in the numerical solution, implies that, in contrast to the case where ξ_m is constant, the transition from small to large initial loop lengths is faster than the usual $l_* \propto t$ growth. This, in turn, means that the spectrum decreases towards low frequencies faster than $\propto f$ as predicted analytically. Having fixed ξ_m to a larger value than $\xi(t_{\rm eq})$, we also predict too little power emitted from small MM loops, thus underestimating the amplitude of the spectrum close to the maximum frequency. All these effects can be observed in Fig. 6.6b.

6.1.5 Complete fundamental spectrum

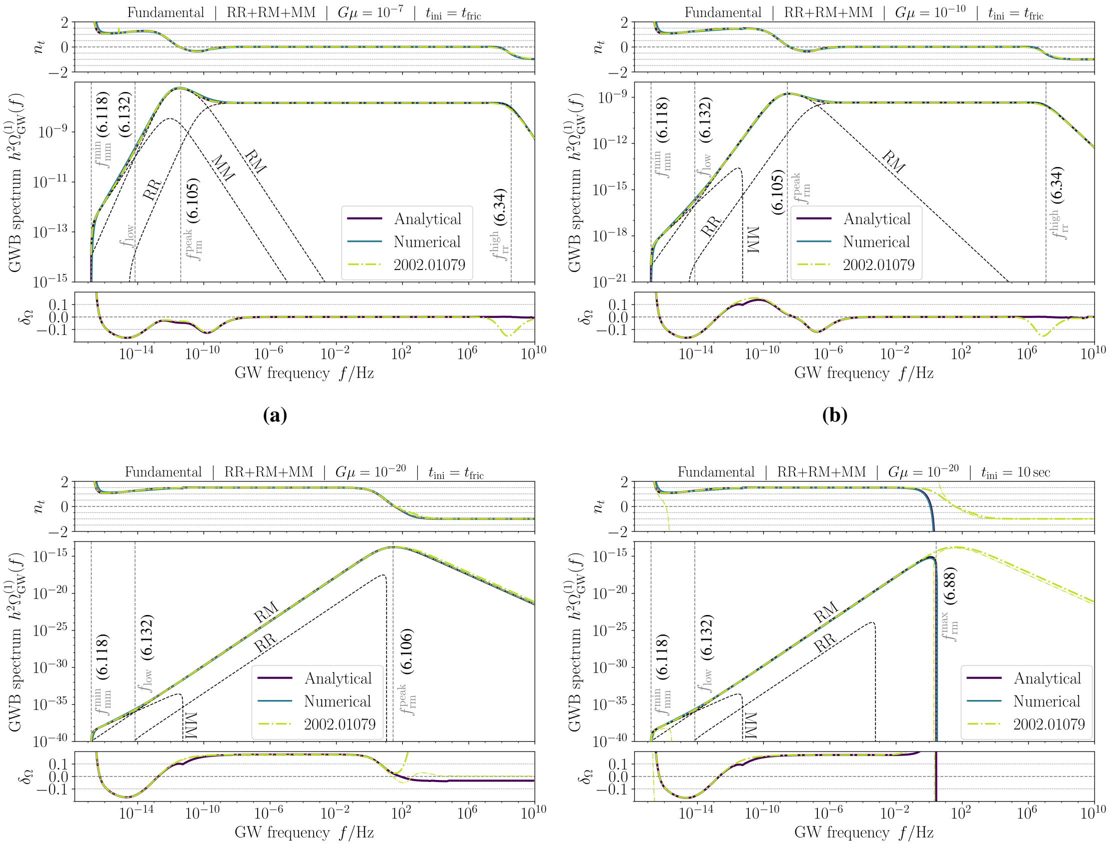
In principle, summing the fundamental RR, RM, and MM spectra will give rise to the complete fundamental spectrum. Before we turn to this sum, recall that both our analytical and numerical results up to now are based on the VOS model, assuming a constant effective number of DOFs. On the numerical side, however, we account for the full time dependence of all time-dependent quantities, such as $a(t)$, $\xi(t)$, and $v_\infty(t)$, based on our numerical solution of the Friedmann and VOS equations and the standard energy composition of our Universe in terms of radiation, matter, and dark energy (keeping the effective numbers of DOFs fixed at their present-day values). In our analytical calculation, on the other hand, we introduce an artificial splitting of the cosmic expansion into an early pure-radiation era and a late pure-matter era, which renders our analytical calculation manageable, but which at the same time also introduces a slight error. We emphasise that, if it were not for this error, our analytical and numerical results would agree with each other perfectly because they would represent the same computation.

Fortunately, there are simple changes that can be applied to find better agreement with the numerical calculation. These corrections enter the RM spectrum only. First, we make the replacement $x_{\rm rm}^{\rm end} \rightarrow x_{\rm m}^0$.[8] This, however, also removes the minimum frequency in the RM spectrum, which would then extend to arbitrarily low frequencies. There should be an overall low-frequency cutoff to the spectrum, associated with the GWs produced by the largest string loops, i.e., the loops produced today. Correspondingly, we also need to cut the RM spectrum off at this universal minimum frequency, which is nothing but the low-frequency cutoff of the MM spectrum $f_{\rm mm}^{\rm min}$. In

---

[7] For extreme values $\chi_{r(m)} \sim 1$ and, thus, extremely short loop lengths at production, the $G\mu$-dependence of C_r and C_m can become relevant and suppress the ratio $\mathcal{A}_{\rm rm}/\mathcal{A}_{\rm mm}$. This can increase the importance of the MM contribution.

[8] Note that the frequency range in which the RM case 1 of Tab. 6.1 applies is, thereby, removed completely.



(a) (b)

(c) (d)

Figure 6.7: Complete fundamental GW spectrum for four qualitatively different cases. Main and lower panels and colour codes are the same as in Fig. 6.1. For the analytical result derived in Ref. [403], we show in panels (c) and (d) two lines: A thin line for the RR+RM+MM spectrum and a thick line for the RM+MM contribution. We do this because their RR spectrum turns negative, leading to large deviations in the full spectrum, while the RM+MM contributions can still provide a good fit for parts of the frequency range. Additionally, the upper panels show the spectral index n_t (cf. Eq. (6.131)) of the numerical and the two analytical spectra. For panel (d), we remark that while the RM and MM spectra of Ref. [403] are independent of $t_{\rm ini}$, the RR spectrum was derived there assuming $t_{\rm ini} = t_{\rm Pl}/G\mu^2$. We have modified this result to make it applicable to other values of $t_{\rm ini}$ as well. Following the discussion in Refs. [403, 411], we also cut their RR and RM spectra off at the frequency $f^{\rm min}_{\rm rr}$.
Figure adapted from Ref. [4].

addition, we assume pure radiation domination before matter–radiation equality to evaluate $t_{\rm eq}$ in the prefactor $\mathcal{A}_{\rm rm}$ such that $t_{\rm eq}^{-3/2}\left(a_{\rm eq}/a_0\right)^3 = \left(2H_r^0\right)^{3/2}$ and the amplitude in Eq. (6.63) becomes

$$\mathcal{A}_{\rm rm} = \sqrt{\frac{128}{3}\pi C_r h^2 \Omega_m \left(\frac{4\Omega_r}{\Omega_m}\right)^{3/4} \left(\frac{G\mu}{\Gamma}\right)^{1/2}}\,. \tag{6.130}$$

Some benchmark spectra, including the above corrections and covering different cases, are shown in Figs. 6.7a to 6.7d. In an additional panel, we show for each of these plots also the spectral index defined as

$$n_t(f) = \frac{\mathrm{d}\ln\left(h^2\Omega(f)\right)}{\mathrm{d}\ln\left(f/\mathrm{Hz}\right)}\,. \tag{6.131}$$

All spectra have a minimum frequency in common, which is given by $f^{\rm min}_{\rm mm}$ in Eq. (6.118). The spectrum roughly rises first with an f^1 power law given by expression Eq. (6.125), which stems from the MM contribution. Afterwards, it smoothly transitions into an $f^{3/2}$ power law that is described in Eq. (6.98) and derives from the RM contribution. This behaviour is visible in all plots. The two power laws intersect at a frequency

$$f_{\text{low}} = \left(\frac{4\mathcal{A}_{\text{mm}}}{9\mathcal{A}_{\text{rm}}}\right)^2 f_m^0 \,, \tag{6.132}$$

such that the two different low-frequency power laws are assumed well below and well above this frequency. Note that f_{low} is, just as $f_{\text{mm}}^{\text{min}}$, independent of Γ and $G\mu$. The range in which the f^1 behaviour is relevant, therefore, stays the same. $f_{\text{rm}}^{\text{peak}}$ increases, however, with decreasing Γ and $G\mu$. The region in which the $f^{3/2}$ power law is important, hence, becomes larger for smaller string tensions. We can observe this in Figs. 6.7a to 6.7d. If we consider late initial times t_{ini} and low string tensions, an overall cutoff $f_{\text{rm}}^{\text{max}}$ given in Eq. (6.88) in the spectrum may arise. In this case, the low-frequency power law suddenly drops to zero towards larger frequencies. This can be observed in Fig. 6.7d. Otherwise, the spectrum will extend to larger frequencies. It peaks then at a frequency $f_{\text{rm}}^{\text{peak}}$, either given by Eq. (6.105) or by Eq. (6.106). This is shown in Figs. 6.7a to 6.7c. Without the $f_{\text{rm}}^{\text{max}}$ cutoff, but still at low string tensions, we might find a maximum frequency $f_{\text{rr}}^{\text{max}}$ (cf. Eq. (6.23)) in the RR contribution. In this case, the RR spectrum will be cut off, and the overall spectrum will be dominated by the RM spectrum. This means that, after reaching the peak of the spectrum, we will find an f^{-1} decay, which is given by Eq. (6.103).[9] We can see such a spectrum in Fig. 6.7c. At larger string tensions, the RR contribution will not be cut off, and instead, we find a plateau with height A_{rr} given in Eq. (6.29). At a frequency $f_{\text{rr}}^{\text{high}}$ given in Eq. (6.34), the plateau ends and turns into an f^{-1}-power law, whose precise form is described by Eq. (6.127). This high-frequency decay as well as the plateau are visible in Figs. 6.7a and 6.7b.

6.2 Total spectrum

So far, we were only concerned with the spectrum from the first harmonic, $\Omega_{\text{GW}}^{(1)}$. The observable result is, however, not the contribution from a single mode, but the sum of all harmonics. As pointed out in Sec. 4.3.3, higher-mode contributions have the convenient property that they may be expressed in terms of the fundamental spectrum. In the following, we want to make use of this relation and derive an explicit analytical approximation to the total spectrum based on our analytical treatment of the fundamental spectrum.

Our starting point is the sum in Eq. (4.90), which we reproduce here for convenience:[10]

$$\Omega_{\text{GW}}\left(f\right) = \frac{1}{H_{n_{\text{max}}}^{(q)}} \sum_{k=1}^{n_{\text{max}}} \frac{1}{k^q}\, \Omega_{\text{GW}}^{(1)}\left(\frac{f}{k}\right) . \tag{6.133}$$

Note that here, and in the following, we use k to denote the mode number, as the index j is already used to label the epochs associated with different effective numbers of degrees of freedom, introduced in Sec. 6.1.2. This differs from earlier chapters, where j was used to label the mode number. It appears that, even in the limit $n_{\text{max}} \to \infty$, this sum cannot be further simplified. Instead, we proceed by approximating the sum as follows: Denoting the summands by $\Phi(k)$, we can generalise the function Φ from mode number $k \in \mathbb{N}$ to $k \in \mathbb{R}_{\geq 1}$. For arbitrary $m, n_{\text{max}} \in \mathbb{N}$ and monotonically decreasing Φ, one can easily show that

$$\sum_{k=1}^{m} \Phi(k) + \int_{m+1}^{n_{\text{max}}+1} \Phi(k)\,\mathrm{d}k \leq \sum_{k=1}^{n_{\text{max}}} \Phi(k) \leq \sum_{k=1}^{m} \Phi(k) + \int_{m}^{n_{\text{max}}} \Phi(k)\,\mathrm{d}k \,. \tag{6.134}$$

---

[9] The high-frequency behaviour for case 2 is not relevant here since, if $f_{\text{rr}}^{\text{max}}$ exists, $f_{\text{rm}}^{\text{start},*}$ will as well. Case 2 will then not apply to asymptotically large frequencies.

[10] For reasons explained in Sec. 4.3.3, we ignore in this treatment, without introducing any inaccuracies, that the earliest GW emission time will be independent of the mode number, but can depend on frequency. We treat the earliest emission time t_{ini} here merely as another free model parameter (see also Sec. 5.1.2).

Choosing a larger value of m will generally improve the approximation. If Φ is monotonically increasing, the upper and lower estimates of the sum are interchanged. Hereafter, we will only write down the explicit expressions for the upper bound. The lower bound is then obtained by the trivial replacement $(n_{\max}, m) \mapsto (n_{\max}+1, m+1)$ in the integrated part. In all of our plots, we show the values between the lower and the upper bound for $m = 1$ as a deep purple band.

For the following considerations, it is important that $q > 1$. Accordingly, if $\Omega_{\mathrm{GW}}^{(1)}(f)$ decreases in f nowhere faster than f^{-1}, the function $\Phi(k) \propto \Omega_{\mathrm{GW}}^{(1)}\left(f/k\right)/k^{q}$ will be monotonically decreasing and Eq. (6.134) can be applied without any changes. As a matter of fact, the f^{-1} decay is, in the absence of a maximum frequency, the fastest decrease we find for the fundamental RR, RM, and MM spectra, occurring at asymptotically large frequencies. Even though this is no longer true in the presence of a maximum frequency, the above formula still provides a good approximation of the actual spectrum.

It should be noted that $\Phi(k)$ represents the analytical expressions derived earlier. Since these expressions may deviate from the numerical results, the band defined by the upper and lower bounds does not necessarily contain the numerically computed total spectrum. Moreover, if one is primarily interested in a single representative value of the spectrum rather than the full band, we recommend using the geometric mean of the upper and lower bounds as an estimate.

6.2.1 General features

To avoid a repetition of reasoning, let us start with a discussion of the ubiquitous features of the GW spectra. For this, consider again the sum over all modes in Eq. (6.133). We can visualise this sum as adding up $n_{\max}$ copies of the fundamental spectrum, which are, with increasing mode number k, shifted towards higher frequencies and smaller amplitudes. It is immediately clear that the minimum frequency of the total spectrum is set by the lowest mode and consequently the same as for the fundamental mode $f_{\min}^{\mathrm{tot}} = f_{\min}$. The maximum frequency is, on the other hand, determined by the largest contributing mode number $f_{\max}^{\mathrm{tot}} = n_{\max} f_{\max}$. It is important to note at this point that while the total minimum frequency is a sensible prediction, the maximum frequency of the total spectrum is artificially introduced by arbitrarily choosing a finite $n_{\max}$ and, accordingly, unphysical. Similarly, the spectrum should, in general, only be considered reliable at frequencies that satisfy $f \ll f_{\max}^{\mathrm{tot}}$.

Apart from minimum and maximum frequencies, all three fundamental spectra approach power-law behaviours in the asymptotic regions, $h^2\Omega_{\mathrm{GW}}^{(1)}(f) \longrightarrow C f^{p}$. In these cases, the sum over harmonics can be evaluated explicitly, namely as

$$h^2\Omega_{\mathrm{GW}}(f) = \frac{C}{H_{n_{\max}}^{(q)}}\left[\zeta(q+p,1) - \zeta(q+p,n_{\max}+1)\right] f^{p}\,, \tag{6.135}$$

where ζ is the Hurwitz zeta function. Note that this is not true in the case of the UHF regime, though, since the different harmonics do not depend only on f/k.

6.2.2 RR loops

We begin the discussion of the total spectrum once again with the gravitational radiation from loops decaying during radiation domination. For the fundamental mode of the spectrum, we found Eq. (6.8). To analytically carry out the integration over the mode number, we need to approximate the term containing $\varphi_2\left(x_r^{\mathrm{ini}}, \chi_r\right)$. For $x_{\mathrm{r}}^{\mathrm{ini}} \gg 1$, we can again use the expansion $\varphi_2(x_r^{\mathrm{ini}}, \chi_r) \to \sqrt{1+\chi_r}\, x_r^{\mathrm{ini}}$, and it turns out that this expression leads to a very good approximation also at intermediate frequencies. This may not come as a surprise, recalling from Sec. 6.1.1 that at those frequencies, the term depending on x_r^{eq} is strongly dominant. In the parameter regime in which φ_2 is relevant and simultaneously $x_r^{\mathrm{ini}} \ll 1$, we need the expansion $\varphi_2(x_r^{\mathrm{ini}}, \chi_r) \to$

$(1+\chi_r)\left(x_r^{\text{ini}}\right)^2 - (1+\chi_r)^2\left(x_r^{\text{ini}}\right)^4$. This is particularly important when $f_{\text{rr}}^{\text{max}}$ exists. Since the shape function deriving from the low-frequency expansion is also only relevant in this regime, we present here only the expanded spectrum, which is additionally easier to evaluate numerically. Combining all of these considerations, we find as an upper limit for the total RR spectrum the formula

$$\begin{aligned} h^2\Omega_{\text{rr}}^{m,\text{upper}}(f) = {} & \frac{1}{H_{n_{\max}}^{(q)}} \sum_{k=1}^{m} \frac{1}{k^q} h^2\Omega_{\text{rr}}^{(1)}\left(\frac{f}{k}\right) + \\ & + \frac{\mathcal{A}_{\text{rr}}}{H_{n_{\max}}^{(q)}} \left\{ S_{\text{rr},1}(k, x_r^{\text{eq}})\Big|_{k_{\text{rr}}^{\min}}^{k_{\text{rr}}^{\max}} - S_{\text{rr},2}(k, \sqrt{1+\chi_r}x_r^{\text{ini}})\Big|_{k_{\text{rr}}^{\min}}^{k_{\text{rr}}^{\text{start}}} - S_{\text{rr},3}\left(k\right)\Big|_{k_{\text{rr}}^{\text{start}}}^{k_{\text{rr}}^{\max}} \right\}, \end{aligned} \tag{6.136}$$

where we introduced the shape functions

$$S_{\text{rr},1}(k,x) = -\frac{{}_2F_1\left[\frac{3}{2}, 1-q, 2-q, -\frac{k}{x}\right]}{(q-1)k^{q-1}}, \tag{6.137}$$

$$S_{\text{rr},2}(k,x) = \begin{cases} \frac{x^3}{k^{q+2}}\left(-\frac{1}{q+2} + \frac{3}{q+4}\frac{x^2}{k^2} - \frac{6}{q+6}\frac{x^4}{k^4}\right) & \text{if } x \ll 1 \\ S_{\text{rr},1}(k,x) & \text{otherwise} \end{cases}, \tag{6.138}$$

$$S_{\text{rr},3}(k) = -\frac{S_{\text{rr}}^{(1)}(\chi_r^{-1})}{(q-1)k^{q-1}}. \tag{6.139}$$

The frequency-dependent mode numbers that appear above are given by

$$k_{\text{rr}}^{\min} = \max\left\{m, f/f_{\text{rr}}^{\max}\right\}, \tag{6.140a}$$

$$k_{\text{rr}}^{\max} = \max\left\{\min\left\{n_{\max}, f/f_{\text{rr}}^{\min}\right\}, k_{\text{rr}}^{\min}\right\}, \tag{6.140b}$$

$$k_{\text{rr}}^{\text{start}} = \min\left\{\max\left\{k_{\text{rr}}^{\min}, f/f_{\text{rr}}^{\text{start},*}\right\}, k_{\text{rr}}^{\max}\right\}. \tag{6.140c}$$

The mode numbers $k_{\text{rr}}^{\min/\max}$ derive, upon the replacement $f \to f/k$, directly from Eq. (6.26). In a similar fashion, $k_{\text{rr}}^{\text{start}}$ corresponds to $f_{\text{rr}}^{\text{start},*}$ derived in Eq. (6.20). At low frequencies ($x \lesssim 10^{-3}$), the evaluation of the hypergeometric function can become numerically difficult. An expansion around $x = 0$ reveals a leading-order term which is independent of the mode number and, therefore, cancels out. Removing it, one may use to a very good approximation

$$S_{\text{rr},1}(k,x) \to \frac{2\Gamma(2-q)}{(q-1)(2q+1)\Gamma(1-q)} \frac{x^{3/2}}{k^{1/2+q}}. \tag{6.141}$$

If the spectrum exhibits no maximum, a good value to start using the low-frequency expression in $S_{\text{rr},2}$ is below $x_r^{\text{ini}} \sim 10^{-8}$. If $f_{\text{rr}}^{\max} > 0$, it is advisable to use this expression only.

Having found a complete expression for the total RR spectrum, we continue with a discussion of its features. While the descriptions of the low-frequency, plateau, and high-frequency regions can be extracted via Eq. (6.135) directly from the fundamental spectrum, for large $n_{\max}$, an intermediate region between the plateau and high-frequency region develops. We will now derive the corresponding power law. For the region below $f_{\text{rr}}^{\text{high}}$ in Eq. (6.34), the fundamental spectrum is well described by the plateau $h^2\Omega_{\text{rr,plateau}}^{(1)} = \mathcal{A}_{\text{rr}}$ in Eq. (6.29). For $f \gg f_{\text{rr}}^{\text{high}}$, the spectrum approaches a high-frequency power law of the form $h^2\Omega_{\text{rr,high}}^{(1)}(f) = \mathcal{A}_{\text{rr}} f_{\text{rr}}^{\text{high}}/f$, which we found in Eq. (6.30). Correspondingly, we can write for the total spectrum (for $1 < q < 2$)

$$\begin{aligned} h^2\Omega_{\mathrm{rr}}(f) =& \frac{1}{H^{(q)}_{n_{\mathrm{max}}}} \sum_{k=1}^{n_{\mathrm{max}}} \frac{1}{k^q} h^2\Omega^{(1)}_{\mathrm{rr}}\left(\frac{f}{k}\right) \simeq \qquad (6.142)\\ \simeq& \frac{1}{H^{(q)}_{n_{\mathrm{max}}}} \left(\sum_{k=1}^{\lfloor k_{\mathrm{high}}(f) \rfloor} \frac{\mathcal{A}_{\mathrm{rr}}}{k^{q-1}} \frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f} + \sum_{k=\lceil k_{\mathrm{high}}(f) \rceil}^{n_{\mathrm{max}}} \frac{\mathcal{A}_{\mathrm{rr}}}{k^q} \right) = \\ =& \frac{\mathcal{A}_{\mathrm{rr}}}{H^{(q)}_{n_{\mathrm{max}}}} \left(\frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f} \zeta\left(q-1, k\right) \Big|_{k=\lfloor k_{\mathrm{high}}(f) \rfloor + 1}^{k=1} + \zeta\left(q, k\right) \Big|_{k=n_{\mathrm{max}}+1}^{k=\lceil k_{\mathrm{high}}(f) \rceil} \right) \qquad (6.143) \end{aligned}$$

with $k_{\mathrm{high}}(f) = f/f^{\mathrm{high}}_{\mathrm{rr}}$. The spectrum can only develop an intermediate power law if there is a region in which $f^{\mathrm{high}}_{\mathrm{rr}} \ll f \ll n_{\mathrm{max}} f^{\mathrm{high}}_{\mathrm{rr}}$. Assuming that this is realised and picking a frequency within this region, we have $k_{\mathrm{high}}(f) \gg 1$. Expanding around this value and approximating $\lfloor k_{\mathrm{high}}(f) \rfloor \simeq \lceil k_{\mathrm{high}}(f) \rceil \simeq k_{\mathrm{high}}(f)$ and $k_{\mathrm{high}}(f) + 1 \simeq k_{\mathrm{high}}(f)$, we find

$$\begin{aligned} h^2\Omega_{\mathrm{rr}}(f) =& \frac{\mathcal{A}_{\mathrm{rr}}}{H^{(q)}_{n_{\mathrm{max}}}} \left(\frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f} \zeta\left(q-1, 1\right) + \left(\frac{1}{2-q} + \frac{1}{q-1}\right) \left(\frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f}\right)^{q-1} - \zeta\left(q, n_{\mathrm{max}}+1\right) \right) \sim \\ \sim& \frac{\mathcal{A}_{\mathrm{rr}}}{H^{(q)}_{n_{\mathrm{max}}}} \frac{1}{(2-q)(q-1)} \left(\frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f}\right)^{q-1} . \qquad (6.144) \end{aligned}$$

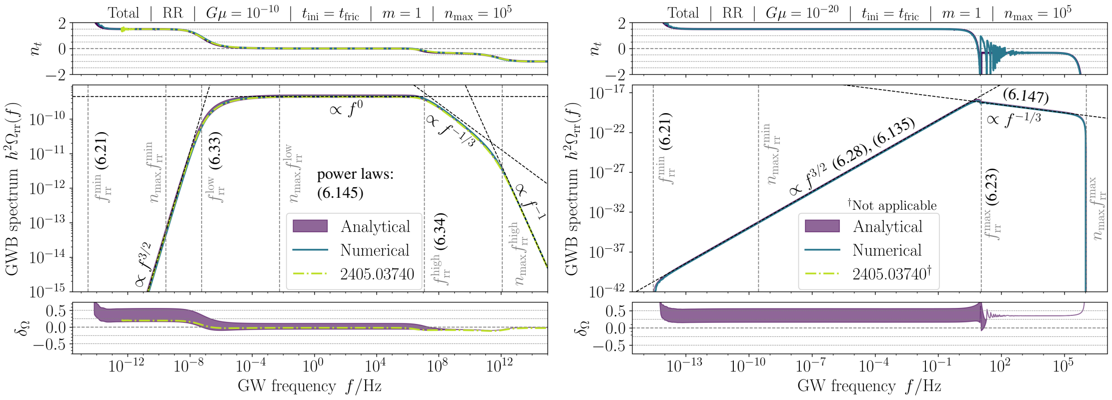
Here, we used that the first term is smaller than the second term by a factor of $(f^{\mathrm{high}}_{\mathrm{rr}}/f)^{2-q}$. The last term is suppressed since $\zeta\left(q, n_{\mathrm{max}}+1\right) \simeq n_{\mathrm{max}}^{1-q}/(q-1)$ and, by assumption, $k_{\mathrm{high}}(f) \ll n_{\mathrm{max}}$. Note, however, that this approximation breaks down as $q \to 2$. In summary, the following power laws can be found in the spectrum in the absence of a maximum frequency:

$$h^2\Omega_{\mathrm{rr}}(f) \simeq \frac{\mathcal{A}_{\mathrm{rr}}}{H^{(q)}_{n_{\mathrm{max}}}} \begin{cases} \zeta\left(q+\frac{3}{2}, k\right) \big|_{k=n_{\mathrm{max}}+1}^{k=1} \left(\frac{f}{f^{\mathrm{eq}}_r}\right)^{3/2} & \text{if } f^{\mathrm{min}}_{\mathrm{rr}} \ll f \ll f^{\mathrm{low}}_{\mathrm{rr}} \\ H^{(q)}_{n_{\mathrm{max}}} & \text{if } f^{\mathrm{low}}_{\mathrm{rr}} \ll f \ll f^{\mathrm{high}}_{\mathrm{rr}} \\ \frac{1}{(2-q)(q-1)} \left(\frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f}\right)^{q-1} & \text{if } f^{\mathrm{high}}_{\mathrm{rr}} \ll f \ll n_{\mathrm{max}} f^{\mathrm{high}}_{\mathrm{rr}} \\ \zeta\left(q-1, k\right) \big|_{k=n_{\mathrm{max}}+1}^{k=1} \left(\frac{f^{\mathrm{high}}_{\mathrm{rr}}}{f}\right) & \text{if } n_{\mathrm{max}} f^{\mathrm{high}}_{\mathrm{rr}} \ll f \end{cases} . \qquad (6.145)$$

A corresponding benchmark spectrum is shown in Fig. 6.8a, which displays our numerical and analytical results alongside the analytical spectrum derived in Ref. [411] on the basis of Ref. [403]. As evident from this plot, all three results are in good agreement. Observe that, as in the case of the fundamental spectrum (cf. Fig. 6.1), our analytical result slightly overestimates the actual spectrum in the low-frequency tail. Due to the approximation we applied to φ_2, our spectrum has the same deviation from the numerical results as Ref. [411] does, when transitioning from the plateau to the high-frequency region. Note that we cut off the spectrum of Ref. [411] at very low frequencies close to $f^{\mathrm{min}}_{\mathrm{rr}}$ as it becomes numerically ill-behaved there.

In the presence of a maximum frequency, the behaviour is different. The spectrum consists only of the low-frequency power law in Eq. (6.28) and has a sharp cutoff at the maximum frequency. We can, therefore, approximate the fundamental spectrum by $h^2\Omega^{(1)}_{\mathrm{rr}} \simeq \mathcal{A}_{\mathrm{rr}} \left(f/f^{\mathrm{eq}}_r\right)^{3/2} \Theta\left(f^{\mathrm{max}}_{\mathrm{rr}} - f\right)$. Hence, we obtain for the total spectrum

$$\begin{aligned} h^2\Omega_{\mathrm{rr}}(f) \simeq& \frac{1}{H^{(q)}_{n_{\mathrm{max}}}} \sum_{k=\lceil f/f^{\mathrm{max}}_{\mathrm{rr}} \rceil}^{n_{\mathrm{max}}} \frac{\mathcal{A}_{\mathrm{rr}}}{k^{q-3/2}} \left(\frac{f}{f^{\mathrm{eq}}_r}\right)^{3/2} = \qquad (6.146)\\ =& \frac{\mathcal{A}_{\mathrm{rr}}}{H^{(q)}_{n_{\mathrm{max}}}} \left(\frac{f}{f^{\mathrm{eq}}_r}\right)^{3/2} \zeta\left(q+3/2, k\right) \Big|_{k=\lceil f/f^{\mathrm{max}}_{\mathrm{rr}} \rceil}^{k=n_{\mathrm{max}}+1} . \end{aligned}$$



(a) **(b)**

Figure 6.8: Total GW spectrum for RR loops and the two qualitatively different cases. Main panels: Result from numerical calculations with a fixed effective number of DOFs (teal, see description at the end of the introduction to this chapter), analytical result derived in this chapter with $m = 1$ and $n_{\rm max} = 10^5$ (deep purple band, showing values between lower and upper bound given in Eq. (6.134)), and analytical result derived in Ref. [411] (light green, dash-dotted). Grey dashed lines show characteristic frequencies of the spectrum, and black dashed lines show different power-law behaviours of the spectrum. Upper panels: Spectral index n_t (cf. Eq. (6.131)) of the three spectra. Lower panels: Relative deviation δ_Ω (cf. Eq. (6.18)) of the two analytical results from the numerical result.
Figure reproduced from Ref. [4].

For $f_{\rm rr}^{\rm max} \ll f \ll n_{\rm max} f_{\rm rr}^{\rm max}$, we can furthermore expand the ζ function, which results in

$$h^2\Omega_{\rm rr}(f) \simeq \frac{\mathcal{A}_{\rm rr}}{\left(q + \frac{1}{2}\right) H_{n_{\rm max}}^{(q)}} \left(\frac{f}{f_r^{\rm eq}}\right)^{3/2} \left(\frac{f}{f_{\rm rr}^{\rm max}}\right)^{-q-1/2} \propto f^{1-q} \,. \tag{6.147}$$

Note that this overestimates the amplitude of the power law by a factor of ~ 3 since, at $f_{\rm rr}^{\rm max}$, the spectrum already deviates from the low-frequency power law.

A benchmark scenario exhibiting the described features is depicted in Fig. 6.8b, in which we compare our analytical and numerical results. Our formulae can reproduce the shape of the numerical spectrum well, but slightly overestimate the amplitude, as was already the case for the fundamental spectrum. Additionally, we also see that the low-frequency and high-frequency power laws describe the spectrum adequately. For the latter power law, we included the mentioned correction factor of 3. Observe that the numerical spectrum exhibits an oscillatory feature most evident in the spectral index n_t. This is of completely analogous origin as the oscillatory feature of low-scale strings discussed in Ch. 5. While our analytical expression correctly reproduces the overall power-law behaviour, it cannot resolve this feature, as is evident from the oscillations in δ_Ω. The result of Ref. [411] cannot be applied in this range of parameter space as it is based on the expression for the fundamental spectrum in Ref. [403], which becomes negative in this case.

UHF regime: For the UHF regime, we can again directly draw conclusions from the results we obtained for the fundamental mode. We simply need to adapt the initial time $t_{\rm ini}$ as described around Eq. (4.94). We have to be careful, though, when calculating the prefactor for the high-frequency power law. From the f^{-2} dependence of Eq. (6.37), one might expect that the different harmonics decay like k^{-q-2}. In reality, they decrease like k^{-q-1} since one factor f^{-1} derives from the lower integration boundary, which is independent of k. Keeping this in mind and applying Eq. (6.135), one finds that the f^{-1} high-frequency power law transitions at the frequency given in Eq. (6.38) to an UHF power law of the form

$$h^2\Omega_{\rm rr,high}^{\delta} \simeq \frac{3\mathcal{A}_{\rm rr}}{2} \left(\frac{G\mu}{\chi_r}\right)^{1/2} \frac{h_r^0}{t_{\rm Pl} f^2} \,. \tag{6.148}$$

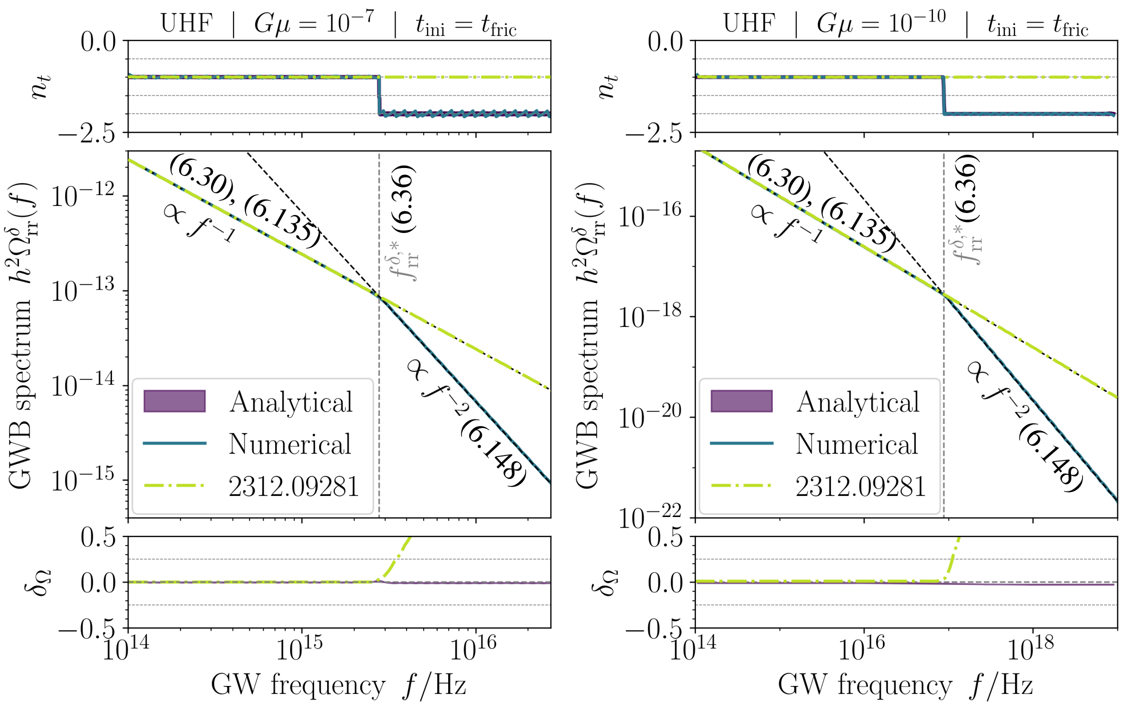
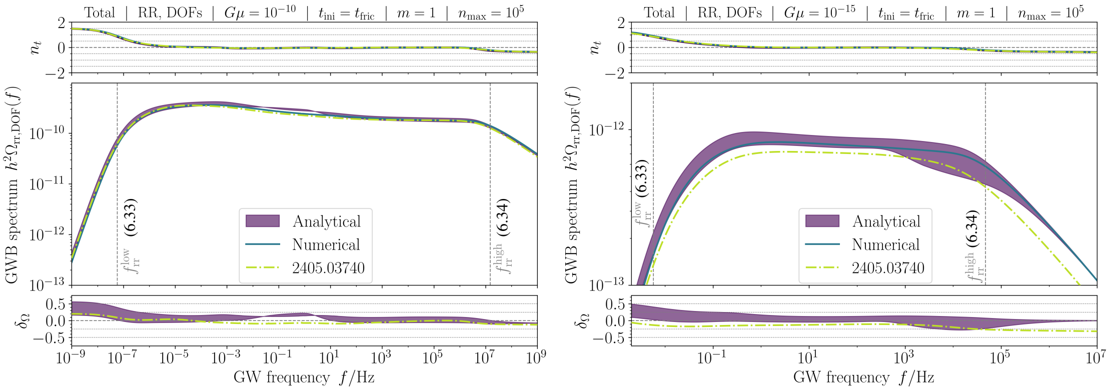


Figure 6.9: UHF regime of the total GW spectrum for the RR case, string tensions of $G\mu = 10^{-7}$ (left) and $G\mu = 10^{-10}$ (right), and initial times $t_{\rm ini} = t_{\rm fric}$. Main panels: Our exact numerical result based on the VOS model accounting for the full time dependence of all relevant quantities and for a fixed effective number of DOFs (teal), our analytical result derived in this chapter for $m = 1$ and $n_{\rm max} = 10^3$ (deep purple band, showing values between lower and upper bound given in Eq. (6.134)) and the numerical result following from Ref. [415]. The grey dashed line shows the frequency at which the UHF effects become relevant. The power laws describing the spectrum below and above this frequency are shown as black dashed lines. Upper panels: Spectral index (cf. Eq. (6.131)) of the three spectra. Lower panels: Relative deviation (cf. Eq. (6.18)) of the analytical result from the numerical result. The visible deviations are due to numerical noise.
Figure reproduced from Ref. [4].

In Fig. 6.9, we can see the transition from the f^{-1} to the f^{-2} power law for two benchmark scenarios. We compare our analytical result to our numerical result as well as to the result of Ref. [415]. For the initial time, we chose, as before, the friction cutoff $t_{\rm ini} = t_{\rm fric}$. As one can see, our analytical and numerical results agree perfectly; visible deviations are only due to numerical inaccuracies. The derived power laws shown in the figure give also an excellent fit to the spectrum. When comparing this result to the fundamental spectrum, one may observe that, in contrast to other features, the transition from high frequencies to ultrahigh frequencies is not washed out by the mode summation. This is because of the mode number independence of the cutoff that we found. This is in contrast to the result of Ref. [415] which discusses the same effect in Appendix A. In this article, the authors allow for GW emission from a k–mode, as soon as a loop of appropriate size had a chance to be formed for the first time. They neglect, however, the fact that loops at later times shrink again to smaller size. In our treatment, we ensured that the relevant condition on the GW wavelength is always fulfilled. In particular, the correction to the initial time found in Ref. [415] is independent of the frequency but only depends on the mode number. For the depicted benchmark scenarios, this would only become relevant for mode numbers $k \gtrsim 10^6$. Since we consider in the plot only mode numbers as large as 10^3, the correction has no effect here and generally underestimates the actual correction due to the UHF effect.



(a) **(b)**

Figure 6.10: Total GW spectra for RR loops, accounting for changes in the effective number of DOFs. The benchmark points show the spectra for tensions of $G\mu = 10^{-10}$ (left panel) and $G\mu = 10^{-15}$ (right panel), and an initial time $t_{\rm ini} = t_{\rm fric}$. The panels and colour code are the same as in Fig. 6.8. This time, our numerical comparison spectrum based on the VOS model is accounting for the full time dependence of all relevant quantities and for a time-dependent effective number of DOFs (see description at the end of the introduction to this chapter).
Figure reproduced from Ref. [4].

6.2.3 Varying degrees of freedom

The previous derivation of the total RR spectrum still assumed a constant effective number of relativistic DOFs. To include variations in this number during radiation domination, we need to adapt the discussion of the previous section and refer to the slightly more complex fundamental spectrum in Sec. 6.1.2. Directly working with this fundamental spectrum and the results we derived in the previous section can be problematic, since the additional splitting of the radiation dominated era will introduce multiple regions in which our expansions of φ_2 can become invalid. We are saved by the fact that the fundamental RR spectrum is, in absence of a maximum frequency, very well approximated if we use the replacement $\varphi_2(x, \chi) \to (1 + \chi_r)^{1/2}\, x$. This gives rise to a fundamental spectrum that can be integrated without any further approximations. In presence of a maximum frequency in the RR spectrum, there is no plateau region and the RR contribution will be strongly subdominant to the RM contribution. Therefore, if a maximum in the RR spectrum occurs, we can comfortably refer to the discussion in the previous section. Let us now turn to the case in which there is no such maximum frequency and investigate the effect of changing DOFs. To use the above approximation consistently, we need to account for the fact that the switch, minimum, and maximum frequencies will be affected as well. Let us summarise these. With the notation used in Sec. 6.1.2, we have as a result of our approximation, indicated by hats, for case A

$$\hat{x}_{A,i}^{\rm start} = \max\left\{\chi_r^{-1}, \sqrt{1+\chi_r}\, x_i^{(i)}\right\}, \qquad \hat{x}_{A,i}^{\rm end} = x_i^{(i+1)}, \tag{6.149}$$

which tells us for the minimum and switch frequency that

$$\hat{f}_{A,i}^{\min} = \frac{f_i^{(i+1)}}{\chi_r}, \qquad \hat{f}_{A,i}^{\rm start,*} = \frac{f_i^{(i)}}{\chi_r\sqrt{1+\chi_r}}. \tag{6.150}$$

The spectrum will no longer contain a maximum frequency but is only non-vanishing if

$$\frac{a_{(i+1)}}{a_{(i)}} > \sqrt{1+\chi_r}\,. \tag{6.151}$$

The simplifications are much stronger in case B. Introducing the frequencies

$$\hat{f}_{B,ij}^{\rm start} = \min\left\{f_i^{(i)}, f_i^{(j)}/\sqrt{1+\chi_r}\right\} \quad \text{and} \quad \hat{f}_{B,ij}^{\rm end} = \max\left\{f_i^{(i+1)}, f_i^{(j+1)}/\sqrt{1+\chi_r}\right\}, \tag{6.152}$$

we can write

$$\hat{x}^{\text{start}}_{B,ij} = f/\hat{f}^{\text{start}}_{B,ij}\,, \qquad \hat{x}^{\text{end}}_{B,ij} = f/\hat{f}^{\text{end}}_{B,ij}, \tag{6.153}$$

and we can also immediately see that the spectrum is only non-vanishing if

$$\hat{f}^{\text{start}}_{B,ij} > \hat{f}^{\text{end}}_{B,ij}\,. \tag{6.154}$$

From this structure, it is clear that neither switch frequencies nor minimum or maximum frequencies exist in the approximated spectra. To make sure that the spectrum still does not extend to arbitrarily low frequencies, we cut it off at $f^{\text{min}}_{\text{rr}}$.

Applying Eq. (6.134) to the spectra that arise by using these simplifications in combination with the fundamental spectrum Eq. (6.48), we find

$$\begin{aligned}
h^2\Omega^{m,\text{upper}}_{\text{rr,DOF}} = {} & \frac{1}{H^{(q)}_{n_{\text{max}}}}\sum_{k=1}^{m}\frac{1}{k^q}h^2\Omega^{(1)}_{\text{rr,DOF}}\left(\frac{f}{k}\right) + \\
& + \frac{\mathcal{A}_{\text{rr}}}{H^{(q)}_{n_{\text{max}}}}\sum_{i=0}^{N-1}\Theta^A_i\mathcal{G}_i\left\{S_{\text{rr},1}\left(k, x_i^{(i+1)}\right)\Big|_m^{k^{\text{max}}_{A,i}} - S_{\text{rr},1}\left(k, \sqrt{1+\chi_r}x_i^{(i)}\right)\Big|_m^{k^{\text{start}}_{A,i}} - S_{\text{rr},3}\left(k\right)\Big|_{k^{\text{start}}_{A,i}}^{k^{\text{max}}_{A,i}}\right\} + \\
& + \frac{\mathcal{A}_{\text{rr}}}{H^{(q)}_{n_{\text{max}}}}\sum_{i=1}^{N-1}\sum_{j=0}^{i-1}\Theta^B_{ij}\mathcal{G}_i^{1/4}\mathcal{G}_j^{3/4}\left\{S_{\text{rr},1}\left(k, \hat{x}^{\text{end}}_{B,ij}\right)\Big|_m^{k^{\text{max}}_{B,ij}} - S_{\text{rr},1}\left(k, \hat{x}^{\text{start}}_{B,ij}\right)\Big|_m^{k^{\text{max}}_{B,ij}}\right\}.
\end{aligned} \tag{6.155}$$

To evaluate the above expression, we need the following frequency-dependent mode numbers

$$k^{\text{max}}_{A,i} = \max\left\{\min\left\{n_{\text{max}}, f/f^{\text{min}}_{A,i}\right\}, m\right\}\,, \tag{6.156a}$$

$$k^{\text{start}}_{A,i} = \min\left\{\max\left\{m, f/\hat{f}^{\text{start},*}_{A,i}\right\}, k^{\text{max}}_{A,i}\right\}\,, \tag{6.156b}$$

$$k^{\text{max}}_{B,ij} = \max\left\{\min\left\{n_{\text{max}}, f/f^{\text{min}}_{\text{rr}}\right\}, m\right\}\,, \tag{6.156c}$$

as well as the Heaviside functions

$$\Theta^A_i = \Theta\left(\frac{a_{(i+1)}}{a_{(i)}} - \sqrt{1+\chi_r}\right) \qquad \text{and} \qquad \Theta^B_{ij} = \Theta\left(\hat{f}^{\text{start}}_{B,ij} - \hat{f}^{\text{end}}_{B,ij}\right)\,. \tag{6.157}$$

In spite of the lack of a closed analytical expression for the time evolution of the effective number of DOFs, the above formulae provide a great approximation of our numerical spectrum; see Fig. 6.10. We observe that the deviations only begin to exceed the $25\,\%$ level in the low-frequency tail of the spectrum. This is unobservable and, hence, unproblematic since, in this part of the spectrum, the RR contribution is strongly subdominant to the RM contribution.

There is a mentionable difference between our result and the one in Ref. [411]. The latter accounts for changes in the DOFs by fitting the analytical GWB to the numerically calculated one. For the result here, we used no information about the GW spectrum to choose our values of $a_{(i)}$ and $\mathcal{G}_i$, since they directly derive from the evolution of g_s and g_ρ. Furthermore, Ref. [411] does not distinguish between effects during loop production time and GW emission time, i.e., between $\mathcal{G}_i$ and $\mathcal{G}_j$.

6.2.4 RM and MM loops

We continue with a discussion of all loops decaying during matter domination. While on a technical level, we still distinguish between RM and MM loops, due to the corrections described in Sec. 6.1.5, it is no longer sensible to consider the resulting GW spectra separately.

For the case of RM loops, we found that the fundamental spectrum is described by Eq. (6.62). In order to carry out the integrals in the estimate Eq. (6.134) analytically, it is necessary to approximate

$\varphi_3(\tilde{x}_m^{\rm ini},\chi_r)$. For large frequencies $\tilde{x}_m^{\rm ini} \gg 1$ we can approximate $\varphi_3(\tilde{x}_m^{\rm ini},\chi_r) \to (1+\chi_r)^{1/3}\tilde{x}_m^{\rm ini}$, while for low frequencies $\tilde{x}_m^{\rm ini} \ll 1$, we have $\varphi_3(\tilde{x}_m^{\rm ini},\chi_r) \to \sqrt{1+\chi_r}(\tilde{x}_m^{\rm ini})^{3/2}$. With this, we are able to find the following expressions for the spectrum:

$$
\begin{aligned}
h^2\Omega_{\rm rm}^{m,\text{upper}}(f) = {} & \frac{1}{H_{n_{\max}}^{(q)}}\sum_{k=1}^{m}\frac{1}{k^q}h^2\Omega_{\rm rm}^{(1)}\left(\frac{f}{k}\right) + \\
& + \frac{\mathcal{A}_{\rm rm}}{H_{n_{\max}}^{(q)}}\left\{S_{{\rm rm},1}\left(k,(1+\chi_r)^{1/3}\tilde{x}_m^{\rm ini}\right)\Big|_{k_{\rm rm}^{\min}}^{k_{\rm rm}^{\rm start}} + S_{{\rm rm},2}\left(k,x_m^{\rm eq}\right)\Big|_{k_{\rm rm}^{\rm start}}^{k_{\rm rm}^{\max}} - S_{{\rm rm},2}\left(k,x_m^0\right)\Big|_{k_{\rm rm}^{\min}}^{k_{\rm rm}^{\max}}\right\}\,,
\end{aligned}
\tag{6.158}
$$

where we introduced the shape functions

$$
S_{{\rm rm},1}(k,x) = \begin{cases} \frac{3}{2(3+2q)k^{3/2+q}\sqrt{x_m^0}}x^3 + F_{\rm RM}(k) & \text{if } x \ll 1 \\ S_{{\rm rm},2}(k,x) & \text{else} \end{cases}\,, \tag{6.159}
$$

$$
S_{{\rm rm},2}(k,x) = \frac{2k^{2-q} + 2(2q-1)(k+x)\left(-x\right)^{1-q}\,{}_2F_1\left[\frac{1}{2},q-1;\frac{3}{2};\frac{k+x}{x}\right]}{\sqrt{x_m^0(k+x)}}\,, \tag{6.160}
$$

as well as

$$
k_{\rm rm}^{\min} = \max\left\{m, f/f_{\rm rm}^{\max}\right\}\,, \tag{6.161a}
$$

$$
k_{\rm rm}^{\max} = \max\left\{\min\left\{n_{\max}, f/f_{\rm mm,min}\right\}, k_{\rm rm}^{\min}\right\}\,, \tag{6.161b}
$$

$$
k_{\rm rm}^{\rm start} = \min\left\{\max\left\{k_{\rm rm}^{\min}, f/f_{\rm rm}^{{\rm start},*}\right\}, k_{\rm rm}^{\max}\right\}\,. \tag{6.161c}
$$

We also introduced the function $F_{\rm RM}(k)$, which only depends on the mode number but not on a frequency ratio x; it need not be specified further as it cancels out. At very high (low) frequencies, the evaluation of the hypergeometric functions can become numerically unstable. Therefore, it is beneficial to use series expansions. While working with the full hypergeometric function can be cumbersome, instead we can directly expand $S_{\rm rm}^{(1)}(x)$ for large (small) x and integrate over the mode number afterwards. In this way, we obtain

$$
S_{{\rm rm},2}\left(k,x\right) \overset{x\to\infty}{\longrightarrow} \frac{3k^{2-q}}{(2-q)\sqrt{x_m^0 x}}\,, \qquad S_{{\rm rm},2}\left(k,x\right) \overset{x\to 0}{\longrightarrow} \frac{3k^{-1/2-q}x^2}{(2+4q)\sqrt{x_m^0}} + F_{\rm RM}(k)\,. \tag{6.162}
$$

Note that the above low-frequency expansion should only be applied simultaneously with the low-frequency expansion in Eq. (6.159). Only in this case, the function $F_{\rm RM}(k)$ is guaranteed to cancel out and can therefore be neglected. If there is no maximum frequency, a good choice to change from the low-frequency expansion to the full spectrum is at $x_m^0 \sim 10^{-5}$. For the high-frequency expansion, we found $x_m^0 \sim 10^{10}$ to be a good value. If a maximum frequency exists, the low-frequency expansion can be applied everywhere.

Most power laws arising in the spectrum derive again via Eq. (6.135) directly from the power laws present in the fundamental spectrum. However, when transitioning from the peak to the high-frequency region, a new power law can develop. Below the peak frequency given in Eq. (6.105) or Eq. (6.106), the spectrum rises approximately like $A(f/f_m^0)^{3/2}$ with the amplitude A given in Eq. (6.98). Above, it decays like $B\,f_m^0/f$, and the amplitude B can be found in Eq. (6.102) or Eq. (6.103). The total spectrum can then be approximated by

$$
h^2\Omega_{\rm rm}\left(f\right) \simeq \frac{1}{H_{n_{\max}}^{(q)}}\sum_{k=1}^{\lfloor k_{\rm rm}^{\rm peak}\rfloor}\frac{B}{k^{q-1}}\frac{f_m^0}{f} + \frac{1}{H_{n_{\max}}^{(q)}}\sum_{k=\lceil k_{\rm rm}^{\rm peak}\rceil}^{n_{\max}}\frac{A}{k^{q+3/2}}\left(\frac{f}{f_m^0}\right)^{3/2}\,, \tag{6.163}
$$

where we introduced $k_{\rm rm}^{\rm peak}(f) = f/f_{\rm rm}^{\rm peak}$. The above sums can be carried out explicitly in terms of Hurwitz zeta functions. Since we are interested in the regime where $f_{\rm rm}^{\rm peak} \ll f \ll n_{\rm max} f_{\rm rm}^{\rm peak}$ and thus $1 \ll k_{\rm rm}^{\rm peak}(f) \ll n_{\rm max}$, we can expand the arising terms and find in total

$$h^2\Omega_{\rm rm}(f) \simeq \frac{f^{1-q}}{H_{n_{\rm max}}^{(q)}} \left(\frac{B}{2-q} \frac{f_m^0}{\left(f_{\rm rm}^{\rm peak}\right)^{2-q}} + \frac{A}{q+1/2} \frac{\left(f_{\rm rm}^{\rm peak}\right)^{q+1/2}}{\left(f_m^0\right)^{3/2}} \right) . \tag{6.164}$$

Within the reasonable parameter range, there is no clear hierarchy between the two remaining terms and we have to take both into account. Due to our approximation, we have overestimated the amplitude of the intermediate power law for the same reason as in the RR case. We can, however, correct for this quite well with a factor of ∼ 4.

The above reasoning can, of course, only be sensibly applied if there is no maximum frequency in the spectrum. If there is a maximum frequency, the fundamental spectrum will be roughly described by the low-frequency power $A(f/f_m^0)^{3/2}$ below the maximum frequency, and it vanishes above the cutoff. This means we can simply take the above expression for the total spectrum, set the high-frequency amplitude $B = 0$, and replace $f_{\rm rm}^{\rm peak}$ by $f_{\rm rm}^{\rm max}$. Let us use the explicit expression $A = 3\mathcal{A}_{\rm rm}/4$ to find

$$h^2\Omega_{\rm rm}(f) \simeq \frac{f^{1-q}}{H_{n_{\rm max}}^{(q)}} \frac{3\mathcal{A}_{\rm rm}}{4q+2} \frac{\left(f_{\rm rm}^{\rm max}\right)^{q+1/2}}{\left(f_m^0\right)^{3/2}} . \tag{6.165}$$

Turning to MM loops, the fundamental spectrum is described by Eq. (6.110). To carry out the integrals in Eq. (6.134), we need to approximate $\varphi_3\left(\tilde{x}_m^{\rm eq}, \chi_m\right)$. Namely, for $\tilde{x}_m^{\rm eq} \gg 1$, we can approximate $\varphi_3\left(\tilde{x}_m^{\rm eq}\right) \to \left(1+\chi_m\right)^{1/3} \tilde{x}_m^{\rm eq}$ and for low frequencies $\varphi_3\left(\tilde{x}_m^{\rm eq}, \chi_m\right) \to \sqrt{1+\chi_m}\left(\tilde{x}_m^{\rm eq}\right)^{3/2}$. This allows us to obtain

$$\begin{aligned} h^2\Omega_{\rm mm}^{m,\text{upper}}(f) = & \frac{1}{H_{n_{\rm max}}^{(q)}} \sum_{k=1}^{m} \frac{1}{k^q} h^2\Omega_{\rm mm}^{(1)}\left(\frac{f}{k}\right) + \qquad (6.166) \\ & + \frac{\mathcal{A}_{\rm mm}}{H_{n_{\rm max}}^{(q)}} \left\{ S_{\rm mm,1}\left(k, x_m^0\right) \Big|_{k_{\rm mm}^{\rm min}}^{k_{\rm mm}^{\rm max}} - S_{\rm mm,3}\left(k, \left(1+\chi_m\right)^{1/3} \tilde{x}_m^{\rm eq}\right) \Big|_{k_{\rm mm}^{\rm min}}^{k_{\rm mm}^{\rm start}} - S_{\rm mm,2}(k) \Big|_{k_{\rm mm}^{\rm start}}^{k_{\rm mm}^{\rm max}} \right\} , \end{aligned}$$

where the shape functions read

$$\begin{aligned} S_{\rm mm,1}(k,x) = & \left(\frac{x}{x_m^0}\right)^2 \frac{k^{1-q}\left({}_2F_1\left(1, 1-q; 2-q; -\frac{k}{x}\right) - 2\left(\frac{k}{x}\right)^2 \ln\left(1+\frac{x}{k}\right) - \left(\frac{k}{x}-1\right)^2\right)}{3-q} \overset{x \ll 1}{\simeq} \\ \simeq & \left(\frac{x}{x_m^0}\right)^2 \left(-\frac{x}{3qk^q} + \frac{x^2}{2(1+q)k^{1+q}}\right) + F_1(x) + F_{\rm MM}(k) , \qquad (6.167) \\ S_{\rm mm,2}(k) = & \frac{k^{3-q} S_{\rm mm}^{(1)}\left(\chi_m^{-1}\right)}{3-q} \overset{\chi_m \gg 1}{\simeq} \frac{k^{3-q}}{3-q} \left(\frac{1}{x_m^0}\right)^2 \left(\frac{\chi_m^{-3}}{3} - \frac{\chi_m^{-4}}{2} - 1\right) = \\ = & \frac{k^{3-q}}{3-q} \left(\frac{1}{x_m^0}\right)^2 \left(\frac{\chi_m^{-3}}{3} - \frac{\chi_m^{-4}}{2}\right) + F_{\rm MM}(k) , \qquad (6.168) \end{aligned}$$

$$\begin{aligned}\tilde{S}_{\mathrm{mm},3}(k,x) =& \frac{k^{3-q}}{(q-3)\,(x_m^0)^2}\Bigg\{1 - {}_2F_1\left(1, 2-\frac{2q}{3}; 3-\frac{2q}{3}; -\left(\frac{k}{x}\right)^{3/2}\right) + 2\ln\left(1+\left(\frac{x}{k}\right)^{3/2}\right) - \\ & - \frac{2}{2q-3}\left(\frac{x}{k}\right)^{3/2}\left(3{}_2F_1\left(1, \frac{2q}{3}-1; \frac{2q}{3}; -\left(\frac{x}{k}\right)^{3/2}\right) + q - 3\right)\Bigg\} \overset{x\ll 1}{\simeq} \\ & \simeq -\frac{2x^{9/2}}{3(3+2q)k^{3/2+q}(x_m^0)^2} + F_2(x) + F_{\mathrm{MM}}(k), \end{aligned} \tag{6.169}$$

$$S_{\mathrm{mm},3}(k,x) = \begin{cases} \tilde{S}_{\mathrm{mm},3}(k,x) & \text{if } x \ll 1 \\ S_{\mathrm{mm},1}(k,x) & \text{if } x \gg 1 \end{cases} \tag{6.170}$$

The different terms are evaluated at the mode numbers

$$k_{\mathrm{mm}}^{\mathrm{min}} = \max\left\{m, f/f_{\mathrm{mm}}^{\mathrm{max}}\right\}, \tag{6.171a}$$

$$k_{\mathrm{mm}}^{\mathrm{max}} = \max\left\{\min\left\{n_{\mathrm{max}}, f/f_{\mathrm{mm}}^{\mathrm{min}}\right\}, k_{\mathrm{mm}}^{\mathrm{min}}\right\}, \tag{6.171b}$$

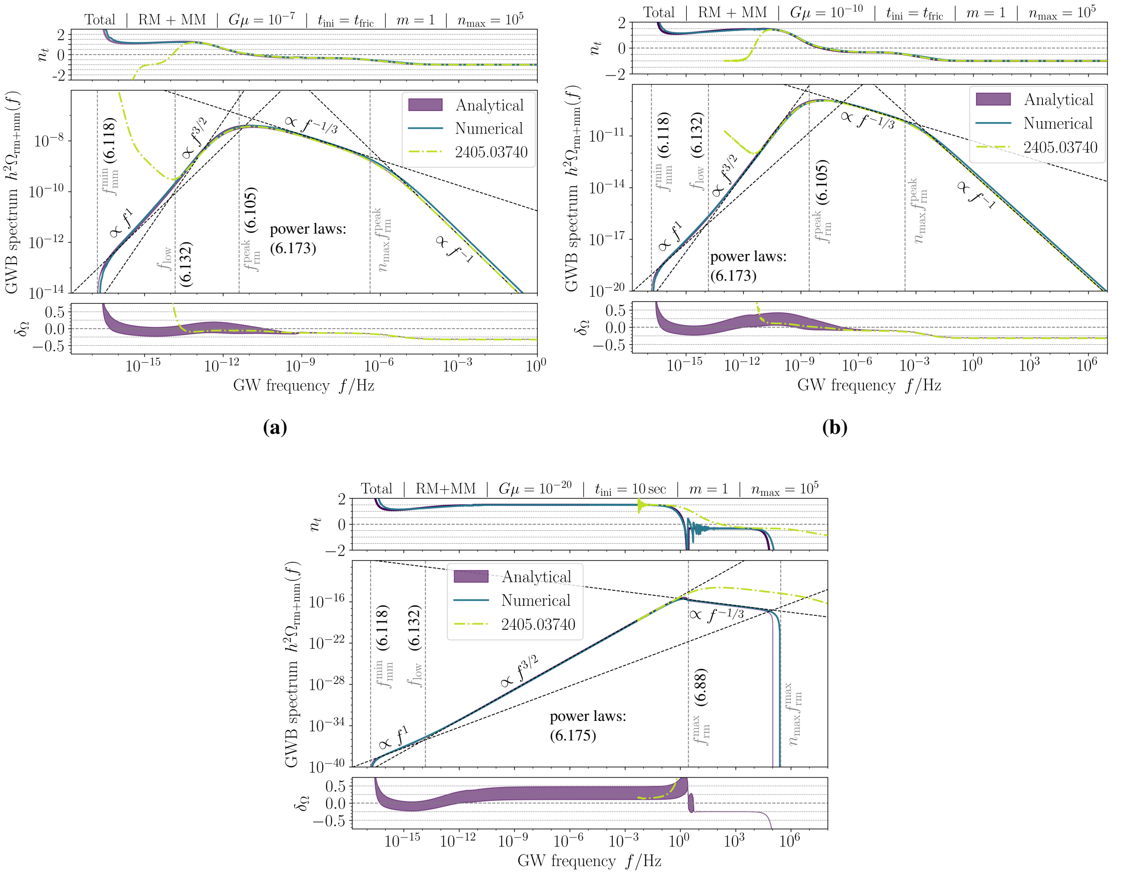
$$k_{\mathrm{mm}}^{\mathrm{start}} = \min\left\{\max\left\{k_{\mathrm{mm}}^{\mathrm{min}}, f/f_{\mathrm{mm}}^{\mathrm{start},*}\right\}, k_{\mathrm{mm}}^{\mathrm{max}}\right\}. \tag{6.171c}$$

A few comments are in order. First, the functions $F_1(x), F_2(x)$ are irrelevant since they cancel out immediately. Furthermore, we find it useful to distinguish four frequency regimes. At very low frequencies (we used $x_m^0 < 10^{-2}$), we use in Eq. (6.166) the low-frequency expression of Eq. (6.170) together with the expanded versions of all the shape functions. In this case, they all contain at leading order a term $F_{\mathrm{MM}}(k)$, which cancels out in the complete spectrum. For larger but still low frequencies (we applied this to $10^{-2} < x_m^0 < 10^3$), we still used the low-frequency expression of Eq. (6.170) but with the whole form of the shape functions, except for $S_{\mathrm{mm},2}$, for which we expanded in $\chi_m \gg 1$. Since this will be true at all frequencies, we use this expression for $S_{\mathrm{mm},2}$ everywhere. Note that, this time, we cannot neglect $F_{\mathrm{MM}}(k)$ since we do not expand the other shape functions and no cancellation will occur: We have to use the form of $S_{\mathrm{mm},2}$ given in the first line of Eq. (6.168). At higher frequencies (we used $10^3 < x_m^0 < 10^6$), we switch to the high-frequency form of Eq. (6.170), using the complete expressions for the shape functions and at very high frequencies ($x_m^0 > 10^6$), we expanded $S_{\mathrm{mm},1}(k,x)$ to avoid problems with the numerical evaluation:

$$S_{\mathrm{mm},1}(k,x) \overset{x\gg 1}{\simeq} \frac{k^{2-q}}{(x_m^0)^2}\left(\frac{x}{2-q} - \frac{2k}{(3-q)^2} - \frac{2k}{3-q}\ln\left(\frac{x}{k}\right) - \frac{3k^2}{(4-q)x}\right). \tag{6.172}$$

Furthermore, note that in cases in which no maximum frequency exists, not only a low-frequency or a high-frequency regime exists, but also an intermediate regime. It turns out, though, that in this case, the entire spectrum is rather well described only using the high-frequency expression given in Eq. (6.170). Since the MM spectrum is in all physically relevant cases, except at very low frequencies, subleading to the RM spectrum, the only relevant power law occurs in the very low-frequency regime. This power law derives with Eq. (6.135) directly from the low-frequency expression for the fundamental spectrum given in Eq. (6.125).

The resulting spectra are shown in Figs. 6.11a to 6.11c and cover the qualitatively different cases. In Fig. 6.11a, neither the RM nor the MM spectrum exhibits a maximum frequency, while for the parameter values in Fig. 6.11b, the MM spectrum does. Since this maximum frequency affects the MM spectrum only at rather large frequencies where it is strongly subleading to the RM contribution, the combined RM+MM spectra look qualitatively the same. For these cases, the



(a)

(b)

(c)

Figure 6.11: Three qualitatively different total GW spectra for RM and MM loops. The panels and colour code are the same as in Fig. 6.8. We cut off the spectra from Ref. [411] given by green dash-dotted lines at low frequencies due to large numerical fluctuations. Note that the steep rise of these spectra towards low frequencies is also a numerical artefact. For panel (c), note also that the RM and MM spectra provided in Ref. [411] are independent of $t_{\rm ini}$.
Figure adapted from Ref. [4].

spectrum can, in summary, be approximated by power laws of the form

$$h^2\Omega_{\rm rm+mm}(f) \simeq \frac{1}{H^{(q)}_{n_{\rm max}}} \begin{cases} \frac{\mathcal{A}_{\rm mm}}{3}\zeta\left(q+1,k\right)\Big|_{k=n_{\rm max}+1}^{k=1} \frac{f}{f_m^0} & \text{if } f_{\rm mm}^{\rm min} \ll f \ll f_{\rm low}, \\ \frac{3\mathcal{A}_{\rm rm}}{4}\zeta\left(q+3/2,k\right)\Big|_{k=n_{\rm max}+1}^{k=1} \left(\frac{f}{f_m^0}\right)^{3/2} & \text{if } f_{\rm low} \ll f \ll f_{\rm rm}^{\rm peak}, \\ \left(\frac{B}{2-q}\frac{f_m^0}{\left(f_{\rm rm}^{\rm peak}\right)^{2-q}} + \frac{A}{q+1/2}\frac{\left(f_{\rm rm}^{\rm peak}\right)^{q+1/2}}{\left(f_m^0\right)^{3/2}}\right)\frac{f^{1-q}}{4} & \text{if } f_{\rm rm}^{\rm peak} \ll f \ll n_{\rm max} f_{\rm rm}^{\rm peak}, \\ B\zeta\left(q-1,n\right)\Big|_{k=n_{\rm max}+1}^{k=1} \frac{f_m^0}{f} & \text{if } n_{\rm max} f_{\rm rm}^{\rm peak} \ll f, \end{cases} \tag{6.173}$$

Here, $A = 3\mathcal{A}_{\rm rm}/4$ and

$$B = 3\mathcal{A}_{\rm rm} \begin{cases} \left(\frac{a_0}{a_{\rm eq}}\right)^{1/4} - 1 \\ \frac{1}{(1+\chi_r)^{1/6}}\left(\frac{a_0}{\tilde{a}_m^{\rm ini}}\right)^{1/4} - 1 \end{cases} \tag{6.174}$$

with the case distinction given in Eqs. (6.102) and (6.103). These power laws are depicted in the mentioned figures and, as can be seen, match the numerical spectrum well. Apart from this, it is clearly visible that at large frequencies, the numerically computed spectrum is, in these cases, well approximated by our analytical calculations as well as by the result found in Ref. [411]. Furthermore, we can see that the expressions of Ref. [411] contain numerical artefacts: They increase towards lower frequencies and become later numerically unstable. We removed the latter part for the purpose of plotting.

A qualitatively different case occurs if not only the MM but also the RM spectrum exhibits a maximum frequency as given in Eq. (6.88). This case is depicted in Fig. 6.11c and the spectrum can be approximated by power laws of the form

$$h^2\Omega_{\mathrm{rm+mm}}(f) \simeq \frac{1}{H^{(q)}_{n_{\max}}} \begin{cases} \frac{\mathcal{A}_{\mathrm{mm}}}{3}\zeta\left(q+1,k\right)\Big|_{k=n_{\max}+1}^{k=1}\frac{f}{f^0_m} & \text{if } f^{\min}_{\mathrm{mm}} \ll f \ll f_{\mathrm{low}}, \\ \frac{3\mathcal{A}_{\mathrm{rm}}}{4}\zeta\left(q+3/2,k\right)\Big|_{k=n_{\max}+1}^{k=1}\left(\frac{f}{f^0_m}\right)^{3/2} & \text{if } f_{\mathrm{low}} \ll f \ll f^{\max}_{\mathrm{rm}}, \\ \frac{3\mathcal{A}_{\mathrm{rm}}}{4q+2}\frac{(f^{\max}_{\mathrm{rm}})^{q+1/2}}{(f^0_m)^{3/2}}f^{1-q} & \text{if } f^{\max}_{\mathrm{rm}} \ll f \ll n_{\max}f^{\max}_{\mathrm{rm}} \end{cases} . \tag{6.175}$$

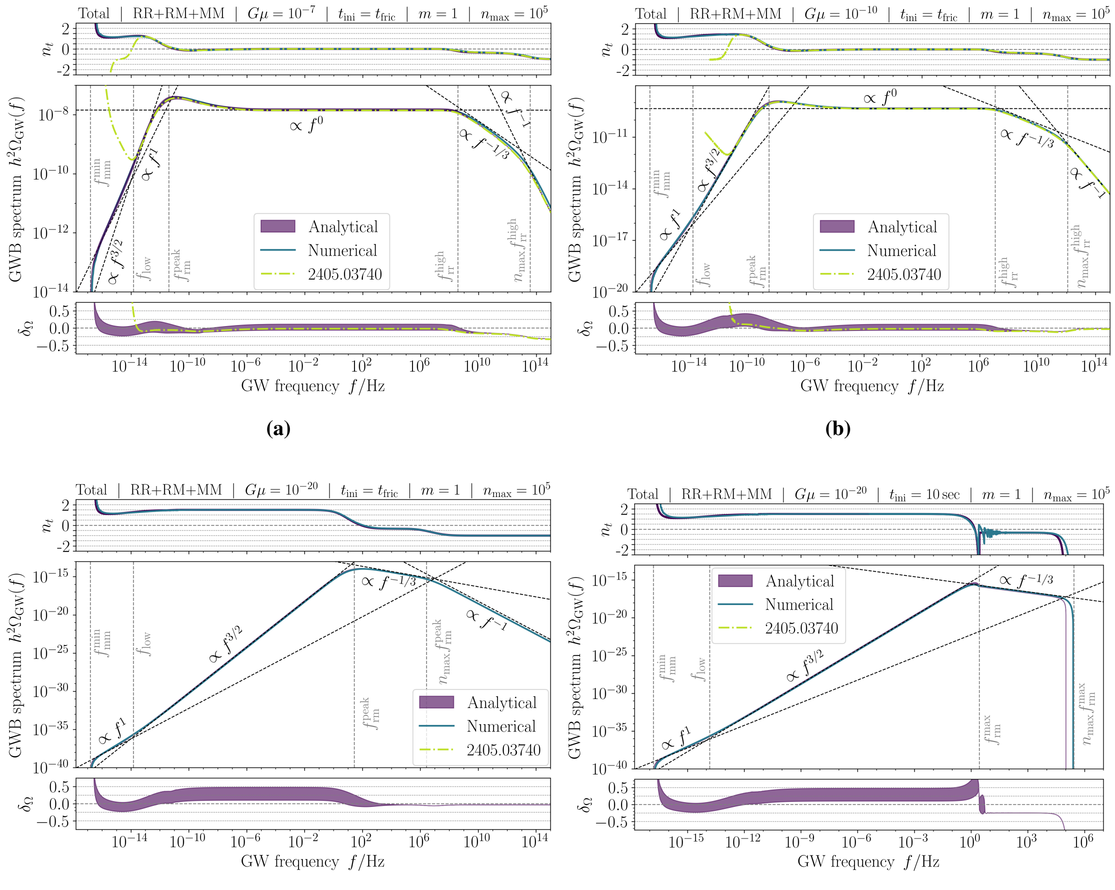
From Fig. 6.11c, we can clearly see that the different power laws capture the result of the numerical computation formidably. Similarly, our full analytical result agrees with the numerical one well, except for the frequency range close to $n_{\max}f^{\max}_{\mathrm{rm}}$. This is not unexpected, since $f^{\max}_{\mathrm{rm}}$ is determined from $\varphi_3\left(\tilde{x}^{\mathrm{ini}}_m, \chi_r\right) = x^0_m$. For our total spectrum, we used, however, the low-frequency expansion of φ_3 and we need to apply it in regions where this approximation starts to break down. Fortunately, as previously discussed, for frequencies close to $n_{\max}f^{\max}_{\mathrm{rm}}$, the spectrum becomes unphysical anyway since $n_{\max}$ is arbitrarily chosen and in principle, we should consider $n_{\max} \to \infty$ as long as the finite string width effects discussed in Secs. 4.3.3, 6.1.1, and 6.2.2 are taken into consideration. More problematically, for very small values in the denominator of $f^{\max}_{\mathrm{rm}}$,[11] the exact position of this maximum frequency is extremely sensitive to the exact parameter choice. Deviations due to our approximation can become large in this case. This is, of course, only the case for a very small parameter region and the power-law expressions give still excellent agreement with the numerical spectrum. Note that the result of Ref. [411] is more or less incapable of capturing the maximum frequency at all, which is inherited from the fact that it derives from a fundamental spectrum that does not exhibit a maximum frequency.

6.2.5 Complete spectrum

We finally turn to a presentation of the complete spectra that can be obtained by combining the above results. A comprehensive summary of these can be found in Sec. 6.4. In Fig. 6.12, we present the four qualitatively different cases for a fixed effective number of relativistic DOFs, each covered by one benchmark scenario. In Fig. 6.12a, none of the three contributions has a maximum frequency, whereas in Fig. 6.12b, the MM contribution exhibits one. Since the MM spectrum is subleading to the two other contributions, there is no visible qualitative difference, though. In Fig. 6.12c, both the MM and the RR spectrum have a maximum frequency, which implies that there is no plateau region and the RM spectrum is, apart from very low frequencies, the dominant one. Finally, in Fig. 6.12d, all three contributions have a maximum frequency. Therefore, this is a GWB spectrum from low-scale strings, as discussed in Ch. 5, and we can clearly identify the characteristic oscillatory feature in the spectral index n_t of the numerical spectrum. Our analytical description does not capture the oscillatory feature, but it does reproduce the general shape of the spectrum. This is visible as subtle oscillations in δ_Ω. For Figs. 6.12a and 6.12b, in which the spectrum contains a plateau region, variations in the effective number of relativistic DOFs during radiation domination will be important. In Fig. 6.13, we show the complete spectra taking this effect into account. We want to emphasise that our analytical expressions, even taking merely $m = 1$, provide a good approximation to the fully numerically computed spectra (except upon closely approaching the minimum or maximum frequencies). Moreover, we are able to analytically

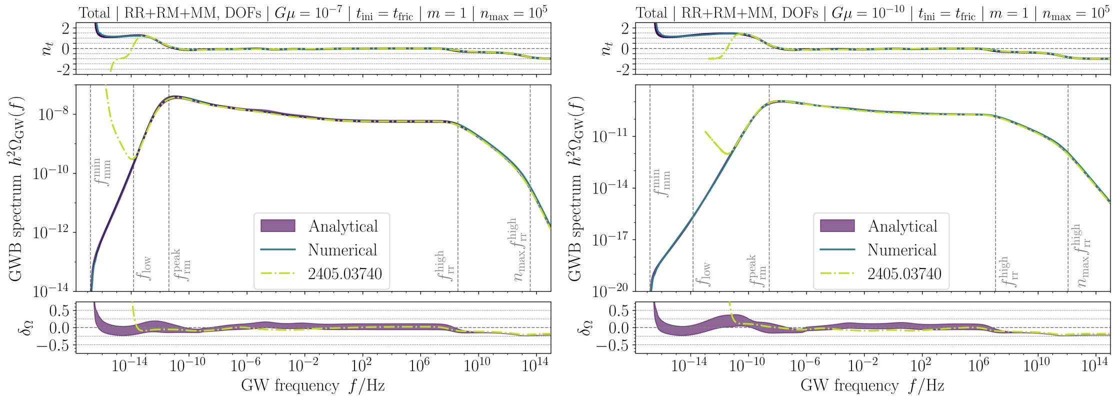
---

[11] In the language of Ch. 5, this corresponds to the parameter range very close to T_{cut}.



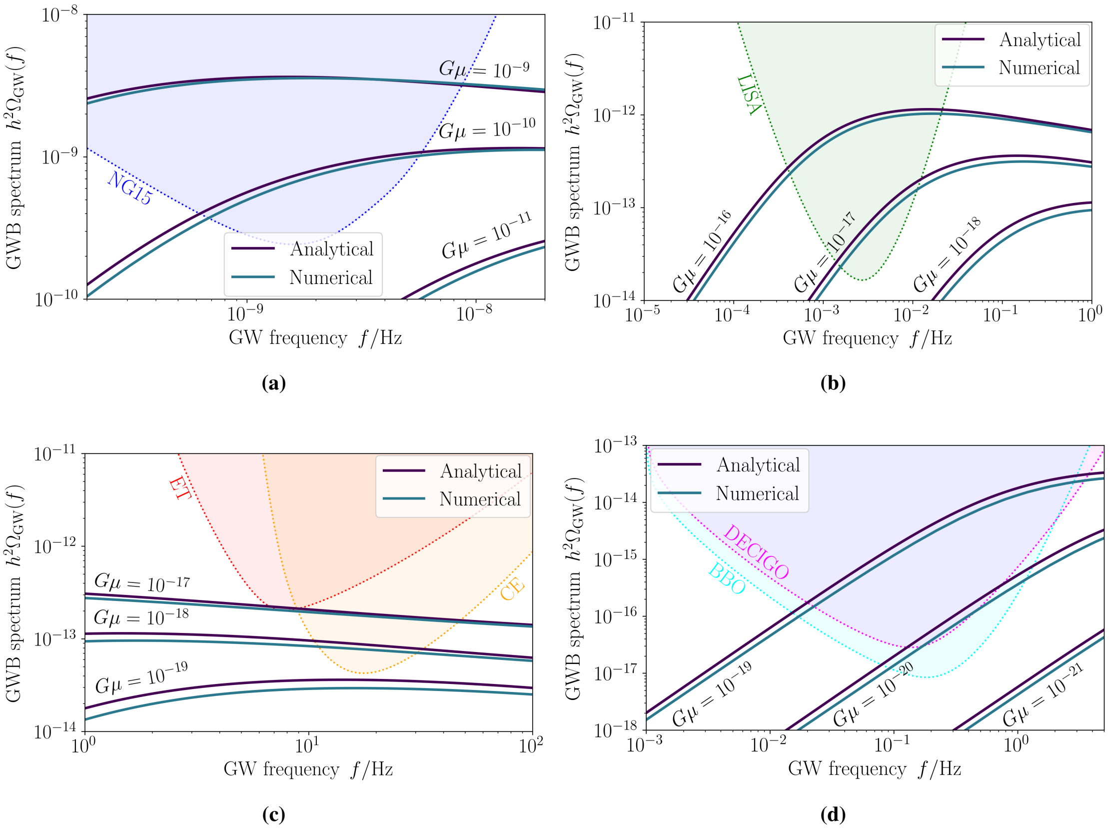
(a) **(b)** **(c)** **(d)**

Figure 6.12: Total GW spectrum with all loop contributions for four qualitatively different cases. Main panels: Our exact numerical result based on the VOS model accounting for the full time dependence of all relevant quantities and for a fixed effective number of DOFs (teal, see description at the end of the introduction to this chapter), our analytical result with $m = 1$ and $n_{\max} = 10^5$ (deep purple band, showing values between lower and upper bound given in Eq. (6.134)), and the analytical result derived in Ref. [411] (light green, dash-dotted). We cut off the latter spectra at low frequencies due to large numerical fluctuations in Figs. 6.12a and 6.12b. Note that the steep rise of these spectra towards low frequencies is also a numerical artefact. Grey dashed lines show characteristic frequencies, and black dashed lines indicate different power-law behaviours. Upper panels: Spectral index (cf. Eq. (6.131)) of the three spectra. Lower panels: Relative deviation (cf. Eq. (6.18)) of the two analytical results from the numerical result.
Figure adapted from Ref. [4].



(a) **(b)**

Figure 6.13: Same as in Fig. 6.12 but variations in the effective number of relativistic DOFs are taken into account.
Figure adapted from Ref. [4].



(a) **(b)**

(c) **(d)**

Figure 6.14: Plots showing the sensitivity curves of NANOGrav for their 15-year data (NG15) in panel 6.14a, LISA in panel 6.14b, Einstein Telescope (ET) and Cosmic Explorer (CE) in panel 6.14c, as well as Big Bang Observer (BBO) and DECIGO in panel 6.14d. In all plots, we show the GWB spectra computed for three different string tensions $G\mu$ such that the lowest spectrum is just out of reach of the respective experiments. We fixed $t_{\rm ini} = t_{\rm fric}$. For each spectrum, we present our numerical spectrum (teal) and the spectrum calculated analytically in our formalism (deep purple line obtained as the geometric mean of lower and upper bound given in Eq. (6.134)) accounting for variations in the effective number of DOFs. Sensitivity curves were taken from Ref. [433] (NG15), and Ref. [434] (all other).
Figure adapted from Ref. [4].

describe for the first time the spectrum for the low-scale case, i.e., for low string tensions and late initial times as given, e.g., in Fig. 6.12d.

The provided expressions for the GWB yield an excellent substitute for the otherwise time-consuming computation necessary to numerically calculate the spectra. This makes our results very useful for data analyses in current and future experiments. In Fig. 6.14, we illustrate this by plotting different spectra (accounting for changes in DOFs) on top of the power-law integrated sensitivity curves for PTAs, ground-based, and space-borne interferometers. Instead of displaying the band spanned by the lower and upper analytical estimates of the spectrum, we show their geometric mean, which is more likely to be used in actual data analyses.

6.3 Discussion

In this chapter, we presented a detailed derivation of analytical formulae for the stochastic GWB produced by a network of stable, local cosmic strings. Our treatment properly describes features of both the spectra from the fundamental mode and the total spectra from all modes. This holds equally true for features that could not be accounted for by expressions previously found in the literature. In particular, this includes maximum and minimum frequencies above and below which the spectrum strictly vanishes. Furthermore, in contrast to earlier work, our formulae allow to analytically compute the GWB from strings with very low tensions for which qualitatively new spectral shapes arise. On top of that, we also provided expressions that analytically approximate the suppression of the spectrum at high frequencies due to an increase in the effective number of DOFs with the temperature in the early Universe. We also quantitatively discussed how the spectra are affected at ultrahigh frequencies that have wavelengths short enough to resolve the finite width of the strings. In difference to previous studies, we found that the finite width starts to affect all mode contributions above a certain critical frequency and is, thus, mode-number-independent.

Apart from providing a detailed understanding of different features of the GW spectra, our expressions for the total spectrum provide an easy way to compute the GWB at all relevant frequencies without having to sum over a huge number of modes explicitly. This reduces computing times significantly. Our results will, therefore, be useful for GW searches with PTAs, ground-based interferometers and space-borne interferometers.

While our studies are restricted to the case of cosmic strings produced in the spontaneous breaking of a local symmetry, current modelling describes the GWB spectrum from stable colour-flux tubes and from cosmic superstrings in certain frequency ranges in essentially the same way. This makes our analytical results also relevant for those. In the future, the procedure laid out in this chapter could be applied in a similar fashion to study the more complicated stochastic GWB from a network of metastable cosmic strings [33, 210].

Comparing the analytical results obtained here to the recent numerical spectra in Fig. 2 of Ref. [364] by eye, we can already tell that our analytical results will provide an accurate approximation up to corrections of at most order 1. In future work [435], we are going to compare our simple analytical results to the spectra of Ref. [364] more rigorously. In some preliminary work, a χ^2 minimisation between our spectra, rescaled with a fudge factor $\mathcal{C}_f$, and the spectra of Ref. [364], weighted by the NANOGrav sensitivity curves from Ref. [433], was performed over the corresponding frequency band. For values of $G\mu \sim 10^{-11} \dots 10^{-9}$, this analysis yields fudge factors $\mathcal{C}_f \sim 0.7 \dots 1.0$, which indicates quantitatively good agreement between our analytical model and the most recent simulations within these parameter and frequency ranges.

6.4 Summary of formulae

Cosmological parameters:

$$h^2\Omega_m = 0.143 \qquad h^2\Omega_r = 4.18\times 10^{-5} \qquad H_r^0 = 2.10\times 10^{-20}\mathrm{Hz} \qquad H_m^0 = 1.23\times 10^{-18}\mathrm{Hz}$$

We split radiation domination into $N = 4$ intervals, in which we assume the effective number of degrees of freedom and thus $\mathcal{G}(t)$ to be constant with $a_{(0)} = a_{\rm ini}$, $a_{(4)} = a_{\rm eq}$, and

$$\begin{array}{cccc} a_{(1)}/a_0 = 1.6\times 10^{-15} & a_{(2)}/a_0 = 3.4\times 10^{-12} & a_{(3)}/a_0 = 2.0\times 10^{-9} & a_{(4)}/a_0 = 2.9\times 10^{-4} \\ \mathcal{G}_0 = 0.39 & \mathcal{G}_1 = 0.43 & \mathcal{G}_2 = 0.83 & \mathcal{G}_3 = 1 \end{array}$$

If $a_{(i)} < a_{\rm ini}$ for any i in the list, then $a_{(i)}$ and $\mathcal{G}_{i-1}$ need to be removed from below $a_{\rm ini}$ until it is the smallest scale factor. After removing r list entries, the above formulae can be applied by relabelling $i \to i-r$ for $i > r$. Correspondingly, radiation domination is then only split into $N = 4-r$ intervals.

Standard values (Secs. 4.2, 4.3.2, 6.1.1, 6.1.3)**:**

$$\begin{array}{cccc} \Gamma = 50 & \xi_r = 0.27 & \xi_m = 0.63 & \alpha = 0.37 \\ v_r = 0.66 & v_m = 0.58 & \tilde{c} = 0.23 & \mathcal{F} = 0.1 \end{array}$$

Prefactors, amplitudes, and other expressions (Secs. 4.3.2, 6.1.1, 6.1.4, 6.1.5)**:**

$$C_\beta = \mathcal{F}\frac{\tilde{c}}{\sqrt{2}\alpha}\frac{v_\beta}{\xi_\beta^4}(\alpha\xi_\beta + \Gamma G\mu)^{3(1-\beta)} \qquad \chi_\beta = \frac{\alpha\xi_\beta}{\Gamma G\mu} \qquad \text{with } \beta = \begin{cases} 1/2 & \text{during RD} \\ 2/3 & \text{during MD} \end{cases}$$

$$\mathcal{A}_{\rm rr} = \frac{128\pi}{9}\,C_r\,h^2\Omega_r\left(\frac{G\mu}{\Gamma}\right)^{1/2} \qquad \mathcal{A}_{\rm rm} = \sqrt{\frac{128}{3}}\pi C_r h^2\Omega_m\left(\frac{4\Omega_r}{\Omega_m}\right)^{3/4}\left(\frac{G\mu}{\Gamma}\right)^{1/2}$$

$$\mathcal{A}_{\rm mm} = 18\pi\,C_m\,h^2\Omega_m\,G\mu$$

Frequencies (Secs. 6.1.1, 6.1.2, 6.1.3, 6.1.4, 6.1.5, 6.2.3)**:**

$$f_r^i = \frac{4H_r^0}{\Gamma G\mu}\frac{a_0}{a(t_i)} \qquad f_k^{(i)} = \frac{4\mathcal{G}_k^{1/2}H_r^0}{\Gamma G\mu}\frac{a_0}{a_{(i)}} \qquad f_m^i = \frac{3H_m^0}{\Gamma G\mu}\left(\frac{a_0}{a(t_i)}\right)^{1/2} \qquad \tilde{f}_m^i = \frac{3H_m^0}{\Gamma G\mu}\left(\frac{a_0}{\tilde{a}_m(t_i)}\right)^{1/2}$$

with $\tilde{a}_m(t)/a_0 = \left(3/2H_m^0 t\right)^{2/3}$.

$$f_{\rm rr}^{\rm min} = \frac{f_r^{\rm eq}}{\chi_r} \qquad f_{\rm rr}^{\rm start,*} = \frac{f_r^{\rm ini}}{\chi_r} \qquad f_{\rm rr}^{\rm max} = \frac{f_r^{\rm eq}}{(1+\chi_r)\left(\frac{a_{\rm ini}}{a_{\rm eq}}\right)^2 - 1}$$

$$f_{A,i}^{\rm min} = \frac{f_i^{(i+1)}}{\chi_r} \qquad \hat{f}_{A,i}^{\rm start,*} = \frac{f_i^{(i)}}{\chi_r\sqrt{1+\chi_r}}$$

$$\hat{f}_{B,ij}^{\rm start} = \min\left\{f_i^{(i)}, f_i^{(j)}/\sqrt{1+\chi_r}\right\} \qquad \hat{f}_{B,ij}^{\rm end} = \max\left\{f_i^{(i+1)}, f_i^{(j+1)}/\sqrt{1+\chi_r}\right\}$$

$$f_{\rm rm}^{\rm max} = \frac{f_m^0}{\left(\frac{\tilde{a}_{\rm m}^{\rm ini}}{a_0}\right)^{3/2}(1+\chi_r) - 1} \qquad f_{\rm rm}^{\rm start,*} = \frac{f_m^{\rm eq}}{\left(\frac{\tilde{a}_{\rm m}^{\rm ini}}{a_{\rm eq}}\right)^{3/2}(1+\chi_r) - 1}$$

$$f_{\rm mm}^{\rm min} = \frac{f_m^0}{\chi_m} \qquad f_{\rm mm}^{\rm start,*} = \frac{\tilde{f}_m^{\rm eq}}{\chi_m} \qquad f_{\rm mm}^{\rm max} = \frac{f_m^0}{(1+\chi_m)\left(\frac{\tilde{a}_m^{\rm eq}}{a_0}\right)^{3/2} - 1}$$

Mode numbers (Secs. 6.2.2, 6.2.3, 6.2.4)**:**

$$\begin{aligned}
k_{\mathrm{rr}}^{\min} &= \max\left\{m, f/f_{\mathrm{rr}}^{\max}\right\} & k_{\mathrm{rr}}^{\max} &= \max\left\{\min\left\{n_{\max}, f/f_{\mathrm{rr}}^{\min}\right\}, k_{\mathrm{rr}}^{\min}\right\} \\
k_{\mathrm{rr}}^{\mathrm{start}} &= \min\left\{\max\left\{k_{\mathrm{rr}}^{\min}, f/f_{\mathrm{rr}}^{\mathrm{start},*}\right\}, k_{\mathrm{rr}}^{\max}\right\} & k_{A,i}^{\max} &= \max\left\{\min\left\{n_{\max}, f/f_{A,i}^{\min}\right\}, m\right\} \\
k_{A,i}^{\mathrm{start}} &= \min\left\{\max\left\{m, f/\hat{f}_{A,i}^{\mathrm{start},*}\right\}, k_{A,i}^{\max}\right\} & k_{B,ij}^{\max} &= \max\left\{\min\left\{n_{\max}, f/f_{\mathrm{rr}}^{\min}\right\}, m\right\} \\
k_{\mathrm{rm}}^{\min} &= \max\left\{m, f/f_{\mathrm{rm}}^{\max}\right\} & k_{\mathrm{rm}}^{\max} &= \max\left\{\min\left\{n_{\max}, f/f_{\mathrm{mm}}^{\min}\right\}, k_{\mathrm{rm}}^{\min}\right\} \\
k_{\mathrm{rm}}^{\mathrm{start}} &= \min\left\{\max\left\{k_{\mathrm{rm}}^{\min}, f/f_{\mathrm{rm}}^{\mathrm{start},*}\right\}, k_{\mathrm{rm}}^{\max}\right\} & k_{\mathrm{mm}}^{\min} &= \max\left\{m, f/f_{\mathrm{mm}}^{\max}\right\} \\
k_{\mathrm{mm}}^{\max} &= \max\left\{\min\left\{n_{\max}, f/f_{\mathrm{mm}}^{\min}\right\}, k_{\mathrm{mm}}^{\min}\right\} & k_{\mathrm{mm}}^{\mathrm{start}} &= \min\left\{\max\left\{k_{\mathrm{mm}}^{\min}, f/f_{\mathrm{mm}}^{\mathrm{start},*}\right\}, k_{\mathrm{mm}}^{\max}\right\}
\end{aligned}$$

Frequency ratios and auxiliary functions (Secs. 6.1.1, 6.1.2, 6.1.3, 6.1.4, 6.2.3)**:**

$$x_r^i = f/f_r^i \qquad x_k^{(i)} = f/f_k^{(i)} \qquad x_m^i = f/f_m^i \qquad \tilde{x}_m^i = f/\tilde{f}_m^i$$

$$\begin{aligned}
\phi\left(y\right) &= 27/2\,y - 1 + 3^{3/2}\left(27/4\,y^2 - y\right)^{1/2} \\
\varphi_2\left(x,\chi\right) &= \left[1/4 + \left(1+\chi\right)x^2\right]^{1/2} - 1/2 \\
\varphi_3\left(x,\chi\right) &= 1/3\left[\phi^{1/3}\left((1+\chi)x^3\right) + \phi^{-1/3}\left((1+\chi)x^3\right) - 1\right]
\end{aligned}$$

$$\begin{aligned}
x_{A,i}^{\mathrm{start}} &= \max\left\{\chi_r^{-1}, \varphi_2(x_i^{(i)}, \chi_r)\right\} & x_{A,i}^{\mathrm{end}} &= x_i^{(i+1)} \\
x_{B,ij}^{\mathrm{start}} &= \max\left\{x_i^{(i)}, \varphi_2(x_i^{(j)}, \chi_r)\right\} & x_{B,ij}^{\mathrm{end}} &= \min\left\{x_i^{(i+1)}, \varphi_2(x_i^{(j+1)}, \chi_r)\right\} \\
x_{\mathrm{rm}}^{\mathrm{start}} &= \max\left\{x_m^{\mathrm{eq}}, \varphi_3\left(\tilde{x}_m^{\mathrm{ini}}, \chi_r\right)\right\} & x_{\mathrm{mm}}^{\mathrm{start}} &= \max\left\{\chi_m^{-1}, \varphi_3\left(\tilde{x}_m^{\mathrm{eq}}, \chi_m\right)\right\}
\end{aligned}$$

$$\hat{x}_{B,ij}^{\mathrm{start(end)}} = f/\hat{f}_{B,ij}^{\mathrm{start(end)}}$$

Shape functions (Secs. 6.1.1, 6.1.3, 6.1.4, 6.2.2, 6.2.4)**:**
We use ${}_2F_1$ to denote the ordinary hypergeometric function and $H_{n_{\max}}^{(q)}$ for the $n_{\max}$-th generalised harmonic number of order q.

$$S_{\mathrm{rr}}^{(1)}\left(x\right) = \left(\frac{x}{1+x}\right)^{3/2} \qquad S_{\mathrm{rm}}^{(1)}\left(x\right) = \frac{2+3\,x}{(x_m^0)^{1/2}\left(1+x\right)^{3/2}}$$

$$S_{\mathrm{mm}}^{(1)}\left(x\right) = \frac{1}{(x_m^0)^2}\left(\frac{x^2+x-1}{1+x} - 2\,\ln\left(1+x\right)\right)$$

$$S_{\mathrm{rr},1}(k,x) = \begin{cases} \frac{2\Gamma(2-q)}{(q-1)(2q+1)\Gamma(1-q)}\,\frac{x^{3/2}}{k^{1/2+q}} & \text{if } x < 10^{-3} \\ -\frac{{}_2F_1\left[\frac{3}{2},1-q,2-q,-\frac{k}{x}\right]}{(q-1)k^{q-1}} & \text{else} \end{cases}$$

$$S_{\mathrm{rr},2}(k,x) = \begin{cases} \frac{x^3}{k^{q+2}}\left(-\frac{1}{q+2} + \frac{3}{q+4}\,\frac{x^2}{k^2} - \frac{6}{q+6}\,\frac{x^4}{k^4}\right) & \text{if } \frac{x}{\sqrt{1+\chi_r}} < 10^{-8} \vee f_{\mathrm{rr}}^{\max} > 0 \\ S_{\mathrm{rr},1}(k,x) & \text{else} \end{cases}$$

$$S_{\mathrm{rr},3}(k) = -\frac{S_{\mathrm{rr}}^{(1)}(\chi_r^{-1})}{(q-1)k^{q-1}}$$

$$S_{\mathrm{rm},1}(k,x) = \begin{cases} \frac{3}{2(3+2q)k^{3/2+q}\sqrt{x_m^0}}x^3 & \text{if } x_m^0 < 10^{-5} \vee f_{\mathrm{rm}}^{\max} > 0 \\ S_{\mathrm{rm},2}(k,x) & \text{else} \end{cases}$$

$$S_{\mathrm{rm},2}(k,x) = \begin{cases} \frac{3k^{-1/2-q}x^2}{(2+4q)\sqrt{x_m^0}} & \text{if } x_m^0 < 10^{-5} \vee f_{\mathrm{rm}}^{\max} > 0 \\ \frac{2k^{2-q}+2(2q-1)(k+x)(-x)^{1-q}\,{}_2F_1\left[\frac{1}{2},q-1;\frac{3}{2};\frac{k+x}{x}\right]}{\sqrt{x_m^0\,(k+x)}} & \text{if } 10^{-5} \le x_m^0 < 10^{10} \wedge f_{\mathrm{rm}}^{\max} < 0 \\ \frac{3k^{2-q}}{(2-q)\sqrt{x_m^0 x}} & \text{else} \end{cases}$$

$$\tilde{S}_{\mathrm{mm},1}(k,x) = \left(\frac{x}{x_m^0}\right)^2 \frac{k^{1-q}\left({}_2F_1\left(1,1-q;2-q;-\frac{k}{x}\right) - 2\left(\frac{k}{x}\right)^2 \ln\left(1+\frac{x}{k}\right) - \left(\frac{k}{x}-1\right)^2\right)}{3-q}$$

$$\tilde{S}_{\mathrm{mm},3}(k,x) = \frac{k^{3-q}}{(q-3)\left(x_m^0\right)^2}\left\{1 - {}_2F_1\left(1,2-\frac{2q}{3};3-\frac{2q}{3};-\left(\frac{k}{x}\right)^{3/2}\right) + \right.$$
$$\left. +2\ln\left(1+\left(\frac{x}{k}\right)^{3/2}\right) - \frac{2}{2q-3}\left(\frac{x}{k}\right)^{3/2}\left(3{}_2F_1\left(1,\frac{2q}{3}-1;\frac{2q}{3};-\left(\frac{x}{k}\right)^{3/2}\right)+q-3\right)\right\}$$

$$S_{\mathrm{mm},1}(k,x) = \begin{cases} \left(\frac{x}{x_m^0}\right)^2\left(-\frac{x}{3qk^q}+\frac{x^2}{2(1+q)k^{1+q}}\right) & \text{if } x_m^0 < 10^{-2} \\ \tilde{S}_{\mathrm{mm},1}(k,x), & \text{if } 10^{-2} \le x_m^0 < 10^6 \\ \frac{k^{2-q}}{(x_m^0)^2}\left(\frac{x}{2-q}-\frac{2k}{(3-q)^2}-\frac{2k}{3-q}\ln\left(\frac{x}{k}\right)-\frac{3k^2}{(4-q)x}\right) & \text{else} \end{cases}$$

$$S_{\mathrm{mm},2}(k) = \begin{cases} \frac{k^{3-q}}{3-q}\left(\frac{1}{x_m^0}\right)^2\left(\frac{\chi_m^{-3}}{3}-\frac{\chi_m^{-4}}{2}\right) & \text{if } x_m^0 < 10^{-2} \\ \frac{k^{3-q}}{3-q}\left(\frac{1}{x_m^0}\right)^2\left(\frac{\chi_m^{-3}}{3}-\frac{\chi_m^{-4}}{2}-1\right) & \text{else} \end{cases}$$

$$S_{\mathrm{mm},3}(k,x) = \begin{cases} -\frac{2x^{9/2}}{3(3+2q)k^{3/2+q}(x_m^0)^2} & \text{if } x_m^0 < 10^{-2} \\ \tilde{S}_{\mathrm{mm},3}(k,x) & \text{if } 10^{-2} \le x_m^0 < 10^3 \\ S_{\mathrm{mm},1}(k,x) & \text{else} \end{cases}$$

RR contribution (Secs. 6.1.2, 6.2.2, 6.2.3)**:**
We denote by N the number of intervals during radiation domination for which we assume the effective number of degrees to be roughly constant (see *Cosmological parameters*)

$$h^2\Omega_{\mathrm{rr}}^{(1)} = \mathcal{A}_{\mathrm{rr}} \sum_{i=0}^{N-1}\left(\mathcal{G}_i\, S_{\mathrm{rr}}^{(1)}(x)\Big|_{x_{A,i}^{\mathrm{start}}}^{x_{A,i}^{\mathrm{end}}} \Theta(x_{A,i}^{\mathrm{end}}-x_{A,i}^{\mathrm{start}}) + \right.$$
$$\left. + \sum_{j=0}^{i-1}\mathcal{G}_j^{3/4}\mathcal{G}_i^{1/4}\, S_{\mathrm{rr}}^{(1)}(x)\Big|_{x_{B,ij}^{\mathrm{start}}}^{x_{B,ij}^{\mathrm{end}}} \Theta(x_{B,ij}^{\mathrm{end}}-x_{B,ij}^{\mathrm{start}})\right)$$

If $f_{\mathrm{rr}}^{\max} \ge 0$:

$$h^2\Omega_{\mathrm{rr}}^{m,\mathrm{upper}} = \frac{1}{H_{n_{\max}}^{(q)}}\sum_{k=1}^{m}\frac{1}{k^q}h^2\Omega_{\mathrm{rr}}^{(1)}\left(\frac{f}{k}\right) +$$
$$+\frac{\mathcal{A}_{\mathrm{rr}}}{H_{n_{\max}}^{(q)}}\left\{S_{\mathrm{rr},1}(k,x_r^{\mathrm{eq}})\Big|_{k_{\mathrm{rr}}^{\min}}^{k_{\mathrm{rr}}^{\max}} - S_{\mathrm{rr},2}(k,\sqrt{1+\chi_r}x_r^{\mathrm{ini}})\Big|_{k_{\mathrm{rr}}^{\min}}^{k_{\mathrm{rr}}^{\mathrm{start}}} - S_{\mathrm{rr},3}\left(k\right)\Big|_{k_{\mathrm{rr}}^{\mathrm{start}}}^{k_{\mathrm{rr}}^{\max}}\right\}$$

If $f_{\mathrm{rr}}^{\max} < 0$:

$$h^2\Omega_{\mathrm{rr}}^{m,\mathrm{upper}} = \frac{1}{H_{n_{\max}}^{(q)}}\sum_{k=1}^{m}\frac{1}{k^q}h^2\Omega_{\mathrm{rr}}^{(1)}\left(\frac{f}{k}\right) +$$
$$+\frac{\mathcal{A}_{\mathrm{rr}}}{H_{n_{\max}}^{(q)}}\sum_{i=0}^{N-1}\Theta_i^A\mathcal{G}_i\left\{S_{\mathrm{rr},1}\left(k,x_i^{(i+1)}\right)\Big|_{m}^{k_{A,i}^{\max}} - S_{\mathrm{rr},1}\left(k,\sqrt{1+\chi_r}x_i^{(i)}\right)\Big|_{m}^{k_{A,i}^{\mathrm{start}}} - S_{\mathrm{rr},3}\left(k\right)\Big|_{k_{A,i}^{\mathrm{start}}}^{k_{A,i}^{\max}}\right\} +$$
$$+\frac{\mathcal{A}_{\mathrm{rr}}}{H_{n_{\max}}^{(q)}}\sum_{i=1}^{N-1}\sum_{j=0}^{i-1}\Theta_{ij}^B\mathcal{G}_i^{1/4}\mathcal{G}_j^{3/4}\left\{S_{\mathrm{rr},1}\left(k,\hat{x}_{B,ij}^{\mathrm{end}}\right)\Big|_{m}^{k_{B,ij}^{\max}} - S_{\mathrm{rr},1}\left(k,\hat{x}_{B,ij}^{\mathrm{start}}\right)\Big|_{m}^{k_{B,ij}^{\max}}\right\}$$

with

$$\Theta_i^A = \Theta\left(\frac{a_{(i+1)}}{a_{(i)}} - \sqrt{1+\chi_r}\right) \qquad \text{and} \qquad \Theta_{ij}^B = \Theta\left(\hat{f}_{B,ij}^{\mathrm{start}} - \hat{f}_{B,ij}^{\mathrm{end}}\right)$$

RM contribution (Secs. 6.1.3, 6.2.4)**:**

$$
\begin{aligned}
h^2\Omega_{\rm rm}^{(1)} &= \mathcal{A}_{\rm rm}\Theta\left(f - f_{\rm mm}^{\rm min}\right)\Theta\left(x_m^0 - x_{\rm rm}^{\rm start}\right)\left[S_{\rm rm}^{(1)}(x_{\rm rm}^{\rm start}) - S_{\rm rm}^{(1)}(x_m^0)\right] \\
h^2\Omega_{\rm rm}^{m,\text{upper}} &= \frac{1}{H_{n_{\rm max}}^{(q)}}\sum_{k=1}^{m}\frac{1}{k^q}h^2\Omega_{\rm rm}^{(1)}\left(\frac{f}{k}\right) + \\
&+ \frac{\mathcal{A}_{\rm rm}}{H_{n_{\rm max}}^{(q)}}\left\{S_{\rm rm,1}\left(k,(1+\chi_r)^{1/3}\tilde{x}_m^{\rm ini}\right)\Big|_{k_{\rm rm}^{\rm min}}^{k_{\rm rm}^{\rm start}} + S_{\rm rm,2}\left(k,x_m^{\rm eq}\right)\Big|_{k_{\rm rm}^{\rm start}}^{k_{\rm rm}^{\rm max}} - S_{\rm rm,2}\left(k,x_m^0\right)\Big|_{k_{\rm rm}^{\rm min}}^{k_{\rm rm}^{\rm max}}\right\}
\end{aligned}
$$

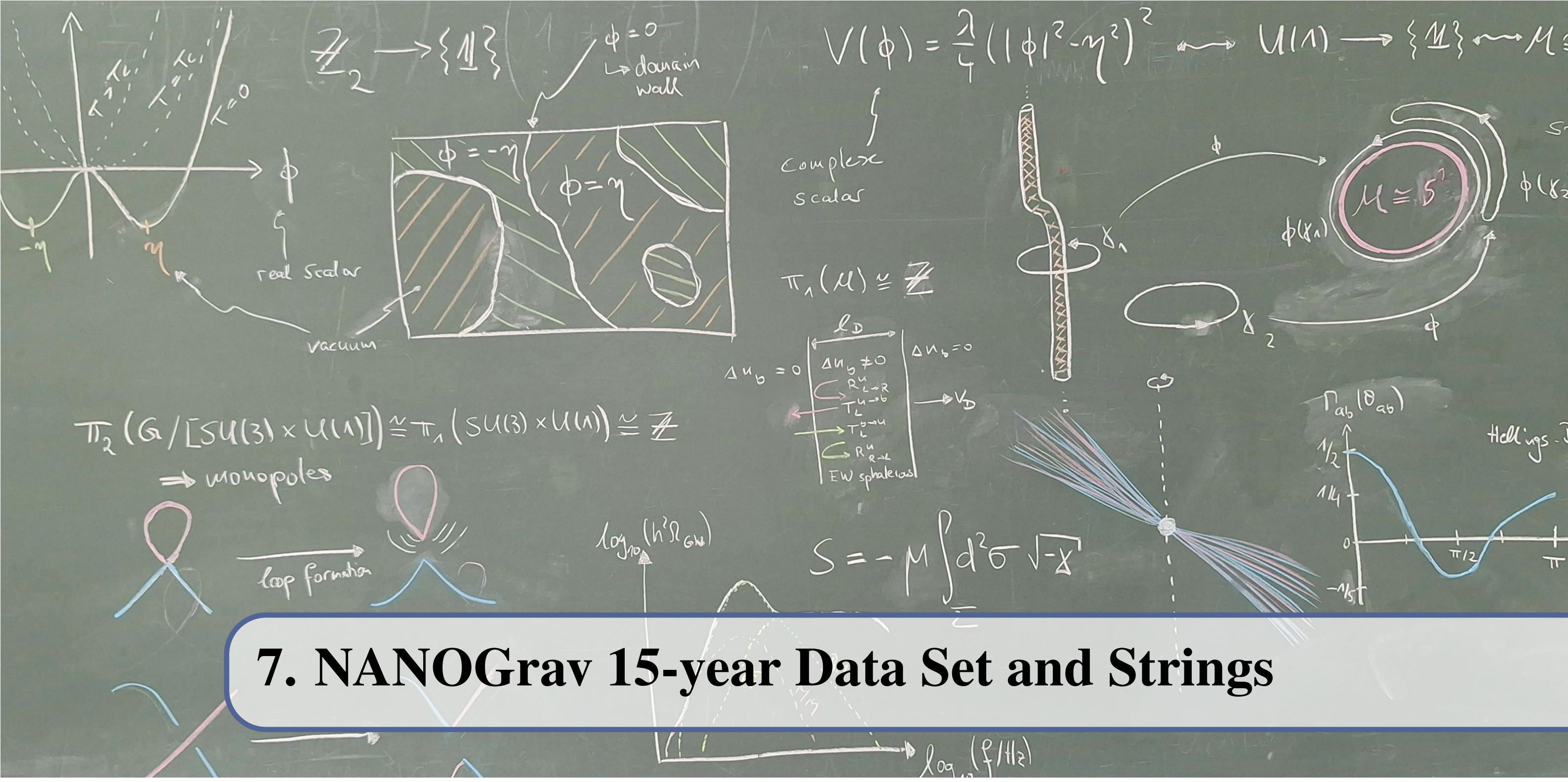
MM contribution (Secs. 6.1.3, 6.1.4, 6.1.5, 6.2.4)**:**

$$
\begin{aligned}
h^2\Omega_{\rm mm}^{(1)} &= \mathcal{A}_{\rm mm}\Theta\left(x_m^0 - x_{\rm mm}^{\rm start}\right)\left[S_{\rm mm}^{(1)}(x_m^0) - S_{\rm mm}^{(1)}(x_{\rm mm}^{\rm start})\right] \\
h^2\Omega_{\rm mm}^{m,\text{upper}} &= \frac{1}{H_{n_{\rm max}}^{(q)}}\sum_{k=1}^{m}\frac{1}{k^q}h^2\Omega_{\rm mm}^{(1)}\left(\frac{f}{k}\right) + \\
&+ \frac{\mathcal{A}_{\rm mm}}{H_{n_{\rm max}}^{(q)}}\left\{S_{\rm mm,1}\left(k,x_m^0\right)\Big|_{k_{\rm mm}^{\rm min}}^{k_{\rm mm}^{\rm max}} - S_{\rm mm,3}\left(k,(1+\chi_m)^{1/3}\tilde{x}_m^{\rm eq}\right)\Big|_{k_{\rm mm}^{\rm min}}^{k_{\rm mm}^{\rm start}} - S_{\rm mm,2}(k)\Big|_{k_{\rm mm}^{\rm start}}^{k_{\rm mm}^{\rm max}}\right\}
\end{aligned}
$$

Complete spectrum:

$$
h^2\Omega_{\rm GW}^{m,\rm upper} = h^2\Omega_{\rm rr}^{m,\rm upper} + h^2\Omega_{\rm rm}^{m,\rm upper} + h^2\Omega_{\rm mm}^{m,\rm upper}
$$

To calculate the complete spectrum, one furthermore needs to choose values for q, m, $n_{\rm max}$, and $t_{\rm ini}$. For our plots, we used $q = 4/3$, $m = 1$, and $n_{\rm max} = 10^5$. A typical choice is $t_{\rm ini} = t_{\rm Pl}/(G\mu)^2$. $h^2\Omega_{\rm GW}^{m,\rm lower}$ is obtained by replacing $(m, n_{\rm max}) \mapsto (m+1, n_{\rm max}+1)$ in the mode numbers. If one is interested in only one value of the spectrum instead of an upper and lower bound, one may refer to either of these bounds or ideally take the geometric mean $h^2\Omega_{\rm GW}^{m} = \sqrt{h^2\Omega_{\rm GW}^{m,\rm upper}h^2\Omega_{\rm GW}^{m,\rm lower}}$.



7. NANOGrav 15-year Data Set and Strings

This chapter is based on Ref. [1]. The detailed contributions of the author to this article are outlined at the beginning of this thesis.

Pulsars, first discovered in 1967 by Bell Burnell et al. [436], can be understood as rapidly rotating and highly magnetised neutron stars [437, 438]. In the simplest model, they are described as magnetic dipoles that emit electromagnetic radiation along their dipole axis, which is a consequence of the misalignment between this axis and the rotation axis (for more sophisticated models, see, e.g. Ref. [439–441]). Due to their short rotational periods, the emitted dipole radiation lies in the range of radio frequencies. It is illustrative to imagine pulsars as lighthouses: Each time the radio beam sweeps over the Earth, the radio signal can be measured as a short pulse in the detector — hence the name *pulsar*. Pulsars can roughly be divided into two classes: canonical pulsars and millisecond pulsars [442]. The former constitute the majority of pulsars. However, we are rather interested in the latter, which are characterised by shorter periods in the range of a few milliseconds and high rotational stability. In particular, this means that the measured radio pulse shapes, averaged over many rotations, remain reproducible over decades [443]. This allows the *times of arrival (TOAs)* of the pulses to be predicted in appropriate timing models. The difference between the measured TOAs and the pulse arrival time predicted by the timing model is known as a *timing residual*. If a GW passes between the pulsar and the detector on Earth, it can perturb the distance the radio pulse has to travel from its emission to its detection. This causes the pulse to arrive slightly earlier or later than predicted by the timing model. Thereby, it will generate non-vanishing timing residuals. Of course, even the best timing models will not be able to capture all pulsar-intrinsic processes completely, which also results in timing residuals that are non-zero. Pulsar-intrinsic noise is, however, expected to be uncorrelated among different pulsars. Therefore, to disentangle GWs from this noise source, one must monitor an entire set of different pulsars, known as a Pulsar Timing Array (PTA).[1] PTAs are sensitive to GWs with nanohertz-frequencies, corresponding to periods of years to decades. Although GWs are not the only possible source of a

---

[1] For recent reviews on PTAs, see Refs. [444, 445]. For a textbook review in the broader context of GW astrophysics and cosmology, see Ref. [404].

correlated signal, a stochastic GWB that is stationary, isotropic, Gaussian, and unpolarised leaves a specific correlation pattern between timing residuals known as the *Hellings–Downs curve*. This pulsar cross-correlation was originally computed in Ref. [446]. In 2023, several PTA collaborations around the world reported evidence for a Hellings–Downs correlation. These include the Australian Parkes PTA (PPTA) [105], the European and the Indian PTA (EPTA, InPTA) [104], the Chinese PTA (CPTA) [103], as well as the North American Nanohertz Observatory for GWs (NANOGrav) [102]. More recently, results from the MeerKAT PTA (MPTA) [447] have also shown evidence in favour of such a correlation, albeit only for certain noise models. Here, we focus on the NANOGrav 15-yr (NG15) data set [448], which finds particularly compelling evidence for a Hellings–Downs correlation with a statistical significance of about $3.5\sigma \ldots 4\sigma$ [102]. Under the assumption that the measured signal is indeed a GWB, we can use the data to test various models that predict a GWB. The most commonly expected source of the signal is a population of inspiraling supermassive black hole binaries (SMBHBs), which are located at the centres of galaxies. This possibility was investigated in detail in Ref. [449] (see also Ref. [445] for a recent review). Alternatively, the GWB could also stem from the early Universe and, as we have seen, in particular from cosmic strings. In the following, we investigate this option in detail. Our discussion is based on the original work in Ref. [1] (for topical studies using earlier PTA data sets, see, e.g., Refs. [450–452]; for complementary discussions, see, e.g., Refs. [407, 426, 428, 453]). We emphasise that, even if the signal in the NG15 data is not caused by a stochastic GWB, parameter values predicting overly large GWB amplitudes, for example due to excessive string tensions, can conflict with the NG15 data. This allows us to place competitive constraints on cosmic string parameters, independently of whether a GWB was observed or not.

The remainder of this chapter is structured as follows. In Sec. 7.1, we review the effect of GWs on the TOAs. In particular, we sketch the derivation of the Hellings–Downs curve. In Sec. 7.2, we summarise the details of the data analysis, including a brief description of the NG15 data set, the model we use to describe the timing residuals, and the statistical tools applied in the Bayesian inference. The main part of this chapter is found in Sec. 7.3. We first summarise the details of a model describing the GWB signal from SMBHBs. This model serves as a common reference in our analysis. We then provide the results of the Bayesian analysis for stable strings, metastable strings, and cosmic superstrings. For each of these categories of strings, this includes a short specification of the model, a description of the results of the model comparison, and a brief interpretation of the reconstructed posterior distributions. In Sec. 7.4, we conclude with a discussion of the implications of our results and future research objectives.

7.1 PTA response to a stochastic GWB

In the following, we first review how a GW passing between the pulsar and the radio telescope on Earth affects the propagation of the pulsar's radio signal. To this end, we largely follow the discussion of Ref. [404].

For the distances between the Earth and the pulsars that are located in our galaxy, we can neglect the effect of cosmic expansion, and describe the GW as a tensor perturbation around a Minkowski background. Since the photon's trajectory passes through vacuum, we can work in the transverse traceless gauge,

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \left(\delta_{ij} + h_{ij}(t, \boldsymbol{x})\right) \mathrm{d}x^i \mathrm{d}x^j \,, \tag{7.1}$$

with ${h^i}_i = 0$ and $\partial_i {h^i}_j = 0$. In the following, we work to linear order in the perturbation, which we assume to satisfy $|h_{ij}| \ll 1$ in the chosen reference frame.

Let us now fix the origin of our coordinate system such that the Earth is located at $\boldsymbol{x} = 0$ and the pulsar at $\boldsymbol{x} = L_a \hat{\boldsymbol{n}}_a$, where $\hat{\boldsymbol{n}}_a$ is a unit vector and the subscript a labels the pulsar that emits the radio signal. Since we work in the transverse traceless gauge, the coordinate positions of Earth and pulsar are fixed, and L_a is not affected by the tensor perturbation. We can describe the photon path as $\boldsymbol{x}_\gamma(t) = x_\gamma(t)\hat{\boldsymbol{n}}_a$. Denoting by t_{em} the time of emission and by t_{obs} the time of observation, we must have $x_\gamma(t_{\text{em}}) = L_a$ and $x_\gamma(t_{\text{obs}}) = 0$. Since photons propagate along null geodesics, which satisfy $(\mathrm{d}s/\mathrm{d}t)^2 = 0$, it follows directly from Eq. (7.1) that

$$0 = 1 - \left(1 + n_a^i n_a^j h_{ij}(t, \boldsymbol{x}_\gamma(t))\right) \left(\frac{\mathrm{d}x_\gamma}{\mathrm{d}t}\right)^2 , \tag{7.2}$$

such that we obtain in linear order of the tensor perturbation

$$\frac{\mathrm{d}x_\gamma}{\mathrm{d}t} = -\left(1 - \frac{n_a^i n_a^j}{2} h_{ij}(t, \boldsymbol{x}_\gamma(t))\right) . \tag{7.3}$$

Here, we used the fact that for the given trajectory the distance to the Earth decreases over time, i.e., $\mathrm{d}x_\gamma/\mathrm{d}t < 0$. Integrating this equation yields

$$L_a = t_{\text{obs}} - t_{\text{em}} - \frac{n_a^i n_a^j}{2} \int_{t_{\text{em}}}^{t_{\text{obs}}} \mathrm{d}t'\, h_{ij}\left(t', \left(t_{\text{em}} + L_a - t'\right) \hat{\boldsymbol{n}}_a\right) , \tag{7.4}$$

where we expanded again in linear order of the tensor perturbation and used that only the zeroth-order observation time $t_{\text{obs}} \simeq t_{\text{em}} + L_a$ and the zeroth-order photon trajectory $x_\gamma(t) \simeq t_{\text{em}} + L_a - t \equiv x_\gamma^{(0)}(t)$ contribute to the linearised integral. We can now consider the light of the next pulse, emitted one pulsar period P_a later at the time $t'_{\text{em}} = t_{\text{em}} + P_a$. In analogy to the above expression, we can compute the arrival time of the pulse t'_{obs} and write

$$t'_{\text{obs}} - t_{\text{obs}} = P_a + \Delta P_a \,, \tag{7.5}$$

where ΔP_a is the deviation in the TOA induced by GWs. From Eq. (7.4) and an analogous expression for t'_{obs}, we find

$$\Delta P_a = \frac{\hat{n}_a^i \hat{n}_a^j}{2} \int_{t_{\text{em}}}^{t_{\text{em}}+L_a} \mathrm{d}t' \left[h_{ij}\left(t' + P_a, x_\gamma^{(0)}(t')\hat{\boldsymbol{n}}_a\right) - h_{ij}\left(t', x_\gamma^{(0)}(t')\hat{\boldsymbol{n}}_a\right)\right] . \tag{7.6}$$

Recalling that PTAs are sensitive to GWs with periods T_a of the order of years, while P_a is of the order of milliseconds, we can expand the first term in the integrand to a very good approximation in leading order around t' and find

$$z_a(t) \equiv \frac{\Delta P_a}{P_a}(t) = \frac{\hat{n}_a^i \hat{n}_a^j}{2} \int_{t-L_a}^{t} \mathrm{d}t' \left[\frac{\partial}{\partial t'} h_{ij}(t', \boldsymbol{x})\right]_{\boldsymbol{x}=x_\gamma^{(0)}(t')\hat{\boldsymbol{n}}_a} . \tag{7.7}$$

Here, we called the observation time $t_{\text{obs}} \to t$. This is an expression for the instantaneous delay in the TOA due to a GW, which is the same as the GW-induced redshift, explaining the usage of the symbol z. For observations continuing over an extended period of time, the GW-induced shifts in the TOAs accumulate to the timing residual R_a:

$$R_a(t) = \int_0^t \mathrm{d}t'\, z_a(t') + R_a(0) \,, \tag{7.8}$$

where $R_a(0)$ is an arbitrary integration constant that we will fix later in a convenient way.

Let us now be more specific about the function h_{ij}. In the transverse traceless gauge, the linearised Einstein equations take the form of a wave equation:

$$\Box h_{ij} = 0 \,. \tag{7.9}$$

A general solution to this equation can easily be obtained by performing a Fourier and polarisation mode decomposition. One finds that

$$h_{ij}(t, \boldsymbol{x}) = \sum_{A\in\{+,\times\}} \int_{-\infty}^{\infty} \mathrm{d}f \oint \mathrm{d}^2\hat{k}\, \tilde{h}_A(f, \hat{\boldsymbol{k}}) e_{ij}^A(\hat{\boldsymbol{k}}) e^{-2\pi i f(t-\hat{\boldsymbol{k}}\cdot\boldsymbol{x})}\,, \tag{7.10}$$

where $\hat{\boldsymbol{k}}$ is a unit vector, $\tilde{h}_A(f, \hat{\boldsymbol{k}}) = \tilde{h}_A^*(-f, \hat{\boldsymbol{k}})$, and the basis of polarisation tensors is given by

$$e_{ij}^{+}(\hat{\boldsymbol{k}}) = \hat{u}^i\hat{u}^j - \hat{v}^i\hat{v}^j \qquad \text{and} \qquad e_{ij}^{\times}(\hat{\boldsymbol{k}}) = \hat{u}^i\hat{v}^j + \hat{v}^i\hat{u}^j\,. \tag{7.11}$$

Here, $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ are vectors that, together with $\hat{\boldsymbol{k}}$, form an orthonormal system. Inserting Eq. (7.10) into Eq. (7.7), one obtains, upon carrying out the integral over t',

$$z(t) = \sum_{A\in\{+,\times\}} \int_{-\infty}^{\infty} \mathrm{d}f \oint \mathrm{d}^2\hat{k}\, \tilde{h}_A(f, \hat{\boldsymbol{k}}) F_a^A(\hat{\boldsymbol{k}}) e^{-2\pi i f t} \left[1 - e^{2\pi i f L_a(1+\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{n}}_a)}\right]\,, \tag{7.12}$$

where we introduced the detector pattern functions

$$F_a^A(\hat{\boldsymbol{k}}) = \frac{\hat{n}_a^i \hat{n}_a^j e_{ij}^A(\hat{\boldsymbol{k}})}{2\left(1 + \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{n}}_a\right)}\,. \tag{7.13}$$

The expression for $z(t)$ consists of two contributions: the two terms in the square brackets. The first term results from evaluating the integral over t' at the observation time, the second from evaluation at the emission time. They are, correspondingly, referred to as *Earth term* and *pulsar term*.

We can now also compute the GW-induced timing residuals, which are, with an appropriate choice of $R_a(0)$ that removes an unobservable constant offset in the residuals, of the form

$$R_a(t) = \frac{i}{2\pi} \sum_{A\in\{+,\times\}} \int_{-\infty}^{\infty} \frac{\mathrm{d}f}{f} \oint \mathrm{d}^2\hat{k}\, \tilde{h}_A(f, \hat{\boldsymbol{k}}) F_a^A(\hat{\boldsymbol{k}}) e^{-2\pi i f t} \left[1 - e^{2\pi i f L_a(1+\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{n}}_a)}\right]\,. \tag{7.14}$$

As explained in the introduction to this chapter, we are interested in a stochastic GWB, and to measure it with PTAs, we focus on the correlation between timing residuals of pulsars $\langle R_a R_b\rangle$. The brackets indicate an ensemble average. However, since our Universe constitutes only one realisation, we assume for practical purposes ergodicity and replace ensemble with time averages. Treating the tensor perturbation as a random variable and assuming a stationary, isotropic, unpolarised background, we can write [454]

$$\left\langle \tilde{h}_A^*(f, \hat{\boldsymbol{k}}) \tilde{h}_{A'}(f', \hat{\boldsymbol{k}}')\right\rangle = \frac{1}{8\pi}\delta(f - f')\delta_{S^2}(\hat{\boldsymbol{k}}, \hat{\boldsymbol{k}}')\delta_{AA'} S_h(f)\,, \tag{7.15}$$

where $S_h(f)$ is the spectral density of the stochastic GWB and satisfies $S_h(f) = S_h(-f)$. The Dirac distribution in the frequency follows from stationarity, which is a good assumption since the GWB from both early-Universe processes and SMBHBs will change over cosmological timescales but remain effectively constant for the duration of any experiment. The assumption that the GWB is unpolarised leads to the Kronecker delta, and is well-justified as we consider a superposition of many randomly oriented sources. Finally, isotropy leads to the Dirac distribution $\delta_{S^2}(\hat{\boldsymbol{k}}, \hat{\boldsymbol{k}}')$, which is normalised to integrate to 1 over the sphere S^2. Isotropy is a good leading-order approximation, in particular for a primordial GWB. This is also true for a GWB from SMBHBs, although it will generally be more anisotropic than those originating from the early Universe. For cosmic strings, the persistence of the network leads to late-time GWB contributions and anisotropy searches may become relevant for this case in the future (see, e.g., Ref. [455]). As of yet, the NG15 data set shows no evidence for anisotropies in the GWB [456] and rules out neither SMBHBs nor early-Universe sources.

Combining Eqs. (7.14) and (7.15) allows us to compute the correlator between timing residuals. One finds

$$\langle R_a(t)R_b(t)\rangle = \int_{-\infty}^{\infty} \frac{\mathrm{d}f}{8\pi^2 f^2} S_h(f) \oint \frac{\mathrm{d}^2\hat{k}}{4\pi} \sum_A F_a^A(\hat{\boldsymbol{k}})F_b^A(\hat{\boldsymbol{k}})\kappa_{ab}(f,\hat{\boldsymbol{k}})\,, \tag{7.16}$$

where we defined

$$\kappa_{ab}(f,\hat{\boldsymbol{k}}) = \left[1 - e^{-2\pi i f L_a(1+\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{n}}_a)}\right]\left[1 - e^{2\pi i f L_b(1+\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{n}}_b)}\right]\,. \tag{7.17}$$

Let us first focus on the term in Eq. (7.17). For the relevant range of frequencies, the pulsars we are interested in are all at distances and separations L satisfying $fL \gg 1$, such that the exponential terms are rapidly oscillating functions and their contribution to the integral is negligible.[2] Therefore, we can replace $\kappa_{ab}(f,\hat{\boldsymbol{k}}) \to 1$ if $a \neq b$ and $\kappa_{aa}(f,\hat{\boldsymbol{k}}) \to 2$. Upon introducing a function that contains all the geometric information of the set up, namely

$$\Gamma_{ab}(\theta_{ab}) = \frac{3}{8\pi}(1+\delta_{ab}) \oint \mathrm{d}^2\hat{k} \sum_A F_a^A(\boldsymbol{k})F_b^A(\boldsymbol{k})\,, \qquad \text{with} \qquad \cos(\theta_{ab}) = \hat{\boldsymbol{n}}_a \cdot \hat{\boldsymbol{n}}_b\,, \tag{7.18}$$

we can write the correlator as the product

$$\langle R_a(t)R_b(t)\rangle = \Gamma_{ab}(\theta_{ab}) \int_0^{\infty} \frac{\mathrm{d}f}{6\pi^2 f^2} S_h(f)\,, \tag{7.19}$$

and via the Wiener–Khinchin theorem as

$$\left\langle R_a(t)R_b(t')\right\rangle = \Gamma_{ab}(\theta_{ab}) \int_0^{\infty} \frac{\mathrm{d}f}{6\pi^2 f^2} S_h(f)\cos\left(2\pi f\left(t-t'\right)\right)\,. \tag{7.20}$$

The second factor is an integral of $S_h(f)$ and describes the overall strength of the timing residual correlation. The first factor, Γ_{ab}, describes how the correlation varies in dependence of the angular separation θ_{ab} of a pulsar pair in the sky and is generally referred to as the overlap reduction function. For the considered GWB, it can be explicitly computed. This was first done in Ref. [446] and the result is the well-known Hellings–Downs (HD) correlation:[3]

$$\Gamma_{ab}(\theta_{ab}) = (1+\delta_{ab})\left[\frac{3}{2}x_{ab}\ln(x_{ab}) - \frac{1}{4}x_{ab} + \frac{1}{2}\right]\,, \quad \text{where} \quad x_{ab} = \frac{1}{2}\left(1-\cos(\theta_{ab})\right)\,. \tag{7.21}$$

The HD curve is plotted in Fig. 7.1. This specific pattern is the hallmark of a stationary, isotropic, and unpolarised stochastic GWB. Clearly, timing residuals have their strongest correlation for pulsars that have small angular separations on the celestial sphere or pulsars that are nearly antipodal. The strongest anticorrelation can be found for pulsars that have an angular separation of nearly $\pi/2$. This is characteristic of a quadrupolar correlation — a result one might naively expect to obtain from GWs. However, the pattern is not perfectly quadrupolar and contains still large contributions from higher multipoles. This is easiest seen by noting that the HD curve reaches its minimum not exactly at $\pi/2$, and satisfies $\Gamma_{ab}(0) \neq \Gamma_{ab}(\pi)$. Mathematically, this can be understood by noting that the denominator of the detector pattern function in Eq. (7.13) introduces a preferred direction. For an extensive and pedagogical discussion of the HD correlation, we refer the interested reader to Ref. [459].

---

[2] For more details, see Refs. [404, 457], and for a case where the approximation is not applicable, see, e.g., Ref. [7].

[3] For a detailed derivation of this result, see, e.g., Ref. [458].

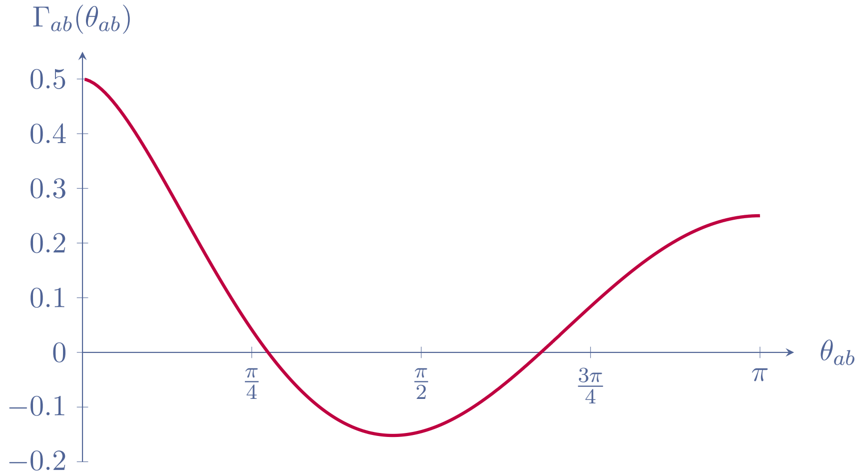


Figure 7.1: Plot of the Hellings–Downs correlation Γ_{ab} given in Eq. (7.21) as a function of the angular pulsar separation θ_{ab} on the celestial sphere. The plot shows only cross-correlations but not the autocorrelation $\Gamma_{aa} = 1$.

Let us now discuss the second part of Eq. (7.20): the spectral density $S_h(f)$. This function contains the spectral information on the GWB and can be expressed in terms of the density parameter $\Omega_{\rm GW}(f)$, which we used throughout previous chapters to describe the GWB spectrum. It is a well-known result in GW physics that the energy density of GWs is, in the transverse traceless gauge, of the form [454, 460, 461]

$$\rho_{\rm GW}(t,\boldsymbol{x}) = \frac{1}{32\pi G}\left\langle \partial_t h_{ij}(t,\boldsymbol{x})\partial_t h^{ij}(t,\boldsymbol{x})\right\rangle . \tag{7.22}$$

Inserting the expansion of $h_{ij}(t,\boldsymbol{x})$ given in Eq. (7.10), and using the description of the GWB in terms of Eq. (7.15), one finds

$$\rho_{\rm GW} = \frac{\pi}{2G}\int_0^\infty {\rm d}f\, f^2 S_h(f)\,. \tag{7.23}$$

Recalling that we defined $\Omega_{\rm GW}(f) = 1/\rho_{\rm crit}\,{\rm d}\rho_{\rm GW}/{\rm d}\ln(f)$ in Eq. (4.86), we can read off the relation

$$\Omega_{\rm GW}(f) = \frac{4\pi^2}{3H_0^2} f^3 S_h(f)\,. \tag{7.24}$$

As we will see in a moment, in practice one works with a Fourier decomposition of the timing residuals. Taking into account that the residuals are measured over a finite observation interval of length $T_{\rm obs}$, we can write

$$R_a(t) = \sum_{i=1}^{\infty}\left\{\alpha_i^{(a)}\sin\left(2\pi f_i t\right) + \beta_i^{(a)}\cos\left(2\pi f_i t\right)\right\}\,, \qquad \text{where} \qquad f_i = \frac{i}{T_{\rm obs}}\,. \tag{7.25}$$

The Fourier coefficients in the series are of the form

$$\alpha_i^{(a)} = \frac{2}{T_{\rm obs}}\int_0^{T_{\rm obs}} {\rm d}t\, R_a(t)\sin(2\pi f_i t) \quad \text{and} \quad \beta_i^{(a)} = \frac{2}{T_{\rm obs}}\int_0^{T_{\rm obs}} {\rm d}t\, R_a(t)\cos(2\pi f_i t)\,, \tag{7.26}$$

and we find from Eq. (7.20) that

$$\left\langle \alpha_i^{(a)}\alpha_j^{(b)}\right\rangle = \left\langle \beta_i^{(a)}\beta_j^{(b)}\right\rangle = \delta_{ij}\Gamma_{ab}\left(\theta_{ab}\right)\frac{S_h(f_i)}{6\pi^2 f_i^2 T_{\rm obs}} \qquad \text{and} \qquad \left\langle \alpha_i^{(a)}\beta_j^{(b)}\right\rangle = 0\,, \tag{7.27}$$

where $T_{\rm obs}$ is introduced as a result of the discretization of the frequency integral with $\Delta f = 1/T_{\rm obs}$. The relevance of this result will become clear in the following section.

7.2 NG15 data analysis

We have seen how a stochastic GWB influences the pulsar timing residuals. We now turn to the question of how the GWB predicted by string models can be tested using the NG15 data set. For a pedagogical introduction to the techniques employed, see Ref. [444].

Data set and timing model: We begin with a very brief description of the data set itself. A more detailed description is provided in Refs. [448, 462]. The data set consists of timing data of 68 pulsars that were recorded between July 2004 and August 2020. For the data analysis, we use a subset of 67 pulsars, excluding one pulsar that was monitored for less than three years. The 67 pulsars were observed with an approximately monthly cadence, except for six pulsars that were observed weekly starting from 2013 at the Green Bank Telescope and from 2015 at the Arecibo Observatory. The TOAs in the data set were obtained from raw data as detailed in Refs. [463–465]. The timing model that was fit to the TOAs takes into account the period of the pulsar and its derivative, the pulsar's location in the sky, its proper motion, and its parallax. It accounts for pulsars that are part of a binary system via five Keplerian binary parameters and, if it improved the fit, a non-Keplerian parameter. The dispersion of the TOAs due to ionised gas between observatory and pulsars was modelled as a step function via DMX parameters following Refs. [463, 466]. The timing model was fitted with the TT(BIPM2019) timescale and the JPL Solar System Ephemeris model DE440 in Ref. [467].

Timing residuals model and likelihood: Let us denote by $\boldsymbol{t}_{\text{obs}}^{\text{TOA}}$ the vector of all the N observed TOAs of a pulsar. The timing model described above provides us with predicted TOAs $\boldsymbol{t}_{\text{pred}}^{\text{TOA}}$ such that the timing residuals can be written as a vector of the form

$$\boldsymbol{\delta t} = \boldsymbol{t}_{\text{obs}}^{\text{TOA}} - \boldsymbol{t}_{\text{pred}}^{\text{TOA}}(\boldsymbol{\xi}_0)\,, \tag{7.28}$$

Here, we used $\boldsymbol{\xi}_0$ to denote the m initial best-fit values of the timing model. These timing residuals, i.e., shifts in the observed TOAs that are not accounted for by the timing model, are due to three main contributions:

$$\boldsymbol{\delta t} = \boldsymbol{t}_{\text{fit}} + \boldsymbol{t}_{\text{wn}} + \boldsymbol{t}_{\text{rn}}\,. \tag{7.29}$$

These are, in the above order: small errors in the fit to the timing model parameters, white noise, and time-correlated stochastic processes called red noise [468].

Let us discuss how we model these different contributions in our analysis. We begin with $\boldsymbol{t}_{\text{fit}}$ and assume that the best-fit parameter values $\boldsymbol{\xi}$ of the timing model, when evaluated for the full GW analysis, only deviate minimally from the initial best-fit values $\boldsymbol{\xi}_0$. Under this assumption, we can expand the predicted TOAs in linear order around $\boldsymbol{\xi}_0$:

$$\boldsymbol{t}_{\text{pred}}^{\text{TOA}}(\boldsymbol{\xi}) = \boldsymbol{t}_{\text{pred}}^{\text{TOA}}(\boldsymbol{\xi}_0) + \boldsymbol{M}\boldsymbol{\epsilon}\,, \quad \text{where} \quad M_{ij} = \frac{\partial t_{\text{pred},i}^{\text{TOA}}}{\partial \xi_j}\bigg|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0} \quad \text{and} \quad \boldsymbol{\epsilon} = \boldsymbol{\xi} - \boldsymbol{\xi}_0\,. \tag{7.30}$$

The $N \times m$ matrix $\boldsymbol{M}$ is referred to as the design matrix.

Next, we discuss the white noise giving rise to $\boldsymbol{t}_{\text{wn}}$. We describe it as $\boldsymbol{t}_{\text{wn}} = \boldsymbol{n}$ in terms of a Gaussian random variable $\boldsymbol{n}$ with vanishing mean. The white noise exhibits a frequency-independent power spectrum and no correlations among different pulsars. The measurement of $t_{\text{obs},i}^{\text{TOA}}$ is associated with a receiver–backend combination that we label I, and we model the white noise covariance as

$$\langle n_i n_j \rangle = \mathcal{F}_I^2 \left(\sigma_i^2 + \mathcal{Q}_I^2\right) \delta_{ij} + \mathcal{J}_I^2 \mathcal{U}_{ij}\,. \tag{7.31}$$

Let us describe the different terms in this white noise model. The TOA uncertainty due to radiometer noise is given by σ_i and multiplied by a parameter $\mathcal{F}_I$ called the "Extra FACtor" (EFAC), which was introduced as a correction for a possible error underestimation. For the possibility of additional white noise that cannot be accounted for as a multiple of the TOA uncertainties, the additional "Extra QUADrature" (EQUAD) term $\mathcal{Q}_I$ was introduced. The last term $\mathcal{J}_I\mathcal{U}_{ij}$ accounts for "Extra CORRelated" white noise (ECORR). This noise arises from the fact that, due to the finite number of pulses folded in each observation epoch, there are random fluctuations in the average pulse profile and thus in the TOAs. This effect is also known as phase jitter. Since in neighbouring frequency bands the recorded TOAs are nearly the same, the phase jitter will be fully correlated during the same observing epoch but uncorrelated between different epochs. This is accounted for with a matrix $\mathcal{U}$ of block-diagonal form, where $\mathcal{U}_{ij} = 1$ if i and j correspond to TOAs from the same observing epoch and $\mathcal{U}_{ij} = 0$ otherwise. The corresponding ECORR parameter is $\mathcal{J}_I$. For our analysis, we set all white-noise parameters to the values obtained by maximising the posterior probability densities obtained in single-pulsar noise studies [462].

Let us now turn to the last term in Eq. (7.29), which accounts for time-correlated stochastic processes. This notably includes a possible contribution from GWs but also from pulsar-intrinsic processes. The time dependence of the noise can be expressed in terms of a Fourier series with a fundamental frequency $1/T_{\rm obs}$, where $T_{\rm obs} = 16.03\,{\rm yr}$ is the timing baseline, defined as the span between the first and last TOAs in the data set. Since our focus lies on low-frequency processes, we truncate the Fourier series at a harmonic number N_f. This allows us to write

$$\boldsymbol{t}_{\rm rn} = \boldsymbol{F}\boldsymbol{\alpha}\,, \tag{7.32}$$

where $\boldsymbol{\alpha} = (\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_{N_f}, \beta_{N_f})^T$ are the $2N_f$ Fourier coefficients, and the Fourier basis is contained in the $N \times 2N_f$ Fourier design matrix

$$\boldsymbol{F} = \begin{pmatrix} \sin(\omega_{\rm obs}t_1) & \cos(\omega_{\rm obs}t_1) & \cdots & \sin(N_f\omega_{\rm obs}t_1) & \cos(N_f\omega_{\rm obs}t_1) \\ \vdots & & \ddots & & \vdots \\ \sin(\omega_{\rm obs}t_N) & \cos(\omega_{\rm obs}t_N) & \cdots & \sin(N_f\omega_{\rm obs}t_N) & \cos(N_f\omega_{\rm obs}t_N) \end{pmatrix}, \tag{7.33}$$

with $\omega_{\rm obs} = 2\pi/T_{\rm obs}$. For pulsar-intrinsic red noise, we restrict to $N_f = 30$ frequency bins ($2\,{\rm nHz} \ldots 59\,{\rm nHz}$), as for higher frequencies white noise is dominant over red noise. For the GWB, we choose $N_f = 14$ harmonics ($2\,{\rm nHz} \ldots 28\,{\rm nHz}$), since fitting a common-spectrum uncorrelated red-noise process (CURN) with a broken power-law spectrum to the NG15 data results in a posterior distribution of the break frequency, which peaks at the 14th frequency bin [102].

For the Fourier coefficients $\boldsymbol{\alpha}$, we assume a Gaussian distribution with vanishing mean and covariance $\left\langle \boldsymbol{\alpha}\boldsymbol{\alpha}^T \right\rangle = \boldsymbol{\phi}$ with components

$$\phi_{ai;bj} = \delta_{ij}\left(\delta_{ab}\varphi_{a,i} + \Gamma_{ab}\rho_i\right)\,, \tag{7.34}$$

where a and b label, as before, pulsars while, here, i and j label mode numbers in the Fourier expansion. The first term in Eq. (7.34) models pulsar-intrinsic red noise without correlations among different pulsars and we describe its spectrum by a power law

$$\varphi_a(f) = \frac{A_a^2}{12\pi^2}\frac{1}{T_{\rm obs}}\left(\frac{f}{1\,{\rm yr}^{-1}}\right)^{-\gamma_a}\,{\rm yr}^3 \qquad \text{and} \qquad \varphi_{a,i} = \varphi_a(i/T_{\rm obs})\,. \tag{7.35}$$

The pulsar-intrinsic red noise is therefore described, for each pulsar, by two parameters that we choose to be $\log_{10} A_a$ with a uniform prior distribution in the range $[-20, -11]$ and γ_a with a uniform prior in the range $[0, 7]$. The second term in Eq. (7.34) models the GWB and we choose the overlap reduction function Γ_{ab} to be the HD correlation given in Eq. (7.21). Comparing this form to the result in Eq. (7.27), it is easily seen that we have the following relation for the power spectral density:

$$\rho(f) = \frac{S_h(f)}{6\pi^2 f^2 T_{\rm obs}} = \frac{H_0^2}{8\pi^4 T_{\rm obs}} \frac{\Omega_{\rm GW}(f)}{f^5}\,, \qquad \text{with} \qquad \rho_i = \rho(i/T_{\rm obs})\,. \tag{7.36}$$

This relation connects the theoretical prediction for $\Omega_{\rm GW}$ — the central quantity of the previous sections — with observational data. In our analysis, we will choose $\Omega_{\rm GW}$ corresponding to various cosmic string models and to a simple SMBHB reference model.

We have now all the ingredients to construct the PTA likelihood. The complete model for the timing residuals is of the form

$$\boldsymbol{\delta t} = \boldsymbol{M\epsilon} + \boldsymbol{n} + \boldsymbol{F\alpha}\,. \tag{7.37}$$

For our analysis, we are only interested in the statistical properties of the noise, not in a specific realisation. This allows us to work with the noise-marginalised likelihood, obtained by integrating out all possible values of $\boldsymbol{\alpha}$ and $\boldsymbol{\epsilon}$. The remaining marginalised likelihood depends only on the parameters appearing in the red-noise covariance matrix $\boldsymbol{\phi}$. These are $\log_{10} A_a$ and γ_a for each pulsar, as well as the model parameters necessary to determine ρ_i. The marginalised likelihood reads [469, 470]

$$p(\boldsymbol{\delta t}|\boldsymbol{\phi}) = \frac{1}{\sqrt{\det\left(2\pi \boldsymbol{C}\right)}} \exp\left(-\frac{1}{2}\boldsymbol{\delta t}^T \boldsymbol{C}^{-1}\boldsymbol{\delta t}\right)\,, \tag{7.38}$$

where $\boldsymbol{C} = \boldsymbol{N} + \boldsymbol{TBT}^T$ is the covariance matrix. Here, $\boldsymbol{N}$ is the covariance matrix of the white noise, $\boldsymbol{T} = \begin{pmatrix}\boldsymbol{M} & \boldsymbol{F}\end{pmatrix}$ is an $N \times (m + 2N_f)$ matrix, and $\boldsymbol{B} = \operatorname{diag}\left(\boldsymbol{\infty}, \boldsymbol{\phi}\right)$ is an $(m + 2N_f) \times (m + 2N_f)$ matrix, where $\boldsymbol{\infty}$ denotes an $m \times m$ diagonal matrix of infinities, corresponding to the assumption of flat priors on the parameters $\boldsymbol{\epsilon}$. These infinities are unproblematic, since we require the inverse of the covariance matrix in which the infinities reduce to zeros.

Bayesian analysis: To test our model against the data, we use Bayesian inference. We start from a prior probability distribution (or prior) $p_M(\Theta)$ that quantifies our degrees of belief in certain parameters Θ of a model M before knowing the data D. We can then use the likelihood, which is the conditional probability distribution $p_M(D|\Theta)$ of the data, given the parameters Θ, to obtain the conditional probability distribution $p_M(\Theta|D)$ of the parameters Θ given the data. The latter is the central object of our following considerations and is referred to as the posterior probability distribution (or posterior). This relation between the different probabilities is made precise by Bayes' theorem, which states that

$$p_M(\Theta|D) = \frac{1}{\mathcal{Z}_M} p_M(D|\Theta) p_M(\Theta)\,, \tag{7.39}$$

where the normalisation

$$\mathcal{Z}_M = \int \mathrm{d}\Theta\, p_M(D|\Theta) p_M(\Theta)\,, \tag{7.40}$$

i.e., the marginalised likelihood, is called the evidence. For our purposes, M is the model of timing residuals given in Eq. (7.37), Θ is the set of parameters needed to specify the covariance matrix $\boldsymbol{\phi}$ given in Eq. (7.34), and the data D are the timing residuals in the NG15 data set. Apart from the prior choices for the CURN parameters $\log_{10} A_a$ and γ_a given before, we will specify the priors on the model parameters when introducing the different string models and the SMBHB model in the next section. To update our degrees of belief and obtain the posterior, we use the likelihood given in Eq. (7.38), which is implemented in `ENTERPRISE` [471] and `ENTERPRISE_EXTENSIONS` [472].

Based on the posterior probability, we identify the regions of our string model parameter space that best fit the signal in the NG15 data. To this end, we marginalise the posteriors over all but one or two parameters of interest. From these marginalised posteriors, we obtain one-dimensional credible intervals (C.I.) or two-dimensional credible regions, respectively. Specifically, for each model that we consider in the next section, we compute the 68% (and sometimes the 95%) credible regions by integrating the posterior density over regions of highest density until the integral reaches the value 0.68 (and 0.95). For a single parameter, this means computing the highest posterior density intervals (HPDIs). Note that HPDIs are constructed in such a way that the posterior density gives the same value when evaluated at the upper and lower bound of the HPDI. Furthermore, the posterior density evaluated at each parameter value inside the HPDI is higher than (or at least as high as) that at any parameter value outside the interval.

For a parameter of interest θ we also compute the Bayes estimator $\langle\theta\rangle$ and its variance σ_θ^2 as

$$\langle\theta\rangle = \int \mathrm{d}\theta\,\theta\,p_M(\theta|D) \qquad \text{and} \qquad \sigma_\theta^2 = \int \mathrm{d}\theta\,\theta^2 p_M(\theta|D) - \langle\theta\rangle^2\,. \tag{7.41}$$

Moreover, we compute the maximum posterior θ_{max}, defined as the parameter value at which the 1D marginalised posterior density $p_M(\theta)$ attains its maximum.

Furthermore, we will perform a model selection: For two models M_1 and M_0, we compute the Bayes factor, which is defined as the evidence ratio

$$BF_{10}(D) = \frac{\mathcal{Z}_{M_1}}{\mathcal{Z}_{M_0}}\,. \tag{7.42}$$

Depending on the value of the Bayes factor, we can interpret model M_1 as favoured over model M_0 or as disfavoured. A useful tool for the interpretation of Bayes factors is the Jeffreys scale [473]: If $BF_{10} < 1$, then model M_1 is disfavoured compared to M_0. If BF_{10} takes values in the intervals $[10^0, 10^{0.5}]$, $[10^{0.5}, 10]$, $[10, 10^{1.5}]$, $[10^{1.5}, 10^2]$ or is even larger, this is interpreted as negligible, substantial, strong, very strong, and decisive evidence favouring model M_1 over M_0, respectively. For the purpose of our analysis, M_0 will be a reference model in which there is no GW signal from strings, but only from SMBHBs as described in the next section.[4] The model M_1 will either contain a GWB signal from strings only, or it will contain a combined signal from strings and SMBHBs.

The latter case offers an additional way to derive parameter bounds. To understand this, let us first consider the case where we want to put a constraint on a single parameter θ associated with a GWB signal from strings. This could, for example, be the string tension. We consider the model M containing the combined GWB signal from the string model and SMBHBs and marginalise the likelihood over all parameters other than θ, such that we are left with the likelihood $p_M(D|\theta)$. For the parameter θ, we can now consider a value θ_0 that lies in the limit, where the strings effectively no longer contribute to the GWB signal, and only the SMBHBs do. Around θ_0, the function $p_M(D|\theta)$ will be constant, corresponding to a plateau of height $p_M(D|\theta_0)$. Consequently, for a parameter region in which $p_M(D|\theta) < p_M(D|\theta_0)$, adding a GWB signal from strings worsens the fit to the NG15 signal compared to when the strings are not contributing. To quantify this, we introduce the likelihood ratio

$$K(\theta) = \frac{p_M(D|\theta)}{p_M(D|\theta_0)}\,, \tag{7.43}$$

---

[4] Note that, comparing all models, e.g., M_1, M_2, to the same reference model M_0 allows us immediately to compare models among each other due to the transitivity of Bayes factors: $BF_{12} = BF_{10}/BF_{20}$.

which we will from now on refer to as the K ratio. We use this ratio to put bounds on our model parameters. Specifically, we view parameter regions for which $K(\theta) < 1/10$ as excluded. We will refer to the corresponding parameter bound as the K bound θ_K, defined as

$$K(\theta_K) = \frac{1}{10} \,. \tag{7.44}$$

It is possible to learn more about the K ratio by rewriting it via Bayes theorem in Eq. (7.39) as

$$K(\theta) = \frac{p_M(\theta|D)}{p_M(\theta)} \frac{p_M(\theta_0)}{p_M(\theta_0|D)} \,. \tag{7.45}$$

Note that for a flat prior, we have $p_M(\theta) = p_M(\theta_0)$ and the K ratio can be written as a ratio of posterior densities. Furthermore, independently of whether the prior is flat or not, we can identify the second ratio as the Savage–Dickey ratio [474], which allows us to identify it with the Bayes factor $BF = \mathcal{Z}_M/\mathcal{Z}_{M_0}$, where M_0 is obtained from M, when removing the contribution to the GWB signal that depends on the parameter θ. For our case, this means that we remove the strings from the model, which makes M_0 our reference SMBHB model. This is, again, a Bayes factor that we compute anyway, and it allows us to evaluate the K ratio as a local posterior-to-prior ratio weighted by the Bayes factor

$$K(\theta) = BF \frac{p_M(\theta|D)}{p_M(\theta)} \,. \tag{7.46}$$

If we are interested in multiple parameters Θ, we proceed in analogy to the one-parameter case and define an upper bound by declaring parameter regions in which $K(\Theta) < 1/10$ as excluded. The K ratio has a clear strength. While posterior densities, and thus bounds derived from them, depend on the prior, the K ratio is defined as a ratio of likelihoods and is therefore unaffected by choices of the prior distribution (see Ref. [475]).

Due to the large number of parameters involved in our models, we use in practice Markov Chain Monte Carlo (MCMC) methods from the package `PTMCMCSampler` [476] to sample from the posterior distribution. We compute the Bayes estimator $\langle\theta\rangle$ and its variance σ_θ^2, both introduced in Eq. (7.41), as the empirical mean and variance of the samples in the MCMC chain. To obtain the marginalised posterior densities that are shown in the plots in the next section, we apply kernel density estimation (KDE) to the MCMC samples using the package `GetDist` [477]. Based on the posterior density acquired in this way, we compute the HPDIs. The K bounds we cite are computed by combining Eqs. (7.44) and (7.46) and using the posteriors resulting from the KDE. The Bayes factors necessary for this and, in fact all Bayes factors we cite, are in practice not based on computing the model evidence of M_0 and M_1 separately, but rather on product space methods [478–480]. To this end, a model indexing variable n_{model} is introduced and sampled together with other variables of models M_0 and M_1. The value of n_{model} decides which model's likelihood is used in a given MCMC sampling step. Using equal priors for the two models $p(M_0) = p(M_1)$, the Bayes factor BF_{10} is determined as the ratio of the number of MCMC samples that n_{model} decided to sample from M_1's likelihood vs. the number of samples associated with M_0. We estimate the uncertainties on these Bayes factors using statistical bootstrapping [481]. In this approach, we resample the original set of MCMC samples with replacement, thus generating multiple realisations from the original set of samples. From these realisations, we obtain a distribution of Bayes factors whose mean and standard deviation are used as the central values and associated errors that we provide below. The values were obtained by generating 5×10^4 bootstrap realisations of the original set of MCMC samples.

The Bayesian inference analyses discussed in the following were implemented via the `ENTERPRISE`-wrapper `PTArcade` [9].

7.3 Results

In the following, we discuss the GWB prediction from various string models in light of the NG15 data. We consider each model individually and in combination with a GWB from SMBHBs. The specific models will be defined in a moment. We also use the GWB from SMBHBs as a reference model (previously M_0) for the computation of Bayes factors. For each model, we present the reconstructed marginalised posterior density of the string parameters. We first briefly describe each string model that we consider and then discuss the results of the data analysis. Before that, we describe the reference model for the GWB signal M_0, namely GWs from SMBHBs (for a recent review, see Ref. [445]).

GWB from SMBHBs — SMBHB: Supermassive black holes (SMBHs) are black holes with masses in the range of $\sim 10^6$ to 10^{10} solar masses. If they accrete large amounts of gas, they can be observed as active galactic nuclei in the electromagnetic spectrum. Direct observations of SMBHs exist due to the Event Horizon Telescope [482, 483]. Observational evidence is consistent with most galaxies containing SMBHs at their centres [484]. In the hierarchical process of structure formation, massive galaxies form in mergers of lighter galaxies [485]. The SMBHs of the lighter progenitor galaxies are expected to move towards the galactic centre of the merged galaxy and form binaries, namely SMBHBs [486]. Based on estimates of the merger rate, most massive galaxies can host more than one pair of SMBHs during their life up to the present day [487]. Very massive SMBHBs that could be observed via their GW signature are expected to be rare [445]. In fact, SMBHBs have not been observed as of yet. Nonetheless, the combined GW signal of a population of thousands to millions of inspiraling SMBHBs leads to a GWB that should be detectable by PTAs [488–490].

Given the narrow width of the frequency range in which PTAs are sensitive, we use a power-law description of the GWB spectrum of SMBHBs (see also Ref. [449]). More precisely, we use for $\rho(f)$ in Eq. (7.36) the form

$$\rho_{\mathrm{BHB}}(f) = \frac{A_{\mathrm{BHB}}^2}{12\pi^2}\frac{1}{T_{\mathrm{obs}}}\left(\frac{f}{\mathrm{yr}^{-1}}\right)^{-\gamma_{\mathrm{BHB}}} \mathrm{yr}^3\,, \tag{7.47}$$

where $\log_{10} A_{\mathrm{BHB}}$ and γ_{BHB} are the model parameters. The normalisation of the GWB mainly depends on the masses and abundance of these SMBHBs, whereas the spectral shape is determined by their evolution on sub-parsec scales [449]. If the binary evolution was exclusively due to emission of gravitational radiation, one would obtain $\gamma_{\mathrm{BHB}} = 13/3$ as shown in Ref. [491]. However, interactions with the local galactic environment, and eccentricities can lead to a flatter spectrum at low frequencies in the PTA band, while the spectrum can also be steeper at high frequencies in the PTA band [492, 493]. Unfortunately, due to the absence of observations, the theoretical prediction of the GWB spectrum from SMBHBs is subject to large uncertainties.

To obtain priors on the amplitude and the spectral index of the power law given in Eq. (7.47), we follow Ref. [494] and base the spectral parameters on simulations of many SMBHB populations. We use the populations from the `GWOnly-Ext` library, which assumes SMBH binary evolution to be exclusively driven by the emission of gravitational radiation as well as circular orbits. The library was generated for Ref. [449] with the software package `holodeck` [495], which is based on semi-analytical models of SMBHB mergers. Each population in the library has an associated GWB spectrum, to which we fit the power law in Eq. (7.47). This provides us with a distribution of the power-law parameters $(\log_{10} A_{\mathrm{BHB}}, \gamma_{\mathrm{BHB}})$, which is well approximated by a two-dimensional Gaussian with mean $\boldsymbol{\mu}_{\mathrm{BHB}}$ and covariance matrix $\boldsymbol{\sigma}^2_{\mathrm{BHB}}$ given by

$$\boldsymbol{\mu}_{\mathrm{BHB}} = \begin{pmatrix} -15.6 \\ 4.7 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\sigma}^2_{\mathrm{BHB}} = \begin{pmatrix} 2.8 & -0.026 \\ -0.026 & 1.2 \end{pmatrix}\,. \tag{7.48}$$

We use this Gaussian distribution as the prior on $(\log_{10} A_{\rm BHB}, \gamma_{\rm BHB})$. This defines the reference model, which we label SMBHB.

Typical GWB spectra obtained from this distribution do not provide a very good fit to the NG15 signal, as their amplitudes are slightly too low and the spectra too steep.[5] It should be kept in mind that sufficiently large eccentricities as well as interactions between the SMBHBs and their galactic environment could flatten the spectrum. Nonetheless, they cannot significantly increase the amplitude. This requires at least one parameter of the binary population to deviate from expectations [449].

When we compare in the following string models to SMBHBs, the large uncertainties in the GWB prediction of both the SMBHB model as well as the string models should be kept in mind. The computed Bayes factors should not be interpreted as clearly favouring one physical origin over another one, but only as one type of spectrum fitting better than the other for the given prior choices.

7.3.1 Stable strings — STABLE

Let us start with the discussion of the GWB from local, topologically stable $U(1)$ strings, which are the central object of discussion of Pt. II of this thesis.

Model description: Let us start with the GWB spectra emitted by the fundamental oscillation mode $\Omega_{\rm GW}^{(1)}(f)$ of the loops. To compute it, we follow the numerical procedure laid out at the end of the introduction to Ch. 6, accounting for changes in the effective number of DOFs during radiation domination. In particular, we use the benchmark parameter values $\alpha \simeq 0.37$, $\tilde{c} \simeq 0.23$, $\Gamma \simeq 50$, and $\mathcal{F} \simeq 0.1$, as previously. To obtain the total GWB spectrum, we evaluate the sum

$$\Omega_{\rm GW}(f) = \sum_{j=1}^{n_{\max}} \frac{\Gamma_j}{\Gamma} \Omega_{\rm GW}^{(1)} \left(\frac{f}{j}\right) . \tag{7.49}$$

To account for our uncertainty in the mode-number dependence of Γ_j, we adopt four different models. We consider one model, STABLE-C, for which the GW emission is due to cusps, one model, STABLE-K, for which GW emission is due to kinks, one "monochromatic" model, STABLE-M, in which only the fundamental mode contributes, and one model, STABLE-N, based on the numerical simulations of Refs. [361, 496], including GWs from cusps and kinks (these models were also used in the analysis of Ref. [452]). Specifically, this means that for the first three models we set, as before (cf. Eq. (4.65)),

$$\Gamma_j = \frac{\Gamma}{H_{n_{\max}}^{(q)}} j^{-q} . \tag{7.50}$$

Our models can then be summarised as

- STABLE-C: $q = 4/3$ and $n_{\max} = 10^5$
- STABLE-K: $q = 5/3$ and $n_{\max} = 10^5$
- STABLE-M: $n_{\max} = 1$, or equivalently, $\Omega_{\rm GW}(f) = \Omega_{\rm GW}^{(1)}(f)$
- STABLE-N: Numerical values of Γ_j obtained from simulations in Refs. [361, 496]

The only free parameter that remains in our model is the string tension. For all stable string models, we adopt a uniform prior on $\log_{10} G\mu$:

$$\log_{10} G\mu \in [-14, -6] . \tag{7.51}$$

---

[5] Note that steepness here refers to $\rho(f)$, not to $\Omega_{\rm GW}(f)$.

Results and discussion: Let us now discuss the results of our Bayesian analysis. We begin with the Bayes factors obtained from comparing the string models against the reference model SMBHB. The results of the model comparison are summarised in Tab. 7.1. As we can see, the GWB signal from all of our four STABLE models fits the NG15 data slightly worse than that from the SMBHB model. Considering a combined signal STABLE+SMBHB from strings and SMBHBs leads to merely a minor improvement in the Bayes factors, which reach values just below 1. This is because adding the GWB from stable strings to that from SMBHBs does not significantly improve the fit, but increases the prior volume. This, in turn, reduces the Bayes factor. The reason that stable string models do not lead to a good fit is straightforward to visualise. In Fig. 7.2, we plot the median GWB spectra of the stable string models. The spectra match the amplitude of the NG15 signal in the lowest frequency bins well, but are too flat to also provide a good fit at higher frequencies. This effect can also be seen in Fig. 6.14a. For string tensions that could lead to a relatively good fit of the spectral shape, the amplitude is too low. For larger string tensions, it is possible to fit the amplitude, but the spectrum is too flat. The stable string spectra either have too low an amplitude or are not sufficiently blue-tilted to fit the NG15 data.

Table 7.1: Bayes factors from the comparison of the models STABLE and STABLE+SMBHB to the reference model SMBHB.

Model	Bayes factor STABLE only	Bayes factor STABLE+SMBHB
STABLE-C	0.277 ± 0.006	0.76 ± 0.01
STABLE-K	0.364 ± 0.008	0.89 ± 0.02
STABLE-M	0.379 ± 0.008	0.84 ± 0.02
STABLE-N	0.307 ± 0.006	0.83 ± 0.01

Next, we consider the reconstructed 1D marginalised posterior density for $\log_{10} G\mu$, shown in Fig. 7.3a. On the left, we see the result for the string models only. The distributions are nearly symmetrical and narrow, with maximum posterior values between $\log_{10} G\mu \sim -10.6$ and -10.1. For more precise values and an overview of different relevant quantities derived from the posterior distribution, see Tab. 7.2. The fall off of the posterior towards smaller and larger string tensions is for the reasons just discussed: in the former case, the spectra have too low amplitudes; in the latter case, they are too flat. It is not surprising that the values of the string tension that maximise the posterior follow the order they do in the plot. For the STABLE-M model, only the fundamental mode contributes, leading, at a given string tension, to the largest GWB amplitude in the nanohertz-band. For the other models, higher loop oscillation modes contribute. As a consequence, the power is redistributed from the lowest mode towards higher modes and thereby to larger frequencies (cf. Eq. (7.49)). This reduces the amplitude in the nanohertz range while also reducing the blue-tilt of the spectrum. For a given string tension, the amplitudes at frequencies relevant for NG15, decrease for the models in the order STABLE-K, STABLE-N, STABLE-C (see also Ref. [452]).

In Fig. 7.3b, we can also see the reconstructed $\log_{10} G\mu$-posterior distribution for a combined GWB signal from the string models and SMBHBs. While still clearly peaked near the same values as without an SMBHB contribution, the posteriors are noticeably asymmetric. In fact, they converge to a constant plateau for string tensions lower than $\log_{10} G\mu \sim -11.5$. As we saw in the plots in Fig. 7.3a, such low string tensions cannot explain the NG15 signal as the corresponding GWB amplitudes are too low. While this is still the case in the combined model, this time, SMBHBs can still give a signal that is sufficiently strong to prevent the posterior density from dropping to zero. For sufficiently low string tensions, the string contribution is completely negligible and the GWB only due to SMBHBs. The posterior becomes independent of $\log_{10} G\mu$ and hence flat. This flat region is exactly the plateau we discussed around Eq. (7.43), when we introduced the K

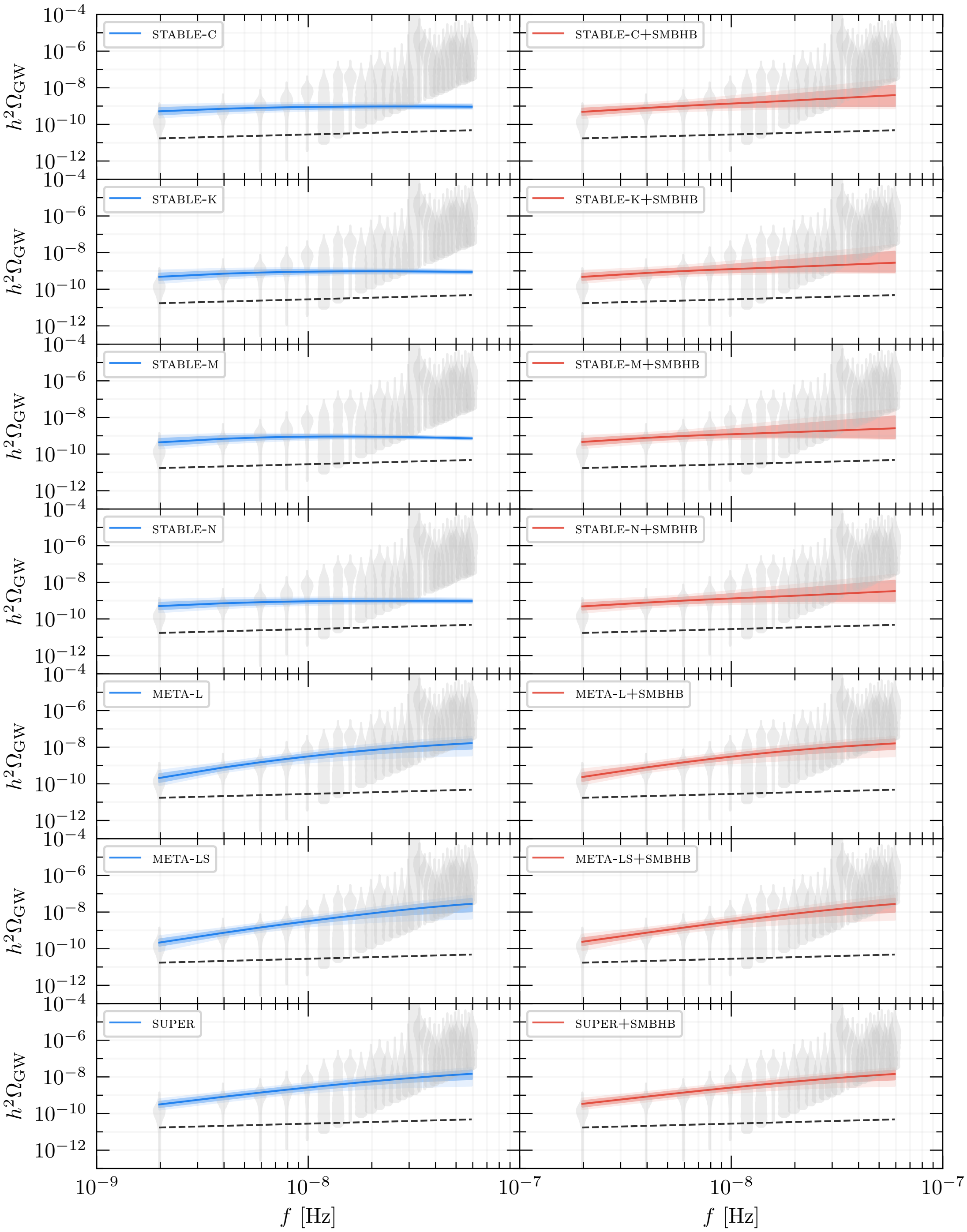
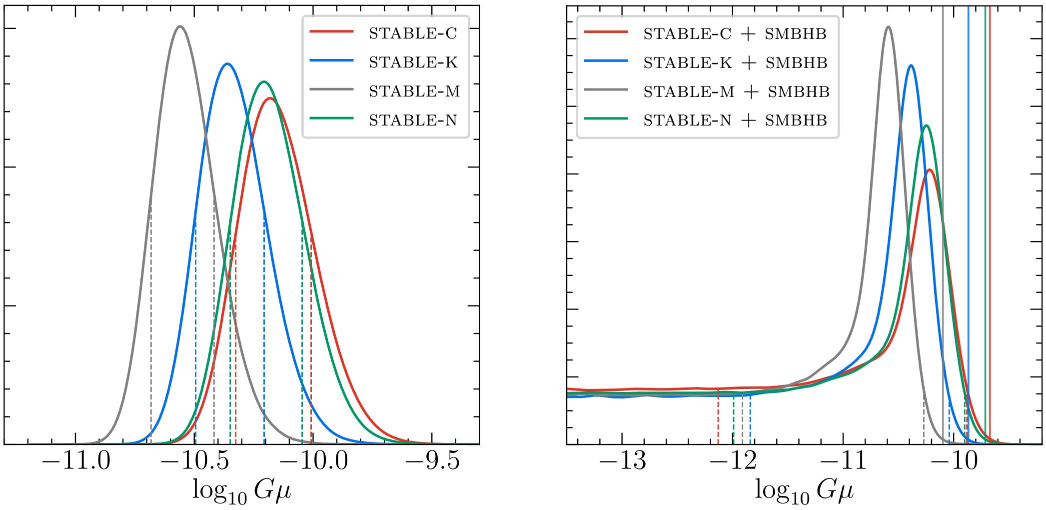


Figure 7.2: Median GWB spectra depicted as solid lines at the centre of shaded regions showing the 68% and 95% equal tailed posterior intervals. Plots are presented for all string models, both individually (left, blue) and in combination with a GWB from SMBHBs (right, red). The distribution of spectral values used to compute the median and intervals is obtained by sampling model parameter values from the posterior distribution. For each set of sampled model parameter values, we compute $h^2\Omega_{\rm GW}$. At a fixed frequency, we have then a sample of $h^2\Omega_{\rm GW}$ values induced from the reconstructed posterior distribution. From this distribution of $h^2\Omega_{\rm GW}$ values, we compute the median and the equal-tailed C.I. To generate the plots, we repeat this procedure at many frequencies between $1/T_{\rm obs}$ and $30/T_{\rm obs}$. Note that, in general, no individual GWB spectrum coincides exactly with the median GWB spectrum that is depicted, although the difference can be marginal. Furthermore, in grey, we show violin plots that depict the one-dimensional marginalised posterior density of $h^2\Omega_{\rm GW}(f_i)$ at frequencies $f_i = i/T_{\rm obs}$ for $i \in \{1, \ldots, 30\}$. As a black dotted line, we also plot the GWB spectrum from SMBHBs described by the power law in Eq. (7.47), using the mean parameters $\boldsymbol{\mu}_{\rm bhb}$ of the two-dimensional Gaussian prior given in Eq. (7.48).

Figure reproduced from Ref. [1].



(a) **(b)**

Figure 7.3: Reconstructed posterior densities for the string-tension parameter $\log_{10} G\mu$. **Left:** Posteriors for the four models with GWB from stable strings only. **Right:** Posteriors for the four models with combined GWB from stable strings and SMBHBs. 68% HPDIs are shown as dashed vertical lines, K bounds are depicted as solid vertical lines. *Figure reproduced from Ref. [1].*

ratio. $G\mu$ values from this plateau region correspond to what we referred to as "θ_0". For string tensions far above the maximum posterior values, there is no such plateau. This is expected as in this region the posterior drops off due to an excessively large amplitude of the GWB signal from strings. Adding SMBHBs cannot reduce this amplitude and therefore does not improve the fit. We can compute the K bound as the value of the string tension at which the posterior has dropped to a tenth of the plateau height. In practice, we use the equivalent procedure outlined in the previous section and find, among the four stable string models, the weakest K bound with the value $\log_{10} G\mu = -9.67$. The precise K bounds for all models are reported in Tab. 7.2 and shown as vertical lines in Fig. 7.3b. Note that, in the absence of the SMBHB signal, there cannot be a plateau region, which is why we can compute the K bound only for the STABLE+SMBHB models but not for the STABLE-only models. We can also derive upper limits from the 95% HPDIs.[6] For STABLE-C+SMBHB, STABLE-K+SMBHB, STABLE-M+SMBHB, and STABLE-N+SMBHB, the 95% HPDIs yield upper bounds of $\log_{10} G\mu < -9.88$, -10.04, -10.26, and -9.90, respectively. This agrees well with the derived values of the K bounds. Additional quantities derived from the posterior density of the STABLE+SMBHB models can be found in Tab. 7.2.

7.3.2 Metastable strings — META

Model description: Next, we turn to metastable strings. Recall from Sec. 2.3, and specifically from Sec. 2.3.2, that metastable strings can unwind through the formation of monopole–antimonopole

---

[6] At this point, we need to make a technical remark. For the STABLE+SMBHB models, both the 68% and 95% HPDIs have lower bounds that fall within the plateau region. For a perfectly flat plateau, this leads to an ambiguously defined HPDI. Instead of including a certain portion of the plateau into the HPDI, we could include another portion of the same width, and obtain a different, but equally valid HPDI. In reality, due to a finite number of samples, the plateau region will exhibit fluctuations, leading to a set of disconnected intervals. To avoid this issue, we restrict to connected intervals and choose the upper bound of the HPDI to the highest string tension at which the posterior reaches the height of the plateau. We then adjust the lower bound such that integrating the posterior over the interval yields 0.68 and 0.95, respectively.

Table 7.2: Various quantities derived from the $1D$ marginalised posterior distributions for the models STABLE and STABLE+SMBHB. K bounds can only be computed for the latter case.

Model	Parameter	Bayes Estimator	Max. Posterior	68% HPDI	K Bound
STABLE only					
STABLE-C	$\log_{10} G\mu$	-10.15 ± 0.16	-10.18	$[-10.33, -10.01]$	—
STABLE-K	$\log_{10} G\mu$	-10.33 ± 0.15	-10.36	$[-10.50, -10.21]$	—
STABLE-M	$\log_{10} G\mu$	-10.53 ± 0.14	-10.56	$[-10.68, -10.42]$	—
STABLE-N	$\log_{10} G\mu$	-10.18 ± 0.15	-10.21	$[-10.35, -10.05]$	—
STABLE+SMBHB					
STABLE-C	$\log_{10} G\mu$	-11.41 ± 1.25	-10.22	$[-12.13, -9.88]$	-9.67
STABLE-K	$\log_{10} G\mu$	-11.34 ± 1.17	-10.38	$[-11.84, -10.04]$	-9.87
STABLE-M	$\log_{10} G\mu$	-11.47 ± 1.09	-10.60	$[-11.91, -10.27]$	-10.10
STABLE-N	$\log_{10} G\mu$	-11.34 ± 1.23	-10.25	$[-11.99, -9.90]$	-9.71

pairs. The rate of pair nucleations per unit time, i.e., the decay rate, reads

$$\Gamma_d = \frac{\mu}{2\pi} e^{-\pi\kappa}\,, \tag{7.52}$$

where the parameter $\sqrt{\kappa}$ can be interpreted as the ratio of monopole mass m_{M} to the square root of the string tension μ,

$$\sqrt{\kappa} = \frac{m_{\mathrm{M}}}{\sqrt{\mu}}\,. \tag{7.53}$$

Due to this decay, string loops break apart and form string segments with monopoles at their ends. As discussed in Sec. 2.3.2, the flux of the monopoles can either be fully confined in the strings or remain partially unconfined. In the latter case, the string segments are expected to primarily evaporate into massless gauge bosons. Correspondingly, we neglect their emission of gravitational radiation. In the former case there are no massless vector bosons left after the production of strings and monopoles. Segments will then mainly decay gravitationally. We can then distinguish two models of metastable strings:

- META-L: Monopoles carry an unconfined flux. We neglect GW emission from segments and consider emission from loops only.
- META-LS: The monopole flux is fully confined in the strings and we consider GW emission from loops and segments.

In both models we compute the GWB following Refs. [213] and [214]. The interested reader is referred to these articles for more details.

We can compute the GWB in the same way as we did for the STABLE models. We perform the computation for loops and segments separately by replacing the loop number density $n(l,t)$ given in Eq. (4.72) via

$$n(l,t) \to n^{(\mathrm{l})}_{\mathrm{meta}}(l,t) \qquad \text{and} \qquad n(l,t) \to n^{(\mathrm{s})}_{\mathrm{meta}}(l,t)\,, \tag{7.54}$$

respectively. For META-L, we work then with the contribution from loops only, for META-LS, we add the loop and segment GWB spectra.

Let us first describe the contribution from metastable-string loops, which is common to both models. The relevant loop number density $n^{(\mathrm{l})}_{\mathrm{meta}}(l,t)$ can be obtained from $n(l,t)$ given in Eq. (4.72), by modifying it with two prefactors:

$$n^{(\mathrm{l})}_{\mathrm{meta}}(l,t) = \Theta(t_s - t_*)E(l,t)n(l,t)\,. \tag{7.55}$$

Here, t_* is, as before, the time of loop production, implicitly given as the solution to Eq. (4.73), and $t_s = \Gamma_d^{-1/2}$ is the time when the string loop decay becomes efficient, the network leaves the scaling regime and enters the decay regime. The Heaviside function takes into account that loop production effectively stops at the time t_s. The other factor is of the form

$$E(l,t) = \exp\left\{-\Gamma_d\left(l(t-t_*) + \frac{1}{2}\Gamma G\mu(t-t_*)^2\right)\right\}\,, \tag{7.56}$$

and accounts for the decay of loops in the network, leading to an exponential suppression.

The decay rate in the exponent can be expressed in terms of $\sqrt{\kappa}$ using Eq. (7.52). We see that the loop number density in Eq. (7.55) is extremely sensitive to this parameter. On the one hand, for $\sqrt{\kappa} \gg 10$ the metastable strings can be treated as quasi-stable, since the network lifetime is longer than the age of Universe. The corresponding GWB spectra will equal those from stable strings. On the other hand, for $\sqrt{\kappa} \sim 1$, the network decays quickly after formation and does not lead to a GWB signal in the PTA band. The non-trivial aspects of a GWB from metastable strings can be captured by choosing the two free parameters of our models, $\log_{10} G\mu$ and $\sqrt{\kappa}$, from uniform priors with ranges

$$\log_{10} G\mu \in [-14, -1.5] \qquad \text{and} \qquad \sqrt{\kappa} \in [7, 9.5]\,. \tag{7.57}$$

Observe that the upper bound on $\log_{10} G\mu$ is chosen larger than for the stable string models. This is because sufficiently strong loop decay reduces the GWB amplitude at nanohertz frequencies and must be compensated by a larger string tension. For all other parameters, we use the same values as we did for stable strings. For the distribution of GW power among different harmonic modes on loops, we use Γ_j as given in Eq. (7.50) with $q = 4/3$.

Let us now turn to the GWB contribution from segments, which is relevant for the model META-LS. For this contribution, we use Γ_j given in Eq. (7.50) with $q = 1$, which amounts to $\Gamma_j \simeq 4/j$. This dependence was derived in Ref. [212], approximating string segments as straight. Let us now summarise the expressions for $n_{\mathrm{meta}}^{(\mathrm{s})}$ found in Refs. [213, 214]. The number density of segments can be decomposed into different contributions.

We start by considering segments produced in the decay of long strings, associated with the loop number density

$$\bar{n}^{(\mathrm{s})}(l,t) = \bar{n}_{\mathrm{rr}}^{(\mathrm{s})}(l,t) + \bar{n}_{\mathrm{rm}}^{(\mathrm{s})}(l,t) + \bar{n}_{\mathrm{mm}}^{(\mathrm{s})}(l,t)\ . \tag{7.58}$$

The first two terms contribute only if the decay regime is entered during the radiation-dominated era, i.e., $t_s < t_{\mathrm{eq}}$. Segments forming from long strings are then associated with the number density

$$\bar{n}_{\mathrm{rr}}^{(\mathrm{s})}(l,t) = \Theta(t_{\mathrm{eq}} - t)\,\Theta(t - t_s)\,\frac{\Gamma_d^2}{\xi_r^2}\frac{(t+t_s)^2}{(t^3 t_s)^{1/2}}e^{-\Gamma_d\left(l(t+t_s)+\frac{1}{2}\Gamma G\mu(t-t_s)(t+3t_s)\right)}\,, \tag{7.59}$$

if the evaluation time t is during radiation domination and with

$$\bar{n}_{\mathrm{rm}}^{(\mathrm{s})}(l,t) = \Theta(t - t_{\mathrm{eq}})\,\Theta(t_{\mathrm{eq}} - t_s)\,\frac{\Gamma_d^2}{\xi_r^2}\left(\frac{t_{\mathrm{eq}}}{t}\right)^2\frac{(t+t_s)^2}{(t_{\mathrm{eq}}^3 t_s)^{1/2}}e^{-\Gamma_d\left(l(t+t_s)+\frac{1}{2}\Gamma G\mu(t-t_s)(t+3t_s)\right)}\,, \tag{7.60}$$

if t falls into the matter-dominated epoch. In contrast, the third term in Eq. (7.58) is non-vanishing only if the decay regime is entered during matter domination, i.e., $t_s > t_{\mathrm{eq}}$. The associated string segment density is then non-vanishing only if evaluated for times in the matter-dominated era with

$$\bar{n}_{\mathrm{mm}}^{(\mathrm{s})}(l,t) = \Theta(t - t_s)\,\Theta(t_s - t_{\mathrm{eq}})\,\frac{\Gamma_d^2}{\xi_m^2}e^{-\Gamma_d\left(lt+\frac{1}{2}\Gamma G\mu(t-t_s)(t+t_s)\right)}\,. \tag{7.61}$$

Working under the simplifying assumption that loops and segments emit gravitational radiation with equal efficiency, we also set here $\Gamma = 50$. For the reduced correlation lengths, we use the values of the attractor solution during matter and radiation domination given in Eq. (4.55), namely $\xi_r = 0.27$ and $\xi_m = 0.63$, respectively.

Segments can not only arise from long strings but also from sub-Hubble loops. Let us first focus on segments that arise directly from loops breaking apart, not from the subsequent breaking of segments into shorter segments. As before, we can write the associated segment number density as

$$\tilde{n}_1^{(\mathrm{s})}\left(l,t\right) = \tilde{n}_{\mathrm{rr},1}^{(\mathrm{s})}\left(l,t\right) + \tilde{n}_{\mathrm{rm},1}^{(\mathrm{s})}\left(l,t\right) + \tilde{n}_{\mathrm{mm},1}^{(\mathrm{s})}\left(l,t\right) \,. \tag{7.62}$$

All three contributions are non-vanishing only if evaluated at times after the decay regime is entered $t > t_s$. Segments formed during radiation domination are associated with the number density

$$\tilde{n}_{\mathrm{rr},1}^{(\mathrm{s})}\left(l,t\right) = \Theta\left(t-t_s\right)\Theta\left(t_{\mathrm{eq}}-t\right)\Theta\left(t_{\mathrm{eq}}-t_*\right)\Gamma_d\left(l\left(t-t_s\right)+\frac{1}{2}\Gamma G\mu\left(t-t_s\right)^2\right)n_{\mathrm{meta}}^{(\mathrm{l})}(l,t) \tag{7.63}$$

if the evaluation time t falls into the radiation-dominated era ($t < t_{\mathrm{eq}}$), and with

$$\tilde{n}_{\mathrm{rm},1}^{(\mathrm{s})}\left(l,t\right) = \Theta\left(t-t_s\right)\Theta\left(t-t_{\mathrm{eq}}\right)\Theta\left(t_{\mathrm{eq}}-t_*\right)\Gamma_d\left(l\left(t-t_s\right)+\frac{1}{2}\Gamma G\mu\left(t-t_s\right)^2\right)n_{\mathrm{meta}}^{(\mathrm{l})}(l,t)\,, \tag{7.64}$$

if evaluated during matter domination ($t > t_{\mathrm{eq}}$). The third contribution is from segments formed during the matter-dominated epoch

$$\begin{aligned}\tilde{n}_{\mathrm{mm},1}^{(\mathrm{s})}\left(l,t\right) = {}& \Theta\left(t-t_s\right)\Theta\left(t_s-t_*\right)\Theta\left(t_*-t_{\mathrm{eq}}\right)\frac{\Gamma_d e^{-\Gamma_d\left(l(t-t_*)+\frac{1}{2}\Gamma G\mu(t-t_*)^2\right)}}{t^2\left(l+\Gamma G\mu t\right)^2} \\ & \times\left\{0.27\left(l\left(t-t_s\right)+\frac{\Gamma G\mu\left(t-t_s\right)^2}{2}\right)+0.45\left(l+\Gamma G\mu t\right)^{1+0.31}\left(F_2\left(t\right)-F_1\left(t\right)-F_2\left(t_s\right)+F_1\left(t_s\right)\right)\right\}.\end{aligned} \tag{7.65}$$

Here, F_2 and F_1 denote the functions

$$F_n\left(x\right) = {}_2F_1\left(n-\beta,-\beta;n+1-\beta;\frac{\Gamma G\mu}{l+\Gamma G\mu\, t}\,x\right)\left(\frac{\Gamma G\mu}{l+\Gamma G\mu\, t}\right)^{n-1}\frac{x^{n-\beta}}{n-\beta}\,, \tag{7.66}$$

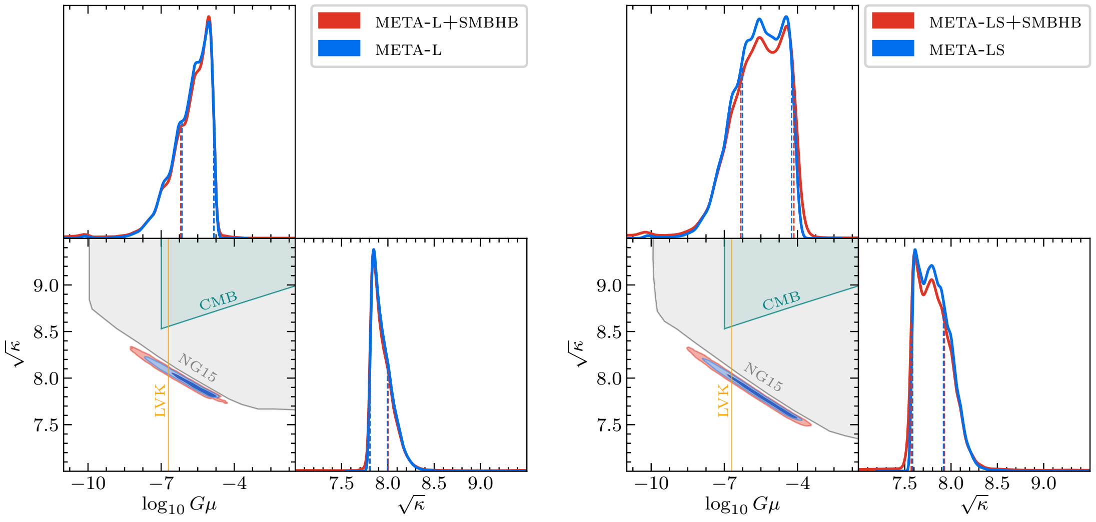
where, as before, ${}_2F_1$ is the Gaussian hypergeometric function and $\beta = 0.31$. The numerical factors 0.27, 0.45, and 0.31 in Eqs. (7.65) and (7.66) occur, because these results, derived in [214], are based on loop number densities from the BOS model [47].

Of course, strings can fragment multiple times. In a first segment generation, we obtain segments directly from loops. The corresponding segment number densities are those given above. In a second generation, the segments from the first generation break apart, in the third generation, the segments from the second generation break apart and so on. While it is possible to compute the total segment number density from all generations (see Ref. [214]), it is numerically demanding. For our data analysis, we follow instead another approach and replace the total segment number densities taking all segment generations into account, by the first-generation segment number densities, rescaled with a fudge factor $\mathcal{C}_f \sim \mathcal{O}\left(10\right)$:

$$\tilde{n}_{\mathrm{rr}}^{(\mathrm{s})}\left(l,t\right)\to\mathcal{C}_f\tilde{n}_{\mathrm{rr},1}^{(\mathrm{s})}\left(l,t\right)\,,\quad \tilde{n}_{\mathrm{rm}}^{(\mathrm{s})}\left(l,t\right)\to\mathcal{C}_f\tilde{n}_{\mathrm{rm},1}^{(\mathrm{s})}\left(l,t\right)\,,\quad \tilde{n}_{\mathrm{mm}}^{(\mathrm{s})}\left(l,t\right)\to\mathcal{C}_f\tilde{n}_{\mathrm{mm},1}^{(\mathrm{s})}\left(l,t\right)\,. \tag{7.67}$$

As shown in Ref. [214], while the dependence of the total and first-generation segment number densities on l and t is different, the above replacement yields comparable GWB spectra. For our computation, we specify $\mathcal{C}_f = 5$, which provides an appropriate value for large regions of parameter space. The number density of all segments is then

$$n_{\mathrm{meta}}^{(\mathrm{s})}(l,t) = \bar{n}^{(\mathrm{s})}(l,t) + \mathcal{C}_f\tilde{n}_1^{(\mathrm{s})}(l,t)\,. \tag{7.68}$$



(a) **(b)**

Figure 7.4: Corner plots showing the reconstructed $2D$ and $1D$ marginalised posterior densities for the metastable string parameters $\log_{10} G\mu$ and $\sqrt{\kappa}$. Results for the GWB from META only are shown in blue, results for the combined GWB from META+SMBHB are shown in red. **Left:** Results for META-L, corresponding to GW emission from loops only. **Right:** Results for META-LS, corresponding to GW emission from loops and segments. 68% HPDIs of the one-dimensional posteriors are depicted as dashed vertical lines. For the two-dimensional posteriors, 68% and 95% highest posterior density regions are shown with a darker and lighter blue or red shading, respectively. The grey shaded regions are excluded by the K bound, corresponding to a K ratio of less than $1/10$ (see Eqs. (7.43), (7.44)). The teal shaded regions are excluded by the CMB bound [51, 392, 393]. Regions corresponding to string tensions larger than indicated by the yellow lines violate the LVK bound $G\mu \lesssim 2 \times 10^{-7}$ [497].
Figure reproduced from Ref. [1].

Results and discussion: With the definitions of our two metastable string models META-L and META-LS in place, we are ready to discuss the results of the Bayesian analysis. We start again with the Bayes factors obtained from comparing the metastable string models to the reference model SMBHB. The results are shown in Tab. 7.3. From these values we conclude that, according to the Jeffreys scale, the NG15 data strongly favour the metastable string models over the SMBHB reference model. The transitivity of Bayes factors implies that the metastable string models also yield a much better fit to the data than any of the stable string models. Adding a GWB signal from SMBHBs to that of metastable strings does not improve the fit to the data, but increases the prior volume and, hence, reduces the Bayes factors.

Table 7.3: Bayes factors from the comparison of the models META and META+SMBHB to the reference model SMBHB.

Model	Bayes factor META only	Bayes factor META+SMBHB
META-L	13.4 ± 0.4	11.1 ± 0.3
META-LS	21.3 ± 0.8	18.9 ± 0.7

The physical reason for the better fit of the metastable strings can be found by considering the reconstructed posterior distributions depicted in Fig. 7.4. For both metastable string models, whether overlaid with an SMBHB contribution or not, the posteriors prefer values of the decay parameter $\sqrt{\kappa} \sim 8$ and string tensions $\log_{10} G\mu \sim -7 \cdots -4$, significantly larger than the tensions preferred by the stable string models. For more precise values obtained from the posteriors, see Tab. 7.4. As touched upon before, this is because the decay of metastable string loops leads, due to the suppression factor in Eq. (7.56), to a reduced GWB amplitude at nanohertz frequencies. The suppression at low frequencies is stronger than at higher frequencies, such that the spectrum is more blue-tilted in the PTA band than the one of stable strings. The reduced amplitude can then be compensated by a larger string tension, which gives an overall GWB signal that better fits the NG15 data. This is also visible in Fig. 7.2 when comparing the median GWB spectra of metastable strings to those of stable strings. As discussed earlier, stable strings can be obtained from metastable strings in the limit of large $\sqrt{\kappa}$. For $\sqrt{\kappa} \gtrsim 9$, the effects of metastable string decay are only relevant at frequencies below the PTA band. And indeed, when considering the 1D posterior of $\log_{10} G\mu$, we see for both metastable string models a very small bump at string tensions that lie around the maximum posterior value of the stable string models, slightly below $\log_{10} G\mu \sim -10$.

Let us also compare the posteriors of the models META-L and META-LS with each other. We first observe that the $\log_{10} G\mu$ posterior of the META-L model is cut off at tensions around $\log_{10} G\mu \sim -5$, while for the META-LS model, we only see a small dip at this tension, and then a secondary peak before the posterior also steeply falls off at $\log_{10} G\mu \sim -4$. Corresponding features are also seen in the $\sqrt{\kappa}$ posterior, which for the META-L model, has a single peak, and, for the META-LS model, a double peak structure. The posterior for META-L sharply drops at $\sqrt{\kappa} \sim 7.8$, whereas the posterior of the META-LS model extends to lower values of $\sqrt{\kappa} \sim 7.5$. Note that features corresponding to large values in the $\log_{10} G\mu$ posterior are associated with small values in the $\sqrt{\kappa}$ posterior. This can be seen from the 2D posterior plot and is clear from our description of the interplay between amplitude enhancement and suppression, as a consequence of larger string tensions and loop fragmentation, respectively. The difference between the 1D posteriors of the metastable string models can be understood as follows. The sharp drop-off in the $\sqrt{\kappa}$ posterior of the META-L model and the dip in the posterior of the META-LS model are both caused by the Heaviside function in Eq. (7.55), which sets loop formation during the decay regime to zero. Since $t_s \sim \Gamma_d^{-1/2} \sim e^{\pi\kappa/2}$, we can see that lower values of $\sqrt{\kappa}$ result in an earlier end of loop production, shifting the low-frequency cutoff in the GWB spectrum to higher frequencies. For values $\sqrt{\kappa} \lesssim 7.8$, this pushes the GWB spectrum from loops out of the PTA band. For the META-LS model, however, there is still the GWB contribution from segments. This second contribution can explain the NG15 signal in parameter regions where the signal from loops is insufficient, and is also responsible for the second peak in the posterior. In reality, the transition between scaling and decay regime will be smooth and not as abrupt as the simplified modelling with a Heaviside function suggests. The sharpness of the drop-off in the posteriors of the META-L model and the dip in those of the META-LS model should be regarded as artefacts of this simplification of the theoretical model. For completeness, let us also remark on the small fluctuations that are visible in the $1D$ marginalised posteriors. These are not due to undersampling or a KDE bandwidth that was chosen too narrowly, but rather a consequence of the grid size that was used for the tabulated data of the metastable string GWB spectra.

Finally, we discuss the observational bounds on the metastable string parameter space as shown in Fig. 7.4. First is the CMB bound. For a metastable string network that has not decayed before recombination, this bound excludes string tensions $\log_{10} G\mu \gtrsim -7$ [51, 392, 393]. We estimate that the metastable string network has completely decayed when the second term in the exponential suppression factor in Eq. (7.56) becomes large. This happens at

$$t_e \simeq \sqrt{\frac{2}{\Gamma G\mu}} t_s \,. \tag{7.69}$$

Accordingly, the bound excludes regions of parameter space satisfying both $\log_{10} G\mu \gtrsim -7$ and $t_e \gtrsim t_{\rm rec} \simeq 3.7 \times 10^5\,\text{yr}$. This CMB-excluded region is shown in teal in Fig. 7.4. We also present the K bound, which excludes the grey-shaded region labelled "NG15" from the parameter space. The NG15 data visibly provide a stronger bound of the parameter space than the CMB bound. In addition, we also plot a yellow vertical line corresponding to the bound by LIGO, Virgo and

KAGRA (LVK) that excludes $\Omega_{\rm GW} \geq 1.7 \times 10^{-8}$ at a frequency of $f_{\rm LVK} = 25\,{\rm Hz}$ [497], and translates into an excluded region of $G\mu \gtrsim 2 \times 10^{-7}$. We see that this bound rules out large parts of the parameter region preferred by the posterior, particularly the 68% highest posterior density regions. However, the LVK constraint does not affect the part of parameter space around $\log_{10} G\mu \sim -7$ and $\sqrt{\kappa} \sim 8$, where the GWB spectra from both metastable string models still provide a good fit to the NG15 data. If the NG15 signal was indeed a GWB signal from such metastable strings, ground-based interferometers should be able to detect it in the foreseeable future. The entire discussion of the LVK bound relies, however, on the assumption that our model correctly predicts the amplitude of the GWB spectrum all the way from the PTA band up to the frequencies of ground-based interferometers. This frequency range spans approximately ten orders of magnitude. In particular at these high frequencies, the GWB from strings depends on the scale factor evolution at times much before Big-Bang nucleosynthesis. At such early times, the expansion of the Universe is rather unconstrained. Deviations from standard Big Bang cosmology, such as an early matter-dominated phase, could suppress the GWB from strings at high frequencies (see Ref. [498]), rendering the LVK bound unproblematic.

Table 7.4: Various quantities derived from the $1D$ marginalised posterior distributions for the models META and META+SMBHB. K bounds can only be computed for the latter case. A dash indicates that the K bound is not reached within the prior range.

Model	Parameter	Bayes Estimator	Max. Posterior	68% HPDI	K Bound
META only					
META-L	$\log_{10} G\mu$	-5.80 ± 0.78	-5.04	$[-6.14, -4.84]$	—
	$\sqrt{\kappa}$	7.95 ± 0.13	7.85	$[7.81, 8.00]$	—
META-LS	$\log_{10} G\mu$	-5.62 ± 1.01	-4.46	$[-6.26, -4.24]$	—
	$\sqrt{\kappa}$	7.83 ± 0.18	7.61	$[7.59, 7.92]$	—
META+SMBHB					
META-L	$\log_{10} G\mu$	-5.90 ± 1.20	-5.05	$[-6.19, -4.83]$	—
	$\sqrt{\kappa}$	7.95 ± 0.18	7.84	$[7.80, 8.00]$	—
META-LS	$\log_{10} G\mu$	-5.70 ± 1.40	-4.44	$[-6.33, -4.15]$	—
	$\sqrt{\kappa}$	7.82 ± 0.23	7.61	$[7.57, 7.93]$	—

7.3.3 Superstrings — SUPER

Model description: The final model we want to consider describes the GWB from cosmic superstrings. These strings are stable but do not emerge from field theory, unlike the strings of the STABLE models. Instead, they are fundamental strings or one-dimensional D-branes of cosmological size that can emerge at the end of brane inflation [424, 425]. For cosmic superstrings, we have, in addition to the string tension, another parameter: the intercommutation probability P. This probability can be significantly smaller than unity (cf. Sec. 4.1.1), which results from the quantum nature of superstrings and the possibility that they may miss each other in the compact extra dimensions. Compared to field-theory strings, the reduction in intercommutation events necessitates much longer network simulations, which have not been performed yet. This leads to large uncertainties in the prediction of the GWB from cosmic superstrings. We follow here the simplest approach, which amounts to rescaling the cosmic string spectrum that is obtained in the case of stable cosmic strings by a prefactor $1/P$:

$$\Omega_{\rm GW}^{\rm super}(f) = \frac{1}{P}\Omega_{\rm GW}^{\rm stable}(f)\,. \tag{7.70}$$

The scaling of $\Omega_{\rm GW}^{\rm super}$ with the intercommutation probability is implied by the VOS model. The loop chopping efficiency scales with the intercommutation probability as $\tilde{c} \propto P^{1/2}$, which results in the universal length scale of the VOS model satisfying $L \propto P^{1/2}$. The reduction of the inter-string distance causes an increase in the energy density of long strings $\rho_\infty = \mu/L^2 \propto P^{-1}$. This scaling carries through to the GWB spectrum. For details, see Refs. [427, 428, 499], which also discuss more sophisticated models of the GWB spectrum from cosmic superstrings. In the PTA band, these are compatible with the simplified model used here.

We choose uniform priors on $\log_{10} G\mu$ and $\log_{10} P$:

$$\log_{10} G\mu \in [-14, -6] \qquad \text{and} \qquad \log_{10} P \in [-3, 0]\,. \tag{7.71}$$

The latter choice is motivated by the theoretically expected range of intercommutation probabilities for cosmic F- and D-superstrings [355]. We obtain the GWB spectra for the SUPER model by rescaling the GWB spectra corresponding to the STABLE-C model specified in Sec. 7.3.1 according to Eq. (7.70).

Results and discussion: Let us now discuss the results of the Bayesian analysis for the superstring model SUPER. We find that the SUPER model yields the best fit to the NG15 data among all string models, as can be seen from the Bayes factors reported in Tab. 7.5. The Bayes factor for the SUPER+SMBHB model is lower than that for the SUPER model since the addition of a GWB from SMBHBs does not significantly improve the fit while increasing the prior volume. Based on the discussion of the stable string models, it is easy to understand why the superstring model leads to a good fit. Recall that for small string tensions $\log_{10} G\mu \lesssim -11$, the GWB spectra of the STABLE models are sufficiently blue-tilted to provide a good fit to the NG15 data. This blue tilt can also be seen in Fig. 6.14a. For ordinary stable strings, such low string tensions result in an amplitude that is too low to explain the NG15 signal. However, for the GWB in the SUPER model, this is not a problem as the amplitude of the spectrum can be pushed to an appropriately large value by rescaling the spectrum with a sufficiently low intercommutation probability. The resulting spectrum is neither too flat nor too weak to provide a good fit to the NG15 data. This can also be seen from the median GWB spectra of the SUPER model in Fig. 7.2, although the improvement over the META models is hardly noticeable by eye.

Table 7.5: Bayes factors from the comparison of the models SUPER and SUPER+SMBHB to the reference model SMBHB.

Model	Bayes factor
SUPER only	46 ± 2
SUPER+SMBHB	37 ± 2

Let us next describe the reconstructed posterior densities for the superstring parameters. These are shown in the corner plot in Fig. 7.5. From the two-dimensional posterior densities, we can clearly see the described interplay of string tension and intercommutation probability when considering the 68% and 95% highest posterior density regions. The nearly diagonal trend indicates that a reduction of $\log_{10} G\mu$ can be compensated by a reduction of the intercommutation probability. Regarding the one-dimensional $\log_{10} P$-posterior, we see that the posterior density is maximised for the lowest intercommutation probability $\log_{10} P = -3$, corresponding to the lower boundary of the prior range. This and other quantities derived from the $1D$ marginalised posterior densities are listed in Tab. 7.6.

From Fig. 7.5, we see that the highest $2D$ posterior density is reached when both $\log_{10} P = -3$ and $\log_{10} G\mu \sim -12$. However, as the LVK bound plotted in yellow indicates, this best-fit region is excluded from parameter space. As discussed in more detail in the previous section on the metastable string models, this constraint assumes a standard scale factor evolution at times far before Big-Bang nucleosynthesis. Deviations in the expansion history can suppress the GWB

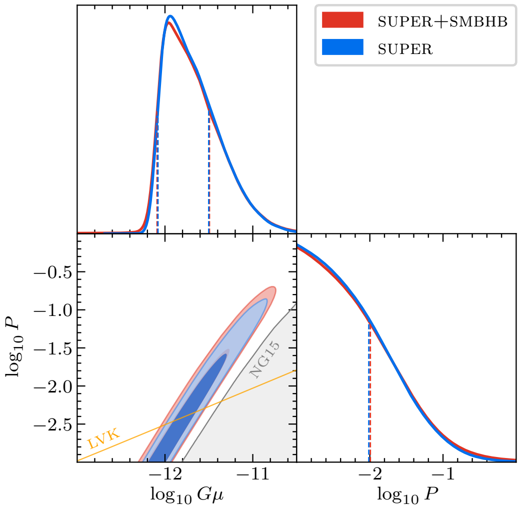


Figure 7.5: Same as in Fig. 7.4, but for the SUPER model.
Figure reproduced from Ref. [1].

spectra of superstrings at high frequencies and thus evade the LVK bound. While the SUPER GWB spectra are, of course, constrained by CMB observations [51], these bounds are too weak to show in the considered part of parameter space. As before, in Fig. 7.5, we also show the parameter region where the K ratio drops below $1/10$ as a grey shaded region. The NG15 data therefore exclude a significant part of the parameter space of the SUPER model. Given the large modelling uncertainties, all these statements refer strictly to the SUPER model, not necessarily to actual superstrings.

Table 7.6: Various quantities derived from the $1D$ marginalised posterior distributions for the models META and META+SMBHB. K bounds can only be computed for the latter case. A dash indicates that the K bound is not reached inside the prior range. An asterisk indicates that a value lies at an upper or lower bound of the prior range.

Model	Parameter	Bayes Estimator	Max. Posterior	68% HPDI	K Bound
SUPER only	$\log_{10} G\mu$	-11.67 ± 0.32	-11.94	$[-12.08, -11.50]$	—
	$\log_{10} P$	-2.23 ± 0.55	-3^*	$[-3^*, -2.01]$	—
SUPER+SMBHB	$\log_{10} G\mu$	-11.68 ± 0.35	-11.96	$[-12.09, -11.49]$	-9.97
	$\log_{10} P$	-2.21 ± 0.57	-3^*	$[-3^*, -1.99]$	—

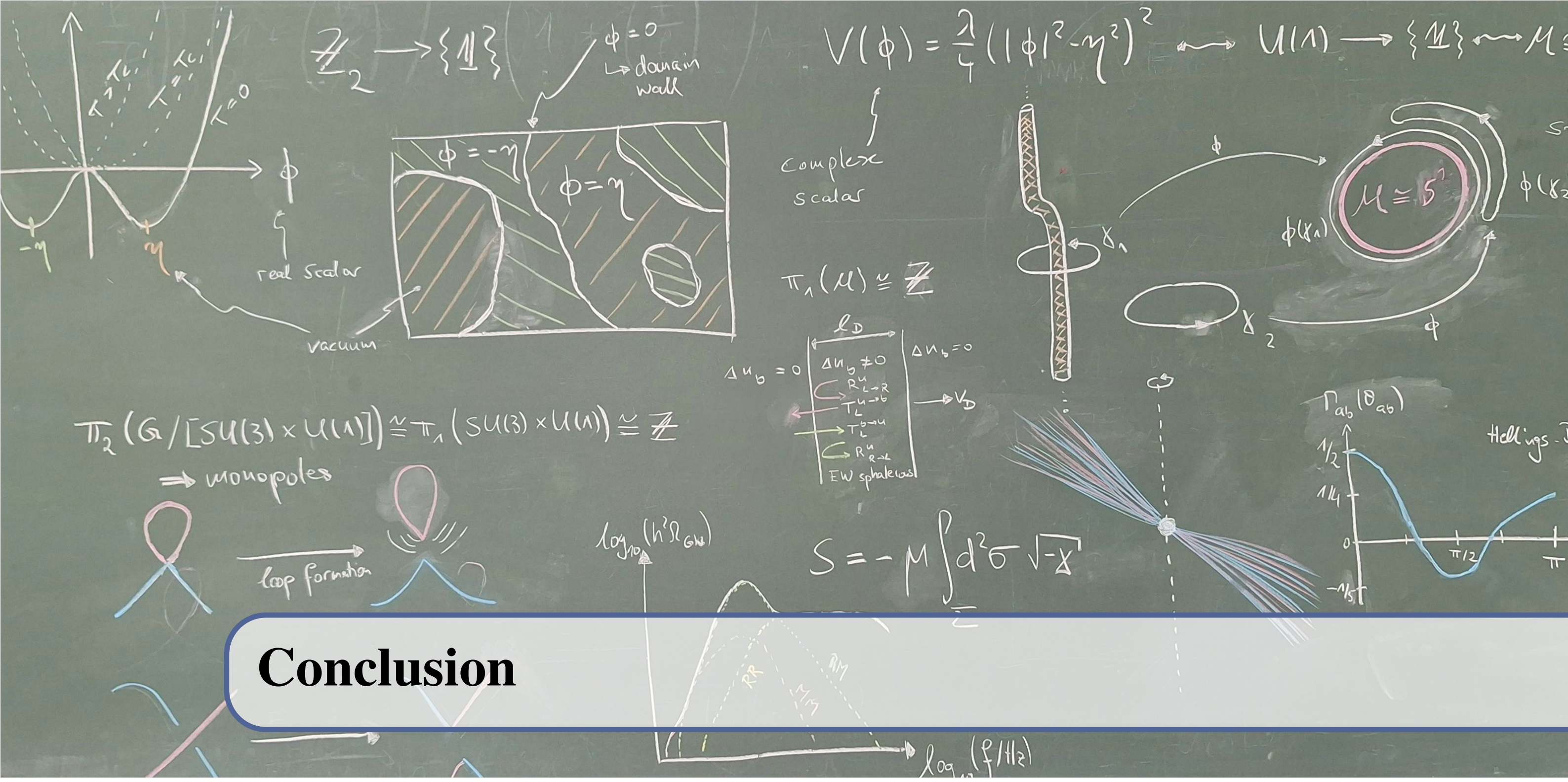
7.4 Discussion

Millisecond pulsars are rapidly spinning neutron stars that emit radio beams, which can be observed as quasi-periodic pulses in radio telescopes on Earth. GWs passing between the pulsar and the Earth lead to slight deviations in the travel time of the radio signal, and therefore to deviations between the expected TOA and the observed TOA of a pulse — a difference referred to as a timing residual. In PTAs, several millisecond pulsars are timed over many years. As we have seen in Sec. 7.1, a stationary, isotropic, Gaussian, and unpolarised GWB gives rise to a specific correlation pattern between the timing residuals of different pulsars, known as the Hellings–Downs correlation. Based on the NANOGrav 15-year data set, whose analysis provides compelling evidence for the observation of such a Hellings–Downs correlation and thus a GWB, we considered various cosmic string models as sources of this signal. These were introduced in Sec. 7.3.

We used four different models of stable cosmic strings to cover model uncertainties. Furthermore, we distinguished between two models of metastable strings, either with the monopole flux fully confined in the strings or with a remaining unconfined flux. In addition to strings of field-theoretical origin, we also considered a simplified model of cosmic superstrings. Based on the spectral information, for each of the models, we performed a comparison to a simple reference model describing the GWB from SMBHBs. We also used Bayesian inference to identify regions of parameter space that lead to good agreement between the modelled spectra and the NG15 signal. None of the four stable string models shows good agreement with the observational data as their spectra are, for those string tensions sufficiently large to match the signal amplitude, too flat in the PTA band. When comparing the stable string models to the GWB from SMBHBs, we find Bayes factors ~ 0.1 if the GWB is purely due to strings and values ~ 1 if the GWB of stable strings is overlaid with the signal from the reference model. The stable string models therefore provide a worse fit to the NG15 data than the SMBHB model by itself. For both metastable strings and superstrings, we are able to find parameter regions that result in good fits to the NG15 data. Model comparison to the reference model gives rise to Bayes factors of $\sim 10 \ldots 20$ in the former case and ~ 45 in the latter case. Adding a GWB contribution from SMBHBs to the string signal reduces the Bayes factors as this does not improve the fit but increases the prior volume. That the metastable string models provide a better fit than the stable string models is a consequence of the decay of string loops. For an appropriate choice of the decay parameter, this can lead to a suppression and blue-tilt of the GWB spectrum. An increased string tension can compensate for the low amplitude of the spectrum. For superstrings, it is possible to consider string tensions sufficiently low to give rise to the correct spectral tilt needed to fit the NG15 data. For ordinary stable strings, this would correspond to signals that are too weak. For cosmic superstrings, the low amplitude can be compensated by a low intercommutation probability. We reemphasise that the computed Bayes factors should not be overinterpreted: They arise from the comparison of models that are subject to large uncertainties. This is true for both the cosmic string models as well as for the SMBHB reference model. Additionally, Bayes factors depend on the choice of prior distribution and thereby suffer, at least to a small amount, from arbitrariness.

For now, the origin of the GWB signal in the NG15 data remains unknown, but upcoming data sets by NANOGrav and other PTAs around the world will improve upon this situation. Nevertheless, a decisive determination of the origin of the GWB signal just based on spectral data from PTAs remains unlikely. On the one hand, this is because PTAs cover only a narrow frequency range, in which many sources give rise to similar spectra. This can also be seen in Ref. [1], which considers models other than cosmic strings that also provide good fits to the data. On the other hand, both string models and more generally other primordial sources of a GWB, as well as SMBHB models, suffer from large theoretical uncertainties. A distinction between the former primordial sources and the latter astrophysical source can be made by using in addition to spectral data also information about anisotropies. The GWB from SMBHBs is the result of much fewer individual sources than that from cosmic strings or other early Universe sources. A measurement of the angular GWB power spectrum could therefore help to rule out astrophysical or primordial sources. In the current NG15 data set, there is no evidence for anisotropies yet. Further information can be gained by investigating a possible polarisation of the GWB and non-Gaussianities.

In spite of the large uncertainties, we stress that, independently of whether the GWB signal in the NG15 data stems from strings or not, we were able to derive competitive constraints on the string model parameters. This is in particular true for the tension of stable strings but also for the parameter space of metastable strings. For superstrings, the large model uncertainties make it harder to estimate how reliable the derived bounds are. A more sophisticated model of the GWB signal from superstrings, however, gives rise to comparable bounds based on the NG15 data [428]. Improved modelling — already available for stable strings [363, 364] — will further strengthen the confidence in parameter bounds derived from future data sets.



Conclusion

In this thesis, we explored various aspects of cosmic defects, focusing particularly on GWs emitted by cosmic strings. We began with a review of cosmological phase transitions in Ch. 1. Symmetries that are restored at early times can be spontaneously broken as the Universe cools down. We considered the symmetry breaking mediated by a Higgs field that acquires a non-vanishing VEV. If the considered theory contains degenerate vacua, this VEV will generally become uncorrelated beyond a certain length. At the boundary of uncorrelated regions, the phase transition might not complete at all, leaving remnants of unbroken symmetry called cosmic defects.

We then systematically classified topologically stable defects in Ch. 2, relating them to non-trivial homotopy groups of the vacuum manifold: monopoles to the second, strings to the first, and domain walls to the zeroth. The study of defects in terms of the homotopy structure alone is incomplete and we complemented it with a discussion of specific field configurations such as Abrikosov–Nielsen–Olesen strings and 't Hooft–Polyakov monopoles. We continued by extending the topological classification to symmetry-breaking chains, introducing metastable defects, which, although stable in a low-energy effective theory, can unwind in the full theory.

An important quantity for the phenomenology of metastable defects is their decay rate. To accurately predict it from microphysics, one must first determine field configurations that describe their unwinding. To this end, we turned to a specific model in which monopoles form at a high energy and strings at a lower scale. In this setup, monopole–antimonopole pairs can nucleate on the strings, causing them to break apart. Previous work [215] developed ansätze to describe the unwinding process. Here, we also proposed an unwinding ansatz, however, in a model with an unconfined $U(1)$ monopole flux, in contrast to prior work. Following Refs. [215, 216], subsequent research may utilise our ansatz to compute the fragmentation rate of strings via instanton methods.

The study of metastable strings opens several promising directions for future work. For example, consider the breaking of a metastable loop due to monopole nucleation. The resulting monopole and antimonopole accelerate towards each other, erasing the string in between, and eventually annihilate. Their relativistic motion generates GWs, which is particularly relevant in light of a possible GWB detection in the near future. The current understanding of the gravitational decay of string segments is based on approximating them as straight open Nambu–Goto strings with point masses at their ends [212]. To assess the appropriateness of modelling segments as straight, one can stay in the Nambu–Goto and point mass approximation and examine the dynamics of a wider class of segments with a non-vanishing impact parameter between the monopoles. Furthermore, the validity of the Nambu–Goto and point mass approximations can be assessed by considering an unwinding ansatz, either that proposed in this work or those given in Refs. [215, 216]. Starting from such an ansatz, the dynamics of string segments can be studied in field theory simulations. The methods of Ref. [500], which explores the gravitational radiation from a colliding confined monopole–antimonopole pair, could provide a basis for such simulations. Another compelling question is whether oscillating, GW-emitting string segments could realistically lead to the formation of primordial black holes, and, if so, what abundance is to be expected to result from a network of metastable strings.

Following our discussion of metastable defects, we concluded the chapter with an overview of other types of defects that do not fall into any of the previously discussed categories. One such class comprises embedded defects, which are generally unstable but can persist in the thermal plasma of the early Universe. We explored this idea in Ch. 3 and used plasma-stabilised embedded domain walls to explain the baryon asymmetry of the Universe via wall-mediated electroweak baryogenesis. This mechanism requires defects inside of which electroweak symmetry is restored after the electroweak phase transition. We focused on an extended electroweak Higgs sector as a sufficiently strong source of CP violation. Fermions are reflected from the moving wall's boundary in a CP-violating manner, leading to a chiral asymmetry that diffuses in front of the wall. For the trailing edge, this is inside the wall where electroweak sphalerons convert the chiral asymmetry into a baryon asymmetry.

With an order-of-magnitude estimate, we demonstrated that a sufficiently long-lived network of domain walls can reproduce the observed baryon-to-entropy ratio. In principle, this mechanism can be applied to any kind of wall that decays early enough to avoid overclosure. We chose to investigate a non-standard type of wall, though. For the first time, we constructed an extension of the SM that gives rise to plasma-stabilised embedded domain walls. The stabilisation mechanism can, in principle, remain effective even after the electroweak phase transition, and the model is designed so that electroweak symmetry remains unbroken inside the walls.

Our proposal illustrates only one interesting application, but plasma-stabilised embedded defects could have a richer phenomenology. For example, they could constitute a relevant source of a GWB. Embedded walls might in parts also collapse into primordial black holes. However, reliable predictions in this context require a detailed study of the defect properties. In particular their tension, width, and the temperature, at which plasma-stabilisation ceases and the defects decay, need to be investigated rigorously. The latter is crucial, as it determines the lifetime of plasma-stabilised defects and distinguishes between those that decay quickly and remain cosmologically irrelevant, and those that persist long enough to have potentially significant consequences. Determining this temperature likely includes examining the backreaction of the out-of-equilibrium defects on the thermal plasma, as well as the role of vacuum fluctuations, which become increasingly important as the temperature decreases. Both have not been considered in previous studies [246, 247, 303].

Moreover, we have seen that plasma effects influence which types of defects are stabilised. In the symmetry-breaking chain we considered, monopoles would have formed in the absence of plasma effects, but including them led to the formation of domain walls. Once the stabilisation has become inefficient and the walls decay, monopoles are expected to reemerge — a transition that does otherwise not happen for defects. Since the correlation length likely grows over time due to wall interactions, the resulting monopoles should be significantly diluted compared to their standard abundance, potentially alleviating a monopole problem. A precise understanding of plasma-stabilised embedded defects and their evolution should therefore be a key objective of future research.

Defects encode valuable information about symmetries restored in the early Universe, specifically the vacuum structure of particle physics. To access this information, one must probe their observational signatures. Among these, the GWB produced by a network of local strings stands out as a promising target, as it could be tested by a number of current and upcoming experiments. With the goal of computing such a GWB in mind, we started our discussion in Ch. 4 by showing that strings arising from a local $U(1)$ symmetry breaking can often be treated as effectively one-dimensional objects in the Nambu–Goto approximation, which allows us to gain insight into string dynamics. Based on the VOS model, we described the evolution of a network of infinitely long or super-Hubble strings and saw that the energy loss into shorter string loops and subsequently into gravitational radiation prevents a string network from overclosing the Universe. We combined the knowledge of the network evolution with a discussion of the decay of individual loops to demonstrate how the GWB from a population of decaying string loops is computed.

In Ch. 5, we applied these techniques to a specific parameter regime characterized by low symmetry-breaking scales. Strings produced at such scales have low tensions, leading to slow decay into gravitational radiation, and the associated loops are produced relatively late and are correspondingly large. Combining these two aspects, we identified regions of parameter space where none of the string loops contributing to the GWB has fully decayed by today. We refer to such strings as low-scale strings and showed that the minimum length in their loop population leads to a maximum frequency for the GWB contribution from the fundamental harmonic loop oscillation mode. In the GWB spectrum from all harmonics, this causes a characteristic oscillatory feature near its peak, with local minima at integer multiples of the frequency of the first minimum. Combined with the specific low- and high-frequency power laws we derived, this results in a distinctive spectral signature that can be probed by future space-based interferometers such as BBO and DECIGO.

To make these predictions, we assumed the initial loop length to be a fixed fraction of the horizon scale. To test the robustness of our results, we repeated the computation using a broader initial loop length distribution adapted to that obtained from large network simulations by BOS. This slightly smoothens the GWB spectra from low-scale strings but leaves the characteristic oscillating feature sufficiently pronounced to be detectable.

String loops can also decay into particle radiation, introducing an additional decay channel alongside gravitational radiation. This generally accelerates loop decay and reduces the viable parameter space for low-scale strings. While it remains debated whether gravitational radiation dominates loop decay at large lengths, a premise we adopted throughout this thesis, it seems widely accepted that below a certain length scale, particle decay becomes dominant. To study this, we modelled loop decay using two limiting scenarios. In the first, loops shorter than the critical length are assumed not to contribute to the GWB, while longer loops evolve solely through GW emission. In the second, we included particle radiation effects in the loop length evolution via nonscaling models for the loop number density, but left gravitational decay unchanged and, in particular, independent of the loop length. We found that sufficiently strong particle decay can significantly shrink the parameter space of low-scale strings. However, across most of the parameter space, both decay channels are, as a consequence of the low string tensions, so inefficient that loop decay is effectively negligible. Under this approximation, we obtained accurate GWB spectra for low-scale strings. Notably, even with a strong reduction of the low-scale-string parameter space due to particle decay, no fine-tuning is required to obtain parameters that give rise to the distinctive GWB signatures future experiments could detect.

We have approached our computations from a phenomenological perspective and treated the initial time of loop formation as a free parameter. However, in a specific particle physics model, this parameter would be fixed. Future studies should consider realistic microscopic models and investigate, which part of our parameter space they correspond to. In this context, further improvements to the GWB prediction can be achieved by accounting for the non-instantaneous onset of loop formation and by exploring potential deviations from the BOS loop length distribution at early times, shortly after the network exits the friction regime.

The possibility that low-scale strings can leave observable imprints was not necessarily anticipated, and current literature remains almost exclusively focused on high-scale strings. Future research should fill this gap and explore the phenomenology of low-scale cosmic strings extensively.

Following our study of this specific parameter regime, we returned to strings emerging from arbitrary scales. To better understand the features and parameter dependencies of their spectra, in Ch. 6, we derived analytical GWB templates. In contrast to earlier work [403, 411], our templates are valid across all frequencies and all relevant string tensions, and capture qualitatively distinct spectral shapes, including those of low-scale strings. They also predict spectral features such as minimum and maximum frequencies correctly, while earlier studies had to introduce those by hand. Moreover, our templates account for changes in the effective number of relativistic degrees of freedom, which are imprinted as step-like features onto the plateau region of the spectrum. Observing them could provide insights into the number of relativistic particle species in the early Universe. Unlike numerical predictions of the GWB spectra from strings, our templates are computationally efficient and especially useful for large parameter scans required in data analyses.

In future work, we will compare our results with those obtained from numerical simulations in Ref. [364], which account for gravitational backreaction. In particular, we will derive, for different experiments and frequency bands, $G\mu$-dependent optimal fudge factors multiplying the spectra. In this way, we can further improve the accuracy of our predictions, while preserving the efficiency of analytical expressions. Our results restrict to the case of stable cosmic strings. However, the applied techniques should extend to the more complex GWBs from metastable strings. Developing analogous analytical templates for these cases would be a valuable contribution to the field.

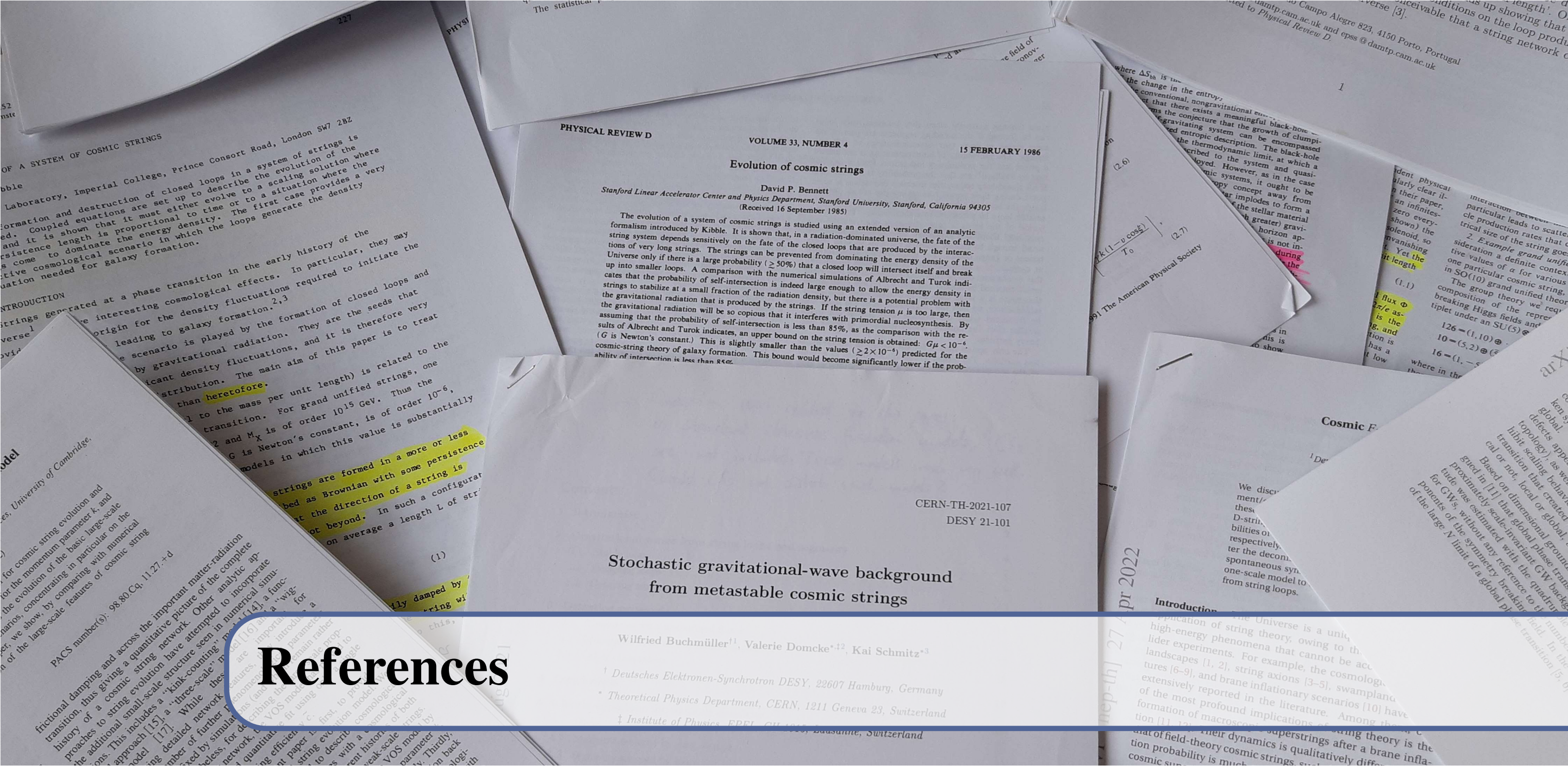
The ultimate goal of predicting GWB spectra is to compare them with experimental data. One way to probe the GWB from cosmic strings is through PTAs. GWs modify the propagation time of radio signals from pulsars to Earth. PTAs monitor the precise arrival times of these pulses from an array of millisecond pulsars. A GWB induces correlated timing shifts between pulsar pairs. For an isotropic, unpolarised, and stationary background, these correlations follow the Hellings–Downs curve, which describes the correlations as a function of the angular separation between pulsars. NANOGrav found evidence for a Hellings–Downs correlation and thus a GWB in their 15-year data set. In our final chapter, Ch. 7, we confronted the theoretically predicted GWB spectra from various cosmic string models with this data set. We used four models for stable strings (accounting for modelling uncertainties), two for metastable strings (with and without a GWB contribution from string segments), and one for superstrings. For each of these models, we performed a Bayesian model comparison with respect to a common reference model and found that all four stable string models provide a significantly worse fit than the metastable string models, whereas the superstring model performs even better. Based on the results of Bayesian parameter estimation, we concluded that the preferred GWB spectra from stable strings were able to match the amplitude of the signal in the data, but were too flat to provide a good fit over the entire frequency band. In contrast, the preferred metastable and superstring spectra matched both the amplitude and spectral slope, being sufficiently blue-tilted. Regardless of whether the signal in the NANOGrav 15-year data set is caused by a GWB, our analysis yields competitive constraints on the parameter space of all string models considered. We emphasise that modelling of metastable and particularly superstrings involves greater uncertainties than that of stable strings, as only for the latter large network simulations exist. Future research should aim to close this gap.

Assuming the observed signal is indeed due to a GWB, modelling uncertainties and the spectral similarities among different sources in the PTA band make it unlikely that the origin of the GWB can be determined from spectral data alone. However, with longer observation times and larger data sets, it should be possible to constrain the anisotropies in the GWB. This would help to decide whether the GWB is of astrophysical origin, i.e., from inspiraling SMBHBs, or whether it is of primordial origin. Astrophysical sources typically produce much larger anisotropies, although long-lived primordial sources, such as cosmic strings, might also generate non-negligible anisotropies. Future studies should investigate the GWB anisotropies expected from networks of stable and metastable strings closely.

Additionally, other future experiments could help to answer the question whether there is a GWB associated with cosmic strings. Since string models are expected to produce a plateau in the spectrum at frequencies much higher than the PTA band, a string-sourced GWB signal should also show in future experiments such as space-based interferometers, e.g., LISA, or ground-based observatories, e.g., Einstein Telescope. For now, the origin of the GWB signal evident in the NANOGrav 15-year data remains unknown, though.

Skepticism about specific models of cosmic defects is certainly justified. However, strong support for the existence of defects is provided by the wide range of GUTs and extensions of the SM from which they naturally emerge. The relevance of defects is, on the one hand, due to their profound connection to the symmetries and vacuum structure of particle physics. On the other hand, it is due to their distinctive observational signatures. In particular, the GWB from cosmic strings offers a promising prospect for a near-future detection through current and upcoming experiments, for example NANOGrav. Probing such a spectrum with different experiments over a large frequency range would, furthermore, allow us to extract valuable information about the expansion history of the Universe.

Many open questions remain. In this thesis we have addressed a number of them, contributing to the broader understanding of the field. Overall, cosmic defects remain a promising and exciting subject for future research.



References

[1] Adeela Afzal et al. “The NANOGrav 15 yr Data Set: Search for Signals from New Physics”. *Astrophys. J. Lett.* 951.1 (2023), L11. DOI: 10.3847/2041-8213/acdc91. ARXIV: 2306.16219.

[2] Tobias Schröder and Robert Brandenberger. “Embedded domain walls and electroweak baryogenesis”. *Phys. Rev. D* 110.4 (2024), 043516. DOI: 10.1103/PhysRevD.110.043516. ARXIV: 2404.13035.

[3] Kai Schmitz and Tobias Schröder. “Gravitational waves from low-scale cosmic strings”. *Phys. Rev. D* 110.6 (2024), 063549. DOI: 10.1103/PhysRevD.110.063549. ARXIV: 2405.10937.

[4] Kai Schmitz and Tobias Schröder. “Gravitational waves from cosmic strings for pedestrians” (2024). ARXIV: 2412.20907.

[5] Kai Schmitz and Tobias Schröder. “Gravitational waves from low-scale cosmic strings without scaling” (2025). ARXIV: 2505.04537.

[6] Martin Enriquez-Rojo and Tobias Schröder. “Asymptotic symmetries and memories of gauge theories in FLRW spacetimes”. *JHEP* 01 (2023), 011. DOI: 10.1007/JHEP01(2023)011. ARXIV: 2207.13726.

[7] Nina Cordes et al. “On the overlap reduction function of pulsar timing array searches for gravitational waves in modified gravity”. *Class. Quant. Grav.* 42.1 (2025), 015003. DOI: 10.1088/1361-6382/ad9881. ARXIV: 2407.04464.

[8] Gabriella Agazie et al. “The NANOGrav 15 yr Data Set: Running of the Spectral Index”. *Astrophys. J. Lett.* 978.2 (2025), L29. DOI: 10.3847/2041-8213/ad99d3. ARXIV: 2408.10166.

[9] Andrea Mitridate et al. “PTArcade” (2023). ARXIV: 2306.16377.

[10] A. Friedman. “Über die Krümmung des Raumes”. *Z. Phys.* 10 (1922), 377–386. DOI: 10.1007/BF01332580.

[11] Georges Lemaitre. “A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae”. *Annales Soc. Sci. Bruxelles A* 47 (1927), 49–59. DOI: 10.1007/s10714-013-1548-3.

[12] Edwin Hubble. “A relation between distance and radial velocity among extra-galactic nebulae”. *Proc. Nat. Acad. Sci.* 15 (1929), 168–173. DOI: 10.1073/pnas.15.3.168.

[13] Julien Lesgourgues and Sergio Pastor. “Massive neutrinos and cosmology”. *Phys. Rept.* 429 (2006), 307–379. DOI: 10.1016/j.physrep.2006.04.001. ARXIV: astro-ph/0603494.

[14] N. Aghanim et al. “Planck 2018 results. VI. Cosmological parameters”. *Astron. Astrophys.* 641 (2020). [Erratum: Astron.Astrophys. 652, C4 (2021)], A6. DOI: 10.1051/0004-6361/201833910. ARXIV: 1807.06209.

[15] Paul Langacker. “Grand Unified Theories and Proton Decay”. *Phys. Rept.* 72 (1981), 185. DOI: 10.1016/0370-1573(81)90059-4.

[16] Stuart Raby. *Supersymmetric Grand Unified Theories: From Quarks to Strings via SUSY GUTs*. Volume 939. Springer, 2017. ISBN: 978-3-319-55255-2. DOI: 10.1007/978-3-319-55255-2.

[17] Djuna Croon et al. “GUT Physics in the era of the LHC”. *Front. in Phys.* 7 (2019), 76. DOI: 10.3389/fphy.2019.00076. ARXIV: 1903.04977.

[18] Ugo Amaldi, Wim de Boer, and Hermann Furstenau. “Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP”. *Phys. Lett. B* 260 (1991), 447–455. DOI: 10.1016/0370-2693(91)91641-8.

[19] John R. Ellis, S. Kelley, and Dimitri V. Nanopoulos. “Precision LEP data, supersymmetric GUTs and string unification”. *Phys. Lett. B* 249 (1990), 441–448. DOI: 10.1016/0370-2693(90)91013-2.

[20] John R. Ellis, S. Kelley, and Dimitri V. Nanopoulos. “Probing the desert using gauge coupling unification”. *Phys. Lett. B* 260 (1991), 131–137. DOI: 10.1016/0370-2693(91)90980-5.

[21] Paul Langacker and Ming-xing Luo. “Implications of precision electroweak experiments for M_t, ρ_0, $\sin^2\theta_W$ and grand unification”. *Phys. Rev. D* 44 (1991), 817–822. DOI: 10.1103/PhysRevD.44.817.

[22] D. A. Kirzhnits. “Weinberg model in the hot universe”. *JETP Lett.* 15 (1972), 529–531.

[23] N. D. Mermin. “The topological theory of defects in ordered media”. *Rev. Mod. Phys.* 51 (1979), 591–648. DOI: 10.1103/RevModPhys.51.591.

[24] A. A. Abrikosov. “On the Magnetic properties of superconductors of the second group”. *Sov. Phys. JETP* 5 (1957), 1174–1182.

[25] Henry Edgar Hall and William Frank Vinen. “The rotation of liquid helium II II. The theory of mutual friction in uniformly rotating helium II”. *Proc. R. Soc. Lond.* A.238 (1956), 215–234. DOI: 10.1098/rspa.1956.0215.

[26] N. D. Mermin and Tin-Lun Ho. “Circulation and Angular Momentum in the A Phase of Superfluid Helium-3”. *Phys. Rev. Lett.* 36 (11 1976), 594–597. DOI: 10.1103/PhysRevLett.36.594.

[27] P. W. Anderson and G. Toulouse. “Phase Slippage without Vortex Cores: Vortex Textures in Superfluid ^{3}He”. *Phys. Rev. Lett.* 38 (9 1977), 508–511. DOI: 10.1103/PhysRevLett.38.508.

[28] Holger Bech Nielsen and P. Olesen. “Vortex Line Models for Dual Strings”. *Nucl. Phys. B* 61 (1973). Edited by J. C. Taylor, 45–61. DOI: 10.1016/0550-3213(73)90350-7.

[29] Gerard 't Hooft. "Magnetic Monopoles in Unified Gauge Theories". *Nucl. Phys. B* 79 (1974). Edited by J. C. Taylor, 276–284. DOI: 10.1016/0550-3213(74)90486-6.

[30] Alexander M. Polyakov. "Particle Spectrum in Quantum Field Theory". *JETP Lett.* 20 (1974). Edited by J. C. Taylor, 194–195.

[31] Ya. B. Zeldovich, I. Yu. Kobzarev, and L. B. Okun. "Cosmological Consequences of the Spontaneous Breakdown of Discrete Symmetry". *Zh. Eksp. Teor. Fiz.* 67 (1974), 3–11.

[32] T. W. B. Kibble. "Topology of Cosmic Domains and Strings". *J. Phys. A* 9 (1976), 1387–1398. DOI: 10.1088/0305-4470/9/8/029.

[33] John Preskill and Alexander Vilenkin. "Decay of metastable topological defects". *Phys. Rev. D* 47 (1993), 2324–2342. DOI: 10.1103/PhysRevD.47.2324. ARXIV: hep-ph/9209210.

[34] Rachel Jeannerot, Jonathan Rocher, and Mairi Sakellariadou. "How generic is cosmic string formation in SUSY GUTs". *Phys. Rev. D* 68 (2003), 103514. DOI: 10.1103/PhysRevD.68.103514. ARXIV: hep-ph/0308134.

[35] David I. Dunsky et al. "GUTs, hybrid topological defects, and gravitational waves". *Phys. Rev. D* 106.7 (2022), 075030. DOI: 10.1103/PhysRevD.106.075030. ARXIV: 2111.08750.

[36] A. Vilenkin and E. P. S. Shellard. *Cosmic Strings and Other Topological Defects*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2000.

[37] Ya. B. Zeldovich and M. Yu. Khlopov. "On the Concentration of Relic Magnetic Monopoles in the Universe". *Phys. Lett. B* 79 (1978), 239–241. DOI: 10.1016/0370-2693(78)90232-0.

[38] John Preskill. "Cosmological Production of Superheavy Magnetic Monopoles". *Phys. Rev. Lett.* 43 (1979), 1365. DOI: 10.1103/PhysRevLett.43.1365.

[39] Alan H. Guth. "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems". *Phys. Rev. D* 23 (1981). Edited by Li-Zhi Fang and R. Ruffini, 347–356. DOI: 10.1103/PhysRevD.23.347.

[40] Andrei D. Linde. "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems". *Phys. Lett. B* 108 (1982). Edited by Li-Zhi Fang and R. Ruffini, 389–393. DOI: 10.1016/0370-2693(82)91219-9.

[41] A. Vilenkin. "Gravitational Field of Vacuum Domain Walls and Strings". *Phys. Rev. D* 23 (1981), 852–857. DOI: 10.1103/PhysRevD.23.852.

[42] George Lazarides and Q. Shafi. "Axion Models with No Domain Wall Problem". *Phys. Lett. B* 115 (1982), 21–25. DOI: 10.1016/0370-2693(82)90506-8.

[43] G. R. Dvali and Goran Senjanovic. "Is there a domain wall problem?" *Phys. Rev. Lett.* 74 (1995), 5178–5181. DOI: 10.1103/PhysRevLett.74.5178. ARXIV: hep-ph/9501387.

[44] S. Ramazanov et al. "Beyond freeze-in: Dark matter via inverse phase transition and gravitational wave signal". *Phys. Rev. D* 105.6 (2022), 063530. DOI: 10.1103/PhysRevD.105.063530. ARXIV: 2104.13722.

[45] T. W. B. Kibble. "Evolution of a system of cosmic strings". *Nucl. Phys. B* 252 (1985). Edited by R. Baier and H. Satz. [Erratum: Nucl.Phys.B 261, 750 (1985)], 227. DOI: 10.1016/0550-3213(85)90439-0.

[46] Vitaly Vanchurin, Ken D. Olum, and Alexander Vilenkin. "Scaling of cosmic string loops". *Phys. Rev. D* 74 (2006), 063527. DOI: 10.1103/PhysRevD.74.063527. ARXIV: gr-qc/0511159.

[47] Jose J. Blanco-Pillado, Ken D. Olum, and Benjamin Shlaer. “The number of cosmic string loops”. *Phys. Rev. D* 89.2 (2014), 023512. DOI: 10.1103/PhysRevD.89.023512. ARXIV: 1309.6637.

[48] Cora Dvorkin, Mark Wyman, and Wayne Hu. “Cosmic String constraints from WMAP and the South Pole Telescope”. *Phys. Rev. D* 84 (2011), 123519. DOI: 10.1103/PhysRevD.84.123519. ARXIV: 1109.4947.

[49] Christophe Ringeval and Francois R. Bouchet. “All Sky CMB Map from Cosmic Strings Integrated Sachs-Wolfe Effect”. *Phys. Rev. D* 86 (2012), 023513. DOI: 10.1103/PhysRevD.86.023513. ARXIV: 1204.5041.

[50] P. A. R. Ade et al. “Planck 2013 results. XXV. Searches for cosmic strings and other topological defects”. *Astron. Astrophys.* 571 (2014), A25. DOI: 10.1051/0004-6361/201321621. ARXIV: 1303.5085.

[51] Tom Charnock et al. “CMB constraints on cosmic strings and superstrings”. *Phys. Rev. D* 93.12 (2016), 123503. DOI: 10.1103/PhysRevD.93.123503. ARXIV: 1603.01275.

[52] J. Richard Gott III. “Gravitational lensing effects of vacuum strings: Exact solutions”. *Astrophys. J.* 288 (1985), 422–427. DOI: 10.1086/162808.

[53] Nick Kaiser and A. Stebbins. “Microwave Anisotropy Due to Cosmic Strings”. *Nature* 310 (1984), 391–393. DOI: 10.1038/310391a0.

[54] Francois R. Bouchet, David P. Bennett, and Albert Stebbins. “Microwave Anisotropy Patterns from Evolving String Networks”. *Nature* 335 (1988), 410–414. DOI: 10.1038/335410a0.

[55] Alexander Vilenkin. “Cosmic strings as gravitational lenses”. *Astrophys. J. Lett.* 282 (1984), L51–L53. DOI: 10.1086/184303.

[56] Katherine J. Mack, Daniel H. Wesley, and Lindsay J. King. “Observing cosmic string loops with gravitational lensing surveys”. *Phys. Rev. D* 76 (2007), 123515. DOI: 10.1103/PhysRevD.76.123515. ARXIV: astro-ph/0702648.

[57] Konrad Kuijken, Xavier Siemens, and Tanmay Vachaspati. “Microlensing by Cosmic Strings”. *Mon. Not. Roy. Astron. Soc.* 384 (2008), 161. DOI: 10.1111/j.1365-2966.2007.12663.x. ARXIV: 0707.2971.

[58] Jolyon K. Bloomfield and David F. Chernoff. “Cosmic String Loop Microlensing”. *Phys. Rev. D* 89.12 (2014), 124003. DOI: 10.1103/PhysRevD.89.124003. ARXIV: 1311.7132.

[59] Robert H. Brandenberger et al. “The 21 cm Signature of Cosmic String Wakes”. *JCAP* 12 (2010), 028. DOI: 10.1088/1475-7516/2010/12/028. ARXIV: 1006.2514.

[60] David Maibach et al. “Extracting the signal of cosmic string wakes from 21-cm observations”. *Phys. Rev. D* 104.12 (2021), 123535. DOI: 10.1103/PhysRevD.104.123535. ARXIV: 2107.07289.

[61] Alexander Vilenkin and Tanmay Vachaspati. “Electromagnetic Radiation from Superconducting Cosmic Strings”. *Phys. Rev. Lett.* 58 (1987), 1041–1044. DOI: 10.1103/PhysRevLett.58.1041.

[62] Pijushpani Bhattacharjee. “Cosmic Strings and Ultrahigh-Energy Cosmic Rays”. *Phys. Rev. D* 40 (1989), 3968. DOI: 10.1103/PhysRevD.40.3968.

[63] David Garfinkle and Tanmay Vachaspati. “Radiation From Kinky, Cuspless Cosmic Loops”. *Phys. Rev. D* 36 (1987), 2229. DOI: 10.1103/PhysRevD.36.2229.

[64] Pierre Auclair, Danièle A. Steer, and Tanmay Vachaspati. "Particle emission and gravitational radiation from cosmic strings: observational constraints". *Phys. Rev. D* 101.8 (2020), 083511. DOI: 10.1103/PhysRevD.101.083511. ARXIV: 1911.12066.

[65] Hiroyuki Tashiro, Eray Sabancilar, and Tanmay Vachaspati. "CMB Distortions from Superconducting Cosmic Strings". *Phys. Rev. D* 85 (2012), 103522. DOI: 10.1103/PhysRevD.85.103522. ARXIV: 1202.2474.

[66] Madeleine Anthonisen et al. "Cosmic Microwave Background Spectral Distortions from Cosmic String Loops". *JCAP* 02 (2016), 047. DOI: 10.1088/1475-7516/2016/02/047. ARXIV: 1509.07998.

[67] Nicklas Ramberg, Wolfram Ratzinger, and Pedro Schwaller. "One μ to rule them all: CMB spectral distortions can probe domain walls, cosmic strings and low scale phase transitions". *JCAP* 02 (2023), 039. DOI: 10.1088/1475-7516/2023/02/039. ARXIV: 2209.14313.

[68] Robert H. Brandenberger, Anne-Christine Davis, and Mark Trodden. "Cosmic strings and electroweak baryogenesis". *Phys. Lett. B* 335 (1994), 123–130. DOI: 10.1016/0370-2693(94)91402-8. ARXIV: hep-ph/9403215.

[69] Robert H. Brandenberger et al. "Local and nonlocal defect mediated electroweak baryogenesis". *Phys. Rev. D* 53 (1996), 4257–4266. DOI: 10.1103/PhysRevD.53.4257. ARXIV: hep-ph/9409281.

[70] Hao Jiao, Robert Brandenberger, and Alexandre Refregier. "Early structure formation from cosmic string loops in light of early JWST observations". *Phys. Rev. D* 108.4 (2023), 043510. DOI: 10.1103/PhysRevD.108.043510. ARXIV: 2304.06429.

[71] Chia-Feng Chang and Yanou Cui. "Stochastic Gravitational Wave Background from Global Cosmic Strings". *Phys. Dark Univ.* 29 (2020), 100604. DOI: 10.1016/j.dark.2020.100604. ARXIV: 1910.04781.

[72] Chia-Feng Chang and Yanou Cui. "Gravitational waves from global cosmic strings and cosmic archaeology". *JHEP* 03 (2022), 114. DOI: 10.1007/JHEP03(2022)114. ARXIV: 2106.09746.

[73] Daniel G. Figueroa et al. "Irreducible background of gravitational waves from a cosmic defect network: update and comparison of numerical techniques". *Phys. Rev. D* 102.10 (2020), 103516. DOI: 10.1103/PhysRevD.102.103516. ARXIV: 2007.03337.

[74] Richard Lynn Davis. "Goldstone Bosons in String Models of Galaxy Formation". *Phys. Rev. D* 32 (1985), 3172. DOI: 10.1103/PhysRevD.32.3172.

[75] Alexander Vilenkin and Tanmay Vachaspati. "Radiation of Goldstone Bosons From Cosmic Strings". *Phys. Rev. D* 35 (1987), 1138. DOI: 10.1103/PhysRevD.35.1138.

[76] Amelia Drew and E. P. S. Shellard. "Radiation from global topological strings using adaptive mesh refinement: Methodology and massless modes". *Phys. Rev. D* 105.6 (2022), 063517. DOI: 10.1103/PhysRevD.105.063517. ARXIV: 1910.01718.

[77] Amelia Drew and E. P. S. Shellard. "Radiation from global topological strings using adaptive mesh refinement: Massive modes". *Phys. Rev. D* 107.4 (2023), 043507. DOI: 10.1103/PhysRevD.107.043507. ARXIV: 2211.10184.

[78] Jorge Baeza-Ballesteros et al. "Gravitational wave emission from a cosmic string loop: Global case". *Phys. Rev. D* 110.4 (2024), 043522. DOI: 10.1103/PhysRevD.110.043522. ARXIV: 2308.08456.

[79] Tanmay Vachaspati, Allen E. Everett, and Alexander Vilenkin. "Radiation From Vacuum Strings and Domain Walls". *Phys. Rev. D* 30 (1984), 2046. DOI: 10.1103/PhysRevD.30.2046.

[80] Robert H. Brandenberger. "On the Decay of Cosmic String Loops". *Nucl. Phys. B* 293 (1987), 812–828. DOI: 10.1016/0550-3213(87)90092-7.

[81] J. J. Blanco-Pillado and Ken D. Olum. "Form of cosmic string cusps". *Phys. Rev. D* 59 (1999). [Erratum: Phys.Rev.D 103, 029902 (2021)], 063508. DOI: 10.1103/PhysRevD.59.063508. ARXIV: gr-qc/9810005.

[82] Tanmay Vachaspati and Alexander Vilenkin. "Formation and Evolution of Cosmic Strings". *Phys. Rev. D* 30 (1984), 2036. DOI: 10.1103/PhysRevD.30.2036.

[83] Thibault Damour and Alexander Vilenkin. "Gravitational wave bursts from cosmic strings". *Phys. Rev. Lett.* 85 (2000), 3761–3764. DOI: 10.1103/PhysRevLett.85.3761. ARXIV: gr-qc/0004075.

[84] Thibault Damour and Alexander Vilenkin. "Gravitational wave bursts from cusps and kinks on cosmic strings". *Phys. Rev. D* 64 (2001), 064008. DOI: 10.1103/PhysRevD.64.064008. ARXIV: gr-qc/0104026.

[85] B. P. Abbott et al. "Constraints on cosmic strings using data from the first Advanced LIGO observing run". *Phys. Rev. D* 97.10 (2018), 102002. DOI: 10.1103/PhysRevD.97.102002. ARXIV: 1712.01168.

[86] R. Abbott et al. "Constraints on Cosmic Strings Using Data from the Third Advanced LIGO–Virgo Observing Run". *Phys. Rev. Lett.* 126.24 (2021), 241102. DOI: 10.1103/PhysRevLett.126.241102. ARXIV: 2101.12248.

[87] Pierre Auclair, Danièle A. Steer, and Tanmay Vachaspati. "Repeated gravitational wave bursts from cosmic strings". *Phys. Rev. D* 108.12 (2023), 123540. DOI: 10.1103/PhysRevD.108.123540. ARXIV: 2306.08331.

[88] Lawrence M. Krauss. "Gravitational waves from global phase transitions". *Phys. Lett. B* 284 (1992), 229–233. DOI: 10.1016/0370-2693(92)90425-4.

[89] Daniel G. Figueroa, Mark Hindmarsh, and Jon Urrestilla. "Exact Scale-Invariant Background of Gravitational Waves from Cosmic Defects". *Phys. Rev. Lett.* 110.10 (2013), 101302. DOI: 10.1103/PhysRevLett.110.101302. ARXIV: 1212.5458.

[90] Disrael Camargo Neves da Cunha, Christophe Ringeval, and François R. Bouchet. "Stochastic gravitational waves from long cosmic strings". *JCAP* 09 (2022), 078. DOI: 10.1088/1475-7516/2022/09/078. ARXIV: 2205.04349.

[91] Wilfried Buchmüller et al. "The Gravitational Wave Spectrum from Cosmological $B - L$ Breaking". *JCAP* 10 (2013), 003. DOI: 10.1088/1475-7516/2013/10/003. ARXIV: 1305.3392.

[92] J. Aasi et al. "Advanced LIGO". *Class. Quant. Grav.* 32 (2015), 074001. DOI: 10.1088/0264-9381/32/7/074001. ARXIV: 1411.4547.

[93] F. Acernese et al. "Advanced Virgo: a second-generation interferometric gravitational wave detector". *Class. Quant. Grav.* 32.2 (2015), 024001. DOI: 10.1088/0264-9381/32/2/024001. ARXIV: 1408.3978.

[94] Kentaro Somiya. "Detector configuration of KAGRA: The Japanese cryogenic gravitational-wave detector". *Class. Quant. Grav.* 29 (2012). Edited by Mark Hannam et al., 124007. DOI: 10.1088/0264-9381/29/12/124007. ARXIV: 1111.7185.

[95] Yoichi Aso et al. “Interferometer design of the KAGRA gravitational wave detector”. *Phys. Rev. D* 88.4 (2013), 043007. DOI: 10.1103/PhysRevD.88.043007. ARXIV: 1306.6747.

[96] Pau Amaro-Seoane et al. *Laser Interferometer Space Antenna*. 2017. ARXIV: 1702.00786.

[97] Naoki Seto, Seiji Kawamura, and Takashi Nakamura. “Possibility of direct measurement of the acceleration of the universe using 0.1-Hz band laser interferometer gravitational wave antenna in space”. *Phys. Rev. Lett.* 87 (2001), 221103. DOI: 10.1103/PhysRevLett.87.221103. ARXIV: astro-ph/0108011.

[98] Vincent Corbin and Neil J. Cornish. “Detecting the cosmic gravitational wave background with the big bang observer”. *Class. Quant. Grav.* 23 (2006), 2435–2446. DOI: 10.1088/0264-9381/23/7/014. ARXIV: gr-qc/0512039.

[99] Maura A. McLaughlin. “The North American Nanohertz Observatory for Gravitational Waves”. *Class. Quant. Grav.* 30 (2013), 224008. DOI: 10.1088/0264-9381/30/22/224008. ARXIV: 1310.0758.

[100] Michael Kramer and David J. Champion. “The European Pulsar Timing Array and the Large European Array for Pulsars”. *Class. Quant. Grav.* 30 (2013), 224009. DOI: 10.1088/0264-9381/30/22/224009.

[101] Chris L. Carilli and S. Rawlings. “Science with the Square Kilometer Array: Motivation, key science projects, standards and assumptions”. *New Astron. Rev.* 48 (2004), 979. DOI: 10.1016/j.newar.2004.09.001. ARXIV: astro-ph/0409274.

[102] Gabriella Agazie et al. “The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background”. *Astrophys. J. Lett.* 951.1 (2023), L8. DOI: 10.3847/2041-8213/acdac6. ARXIV: 2306.16213.

[103] Heng Xu et al. “Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I”. *Res. Astron. Astrophys.* 23.7 (2023), 075024. DOI: 10.1088/1674-4527/acdfa5. ARXIV: 2306.16216.

[104] J. Antoniadis et al. “The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals”. *Astron. Astrophys.* 678 (2023), A50. DOI: 10.1051/0004-6361/202346844. ARXIV: 2306.16214.

[105] Daniel J. Reardon et al. “Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array”. *Astrophys. J. Lett.* 951.1 (2023), L6. DOI: 10.3847/2041-8213/acdd02. ARXIV: 2306.16215.

[106] Yoichiro Nambu. “Quasiparticles and Gauge Invariance in the Theory of Superconductivity”. *Phys. Rev.* 117 (1960). Edited by J. C. Taylor, 648–663. DOI: 10.1103/PhysRev.117.648.

[107] J. Goldstone. “Field Theories with Superconductor Solutions”. *Nuovo Cim.* 19 (1961), 154–164. DOI: 10.1007/BF02812722.

[108] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. “Broken Symmetries”. *Phys. Rev.* 127 (1962), 965–970. DOI: 10.1103/PhysRev.127.965.

[109] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. *Phys. Rev. Lett.* 13 (1964). Edited by J. C. Taylor, 508–509. DOI: 10.1103/PhysRevLett.13.508.

[110] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. *Phys. Rev. Lett.* 13 (1964). Edited by J. C. Taylor, 321–323. DOI: 10.1103/PhysRevLett.13.321.

[111] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. *Phys. Rev. Lett.* 13 (1964). Edited by J. C. Taylor, 585–587. DOI: 10.1103/PhysRevLett.13.585.

[112] Julian S. Schwinger. “Gauge Invariance and Mass”. *Phys. Rev.* 125 (1962). Edited by J. C. Taylor, 397–398. DOI: 10.1103/PhysRev.125.397.

[113] Julian S. Schwinger. “Gauge Invariance and Mass. 2.” *Phys. Rev.* 128 (1962), 2425–2429. DOI: 10.1103/PhysRev.128.2425.

[114] Philip W. Anderson. “Plasmons, Gauge Invariance, and Mass”. *Phys. Rev.* 130 (1963). Edited by J. C. Taylor, 439–442. DOI: 10.1103/PhysRev.130.439.

[115] Steven Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, 2013. ISBN: 978-1-139-63247-8. DOI: 10.1017/CBO9781139644174.

[116] Aron J. Beekman, Louk Rademaker, and Jasper van Wezel. “An Introduction to Spontaneous Symmetry Breaking”. *SciPost Phys. Lect. Notes* 11 (2019), 1. DOI: 10.21468/SciPostPhysLectNotes.11. ARXIV: 1909.01820.

[117] John Lee. *Introduction to Smooth Manifolds*. 2nd edition. Graduate Texts in Mathematics. Springer New York, 2012. ISBN: 978-1-4419-9982-5. DOI: 10.1007/978-1-4419-9982-5.

[118] V. Mukhanov. *Physical Foundations of Cosmology*. Oxford: Cambridge University Press, 2005. ISBN: 978-0-521-56398-7. DOI: 10.1017/CBO9780511790553.

[119] J. I. Kapusta and Charles Gale. *Finite-temperature field theory: Principles and applications*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011. ISBN: 978-0-521-17322-3. DOI: 10.1017/CBO9780511535130.

[120] Mariano Quiros. “Finite temperature field theory and phase transitions”. *ICTP Summer School in High-Energy Physics and Cosmology*. 1999, 187–259. ARXIV: hep-ph/9901312.

[121] Robert H. Brandenberger. “Quantum Field Theory Methods and Inflationary Universe Models”. *Rev. Mod. Phys.* 57 (1985), 1. DOI: 10.1103/RevModPhys.57.1.

[122] Mikko Laine and Aleksi Vuorinen. *Basics of Thermal Field Theory*. Volume 925. Springer, 2016. DOI: 10.1007/978-3-319-31933-9. ARXIV: 1701.01554.

[123] Isabella Masina and Mariano Quiros. “An introduction to effective potential methods in field theory” (2025). ARXIV: 2501.12713.

[124] G. Jona-Lasinio. “Relativistic field theories with symmetry breaking solutions”. *Nuovo Cim.* 34 (1964), 1790–1795. DOI: 10.1007/BF02750573.

[125] K. Symanzik. “Renormalizable models with simple symmetry breaking. 1. Symmetry breaking by a source term”. *Commun. Math. Phys.* 16 (1970), 48–80. DOI: 10.1007/BF01645494.

[126] Sidney R. Coleman and Erick J. Weinberg. “Radiative Corrections as the Origin of Spontaneous Symmetry Breaking”. *Phys. Rev. D* 7 (1973), 1888–1910. DOI: 10.1103/PhysRevD.7.1888.

[127] Steven Weinberg. “Perturbative Calculations of Symmetry Breaking”. *Phys. Rev. D* 7 (1973), 2887–2910. DOI: 10.1103/PhysRevD.7.2887.

[128] Anders Andreassen, William Frost, and Matthew D. Schwartz. “Consistent Use of Effective Potentials”. *Phys. Rev. D* 91.1 (2015), 016009. DOI: 10.1103/PhysRevD.91.016009. ARXIV: 1408.0287.

[129] Peter Athron et al. “How arbitrary are perturbative calculations of the electroweak phase transition?” *JHEP* 01 (2023), 050. DOI: 10.1007/JHEP01(2023)050. ARXIV: 2208.01319.

[130] Debanjan Balui et al. “Gauge Invariant Effective Potential” (2025). ARXIV: 2502.17156.

[131] Steven Weinberg. “Gauge and Global Symmetries at High Temperature”. *Phys. Rev. D* 9 (1974), 3357–3378. DOI: 10.1103/PhysRevD.9.3357.

[132] Claude W. Bernard. “Feynman Rules for Gauge Theories at Finite Temperature”. *Phys. Rev. D* 9 (1974), 3312. DOI: 10.1103/PhysRevD.9.3312.

[133] L. Dolan and R. Jackiw. “Symmetry Behavior at Finite Temperature”. *Phys. Rev. D* 9 (1974), 3320–3341. DOI: 10.1103/PhysRevD.9.3320.

[134] D. A. Kirzhnits and Andrei D. Linde. “A Relativistic phase transition”. *Zh. Eksp. Teor. Fiz.* 67 (1974), 1263–1275.

[135] S. L. Glashow. “Partial Symmetries of Weak Interactions”. *Nucl. Phys.* 22 (1961), 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[136] Abdus Salam and John Clive Ward. “Electromagnetic and weak interactions”. *Phys. Lett.* 13 (1964), 168–171. DOI: 10.1016/0031-9163(64)90711-5.

[137] Steven Weinberg. “A Model of Leptons”. *Phys. Rev. Lett.* 19 (1967), 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[138] Abdus Salam. “Weak and Electromagnetic Interactions”. *Conf. Proc. C* 680519 (1968), 367–377. DOI: 10.1142/9789812795915˙0034.

[139] Nathan Herring, Shuyang Cao, and Daniel Boyanovsky. “Is the finite temperature effective potential effective for dynamics?” *Phys. Rev. D* 111.1 (2025), 016028. DOI: 10.1103/PhysRevD.111.016028. ARXIV: 2410.21633.

[140] Rabindra N. Mohapatra and Goran Senjanovic. “Broken Symmetries at High Temperature”. *Phys. Rev. D* 20 (1979), 3390–3398. DOI: 10.1103/PhysRevD.20.3390.

[141] Patrick Meade and Harikrishnan Ramani. “Unrestored Electroweak Symmetry”. *Phys. Rev. Lett.* 122.4 (2019), 041802. DOI: 10.1103/PhysRevLett.122.041802. ARXIV: 1807.07578.

[142] Anupam Mazumdar and Graham White. “Review of cosmic phase transitions: their significance and experimental signatures”. *Rept. Prog. Phys.* 82.7 (2019), 076901. DOI: 10.1088/1361-6633/ab1f55. ARXIV: 1811.01948.

[143] Mark Hindmarsh, Anne-Christine Davis, and Robert H. Brandenberger. “Formation of topological defects in first order phase transitions”. *Phys. Rev. D* 49 (1994), 1944–1950. DOI: 10.1103/PhysRevD.49.1944. ARXIV: hep-ph/9307203.

[144] Andrew A. de Laix and Tanmay Vachaspati. “On random bubble lattices”. *Phys. Rev. D* 59 (1999), 045017. DOI: 10.1103/PhysRevD.59.045017. ARXIV: hep-ph/9802423.

[145] Michela D’Onofrio and Kari Rummukainen. “Standard model cross-over on the lattice”. *Phys. Rev. D* 93.2 (2016), 025003. DOI: 10.1103/PhysRevD.93.025003. ARXIV: 1508.07161.

[146] Y. Aoki et al. “The Order of the quantum chromodynamics transition predicted by the standard model of particle physics”. *Nature* 443 (2006), 675–678. DOI: 10.1038/nature05120. ARXIV: hep-lat/0611014.

[147] M. Hindmarsh and A. Rajantie. “Defect formation and local gauge invariance”. *Phys. Rev. Lett.* 85 (2000), 4660–4663. DOI: 10.1103/PhysRevLett.85.4660. ARXIV: cond-mat/0007361.

[148] M. Hindmarsh and A. Rajantie. “Phase transition dynamics in the hot Abelian Higgs model”. *Phys. Rev. D* 64 (2001), 065016. DOI: 10.1103/PhysRevD.64.065016. ARXIV: hep-ph/0103311.

[149] Jose J. Blanco-Pillado, Ken D. Olum, and Alexander Vilenkin. “Cosmic string formation by flux trapping”. *Phys. Rev. D* 76 (2007), 103520. DOI: 10.1103/PhysRevD.76.103520. ARXIV: 0706.1577.

[150] Tanmay Vachaspati. *Kinks and Domain Walls : An Introduction to Classical and Quantum Solitons*. Oxford University Press, 2007. ISBN: 978-1-009-29045-6. DOI: 10.1017/9781009290456.

[151] W. H. Zurek. “Cosmological Experiments in Superfluid Helium?” *Nature* 317 (1985), 505–508. DOI: 10.1038/317505a0.

[152] W. H. Zurek. “Cosmic strings in laboratory superfluids and the topological remnants of other phase transitions”. *Acta Phys. Polon. B* 24 (1993), 1301–1311.

[153] W. H. Zurek. “Cosmological experiments in condensed matter systems”. *Phys. Rept.* 276 (1996), 177–221. DOI: 10.1016/S0370-1573(96)00009-9. ARXIV: cond-mat/9607135.

[154] M. E. Fisher. “Renormalization group theory: Its basis and formulation in statistical physics”. *Rev. Mod. Phys.* 70 (1998). Edited by Tian Yu Cao, 653–681. DOI: 10.1103/RevModPhys.70.653.

[155] Adolfo del Campo and Wojciech H. Zurek. “Universality of phase transition dynamics: Topological Defects from Symmetry Breaking”. *Int. J. Mod. Phys. A* 29.8 (2014). Edited by Jerome Gauntlett, 1430018. DOI: 10.1142/S0217751X1430018X. ARXIV: 1310.1600.

[156] Roger F. Dashen, Brosl Hasslacher, and Andre Neveu. “Nonperturbative Methods and Extended Hadron Models in Field Theory 2. Two-Dimensional Models and Extended Hadrons”. *Phys. Rev. D* 10 (1974), 4130–4138. DOI: 10.1103/PhysRevD.10.4130.

[157] T. H. R. Skyrme. “Particle states of a quantized meson field”. *Proc. Roy. Soc. Lond. A* 262 (1961), 237–245. DOI: 10.1098/rspa.1961.0115.

[158] G. H. Derrick. “Comments on nonlinear wave equations as models for elementary particles”. *J. Math. Phys.* 5 (1964), 1252–1254. DOI: 10.1063/1.1704233.

[159] M. Nakahara. *Geometry, topology and physics*. 2nd edition. CRC Press, 2003. DOI: 10.1201/9781315275826.

[160] V. L. Ginzburg and L. D. Landau. “On the Theory of superconductivity”. *Zh. Eksp. Teor. Fiz.* 20 (1950). Edited by D. ter Haar, 1064–1082. DOI: 10.1016/b978-0-08-010586-4.50078-x.

[161] E. B. Bogomolny. “Stability of Classical Solutions”. *Sov. J. Nucl. Phys.* 24 (1976), 449.

[162] Laurence Jacobs and Claudio Rebbi. “Interaction Energy of Superconducting Vortices”. *Phys. Rev. B* 19 (1979), 4486–4494. DOI: 10.1103/PhysRevB.19.4486.

[163] H. J. de Vega and F. A. Schaposnik. “A Classical Vortex Solution of the Abelian Higgs Model”. *Phys. Rev. D* 14 (1976), 1100–1106. DOI: 10.1103/PhysRevD.14.1100.

[164] Leandros Perivolaropoulos. “Asymptotics of Nielsen-Olesen vortices”. *Phys. Rev. D* 48 (1993), 5961–5962. DOI: 10.1103/PhysRevD.48.5961. ARXIV: hep-ph/9310264.

[165] Christopher T. Hill, Hardy M. Hodges, and Michael S. Turner. “Bosonic Superconducting Cosmic Strings”. *Phys. Rev. D* 37 (1988), 263. DOI: 10.1103/PhysRevD.37.263.

[166] A. Vilenkin and A. E. Everett. “Cosmic Strings and Domain Walls in Models with Goldstone and PseudoGoldstone Bosons”. *Phys. Rev. Lett.* 48 (1982), 1867–1870. DOI: 10.1103/PhysRevLett.48.1867.

[167] Alexander Vilenkin. “Cosmic Strings and Domain Walls”. *Phys. Rept.* 121 (1985), 263–315. DOI: 10.1016/0370-1573(85)90033-X.

[168] P. Sikivie. “Of Axions, Domain Walls and the Early Universe”. *Phys. Rev. Lett.* 48 (1982), 1156–1159. DOI: 10.1103/PhysRevLett.48.1156.

[169] Masahiro Kawasaki and Kazunori Nakayama. “Axions: Theory and Cosmological Role”. *Ann. Rev. Nucl. Part. Sci.* 63 (2013), 69–95. DOI: 10.1146/annurev-nucl-102212-170536. ARXIV: 1301.1123.

[170] Masahiro Kawasaki, Ken’ichi Saikawa, and Toyokazu Sekiguchi. “Axion dark matter from topological defects”. *Phys. Rev. D* 91.6 (2015), 065014. DOI: 10.1103/PhysRevD.91.065014. ARXIV: 1412.0789.

[171] Qaisar Shafi and Alexander Vilenkin. “Spontaneously Broken Global Symmetries and Cosmology”. *Phys. Rev. D* 29 (1984), 1870. DOI: 10.1103/PhysRevD.29.1870.

[172] Bowen Fu, Anish Ghoshal, and Stephen F. King. “Cosmic string gravitational waves from global $U(1)_{B-L}$ symmetry breaking as a probe of the type I seesaw scale”. *JHEP* 11 (2023), 071. DOI: 10.1007/JHEP11(2023)071. ARXIV: 2306.07334.

[173] H. J. de Vega and F. A. Schaposnik. “Electrically Charged Vortices in Nonabelian Gauge Theories With Chern-simons Term”. *Phys. Rev. Lett.* 56 (1986), 2564. DOI: 10.1103/PhysRevLett.56.2564.

[174] Mukunda Aryal and Allen E. Everett. “Properties of $Z(2)$ Strings”. *Phys. Rev. D* 35 (1987), 3105. DOI: 10.1103/PhysRevD.35.3105.

[175] T. W. B. Kibble, George Lazarides, and Q. Shafi. “Strings in SO(10)”. *Phys. Lett. B* 113 (1982), 237–239. DOI: 10.1016/0370-2693(82)90829-2.

[176] H. J. de Vega and F. A. Schaposnik. “Vortices and electrically charged vortices in nonAbelian gauge theories”. *Phys. Rev. D* 34 (1986), 3206–3213. DOI: 10.1103/PhysRevD.34.3206.

[177] A. E. Everett. “New Mechanism for Superconductivity in Cosmic Strings”. *Phys. Rev. Lett.* 61 (1988), 1807–1810. DOI: 10.1103/PhysRevLett.61.1807.

[178] Mark G. Alford et al. “The Interactions and Excitations of Nonabelian Vortices”. *Phys. Rev. Lett.* 64 (1990). [Erratum: Phys.Rev.Lett. 65, 668 (1990)], 1632. DOI: 10.1103/PhysRevLett.64.1632.

[179] Mark G. Alford et al. “Zero modes of nonabelian vortices”. *Nucl. Phys. B* 349 (1991), 414–438. DOI: 10.1016/0550-3213(91)90331-Q.

[180] John Preskill and Lawrence M. Krauss. “Local Discrete Symmetry and Quantum Mechanical Hair”. *Nucl. Phys. B* 341 (1990), 50–100. DOI: 10.1016/0550-3213(90)90262-C.

[181] Mark Hindmarsh. “Massless modes on cosmic strings”. *Physica B* 178.1-4 (1992), 47–55. DOI: 10.1016/0921-4526(92)90178-U.

[182] Martin Bucher, Hoi-Kwong Lo, and John Preskill. “Topological approach to Alice electrodynamics”. *Nucl. Phys. B* 386 (1992), 3–26. DOI: 10.1016/0550-3213(92)90173-9. ARXIV: hep-th/9112039.

[183] Alexander M. Polyakov. “Isomeric States of Quantum Fields”. *Zh. Eksp. Teor. Fiz.* 68 (1975), 1975.

[184] Thomas W. Kirkman and Cosmas K. Zachos. “Asymptotic Analysis of the Monopole Structure”. *Phys. Rev. D* 24 (1981), 999. DOI: 10.1103/PhysRevD.24.999.

[185] Paul Adrien Maurice Dirac. “Quantised singularities in the electromagnetic field”. *Proc. Roy. Soc. Lond. A* 133.821 (1931), 60–72. DOI: 10.1098/rspa.1931.0130.

[186] Tai Tsun Wu and Chen Ning Yang. “Concept of Nonintegrable Phase Factors and Global Formulation of Gauge Fields”. *Phys. Rev. D* 12 (1975). Edited by Jong-Ping Hsu and D. Fine, 3845–3857. DOI: 10.1103/PhysRevD.12.3845.

[187] E. B. Bogomolny and A. I. Vainshtein. "Stability of Strings in Gauge Abelian Theory". *Sov. J. Nucl. Phys.* 23 (1976), 588–591.

[188] M. K. Prasad and Charles M. Sommerfield. "An Exact Classical Solution for the 't Hooft Monopole and the Julia-Zee Dyon". *Phys. Rev. Lett.* 35 (1975), 760–762. DOI: 10.1103/PhysRevLett.35.760.

[189] E. B. Bogomolny and M. S. Marinov. "Calculation of the Monopole Mass in Gauge Theory". *Yad. Fiz.* 23 (1976), 676–680.

[190] P. Forgacs, Z. Horvath, and L. Palla. "Exact Multi - Monopole Solutions in the Bogomolny-Prasad-Sommerfield Limit". *Phys. Lett. B* 99 (1981). [Erratum: Phys.Lett.B 101, 457 (1981)], 232. DOI: 10.1016/0370-2693(81)91115-1.

[191] Chen Zhang et al. "On the cosmological abundance of magnetic monopoles". *JHEP* 08 (2024), 220. DOI: 10.1007/JHEP08(2024)220. ARXIV: 2404.04926.

[192] Erick J. Weinberg. "Constraints on the Annihilation of Primordial Magnetic Monopoles". *Phys. Lett. B* 126 (1983), 441–444. DOI: 10.1016/0370-2693(83)90359-3.

[193] Paul Langacker and So-Young Pi. "Magnetic Monopoles in Grand Unified Theories". *Phys. Rev. Lett.* 45 (1980), 1. DOI: 10.1103/PhysRevLett.45.1.

[194] G. R. Dvali, Alejandra Melfo, and Goran Senjanovic. "Is There a monopole problem?" *Phys. Rev. Lett.* 75 (1995), 4559–4562. DOI: 10.1103/PhysRevLett.75.4559. ARXIV: hep-ph/9507230.

[195] Borut Bajc, Antonio Riotto, and Goran Senjanovic. "Large lepton number of the universe and the fate of topological defects". *Phys. Rev. Lett.* 81 (1998), 1355–1358. DOI: 10.1103/PhysRevLett.81.1355. ARXIV: hep-ph/9710415.

[196] G. R. Dvali, Hong Liu, and Tanmay Vachaspati. "Sweeping away the monopole problem". *Phys. Rev. Lett.* 80 (1998), 2281–2284. DOI: 10.1103/PhysRevLett.80.2281. ARXIV: hep-ph/9710301.

[197] Dejan Stojkovic and Katherine Freese. "A Black hole solution to the cosmological monopole problem". *Phys. Lett. B* 606 (2005), 251–257. DOI: 10.1016/j.physletb.2004.12.019. ARXIV: hep-ph/0403248.

[198] Manuel Barriola and Alexander Vilenkin. "Gravitational Field of a Global Monopole". *Phys. Rev. Lett.* 63 (1989), 341. DOI: 10.1103/PhysRevLett.63.341.

[199] D. P. Bennett and S. H. Rhie. "Cosmological evolution of global monopoles and the origin of large scale structure". *Phys. Rev. Lett.* 65 (1990), 1709–1712. DOI: 10.1103/PhysRevLett.65.1709.

[200] Leandros Perivolaropoulos. "Instabilities and interactions of global topological defects". *Nucl. Phys. B* 375 (1992), 665–693. DOI: 10.1016/0550-3213(92)90115-R.

[201] C. J. A. P. Martins and A. Achucarro. "Evolution of local and global monopole networks". *Phys. Rev. D* 78 (2008), 083541. DOI: 10.1103/PhysRevD.78.083541.

[202] Asier Lopez-Eiguren et al. "Cosmic Microwave Background constraints for global strings and global monopoles". *JCAP* 07 (2017), 026. DOI: 10.1088/1475-7516/2017/07/026. ARXIV: 1705.04154.

[203] Robert Brandenberger and Hao Jiao. "Cosmic Textures and Global Monopoles as Seeds for Super-Massive Black Holes". *JCAP* 02 (2020), 002. DOI: 10.1088/1475-7516/2020/02/002. ARXIV: 1908.04585.

[204] T. W. B. Kibble, George Lazarides, and Q. Shafi. “Walls Bounded by Strings”. *Phys. Rev. D* 26 (1982), 435. DOI: 10.1103/PhysRevD.26.435.

[205] Allen E. Everett and Alexander Vilenkin. “Left-right Symmetric Theories and Vacuum Domain Walls and Strings”. *Nucl. Phys. B* 207 (1982), 43–53. DOI: 10.1016/0550-3213(82)90135-3.

[206] F. Alexander Bais. “The Topology of Monopoles Crossing a Phase Boundary”. *Phys. Lett. B* 98 (1981), 437–440. DOI: 10.1016/0370-2693(81)90447-0.

[207] Edmund J. Copeland et al. “Monopoles Connected by Strings and the Monopole Problem”. *Nucl. Phys. B* 298 (1988), 445–457. DOI: 10.1016/0550-3213(88)90350-1.

[208] T. W. B. Kibble and Tanmay Vachaspati. “Monopoles on strings”. *J. Phys. G* 42.9 (2015), 094002. DOI: 10.1088/0954-3899/42/9/094002. ARXIV: 1506.02022.

[209] Wilfried Buchmuller. “Metastable strings and dumbbells in supersymmetric hybrid inflation”. *JHEP* 04 (2021), 168. DOI: 10.1007/JHEP04(2021)168. ARXIV: 2102.08923.

[210] Wilfried Buchmüller, Valerie Domcke, and Kai Schmitz. “Metastable cosmic strings”. *JCAP* 11 (2023), 020. DOI: 10.1088/1475-7516/2023/11/020. ARXIV: 2307.04691.

[211] A. Vilenkin. “Cosmological Evolution of Monopoles Connected by Strings”. *Nucl. Phys. B* 196 (1982), 240–258. DOI: 10.1016/0550-3213(82)90037-2.

[212] Xavier Martin and Alexander Vilenkin. “Gravitational radiation from monopoles connected by strings”. *Phys. Rev. D* 55 (1997), 6054–6060. DOI: 10.1103/PhysRevD.55.6054. ARXIV: gr-qc/9612008.

[213] Louis Leblond, Benjamin Shlaer, and Xavier Siemens. “Gravitational Waves from Broken Cosmic Strings: The Bursts and the Beads”. *Phys. Rev. D* 79 (2009), 123519. DOI: 10.1103/PhysRevD.79.123519. ARXIV: 0903.4686.

[214] Wilfried Buchmüller, Valerie Domcke, and Kai Schmitz. “Stochastic gravitational-wave background from metastable cosmic strings”. *JCAP* 12.12 (2021), 006. DOI: 10.1088/1475-7516/2021/12/006. ARXIV: 2107.04578.

[215] M. Shifman and Alexei Yung. “Metastable strings in Abelian Higgs models embedded in nonAbelian theories: Calculating the decay rate”. *Phys. Rev. D* 66 (2002), 045012. DOI: 10.1103/PhysRevD.66.045012. ARXIV: hep-th/0205025.

[216] Akifumi Chitose et al. “Revisiting metastable cosmic string breaking”. *JHEP* 04 (2024), 068. DOI: 10.1007/JHEP04(2024)068. ARXIV: 2312.15662.

[217] Akifumi Chitose et al. “Cosmic Strings in Multi-Step Symmetry Breaking” (2025). ARXIV: 2506.15194.

[218] George Lazarides, Qaisar Shafi, and Amit Tiwari. “Composite topological structures in SO(10)”. *JHEP* 05 (2023), 119. DOI: 10.1007/JHEP05(2023)119. ARXIV: 2303.15159.

[219] Richard Lynn Davis. “Texture: A cosmological topological defect”. *Phys. Rev. D* 35 (1987), 3705. DOI: 10.1103/PhysRevD.35.3705.

[220] Neil Turok. “Global Texture as the Origin of Cosmic Structure”. *Phys. Rev. Lett.* 63 (1989), 2625. DOI: 10.1103/PhysRevLett.63.2625.

[221] Neil Turok and David Spergel. “Global Texture and the Microwave Background”. *Phys. Rev. Lett.* 64 (1990), 2736. DOI: 10.1103/PhysRevLett.64.2736.

[222] Joanes Lizarraga et al. “Constraining topological defects with temperature and polarization anisotropies”. *Phys. Rev. D* 90.10 (2014), 103504. DOI: 10.1103/PhysRevD.90.103504. ARXIV: 1408.4126.

[223] Katherine Jones-Smith, Lawrence M. Krauss, and Harsh Mathur. “A Nearly Scale Invariant Spectrum of Gravitational Radiation from Global Phase Transitions”. *Phys. Rev. Lett.* 100 (2008), 131302. DOI: 10.1103/PhysRevLett.100.131302. ARXIV: 0712.0778.

[224] Elisa Fenu et al. “Gravitational waves from self-ordering scalar fields”. *JCAP* 10 (2009), 005. DOI: 10.1088/1475-7516/2009/10/005. ARXIV: 0908.0425.

[225] M. S. Madsen. “The Isotropization of cosmological textures”. *Phys. Lett. B* 312 (1993), 75–78. DOI: 10.1016/0370-2693(93)90490-9.

[226] Robert Brandenberger, Bryce Cyr, and Hao Jiao. “Cosmic Rays and Spectral Distortions from Collapsing Textures”. *JCAP* 09 (2020), 035. DOI: 10.1088/1475-7516/2020/09/035. ARXIV: 2005.11099.

[227] Neil Turok and John Zadrozny. “Dynamical generation of baryons at the electroweak transition”. *Phys. Rev. Lett.* 65 (1990), 2331–2334. DOI: 10.1103/PhysRevLett.65.2331.

[228] T. Vachaspati and A. Achucarro. “Semilocal cosmic strings”. *Phys. Rev. D* 44 (1991), 3067–3071. DOI: 10.1103/PhysRevD.44.3067.

[229] Mark Hindmarsh. “Existence and stability of semilocal strings”. *Phys. Rev. Lett.* 68 (1992), 1263–1266. DOI: 10.1103/PhysRevLett.68.1263.

[230] Ana Achucarro et al. “Dynamical simulations of semilocal strings”. *Nucl. Phys. B* 388 (1992), 435–456. DOI: 10.1016/0550-3213(92)90621-H.

[231] John Preskill. “Semilocal defects”. *Phys. Rev. D* 46 (1992), 4218–4231. DOI: 10.1103/PhysRevD.46.4218. ARXIV: hep-ph/9206216.

[232] Ana Achucarro and Tanmay Vachaspati. “Semilocal and electroweak strings”. *Phys. Rept.* 327 (2000), 347–426. DOI: 10.1016/S0370-1573(99)00103-9. ARXIV: hep-ph/9904229.

[233] Manuel Barriola, Tanmay Vachaspati, and Martin Bucher. “Embedded defects”. *Phys. Rev. D* 50 (1994), 2819–2825. DOI: 10.1103/PhysRevD.50.2819. ARXIV: hep-th/9306120.

[234] Anne-Christine Davis and Nathan F. Lepora. “Embedded defects and symmetry breaking in flipped SU(5)”. *Phys. Rev. D* 52 (1995), 7265–7275. DOI: 10.1103/PhysRevD.52.7265. ARXIV: hep-ph/9504411.

[235] Nathan F. Lepora and Anne-Christine Davis. “Examples of embedded defects (in particle physics and condensed matter)”. *Phys. Rev. D* 58 (1998), 125028. DOI: 10.1103/PhysRevD.58.125028. ARXIV: hep-ph/9507466.

[236] Nathan F. Lepora and T. W. B. Kibble. “Classifying vortex solutions to gauge theories”. *Phys. Rev. D* 59 (1999), 125019. DOI: 10.1103/PhysRevD.59.125019. ARXIV: hep-th/9904177.

[237] Yoichiro Nambu. “String-Like Configurations in the Weinberg-Salam Theory”. *Nucl. Phys. B* 130 (1977). Edited by T. Eguchi, 505. DOI: 10.1016/0550-3213(77)90252-8.

[238] Tanmay Vachaspati. “Vortex solutions in the Weinberg-Salam model”. *Phys. Rev. Lett.* 68 (1992). [Erratum: Phys.Rev.Lett. 69, 216 (1992)], 1977–1980. DOI: 10.1103/PhysRevLett.68.1977.

[239] Tanmay Vachaspati. “Electroweak strings”. *Nucl. Phys. B* 397 (1993), 648–671. DOI: 10.1016/0550-3213(93)90189-V.

[240] M. James, L. Perivolaropoulos, and T. Vachaspati. “Stability of electroweak strings”. *Phys. Rev. D* 46 (1992), R5232–R5235. DOI: 10.1103/PhysRevD.46.R5232.

[241] Tanmay Vachaspati and Richard Watkins. “Bound states can stabilize electroweak strings”. *Phys. Lett. B* 318 (1993), 163–168. DOI: 10.1016/0370-2693(93)91800-3. ARXIV: hep-ph/9211284.

[242] Richard Holman et al. “Metastable cosmic strings in realistic models”. *Phys. Rev. D* 46 (1992), 5352–5359. DOI: 10.1103/PhysRevD.46.5352. ARXIV: hep-ph/9208245.

[243] Leandros Perivolaropoulos. “Stable, spinning embedded vortices”. *Phys. Rev. D* 50 (1994), 962–966. DOI: 10.1103/PhysRevD.50.962. ARXIV: hep-ph/9403298.

[244] Jaume Garriga and Xavi Montes. “Stability of Z strings in strong magnetic fields”. *Phys. Rev. Lett.* 75 (1995), 2268–2271. DOI: 10.1103/PhysRevLett.75.2268. ARXIV: hep-ph/9505424.

[245] Minos Axenides and Leandros Perivolaropoulos. “Topological defects with nonsymmetric core”. *Phys. Rev. D* 56 (1997), 1973–1981. DOI: 10.1103/PhysRevD.56.1973. ARXIV: hep-ph/9702221.

[246] M. Nagasawa and Robert H. Brandenberger. “Stabilization of embedded defects by plasma effects”. *Phys. Lett. B* 467 (1999), 205–210. DOI: 10.1016/S0370-2693(99)01140-5. ARXIV: hep-ph/9904261.

[247] Michiyasu Nagasawa and Robert Brandenberger. “Stabilization of the electroweak Z string in the early universe”. *Phys. Rev. D* 67 (2003), 043504. DOI: 10.1103/PhysRevD.67.043504. ARXIV: hep-ph/0207246.

[248] Edward Witten. “Superconducting Strings”. *Nucl. Phys. B* 249 (1985), 557–592. DOI: 10.1016/0550-3213(85)90022-7.

[249] T. Vachaspati. “Magnetic fields from cosmological phase transitions”. *Phys. Lett. B* 265 (1991), 258–261. DOI: 10.1016/0370-2693(91)90051-Q.

[250] Robert H. Brandenberger et al. “Superconducting cosmic strings and primordial magnetic fields”. *Phys. Lett. B* 293 (1992), 287–293. DOI: 10.1016/0370-2693(92)90885-8. ARXIV: hep-ph/9206232.

[251] Alejandra Kandus, Kerstin E. Kunze, and Christos G. Tsagas. “Primordial magnetogenesis”. *Phys. Rept.* 505 (2011), 1–58. DOI: 10.1016/j.physrep.2011.03.001. ARXIV: 1007.3891.

[252] Tanmay Vachaspati. “Cosmic Sparks from Superconducting Strings”. *Phys. Rev. Lett.* 101 (2008), 141301. DOI: 10.1103/PhysRevLett.101.141301. ARXIV: 0802.0711.

[253] Yun-Wei Yu et al. “Implications of fast radio bursts for superconducting cosmic strings”. *JCAP* 11 (2014), 040. DOI: 10.1088/1475-7516/2014/11/040. ARXIV: 1409.5516.

[254] Pijushpani Bhattacharjee and Gunter Sigl. “Origin and propagation of extremely high-energy cosmic rays”. *Phys. Rept.* 327 (2000), 109–247. DOI: 10.1016/S0370-1573(99)00101-5. ARXIV: astro-ph/9811011.

[255] Veniamin Berezinsky et al. “UHE neutrinos from superconducting cosmic strings”. *Phys. Rev. D* 80 (2009), 023014. DOI: 10.1103/PhysRevD.80.023014. ARXIV: 0901.0527.

[256] William T. Emond, Sabir Ramazanov, and Rome Samanta. “Gravitational waves from melting cosmic strings”. *JCAP* 01.01 (2022), 057. DOI: 10.1088/1475-7516/2022/01/057. ARXIV: 2108.05377.

[257] E. Babichev et al. “Gravitational shine of dark domain walls”. *JCAP* 04.04 (2022), 028. DOI: 10.1088/1475-7516/2022/04/028. ARXIV: 2112.12608.

[258] E. Babichev et al. “NANOGrav spectral index γ=3 from melting domain walls”. *Phys. Rev. D* 108.12 (2023), 123529. DOI: 10.1103/PhysRevD.108.123529. ARXIV: 2307.04582.

[259] Xiu-Fei Li. “Probing the high temperature symmetry breaking with gravitational waves from domain walls” (2023). ARXIV: 2307.03163.

[260] I. Dankovsky et al. “Numerical analysis of melting domain walls and their gravitational waves”. *JCAP* 02 (2025), 064. DOI: 10.1088/1475-7516/2025/02/064. ARXIV: 2410.21971.

[261] Gerald Rosen. “Charged Particlelike Solutions to Nonlinear Complex Scalar Field Theories”. *J. Math. Phys.* 9.7 (1968), 999–1002. DOI: 10.1063/1.1664694.

[262] T. D. Lee and G. C. Wick. “Vacuum Stability and Vacuum Excitation in a Spin 0 Field Theory”. *Phys. Rev. D* 9 (1974), 2291–2316. DOI: 10.1103/PhysRevD.9.2291.

[263] Sidney R. Coleman. “Q-balls”. *Nucl. Phys. B* 262.2 (1985). [Addendum: Nucl.Phys.B 269, 744 (1986)], 263. DOI: 10.1016/0550-3213(86)90520-1.

[264] R. Friedberg, T. D. Lee, and A. Sirlin. “A Class of Scalar-Field Soliton Solutions in Three Space Dimensions”. *Phys. Rev. D* 13 (1976), 2739–2761. DOI: 10.1103/PhysRevD.13.2739.

[265] R. Friedberg, T. D. Lee, and A. Sirlin. “Gauge Field Nontopological Solitons in Three Space Dimensions. 1.” *Nucl. Phys. B* 115 (1976), 1–31. DOI: 10.1016/0550-3213(76)90274-1.

[266] R. Friedberg, T. D. Lee, and A. Sirlin. “Gauge Field Nontopological Solitons in Three Space Dimensions. 2.” *Nucl. Phys. B* 115 (1976), 32–47. DOI: 10.1016/0550-3213(76)90275-3.

[267] R. Friedberg and T. D. Lee. “Fermion Field Nontopological Solitons. 1.” *Phys. Rev. D* 15 (1977), 1694. DOI: 10.1103/PhysRevD.15.1694.

[268] R. Friedberg and T. D. Lee. “Fermion Field Nontopological Solitons. 2. Models for Hadrons”. *Phys. Rev. D* 16 (1977), 1096. DOI: 10.1103/PhysRevD.16.1096.

[269] I. L. Bogolyubsky and V. G. Makhankov. “On the Pulsed Soliton Lifetime in Two Classical Relativistic Theory Models”. *JETP Lett.* 24 (1976), 12.

[270] I. L. Bogolyubsky and V. G. Makhankov. “Dynamics of Heavy Spherically-Symmetric Pulsons”. *Pisma Zh. Eksp. Teor. Fiz.* 25 (1977), 120–123.

[271] Marcelo Gleiser. “Pseudostable bubbles”. *Phys. Rev. D* 49 (1994), 2978–2981. DOI: 10.1103/PhysRevD.49.2978. ARXIV: hep-ph/9308279.

[272] Edward W. Kolb and Igor I. Tkachev. “Nonlinear axion dynamics and formation of cosmological pseudosolitons”. *Phys. Rev. D* 49 (1994), 5040–5051. DOI: 10.1103/PhysRevD.49.5040. ARXIV: astro-ph/9311037.

[273] Edmund J. Copeland, M. Gleiser, and H. -R. Muller. “Oscillons: Resonant configurations during bubble collapse”. *Phys. Rev. D* 52 (1995), 1920–1933. DOI: 10.1103/PhysRevD.52.1920. ARXIV: hep-ph/9503217.

[274] Alexander Kusenko and Mikhail E. Shaposhnikov. “Supersymmetric Q balls as dark matter”. *Phys. Lett. B* 418 (1998), 46–54. DOI: 10.1016/S0370-2693(97)01375-0. ARXIV: hep-ph/9709492.

[275] Yang Bai, Andrew J. Long, and Sida Lu. “Tests of Dark MACHOs: Lensing, Accretion, and Glow”. *JCAP* 09 (2020), 044. DOI: 10.1088/1475-7516/2020/09/044. ARXIV: 2003.13182.

[276] Siyu Jiang, Fa Peng Huang, and Pyungwon Ko. “Gauged Q-ball dark matter through a cosmological first-order phase transition”. *JHEP* 07 (2024), 053. DOI: 10.1007/JHEP07(2024)053. ARXIV: 2404.16509.

[277] Shuang-Yong Zhou et al. “Gravitational Waves from Oscillon Preheating”. *JHEP* 10 (2013), 026. DOI: 10.1007/JHEP10(2013)026. ARXIV: 1304.6094.

[278] Stefan Antusch, Francesco Cefala, and Stefano Orani. “What can we learn from the stochastic gravitational wave background produced by oscillons?” *JCAP* 03 (2018), 032. DOI: 10.1088/1475-7516/2018/03/032. ARXIV: 1712.03231.

[279] Mustafa A. Amin et al. “Gravitational waves from asymmetric oscillon dynamics?” *Phys. Rev. D* 98 (2018), 024040. DOI: 10.1103/PhysRevD.98.024040. ARXIV: 1803.08047.

[280] Kaloian D. Lozanov and Mustafa A. Amin. “Gravitational perturbations from oscillons and transients after inflation”. *Phys. Rev. D* 99.12 (2019), 123504. DOI: 10.1103/PhysRevD.99.123504. ARXIV: 1902.06736.

[281] Takashi Hiramatsu, Evangelos I. Sfakianakis, and Masahide Yamaguchi. “Gravitational wave spectra from oscillon formation after inflation”. *JHEP* 03 (2021), 021. DOI: 10.1007/JHEP03(2021)021. ARXIV: 2011.12201.

[282] Shinta Kasuya and Masahiro Kawasaki. “Baryogenesis from the gauge-mediation type Q-ball and the new type of Q-ball as the dark matter”. *Phys. Rev. D* 89.10 (2014), 103534. DOI: 10.1103/PhysRevD.89.103534. ARXIV: 1402.4546.

[283] Yang Bai et al. “Catalyzed baryogenesis”. *JHEP* 10 (2021), 147. DOI: 10.1007/JHEP10(2021)147. ARXIV: 2106.12589.

[284] E. Ya. Nugaev and A. V. Shkerin. “Review of Nontopological Solitons in Theories with $U(1)$-Symmetry”. *J. Exp. Theor. Phys.* 130.2 (2020), 301–320. DOI: 10.1134/S1063776120020077. ARXIV: 1905.05146.

[285] Shuang-Yong Zhou. “Non-topological solitons and quasi-solitons”. *Rept. Prog. Phys.* 88.4 (2025), 046901. DOI: 10.1088/1361-6633/adc69e. ARXIV: 2411.16604.

[286] N. Aghanim et al. “Planck 2018 results. I. Overview and the cosmological legacy of Planck”. *Astron. Astrophys.* 641 (2020), A1. DOI: 10.1051/0004-6361/201833880. ARXIV: 1807.06205.

[287] Brian D. Fields et al. “Big-Bang Nucleosynthesis after Planck”. *JCAP* 03 (2020). [Erratum: JCAP 11, E02 (2020)], 010. DOI: 10.1088/1475-7516/2020/03/010. ARXIV: 1912.01132.

[288] A. D. Sakharov. “Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe”. *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[289] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov. “On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe”. *Phys. Lett. B* 155 (1985), 36. DOI: 10.1016/0370-2693(85)91028-7.

[290] Andrew G. Cohen, D. B. Kaplan, and A. E. Nelson. “Progress in electroweak baryogenesis”. *Ann. Rev. Nucl. Part. Sci.* 43 (1993), 27–70. DOI: 10.1146/annurev.ns.43.120193.000331. ARXIV: hep-ph/9302210.

[291] Mark Trodden. “Electroweak baryogenesis”. *Rev. Mod. Phys.* 71 (1999), 1463–1500. DOI: 10.1103/RevModPhys.71.1463. ARXIV: hep-ph/9803479.

[292] Antonio Riotto and Mark Trodden. “Recent progress in baryogenesis”. *Ann. Rev. Nucl. Part. Sci.* 49 (1999), 35–75. DOI: 10.1146/annurev.nucl.49.1.35. ARXIV: hep-ph/9901362.

[293] David E. Morrissey and Michael J. Ramsey-Musolf. “Electroweak baryogenesis”. *New J. Phys.* 14 (2012), 125003. DOI: 10.1088/1367-2630/14/12/125003. ARXIV: 1206.2942.

[294] Dietrich Bodeker and Wilfried Buchmuller. “Baryogenesis from the weak scale to the grand unification scale”. *Rev. Mod. Phys.* 93.3 (2021), 035004. DOI: 10.1103/RevModPhys.93.035004. ARXIV: 2009.07294.

[295] Michael J. Ramsey-Musolf. "The electroweak phase transition: a collider target". *JHEP* 09 (2020), 179. DOI: 10.1007/JHEP09(2020)179. ARXIV: 1912.07189.

[296] Neil Turok and John Zadrozny. "Phase transitions in the two doublet model". *Nucl. Phys. B* 369 (1992), 729–742. DOI: 10.1016/0550-3213(92)90284-I.

[297] A. T. Davies et al. "Baryogenesis constraints on two Higgs doublet models". *Phys. Lett. B* 336 (1994), 464–470. DOI: 10.1016/0370-2693(94)90559-2.

[298] Jens O. Andersen et al. "Nonperturbative Analysis of the Electroweak Phase Transition in the Two Higgs Doublet Model". *Phys. Rev. Lett.* 121.19 (2018), 191802. DOI: 10.1103/PhysRevLett.121.191802. ARXIV: 1711.09849.

[299] Mark Trodden, Anne-Christine Davis, and Robert H. Brandenberger. "Particle physics models, topological defects and electroweak baryogenesis". *Phys. Lett. B* 349 (1995), 131–136. DOI: 10.1016/0370-2693(95)00214-6. ARXIV: hep-ph/9412266.

[300] Tomislav Prokopec et al. "Dynamical breaking of CPT symmetry in defect networks and baryogenesis". *Phys. Lett. B* 384 (1996), 175–179. DOI: 10.1016/0370-2693(96)00602-8. ARXIV: hep-ph/9511349.

[301] Robert H. Brandenberger, Wessyl Kelly, and Masahide Yamaguchi. "Electroweak baryogenesis with embedded domain walls". *Prog. Theor. Phys.* 117 (2007), 823–834. DOI: 10.1143/PTP.117.823. ARXIV: hep-ph/0503211.

[302] James M. Cline et al. "String mediated electroweak baryogenesis: A Critical analysis". *Phys. Rev. D* 59 (1999), 065014. DOI: 10.1103/PhysRevD.59.065014. ARXIV: hep-ph/9810261.

[303] Johanna Karouby and Robert Brandenberger. "Effects of a Thermal Bath of Photons on Embedded String Stability". *Phys. Rev. D* 85 (2012), 107702. DOI: 10.1103/PhysRevD.85.107702. ARXIV: 1203.0073.

[304] K. Jansen and M. Laine. "Inverse symmetry breaking with 4-D lattice simulations". *Phys. Lett. B* 435 (1998), 166–174. DOI: 10.1016/S0370-2693(98)00775-8. ARXIV: hep-lat/9805024.

[305] G. Bimonte et al. "Inverse symmetry breaking on the lattice: An Accurate MC study". *Nucl. Phys. B* 559 (1999), 103–122. DOI: 10.1016/S0550-3213(99)00421-6. ARXIV: hep-lat/9903027.

[306] Iason Baldes and Géraldine Servant. "High scale electroweak phase transition: baryogenesis \& symmetry non-restoration". *JHEP* 10 (2018), 053. DOI: 10.1007/JHEP10(2018)053. ARXIV: 1807.08770.

[307] Xinmin Zhang, Tao Huang, and Robert H. Brandenberger. "Pion and eta strings". *Phys. Rev. D* 58 (1998), 027702. DOI: 10.1103/PhysRevD.58.027702. ARXIV: hep-ph/9711452.

[308] Robert H. Brandenberger and Xin-min Zhang. "Anomalous global strings and primordial magnetic fields". *Phys. Rev. D* 59 (1999), 081301. DOI: 10.1103/PhysRevD.59.081301. ARXIV: hep-ph/9808306.

[309] Gerard 't Hooft. "Symmetry Breaking Through Bell-Jackiw Anomalies". *Phys. Rev. Lett.* 37 (1976). Edited by Mikhail A. Shifman, 8–11. DOI: 10.1103/PhysRevLett.37.8.

[310] Frans R. Klinkhamer and N. S. Manton. "A Saddle Point Solution in the Weinberg-Salam Theory". *Phys. Rev. D* 30 (1984), 2212. DOI: 10.1103/PhysRevD.30.2212.

[311] Michael Joyce, Tomislav Prokopec, and Neil Turok. "Nonlocal electroweak baryogenesis. Part 1: Thin wall regime". *Phys. Rev. D* 53 (1996), 2930–2957. DOI: 10.1103/PhysRevD.53.2930. ARXIV: hep-ph/9410281.

[312] Jan Ambjorn et al. “Sphaleron transitions and baryon asymmetry: A Numerical real time analysis”. *Nucl. Phys. B* 353 (1991), 346–378. DOI: 10.1016/0550-3213(91)90341-T.

[313] Dietrich Bodeker. “On the effective dynamics of soft nonAbelian gauge fields at finite temperature”. *Phys. Lett. B* 426 (1998), 351–360. DOI: 10.1016/S0370-2693(98)00279-2. ARXIV: hep-ph/9801430.

[314] Guy D. Moore. “Sphaleron rate in the symmetric electroweak phase”. *Phys. Rev. D* 62 (2000), 085011. DOI: 10.1103/PhysRevD.62.085011. ARXIV: hep-ph/0001216.

[315] Andrew G. Cohen, D. B. Kaplan, and A. E. Nelson. “Spontaneous baryogenesis at the weak phase transition”. *Phys. Lett. B* 263 (1991), 86–92. DOI: 10.1016/0370-2693(91)91711-4.

[316] Neil Turok and John Zadrozny. “Electroweak baryogenesis in the two doublet model”. *Nucl. Phys. B* 358 (1991), 471–493. DOI: 10.1016/0550-3213(91)90356-3.

[317] A. E. Nelson, D. B. Kaplan, and Andrew G. Cohen. “Why there is something rather than nothing: Matter from weak interactions”. *Nucl. Phys. B* 373 (1992), 453–478. DOI: 10.1016/0550-3213(92)90440-M.

[318] Andrew G. Cohen, D. B. Kaplan, and A. E. Nelson. “Diffusion enhances spontaneous electroweak baryogenesis”. *Phys. Lett. B* 336 (1994), 41–47. DOI: 10.1016/0370-2693(94)00935-X. ARXIV: hep-ph/9406345.

[319] Michael Joyce, Tomislav Prokopec, and Neil Turok. “Electroweak baryogenesis from a classical force”. *Phys. Rev. Lett.* 75 (1995). [Erratum: Phys.Rev.Lett. 75, 3375 (1995)], 1695–1698. DOI: 10.1103/PhysRevLett.75.1695. ARXIV: hep-ph/9408339.

[320] Michael Joyce, Tomislav Prokopec, and Neil Turok. “Nonlocal electroweak baryogenesis. Part 2: The Classical regime”. *Phys. Rev. D* 53 (1996), 2958–2980. DOI: 10.1103/PhysRevD.53.2958. ARXIV: hep-ph/9410282.

[321] James M. Cline, Kimmo Kainulainen, and Axel P. Vischer. “Dynamics of two Higgs doublet CP violation and baryogenesis at the electroweak phase transition”. *Phys. Rev. D* 54 (1996), 2451–2472. DOI: 10.1103/PhysRevD.54.2451. ARXIV: hep-ph/9506284.

[322] Björn Garbrecht. “Why is there more matter than antimatter? Calculational methods for leptogenesis and electroweak baryogenesis”. *Prog. Part. Nucl. Phys.* 110 (2020), 103727. DOI: 10.1016/j.ppnp.2019.103727. ARXIV: 1812.02651.

[323] Yuan-Zhen Li, Michael J. Ramsey-Musolf, and Jiang-Hao Yu. “Does the Electron EDM Preclude Electroweak Baryogenesis ?” (2024). ARXIV: 2404.19197.

[324] Theodore Garagounis and Mark Hindmarsh. “Scaling in numerical simulations of domain walls”. *Phys. Rev. D* 68 (2003), 103506. DOI: 10.1103/PhysRevD.68.103506. ARXIV: hep-ph/0212359.

[325] A. M. M. Leite and C. J. A. P. Martins. “Scaling Properties of Domain Wall Networks”. *Phys. Rev. D* 84 (2011), 103523. DOI: 10.1103/PhysRevD.84.103523. ARXIV: 1110.3486.

[326] C. J. A. P. Martins et al. “Extending the velocity-dependent one-scale model for domain walls”. *Phys. Rev. D* 93.4 (2016), 043534. DOI: 10.1103/PhysRevD.93.043534. ARXIV: 1602.01322.

[327] D. Grüber, L. Sousa, and P. P. Avelino. “Stochastic gravitational wave background generated by domain wall networks”. *Phys. Rev. D* 110.2 (2024), 023505. DOI: 10.1103/PhysRevD.110.023505. ARXIV: 2403.09816.

[328] Rabindra N. Mohapatra and Jogesh C. Pati. “Left-Right Gauge Symmetry and an Isoconjugate Model of CP Violation”. *Phys. Rev. D* 11 (1975), 566–571. DOI: 10.1103/PhysRevD.11.566.

[329] R. N. Mohapatra and Jogesh C. Pati. “A Natural Left-Right Symmetry”. *Phys. Rev. D* 11 (1975), 2558. DOI: 10.1103/PhysRevD.11.2558.

[330] G. Senjanovic and Rabindra N. Mohapatra. “Exact Left-Right Symmetry and Spontaneous Violation of Parity”. *Phys. Rev. D* 12 (1975), 1502. DOI: 10.1103/PhysRevD.12.1502.

[331] Graciela B. Gelmini, Marcelo Gleiser, and Edward W. Kolb. “Cosmology of Biased Discrete Symmetry Breaking”. *Phys. Rev. D* 39 (1989), 1558. DOI: 10.1103/PhysRevD.39.1558.

[332] D. Coulson, Z. Lalak, and Burt A. Ovrut. “Biased domain walls”. *Phys. Rev. D* 53 (1996), 4237–4246. DOI: 10.1103/PhysRevD.53.4237.

[333] Sebastian E. Larsson, Subir Sarkar, and Peter L. White. “Evading the cosmological domain wall problem”. *Phys. Rev. D* 55 (1997), 5129–5135. DOI: 10.1103/PhysRevD.55.5129. ARXIV: hep-ph/9608319.

[334] P. P. Avelino, C. J. A. P. Martins, and L. Sousa. “Dynamics of Biased Domain Walls and the Devaluation Mechanism”. *Phys. Rev. D* 78 (2008), 043521. DOI: 10.1103/PhysRevD.78.043521. ARXIV: 0805.4013.

[335] D. Förster. “Dynamics of Relativistic Vortex Lines and their Relation to Dual Theory”. *Nucl. Phys. B* 81 (1974), 84–92. DOI: 10.1016/0550-3213(74)90008-X.

[336] Y. Nambu. “Quark model and the factorization of the Veneziano amplitude”. *Proceedings of the International Conference on Symmetries and Quark Models* 65 (1969), 269–277.

[337] Tetsuo Goto. “Relativistic quantum mechanics of one-dimensional mechanical continuum and subsidiary condition of dual resonance model”. *Prog. Theor. Phys.* 46 (1971), 1560–1569. DOI: 10.1143/PTP.46.1560.

[338] Carlo Maccaferri, Fabio Marino, and Beniamino Valsesia. “Introduction to string theory”. *SciPost Phys. Lect. Notes* 90 (2025), 1. DOI: 10.21468/SciPostPhysLectNotes.90. ARXIV: 2311.18111.

[339] Kei-ichi Maeda and Neil Turok. “Finite Width Corrections to the Nambu Action for the Nielsen-Olesen String”. *Phys. Lett. B* 202 (1988), 376–380. DOI: 10.1016/0370-2693(88)90488-1.

[340] R. Gregory. “Effective Action for a Cosmic String”. *Phys. Lett. B* 206 (1988), 199–204. DOI: 10.1016/0370-2693(88)91492-X.

[341] Malcolm R. Anderson et al. “Effective action and motion of a cosmic string”. *Phys. Rev. D* 56 (1997), 8014–8028. DOI: 10.1103/PhysRevD.56.8014. ARXIV: hep-ph/9707324.

[342] Neil Turok. “Grand Unified Strings and Galaxy Formation”. *Nucl. Phys. B* 242 (1984), 520–541. DOI: 10.1016/0550-3213(84)90407-3.

[343] Ken D. Olum and J. J. Blanco-Pillado. “Field theory simulation of Abelian Higgs cosmic string cusps”. *Phys. Rev. D* 60 (1999), 023503. DOI: 10.1103/PhysRevD.60.023503. ARXIV: gr-qc/9812040.

[344] Jose J. Blanco-Pillado et al. “Nambu-Goto dynamics of field theory cosmic string loops”. *JCAP* 05 (2023), 035. DOI: 10.1088/1475-7516/2023/05/035. ARXIV: 2302.03717.

[345] P. J. Ruback. “Vortex String Motion in the Abelian Higgs Model”. *Nucl. Phys. B* 296 (1988), 669–678. DOI: 10.1016/0550-3213(88)90038-7.

[346] N. S. Manton. “A Remark on the Scattering of BPS Monopoles”. *Phys. Lett. B* 110 (1982), 54–56. DOI: 10.1016/0370-2693(82)90950-9.

[347] D. Stuart. “The Geodesic approximation for the Yang-Mills Higgs equations”. *Commun. Math. Phys.* 166 (1994), 149–190. DOI: 10.1007/BF02099305.

[348] L. M. A. Bettencourt and T. W. B. Kibble. “Nonintercommuting configurations in the collisions of type I U(1) cosmic strings”. *Phys. Lett. B* 332 (1994), 297–304. DOI: 10.1016/0370-2693(94)91257-2. ARXIV: hep-ph/9405221.

[349] Luis M. A. Bettencourt, Pablo Laguna, and Richard A. Matzner. “Nonintercommuting cosmic strings”. *Phys. Rev. Lett.* 78 (1997), 2066–2069. DOI: 10.1103/PhysRevLett.78.2066. ARXIV: hep-ph/9612350.

[350] E. P. S. Shellard. “Cosmic String Interactions”. *Nucl. Phys. B* 283 (1987), 624–656. DOI: 10.1016/0550-3213(87)90290-2.

[351] Richard A. Matzner. “Interaction of U(1) cosmic strings: Numerical intercommutation”. *Comput. Phys.* 2.5 (1988), 51–64. DOI: 10.1063/1.168306.

[352] Amihay Hanany and Koji Hashimoto. “Reconnection of colliding cosmic strings”. *JHEP* 06 (2005), 021. DOI: 10.1088/1126-6708/2005/06/021. ARXIV: hep-th/0501031.

[353] Ana Achúcarro and Roland de Putter. “Effective nonintercommutation of local cosmic strings at high collision speeds”. *Phys. Rev. D* 74 (12 2006), 121701. DOI: 10.1103/PhysRevD.74.121701.

[354] Minoru Eto et al. “Universal Reconnection of Non-Abelian Cosmic Strings”. *Phys. Rev. Lett.* 98 (9 2007), 091602. DOI: 10.1103/PhysRevLett.98.091602.

[355] Mark G. Jackson, Nicholas T. Jones, and Joseph Polchinski. “Collisions of cosmic F and D-strings”. *JHEP* 10 (2005), 013. DOI: 10.1088/1126-6708/2005/10/013. ARXIV: hep-th/0405229.

[356] Masaki Yamada and Kazuya Yonekura. “Cosmic strings from pure Yang–Mills theory”. *Phys. Rev. D* 106.12 (2022), 123515. DOI: 10.1103/PhysRevD.106.123515. ARXIV: 2204.13123.

[357] Masaki Yamada and Kazuya Yonekura. “Cosmic F- and D-strings from pure Yang–Mills theory”. *Phys. Lett. B* 838 (2023), 137724. DOI: 10.1016/j.physletb.2023.137724. ARXIV: 2204.13125.

[358] A. Achucarro and G. J. Verbiest. “Higher order intercommutations in Cosmic String Collisions”. *Phys. Rev. Lett.* 105 (2010), 021601. DOI: 10.1103/PhysRevLett.105.021601. ARXIV: 1006.0979.

[359] G. J. Verbiest and A. Achucarro. “High speed collision and reconnection of Abelian Higgs strings in the deep type-II regime”. *Phys. Rev. D* 84 (2011), 105036. DOI: 10.1103/PhysRevD.84.105036. ARXIV: 1106.4666.

[360] Tanmay Vachaspati and Alexander Vilenkin. “Gravitational Radiation from Cosmic Strings”. *Phys. Rev. D* 31 (1985), 3052. DOI: 10.1103/PhysRevD.31.3052.

[361] Jose J. Blanco-Pillado, Ken D. Olum, and Benjamin Shlaer. “Cosmic string loop shapes”. *Phys. Rev. D* 92.6 (2015), 063528. DOI: 10.1103/PhysRevD.92.063528. ARXIV: 1508.02693.

[362] E. J. Copeland and T. W. B. Kibble. “Kinks and small-scale structure on cosmic strings”. *Phys. Rev. D* 80 (2009), 123523. DOI: 10.1103/PhysRevD.80.123523. ARXIV: 0909.1960.

[363] Jeremy M. Wachter et al. “Numerical gravitational backreaction on cosmic string loops from simulation” (2024). ARXIV: 2411.10366.

[364] Jeremy M. Wachter, Ken D. Olum, and Jose J. Blanco-Pillado. “More accurate gravitational wave backgrounds from cosmic strings” (2024). ARXIV: 2411.16590.

[365] T. W. B. Kibble and Neil Turok. “Selfintersection of Cosmic Strings”. *Phys. Lett. B* 116 (1982), 141–143. DOI: 10.1016/0370-2693(82)90993-5.

[366] A. L. Chen, D. A. DiCarlo, and S. A. Hotes. “Selfintersections in a Three Parameter Space of Cosmic Strings”. *Phys. Rev. D* 37 (1988), 863. DOI: 10.1103/PhysRevD.37.863.

[367] David Delaney, Kimberly Engle, and Xania Scheick. “The General Two Harmonic Cosmic String”. *Phys. Rev. D* 41 (1990), 1775. DOI: 10.1103/PhysRevD.41.1775.

[368] C. J. Burden and L. J. Tassie. “Additional Rigidly Rotating Solutions in the String Model of Hadrons”. *Austral. J. Phys.* 37 (1984), 1–7. DOI: 10.1071/PH840001.

[369] C. J. Burden. “Gravitational Radiation From a Particular Class of Cosmic Strings”. *Phys. Lett. B* 164 (1985), 277–281. DOI: 10.1016/0370-2693(85)90326-0.

[370] Bruce Allen, Paul Casper, and Adrian Ottewill. “Analytic results for the gravitational radiation from a class of cosmic string loops”. *Phys. Rev. D* 50 (1994), 3703–3712. DOI: 10.1103/PhysRevD.50.3703. ARXIV: gr-qc/9405037.

[371] David B. DeLaney and Robert W. Brown. “A Product Representation for the Harmonic Series of a Unit Vector: A String Application”. *Phys. Rev. Lett.* 63 (1989), 474. DOI: 10.1103/PhysRevLett.63.474.

[372] R. W. Brown et al. “Closed strings with low harmonics and kinks”. *Phys. Rev. D* 48 (1993), 2548–2562. DOI: 10.1103/PhysRevD.48.2548. ARXIV: hep-th/9307077.

[373] Xavier A. Siemens and T. W. B. Kibble. “High harmonic configurations of cosmic strings: An Analysis of selfintersections”. *Nucl. Phys. B* 438 (1995), 307–319. DOI: 10.1016/0550-3213(94)00592-3. ARXIV: hep-ph/9412216.

[374] Despoina Pazouli. “High-harmonic cosmic strings and gravitational waves”. PhD thesis. Nottingham U., Nottingham U., 2020. ARXIV: 2108.08242.

[375] Neil Turok and Pijushpani Bhattacharjee. “Strechting Cosmic Strings”. *Phys. Rev. D* 29 (1984), 1557. DOI: 10.1103/PhysRevD.29.1557.

[376] A. Vilenkin. “Cosmic Strings”. *Phys. Rev. D* 24 (1981), 2082–2089. DOI: 10.1103/PhysRevD.24.2082.

[377] C. J. A. P. Martins and E. P. S. Shellard. “Quantitative string evolution”. *Phys. Rev. D* 54 (1996), 2535–2556. DOI: 10.1103/PhysRevD.54.2535. ARXIV: hep-ph/9602271.

[378] C. J. A. P. Martins and E. P. S. Shellard. “Extending the velocity-dependent one-scale string evolution model”. *Phys. Rev. D* 65 (4 2002), 043514. DOI: 10.1103/PhysRevD.65.043514.

[379] C. J. A. P. Martins, J. N. Moore, and E. P. S. Shellard. “A Unified model for vortex string network evolution”. *Phys. Rev. Lett.* 92 (2004), 251601. DOI: 10.1103/PhysRevLett.92.251601. ARXIV: hep-ph/0310255.

[380] Ken’ichi Saikawa and Satoshi Shirai. “Precise WIMP Dark Matter Abundance and Standard Model Thermodynamics”. *JCAP* 08 (2020), 011. DOI: 10.1088/1475-7516/2020/08/011. ARXIV: 2005.03544.

[381] R. M. Rohm. “Some current problems in particle physics beyond the standard model”. PhD thesis. Princeton University, 1985.

[382] P. de Sousa Gerbert and R. Jackiw. “Classical and Quantum Scattering on a Spinning Cone”. *Commun. Math. Phys.* 124 (1989), 229. DOI: 10.1007/BF01219196.

[383] Mark G. Alford and Frank Wilczek. “Aharonov-Bohm Interaction of Cosmic Strings with Matter”. *Phys. Rev. Lett.* 62 (1989), 1071. DOI: 10.1103/PhysRevLett.62.1071.

[384] A. Vilenkin. “Cosmic string dynamics with friction”. *Phys. Rev. D* 43 (1991), 1060–1062. DOI: 10.1103/PhysRevD.43.1060.

[385] C. J. A. P. Martins and E. P. S. Shellard. “Extending the velocity dependent one scale string evolution model”. *Phys. Rev. D* 65 (2002), 043514. DOI: 10.1103/PhysRevD.65.043514. ARXIV: hep-ph/0003298.

[386] S. Mukovnikov and L. Sousa. “Ultra-high frequency gravitational waves from cosmic strings with friction” (2024). ARXIV: 2404.13213.

[387] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: John Wiley and Sons, 1972. ISBN: 978-0-471-92567-5.

[388] Bruce Allen and E. P. S. Shellard. “Gravitational radiation from cosmic strings”. *Phys. Rev. D* 45 (1992), 1898–1912. DOI: 10.1103/PhysRevD.45.1898.

[389] Malcolm R. Anderson. “Self-similar evaporation of a rigidly-rotating cosmic string loop”. *Class. Quant. Grav.* 22 (2005), 2539–2568. DOI: 10.1088/0264-9381/22/13/002. ARXIV: gr-qc/0505160.

[390] Jose J. Blanco-Pillado and Ken D. Olum. “Stochastic gravitational wave background from smoothed cosmic string loops”. *Phys. Rev. D* 96.10 (2017), 104046. DOI: 10.1103/PhysRevD.96.104046. ARXIV: 1709.02693.

[391] P. Binetruy et al. “Gravitational Wave Bursts from Cosmic Superstrings with Y-junctions”. *Phys. Rev. D* 80 (2009), 123510. DOI: 10.1103/PhysRevD.80.123510. ARXIV: 0907.4522.

[392] P. A. R. Ade et al. “Planck 2015 results. XIII. Cosmological parameters”. *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. ARXIV: 1502.01589.

[393] Joanes Lizarraga et al. “New CMB constraints for Abelian Higgs cosmic strings”. *JCAP* 10 (2016), 042. DOI: 10.1088/1475-7516/2016/10/042. ARXIV: 1609.03386.

[394] Pierre Auclair et al. “Probing the gravitational wave background from cosmic strings with LISA”. *JCAP* 04 (2020), 034. DOI: 10.1088/1475-7516/2020/04/034. ARXIV: 1909.00819.

[395] David P. Bennett and Francois R. Bouchet. “High resolution simulations of cosmic string evolution. 1. Network evolution”. *Phys. Rev. D* 41 (1990), 2408. DOI: 10.1103/PhysRevD.41.2408.

[396] L. Sousa and P. P. Avelino. “Stochastic Gravitational Wave Background generated by Cosmic String Networks: Velocity-Dependent One-Scale model versus Scale-Invariant Evolution”. *Phys. Rev. D* 88.2 (2013), 023516. DOI: 10.1103/PhysRevD.88.023516. ARXIV: 1304.2445.

[397] Larissa Lorenz, Christophe Ringeval, and Mairi Sakellariadou. “Cosmic string loop distribution on all length scales and at any redshift”. *JCAP* 10 (2010), 003. DOI: 10.1088/1475-7516/2010/10/003. ARXIV: 1006.0931.

[398] Christophe Ringeval, Mairi Sakellariadou, and Francois Bouchet. “Cosmological evolution of cosmic string loops”. *JCAP* 02 (2007), 023. DOI: 10.1088/1475-7516/2007/02/023. ARXIV: astro-ph/0511646.

[399] Christophe Ringeval and Teruaki Suyama. “Stochastic gravitational waves from cosmic string loops in scaling”. *JCAP* 12 (2017), 027. DOI: 10.1088/1475-7516/2017/12/027. ARXIV: 1709.03845.

[400] Pierre G. Auclair. “Impact of the small-scale structure on the Stochastic Background of Gravitational Waves from cosmic strings”. *JCAP* 11 (2020), 050. DOI: 10.1088/1475-7516/2020/11/050. ARXIV: 2009.00334.

[401] Pierre Auclair et al. "Cosmic string loop production functions". *JCAP* 06 (2019), 015. DOI: 10.1088/1475-7516/2019/06/015. ARXIV: 1903.06685.

[402] Jose J. Blanco-Pillado and Ken D. Olum. "Direct determination of cosmic string loop density from simulations". *Phys. Rev. D* 101.10 (2020), 103018. DOI: 10.1103/PhysRevD.101.103018. ARXIV: 1912.10017.

[403] Lara Sousa, Pedro P. Avelino, and Guilherme S. F. Guedes. "Full analytical approximation to the stochastic gravitational wave background generated by cosmic string networks". *Phys. Rev. D* 101.10 (2020), 103508. DOI: 10.1103/PhysRevD.101.103508. ARXIV: 2002.01079.

[404] Michele Maggiore. *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*. Oxford University Press, 2018. ISBN: 978-0-19-857089-9.

[405] S. A. Sanidas, R. A. Battye, and B. W. Stappers. "Constraints on cosmic string tension imposed by the limit on the stochastic gravitational wave background from the European Pulsar Timing Array". *Phys. Rev. D* 85 (2012), 122003. DOI: 10.1103/PhysRevD.85.122003. ARXIV: 1201.2419.

[406] Mark Hindmarsh et al. "Loop decay in Abelian-Higgs string networks". *Phys. Rev. D* 104.4 (2021), 043519. DOI: 10.1103/PhysRevD.104.043519. ARXIV: 2103.16248.

[407] Jun'ya Kume and Mark Hindmarsh. "Revised bounds on local cosmic strings from NANOGrav observations". *JCAP* 12 (2024), 001. DOI: 10.1088/1475-7516/2024/12/001. ARXIV: 2404.02705.

[408] Hanyu Cheng and Luca Visinelli. "Future targets for light gauge bosons from cosmic strings". *Phys. Dark Univ.* 46 (2024), 101667. DOI: 10.1016/j.dark.2024.101667. ARXIV: 2408.16334.

[409] Jorge Baeza-Ballesteros et al. "Gravitational Wave Emission from Cosmic String Loops, II: Local Case" (2024). ARXIV: 2408.02364.

[410] Zizhuo Zhao, Ligong Bian, and Jing Shu. "Testing Nambu-Goto approximation of cosmic string by lattice field simulations" (2025). ARXIV: 2507.00685.

[411] Jose J. Blanco-Pillado et al. "Gravitational waves from cosmic strings in LISA: reconstruction pipeline and physics interpretation" (2024). ARXIV: 2405.03740.

[412] Yann Gouttenoire, Géraldine Servant, and Peera Simakachorn. "Beyond the Standard Models with Cosmic Strings". *JCAP* 07 (2020), 032. DOI: 10.1088/1475-7516/2020/07/032. ARXIV: 1912.02569.

[413] Jeff A. Dror et al. "Testing the Seesaw Mechanism and Leptogenesis with Gravitational Waves". *Phys. Rev. Lett.* 124.4 (2020), 041804. DOI: 10.1103/PhysRevLett.124.041804. ARXIV: 1908.03227.

[414] Wilfried Buchmüller et al. "Probing the scale of grand unification with gravitational waves". *Phys. Lett. B* 809 (2020), 135764. DOI: 10.1016/j.physletb.2020.135764. ARXIV: 1912.03695.

[415] Géraldine Servant and Peera Simakachorn. "Ultrahigh frequency primordial gravitational waves beyond the kHz: The case of cosmic strings". *Phys. Rev. D* 109.10 (2024), 103538. DOI: 10.1103/PhysRevD.109.103538. ARXIV: 2312.09281.

[416] Daiju Matsunami et al. "Decay of Cosmic String Loops Due to Particle Radiation". *Phys. Rev. Lett.* 122.20 (2019), 201301. DOI: 10.1103/PhysRevLett.122.201301. ARXIV: 1903.05102.

[417] P. Binetruy et al. "Proliferation of sharp kinks on cosmic (super-)string loops with junctions". *Phys. Rev. D* 82 (2010), 083524. DOI: 10.1103/PhysRevD.82.083524. ARXIV: 1005.2426.

[418] P. Binetruy et al. “Gravitational wave signatures from kink proliferation on cosmic (super-) strings”. *Phys. Rev. D* 82 (2010), 126007. DOI: 10.1103/PhysRevD.82.126007. ARXIV: 1009.2484.

[419] Kai Schmitz. “New Sensitivity Curves for Gravitational-Wave Signals from Cosmological Phase Transitions”. *JHEP* 01 (2021), 097. DOI: 10.1007/JHEP01(2021)097. ARXIV: 2002.04615.

[420] Thibault Damour and Alexander Vilenkin. “Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows”. *Phys. Rev. D* 71 (2005), 063510. DOI: 10.1103/PhysRevD.71.063510. ARXIV: hep-th/0410222.

[421] Xavier Siemens et al. “Gravitational wave bursts from cosmic (super)strings: Quantitative analysis and constraints”. *Phys. Rev. D* 73 (2006), 105001. DOI: 10.1103/PhysRevD.73.105001. ARXIV: gr-qc/0603115.

[422] M. C. Guzzetti et al. “Gravitational waves from inflation”. *Riv. Nuovo Cim.* 39.9 (2016), 399–495. DOI: 10.1393/ncr/i2016-10127-1. ARXIV: 1605.01615.

[423] Chiara Caprini et al. “Gravitational waves from first-order phase transitions in LISA: reconstruction pipeline and physics interpretation”. *JCAP* 10 (2024), 020. DOI: 10.1088/1475-7516/2024/10/020. ARXIV: 2403.03723.

[424] Gia Dvali and Alexander Vilenkin. “Formation and evolution of cosmic D strings”. *JCAP* 03 (2004), 010. DOI: 10.1088/1475-7516/2004/03/010. ARXIV: hep-th/0312007.

[425] Edmund J. Copeland, Robert C. Myers, and Joseph Polchinski. “Cosmic F and D strings”. *JHEP* 06 (2004), 013. DOI: 10.1088/1126-6708/2004/06/013. ARXIV: hep-th/0312067.

[426] John Ellis et al. “What is the source of the PTA GW signal?” *Phys. Rev. D* 109.2 (2024), 023522. DOI: 10.1103/PhysRevD.109.023522. ARXIV: 2308.08546.

[427] L. Sousa and P. P. Avelino. “Probing Cosmic Superstrings with Gravitational Waves”. *Phys. Rev. D* 94.6 (2016), 063529. DOI: 10.1103/PhysRevD.94.063529. ARXIV: 1606.05585.

[428] Anastasios Avgoustidis et al. “The stochastic gravitational wave background from cosmic superstrings” (2025). ARXIV: 2503.10361.

[429] R. A. Battye, R. R. Caldwell, and E. P. S. Shellard. “Gravitational waves from cosmic strings”. *Conference on Topological Defects and CMB*. 1997, 11–31. ARXIV: astro-ph/9706013.

[430] Yanou Cui et al. “Probing the pre-BBN universe with gravitational waves from cosmic strings”. *JHEP* 01 (2019), 081. DOI: 10.1007/JHEP01(2019)081. ARXIV: 1808.08968.

[431] Robert R. Caldwell, Tristan L. Smith, and Devin G. E. Walker. “Using a Primordial Gravitational Wave Background to Illuminate New Physics”. *Phys. Rev. D* 100.4 (2019), 043513. DOI: 10.1103/PhysRevD.100.043513. ARXIV: 1812.07577.

[432] Stefan Antusch et al. “Probing SUSY at Gravitational Wave Observatories” (2024). ARXIV: 2405.03746.

[433] The NANOGrav Collaboration. *Noise Spectra and Stochastic Background Sensitivity Curve for the NG15-year Dataset*. Version 1.0.0. Zenodo, 2023. DOI: 10.5281/zenodo.8092346.

[434] Kai Schmitz. *New Sensitivity Curves for Gravitational-Wave Experiments*. Version v1. Zenodo, 2020. DOI: 10.5281/zenodo.3689582.

[435] Olivia Bitcon, Kai Schmitz, and Tobias Schröder. in preparation. 2025.

[436] Hewish, A. and Bell, S. J. and Pilkington, J. D. H. and Scott, P. F. and Collins, R. A. "Observation of a Rapidly Pulsating Radio Source". *Nature* 217.5130 (1968), 709–713. DOI: 10.1038/217709a0.

[437] F. Pacini. "Rotating Neutron Stars, Pulsars and Supernova Remnants". *Nature* 219.5150 (1968), 145–146. DOI: 10.1038/219145a0.

[438] T. Gold. "Rotating neutron stars as the origin of the pulsating radio sources". *Nature* 218 (1968), 731–732. DOI: 10.1038/218731a0.

[439] Lorimer, D. R. and Kramer, M. *Handbook of Pulsar Astronomy*. Volume 4. Cambridge University Press, 2004.

[440] Duncan R. Lorimer. "Binary and Millisecond Pulsars". *Living Reviews in Relativity* 11.8 (1 2008). DOI: 10.12942/lrr-2008-8.

[441] Lyne, Andrew and Graham-Smith, Francis. *Pulsar Astronomy*. Cambridge University Press, 2012.

[442] D. C. Backer et al. "A millisecond pulsar". *Nature* 300.5893 (1982), 615–618. DOI: 10.1038/300615a0.

[443] Helfand, D. J. and Manchester, R. N. and Taylor, J. H. "Observations of pulsar radio emission. III. Stability of integrated profiles." *Astrophys. J.* 198 (1975), 661–670. DOI: –10.1086/153644˝.

[444] Stephen R. Taylor. "The Nanohertz Gravitational Wave Astronomer" (2021). ARXIV: 2105.13270.

[445] Luke Zoltan Kelley. "Pulsar Timing Arrays" (2025). ARXIV: 2505.00797.

[446] R. W. Hellings and G. S. Downs. "Upper limits on the isotropic gravitational radiation background from pulsar timing analysis". *Astrophys. J. Lett.* 265 (1983), L39–L42. DOI: 10.1086/183954.

[447] Matthew T. Miles et al. "The MeerKAT Pulsar Timing Array: the first search for gravitational waves with the MeerKAT radio telescope". *Mon. Not. Roy. Astron. Soc.* 536.2 (2024), 1489–1500. DOI: 10.1093/mnras/stae2571. ARXIV: 2412.01153.

[448] Gabriella Agazie et al. "The NANOGrav 15 yr Data Set: Observations and Timing of 68 Millisecond Pulsars". *Astrophys. J. Lett.* 951.1 (2023), L9. DOI: 10.3847/2041-8213/acda9a. ARXIV: 2306.16217.

[449] Gabriella Agazie et al. "The NANOGrav 15 yr Data Set: Constraints on Supermassive Black Hole Binaries from the Gravitational-wave Background". *Astrophys. J. Lett.* 952.2 (2023), L37. DOI: 10.3847/2041-8213/ace18b. ARXIV: 2306.16220.

[450] Simone Blasi, Vedran Brdar, and Kai Schmitz. "Has NANOGrav found first evidence for cosmic strings?" *Phys. Rev. Lett.* 126.4 (2021), 041305. DOI: 10.1103/PhysRevLett.126.041305. ARXIV: 2009.06607.

[451] John Ellis and Marek Lewicki. "Cosmic String Interpretation of NANOGrav Pulsar Timing Data". *Phys. Rev. Lett.* 126.4 (2021), 041304. DOI: 10.1103/PhysRevLett.126.041304. ARXIV: 2009.06555.

[452] Jose J. Blanco-Pillado, Ken D. Olum, and Jeremy M. Wachter. "Comparison of cosmic string and superstring models to NANOGrav 12.5-year results". *Phys. Rev. D* 103.10 (2021), 103512. DOI: 10.1103/PhysRevD.103.103512. ARXIV: 2102.08194.

[453] Daniel G. Figueroa et al. "Cosmological Background Interpretation of Pulsar Timing Array Data". *Phys. Rev. Lett.* 132.17 (2024), 171002. DOI: 10.1103/PhysRevLett.132.171002. ARXIV: 2307.02399.

[454] Michele Maggiore. *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford University Press, 2007. ISBN: 978-0-19-852074-0. DOI: 10.1093/acprof:oso/9780198570745.001.0001.

[455] Alexander C. Jenkins and Mairi Sakellariadou. "Anisotropies in the stochastic gravitational-wave background: Formalism and the cosmic string case". *Phys. Rev. D* 98.6 (2018), 063509. DOI: 10.1103/PhysRevD.98.063509. ARXIV: 1802.06046.

[456] Gabriella Agazie et al. "The NANOGrav 15 yr Data Set: Search for Anisotropy in the Gravitational-wave Background". *Astrophys. J. Lett.* 956.1 (2023), L3. DOI: 10.3847/2041-8213/acf4fd. ARXIV: 2306.16221.

[457] Chiara M. F. Mingarelli and Trevor Sidery. "Effect of small interpulsar distances in stochastic gravitational wave background searches with pulsar timing arrays". *Phys. Rev. D* 90.6 (2014), 062011. DOI: 10.1103/PhysRevD.90.062011. ARXIV: 1408.6840.

[458] Melissa Anholm et al. "Optimal strategies for gravitational wave stochastic background searches in pulsar timing data". *Phys. Rev. D* 79 (2009), 084030. DOI: 10.1103/PhysRevD.79.084030. ARXIV: 0809.0701.

[459] Joseph D. Romano and Bruce Allen. "Answers to frequently asked questions about the pulsar timing array Hellings and Downs curve". *Class. Quant. Grav.* 41.17 (2024), 175008. DOI: 10.1088/1361-6382/ad4c4c. ARXIV: 2308.05847.

[460] Richard A. Isaacson. "Gravitational Radiation in the Limit of High Frequency. I. The Linear Approximation and Geometrical Optics". *Phys. Rev.* 166 (1968), 1263–1271. DOI: 10.1103/PhysRev.166.1263.

[461] Richard A. Isaacson. "Gravitational Radiation in the Limit of High Frequency. II. Nonlinear Terms and the Effective Stress Tensor". *Phys. Rev.* 166 (1968), 1272–1279. DOI: 10.1103/PhysRev.166.1272.

[462] Gabriella Agazie et al. "The NANOGrav 15 yr Data Set: Detector Characterization and Noise Budget". *Astrophys. J. Lett.* 951.1 (2023), L10. DOI: 10.3847/2041-8213/acda88. ARXIV: 2306.16218.

[463] Zaven Arzoumanian, Adam Brazier, S. Burke-Spolaor, et al. "The NANOGrav Nine-year Data Set: Observations, Arrival Time Measurements, and Analysis of 37 Millisecond Pulsars". *Astrophys. J.* 813.1 (2015), 65. DOI: 10.1088/0004-637X/813/1/65. ARXIV: 1505.07540.

[464] Zaven Arzoumanian, Adam Brazier, S. Burke-Spolaor, et al. "The NANOGrav 11-year Data Set: High-precision timing of 45 Millisecond Pulsars". *Astrophys. J. Suppl.* 235.2 (2018), 37. DOI: 10.3847/1538-4365/aab5b0. ARXIV: 1801.01837.

[465] Md F. Alam, Z. Arzoumanian, P. T. Baker, et al. "The NANOGrav 12.5 yr Data Set: Observations and Narrowband Timing of 47 Millisecond Pulsars". *Astrophys. J. Suppl.* 252.1 (2021), 4. DOI: 10.3847/1538-4365/abc6a0. ARXIV: 2005.06490.

[466] M. L. Jones, M. A. McLaughlin, M. T. Lam, et al. "The NANOGrav Nine-year Data Set: Measurement and Analysis of Variations in Dispersion Measures". *Astrophys. J.* 841.2 (2017), 125. DOI: 10.3847/1538-4357/aa73df. ARXIV: 1612.03187.

[467] Ryan S. Park et al. "The JPL Planetary and Lunar Ephemerides DE440 and DE441". *"Astrophys. J."* 161.3, 105 (2021), 105. DOI: 10.3847/1538-3881/abd414.

[468] M. Vallisneri et al. "Modeling the uncertainties of solar-system ephemerides for robust gravitational-wave searches with pulsar timing arrays" (2020). DOI: 10.3847/1538-4357/ab7b67. ARXIV: 2001.00595.

[469] Rutger van Haasteren and Yuri Levin. “Understanding and analysing time-correlated stochastic signals in pulsar timing”. *Monthly Notices of the Royal Astronomical Society* 428.2 (2012), 1147–1159. DOI: 10.1093/mnras/sts097.

[470] Lindley Lentati et al. “Hyper-efficient model-independent Bayesian method for the analysis of pulsar timing data”. *Physical Review D* 87.10 (2013). DOI: 10.1103/physrevd.87.104021.

[471] Justin A. Ellis et al. *ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust PulsaR Inference SuitE*. Astrophysics Source Code Library, record ascl:1912.015. 2019.

[472] Stephen R. Taylor et al. *enterprise_extensions*. v2.3.3. 2021. `https://github.com/nanograv/enterprise_extensions`.

[473] Harold Jeffreys. *Theory of Probability*. Oxford University Press, 1961. ISBN: 9780191589676.

[474] James M. Dickey. “The Weighted Likelihood Ratio, Linear Hypotheses on Normal Location Parameters”. *The Annals of Mathematical Statistics* 42.1 (1971), 204–223.

[475] Adelchi Azzalini. *Statistical Inference: Based on the likelihood*. Boca Raton, London, New York, Washington D.C.: Chapman & Hall/CRC, 1996. ISBN: 978-0412606502.

[476] Justin Ellis and Rutger van Haasteren. *jellis18/PTMCMCSampler: Official Release*. 2017. DOI: 10.5281/zenodo.1037579.

[477] Antony Lewis. “GetDist: a Python package for analysing Monte Carlo samples” (2019). ARXIV: 1910.13970. `https://getdist.readthedocs.io`.

[478] Bradley P. Carlin and Siddhartha Chib. “Bayesian Model Choice via Markov Chain Monte Carlo Methods”. *Journal of the Royal Statistical Society. Series B (Methodological)* 57.3 (1995), 473–484.

[479] Simon J. Godsill. “On the Relationship between Markov Chain Monte Carlo Methods for Model Uncertainty”. *Journal of Computational and Graphical Statistics* 10.2 (2001), 230–248.

[480] S. Hee et al. “Bayesian model selection without evidences: application to the dark energy equation-of-state”. *Monthly Notices of the Royal Astronomical Society* 455.3 (2015), 2461–2473. DOI: 10.1093/mnras/stv2217.

[481] B. Efron and R. Tibshirani. “An introduction to the bootstrap”. *Statist. Sci.* 57.1 (1986), 54–75. DOI: 10.1214/ss/1177013815.

[482] Kazunori Akiyama et al. “First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole”. *Astrophys. J. Lett.* 875 (2019), L1. DOI: 10.3847/2041-8213/ab0ec7. ARXIV: 1906.11238.

[483] Kazunori Akiyama et al. “First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way”. *Astrophys. J. Lett.* 930.2 (2022), L12. DOI: 10.3847/2041-8213/ac6674. ARXIV: 2311.08680.

[484] John Kormendy and Luis C. Ho. “Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies”. *Ann. Rev. Astron. Astrophys.* 51 (2013), 511–653. DOI: 10.1146/annurev-astro-082708-101811. ARXIV: 1304.7762.

[485] S. D. M. White and M. J. Rees. “Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering”. *Monthly Notices of the Royal Astronomical Society* 183.3 (1978), 341–358. DOI: 10.1093/mnras/183.3.341.

[486] M. C. Begelman, R. D. Blandford, and M. J. Rees. “Massive black hole binaries in active galactic nuclei”. *Nature* 287.5780 (1980), 307–309. DOI: 10.1038/287307a0.

[487] Vicente Rodriguez-Gomez et al. “The merger rate of galaxies in the Illustris Simulation: a comparison with observations and semi-empirical models”. *Mon. Not. Roy. Astron. Soc.* 449.1 (2015), 49–64. DOI: 10.1093/mnras/stv264. ARXIV: 1502.01339.

[488] Mohan Rajagopal and Roger W. Romani. “Ultra–Low-Frequency Gravitational Radiation from Massive Black Hole Binaries”. *Astrophys. J.* 446 (1995), 543. DOI: 10.1086/175813. ARXIV: astro-ph/9412038.

[489] Andrew H. Jaffe and Donald C. Backer. “Gravitational waves probe the coalescence rate of massive black hole binaries”. *Astrophys. J.* 583 (2003), 616–631. DOI: 10.1086/345443. ARXIV: astro-ph/0210148.

[490] J. Stuart B. Wyithe and Abraham Loeb. “Low-Frequency Gravitational Waves from Massive Black Hole Binaries: Predictions for LISA and Pulsar Timing Arrays”. *Astrophys. J.* 590.2 (2003), 691–706. DOI: 10.1086/375187. ARXIV: astro-ph/0211556.

[491] E. S. Phinney. “A Practical theorem on gravitational wave backgrounds” (2001). ARXIV: astro-ph/0108028.

[492] Alberto Sesana, Alberto Vecchio, and Carlo Nicola Colacino. “The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays”. *Mon. Not. Roy. Astron. Soc.* 390 (2008), 192. DOI: 10.1111/j.1365-2966.2008.13682.x. ARXIV: 0804.4476.

[493] B. Kocsis and A. Sesana. “Gas-driven massive black hole binaries: signatures in the nHz gravitational wave background”. *Monthly Notices of the Royal Astronomical Society* 411.3 (2011), 1467–1479. DOI: 10.1111/j.1365-2966.2010.17782.x. ARXIV: 1002.0584.

[494] H. Middleton et al. “Massive black hole binary systems and the NANOGrav 12.5 yr results”. *Mon. Not. Roy. Astron. Soc.* 502.1 (2021), L99–L103. DOI: 10.1093/mnrasl/slab008. ARXIV: 2011.01246.

[495] Kelley, L. Z. et al. *holodeck (unpublished)*. `https://holodeck-gw.readthedocs.io/en/main/index.html`.

[496] Jose J. Blanco-Pillado, Ken D. Olum, and Benjamin Shlaer. “Large parallel cosmic string simulations: New results on loop production”. *Phys. Rev. D* 83 (2011), 083514. DOI: 10.1103/PhysRevD.83.083514. ARXIV: 1101.5173.

[497] R. Abbott, T. D. Abbott, S. Abraham, et al. “Upper limits on the isotropic gravitational-wave background from Advanced LIGO and Advanced Virgo’s third observing run”. *Phys. Rev. D* 104.2 (2021), 022004. DOI: 10.1103/PhysRevD.104.022004. ARXIV: 2101.12130.

[498] Rouzbeh Allahverdi et al. “The First Three Seconds: a Review of Possible Expansion Histories of the Early Universe”. *Open J. Astrophys.* 4 (2021), astro.2006.16182. DOI: 10.21105/astro.2006.16182. ARXIV: 2006.16182.

[499] Anastasios Avgoustidis and E. P. S. Shellard. “Effect of reconnection probability on cosmic (super)string network density”. *Phys. Rev. D* 73 (2006), 041301. DOI: 10.1103/PhysRevD.73.041301. ARXIV: astro-ph/0512582.

[500] Gia Dvali, Juan Sebastián Valbuena-Bermúdez, and Michael Zantedeschi. “Dynamics of confined monopoles and similarities with confined quarks”. *Phys. Rev. D* 107.7 (2023), 076003. DOI: 10.1103/PhysRevD.107.076003. ARXIV: 2210.14947.