

## TASI lectures on Phase Transitions, Baryogenesis, and Gravitational Waves

---

**Djuna Croon<sup>a</sup>**

<sup>a</sup>*Institute for Particle Physics Phenomenology, Department of Physics, Durham University, Durham DH1 3LE, U.K.*

*E-mail:* [djuna.1.croon@durham.ac.uk](mailto:djuna.1.croon@durham.ac.uk)

These lectures, presented at the 2022 TASI summer school, give an introductory overview of first-order phase transitions in the early Universe, baryogenesis, and the resulting gravitational wave phenomenology. We introduce thermal field theory via the imaginary time formalism, and comment on the pitfalls of 1-loop calculations and alternative approaches. Then, we discuss how to calculate the false vacuum decay rate in first order phase transitions, of which we give various examples in theories beyond the Standard Model. Baryogenesis is presented via the Sakharov conditions, and how they are met in important classes of examples. Finally, we explore gravitational waves from the early Universe, first reviewing the basics of gravitational wave generation and then focusing on the specific example of first order phase transitions.

*2022 Theoretical Advanced Study Institute in Particle Theory (TASI2022)*

*6 June - 1 July, 2022*

*University of Colorado, Boulder*

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	These lecture notes	3
1.2	What do (we think) we know about the early Universe?	3
<b>2</b>	<b>Field theory at finite temperature</b>	<b>5</b>
2.1	Imaginary time formalism	5
2.1.1	The partition function	5
2.1.2	Scalar field potential	9
2.1.3	Loop expansion	11
2.2	IR-divergences and non-perturbativity at high temperatures	13
2.2.1	Daisy resummation	14
2.2.2	Dimensional reduction	14
2.2.3	Functional methods	15
<b>3</b>	<b>First order phase transitions</b>	<b>16</b>
3.1	False vacuum decay	17
3.1.1	Reminder: tunneling in quantum mechanics	17
3.1.2	Semi-classical methods in field theory	18
3.1.3	The direct method	24
3.2	Bubble expansion	26
3.3	Examples of first order phase transitions	28
3.3.1	The Electroweak Phase Transition	28
3.3.2	Confinement and chiral symmetry breaking	30
3.3.3	Hidden sectors	31
<b>4</b>	<b>Baryogenesis</b>	<b>32</b>
4.1	The Sakharov conditions	33
4.2	Models of baryogenesis	35
4.2.1	Electroweak baryogenesis	35
4.2.2	Spontaneous baryogenesis	37
4.2.3	Affleck-Dine baryogenesis	39
4.2.4	Leptogenesis	39
<b>5</b>	<b>Gravitational waves from the Early Universe</b>	<b>40</b>
5.1	Gravitational wave basics	41
5.2	Stochastic gravitational wave backgrounds	43
5.3	Gravitational waves from phase transitions	45
<b>6</b>	<b>Conclusions and outlook</b>	<b>47</b>

## 1. Introduction

### 1.1 These lecture notes

These notes are an (only slightly) extended version of four 75-minute lectures I gave at the TASI 2022 summer school. They cover aspects of early Universe physics, focused in particular on thermal phase transitions, their connection to models of baryogenesis, and their potential gravitational wave phenomenology. Clearly, each of these topics is deserving of its own lecture course, over a term or over several terms. As such these notes should only be considered a fleeting introduction to the topics in question. I have added many suggestions for further reading throughout the text.

The topics discussed in these notes are areas of high research activity. In this lecture series, I have prioritised developing an intuitive picture of the dynamics in the early Universe over doing full justice to the latest technical developments. Any reader interested in pursuing research at the cutting edge is recommended to consult additional educational texts and research papers. I would like to highlight two areas in particular in which these notes omit comprehensive detail. Firstly, these notes give a very brief introduction to thermal field theory, focusing mainly on the 4-D one-loop imaginary time formalism. While a helpful starting point, it is known that this formalism of finite temperature field theory is plagued by large theoretical uncertainty [1]. Though the formalism is most well-known and used, these flaws cannot be ignored in phenomenological calculations. Dimensional reduction [2] can be applied to thermal field theory and allows for more accurate calculations in most cases. I have included a short description of the spirit of such calculations, but do not attempt to give a complete description of the ongoing research in this area. Pedagogical texts on this topic (including worked examples as well as code packages) exist in the literature, and I have provided examples in the references.

Moreover, I have not provided a detailed overview on the dynamics of the primordial plasma after bubble nucleation, including the calculation of the bubble wall speed and the full computation of the baryon asymmetry generated across the bubble wall. These notes focus instead on general aspects of baryogenesis and broad classes of models in the literature. The reader is encouraged to consult the many research papers and excellent topical reviews cited in the section. A thorough overview of the current status of the technical challenges of these calculations was provided recently in [3].

I hope these lecture notes serve as a valuable first introduction to the captivating subjects of phase transitions and baryogenesis, as well as the ongoing exciting research endeavors dedicated to exploring them. This work connects to fundamental open problems in physics today, and has phenomenological prospects for the near future. The advent of gravitational wave astronomy has delivered a resurgence of interest in these topics. As we will see, there are still plenty of important research questions to answer before potential signals can be reliably understood in terms of early Universe physics.

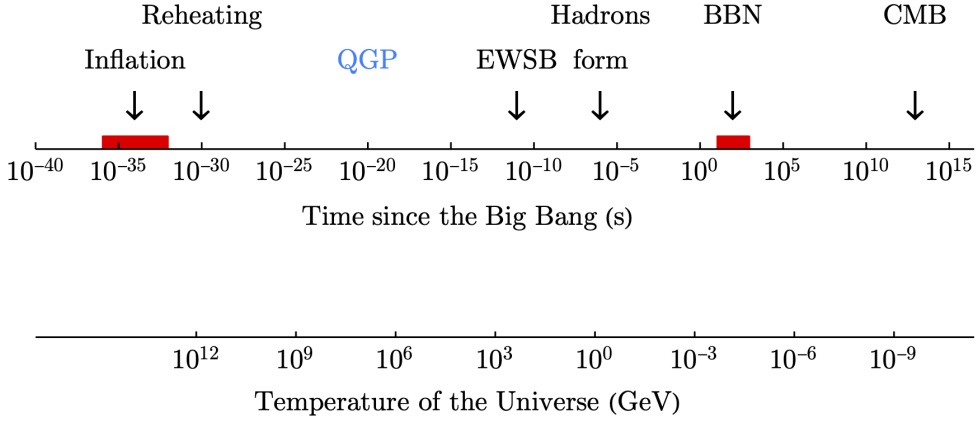
### 1.2 What do (we think) we know about the early Universe?

Before we dive into the specialised topics of baryogenesis, cosmological phase transitions, and their (gravitational wave) phenomenology, let us first pause and consider the things we think we already know about the early Universe. Assuming the Universe was radiation dominated for most

of the period of interest,

$$H = \frac{1}{2t} = \sqrt{\frac{8\pi}{3} \frac{\rho_{\text{rad}}}{M_p^2}} = \sqrt{\frac{8\pi}{3} \frac{\pi^2}{30} g_* \frac{T^4}{M_p^2}} = \sqrt{\frac{8\pi^3}{90} g_*} \frac{T^2}{M_p} \quad (1)$$

such that time and temperature can be straightforwardly related. Then, we might draw an approximate timeline of the early Universe, assuming the SM degrees of freedom (note that this assumes radiation domination, and therefore stops being accurate at matter-radiation equality):



The first event most cosmologists (though not all!) will place on the cosmic timeline is that of cosmic inflation, the exponential expansion of space which provides the initial conditions for structure formation (as well as resolving a number of problems with the Big Bang model it replaced). Unfortunately, neither the theory nor any observations to date pinpoint a scale for cosmic inflation, and in principle it could have taken place at temperatures anywhere between a couple of orders of magnitude below the Planck scale  $M_p = 1.2 \times 10^{19}$  GeV and Big Bang Nucleosynthesis (BBN), occurring at  $O(1)$  MeV, the very end of our early Universe timeline.

After inflation, the particle(s) responsible for the process decay to other particles, including those of the Standard Model (SM). If the resulting energetic plasma has a high enough temperature, quantum chromodynamics (QCD) is not confined and quarks and gluons were free. In the SM, this confinement happens at around 1 GeV in a phase transition associated with the breaking of the (global) chiral symmetry in the quark sector. This is preceded by the breaking of the electroweak gauge symmetry of the SM, which happens when the Higgs boson obtained its vacuum expectation value (VEV) at around 100 GeV. The latter process gave mass to the fundamental particles, the former accounts for most of the mass of the composite particles.

During BBN, light elements are formed from the hadrons: hydrogen and helium, deuterium, helium-3, helium-4, and lithium-7. The primordial plasma is still ionized until recombination, many decades later. Then, the electrons bind up into nuclei, and the plasma finally becomes transparent to photons. These first photons constitute the Cosmic Microwave Background (CMB). Note that the conversion timeline above breaks down before this time, as matter-radiation equality happened before recombination.

Several important events cannot yet be placed on the cosmic timeline. The thermal evolution (including a possible mass mechanism) of the sector containing dark matter is an example. Even the cosmic history of luminous matter cannot be completely determined without answering what biased it over antimatter in the early Universe. Part of the problem is that our earliest experimental information comes from the CMB. We simply cannot look back any further with photons, because before recombination the plasma was opaque to photons. Photons released in any big event in the early Universe are therefore reabsorbed, and with them the precious secrets they hold.

The same is not true for gravitons. Gravitational radiation couples much more weakly than photons, and a “loud” source of such radiation in the early Universe can in principle be detected by experiments today. Some of the missing events in our cosmic origins story have been connected to such sources. Therefore, a detection of primordial gravitational waves would revolutionise our understanding of the Universe.

In these lecture notes, we will delve into the theory of thermal phase transitions, to explore possible answers to these questions. We will also study their phenomenology, with a particular eye to upcoming gravitational wave experiments. Other sources of gravitational waves from the early Universe have been suggested, including inflation and (pre-)heating, scalar field fragmentation, and cosmic defects. All of these topics are growing areas of research, and characterising the phenomenology is a major task ahead of the first generation of space-based interferometers.

As we focus on thermal phase transitions here, we will start with a very brief introduction to the basics of thermal field theory. Then we will study tunnelling and the dynamics of bubble nucleation. We will finish by studying gravitational waves, with some general aspects as well as the spectra of phase transitions in particular.

## 2. Field theory at finite temperature

To start with, we need to learn more about quantum field theory at nonzero temperature. Some of our intuition about zero-temperature field theory is applicable to finite temperature, but there are some important differences to consider. Our most important object of interest is the scalar potential, which gets thermal corrections from all particles which couple to it. Because we are mostly interested in equilibrium quantities, we will primarily use the imaginary time formalism. At the end, we will address different techniques to deal with quantum corrections.

The overview in this section is largely based on the textbook [4] and the lecture notes [5, 6]. Since this is a brief overview only, if you are interested in learning more about finite temperature field theory you are encouraged to consider these works (or others such as [7, 8]), as well as the research articles cited in the text.

### 2.1 Imaginary time formalism

#### 2.1.1 The partition function

The object we will be basing our thermal field theory on is the grand canonical partition function,

$$\mathcal{Z}(T) \equiv \text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}], \quad \text{where} \quad \beta \equiv \frac{1}{T}, \quad (2)$$

where  $\hat{H}$  is the Hamiltonian operator of the theory,  $\mu$  is the chemical potential,  $\hat{N}$  is the particle number, the trace is taken over the full Hilbert space, and where we use units  $k_B = 1$  (we will generally use particle physics units in these lecture notes, unless otherwise indicated). In the following we will ignore the role of the chemical potential, usually a good approximation in early Universe physics. One way in which you can see this is that chemical potentials correspond to conserved particle numbers. In the early Universe, even when baryon number is conserved (after electroweak symmetry breaking at  $T \sim 100$  GeV),  $\mu_B/T \ll 1$ .

The grand canonical partition function ensures that the trace of the density matrix  $\rho = e^{-\beta(\hat{H}-\mu\hat{N})}/\mathcal{Z}$  equals one, as it must. It can be used to determine all the thermodynamic properties of our early Universe system.  $\mathcal{Z}(T)$  can be written in various ways: as a transition amplitude, which can in turn be expressed as a path integral, i.e.

$$\begin{aligned}\mathcal{Z}(T) &= \sum_a \int d\phi_a \langle \phi_a | e^{-\beta\hat{H}} | \phi_a \rangle \\ &= \int_{\phi(0,\mathbf{x})=\phi(\beta,\mathbf{x})} \mathcal{D}\phi \exp(-S_E); \quad \text{with} \\ S_E &= \int_0^\beta d\tau \int_V \mathcal{L}_E\end{aligned}\tag{3}$$

where the subscript  $E$  denotes Euclidean signature. Sometimes powers of  $\hbar$  are kept explicit in this expression, but in these notes we will generally use  $\hbar = 1$ . The finite temperature object (3) can be obtained from the usual partition functions by the following steps: carrying out a Wick rotation  $t \rightarrow i\tau$ , restricting  $\tau$  to the interval  $\{0, \beta\}$ , and requiring periodicity of the fields:  $\phi(0, \mathbf{x}) = \phi(\beta, \mathbf{x})$ .<sup>1</sup> The periodicity is required by the Kubo-Martin-Schwinger (KMS) relation [11, 12]. It follows from the trace operation in (2), which has a cyclic property, such that  $\phi_a(\mathbf{x}) = \phi(0, \mathbf{x}) = \phi(\beta, \mathbf{x})$ , and the fact that two functions of the Hamiltonian operator commute.

A direct consequence of this periodicity is that there is a discrete set of Fourier modes of the theory, called the Matsubara modes [13]. A scalar field can for example be expanded as:

$$\begin{aligned}\phi(x) &= T \sum_{\omega_n} \int_{\mathbf{k}} e^{ik \cdot x} \phi(k) = T \sum_{\omega_n} \int_{\mathbf{k}} e^{i(\omega_k \tau - \mathbf{k} \cdot \mathbf{x})} \phi(k) \\ \omega_n &= 2n\pi T\end{aligned}\tag{4}$$

where  $k = (\omega_n, \mathbf{k})$ . It is easy to show that this satisfies the periodic boundary condition in (3). You may note that there is another choice of Matsubara modes which satisfies the periodic boundary conditions, though in this case they are anti-periodic ( $\psi(0, \mathbf{x}) = -\psi(\beta, \mathbf{x})$ ). This choice corresponds to fermionic fields, for which the Matsubara frequencies are odd:  $\omega_n = (2n+1)\pi T$ . Gauge bosons have even Matsubara frequencies, like scalars.

Let us first proceed with a real, free scalar field, with Lagrangian

$$\mathcal{L}_E = \frac{1}{2} \partial_\mu \phi \partial_\mu \phi + \frac{1}{2} m^2 \phi^2\tag{5}$$

<sup>1</sup>Note that the results in this section can be obtained in the real time formalism [9, 10] as well, but we will focus on the imaginary time formalism here as it gives some neat insights.

The field  $\phi$  has expansion (4). In principle,  $n$  is an integer which can take both positive and negative values, but because

$$\phi(\omega_n, \mathbf{k})^* = \phi(-\omega_n, -\mathbf{k}) \quad (6)$$

which is necessary such that  $\phi(x)$  is real, only the non-negative integers end up being important in this particular case. This also implies of course that the zero mode has to be real.

Then we can write the partition function in terms of the mode expansion,<sup>2</sup>

$$\mathcal{Z}(T) = \int \mathcal{D}\phi \exp \left( -T \sum_{\omega_n} \int_{\mathbf{k}} \frac{1}{2} \left( \mathbf{k}^2 + \omega_n^2 + m^2 \right) |\phi(k)|^2 \right) \quad (7)$$

which (almost) looks like a type of integral we know how to compute: a Gaussian integral.

Remember, Gaussian integrals are integrals of the form

$$\int_{-\infty}^{\infty} e^{-ax^2/2 + Jx} dx = \sqrt{\frac{2\pi}{a}} e^{J^2/2a}$$

or, for multi-dimensional integrals (of which functional integrals are a special case),

$$\int_{-\infty}^{\infty} e^{-\mathbf{x}^\dagger A \mathbf{x}/2 + \mathbf{J}^\dagger \cdot \mathbf{x}} d\mathbf{x} = \sqrt{\frac{(2\pi)^n}{\det A}} e^{\mathbf{J}^\dagger A^{-1} \mathbf{J}/2}$$

see standard textbooks.

Our expression (7) is a Gaussian integral except for the integration over  $\mathbf{k}$  in the exponent. It turns out that it is easier to compute this expression at discrete volume  $V$ , and then take the continuum limit at the end. This momentum is discrete:  $k_i = 2\pi n_i / L_i$  where  $i$  runs over the spatial dimensions, and integrals become products or sums.

We also need to change variables in the integration measure  $\mathcal{D}\phi(x)$  to use our Matsubara mode expansion. This change of variables is captured in a Jacobian. We have,

$$\mathcal{D}\phi(x) = \underbrace{\left| \det \frac{\delta\phi(x)}{\delta\phi(k)} \right|}_{\text{Jacobian}} \overbrace{d\phi(0, k)}^{\text{zero mode}} \underbrace{\prod_{n \geq 1} d\phi(\omega_n, k)}_{\text{non-zero modes}} \quad (8)$$

<sup>2</sup>This stems from the fact that the quadratic form is

$$\begin{aligned} \int_0^\beta d\tau \int_{\mathbf{x}} \phi_1(\tau, \mathbf{x}) \phi_2(\tau, \mathbf{x}) &= T \int_0^\beta d\tau \int_{\mathbf{x}} \sum_{\omega_n} \sum_{\omega_m} \int_{\mathbf{k}} \int_{\mathbf{p}} \phi_1(\omega_m, \mathbf{p}) \phi_2(\omega_n, \mathbf{k}) e^{i(\mathbf{k}+\mathbf{p}) \cdot \mathbf{x}} \\ &= T \sum_{\omega_n} \sum_{\omega_m} \int_{\mathbf{k}} \int_{\mathbf{p}} \phi_1(\omega_m, \mathbf{p}) \phi_2(\omega_n, \mathbf{k}) \delta_{m, -n} \delta(\mathbf{k} - \mathbf{p}) \\ &= T \sum_{\omega_n} \int_{\mathbf{k}} \phi_1(-\omega_n, -\mathbf{k}) \phi_2(\omega_n, \mathbf{k}) \end{aligned}$$

in combination with (6). This also gives the sign of the  $\mathbf{k}^2 + \omega_n^2$  terms.

such that our expression becomes, writing  $\omega^2 \equiv \mathbf{k}^2 + m^2$  (suppressing a possible subscript  $k$  for convenience, but we should remember that  $\omega$  depends on  $\mathbf{k}$ )

$$\begin{aligned}
 \mathcal{Z}(T) &= \int \mathcal{D}\phi \exp \left( -\frac{1}{2} T \sum_{\omega_n} \frac{1}{V} \sum_{\mathbf{k}} \left( \omega^2 + \omega_n^2 \right) |\phi(\omega_n, \mathbf{k})|^2 \right) \\
 &= \int \mathcal{D}\phi \prod_{\mathbf{k}} \exp \left( -\frac{T}{2V} \sum_{n=-\infty}^{\infty} \left( \omega^2 + \omega_n^2 \right) |\phi(\omega_n, \mathbf{k})|^2 \right) \\
 &= \int \mathcal{D}\phi \prod_{\mathbf{k}} \exp \left( -\frac{T\omega^2}{2V} |\phi(0, \mathbf{k})|^2 - \frac{T}{V} \sum_{n=1}^{\infty} \left( \omega^2 + \omega_n^2 \right) |\phi(\omega_n, \mathbf{k})|^2 \right) \\
 &= \prod_{\mathbf{k}} \left| \det \frac{\delta\phi(x)}{\delta\phi(k)} \right| \int d\phi(0, \mathbf{k}) \prod_{n \geq 1} d\phi(\omega_n, \mathbf{k}) \exp \left( -\frac{T\omega^2}{2V} |\phi(0, \mathbf{k})|^2 - \frac{T}{V} \sum_{n=1}^{\infty} \left( \omega^2 + \omega_n^2 \right) |\phi(\omega_n, \mathbf{k})|^2 \right) \\
 &= \prod_{\mathbf{k}} \left| \det \frac{\delta\phi(x)}{\delta\phi(k)} \right| \underbrace{\sqrt{\frac{2\pi V}{T\omega^2}}}_{\text{zero mode}} \times \underbrace{\prod_{n \geq 1} \frac{\pi V}{T(\omega^2 + \omega_n^2)}}_{\text{nonzero modes}}
 \end{aligned} \tag{9}$$

where we carried out the Gaussian integral in the last line.

Now the determinant term can be found by matching the two computations: the computation using Matsubara modes  $\phi(k)$  and the one using the “full theory” with  $\phi(x)$ . I will not do this computation here, but simply quote the result (see e.g. [6] for more details)

$$\left| \det \frac{\delta\phi(x)}{\delta\phi(k)} \right| = \frac{T}{2\pi} \sqrt{\frac{2\pi T}{V}} \prod_{n=1}^{\infty} \frac{T\omega_n^2}{\pi V} \tag{10}$$

such that we have

$$\begin{aligned}
 \mathcal{Z}(T) &= \prod_{\mathbf{k}} \frac{T}{2\pi} \sqrt{\frac{2\pi T}{V}} \sqrt{\frac{2\pi V}{T\omega^2}} \prod_{n \geq 1} \frac{\omega_n^2}{\omega^2 + \omega_n^2} \\
 &= \prod_{\mathbf{k}} \frac{T}{\omega} \prod_{n \geq 1} \frac{\omega_n^2}{\omega^2 + \omega_n^2} \\
 &= \prod_{\mathbf{k}} T \prod_{n=-\infty}^{\infty} (\omega_n^2 + \omega^2)^{-\frac{1}{2}} \prod_{n'=-\infty}^{\infty} (\omega_{n'}^2)^{\frac{1}{2}} \\
 &= \exp \left\{ \sum_{\mathbf{k}} \left[ \log T - \frac{1}{2} \sum_n \log(\omega_n^2 + \omega^2) + \sum_{n'} \log \omega_{n'} \right] \right\}
 \end{aligned} \tag{11}$$

where  $n'$  implies that the zero mode has been omitted from the sum.



Alternatively, we can also write the products like expressions we know:

$$\begin{aligned}
 \prod_{n=1}^{\infty} \frac{\omega_n^2}{\omega_n^2 + \omega^2} &= \frac{1}{\prod_{n=1}^{\infty} \left[ 1 + \frac{(\omega/2\pi T)^2}{n^2} \right]} \\
 &= \frac{\omega/2T}{\sinh(\omega/2T)} \\
 &= \frac{\omega/T}{e^{\omega/2T} - e^{-\omega/2T}} \\
 &= \frac{\omega}{T} \frac{e^{-\omega/2T}}{1 - e^{-\omega/T}} \\
 &= \frac{\omega}{T} \exp \left\{ -\frac{1}{T} \left[ \frac{\omega}{2} + T \ln(1 - e^{-\omega/T}) \right] \right\}
 \end{aligned} \tag{12}$$

using the identity

$$\frac{\sinh \pi x}{\pi x} = \prod_{n=1}^{\infty} \left[ 1 + \frac{x^2}{n^2} \right] = \frac{e^{\pi x} - e^{-\pi x}}{2\pi x}.$$

Now we are finally ready to take the infinite volume limit:

$$\frac{1}{V} \sum_{\mathbf{k}} \rightarrow \int \frac{d^3 k}{(2\pi)^3}$$

to obtain, using (11),

$$\mathcal{Z}(T) = \exp \left\{ V \int \frac{d^3 k}{(2\pi)^3} \left[ \log T - \frac{1}{2} \sum_n \log(\omega_n^2 + \omega^2) + \sum_{n'} \log \omega_{n'} \right] \right\}. \tag{13}$$

Alternatively, using (12),

$$\begin{aligned}
 \mathcal{Z}(T) &= \prod_{\mathbf{k}} \frac{T}{\omega} \prod_{n \geq 1} \frac{\omega_n^2}{\omega^2 + \omega_n^2} \\
 &= \prod_{\mathbf{k}} \exp \left[ -\frac{1}{T} \left( \frac{\omega}{2} + T \ln(1 - e^{-\omega/T}) \right) \right] \\
 &= \exp \left[ -\frac{V}{T} \int \frac{d^3 k}{(2\pi)^3} \left( \frac{\omega}{2} + T \ln(1 - e^{-\omega/T}) \right) \right]
 \end{aligned} \tag{14}$$

where the last expression is again in the infinite volume limit.<sup>3</sup>

### 2.1.2 Scalar field potential

From the partition function one can derive the following quantities

---

<sup>3</sup>Note that the derivative of the argument of the exponential is proportional to the Bose-Einstein distribution function

$$n_B(\omega) = \frac{1}{e^{\omega/T} - 1} \tag{15}$$

since  $d \ln(1 - e^{-x})/dx = (e^x - 1)^{-1}$ . As we will see, this implies that the Bose-Einstein distribution function appears in the average energy and the entropy.

- The free energy

$$F = -T \ln \mathcal{Z}; \quad (16)$$

- The average (internal) energy

$$E = \frac{1}{\mathcal{Z}} \text{Tr}[\hat{H} e^{-\beta \hat{H}}] \quad (17)$$

where  $\hat{H}$  is again the Hamiltonian operator;

- The entropy

$$S = -\frac{\partial F}{\partial T} = \ln \mathcal{Z} + \frac{1}{T \mathcal{Z}} \text{Tr}[\hat{H} e^{-\beta \hat{H}}] = -\frac{F}{T} + \frac{E}{T}. \quad (18)$$

We are especially interested in the free energy density, since we want to know how finite temperature affects our scalar field potential. The free energy density for our scalar field is given by (using (14))

$$\begin{aligned} \lim_{V \rightarrow \infty} \frac{F}{V} &= \int \frac{d^3 k}{(2\pi)^3} \left( \frac{\omega}{2} + T \ln \left( 1 - e^{-\omega/T} \right) \right) \\ &= \int \frac{d^3 k}{(2\pi)^3} \left( \frac{\sqrt{m^2 + \mathbf{k}^2}}{2} + T \ln \left( 1 - e^{-\sqrt{m^2 + \mathbf{k}^2}/T} \right) \right) \\ &\equiv J_0(m) + \tilde{J}_B(m, T) \end{aligned} \quad (19)$$

It is immediately obvious that the first term in this integral is UV-divergent. However, it is also independent of temperature, and therefore not our main concern here. The second term is given by<sup>4</sup>

$$\begin{aligned} \tilde{J}_B(m, T) &= T \int \frac{d^3 k}{(2\pi)^3} \ln \left( 1 - e^{-\sqrt{m^2 + \mathbf{k}^2}/T} \right) \\ &= \frac{T}{2\pi^2} \int d|\mathbf{k}| \mathbf{k}^2 \ln \left( 1 - e^{-\sqrt{m^2 + \mathbf{k}^2}/T} \right) \\ &= \frac{T^4}{2\pi^2} \int dx x^2 \ln \left( 1 - e^{-\sqrt{(m/T)^2 + x^2}} \right) \end{aligned} \quad (20)$$

This term is not by itself illuminating, but we can get some more insight by doing high and low temperature expansions. For example, in the high-temperature case  $m \ll T$ ,<sup>5</sup>

$$\tilde{J}_B(m, T) = \frac{T^4}{2\pi^2} \left( -\frac{\pi^4}{45} + \frac{\pi^2 m^2}{12T^2} - \frac{\pi |m|^3}{6T^3} - \frac{m^4}{32T^4} \left[ \ln \left( \frac{e^{\gamma_E}}{16\pi^2} \frac{m^2}{T^2} \right) - \frac{3}{2} \right] + \mathcal{O} \left( \frac{m^6}{T^6} \right) \right). \quad (21)$$

where  $\gamma_E \approx 0.5772$  is the Euler-Mascheroni constant. Note that this high-temperature expansion is not a proper Taylor expansion, because the logarithm has a branch cut nearby  $m/T = 0$ , so strictly mathematically speaking the expansion around this point has zero radius of convergence (hence also the absolute value around the cubic term; it should be interpreted as  $|m| = \sqrt{m^2}$ ). Nevertheless, it helps us build some physical intuition.<sup>6</sup>

<sup>4</sup>Usually  $J_i$  are defined without the factor  $T^4/2\pi^2$ , such that  $\tilde{J}_i = T^4/2\pi^2 J_i$ .

<sup>5</sup>In practise, this expansion is numerically reasonably accurate to about  $m/T = 2$  at quartic order.

<sup>6</sup>For a derivation of this expansion, see Ref. [6] or [14]. The upshot is that the cubic term comes from the Matsubara zero mode.

We looked at a scalar field, but actually all bosonic fields contribute to the free energy density through a function like (20). For fermions, the derivation is a little different. The difference comes from the fermionic Matsubara frequencies; these are odd in  $n$ :  $\omega_n = (2n+1)\pi T$ . We will not repeat this calculation, but simply quote the result

$$\begin{aligned} \left( \lim_{V \rightarrow \infty} \frac{F}{V} \right)_{\text{fermions}} &= \int \frac{d^3 k}{(2\pi)^3} \left( \frac{\omega}{2} + T \ln \left( 1 + e^{-\omega/T} \right) \right) \\ &= \int \frac{d^3 k}{(2\pi)^3} \left( \frac{\sqrt{m^2 + \mathbf{k}^2}}{2} + T \ln \left( 1 + e^{-\sqrt{m^2 + \mathbf{k}^2}/T} \right) \right) \\ &\equiv J_0(m) + J_F(m, T) \end{aligned} \quad (22)$$

where we have a temperature-independent vacuum energy again, and a temperature-dependent fermionic contribution (note that the derivative of the log gives the Fermi-Dirac distribution  $n_F(\omega) = (e^{\omega/T} + 1)^{-1}$ ). The high-temperature expansion of this result is given by

$$\tilde{J}_F = \frac{T^4}{2\pi^2} \left( -\frac{7\pi^4}{360} + \frac{\pi^2 m^2}{24T^2} - \frac{m^4}{32T^4} \left[ \ln \left( \frac{e^{\gamma_E} m^2}{\pi^2 T^2} \right) - \frac{3}{2} \right] + O \left( \frac{m^6}{T^6} \right) \right). \quad (23)$$

### 2.1.3 Loop expansion

At this point you might wonder how we can use our result (20) in practice. We are interested in deriving contributions to a scalar field potential, but (20) looks field-independent. The crux is that we have been considering a free theory, whereas field-dependent contributions arise in an interacting theory. In an interacting theory, carrying out the functional integral – as we did in (9) – would not have lead to a closed-form expression. However, for small couplings, we can still find an expression for the effective potential in perturbation theory.

The calculation of the 1-loop effective potential using a path integral can be done in the following way, first introduced by [15],

$$V_{\text{eff}}^{1\text{-loop}}(\bar{\phi}) = -\frac{1}{2}i \int \frac{d^4 p}{(2\pi)^4} \ln \left( \frac{i\mathcal{D}^{-1}(\bar{\phi}; p)}{i\mathcal{D}^{-1}(0; p)} \right) \quad (24)$$

where  $\mathcal{D}^{-1}$  is the inverse tree-level propagator obtained from a free field Lagrangian. In particular,  $\mathcal{D}^{-1}(\bar{\phi}; p)$  comes from the shift  $\phi \rightarrow \bar{\phi} + \varphi$  (where  $\bar{\phi}$  is a constant field), which leads to an effective mass for  $\varphi$  in the Lagrangian<sup>7</sup>

$$m_{\text{eff}}^2(\bar{\phi}) \equiv \left. \frac{\partial^2 V(\phi)}{\partial \phi^2} \right|_{\phi=\bar{\phi}} \quad (25)$$

such that

$$V_{\text{eff}}^{1\text{-loop}}(\bar{\phi}) = -\frac{1}{2}i \int \frac{d^4 p}{(2\pi)^4} \ln \left( \frac{p^2 - m^2(\phi)}{p^2 - m^2} \right). \quad (26)$$

You can check that this expression is equivalent to the Coleman-Weinberg potential derived in an interacting scalar theory. This is a result originally derived in zero-temperature field theory, but nothing stops us from applying it to our finite temperature formalism.

<sup>7</sup>Note that in a gauge theory, this requires fixing a gauge.

In fact, if we had actually calculated the one loop diagrams with the Feynman rules for the imaginary time formalism:

$$\begin{aligned}
 \text{Boson propagator : } & \frac{i}{p^2 - m^2}; \quad p^\mu = [2ni\pi\beta^{-1}, \mathbf{p}] \\
 \text{Fermion propagator : } & \frac{i}{\gamma \cdot p - m}; \quad p^\mu = [(2n+1)i\pi\beta^{-1}, \mathbf{p}] \\
 \text{Loop integral : } & \frac{i}{\beta} \sum_{n=-\infty}^{\infty} \int \frac{d^3 p}{(2\pi)^3} \\
 \text{Vertex function : } & -i\beta(2\pi)^3 \delta_{\Sigma} \omega_i \delta^{(3)} \left( \sum_i \mathbf{p}_i \right)
 \end{aligned} \tag{27}$$

that would have yielded the same result as (20) with a shifted mass. For example, if we consider a scalar field theory, with  $V_{\text{tree}} = m^2 \phi^2/2 + \lambda \phi^4/4!$ , then the sum of all 1-loop diagrams gives the 1-loop potential

$$\begin{aligned}
 V_{T \neq 0, 1\text{-loop}}(\phi) &= -T \sum_{n=1}^{\infty} \sum_m \int \frac{d^3 p}{(2\pi)^3} \frac{(-1)^n}{2n} \left[ \frac{\lambda \phi^2/2}{\mathbf{p}^2 + \omega_m^2 + m^2} \right]^n \\
 &= \frac{T}{2} \sum_m \int \frac{d^3 p}{(2\pi)^3} \ln \left[ 1 + \frac{\lambda \phi^2/2}{\mathbf{p}^2 + \omega_m^2 + m^2} \right] \\
 &= \frac{T}{2} \sum_m \int \frac{d^3 p}{(2\pi)^3} \ln \left[ \frac{\mathbf{p}^2 + \omega_m^2 + m_{\text{eff}}^2(\phi)}{\mathbf{p}^2 + \omega_m^2 + m^2} \right] \\
 &= \frac{T}{2} \sum_m \int \frac{d^3 p}{(2\pi)^3} \ln [\mathbf{p}^2 + \omega_m^2 + m_{\text{eff}}^2(\phi)] - \ln [\mathbf{p}^2 + \omega_m^2 + m^2]
 \end{aligned} \tag{28}$$

where  $n$  counts the number of vertices,  $m$  labels the Matsubara index, and  $m_{\text{eff}}$  is defined as in (25). The last term in (28) is field-independent, and the first term can be compared to the field-dependent part of e.g. (13) with  $\omega^2 = \mathbf{k}^2 + m_{\text{eff}}^2(\phi)$ .

All of this means that while the first term in (21) is field-independent, and therefore not of interest to us, the second term becomes field-dependent through the effective mass. For example, if my potential is  $V(\phi) = \mu^2 \phi^2/2 + \lambda \phi^4/4!$ , then  $m_{\text{eff}}^2(\phi) = \mu^2 + \lambda \phi^2/2$ .

Now we understand how to interpret our thermal functions in perturbation theory, I would like to draw your attention to the appearance of the cubic term with the opposite sign in (21). If our potential is indeed  $V(\phi) = \mu^2 \phi^2/2 + \lambda \phi^4/4!$ , you can see that for the right values of the parameters, a temperature-dependent quartic generates an extra minimum at  $\phi \neq 0$ , separated by a potential barrier:

$$V(\phi)_{T \neq 0} \sim \frac{m_{T \neq 0}^2}{2} \phi^2 - \frac{\lambda_3}{3} \phi^3 + \frac{\lambda_{T \neq 0}}{4} \phi^4 + O\left(\frac{\phi^6}{T^2}\right) \tag{29}$$

where  $\lambda_3 \propto T$ . This scalar potential can have two distinct minima. You can obviously work out for what coefficients there are two minima by extremising this function. If we do this, we find that there are three distinct solutions for  $\lambda_3^2 > 4m_{T \neq 0}^2 \lambda_{T \neq 0}$ .

Notice how there is no effective cubic term in this expansion for the fermionic thermal function. This will be important when we discuss first order phase transitions, because the effective cubic

interaction from boson loops can give rise to a potential barrier, whereas no equivalent thing can be said for fermion loops.

## 2.2 IR-divergences and non-perturbativity at high temperatures

Since we rely on perturbation theory to construct our thermal scalar potential, a few remarks about perturbativity at high temperatures should be made. To keep track of the perturbative expansion in a theory with a number of couplings, a common approach is to assume a power counting relation between those couplings. For the Standard Model, one might adopt, [16]

$$g'^2 \sim g_Y^2 \sim \lambda \sim g^2, \quad (30)$$

where  $g$  and  $g'$  are the weak gauge couplings,  $g_Y$  is the Yukawa coupling between the Higgs and the top quark, and  $\lambda$  is the quartic self-coupling in the Higgs potential.

At high temperatures, loop integrals such as in (28) include a Matsubara frequency sum. The evaluation of this sum implies that in practice, the effective loop expansion parameter involves the bosonic distribution function:

$$g^2 \rightarrow g^2 n_B(E, T) = \frac{g^2}{e^{E/T} - 1} \approx \frac{g^2 T}{E} \geq \frac{g^2 T}{m}. \quad (31)$$

Then, at high temperatures  $m \ll g^2 T$ , even a perturbative theory (at zero temperature) can feature non-perturbativity, for example for the bosonic zero mode. This is called Linde's infrared problem [17]. Note that this problem exists for bosonic modes only – the Fermi-Dirac distribution has a different high-temperature expansion.

Let us look at an example. If we evaluate the following diagram (here the dashed lines are a zero mode, whereas the double-dashed lines correspond to non-zero modes), we find [14, 18]<sup>8</sup>

$$\begin{aligned} \text{Diagram} &\propto \lambda^{N+1} \left[ \int_p \frac{T}{(p^2 + m^2)^{N+1}} \right] \left[ \oint_Q' \frac{1}{Q^2} \right]^N \\ &\propto g^{2N+2} \left[ \int_p \frac{T}{(p^2 + m^2)^{N+1}} \right] \left[ \oint_Q' \frac{1}{Q^2} \right]^N \\ &\propto g^2 m T \left( \frac{gT}{m} \right)^{2N}, \end{aligned} \quad (32)$$

where the first integral is over all the internal zero mode propagators and the second integral is over the loops of non-zero modes (indicated by the prime):

$$\int_p \equiv \left( \frac{\bar{\mu}^2 e_E^\gamma}{4\pi} \right)^\epsilon \int \frac{d^d p}{(2\pi)^d}, \quad \text{and} \quad \oint_P' \equiv T \sum_{\omega_n \neq 0} \int_p. \quad (33)$$

It is obvious that this diagram has an IR-divergence: in the limit  $m \rightarrow 0$ , it goes to infinity for any  $N > 0$ . More generally, when  $T \sim m/g$ , perturbation theory breaks down. This regime of temperatures is usually relevant for the study of phase transitions, so it would appear that thermal field theory has limited predictivity in the precise context we are interested in. Luckily, there are a few different ways of dealing with this apparent breakdown of perturbativity. We will discuss a few examples below.

<sup>8</sup>Note there is a typo in the evaluation of this diagram in [1].

### 2.2.1 Daisy resummation

The IR-divergence of the diagram above can be (partially) cured if the screening mass of the zero mode is taken into account. In a method first worked out by Parwani and Arnold and Espinosa in Refs. [19] and [20], one calculates the thermal corrections of the non-zero modes to find the corrected mass of the zero mode. Then, one can use this corrected mass to compute the contributions of the zero mode,

$$m^2 \rightarrow m_T^2 = m^2 + \#g^2T^2 \quad (34)$$

where  $\#$  is a numerical coefficient which depends on the multiplicity of the bosonic field. With the replacement (34), the problematic diagram (32) becomes  $\propto g^3T^4$  independently of  $N$ , so we can conclude that IR-divergences change the nature of the coupling expansion. Addressing them in this way is called daisy resummation (or sometimes ring resummation), because the diagram looks like a flower.

Daisy resummation also impacts the effective effective cubic interaction in high-temperature expansion, which is often critical to the order of the transition. We have

$$J_B(m, T) \ni -\frac{T}{12\pi}|m|^3 \rightarrow -\frac{T}{12\pi}|m^2 + \#g^2T^2|^3 \quad (35)$$

such that if the self-energy is large, the corrected expression behaves less as a cubic in  $\phi$  and therefore the phase transition may cease to be first order.

To take into account the daisy resummation, we need to work out the proper form of (34), and then self-consistently take it into account in our finite temperature field theory. The thermal screening masses (also called Debye masses) are often denoted by  $\Pi$ , and are calculated thermal corrections of the non-zero modes to the zero modes of the theory (in a gauge eigenstate basis of the theory, if a gauge theory is used). For more details about the calculation of the screening masses, see e.g. [1].

### 2.2.2 Dimensional reduction

An alternative strategy is that of dimensional reduction. The idea of this method dates back to the 1980s [2, 21–23]; it was first applied to the EWPT in the 90s [16] (see also [24–26], and a recent update for the EWPT [27]).<sup>9</sup> Dimensional reduction is a powerful tool when dealing with systems where the physical effects of certain dimensions can be considered negligible on specific scales. As we have seen, in finite temperature field theory, the behaviour of a system is predominantly determined by its static properties and the dynamics along the ‘imaginary time’ dimension. The latter dynamics effectively decouples on scales  $r > \beta$ . Then, dimensional reduction leaves a lower-dimensional effective field theory which maintains the important physics of the original high-temperature system but is much simpler to deal with.

For the example of a gauge theory with coupling  $g$ , dimensional reduction takes advantage of a hierarchy of scales

$$g^2T/\pi \ll gT \ll \pi T, \quad (36)$$

<sup>9</sup>For recent reviews of dimensional reduction, see [1, 28] and references therein.

<i>Start: <math>(d + 1)</math>-dimensional theory</i>				
Scale	Validity	Dimension	Lagrangian	Fields
Hard	$\pi T$	$d + 1$	$\mathcal{L}$	$A_\mu, \phi, \psi_i$
		$\downarrow$ <i>Integrate out <math>n \neq 0</math> modes and fermions</i>		
Soft	$gT$	$d$	$\mathcal{L}_{3d}$	$A_r, A_0, \phi$
		$\downarrow$ <i>Integrate out temporal adjoint scalars <math>A_0</math></i>		
Ultrasoft	$g^2 T / \pi$	$d$	$\tilde{\mathcal{L}}_{3d}$	$A_r, \phi$
<i>End: <math>d</math>-dimensional pure gauge theory</i>				

**Table 1:** Dimensional reduction of a  $(d + 1)$ -dimensional theory with scalar field  $\phi$ , gauge boson  $A_\mu$ , and fermion  $\psi$  based on the scale hierarchy at high temperature. In the first step, one integrates out all hard modes; in the second one also integrates out the temporal adjoint scalar  $A_0$ . At the ultrasoft scale only the spatial gauge field  $A_r$  and the scalar undergoing the phase transition  $\phi$  remains. Adjusted from [1].

to distinguish between a hard scale ( $p \sim \pi T$ , where  $p = |p|$  denotes a momentum scale of particles in the heat bath) at which the theory is perturbative, an ultrasoft scale ( $p \sim g^2 T$ ) where the theory is non-perturbative, and a soft scale ( $p \sim gT$ ) in between. Table 1 shows the dimensional reduction of a toy model theory containing a scalar, a fermion, and one gauge field.<sup>10</sup>

The philosophy of dimensional reduction is to treat perturbative modes perturbatively and nonperturbative modes nonperturbatively. Hence, one integrates out certain modes (fermionic modes and non-zero bosonic Matsubara modes, as well as temporal parts of the gauge modes) perturbatively, whereas the degrees of freedom at the ultrasoft scale can be treated with non-perturbative lattice studies, as was first explored in [29–31]. In some cases, however, higher loop perturbative studies of the ultrasoft degrees of freedom also give accurate results [32–35] (a mathematica package has been developed to perform matching at the next-to-leading order [36]).

### 2.2.3 Functional methods

Lastly, one can use functional methods to deal with renormalisation both at zero and at finite temperature. Such methods do not engage in a loop expansion and are therefore in principle more appropriate in regimes where perturbativity breaks down.

An example of such a method is the functional renormalisation group (FRG). The philosophy of this method is to introduce a regulator function  $R_k$  which ensures fluctuations are taken into account as the theory flows from a UV input scale  $k = \Lambda$  to an IR scale  $k = 0$ . To do so, the action is supplemented with a mass-like term

$$\Delta S_k[\phi] = \frac{1}{2} \int_{\mathbf{p}} \phi[-\mathbf{p}] R_k(\mathbf{p}) \phi[\mathbf{p}] \quad (37)$$

which ensures fluctuations with  $|\mathbf{p}| \ll k$  are suppressed, while fluctuations with  $|\mathbf{p}| \gg k$  are suppressed by  $\partial_k R_k$  in the flow. The addition of (37) implies that the partition function develops a dependence on  $k$ . Then, the average effective action<sup>11</sup> can be defined in terms of the scale dependent

<sup>10</sup>The scalar typically sits between the soft and the ultrasoft scale,  $m_\phi \sim g^{3/2} T / \sqrt{\pi}$ , implying it is still just perturbative. We thank Oli Gould for pointing this out.

<sup>11</sup>Then, rather than the Helmholtz free energy  $F_k = -T \ln \mathcal{Z}_k$  as above, the average effective action is relevant to FRG.

partition function  $Z_k[J]$ ,

$$\Gamma_k[\phi] = -\ln Z_k[J] + \int_x J\phi - \Delta S_k[\phi] \quad (38)$$

where  $J$  is an external source. From this, one can derive a flow equation usually referred to as the Wetterich equation [37] by simply taking the partial derivative with respect to  $k$ :

$$\partial_k \Gamma_k[\phi] = \frac{1}{2} \text{Tr} \left[ \frac{\partial_k R_k}{\Gamma_k^{(2)}[\phi] + R_k} \right] \quad (39)$$

where  $\Gamma_k^{(2)}$  is the second order functional derivative of  $\Gamma_k[\phi]$ . Note that this flow equation is formally exact: no expansion was used to derive it. However, in practise one needs to invoke an approximation scheme to solve it.

The flow equation (39) can be used in both the zero temperature and the finite temperature regimes. Due to its non-perturbative nature it has been applied to calculate observables in the QCD sector of the SM, with good correspondence to lattice studies.<sup>12</sup> A method to apply functional renormalisation techniques to the problem of false vacuum decay was first proposed in [39] and is still under development.

### 3. First order phase transitions

Having developed a basic understanding of field theory at finite temperature, and some of the techniques we can apply to calculate corrections to a scalar potential, we now turn to an important application: phase transitions in the early Universe. Different types of phase transitions exist, and can be classified by their order, a measure of how discontinuous the process is. The order of the phase transition is also an important aspect of its phenomenology. A first order phase transition is formally defined as a phase transition in which the free energy is non-analytic (its derivatives are discontinuous). Namely, if  $\partial F/\partial T$  is discontinuous, the average energy, defined in the previous section is also discontinuous:

$$\begin{aligned} E &= \frac{1}{Z} \text{Tr}[\hat{H} e^{-\beta \hat{H}}] \\ &= \frac{T^2}{Z} \frac{\partial}{\partial T} Z \\ &= T^2 \frac{\partial}{\partial T} \log Z \\ &= F - T \frac{\partial F}{\partial T} \end{aligned} \quad (40)$$

where we have assumed zero chemical potential, as we're doing throughout. The difference in average energy (or the difference in average energy density) is called the latent heat, so the statements that latent heat is released and that the free energy is non-analytic are equivalent.

<sup>12</sup>See [38] for a comprehensive recent review of the functional renormalisation group and its applications.



### 3.1 False vacuum decay

In field theory, the discontinuity in the derivative of the free energy can be described through the (potential) energy surface of a scalar order parameter, which has a local minimum in which the field is initially located, and a global minimum to which it tunnels. This description can be used at zero temperature as well as at finite temperature – in the latter case, we use periodicity in Euclidean time in our description of false vacuum decay.

#### 3.1.1 Reminder: tunneling in quantum mechanics

To warm up, let's remind ourselves of tunneling through a barrier in quantum mechanics. The semi-classical (i.e. small  $\hbar$ ) approximation to this problem shares many similarities with the most popular way of calculating the false vacuum decay in quantum field theory.

In classical mechanics, a particle with insufficient energy to surmount a potential barrier would be completely reflected back. In quantum mechanics, there is a non-zero probability that the particle can tunnel through the barrier, appearing on the other side even though its energy is less than the barrier's height. The Wentzel-Kramers-Brillouin (WKB) approximation provides an approximate solution to calculate this tunneling probability. The key idea behind this approximation is to treat the particle's motion classically outside the barrier, where the potential energy is lower, but quantum mechanically inside the barrier, where the potential energy is higher. The transition between classical and quantum regions occurs in the vicinity of the barrier.

To apply the WKB approximation, we start by dividing the problem into three regions: the incoming region, the barrier region, and the transmitted region. Imagine our barrier region is between  $x_0$  and  $x_1$ . In the incoming and transmitted regions, the particle's motion is described by classical trajectories. Inside the barrier region, the particle's motion must be treated quantum mechanically. Our state  $\psi$  satisfies the following equation of motion, the time-independent Schrödinger equation,

$$\frac{d^2\psi}{dx^2} - \frac{2m(V(x) - E)}{\hbar^2}\psi = 0. \quad (41)$$

For slowly varying  $V(x)$ , the WKB-approximation to the wave-function is

$$\psi(x) \propto \frac{1}{\sqrt{k(x)}} \exp\left(\pm i \int_{x_0}^x dx' k(x')\right) \quad \text{where} \quad k(x) = \sqrt{2m(E - V(x))}/\hbar. \quad (42)$$

Here  $\int k(x)dx$  is the classical action. For  $E > V(x)$ , as is the case in the classical regions, this is an oscillating function, but for  $E < V(x)$ , the action is complex and  $\psi$  is exponentially decreasing.<sup>13</sup> The sign inside the exponential depends on the direction of propagation.

In the WKB approximation, the probability of a particle tunneling through the barrier (versus being reflected back) is essentially determined by the total decrease of the amplitude in the tunneling region. For a particle coming in from  $x < x_0$ , the tunneling amplitude to  $x > x_1$  is given by

$$\gamma_{\text{WKB}} \propto e^{-\frac{1}{\hbar} \int_{x_0}^{x_1} \sqrt{2m(V(x)-E)} dx} + \mathcal{O}(\hbar), \quad (43)$$

<sup>13</sup>At the boundary of the classical and tunneling regions, where  $V(x_0) = E$ , the approximation (42) clearly breaks down. Expanding around  $x_0$ , in the vicinity of the the boundary the Schrödinger equation becomes,  $\frac{d^2\psi}{dx^2} - U_1(x - x_0)\psi + \mathcal{O}(x - x_0)^2 = 0$ , the form of which is known as the Airy equation. By matching the wavefunction (42) and its derivative in both regions to the Airy functions at the boundary, one can relate the coefficients.

such that the tunneling probability exponentially decreases with both the height and width of the barrier.

Analogously to tunneling in quantum mechanics, false vacuum decay rate in quantum field theory is associated with an imaginary part of the action. However, path integrals are manifestly real, and therefore recovering the false vacuum decay rate is not a trivial problem. Two main methods have been proposed in the literature:

- The popular semi-classical method accredited to Callan and Coleman [40, 41], which effectively deforms the integration contour;
- The direct method, developed in [42, 43].

We will first focus on the former, but also draw a quick comparison to the latter, which may in principle be used in non-perturbative calculations [39]. As ever, I encourage the reader to explore the rich topic of false vacuum decay in the wider literature, beyond the cursory review I can offer here. On the topic of phenomenological calculations of first order phase transitions in the early Universe in particular, I highly recommend the recent work [44].

### 3.1.2 Semi-classical methods in field theory

The field theory analogue of the WKB approximation in quantum mechanics was developed by Callan and Coleman in the 1970s [40, 41].<sup>14</sup> These two gentlemen proposed that we go to Euclidean space, as we did in the previous section. The analytical continuation from Minkowski spacetime to Euclidean spacetime is useful because in Euclidean spacetime, where the action is real and positive-definite, the path integral is convergent and amenable to approximation techniques like the saddle point approximation. A real-time formalism was worked out in Ref. [47].

Let us first discuss false vacuum decay at zero temperature. To use a saddle point approximation, the next step is to expand

$$\phi = \bar{\phi} + \varphi, \quad (44)$$

where  $\bar{\phi}$  is a solution to the classical EOM. We use  $\hbar = 1$  throughout these lecture notes, but if we were to keep it explicit, the expansion would be  $\phi = \bar{\phi} + \hbar^{1/2}\varphi$ . If  $\bar{\phi}$  minimises the Euclidean action, then it will dominate the path integral. Thus, we can expand the path integral around  $\phi = \bar{\phi}$  as

$$\begin{aligned} \mathcal{Z} &= \int \mathcal{D}\phi e^{-S_E[\phi]} \\ &= \int_C \mathcal{D}\varphi e^{-S_E[\bar{\phi}] - \frac{1}{2}S_E''[\bar{\phi}]\varphi^2 + \dots} \end{aligned} \quad (45)$$

where the linear term vanishes because  $\bar{\phi}$  is a solution to the classical equations of motion (in actuality there is usually more than one solution  $\phi_n$ , in which case we need to sum over  $n$ ). The integration contour  $C$  is the steepest descent contour through the saddle point. The second order expression is now a Gaussian integral, which we can evaluate to obtain

$$\mathcal{Z} \sim \mathcal{Z}_G = e^{-S_E[\bar{\phi}]} (\det S_E''[\bar{\phi}])^{-1/2} \quad (46)$$

<sup>14</sup>For some recent reviews, see [45, 46].

(see the note on Gaussian integrals above). As this expression only depends on  $\bar{\phi}$ , in principle our task is simple: just find the solution(s) to the classical equations of motion and substitute them back into (46) to find the free energy and related quantities. Let us do that and then consider how to relate our findings to the false vacuum decay rate. To find the equations of motion we extremise the Euclidean action. For a scalar field, we might have:

$$\nabla^2 \bar{\phi} - \frac{dV(\bar{\phi})}{d\bar{\phi}} = 0. \quad (47)$$

Because solutions that extremise the energy are often expected to be spherically symmetric, and this is indeed also the case for the tunneling solutions [48], the only relevant coordinate is the radial one:  $r = (\sum_i^d x_i^2)^{1/2}$ . Therefore the classical equations of motion (in  $d$  dimensions) become,

$$\frac{d^2 \bar{\phi}}{dr^2} + \frac{d-1}{r} \frac{d\bar{\phi}}{dr} = \frac{dV(\bar{\phi})}{d\bar{\phi}}. \quad (48)$$

At zero temperature, the relevant number of dimensions is obviously  $d = 4$ : the field configuration describing tunnelling has  $O(4)$  symmetry. At finite temperature, for length scales  $r > \beta$ , the relevant number of dimensions is  $d = 3$ .

The decision of *which* saddle points to consider is an important aspect of the semi-classical approximation for false vacuum decay. In particular, if all saddle points are taken into account, the path integral is manifestly real. As we expect false vacuum decay to be associated with an imaginary contribution, we have to be selective. So, let us discuss the solutions of (48) for a typical situation we would like to consider: a double well potential. First of all, there are two trivial solutions, which are the static field values corresponding to the bottoms of the wells: for these field configurations all terms vanish:  $\partial_r^2 \phi = 0$ ,  $\partial_r \phi = 0$  and  $V' = 0$ . We can call these solutions  $\phi_{\text{FV}}$  and  $\phi_{\text{TV}}$  for the less deep (false vacuum) and deeper vacuum (true vacuum) respectively. Then there are a number of other solutions, which depend on the boundary conditions we impose. We are interested in field configurations that interpolate between  $\phi_{\text{FV}}$  and some field value on the other side of the barrier (note that this need not be  $\phi_{\text{TV}}$ ), so a natural choice is to impose  $\phi_{\text{FV}}$  at  $r = \pm\infty$ . As we are only interested in reaching the other side of the barrier, and not at which field value exactly this happens, we may not want to impose a similar Dirichlet boundary condition at the origin. We can instead impose finite energy at the origin to give us the second boundary condition,  $\partial_r \phi|_{r=0} = 0$ . Thus, it seems reasonable that the boundary conditions of a metastable state of the false vacuum would be,

$$\begin{aligned} \bar{\phi}(r \rightarrow \pm\infty) &= \phi_{\text{FV}}, \\ \left. \frac{d\bar{\phi}}{dr} \right|_{r=0} &= 0. \end{aligned} \quad (49)$$

Indeed, you can find (48) and (49) in many papers describing cosmic phase transitions. They describe the *bounce*  $\bar{\phi}_b$ , an  $O(d)$  field configuration which starts (and ends) in the false vacuum, and reaches the other side of the potential barrier (say  $\bar{\phi} = \phi_e$ ) with zero velocity ( $\partial_\mu \bar{\phi} = 0$ ). But they also still describe the static solution  $\bar{\phi} = \phi_{\text{FV}}$ , so we are stuck with that saddle point too.

Now, how do we find the false vacuum decay rate from  $\mathcal{Z}$ ? A simple way to get intuition for that (courtesy of [6]),<sup>15</sup> is to imagine a metastable (zero-temperature) state and time-evolving it

<sup>15</sup>For a more rigorous argument, see the original paper by Callan and Coleman [41].

forward. We have,

$$\begin{aligned} |\phi(t, \mathbf{x})\rangle &= e^{-iEt} |\phi(0, \mathbf{x})\rangle = e^{-i(\text{Re}(E) + i\text{Im}(E))t} |\phi(0, \mathbf{x})\rangle \\ &\rightarrow \langle \phi(t, \mathbf{x}) | \phi(t, \mathbf{x}) \rangle = e^{2\text{Im}(E)t} \langle \phi(0, \mathbf{x}) | \phi(0, \mathbf{x}) \rangle \end{aligned} \quad (50)$$

such that the decay rate is given by  $\Gamma = -2\text{Im}(E)$  – as expected, metastability is associated with imaginary energy (eigenvalues). Analogously, for a thermal state, we might expect something like  $\Gamma = -2\text{Im}(F)$ . This would mean that

$$\begin{aligned} \Gamma &= -2\text{Im}(F) \\ &= 2T\text{Im} \left[ \ln \left( \mathcal{Z}_0 + \int \mathcal{D}\varphi e^{-S_E[\bar{\phi}_b] - \frac{1}{2}S_E''[\bar{\phi}_b]\varphi^2 + \dots} \right) \right] \\ &\sim 2T\text{Im} [\ln \mathcal{Z}_0] + 2\frac{T}{\mathcal{Z}_0} \text{Im} \left[ \int \mathcal{D}\varphi e^{-S_E[\bar{\phi}_b] - \frac{1}{2}S_E''[\bar{\phi}_b]\varphi^2 + \dots} \right] \\ &\sim \frac{T}{\mathcal{Z}_0} \text{Im} \left[ (\det S_E''[\bar{\phi}])^{-1/2} e^{-S_E[\bar{\phi}_b]} \right] \\ &\equiv A(T) e^{-S_E[\bar{\phi}_b]}. \end{aligned} \quad (51)$$

where we have restricted ourselves to saddle points with the boundary conditions (49),  $\mathcal{Z}_0 \sim \mathcal{Z}_G[\phi_{\text{FV}}]$  is the contribution from the solution  $\bar{\phi} = \phi_{\text{FV}}$  which stays in the false vacuum, which we expect to dominate over  $\mathcal{Z}[\bar{\phi}_b]$  (such that we could expand the log). For the other saddle point, we have evaluated the Gaussian integral, including a factor 1/2 which we will comment on below.<sup>16</sup> We have also grouped together the prefactor  $A(T)$ :

$$A(T) = \frac{T}{\mathcal{Z}_0} \text{Im} \left[ (\det S_E''[\bar{\phi}_b])^{-1/2} \right] \quad (52)$$

At finite temperature, a Euclidean action we may be interested in is the scalar field action<sup>17</sup>

$$S_E = \int_0^\beta d\tau \int d^3x_E [(\partial_{x_E}\phi)^2 + V(\phi, T)] \quad (53)$$

in which the subscript  $E$  denotes Euclidean and the integral is over Euclidean space. The coordinates in Euclidean space are  $\mathbf{x}$  as before, and  $\tau = it$ . As mentioned above, on length scales large compared to  $\beta$ , the system behaves like it is approximately 3-dimensional. In this case,

$$\begin{aligned} S_E[\phi] &= \int_\tau \int_{\mathbf{x}} L_E \\ &= \left( \frac{1}{T} \right) \int_{\mathbf{x}} L_E \\ &= \frac{S_{E,3}}{T} \end{aligned} \quad (54)$$

and thermal fluctuations assist in the phase transition (as mentioned, in this case one should also take  $d = 3$  in (48)). In the following I will often just use  $S_E$  to mean  $S_{E,3}$ , but it will be clear

<sup>16</sup>This seemingly innocent factor of 2 was derived in [49], elaborated on in [41] and more recently in [43].

<sup>17</sup>One should be careful to avoid double-counting, as per the discussion in [1, 39]

from context that we are in a thermal situation. That's also the situation we are most interested in in the early Universe. It is usually the case that the nucleation rate (51) increases exponentially as the temperature decreases, because the ratio  $S_E/T$  decreases with decreasing temperature. If the barrier disappears,  $S_E/T$  tends to zero.

We have yet to determine the prefactor  $A(T)$  in (51). If we were interested in the decay rate per unit volume  $\Gamma/\mathcal{V}$ , which we indeed often are, dimensional analysis tells us that it should be  $A(T)/\mathcal{V} \sim T^4$ , which is indeed an approximation that is sometimes made – the argument is that because it features in the argument of the exponential,  $S_E/T$  is far more important and therefore dimensional analysis is sufficient. However, an understanding the origin of the prefactor lends some insight to false vacuum decay rate calculations in general. Namely, in our exploration of the prefactor we will discover what makes the free-energy we are calculating imaginary, leading to a nonzero false vacuum decay rate.

Let's go back to our field expansion. We may expand our field (44) further into an eigenbasis of an operator:

$$\phi = \bar{\phi} + \sum_i \zeta_i \varphi_i. \quad (55)$$

Given the expansion (45), good choice is the operator  $S''_E[\bar{\phi}] = (-\square + V''[\bar{\phi}])$ , which characterises the size of fluctuations. Note that in this case the result of our Gaussian integral in (46) is equivalent to

$$\mathcal{Z} = e^{-S_E[\bar{\phi}]} \prod_i (\lambda_i)^{-1/2} \quad (56)$$

where  $\lambda_i$  are the eigenvalues of the operator  $S''_E[\bar{\phi}]$ :

$$S''_E[\bar{\phi}] \varphi_i = (-\square + V''[\bar{\phi}]) \varphi_i = \lambda_i \varphi_i. \quad (57)$$

The operator  $S''_E[\bar{\phi}]$  has an interesting set of eigenvalues. In particular, it (in principle) has zero modes ( $\lambda = 0$ ). The zero modes are given by  $\partial_\mu \bar{\phi}$ , which we can easily check

$$S''_E[\bar{\phi}] \partial_\mu \bar{\phi} = \partial_\mu (S'_E[\bar{\phi}]) = 0 \quad (58)$$

where  $S'_E[\bar{\phi}] = 0$  by the definition of  $\bar{\phi}$ . A seemingly trivial point is that if the field  $\phi$  is space-time independent – such as the false vacuum saddle point ( $\bar{\phi} = \phi_{\text{FV}}$ ) – no zero modes exist.

Since the bounce solution has four zero modes, it is clear that this is not the ground state, as that must be unique. So there must be at least one lower energy state: a negative mode  $\lambda_-$ . In an analogy with simple harmonic oscillators (SHO), we can motivate that there is just a single negative mode. The spectrum of a SHO contains a ground state without nodes, and a first excited state with one node. As our zero mode has one node (where the bounce happens, at  $\bar{\phi}(r=0) \equiv \bar{\phi}_e$ , as there  $\partial_\mu \bar{\phi} = 0$ ), it must be the first excited state. The existence of the negative mode also implies that  $F$  is imaginary. This can be seen directly from (56): the single negative mode under the square root renders the entire product imaginary.

We should make a few comments about this negative mode, and its implications for the path integral. First, its existence implies that the Gaussian integral around the bounce solution (45) is

divergent when the integration contour  $C$  is real. This is the first hint that to get something sensible, we must perform an analytic continuation of the integration contour. The negative mode arises from the fact that the bounce solution is not a minimum of the Euclidean action, but a saddle point: among the family of solutions which penetrate the barrier – which includes besides the bounce also configurations which return to the false vacuum before or after reaching the other side of the barrier, and the limiting case  $\phi_{\text{FV}}$  – the bounce is a maximum. This means that if our integration contour includes all saddle points in this family, the bounce gives a subdominant contribution.

So, it turns out that to compute the false vacuum decay rate, the integration contour must be chosen carefully. The appropriate contour is complex, and along the steepest descent contour through the false vacuum saddle point:  $C_F$ . This corresponds to the physical situation we have in mind when calculating the false vacuum decay rate: choosing this contour implies that the resonance associated with the false vacuum dominates, even if  $\phi_{\text{FV}}$  is not the dominant saddle point (it is easy to see that  $\phi_{\text{TV}}$  will give a greater contribution). Integrating along this contour results in a complex path integral, with the imaginary part associated with the decay rate:  $\Gamma \propto \text{Im} \ln \mathcal{Z} \sim \text{Im} \mathcal{Z} / \text{Re} \mathcal{Z}$ . However, the imaginary part comes from a part of the contour far away from the FV saddle point, such that the *Gaussian* path integral evaluated on the false vacuum saddle point is real –  $\text{Im} [\ln \mathcal{Z}_G[\phi_{\text{FV}}]] = 0$ .

As shown in [43], the imaginary part of the path integral over  $C_F$  is equivalent to *half* of the steepest descent contour passing through the bounce saddle point (modulo a sign).<sup>18</sup> This complex contour  $C_b$  leads to a finite path integral despite the negative eigenvalue, with an imaginary part:

$$\text{Im} \mathcal{Z}_G = \text{Im} \int_{C_b} \mathcal{D}\varphi e^{-S_E[\bar{\phi}] + \frac{1}{2} |\lambda_-| \varphi^2} = \frac{1}{2} e^{-S_E[\bar{\phi}]} (\det |\lambda_-|)^{-1/2}. \quad (59)$$

This choice of contour is the justification for restricting our analysis to the bounce and the FV saddle points (note that we still need the latter to compute  $\text{Re} \mathcal{Z} = \mathcal{Z}_0$ ), which we imposed through the boundary conditions (49). A more detailed analysis of the appropriate integration contour in can be found in [43].

With the right choice of integration contour (or equivalently the boundary conditions (49) for the saddle points), (51) is our false vacuum decay rate. We can clarify this expression some more by analysing the zero modes. Let us first separate them out of (55):

$$\phi = \bar{\phi} + N \zeta^\mu \partial_\mu \phi + \sum_i' \zeta_i \varphi_i \quad (60)$$

where the prime indicates that the zero modes are to be left out (and as before we have set  $\hbar = 1$ ). Here  $N$  indicates that the modes are not properly normalised yet, because

$$\begin{aligned} \langle \partial_\mu \phi | \partial_\nu \phi \rangle &= \frac{1}{4} \delta_{\mu\nu} \int d^4x (\partial_\rho \phi)^2 \\ &= \delta_{\mu\nu} S_E[\phi] \end{aligned} \quad (61)$$

<sup>18</sup>Note that this makes (51) rather subtle, as it would be incorrect to state  $2\text{Im} [\ln \mathcal{Z}_0] + 2 \frac{T}{\mathcal{Z}_0} \text{Im} \left[ \int \mathcal{D}\varphi e^{-S_E[\bar{\phi}_b] - \frac{1}{2} S_E''[\bar{\phi}_b] \varphi^2 + \dots} \right] \sim 2\text{Im} [\ln \mathcal{Z}_G[\phi_{\text{FV}}]] + 2 \frac{T}{\mathcal{Z}_0} \text{Im} [\mathcal{Z}_G[\phi_b]]$  – this would imply ignoring an imaginary contribution which is of the same order as the bounce contribution.

where the last line is using that for solutions to the EOM, multiplying by  $\partial_\mu \phi$  a sort of virial theorem holds:  $\frac{1}{2}(\partial_\mu \bar{\phi})^2 = V(\bar{\phi})$ . So, the normalisation of the zero modes should be  $N = \sqrt{2\pi/S_E[\bar{\phi}_b]}$  if we choose modes to be normalised such that  $\langle \phi_i | \phi_j \rangle = 2\pi \delta_{ij}$ . Moreover, the zero modes, being space-time translation of the bounce, are proportional to the fourth root of the space-time volume  $\mathcal{V}^{1/4}$ .

All in all, for four space-time dimensions (and thus four zero modes), we find

$$\begin{aligned} \Gamma &\sim \frac{T}{\mathcal{Z}_0} \text{Im} \left[ (\det S''_E[\bar{\phi}])^{-1/2} e^{-S_E[\bar{\phi}]} \right] \\ &\sim T \left| \frac{\det S''_E[\bar{\phi}_b]}{\det S''_E[\bar{\phi}_{\text{FV}}]} \right|^{-1/2} e^{-S_E[\bar{\phi}]} \\ &\sim T \mathcal{V} \left( \frac{S_E[\bar{\phi}_b]}{2\pi} \right)^2 \left| \frac{\det' S''_E[\bar{\phi}_b]}{\det S''_E[\bar{\phi}_{\text{FV}}]} \right|^{-1/2} e^{-S_E[\bar{\phi}]} \end{aligned} \quad (62)$$

where once again the prime denotes the zero modes are taken out of the determinant, and where we have written  $\mathcal{Z}_0 = (\det S''_E[\bar{\phi}_{\text{FV}}])^{-1/2}$ . At high temperatures, there is effectively one fewer zero mode (i.e. not  $\partial_\tau \phi$ ). Therefore,

$$-2\text{Im}F \sim T \mathcal{V} \left( \frac{S_E[\bar{\phi}_b]}{2\pi T} \right)^{3/2} \left| \frac{\det' S''_E[\bar{\phi}_b]}{\det S''_E[\bar{\phi}_{\text{FV}}]} \right|^{-1/2} e^{-S_E[\bar{\phi}]/T} \quad (63)$$

It also turns out the relation for  $\Gamma$  does not correspond exactly to our naive guess at high temperatures, but needs to be modified to [50]

$$\begin{aligned} \Gamma &= -\frac{\sqrt{|\lambda_-|}}{\pi T} \text{Im}F \\ &= \frac{\sqrt{|\lambda_-|}}{2\pi} \mathcal{V} \left( \frac{S_E[\bar{\phi}_b]}{2\pi T} \right)^{3/2} \left| \frac{\det' S''_E[\bar{\phi}_b]}{\det S''_E[\bar{\phi}_{\text{FV}}]} \right|^{-1/2} e^{-S_E[\bar{\phi}]/T} \end{aligned} \quad (64)$$

where  $\lambda_-$  is the eigenvalue of the negative mode. This expression captures the classical growth of the unstable mode at temperature  $T$ .<sup>19</sup>

From this, we can derive the thermal parameters of the phase transition:

- The nucleation temperature (or onset temperature) of the phase transition needs to be defined by convention. For example, it may be defined as the temperature at which one can expect one bubble per Hubble volume:  $\Gamma \times H^{-1} = H^3$ . Solving for  $S_E/T$ , this implies  $S_E/T \sim -3 + \log A(T)/T^4 - 4 \log T/M_p$ . One may also be interested in the temperature of percolation, after the growth of the nucleated bubbles and closer to the completion of the phase transition. This rate will also depend on the dynamics after nucleation.
- The parameter  $\beta$  parametrises the rate of change of the nucleation rate, and is usually defined as  $\beta = (1/\Gamma) \times d\Gamma/dt$ . This has the dimensions of a rate, and is usually normalised to the Hubble rate,  $\beta/H$ .

<sup>19</sup>The expression (64) assumes the calculation can be done in equilibrium. Non-equilibrium physics is expected to enter the unstable (negative mode), leading to modifications in the nucleation prefactor [49]. See also [51, 52].



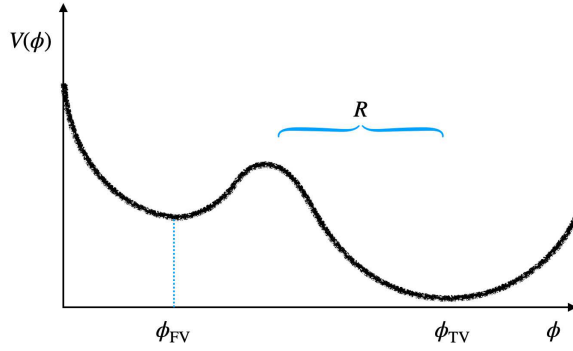
- The parameter  $\alpha$  gives the latent heat at the nucleation temperature, usually normalised to the radiation energy density in the plasma at the nucleation time. Different definitions exist, but it is usually defined in terms of the trace anomaly:  $\alpha = (\Delta V - 1/4 d\Delta V/d \ln T)/\rho_{\text{rad}}$  where  $\Delta V$  is the potential difference between the vacua, evaluated at the nucleation temperature.

We will encounter these parameters further on in these notes, when characterising the phenomenology of the phase transition.

### 3.1.3 The direct method

The method we just studied is the most commonly applied method of calculating false vacuum decay, and it relies on a saddle point approximation and a particular analytic continuation. There is however an alternative method, developed by Andreassen, Farhi, Frost and Schwartz in 2016 [42, 43]. Let's have a brief look at how this works.

First, we need to go back to the problem we wanted to address. We are interested in the decay of a field initially localised in the false vacuum,  $\phi_i = \phi_{\text{FV}}$ , which decays to the other side of the barrier, a (generally multi-dimensional) region which we can call  $R$ :



This decay rate is given by

$$\Gamma = \lim_{\substack{t/t_{\text{slosh}} \rightarrow \infty \\ t/t_{\text{NL}} \rightarrow 0}} \frac{1}{P_{\text{FV}}(t)} \frac{dP_R(t)}{dt} \quad (65)$$

$$P_R(t) = \int_R d\phi_f |\langle \phi_f; t | \phi_{\text{FV}}; 0 \rangle|^2$$

with the latter the probability that the state originally at  $\phi_{\text{FV}}$  is found somewhere in  $R$  at time  $t$ . Here  $t_{\text{slosh}}$  characterises the time-scale for movement within the region of the false vacuum, and  $t_{\text{NL}}$  is the timescale for the transmitted wave-function to start propagating back to the false vacuum. Thus, (65) is applicable for timescales long compared to movements within the neighbourhood of the false vacuum, but short compared to timescales where the wave-function might be transmitted back. This may seem trivial, but is actually a very important insight underpinning false vacuum decay rate calculations. Namely, the more naive limit  $T \rightarrow \infty$  would pick out a saddle point called the shot – which starts and ends at  $\phi_{\text{FV}}$  but spends most of its time on the other side of the barrier, and will therefore be associated with the energy of the true vacuum – unless the path integral contour is deformed, as we did in the previous subsection. If the shot dominates, one finds that the false vacuum decay rate is zero.



After tunneling, when our state emerges on the other side of the barrier, it will do so on a surface in  $\phi$  which we can call  $\Sigma$ . The bounce conserves energy, and therefore it will appear on the surface  $\Sigma$  with  $U[\phi_\Sigma] = U[\phi_{\text{FV}}]$ . Here  $U$  includes not only the potential energy but also the gradient energy:

$$U[\phi] = \int d^3x [(\nabla\phi)^2 + V(\phi)] \quad (66)$$

Then, the two-point function in (65) can be written as a Feynman propagator, which we can split in two parts, before and after it crosses the surface  $\Sigma$ :

$$\begin{aligned} |\langle\phi_f; t|\phi_{\text{FV}}; 0\rangle|^2 &= |D_F(\phi_{\text{FV}}, 0; \phi_f, t)|^2 \\ &= \int_\Sigma db \int_0^t dt' D_F(\phi_{\text{FV}}, 0; b, t') D_F(b, t'; \phi_f, t) \delta(t_b[\phi] - t') \\ &= \int_\Sigma db \int_0^t dt' D_F(\phi_{\text{FV}}, 0; b, t') D_F(b, 0; \phi_f, t - t') \delta(t_b[\phi] - t') \end{aligned} \quad (67)$$

where

$$D_F(b, 0; \phi_f, t) = \mathcal{N} \int_{\phi(0)=b}^{\phi(t)=\phi_f} \mathcal{D}\phi e^{iS[\phi]} \quad (68)$$

where  $\mathcal{N}$  is just a prefactor. Here  $t_b[\phi]$  gives the time at which  $\phi$  first crosses the point  $b$  on the surface  $\Sigma$ , on the other side of the barrier; the delta function ensures that this happens *at least once*.

The two-point function (67) can be simplified by a number of manipulations detailed in [42, 43]. The result is not convergent for real  $t$ , a problem that can be dealt with through the use of an imaginary integration path, or equivalently through analytic continuation to Euclidean space. The analytic continuation of the delta function requires a careful analysis which we will not repeat it here. We simply quote the result:

$$\Gamma = \frac{\mathcal{N}\mathcal{N}^*}{P_{\text{FV}}(t)} \text{Im} \left( \int_\Sigma db \int_{\phi(-\tau)=\phi_F}^{\phi(\tau)=\phi_f} \mathcal{D}\phi e^{-S_E[\phi]} \delta(\tau_b[\phi]) \right), \quad (69)$$

where the integral in the brackets needs to be evaluated for real  $\tau$ , and then analytically continued to  $\tau \rightarrow it$  to yield something purely imaginary.

The final expression for the decay rate  $\gamma_{\text{FV}}$  may be written as

$$\Gamma = \lim_{T \rightarrow \infty} \left| 2 \text{Im} \left( \frac{\int \mathcal{D}\phi e^{-S_E[\phi]} \delta(\tau_b[\phi])}{\int \mathcal{D}\phi e^{-S_E[\phi]}} \right) \right|_{\tau=it} \quad (70)$$

$$= \lim_{T \rightarrow \infty} \left| 2 \text{Im} \left( \frac{1}{\tau} \frac{\int_{\text{hits } b} \mathcal{D}\phi e^{-S_E[\phi]}}{\int \mathcal{D}\phi e^{-S_E[\phi]}} \right) \right|_{\tau=it} \quad (71)$$

where the path integrals have boundary conditions  $\phi(\pm\tau) = \phi_F$ . The meaning of  $\text{Im}()_{\tau=it}$  is to evaluate the content inside the parentheses for real  $\tau$ , analytically continue to imaginary  $\tau = it$ , and then take the imaginary part.

The direct method for calculating false vacuum decay offers an appealing feature – it provides a non-perturbative description of the decay process. Unlike the semi-classical approach, which

relies on a saddle point approximation and careful selection of the integration contour, the direct method should in principle allow for the determination of the false vacuum decay rate without such approximations. However, one still needs to apply boundary conditions to the path integral, as well as the condition that the numerator hits the point  $b$ . A saddle point approximation, as we relied on extensively in the previous section, is the easiest way of doing this. In this case, as discussed in Ref. [43], there are three saddle points which satisfy the boundary conditions  $\phi(\pm\tau) = \phi_F$ : the (constant) false vacuum solution, the bounce, and the shot. The latter is a field profile which rapidly transitions to the true vacuum, stays there for a long time, and then rapidly transitions back.

Only the bounce and shot cross  $b$  and contribute to the numerator of (71). To see which one dominates, we can examine the  $\tau$ -dependence of the two saddle points. The shot rapidly transitions from the false vacuum to the true vacuum and back, resulting in a  $\tau$ -independent term (from the rapid transition) plus a term linear in  $\tau$  proportional to the true vacuum energy. On the other hand, the bounce also has a quick transition to the turning point and back, but spends the most amount of time in the false vacuum, therefore contributing – besides a  $\tau$ -independent term – a term linear in  $\tau$  proportional to the false vacuum energy:

$$S_{\text{shot}} = \tau E_{\text{TV}} + S_S^0 \quad (72)$$

$$S_{\text{bounce}} = \tau E_{\text{FV}} + S_B^0. \quad (73)$$

Importantly, these need to be evaluated for imaginary  $\tau$  – it is real time we are interested in, after all. In this case, the terms linear in  $\tau$  are both purely imaginary and play no role. In contrast, the two pieces  $S_S^0$  and  $S_B^0$  remain purely real when rotated back to imaginary  $\tau$ . Since the shot moves faster than the bounce,  $S_S^0 > S_B^0$ . Therefore, the shot is exponentially suppressed with respect to the bounce, and the latter dominates the path integral.

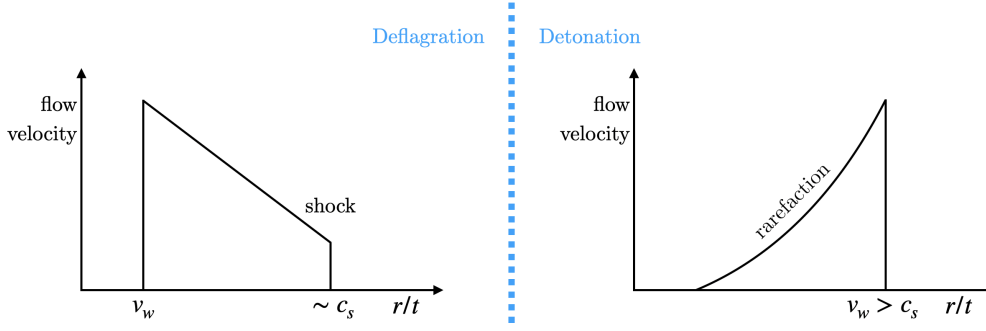
Thus it can be seen that the direct method provides an alternative route towards the calculation of the false vacuum decay rate. It has been shown that applying the saddle point approximation to (71) yields an equivalent result to the semi-classical method.

### 3.2 Bubble expansion

We will briefly comment on what happens to the critical bubbles after they have nucleated. At this point, there are several forces on the bubble wall: the latent heat of the phase transition results in an outward pressure (which we may call the vacuum pressure), bubble wall tension, and friction on the bubble wall, caused by interactions with the particles in the plasma.

If these forces are balanced at some point before the bubble wall has traversed the inter-bubble distance, an asymptotic bubble wall speed is reached. What that bubble wall speed is has phenomenological consequences both for how the latent heat is dissipated and for processes which are concurrent with the phase transition, such as baryogenesis.

Qualitatively, one can distinguish two different scenarios: the detonation scenario, in which the wall speed of the bubbles  $v_w$  (sometimes called the wall velocity, despite the assumption that bubbles expand spherically and only the magnitude is therefore relevant) is larger than the speed of sound  $c_s = 1/\sqrt{3}$  in the relativistic plasma, and deflagrations, where it is smaller. The situations are different in fluid flow profiles, as illustrated in Fig. 1. Note that in the rest frame of the medium, the fluid velocity needs to be zero both at infinity and in the center of the bubble. In the deflagration



**Figure 1:** Fluid profiles for deflagrations and detonations [53].

case, a shock front of cosmological fluid travels in front of the bubble wall. In the detonation case, the fluid flow profile does not form a shockfront, but rather a rarefaction behind the bubble wall. The classic reference [53] contains a more in-depth discussion and plots of the fluid profile.

For gravitational wave generation, it is important to know whether the bubble wall velocity  $v_w \rightarrow 1$ , and moreover, the Lorentz factor  $\gamma_w = (1 - v_w^2)^{-1} \rightarrow \infty$ . Such phase transitions are often referred to as runaway transitions. Implicitly, this means that most of the energy released in the phase transition has gone into accelerating the bubble walls, and has not been leaked to the plasma via friction. In this scenario the contribution to the gravitational wave spectrum coming from the bubble walls dominate over the processes in the plasma. Up until relatively recently, it was believed that runaway transitions could be realised in generic models of phase transitions. However, this insight was based on an incomplete estimation of the transition radiation of gauge fields in the bubble wall. Namely, the calculation of the 1-to-1 pressure for relativistic bubble walls, done in 2009 by Bodeker and Moore, found that

$$P_{\text{thermal}} < P_{\text{vacuum}}, \quad \text{1-to-1 pressure [54]}, \quad (74)$$

(where  $P_{\text{thermal}}$  is the pressure generated by interactions with the plasma) was possible and therefore runaway transitions could occur in generic theories of particle physics. However, it turns out that when a gauge boson gains a mass in the phase transition (as is the case when the phase transition results from the spontaneous breaking of a gauge theory), a bunch of soft gauge bosons may be emitted in the process. A repeated calculation by the same authors in 2017 took into account an extra soft boson or multiple soft bosons in the final state, and found that

$$P_{\text{thermal}} \propto \gamma_w, \quad \text{1-to-2 or 1-to-n pressure [55]}, \quad (75)$$

and therefore a terminal velocity is always reached. Even more recently, other authors have used a different approach towards the 1-to-n resummation of all diagrams. In these works, it was found that the scaling of  $P_{\text{thermal}}$  scales with an even higher power of  $\gamma_w$ :

$$P_{\text{thermal}} \propto \gamma_w^2, \quad \text{1-to-2 or 1-to-n pressure [56]}. \quad (76)$$

There is no current consensus on what the appropriate scaling of  $P_{\text{thermal}}$  with  $\gamma_w$  should be (see also e.g. [57–61]). However, in both cases (75) and (76) runaway phase transitions are not realised for models of spontaneous symmetry breaking of a gauge theory, which includes many important scenarios such as EWSB.

### 3.3 Examples of first order phase transitions

In the SM, there are two phase transitions in the early Universe: the phase transition related to the breaking of the electroweak symmetry, and that related to the confinement of the QCD sector. Alas, with the measured values of masses in couplings, we do not expect a first order phase transition took place in either instance. While indeed the weak gauge bosons and the Higgs self-coupling introduce an effective cubic coupling in the Higgs potential at high temperatures, this term comes with a coefficient that is too small to separate the vacua at finite temperature and that at zero temperature.<sup>20</sup> For the case of QCD, confinement leads to a spontaneous breaking of the global chiral symmetry. The order of this process is somewhat more subtle to study (as we will see below), but through non-perturbative lattice simulations the consensus is that this transition was also not first-order. Nevertheless, there are many theories beyond the SM (BSM) in which a first order phase transition features, some of which are extra interesting because they also address a big open question in particle cosmology. In this section we will study some examples.

#### 3.3.1 The Electroweak Phase Transition

Electroweak symmetry breaking (EWSB) is an example of a broad class of phase transitions, related to the spontaneous breaking of a gauge symmetry. As EWSB happened in the context of the SM, and because it can be related to baryogenesis (as we will study in the next section), it is natural that the electroweak phase transition (EWPT) is a well-studied topic.

We should begin our study of the EWPT by computing the thermal corrections to the potential of the Higgs boson  $\phi$ :<sup>21</sup> Using the 1-loop formalism and daisy resummation we introduced in the section above, we have:

$$V_{\text{eff}}(\phi, T, \bar{\mu}) = V_{\text{tree}} + V_{1\text{-loop}} . \quad (77)$$

where  $\bar{\mu}$  is the renormalisation scale. Because the Higgs is an SU(2) doublet, the contributions to the 1-loop potential in the direction that develops the VEV include both the physical Higgs field and the (would-be) goldstone modes. Of course, there are also corrections from gauge bosons and fermions which couple to the Higgs. The one-loop correction can be further divided into a zero-temperature Coleman-Weinberg piece and a thermal piece

$$V_{1\text{-loop}} = V_{\text{CW}} + V_{\text{T}} + V_{\text{daisy}} . \quad (78)$$

We are mostly interested in the thermal piece here, which we can evaluate using the thermal

<sup>20</sup>In fact, technically the transition is not even second-order: there is no discontinuity in the thermodynamic observables. This type of smooth change in the ground-state is called a crossover.

<sup>21</sup>For a detailed overview of this computation, see [1].

functions we derived above:

$$\begin{aligned}
 V_T &= \frac{T^4}{2\pi^2} \sum_{i \ni \phi, \chi, W, Z} c_i J_B(m_i) + \frac{T^4}{2\pi^2} \sum_{j \ni t} c_j J_F(m_j) \\
 &= \frac{T^4}{2\pi^2} (J_{1\text{-loop}}(m_\phi) - 4N_c J_{1\text{-loop}}(m_t)) \\
 &\quad + \frac{T^4}{2\pi^2} \underbrace{(2J_{1\text{-loop}}(m_2^+) + 2J_{1\text{-loop}}(m_2^-) + J_{1\text{-loop}}(m_1^+) + J_{1\text{-loop}}(m_1^-))}_{\text{goldstone modes}} \\
 &\quad + \frac{T^4}{2\pi^2} \underbrace{((D-1)(2J_{1\text{-loop}}(m_W) + J_{1\text{-loop}}(m_Z) + J_{1\text{-loop}}(m_\gamma)))}_{\text{gauge bosons}}.
 \end{aligned}$$

with the field-dependent mass eigenvalues, given by

$$\begin{aligned}
 \text{Higgs boson} \quad m_\phi^2 &= m_h^2 + 3\lambda\phi^2, \\
 \text{goldstone modes} \quad \begin{cases} (m_1^\pm)^2 &= \frac{1}{2} \left( m_\chi^2 \pm \sqrt{m_\chi^2(m_\chi^2 - \xi_1 g'^2 \phi^2 - \xi_2 \lambda_{\phi h}^2 \phi^2)} \right), \\ (m_2^\pm)^2 &= (m_3^\pm)^2 = \frac{1}{2} \left( m_\chi^2 \pm \sqrt{m_\chi^2(m_\chi^2 - \xi_2 g^2 \phi^2)} \right), \end{cases} \\
 \text{gauge bosons} \quad \begin{cases} m_W^2 &= \frac{1}{4} g^2 \phi^2, \\ m_Z^2 &= \frac{1}{4} (g'^2 + g^2) \phi^2, \\ m_\gamma^2 &= 0, \end{cases} \\
 \text{top quark} \quad m_t^2 &= \frac{1}{2} g_Y^2 \phi^2,
 \end{aligned} \tag{79}$$

where  $m_\chi^2 = m_h^2 + \lambda\phi^2$  is the Goldstone mode mass eigenvalue in the Landau gauge. Here  $g$  is the SU(2) gauge coupling,  $g'$  the U(1) gauge coupling, and  $g_Y$  the top Yukawa coupling. As the latter dominates over the other Yukawa couplings, it is usually sufficient to consider the top quark only.

At high temperature, the bosonic contributions to the thermal potential can introduce a barrier between the symmetric and the broken phase, as can be seen by expanding the thermal functions at high temperatures as we saw above:

$$J_B\left(\frac{m}{T}\right) \sim \frac{1}{24} \frac{m^2}{T^2} - \frac{1}{12\pi^2} \frac{|m|^3}{T^3}. \tag{80}$$

It turns out that in the SM, the resulting cubic is too small to lead to a first order phase transition (which would require something like positive Lagrangian parameter  $m^2$  and  $\lambda_3^2 > 4m^2\lambda$ ). If the Higgs mass was smaller, it might have been. A lattice study based on the dimensionally reduced method introduced earlier on in these notes demonstrated that this would be the case for  $m_h \lesssim 72$  GeV [30, 62–64].

While the EWPT is not a first order transition in the SM, it can be in models with an extended scalar sector. This includes theories with extra scalar singlets, doublets, or triplets coupling to the Higgs boson. We can demonstrate that the order of the phase transition can be changed through a simple example in which the theory has a heavier scalar singlet, coupling to the Higgs boson like

$$\mathcal{L} \ni \tilde{m}\phi^2 s + \lambda_{\phi s}\phi^2 s^2. \tag{81}$$

At scales (much) below the mass of the singlet, this leads (besides corrections to the mass and the quartic) to an effective dimension-6 operator in the Higgs potential, which is now of the form,<sup>22</sup>

$$V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4 + \frac{\phi^6}{\Lambda^2}. \quad (82)$$

With this potential, it is even possible to have a tree-level barrier, as we can accomodate a positive mass squared parameter  $m^2$  and a negative self coupling parameter  $\lambda$ . Note that this does not mean the effective self coupling is negative in the Higgs vacuum, just like the negative Higgs mass parameter does not mean the physical Higgs mass is negative in the SM. Fixing the physical Higgs vev and Higgs mass in the zero temperature vacuum, the phase transition now depends on the parameter  $\Lambda$ . In this effective theory, previously studied in [1, 69–73], this parameter controls the strength of the phase transition, including its potential to lead to observable phenomenology.

### 3.3.2 Confinement and chiral symmetry breaking

The EWPT is an example of a phase transition related to the breaking of a gauge symmetry. But this is not the only type of phase transition that we know to have occurred in the early Universe: the confinement of QCD is another kind. Above the confinement scale, the SM has a global  $SU_L(N_f) \times SU_R(N_f)$  symmetry, where  $N_f$  is the number of dynamical flavours in the primordial plasma. The confinement of free quarks into hadrons lead to the breaking of this preceding chiral symmetry to its diagonal subgroup:

$$SU_L(N_f) \times SU_R(N_f) \rightarrow SU_V(N_f).$$

For SM QCD, at the temperature of confinement  $T_{\text{QCD}} \sim 100\text{MeV} - \text{GeV}$  (the exact temperature is a target of lattice simulations; the most recent result is 157 MeV [74]), the number of dynamical quarks in the plasma is  $N_f = 2$  or  $N_f = 3$ , depending on how you count the strange quark with mass  $m_s \sim 95\text{ MeV}$  (sometimes this is denoted as  $N_f = 2 + 1$ ).

The order of the chiral symmetry breaking phase transition has been the topic of much debate. Low- and high energy effective theories become unreliable in the precise regime of the phase transition. An old and much cited argument by Pisarski and Wilczek (PW) [75] is based on the linear  $\Sigma$ -model – a low energy theory of the quark bilinears in QCD which form a condensate  $\Sigma_{ij} \sim \langle \bar{\psi}_{Rj} \psi_{Li} \rangle$ . The field  $\Sigma_{ij}$  can be decomposed into the pions and mesons (the light composite states) of QCD:

$$\Sigma_{ij} = \frac{\varphi + i\eta'}{\sqrt{2N_F}} \delta_{ij} + X^a T_{ij}^a + i\pi^a T_{ij}^a. \quad (83)$$

In this model, one studies a  $\Sigma$  field Lagrangian invariant under  $SU_L(N_f) \times SU_R(N_f)$ :

$$V(\Sigma) = m_\Sigma^2 \text{Tr} [\Sigma \Sigma^\dagger] - \mu_\Sigma (\det \Sigma + h.c.) + \frac{\lambda}{2} \left( \text{Tr} [\Sigma \Sigma^\dagger] \right)^2 + \frac{\kappa}{2} \text{Tr} [\Sigma \Sigma^\dagger \Sigma \Sigma^\dagger] \quad (84)$$

where the  $\mu_\Sigma$  term is generated by instanton effects, and gives the  $\eta'$  a mass.<sup>23</sup> Then, chiral symmetry breaking corresponds to the real diagonal direction obtaining a vacuum expectation value,  $\langle \varphi \rangle \sim f_\pi$ .

<sup>22</sup>A complete basis of operators includes besides this potential operator also derivative couplings, see e.g. [65–68].

<sup>23</sup>See [76] for some further discussion.

PW derived the  $\beta$  functions for the couplings  $\lambda$  and  $\kappa$  in this theory, in  $4 - \epsilon$  dimensions. Analysing these  $\beta$  functions, they found that no infrared stable fixed point exists for  $N_f \geq 3$ . In RGE analyses, IR fixed points are associated with continuous (second order) phase transitions [77]. Close to the fixed point, all physical quantities can be expressed as a power law in  $(T - T_c)/T_c$ . For a first order (non-continuous) phase transition, no such scaling behaviour is observed; the thermodynamic variables instead exhibit discontinuities. The conclusion was therefore that for  $N_f \geq 3$  the phase transition must be first order, realised by fluctuations.

Needless to say, and also recognised by PW at the time, the argument is not water-tight. The linear  $\Sigma$  model is a low-energy effective theory whose validity should break down right around the confinement scale, due to the appearance of free quarks in the spectrum. More importantly, the  $4 - \epsilon$  expansion might not be appropriate for an effectively 3-dimensional theory at finite temperature. And then the anomaly term  $\mu_\Sigma$  needs to be taken into account away from the limit  $N_c \rightarrow \infty$ . Lattice studies have sought to shed light on this issue, with some studies in the 90s agreeing with PW [78], but some more recent studies not finding evidence for a first order phase transition up until  $N_f = 6$  [79]. For SM QCD, the consensus is that the transition is a crossover (at zero chemical potential, see e.g. [80–82]). However, it is still unclear whether a modified version in which the chiral symmetry is larger at the time of breaking does lead to a first order transition, a possibility that has been studied by model builders in the literature (e.g. [76, 83–85]).

Lastly, let us comment on the possibility that  $N_f = 0$ , such that the phase transition is pure-Yang Mills. This phase transition obviously has little to do with chiral symmetry breaking, and is simpler to study on the lattice. In this case, the evidence points to a first order phase transition [82, 86–89].

### 3.3.3 Hidden sectors

A possibility which has become a popular topic of study in the previous few years, is that a phase transition in a hidden sector took place in the early Universe. For example, a dark Higgs mechanism is a plausible way of generating mass in a hidden sector. Unlike the Higgs sector of the Standard Model, which has been extensively studied and constrained by collider experiments, we have very limited knowledge about the masses and couplings in such a hidden sector. As a result, there is a higher degree of freedom for speculation and exploration in understanding its properties and implications.

Hidden sector phase transitions have garnered significant attention due to the possibility of producing detectable gravitational wave signals. It is an attractive idea that even with a secluded hidden sector – with no (appreciable) non-gravitational interactions with the SM – there is still a discovery potential via gravitational waves, although the information that can be gleaned from such signals is typically relatively limited, as we will see in the last section. With our results from the previous sections, we can intuit some of the general rules of thumb for the strength of the phase transition, which relates to its observability:

- A significant latent heat needs to be released for the phase transition to lead to an observable gravitational wave signal, which implies that  $\Delta V$  should be of order  $T^4$  at the nucleation temperature;
- For a thermally induced barrier, the more bosons in the theory (which gain appreciable mass in the phase transition), the greater the effective cubic term in the potential and the stronger



the phase transition;

- The shallower the order parameter potential (i.e. the larger the ratio  $v/\Lambda$ , where  $V \propto \Lambda^4$ ), the stronger the transition [90].

Indeed, a gravitational wave signal from a first order phase transition could only be observable if a large fraction of the energy in the primordial plasma participated in the phase transition. Though this does not directly point to dark matter, many proposals nevertheless connect the hidden sector phase transition to an appreciable relic abundance.

#### 4. Baryogenesis

It is almost self-evident that our Universe must contain more matter than anti-matter, for matter and anti-matter can mutually annihilate. The matter-antimatter asymmetry of the Universe (or the BAU, the baryon asymmetry of the Universe) can be measured in two different ways. The first is using the light element abundances formed during BBN, as these processes are sensitive to the baryon-to-photon ratio – in particular, more  $^4\text{He}$  is formed for higher ratios (see, e.g. [91]). The second is using the positions of the peaks in the CMB (most recently by the Planck collaboration [92]), as the baryon-to-photon ratio reduces the sound speed, moving the peaks to smaller angular scales. The two measurements agree.<sup>24</sup>

As the baryon number density redshifts, it is convention to normalise the BAU to something else that redshifts at the same rate, to obtain a constant number. There are two different ways of doing this. The first is normalised to the photon number density:

$$\eta = \frac{n_B - n_{\bar{B}}}{n_\gamma} \sim 6.1 \times 10^{-10} \quad (85)$$

(sometimes  $n_B$  is used to denote the net baryon number) and the second is normalised to the entropy density:

$$\eta = \frac{n_B - n_{\bar{B}}}{s} \sim 8.6 \times 10^{-11}. \quad (86)$$

The BAU could have formed across a vast range of scales in the early Universe, only bracketed by BBN – as obvious from the above, to be consistent with the observed light element abundances today – and the end of inflation, such that the BAU is not diluted.<sup>25</sup> However, as we will see, in many models the BAU must have formed before EWSB (at  $T \sim 100$  GeV), because the weak sphaleron processes form an essential component.

There is no general theory to explain production of the BAU, or *baryogenesis*; a vast number of proposals exist. In this section we will study the general aspects that such theories must necessarily address, and then we move on to study a few notable examples. Some recommended reviews of baryogenesis, including lectures given at previous instalments of TASI, are [93–95].

<sup>24</sup>Note, however, the lithium problem: a tension between the observed abundance of  $^7\text{Li}$  in the universe and CMB data.

<sup>25</sup>Additionally, there are constraints on isocurvature perturbations in the CMB from rolling fields during inflation.



#### 4.1 The Sakharov conditions

Andrei Sakharov, in 1967, formulated three conditions which theories that explain the BAU must (generally) satisfy. The Sakharov conditions continue to be widely employed nowadays, and they offer a practical way of framing our discussion. Models of baryogenesis must:

1. Violate baryon number conservation. This one is probably very obvious: if a theory conserves baryon number, it cannot create more baryons than anti-baryons.
2. Violate C- and CP symmetry. The former is such that the reactions producing baryons are not balanced by the reactions that produce anti-baryons. The latter is such that the theory does not produce the same number of left-handed baryons and right-handed anti-baryons (and vice versa).
3. Out-of-equilibrium reactions. If the theory establishes equilibrium, thermal processes are balanced by their reverse processes, including the processes producing baryons. There are exceptions to this condition however, an important one being spontaneous baryogenesis which we will discuss below.

It is the last condition which gives the special role a first order phase transition can play in a theory of baryogenesis. In a first order phase transition, the baryon asymmetry can be created around the expanding bubble walls. There are other ways, however, the simplest being out-of-equilibrium decays.

As certain Sakharov conditions are fulfilled within the framework of the SM, it is not immediately obvious that the BAU is a mystery. Therefore, it is instructive to investigate the extent to which these conditions are met within the SM:

1. The Standard Model features baryon number violation via the Bell-Jackiw chiral anomalies. Classically, the electroweak Lagrangian conserves baryon number, i.e.  $\partial^\mu j_\mu^B = 0$ , as quarks always appear in bilinears:

$$j_B^\mu = \frac{1}{3} \sum_i \bar{Q}_L^i \gamma^\mu Q_L^i + \bar{u}_R^i \gamma^\mu u_R^i + \bar{d}_R^i \gamma^\mu d_R^i \quad (87)$$

However, quantum mechanically, this changes to

$$\partial^\mu j_\mu^B = \frac{g^2 c}{16\pi^2} F^{\mu\nu a} \tilde{F}_{\mu\nu}^a \quad (88)$$

where  $c \propto \hbar$  is a constant ( $c \rightarrow 0$  in the classical limit) and where  $F_{\mu\nu}^a$  (with  $\tilde{F}_{\mu\nu}^a = \epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma a}$ ) is the electroweak gauge field strength.

It turns out that *sphalerons*, solutions to the electroweak field equations, mediate non-perturbative processes which can change anti-leptons into baryons and vice-versa. An example of a sphaleron process would be:

$$\bar{\nu}_e \bar{\nu}_\mu \bar{\nu}_\tau \rightarrow u_L d_L d_L c_L b_L d_L t_L b_L b_L.$$

This process has lepton number  $\Delta L = 3$ , baryon number  $\Delta B = 3$ ,  $\Delta(B - L) = 0$  – sphalerons violate  $B + L$  and conserve  $B - L$ . Sphalerons only couple to left-handed quarks as they are part of the weak sector of the SM.

In the broken phase, sphaleron transitions have a rate

$$\Gamma_{\text{sph}} \propto e^{-E_{\text{sph}}(T)/T} \quad (89)$$

where  $E_{\text{sph}}(T)$  is the sphaleron energy, which depends on the masses of the weak gauge bosons (and therefore on the vev of the Higgs field  $\phi$ ). The transitions are in equilibrium for  $\Gamma_{\text{sph}} > H$ , which happens before EWSB. In equilibrium, sphalerons realise

$$3\mu_B + \sum_i^3 \mu_{iL} = 0 \quad (90)$$

where the sum is over all lepton generations. Thus, they communicate a lepton asymmetry to a baryon asymmetry and vice versa. After EWSB, the sphaleron energy becomes large, and the transitions are exponentially shut off.

2. The weak sector of the SM violates C-symmetry (in particular in the absence of right-handed neutrinos). The SM also features CP violation through the phases in the CKM and PMNS matrix. Because in general these phases are not invariant under redefinitions of the quark fields, a popular invariant to use is the Jarlskog invariant [96]:

$$J = c_{12}c_{13}^2c_{23}s_{12}s_{13}s_{23}\sin\delta \sim \mathcal{O}(10^{-5}) \quad \text{CKM matrix}$$

where the  $c$ 's and  $s$ 's are cosines and sines of angles in the CKM matrix, and  $\delta$  is the complex phase. In the PMNS matrix there is currently only an upper limit on  $J$ , though the phase might be measured in future experiments such as DUNE.

We can estimate the amount of CP violation realised by the complex phase in the CKM matrix as  $J \times (m_c - m_u)(m_t - m_c)(m_t - m_u)(m_s - m_d)(m_b - m_s)(m_b - m_d)/T^{12} \sim 10^{-20} \ll \eta$ . Thus, the amount of CP violation in the CKM matrix is not large enough to realise baryogenesis within the SM.

3. In the previous section, we have already established that we do not expect a first order phase transition in the context of the SM, but we might investigate whether out-of-equilibrium decays are expected to happen. In the early Universe, particles of mass  $m$  and (dominant) coupling  $g$  feature out-of-equilibrium decays if

$$\Gamma \sim \frac{g^2 m}{4\pi\gamma} < H \quad (91)$$

where  $\gamma = \langle E \rangle / m \sim 1$  for  $m > T$ . Assuming radiation domination, this implies an upper limit for the involved coupling of

$$g < \sqrt{\frac{4\pi m}{\sqrt{g_*} M_p}} \quad (92)$$

if we conservatively assume the decay happens when  $m \sim T$ . This means that for particles with masses around the weak scale, the couplings would have to be extremely small to lead to out-of-equilibrium decays.

Thus, it is not possible to explain the BAU within the SM, and its existence can be taken a piece of evidence that new physics is needed. BSM theories can realise baryogenesis in a variety of different ways, often taking advantage of one (or more) of the conditions already present in the SM. In the following we will discuss a few important examples.

## 4.2 Models of baryogenesis

After examining the general conditions of baryogenesis, we will now delve into a few specific examples to illustrate the concept further. It is important to note that these examples are not intended to provide an exhaustive analysis. Nonetheless, they will help develop a better understanding of how the Sakharov conditions can be met in the early Universe and how new physics can be employed to explain the Baryon Asymmetry of the Universe (BAU).

### 4.2.1 Electroweak baryogenesis

In theories of electroweak baryogenesis (EWBG),<sup>26</sup> the Sakharov conditions for the generation of the matter-antimatter asymmetry are satisfied at the electroweak (EW) scale. The EWPT is a first order process, in which bubbles with  $\langle h \rangle = v_h$  are nucleated in a background with  $\langle h \rangle = 0$ . As the EWPT is not first order in the SM, models of EWBG typically include an extended scalar sector (including an extra singlet, for example, or a second Higgs doublet).

Besides the first order phase transition, the CP violation of the SM also needs to be enhanced. Usually, the CP-violating source term is given by the complex spatially varying mass term of a fermionic field:

$$\mathcal{L} \ni -m(z)\bar{\psi}_L\psi_R - m^*(z)\bar{\psi}_R\psi_L \quad \text{with} \quad m(z) = |m(z)|e^{i\theta(z)} = m_R(z) + im_I(z) \quad (93)$$

where  $z$  is the coordinate perpendicular to the bubble wall. In this case, when the fermion is decomposed into a left-handed and a right-handed chiral component, the Dirac equation will split into two coupled EOM with opposite signs in front of the  $\partial_z$  propagation term, meaning that the chiral components propagate differently in the  $z$  direction. Note that electric dipole moment (EDM) constraints, in particular of the electron, severely constrain what CP violating sources are still allowed.

The SM does already contain processes which can violate baryon number: the EW sphaleron transitions. EWBG takes advantage of these transitions, which only occur for  $T \gtrsim v_h$ , so only outside of the expanding bubbles.

The moving bubble walls separate particles in the plasma, see Fig. 2. After scattering with the bubble wall, the particles diffuse in the plasma. The typical diffusion time can be estimated by equating the diffusion length  $d_{\text{diff}} = \sqrt{\bar{D}t}$  (where  $\bar{D}$  is an effective diffusion constant, from Fick's law), with the distance the wall has moved in a particular time interval  $d_{\text{wall}} = v_w t$ :

$$\sqrt{\bar{D}t} = v_w t \quad \rightarrow \quad t_{\text{diff}} = \frac{\bar{D}}{v_w^2}. \quad (94)$$

<sup>26</sup>For reviews, see e.g. [97].

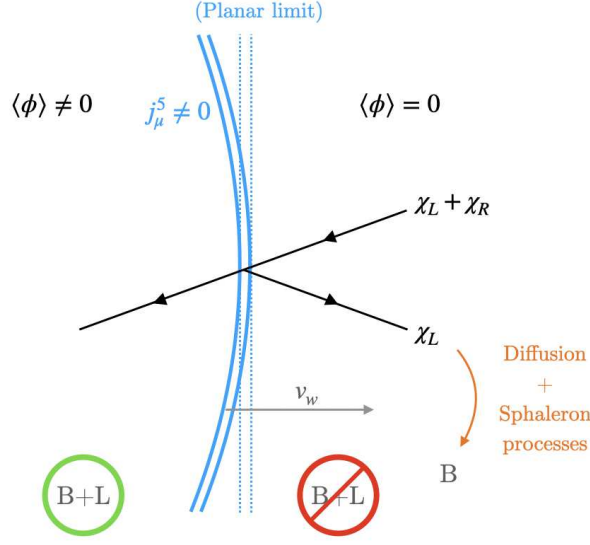


Figure 2

The effective diffusion time  $t_{\text{diff}}$  determines which processes need to be taken into account when considering the CP violating number densities in front of the bubble wall. For example, besides the CP-violating source, there may be interactions which dilute the number density.

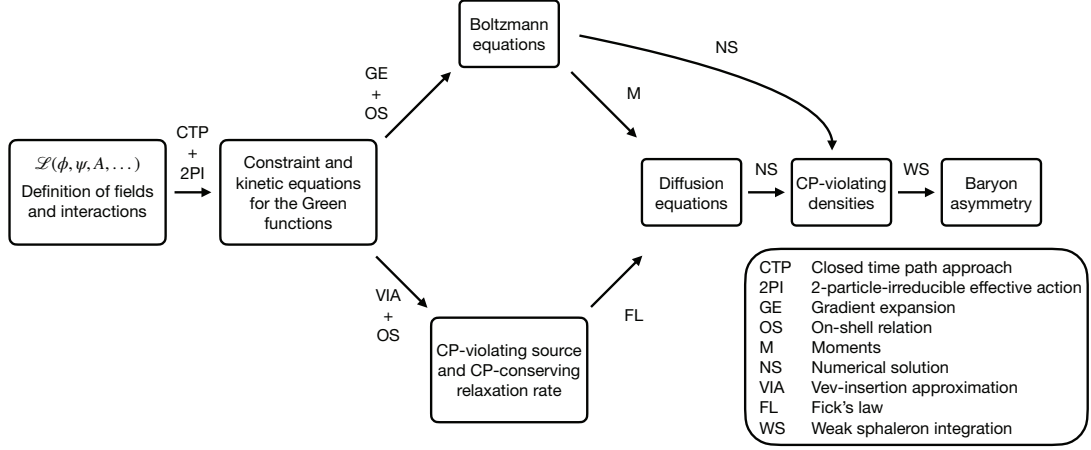
In front of the bubble walls, then, the EW sphalerons convert the CP asymmetry (in particular, the asymmetry in left-handed particles over their anti-particles  $n_L - \bar{n}_L$ , as sphalerons are configurations which interpolate between distinct  $\text{SU}(2)_L$  vacua) into a baryon asymmetry:

$$\partial_\mu j_B^\mu = -\frac{N_f}{2} [c_1 n_B + c_2 n_L] \quad (95)$$

where  $c_{1,2}$  are constants which depend on the weak sphaleron rate. The weak sphaleron rate is much slower than both the timescales of interactions with the bubble wall producing a net  $n_L$  and the diffusion ahead of the bubble wall, so the steps are effectively decoupled (one first computes  $n_L$  and then evaluates the sphaleron production of  $n_B$ ).

The EW sphalerons are active before EWSB: the transitions are proportional to  $\Gamma_{\text{sph}} \propto \exp(-E_{\text{sph}}(T)/T)$  where  $E_{\text{sph}}(T)$  is the sphaleron energy (which grows as  $T$  falls), which is proportional to the masses of the EW gauge bosons. The sphalerons are in equilibrium if  $\Gamma_{\text{sph}} > H$ . Therefore, the sphaleron rate is suppressed if  $v_h/T \gtrsim 1$ , in the interiors of the bubbles. This needs to be the case, because otherwise the net baryon number created in front of the expanding bubble walls is erased again inside the bubbles. The sharp turnoff of sphaleron processes across the wall solidifies the BAU.

Figure 3 gives an overview of the calculation of the BAU in EWBG. As is seen in this figure, there are two different methods in which the CP-violation in front of the bubble wall is established. There is the well-established semi-classical method [98–100] – the top route in the diagram. This method relies on the typical de Broglie wavelengths of the particles being small compared to the bubble wall width ( $L_{dB}/L_w \ll 1$ ), and an on-shell (OS) relation. The alternative approach to



**Figure 3:** Schematic overview of the computation of the baryon asymmetry in EWBG calculations. From [3].

derive the diffusion equations uses the so-called VEV-insertion approximation (VIA) [101–104], in which the VEV-dependent part of the mass is treated as a perturbation on top of the thermal masses. The VIA relates the self-energy and the Green’s function. Then, the CPV source and relaxation rate enter into the diffusion equations obtained from Fick’s law. For  $O(1)$  couplings between the CP-violating particle and the scalar field the VEV expansion should break down, though the VIA is still a topic of active debate outside this regime [105–110].

#### 4.2.2 Spontaneous baryogenesis

The mechanism called spontaneous baryogenesis [111] is an interesting counter-example to the universal validity of the Sakharov conditions, as it can happen in thermal equilibrium. There are a lot of different implementations of the mechanism, which share a few basic features. Namely, there is a term in the Lagrangian of the form

$$\mathcal{L} \ni \frac{\partial_\mu \phi}{f} j_B^\mu \quad (96)$$

where  $\phi$  is a scalar particle, usually a pseudo-Nambu Goldstone boson of the breaking of a global  $U(1)$  symmetry (could be gauged baryon number),  $f$  is some scale (if  $\phi$  is a pNGB it is the corresponding decay constant), and where  $j_B^\mu$  is the baryon current (87).

Then, one assumes  $\phi$  is spatially homogeneous, for example as it is smoothed out by inflation, which means

$$\partial_\mu \phi \sim \partial_t \phi \quad \rightarrow \quad \partial_\mu \phi j_B^\mu \sim \partial_t \phi j_B^0 = \partial_t \phi n_B \quad (97)$$

recognising that the 0-component of a current is a number density, and

$$n_B \equiv \sum_i n_i - \bar{n}_i. \quad (98)$$

Then, if  $\phi$  starts to roll, it acts like a baryon chemical potential

$$\frac{\dot{\phi}}{f} = \mu_B \quad \rightarrow \quad \mathcal{L} \ni \mu_B n_B. \quad (99)$$

and breaks CPT dynamically. We can further write

$$\begin{aligned} n_i - \bar{n}_i &= g_i \int \frac{d^3 p}{(2\pi)^3} \left[ \frac{1}{e^{(p-\mu_i)/T} + 1} - \frac{1}{e^{(p+\mu_i)/T} + 1} \right] \\ &= \frac{g_i}{6} \mu_i T^2 \left( 1 + O\left(\frac{\mu_i^2}{T^2}\right) \right). \end{aligned} \quad (100)$$

In the presence of B-violating interactions which freeze out at some  $T_f$ , another necessary ingredient for spontaneous baryogenesis, the final baryon number is thus given by,

$$n_B \propto \left. \frac{\dot{\phi}}{f} T^2 \right|_{T=T_f}. \quad (101)$$

Note that eventually the system falls to its ground state which has  $\dot{\phi} = 0$ , such that if the baryon number violating interactions do not freeze out before this, eventually the Universe would relax to  $B = 0$ .

A variation to spontaneous baryogenesis is QCD baryogenesis [83, 84], which relies on an early cosmological period of confinement. Such a period could be realised by a modification to the gluon kinetic term, i.e.

$$-\frac{1}{4} \left( \frac{1}{g_{s,0}^2} + \frac{S}{M_*} \right) G_{\mu\nu}^a G^{a\mu\nu} \quad (102)$$

where  $S$  is a SM singlet,  $M_*$  some mass scale, and  $g_{s,0}$  the strong coupling in the absence of  $S$ . Now, one can derive the running of the effective strong coupling, which leads to a confinement scale

$$\Lambda_{\text{QCD}}(v_S) = \Lambda_0 e^{\frac{24\pi^2}{2N_f - 33} \frac{v_S}{M_*}} \quad (103)$$

where  $\Lambda_0$  is the QCD scale for  $v_S = 0$  (note that this does not need to coincide with the SM value; we may live in a vacuum with  $v_S \neq 0$ ). Now for  $v_S < 0$  QCD confines earlier, and may even confine before EWSB. If this is the case, QCD confinement could be a first order phase transition, as all the SM quarks are light before it happens. Moreover, confinement *causes* EWSB, as the quark yukawa couplings give a linear term in the Higgs potential,

$$V(h) \ni y \langle \bar{q} q \rangle h \sim y \Lambda_{\text{QCD}}^3 h. \quad (104)$$

This means that the sphaleron processes are active before confinement, but freeze out at the high confinement scale.

Now, the QCD axion may play the role of the field  $\phi$  in (96). In QCD axion models, the strong sector of the SM may feature CP violation before confinement. At the QCD scale, the axion potential switches on, and dynamically relaxes the CP-violating  $\bar{\theta}$ -angle to zero. The action potential can be related to the  $\langle G\tilde{G} \rangle$  condensate as

$$\frac{\alpha_s}{8\pi} \langle G\tilde{G} \rangle = m_a^2 f_a^2 \sin \bar{\theta} \quad (105)$$

Then, the CP-violation in the strong sector could be communicated to the weak sector through the  $\eta'$  meson, which talks to both  $G\tilde{G}$  and  $W\tilde{W}$ .

Finally, the resulting chemical potential can then be estimated as

$$\mu_B \propto \frac{1}{f_\pi^2 m_\eta^2} \frac{d}{dt} [\sin \bar{\theta} m_a^2 f_a^2] \quad (106)$$

in particular as  $m_a(T)$  evolves, this will be non-zero.

Of course, we live in a vacuum with  $\Lambda_{\text{QCD}} \sim 1 \text{ GeV}$ , which we know from e.g. meson masses. Therefore, the period of early confinement needs to be followed by either a deconfinement phase transition, or a relaxation processes. Details of how that may occur in a minimal model can be found in [84, 112].

#### 4.2.3 Affleck-Dine baryogenesis

The Affleck-Dine mechanism provides a way of generating the BAU during and after inflation [113]. The mechanism requires the existence of flat directions with non-zero baryon number: scalar fields charged under  $B$  (or  $L$ , or  $B - L$ ) with a shallow potential, which are typically motivated in the context of supersymmetry. In its simplest form, the Affleck-Dine field is a complex scalar field, with a potential that looks like

$$V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4 + \frac{1}{4}\lambda'(\phi^4 - \phi^{*4}). \quad (107)$$

Clearly, the last term breaks the  $U(1)$  symmetry otherwise associated with the field. If  $V(\phi)^{1/4} \ll H$  during inflation, the field is stabilised by Hubble friction: the field is not in thermal equilibrium. If this happens at nonzero  $\langle\phi\rangle$ , it is clear CP is also broken.

As inflation ends and the universe reheats, the complex  $\phi$  field begins to oscillate around its potential minima. Due to the  $\lambda'$  term, the Affleck-Dine field experiences a complex phase evolution as it oscillates. This rotation gives rise to the creation of a nonzero baryon number density (the zeroth component of the baryon current carried by  $\phi$ ),

$$n_B = ic_B (\dot{\phi}^* \phi - \phi^* \dot{\phi}) \quad (108)$$

where  $c_B$  is the  $B$  charge of the Affleck-Dine field. As the amplitude of oscillations redshifts away, the first few oscillations give the dominant contribution to the generated baryon asymmetry.

For high scale inflation, the electroweak sphalerons are still active after the creation of baryon number by the Affleck-Dine mechanism. In this case, the sphalerons will redistribute some of the created baryon number into lepton number [114]. However, as the electroweak sphalerons only violate  $B + L$ , a net asymmetry remains.

#### 4.2.4 Leptogenesis

To finish this section, let us briefly consider a class of models of baryogenesis where the departure from thermal equilibrium comes from decays. In theories of leptogenesis, decays of heavy sterile neutrinos create a lepton number asymmetry, which can then be converted to a baryon number by the electroweak sphalerons [115]. In these scenarios, the heavy sterile neutrinos can play a double role, as they can also function to explain the light neutrino masses via a type-1 seesaw



model. The SM Lagrangian is supplemented a Majorana mass for the sterile neutrinos  $N$  as well as by an interaction<sup>27</sup>

$$\mathcal{L} \ni -\frac{1}{2}M_N N^2 - \lambda L H N + h.c. \quad (109)$$

where  $\lambda$  is a coupling,  $L$  a left-handed SM lepton,  $H$  the Higgs doublet, and  $N$  the sterile neutrino. This interaction can give rise to the following decays,

$$N \rightarrow LH \quad \text{or} \quad N \rightarrow \bar{L}\bar{H}.$$

the first of which violates lepton number by  $\Delta L = 1$ , and the second violates it by  $\Delta L = -1$ . It is clear that if these processes occur at the same rate, the net lepton number generated is zero. However, CP-violation comes from the interference between the tree-level decay diagram and the 1-loop diagram, such that  $\Gamma(N \rightarrow HL) < \Gamma(N \rightarrow \bar{H}\bar{L})$ . In particular, in the scenario where there are three right handed neutrinos with  $M_1 \ll M_2, M_3$ , the interference between the different diagrams implies,

$$\begin{aligned} \epsilon &= \frac{\Gamma(N_1 \rightarrow HL) - \Gamma(N_1 \rightarrow \bar{H}\bar{L})}{\Gamma(N_1 \rightarrow HL) + \Gamma(N_1 \rightarrow \bar{H}\bar{L})} \\ &\propto \frac{\sum_{a=2,3} \text{Im}[(\lambda^\dagger \lambda)_{a1}^2] \frac{M_1}{M_a}}{(\lambda^\dagger \lambda)_{11}}. \end{aligned} \quad (110)$$

The produced lepton asymmetry is not the end of the story, as the decay products can thermalise and give rise to so-called *washout* processes, include lepton-number-violating scatterings, decays, and inverse decays, which can reduce the produced asymmetry. It is possible to set constraints on the masses in couplings in the neutrino sector from the avoidance of washout and the necessity of out-of-equilibrium decays.

The resulting baryon number can be expressed in terms of the final baryon number, noting that the EW sphalerons will drive  $B + L \rightarrow 0$ . This implies  $B_i = (B_i + L_i)/2 + (B_i - L_i)/2 \rightarrow (B_f - L_f)/2 = B_f$  such that if leptogenesis produces an initial  $L_i$ , the resulting baryon number after sphaleron processes is  $B_f \sim -L_i/2$  (a more detailed calculation shows a slight departure from this relation, see [114]).

## 5. Gravitational waves from the Early Universe

Much of the present interest in primordial first order phase transitions can be explained by the prospect of detecting a stochastic gravitational wave background in the near future. As we will see, first order phase transitions represent one of the early Universe phenomena that could give rise to such a spectrum. If this signal were to be detected, it would provide us with the earliest known probe of the Universe. In this section, we will explore the anticipated phenomenology by first revisiting the fundamental aspects of gravitational wave sources and detection, followed by an examination of gravitational wave backgrounds.

<sup>27</sup>I am using simplified notation here, for example suppressing generation indices. For a more complete treatment, please refer to reviews such as [116–118].



## 5.1 Gravitational wave basics

The simplest way to get some intuition for gravitational waves is through linearized gravity,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad (111)$$

where  $\eta_{\mu\nu}$  is the Minkowski metric of flat space and  $h_{\mu\nu}$  is a small perturbation. Here the Greek indices run over space-time dimensions:  $\mu, \nu = 0, 1, 2, 3$ . From Einstein's field equations (to linear order in  $h_{\mu\nu}$ ), one can obtain a wave equation for the (massless) metric perturbation  $h_{\mu\nu}$ , where the source term is given by the *energy-momentum tensor* of the GW-source:<sup>28</sup>

$$\square h_{\mu\nu} = -16\pi G_N T_{\mu\nu}, \quad (112)$$

where the  $\square$  operator is the d'Alembertian:  $\square = \partial_\mu \partial^\mu$ .<sup>29</sup> To derive this equation, one has to assume the Lorentz gauge,  $\partial^\nu h_{\mu\nu} = 0$ .

One thing that we can immediately examine is the what degrees of freedom live in  $h_{\mu\nu}$ . Because  $h_{\mu\nu}$  is a symmetric rank 2 tensor, it has in principle 10 independent components. However, not all of these correspond to physical degrees of freedom: four of these are gauge degrees of freedom (corresponding to coordinate transformations). In fact, in deriving the wave equation, we have already had to assume the Lorentz gauge, fixing four components. If we are interested in the propagation of the wave, we solve the wave equation *outside the source*,  $\square h_{\mu\nu} = 0$ . This can be used to determine four more components, so we should end up with  $10 - 2 \times 4 = 2$  propagating degrees of freedom. Typically, one fixes the gauge further into the transverse-traceless gauge (TT, meaning  $h^\mu_\mu = 0$ ,  $h_{0\mu} = 0$ , and  $\partial^i h_{ij} = 0$ ) and the two independent components are the + and  $\times$  polarisations of the wave. We can find an example: if we consider  $h_{\mu\nu}$  to be a monochromatic wave traveling in the  $z$ -direction, it will take the following form in the TT gauge:

$$h_{\mu\nu}^{\text{TT}} = \left[ h_+ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + h_\times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right] e^{ikz}. \quad (113)$$

The TT components are actually the only radiative components of  $h_{\mu\nu}$ ; the fact that further components obey a wave equation in the Lorentz gauge is a gauge artifact.

Just as in electrodynamics, the linear wave equation (112) can be solved by a retarded Green's function:

$$h_{ij}^{\text{TT}}(t, \mathbf{x}) = 4G_N \Lambda_{ij,kl} \int d^3x' \frac{1}{|\mathbf{x} - \mathbf{x}'|} T_{kl}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}') \quad (114)$$

where  $\Lambda_{ij,kl} = P_{ik}P_{jl} - P_{ij}P_{kl}/2$  with  $P_{ij}(\hat{\mathbf{n}}) = \delta_{ij} - n_i n_j$  (where  $\hat{\mathbf{n}}$  is a unit vector in the direction of propagation, e.g.  $\hat{\mathbf{n}} = (0, 0, 1)$  in our case above) is the transverse-traceless projector. As

<sup>28</sup>This is actually a wave equation for the trace-reversed metric perturbation  $\tilde{h}_{\mu\nu} = h_{\mu\nu} - \eta_{\mu\nu} h/2$  where  $h = h^\mu_\mu$  (note that this means that  $\tilde{h} = -h$ ) and I have dropped the tilde, but for our purposes the difference is unimportant. See e.g. Refs. [119, 120] for more detailed introductions to this equation and the solutions discussed in this section.

<sup>29</sup>Contrary to some of the literature on gravitational waves, I will continue to use particle physics units  $c = 1$  but  $G_N \neq 1$ .

this solution is in the TT gauge, it only features spatial dimensions ( $i, j = 1, 2, 3$ ); the temporal contribution can be related to it using energy-momentum conservation.

The solution to the wave equation (114) is not in general not very insightful. However, in many cases, we can learn a lot from its limiting behaviour. For example, if the distance  $r$  to the source is large compared to the typical size of the source, we can replace  $|\mathbf{x} - \mathbf{x}'| \approx r$ :

$$h_{ij}^{\text{TT}}(t, \mathbf{x}) = \frac{4G_N}{r} \Lambda_{ij,kl} \int d^3x' T_{kl}(t - r, \mathbf{x}'). \quad (115)$$

The replacement in the time argument of  $T_{kl}$  corresponds to the assumption that the motion inside the source is non-relativistic (i.e.  $v \ll 1$ ). With a bit of massaging (taking advantage of the conservation of the energy-momentum tensor to write  $\int d^3x T_{ij} = \int d^3x \partial_t^2 (T_{00} x_i x_j) / 2 + \text{total derivatives}$ ), (115) can be recognized as the first term in the expansion in spherical harmonics: the quadrupole (i.e.  $h_{ij} = h_{ij}^{[\text{quad}]} + \dots$ ). It can be rewritten as

$$h_{ij}^{[\text{quad}]} = \frac{2G}{r} \Lambda_{ij,kl} \ddot{Q}_{kl} \quad \text{where} \quad Q_{ij} = \int d^3x \rho(t, \mathbf{x}) \left( x_i x_j - \frac{1}{3} r^2 \delta_{ij} \right) \quad (116)$$

which, as we anticipated, depends only on spatial dimensions.  $Q_{ij}$  is the *quadrupole moment* tensor of the source (the dots denote time derivatives), and  $\rho(t, \mathbf{x})$  is a time-varying mass density distribution.

We made some assumptions in this derivation, most importantly the validity of the linearised version of the Einstein equations. This linearised version is not always valid; in particular, as gravitational waves carry energy,  $h_{\mu\nu}$  itself should affect the metric. However, it turns out that it is straightforward to generalise the derivation, and all that it will cost is a replacement of the stress-energy tensor by  $T_{\mu\nu} + t_{\mu\nu}$ , where  $t_{\mu\nu}$  is a pseudo-tensor (constructed from  $h_{\mu\nu}$ ) which takes into account the backreaction of  $h_{\mu\nu}$  on the metric. In situations where self-gravity dominates,  $t_{\mu\nu}$  describes the gravitational binding energy. Thus, the quadrupole formula can be applied to situations like binary mergers.

Let us consider the result (116) for a moment. Unlike electromagnetism, which features dipole radiation, gravitational waves are generated by accelerated sources with nonzero quadrupole moments (or, in principle, higher order multipoles). One way to explain the difference is using the fact that in electromagnetism you have opposite charges (positive and negative electric charge), whereas in gravity, the charge (mass) is positive definite. Dipole moments describe the distribution of charge away from some centre in some direction, so if the charges have equal sign and the center of mass is picked as the center, they vanish. In contrast, quadrupole moments do not depend directly on the sign of the charges.<sup>30</sup>

There are a few practical lessons we can learn from the quadrupole formula for a single source (116). An obvious one, which nonetheless gives some good first intuition, is that spherically symmetric systems do not generate gravitational waves. Moreover, static or uniformly moving systems do not generate gravitational waves either: acceleration is needed.

To get some intuition for the types of sources that generate observable gravitational waves, we can make a very naive estimate of the size of  $h_{\text{quad}}$ . Using dimensional analysis,  $\ddot{Q}$  should have mass dimension one. Let us take it to be  $\ddot{Q} = 60 M_\odot$ , the mass of a very large black hole. For  $r$ ,

<sup>30</sup>In general relativity, the monopole describes the total amount of mass.

let us take  $r = 140 \times 10^6$  ly, the distance to the first binary NS merger measured by LIGO – one of the more nearby ones. These two numbers imply together that  $h_{\text{quad}} \sim 10^{-19}$ : a very small, dimensionless number, despite the large “source” we chose. The reason is of course that Newton’s constant makes for a very small coupling constant. This is why only gravitational waves from enormous and relatively nearby sources can be observed.

On that topic of detection, let us make one last remark about (116). Gravitational waves get weaker with  $1/r$ , not with  $1/r^2$ , like the gravitational force itself. The total *energy* of the quadrupole radiation would fall off with  $1/r^2$ , but that is not what gets detected. Unlike photon detectors, which absorb EM radiation, gravitational wave detectors measure the displacement of test masses by the action of a gravitational waves. Therefore, the signal is proportional to the amplitude of the wave, and only falls off as  $1/r$ . This allows future gravitational wave detectors to look back at very high redshifts compared to EM telescopes.

## 5.2 Stochastic gravitational wave backgrounds

The results in the previous section apply to transient signals: binary mergers of compact objects which are “loud” enough to be individually observable, and perhaps other sources such as supernova explosions. While it is still true that for an observable signal, an energetic changing quadrupole moment is needed, gravitational waves from the early Universe fall into a different category: that of stochastic gravitational wave backgrounds (SGWB). This category includes gravitational waves from first order phase transitions, as we studied earlier in these notes. Much like the CMB, such gravitational wave sources are observable via their power spectrum.<sup>31</sup>

To compute the SGWB power spectrum, we will use two-point correlators. For example, to find the energy in a stochastic gravitational wave background (SGWB), one might expect to need to compute the two-point correlator of the energy-momentum tensor (and in particular its TT part):  $\langle T_{ij}(x)T^{ij}(y) \rangle$ . It is useful to use a plane wave decomposition,

$$\begin{aligned} h_{ij}(t, x) &= \int_{-\infty}^{\infty} df \int_{S^2} d^2\Omega_k h_{ij}(f, k) e^{2\pi i f (t - k \cdot x/c)} \\ &= \sum_{A=+, \times} \int_{-\infty}^{\infty} df \int_{S^2} d^2\Omega_k e_{ij}^A h_A(f, k) e^{2\pi i f (t - k \cdot x/c)} \end{aligned} \quad (117)$$

For a stochastic background, the Fourier coefficients  $h_A(f, k)$  are random variables; their ensemble averages are more interesting. The  $e_{ij}^A$  are again polarisation tensors. Since our stochastic background will consist of waves with all possible propagation directions (labeled by  $k$ ), we can no longer use the transverse-traceless gauge to assume  $i, j = 1, 2$ . Instead, the indices need to go from 1 to 3.

We are interested in cosmological sources, such that our stochastic background is stationary (in time), approximately Gaussian (i.e. all  $n$ -point functions can be expressed in terms of two-point functions  $\langle h_A h_{A'} \rangle$ ), and isotropic. Most cosmological stochastic backgrounds are also unpolarised, although there are some examples which have definite polarisation (gauge boson production after axion inflation, for example). The SGWB due to first order phase transitions is not expected to be

<sup>31</sup>Note that binary mergers also contribute a stochastic gravitational wave background, which is the sum of all binary merger signals which are cannot individually be resolved.

polarised, such that the two point function of Fourier coefficients  $h_A(f, k)$  can be characterised by the so-called spectral density  $S_h$ :

$$\langle h_A^*(f, \hat{\mathbf{n}}) h_{A'}(f', \hat{\mathbf{n}}') \rangle = \overbrace{\delta(f - f')}^{\text{stationarity}} \overbrace{\frac{\delta^{(2)}(\hat{\mathbf{n}}, \hat{\mathbf{n}}')}{4\pi}}^{\text{isotropy}} \overbrace{\delta_{AA'}}^{\text{unpolarised}} \frac{1}{2} S_h(f) \quad (118)$$

where the  $4\pi$  is just a normalisation choice, such that integrating (118) over  $\hat{\mathbf{n}}$  and  $\hat{\mathbf{n}}'$  leads to  $\delta(f, f') \delta_{AA'} S_h(f)/2$  (as  $d^2\hat{\mathbf{n}} = d \cos \theta d\phi$ ). The same goes for the factor  $1/2$ : a normalisation choice.

The correlator (118) implies that

$$\langle h_{ij}(t) h^{ij}(t) \rangle = 4 \int_0^\infty df S_h(f) \quad (119)$$

where the 4 follows from the normalisation of the polarisation tensor. In experiments, the spectral density  $S_h$  can be compared with the strain noise  $S_n$ . An equivalent definition that is sometimes used is the signal characteristic strain amplitude  $h_c$ , which for a stochastic gravitational wave background is defined by

$$\langle h_{ij} h^{ij} \rangle = 2 \int_{f=0}^{f=\infty} d \log f h_c^2(f) \quad (120)$$

such that  $S_f$  and  $h_c$  are related by

$$S_f = \frac{1}{2} \frac{h_c^2}{f}. \quad (121)$$

For cosmological sources we are often interested in the energy density of the gravitational wave background. This is given by<sup>32</sup>

$$\rho_{\text{GW}} = \frac{1}{32\pi G_N} \langle \dot{h}_{ij}^2(t, \mathbf{x}) \rangle. \quad (122)$$

For cosmological sources, the energy density in the SGWB is often written in terms of the critical energy density  $\rho_c$  using the parameter  $\Omega_{\text{GW}}(f)$ ,

$$\begin{aligned} \rho_{\text{GW}} &= \int_{f=0}^{f=\infty} d \log f \frac{d\rho_{\text{GW}}}{d \log f} \\ &\equiv \rho_c \int_{f=0}^{f=\infty} d \log f \Omega_{\text{GW}}(f) \end{aligned} \quad (123)$$

where  $\rho_c = 3H_0^2/8\pi G_N$  is the critical energy density. This definition of  $\Omega_{\text{GW}}(f)$  can be a little confusing; as can be seen, it is the quantity which, *when integrated over log f*, yields the fractional energy density in gravitational waves  $\Omega_{\text{GW}} = \rho_{\text{GW}}/\rho_c$ .

<sup>32</sup>Fun fact: for some time, it was controversial whether gravitational waves carried energy or not. If you're curious, look into the "sticky bead" gedanken experiment.

### 5.3 Gravitational waves from phase transitions

As we have seen, first order phase transitions are described by the nucleation of spherical field configurations, “bubbles”, of true vacuum in a background of false vacuum. Critically sized bubbles subsequently expand. However, we have also seen that spherically symmetric configurations cannot generate gravitational waves. It is only when expanding bubbles collide, and when the accompanying plasma shells collide, that gravitational waves are sourced [121, 122].

Thus, there are three main sources of gravitational wave generation during (and after) a first order phase transition: the collisions of the walls of the bubble, the plasma shells which move at the speed of sound, and the turbulence in the plasma which results from the decay of this motion. Which contribution to the spectrum is most important depends on how the latent heat of the phase transition is released. The collisions of the scalar field shells themselves carry most energy in the so-called runaway transitions discussed in an earlier section, where  $\gamma_w \rightarrow \infty$ . Away from this limit, the latent heat goes into acoustic motion of the plasma, which has a finite lifetime before it decays into turbulence.

The calculation of the GW spectra has been approached both analytically (e.g. [123–125]) and through simulations (e.g. [126–128]). For the scalar field contribution, an oft-employed approximation is called the envelop approximation, and considers the bubble walls to be infinitesimally thin and to disappear immediately upon collision. In hydrodynamical simulations, correlators of the  $T_{\mu\nu}$  can be extracted from the relativistic plasma during and after the transition.

We have previously looked at two different scenarios for bubble expansion: deflagration and detonation, distinguished by the bubble wall speed compared to the speed of sound in the plasma. The most relevant situation for gravitational wave production is the detonation case, both because this scenario produces the largest spectra and because this scenario is most likely realised for interesting scenarios.<sup>33</sup> The envelope approximation works particularly well for detonations, because the kinetic-energy density is concentrated in a thin shell near the bubble wall.<sup>34</sup> Moreover, because of the supersonic wall speed, the collisions do not affect the propagation of the parts of the bubble wall which have not collided yet.

Even without doing a detailed calculation of the gravitational wave spectrum, it is possible to make a few estimates. This is because the typical frequency of a gravitational wave signal that originated in the early Universe today is largely determined by cosmological redshift. In other words,

$$f_0 = \frac{a(t_*)}{a(t_0)} f_* \sim 7.8 \times 10^{-16} \left( \frac{106}{g_*(T_*)} \right)^{1/3} \left( \frac{100 \text{ GeV}}{T_*} \right) f_* \quad (124)$$

where the subscript 0 indicates a value today and the \* indicates the value at emission. Here  $g_*$  is the number of degrees of freedom. The second line follows from the conservation of entropy during radiation domination (s.t.  $g_*(T_*)T_*^3 a_*^3 = g_{*,0}T_0^3 a_0^3 \rightarrow a_* = (T_0/T_*)(g_0/g_*)^{1/3}a_0$ ).

<sup>33</sup>There is friction on the bubble wall from interactions with the plasma, and transition radiation. However, this is not typically large enough of a break to realise  $v_w < c_s$ .

<sup>34</sup>A fully analytic calculation the reader is referred to [123].

The typical frequency at production  $f_*$  can be estimated using the parameter  $\beta$ , a typical rate of the phase transition. Remember that the false vacuum decay rate was proportional to  $e^{-S_E/T}$ . We can expand the exponentiated quantity around a particular time  $t'$ ,

$$\frac{S_E}{T} = \frac{S_E}{T} \Big|_{t=t'} + (t-t') \underbrace{\frac{d}{dt} \frac{S_E}{T} \Big|_{t=t'}}_{-\beta} + O(t-t')^2 \quad (125)$$

defining the parameter  $\beta = d/dt S_E/T$ , such that

$$\Gamma = \Gamma(t') e^{\beta(t-t')}. \quad (126)$$

If the time  $t'$  is chosen to be some characteristic time during the transition (for example the nucleation time),  $\beta$  is often said to be the rate of the phase transition, and  $\beta^{-1}$  its typical duration.

Associating  $f_*$  with  $\beta$ ,  $\beta \sim f_*$ , we can write

$$\begin{aligned} f_0 &\sim 7.8 \times 10^{-16} \left( \frac{106}{g_*(T_*)} \right)^{1/3} \left( \frac{100 \text{ GeV}}{T_*} \right) H_* \frac{\beta}{H_*} \\ &\sim 10^{-5} \left( \frac{g_*(T_*)}{106} \right)^{1/6} \left( \frac{T_*}{100 \text{ GeV}} \right) \frac{\beta}{H_*} \text{ Hz} \end{aligned} \quad (127)$$

using that  $\beta$  is usually normalised to the Hubble rate at the time of the transition. A more detailed calculation takes into account wall velocities away from  $v_w = 1$ , as well as the fraction of space-time volume which has not decayed yet to estimate the bubble radius at collision to be  $R = (8\pi)^{1/3} v_w / \beta$  and  $f \sim R^{-1}$ , see [129, 130].

The ratio  $\beta/H$  can be calculated for a given transition, and is usually  $O(10 - 10^3)$ . Note that if  $\beta \lesssim H$ , the phase transition is not expected to complete. Fixing  $\beta/H = 100$ , and using  $g_* = 106$  as in the SM (noting that  $f_0$  does not depend on the latter strongly), we can estimate the frequencies of phase transitions happening in the early Universe. For example, a phase transition happening at the weak scale,  $T_* = 100 \text{ GeV}$ , implies a gravitational wave spectrum today peaking at  $f_0 \sim 10^{-3} \text{ Hz}$ , whereas a phase transition happening at  $T_* = 10^6 \text{ GeV}$  would peak at  $f_0 \sim 10 \text{ Hz}$ . We can compare this with the typical frequency windows of experiments from which you can see that weak-scale phase transitions are in the range of space-based interferometers, explaining the interest into probing electroweak baryogenesis with LISA [131]. To explain the very recently confirmed low-frequency SGWB in pulsar timing arrays [132–135], the temperature of the phase transition would have to be much lower.

In these notes, we will not attempt to derive the amplitude of the SGWB from phase transitions, referring instead to results of simulations. Assuming the sound wave spectrum dominates these give [128, 136],

$$h^2 \Omega_{\text{sw}}(f) = 8.5 \times 10^{-6} \left( \frac{100}{g_*} \right)^{1/3} \kappa_f^2 \frac{\alpha^2}{(1+\alpha)^2} \left( \frac{H_*}{\beta} \right) S_{\text{sw}}(f) \left( 1 - \frac{1}{\sqrt{1+2H_* t_{\text{sw}}}} \right), \quad (128)$$

$$\kappa_f = \frac{\alpha}{\alpha + 0.083\sqrt{\alpha} + 0.73}, \quad (129)$$



where  $S_{\text{sw}}(f)$  encodes the frequency dependence; at the peak,  $S_{\text{sw}}(f) = 1$ . Here  $t_{\text{sw}}$  denotes the timescale on which acoustic waves are active, as eventually the plasma shells decay into turbulent motion. One would expect this to happen on timescales [137, 138]

$$H_* t_{\text{sw}} = \frac{2(8\pi)^{1/3} \sqrt{1+\alpha}}{\sqrt{3\alpha} \kappa_f \beta / H_*}, \quad (130)$$

such that the faster the transition, the smaller the bubbles are at coalescence, and the faster the decay of the motion in the plasma into turbulence.

Lastly, about the spectral shape, we can make the following comments, borrowing an argument from [139]. We expect the energy density to go like  $\rho_{\text{GW}} \propto \langle \dot{h}_{ij} \dot{h}_{ij} \rangle \propto f^2 P_s(f)$ , where we have assumed  $\dot{h} \simeq f h$  and where  $P_s(f)$  is some power spectrum. At frequencies smaller than the inverse of a characteristic length of the phase transition (for instance the inverse size of the bubbles at collision),  $f < f_{\text{peak}}$ , we do not expect the power spectrum  $P_s(f)$  to scale with  $f$ . Then, we find

$$\Omega_{\text{GW}}(f) \propto \frac{d\rho_{\text{GW}}}{d \log f} \propto f^3 \quad f \ll f_{\text{peak}}. \quad (131)$$

At frequencies larger than the characteristic length,  $P_s(f)$  has a negative power, to keep the total energy in gravitational waves finite. How quickly  $\Omega_{\text{GW}}(f)$  falls off depends on how coherent the source is, and typically differs for gravitational waves generated by plasma motion and scalar field shells.

## 6. Conclusions and outlook

The prospect of detecting gravitational wave signals from the pre-BBN Universe has reinvigorated research into cosmological first order phase transitions and related models of baryogenesis. These are mature research directions, with seminal works originating in the 1970s, 80s, and 90s. However, many significant technical challenges still lay ahead in characterising primordial gravitational wave spectra accurately, some of which we have identified in these notes.

These notes started with an introductory overview of thermal field theory and a relatively in-depth discussion (in comparison to other topics in this work, certainly not in comparison with the literature) of false vacuum decay. I have found that studying these topics helped me develop a deeper understanding and greater appreciation of field theory. In my opinion, there are still great opportunities for technical developments in these areas, in particular in situations where the usual perturbative expansions are not reliable. This technical progress is essential to complement the many model-building efforts that are already flourishing, as it is required to make accurate phenomenological predictions. The same is true for calculations of the baryon asymmetry generated during first order phase transitions, though we spend less time on such in-depth calculations here.

As gravitational waves were not introduced by any other lecturer at TASI 2022, I included a quick introduction to the topic. This was meant to develop some intuition for the types of events that could generate gravitational waves, before moving on to stochastic backgrounds such as generated by primordial sources. Research on the gravitational wave phenomenology of first order phase transitions has to great extent been driven by enormous progress on hydrodynamical simulations of early-Universe bubble dynamics in the last decade. These notes have not focused



on this aspect of the predictions – primarily because this is not my area of expertise – but there is a substantial amount of important ongoing research in this field, including on accurately capturing magnetohydrodynamic turbulence.

In conclusion, the study of gravitational waves from phase transitions is a rapidly expanding field. The significant technical challenges are matched by the promise of this program, as it has the potential to provide profound insights into our understanding of the universe and its history. I hope that these notes are useful to anyone interested in contributing to this active area of research.

## Acknowledgments

I am very thankful to the scientific organisers of the 2022 TASI summer school (JiJi Fan, Stefania Gori, Lian-Tao Wang) for inviting me to give these lectures and encouraging me to write up these lecture notes. I am also grateful to the local TASI organisers (Tom DeGrand, Oliver DeWolfe, and Ethan Neil) for ensuring the school ran smoothly. I thank the students attending TASI 2022 for their enthusiasm and many questions.

I am indebted to my teachers and collaborators – who I will not list individually, for fear of accidental omissions – for teaching me many aspects relevant to these lectures over the years. Specifically, I am grateful to Oli Gould, Nell Hall, and Philipp Schicho, who provided insightful comments on these notes. Lastly, I am grateful to my partner for accompanying me to Boulder to give these lectures a couple of months before our son was born.

## References

- [1] D. Croon, O. Gould, P. Schicho, T.V.I. Tenkanen and G. White, *Theoretical uncertainties for cosmological first-order phase transitions*, *JHEP* **04** (2021) 055 [[2009.10080](#)].
- [2] P.H. Ginsparg, *First Order and Second Order Phase Transitions in Gauge Theories at Finite Temperature*, *Nucl. Phys. B* **170** (1980) 388.
- [3] P. Asadi et al., *Early-Universe Model Building*, [2203.06680](#).
- [4] J.I. Kapusta and C. Gale, *Finite-temperature field theory: Principles and applications*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (2011), [10.1017/CBO9780511535130](#).
- [5] M. Quiros, *Finite temperature field theory and phase transitions*, in *ICTP Summer School in High-Energy Physics and Cosmology*, pp. 187–259, 1, 1999 [[hep-ph/9901312](#)].
- [6] M. Laine and A. Vuorinen, *Basics of Thermal Field Theory*, vol. 925, Springer (2016), [10.1007/978-3-319-31933-9](#), [[1701.01554](#)].
- [7] A. Das, *Finite temperature field theory*, World scientific (1997).
- [8] M.L. Bellac, *Thermal Field Theory*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (3, 2011), [10.1017/CBO9780511721700](#).

- [9] L.V. Keldysh et al., *Diagram technique for nonequilibrium processes*, *Sov. Phys. JETP* **20** (1965) 1018.
- [10] J.S. Schwinger, *Brownian motion of a quantum oscillator*, *J. Math. Phys.* **2** (1961) 407.
- [11] R. Kubo, *Statistical mechanical theory of irreversible processes. I. General theory and simple applications in magnetic and conduction problems*, *J. Phys. Soc. Jap.* **12** (1957) 570.
- [12] P.C. Martin and J.S. Schwinger, *Theory of many particle systems. 1.*, *Phys. Rev.* **115** (1959) 1342.
- [13] T. Matsubara, *A New approach to quantum statistical mechanics*, *Prog. Theor. Phys.* **14** (1955) 351.
- [14] L. Dolan and R. Jackiw, *Symmetry Behavior at Finite Temperature*, *Phys. Rev. D* **9** (1974) 3320.
- [15] R. Jackiw, *Functional evaluation of the effective potential*, *Phys. Rev. D* **9** (1974) 1686.
- [16] K. Kajantie, M. Laine, K. Rummukainen and M.E. Shaposhnikov, *Generic rules for high temperature dimensional reduction and their application to the standard model*, *Nucl. Phys. B* **458** (1996) 90 [[hep-ph/9508379](#)].
- [17] A.D. Linde, *Infrared Problem in Thermodynamics of the Yang-Mills Gas*, *Phys. Lett. B* **96** (1980) 289.
- [18] D.A. Kirzhnits and A.D. Linde, *Symmetry Behavior in Gauge Theories*, *Annals Phys.* **101** (1976) 195.
- [19] R.R. Parwani, *Resummation in a hot scalar field theory*, *Phys. Rev. D* **45** (1992) 4695 [[hep-ph/9204216](#)].
- [20] P.B. Arnold and O. Espinosa, *The Effective potential and first order phase transitions: Beyond leading-order*, *Phys. Rev. D* **47** (1993) 3546 [[hep-ph/9212235](#)].
- [21] T. Appelquist and R.D. Pisarski, *High-Temperature Yang-Mills Theories and Three-Dimensional Quantum Chromodynamics*, *Phys. Rev. D* **23** (1981) 2305.
- [22] S. Nadkarni, *Dimensional Reduction in Hot QCD*, *Phys. Rev. D* **27** (1983) 917.
- [23] N.P. Landsman, *Limitations to Dimensional Reduction at High Temperature*, *Nucl. Phys. B* **322** (1989) 498.
- [24] K. Farakos, K. Kajantie, K. Rummukainen and M.E. Shaposhnikov, *3-D physics and the electroweak phase transition: Perturbation theory*, *Nucl. Phys.* **B425** (1994) 67 [[hep-ph/9404201](#)].
- [25] E. Braaten and A. Nieto, *Effective field theory approach to high temperature thermodynamics*, *Phys. Rev.* **D51** (1995) 6990 [[hep-ph/9501375](#)].

- [26] E. Braaten and A. Nieto, *Free energy of QCD at high temperature*, *Phys. Rev. D* **53** (1996) 3421 [[hep-ph/9510408](#)].
- [27] O. Gould, S. Güyer and K. Rummukainen, *First-order electroweak phase transitions: A nonperturbative update*, *Phys. Rev. D* **106** (2022) 114507 [[2205.07238](#)].
- [28] P.M. Schicho, T.V.I. Tenkanen and J. Österman, *Robust approach to thermal resummation: Standard Model meets a singlet*, *JHEP* **06** (2021) 130 [[2102.11145](#)].
- [29] K. Farakos, K. Kajantie, K. Rummukainen and M.E. Shaposhnikov, *3d physics and the electroweak phase transition: A Framework for lattice Monte Carlo analysis*, *Nucl. Phys.* **B442** (1995) 317 [[hep-lat/9412091](#)].
- [30] K. Kajantie, M. Laine, K. Rummukainen and M.E. Shaposhnikov, *The Electroweak phase transition: A Nonperturbative analysis*, *Nucl. Phys. B* **466** (1996) 189 [[hep-lat/9510020](#)].
- [31] K. Kajantie, M. Laine, K. Rummukainen and M.E. Shaposhnikov, *A Nonperturbative analysis of the finite  $T$  phase transition in  $SU(2) \times U(1)$  electroweak theory*, *Nucl. Phys.* **B493** (1997) 413 [[hep-lat/9612006](#)].
- [32] M. Laine, P. Schicho and Y. Schröder, *Soft thermal contributions to 3-loop gauge coupling*, *JHEP* **05** (2018) 037 [[1803.08689](#)].
- [33] L. Niemi, P. Schicho and T.V.I. Tenkanen, *Singlet-assisted electroweak phase transition at two loops*, *Phys. Rev. D* **103** (2021) 115035 [[2103.07467](#)].
- [34] O. Gould and T.V.I. Tenkanen, *On the perturbative expansion at high temperature and implications for cosmological phase transitions*, *JHEP* **06** (2021) 069 [[2104.04399](#)].
- [35] A. Ekstedt, O. Gould and J. Löfgren, *Radiative first-order phase transitions to next-to-next-to-leading order*, *Phys. Rev. D* **106** (2022) 036012 [[2205.07241](#)].
- [36] A. Ekstedt, P. Schicho and T.V.I. Tenkanen, *DRalgo: A package for effective field theory approach for thermal phase transitions*, *Comput. Phys. Commun.* **288** (2023) 108725 [[2205.08815](#)].
- [37] C. Wetterich, *Exact evolution equation for the effective potential*, *Phys. Lett. B* **301** (1993) 90 [[1710.05815](#)].
- [38] N. Dupuis, L. Canet, A. Eichhorn, W. Metzner, J.M. Pawłowski, M. Tissier et al., *The nonperturbative functional renormalization group and its applications*, [2006.04853](#).
- [39] D. Croon, E. Hall and H. Murayama, *Non-perturbative methods for false vacuum decay*, [2104.10687](#).
- [40] S.R. Coleman, *The Fate of the False Vacuum. I. Semiclassical Theory*, *Phys. Rev. D* **15** (1977) 2929.

- [41] C.G. Callan, Jr. and S.R. Coleman, *The Fate of the False Vacuum. 2. First Quantum Corrections*, *Phys. Rev. D* **16** (1977) 1762.
- [42] A. Andreassen, D. Farhi, W. Frost and M.D. Schwartz, *Direct Approach to Quantum Tunneling*, *Phys. Rev. Lett.* **117** (2016) 231601 [[1602.01102](#)].
- [43] A. Andreassen, D. Farhi, W. Frost and M.D. Schwartz, *Precision decay rate calculations in quantum field theory*, *Phys. Rev. D* **95** (2017) 085011 [[1604.06090](#)].
- [44] P. Athron, C. Balázs, A. Fowlie, L. Morris and L. Wu, *Cosmological phase transitions: from perturbative particle physics to gravitational waves*, [2305.02357](#).
- [45] W.-Y. Ai, *Aspects of false vacuum decay*, Ph.D. thesis, Munich, Tech. U., 2019.
- [46] F. Devoto, S. Devoto, L. Di Luzio and G. Ridolfi, *False vacuum decay: an introductory review*, *J. Phys. G* **49** (2022) 103001 [[2205.03140](#)].
- [47] W.-Y. Ai, B. Garbrecht and C. Tamarit, *Functional methods for false vacuum decay in real time*, *JHEP* **12** (2019) 095 [[1905.04236](#)].
- [48] S.R. Coleman, V. Glaser and A. Martin, *Action Minima Among Solutions to a Class of Euclidean Scalar Field Equations*, *Commun. Math. Phys.* **58** (1978) 211.
- [49] J. Langer, *Theory of the condensation point*, *Annals of Physics* **41** (1967) 108.
- [50] I. Affleck, *Quantum Statistical Metastability*, *Phys. Rev. Lett.* **46** (1981) 388.
- [51] O. Gould and J. Hirvonen, *Effective field theory approach to thermal bubble nucleation*, *Phys. Rev. D* **104** (2021) 096015 [[2108.04377](#)].
- [52] A. Ekstedt, *Bubble Nucleation to All Orders*, [2201.07331](#).
- [53] M. Kamionkowski, A. Kosowsky and M.S. Turner, *Gravitational radiation from first order phase transitions*, *Phys. Rev. D* **49** (1994) 2837 [[astro-ph/9310044](#)].
- [54] D. Bodeker and G.D. Moore, *Can electroweak bubble walls run away?*, *JCAP* **05** (2009) 009 [[0903.4099](#)].
- [55] D. Bodeker and G.D. Moore, *Electroweak Bubble Wall Speed Limit*, *JCAP* **05** (2017) 025 [[1703.08215](#)].
- [56] S. Höche, J. Kozaczuk, A.J. Long, J. Turner and Y. Wang, *Towards an all-orders calculation of the electroweak bubble wall velocity*, *JCAP* **03** (2021) 009 [[2007.10343](#)].
- [57] A. Azatov and M. Vanvlasselaer, *Bubble wall velocity: heavy physics effects*, *JCAP* **01** (2021) 058 [[2010.02590](#)].
- [58] Y. Gouttenoire, R. Jinno and F. Sala, *Friction pressure on relativistic bubble walls*, *JHEP* **05** (2022) 004 [[2112.07686](#)].

- [59] S. De Curtis, L.D. Rose, A. Guiggiani, A.G. Muyor and G. Panico, *Bubble wall dynamics at the electroweak phase transition*, *JHEP* **03** (2022) 163 [2201.08220].
- [60] M. Lewicki, V. Vaskonen and H. Veermäe, *Bubble dynamics in fluids with N-body simulations*, 2205.05667.
- [61] B. Laurent and J.M. Cline, *First principles determination of bubble wall velocity*, *Phys. Rev. D* **106** (2022) 023501 [2204.13120].
- [62] K. Kajantie, M. Laine, K. Rummukainen and M.E. Shaposhnikov, *Is there a hot electroweak phase transition at  $m_H \gtrsim m_W$ ?*, *Phys. Rev. Lett.* **77** (1996) 2887 [hep-ph/9605288].
- [63] M. Laine and K. Rummukainen, *What's new with the electroweak phase transition?*, *Nucl. Phys. B Proc. Suppl.* **73** (1999) 180 [hep-lat/9809045].
- [64] M. Laine, *Electroweak phase transition beyond the standard model*, in *4th International Conference on Strong and Electroweak Matter*, pp. 58–69, 6, 2000, DOI [hep-ph/0010275].
- [65] W. Buchmüller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys.* **B268** (1986) 621.
- [66] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [1008.4884].
- [67] J. de Blas, M. Chala, M. Perez-Victoria and J. Santiago, *Observable Effects of General New Scalar Particles*, *JHEP* **04** (2015) 078 [1412.8480].
- [68] D. Marzocca et al., *BSM Benchmarks for Effective Field Theories in Higgs and Electroweak Physics*, (2020) [2009.01249].
- [69] C. Grojean, G. Servant and J.D. Wells, *First-order electroweak phase transition in the standard model with a low cutoff*, *Phys. Rev. D* **71** (2005) 036001 [hep-ph/0407019].
- [70] D. Bödeker, L. Fromme, S.J. Huber and M. Seniuch, *The Baryon asymmetry in the standard model with a low cut-off*, *JHEP* **02** (2005) 026 [hep-ph/0412366].
- [71] C. Delaunay, C. Grojean and J.D. Wells, *Dynamics of Non-renormalizable Electroweak Symmetry Breaking*, *JHEP* **04** (2008) 029 [0711.2511].
- [72] R.-G. Cai, M. Sasaki and S.-J. Wang, *The gravitational waves from the first-order phase transition with a dimension-six operator*, *JCAP* **08** (2017) 004 [1707.03001].
- [73] M. Chala, C. Krause and G. Nardini, *Signals of the electroweak phase transition at colliders and gravitational wave observatories*, *JHEP* **07** (2018) 062 [1802.02168].
- [74] S. Borsanyi, Z. Fodor, J.N. Guenther, R. Kara, S.D. Katz, P. Parotto et al., *QCD Crossover at Finite Chemical Potential from Lattice Simulations*, *Phys. Rev. Lett.* **125** (2020) 052001 [2002.02821].

- [75] R.D. Pisarski and F. Wilczek, *Remarks on the Chiral Phase Transition in Chromodynamics*, *Phys. Rev. D* **29** (1984) 338.
- [76] D. Croon, R. Houtz and V. Sanz, *Dynamical Axions and Gravitational Waves*, *JHEP* **07** (2019) 146 [[1904.10967](#)].
- [77] P. Bak, S. Krinsky and D. Mukamel, *First-Order Transitions, Symmetry, and the epsilon Expansion*, *Phys. Rev. Lett.* **36** (1976) 52.
- [78] Y. Iwasaki, K. Kanaya, S. Sakai and T. Yoshie, *Chiral phase transition in lattice QCD with Wilson quarks*, *Z. Phys. C* **71** (1996) 337 [[hep-lat/9504019](#)].
- [79] F. Cuteri, O. Philipsen and A. Sciarra, *On the order of the QCD chiral phase transition for different numbers of quark flavours*, *JHEP* **11** (2021) 141 [[2107.12739](#)].
- [80] HOTQCD collaboration, *Chiral crossover in QCD at zero and non-zero chemical potentials*, *Phys. Lett. B* **795** (2019) 15 [[1812.08235](#)].
- [81] Y. Aoki, S. Borsanyi, S. Durr, Z. Fodor, S.D. Katz, S. Krieg et al., *The QCD transition temperature: results with physical masses in the continuum limit II.*, *JHEP* **06** (2009) 088 [[0903.4155](#)].
- [82] Y. Aoki, G. Endrodi, Z. Fodor, S.D. Katz and K.K. Szabo, *The Order of the quantum chromodynamics transition predicted by the standard model of particle physics*, *Nature* **443** (2006) 675 [[hep-lat/0611014](#)].
- [83] S. Ipek and T.M.P. Tait, *Early Cosmological Period of QCD Confinement*, *Phys. Rev. Lett.* **122** (2019) 112001 [[1811.00559](#)].
- [84] D. Croon, J.N. Howard, S. Ipek and T.M.P. Tait, *QCD baryogenesis*, *Phys. Rev. D* **101** (2020) 055042 [[1911.01432](#)].
- [85] D. Berger, S. Ipek, T.M.P. Tait and M. Waterbury, *Dark Matter Freeze Out during an Early Cosmological Period of QCD Confinement*, *JHEP* **07** (2020) 192 [[2004.06727](#)].
- [86] B. Svetitsky and L.G. Yaffe, *Critical Behavior at Finite Temperature Confinement Transitions*, *Nucl. Phys. B* **210** (1982) 423.
- [87] O. Kaczmarek, F. Karsch, E. Laermann and M. Lutgemeier, *Heavy quark potentials in quenched QCD at high temperature*, *Phys. Rev. D* **62** (2000) 034021 [[hep-lat/9908010](#)].
- [88] C. Alexandrou, A. Borici, A. Feo, P. de Forcrand, A. Galli, F. Jegerlehner et al., *The Deconfinement phase transition in one flavor QCD*, *Phys. Rev. D* **60** (1999) 034504 [[hep-lat/9811028](#)].
- [89] WHOT-QCD collaboration, *Phase structure of finite temperature QCD in the heavy quark region*, *Phys. Rev. D* **84** (2011) 054502 [[1106.0974](#)].



- [90] D. Croon, V. Sanz and G. White, *Model Discrimination in Gravitational Wave spectra from Dark Phase Transitions*, *JHEP* **08** (2018) 203 [[1806.02332](#)].
- [91] B.D. Fields, K.A. Olive, T.-H. Yeh and C. Young, *Big-Bang Nucleosynthesis after Planck*, *JCAP* **03** (2020) 010 [[1912.01132](#)].
- [92] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [[1807.06209](#)].
- [93] A. Riotto and M. Trodden, *Recent progress in baryogenesis*, *Ann. Rev. Nucl. Part. Sci.* **49** (1999) 35 [[hep-ph/9901362](#)].
- [94] M. Dine and A. Kusenko, *The Origin of the matter - antimatter asymmetry*, *Rev. Mod. Phys.* **76** (2003) 1 [[hep-ph/0303065](#)].
- [95] J.M. Cline, *TASI Lectures on Early Universe Cosmology: Inflation, Baryogenesis and Dark Matter*, *PoS TASI2018* (2019) 001 [[1807.08749](#)].
- [96] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation*, *Phys. Rev. Lett.* **55** (1985) 1039.
- [97] D.E. Morrissey and M.J. Ramsey-Musolf, *Electroweak baryogenesis*, *New J. Phys.* **14** (2012) 125003 [[1206.2942](#)].
- [98] K. Kainulainen, T. Prokopec, M.G. Schmidt and S. Weinstock, *First principle derivation of semiclassical force for electroweak baryogenesis*, *JHEP* **06** (2001) 031 [[hep-ph/0105295](#)].
- [99] T. Prokopec, M.G. Schmidt and S. Weinstock, *Transport equations for chiral fermions to order  $\hbar$  and electroweak baryogenesis. Part I*, *Annals Phys.* **314** (2004) 208 [[hep-ph/0312110](#)].
- [100] T. Prokopec, M.G. Schmidt and S. Weinstock, *Transport equations for chiral fermions to order  $\hbar$  and electroweak baryogenesis. Part II*, *Annals Phys.* **314** (2004) 267 [[hep-ph/0406140](#)].
- [101] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, *Phys. Rev. D* **51** (1995) 379 [[hep-ph/9404302](#)].
- [102] P. Huet and A.E. Nelson, *Electroweak baryogenesis in supersymmetric models*, *Phys. Rev. D* **53** (1996) 4578 [[hep-ph/9506477](#)].
- [103] A. Riotto, *Towards a nonequilibrium quantum field theory approach to electroweak baryogenesis*, *Phys. Rev. D* **53** (1996) 5834 [[hep-ph/9510271](#)].
- [104] C. Lee, V. Cirigliano and M.J. Ramsey-Musolf, *Resonant relaxation in electroweak baryogenesis*, *Phys. Rev. D* **71** (2005) 075010 [[hep-ph/0412354](#)].
- [105] M. Postma and J. van de Vis, *Source terms for electroweak baryogenesis in the vev-insertion approximation beyond leading order*, *JHEP* **02** (2020) 090 [[1910.11794](#)].



- [106] J.M. Cline and K. Kainulainen, *Electroweak baryogenesis at high bubble wall velocities*, *Phys. Rev. D* **101** (2020) 063525 [2001.00568].
- [107] M. Postma, *A different perspective on the vev insertion approximation for electroweak baryogenesis*, *JHEP* **09** (2021) 055 [2107.05971].
- [108] J.M. Cline and B. Laurent, *Electroweak baryogenesis from light fermion sources: A critical study*, *Phys. Rev. D* **104** (2021) 083507 [2108.04249].
- [109] K. Kainulainen, *CP-violating transport theory for electroweak baryogenesis with thermal corrections*, *JCAP* **11** (2021) 042 [2108.08336].
- [110] M. Postma, J. van de Vis and G. White, *Resummation and cancellation of the VIA source in electroweak baryogenesis*, *JHEP* **12** (2022) 121 [2206.01120].
- [111] A.G. Cohen and D.B. Kaplan, *Thermodynamic Generation of the Baryon Asymmetry*, *Phys. Lett. B* **199** (1987) 251.
- [112] L. Sagunski, P. Schicho and D. Schmitt, *Supercool exit: Gravitational waves from QCD-triggered conformal symmetry breaking*, *Phys. Rev. D* **107** (2023) 123512 [2303.02450].
- [113] I. Affleck and M. Dine, *A new mechanism for baryogenesis*, *Nuclear Physics B* **249** (1985) 361.
- [114] J.A. Harvey and M.S. Turner, *Cosmological baryon and lepton number in the presence of electroweak fermion number violation*, *Phys. Rev. D* **42** (1990) 3344.
- [115] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, *Phys. Lett. B* **174** (1986) 45.
- [116] M.A. Luty, *Baryogenesis via leptogenesis*, *Phys. Rev. D* **45** (1992) 455.
- [117] W. Buchmuller, P. Di Bari and M. Plumacher, *Leptogenesis for pedestrians*, *Annals Phys.* **315** (2005) 305 [hep-ph/0401240].
- [118] S. Davidson, E. Nardi and Y. Nir, *Leptogenesis*, *Phys. Rept.* **466** (2008) 105 [0802.2962].
- [119] E.E. Flanagan and S.A. Hughes, *The Basics of gravitational wave theory*, *New J. Phys.* **7** (2005) 204 [gr-qc/0501041].
- [120] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*, Oxford Master Series in Physics, Oxford University Press (2007).
- [121] E. Witten, *Cosmic Separation of Phases*, *Phys. Rev. D* **30** (1984) 272.
- [122] C.J. Hogan, *Gravitational radiation from cosmological phase transitions*, *Mon. Not. Roy. Astron. Soc.* **218** (1986) 629.

- [123] R. Jinno and M. Takimoto, *Gravitational waves from bubble collisions: An analytic derivation*, *Phys. Rev. D* **95** (2017) 024009 [[1605.01403](#)].
- [124] M. Hindmarsh, *Sound shell model for acoustic gravitational wave production at a first-order phase transition in the early Universe*, *Phys. Rev. Lett.* **120** (2018) 071301 [[1608.04735](#)].
- [125] M. Hindmarsh and M. Hijazi, *Gravitational waves from first order cosmological phase transitions in the Sound Shell Model*, *JCAP* **12** (2019) 062 [[1909.10040](#)].
- [126] M. Hindmarsh, S.J. Huber, K. Rummukainen and D.J. Weir, *Gravitational waves from the sound of a first order phase transition*, *Phys. Rev. Lett.* **112** (2014) 041301 [[1304.2433](#)].
- [127] M. Hindmarsh, S.J. Huber, K. Rummukainen and D.J. Weir, *Numerical simulations of acoustically generated gravitational waves at a first order phase transition*, *Phys. Rev. D* **92** (2015) 123009 [[1504.03291](#)].
- [128] M. Hindmarsh, S.J. Huber, K. Rummukainen and D.J. Weir, *Shape of the acoustic gravitational wave power spectrum from a first order phase transition*, *Phys. Rev. D* **96** (2017) 103520 [[1704.05871](#)].
- [129] K. Enqvist, J. Ignatius, K. Kajantie and K. Rummukainen, *Nucleation and bubble growth in a first order cosmological electroweak phase transition*, *Phys. Rev. D* **45** (1992) 3415.
- [130] A.H. Guth and E.J. Weinberg, *Cosmological Consequences of a First Order Phase Transition in the SU(5) Grand Unified Model*, *Phys. Rev. D* **23** (1981) 876.
- [131] LISA COSMOLOGY WORKING GROUP collaboration, *Cosmology with the Laser Interferometer Space Antenna*, [2204.05434](#).
- [132] NANOGrav collaboration, *The NANOGrav 15-year Data Set: Evidence for a Gravitational-Wave Background*, [2306.16213](#).
- [133] J. Antoniadis et al., *The second data release from the European Pulsar Timing Array III. Search for gravitational wave signals*, [2306.16214](#).
- [134] D.J. Reardon et al., *Search for an isotropic gravitational-wave background with the Parkes Pulsar Timing Array*, [2306.16215](#).
- [135] H. Xu et al., *Searching for the nano-Hertz stochastic gravitational wave background with the Chinese Pulsar Timing Array Data Release I*, [2306.16216](#).
- [136] D.J. Weir, *Gravitational waves from a first order electroweak phase transition: a brief review*, *Phil. Trans. Roy. Soc. Lond. A* **376** (2018) 20170126 [[1705.01783](#)].
- [137] J. Ellis, M. Lewicki, J.M. No and V. Vaskonen, *Gravitational wave energy budget in strongly supercooled phase transitions*, *JCAP* **06** (2019) 024 [[1903.09642](#)].
- [138] H.-K. Guo, K. Sinha, D. Vagie and G. White, *Phase Transitions in an Expanding Universe: Stochastic Gravitational Waves in Standard and Non-Standard Histories*, *JCAP* **01** (2021) 001 [[2007.08537](#)].

- [139] C. Caprini, R. Durrer, T. Konstandin and G. Servant, *General Properties of the Gravitational Wave Spectrum from Phase Transitions*, *Phys. Rev. D* **79** (2009) 083519 [[0901.1661](#)].