Review

Review of cosmic phase transitions: their significance and experimental signatures

Anupam Mazumdar\textsuperscript{1,3} and Graham White\textsuperscript{2}

\textsuperscript{1} Van Swinderen Institute, University of Groningen, 9747 AG, Groningen, The Netherlands
\textsuperscript{2} TRIUMF, 4004 Wesbrook Mall, Vancouver, British Columbia V6T 2A3, Canada

E-mail: anupam.mazumdar@rug.nl

Received 18 July 2017, revised 8 March 2019
Accepted for publication 3 May 2019
Published 25 June 2019

Abstract

The study of cosmic phase transitions are of central interest in modern cosmology. In the standard model of cosmology the Universe begins in a very hot state, right after the end of inflation via the process of reheating/preheating, and cools to its present temperature as the Universe expands. Both new and existing physics at any scale can be responsible for catalyzing either first, second or cross over phase transition, which could be either thermal or non-thermal with a potential observable imprints. Thus this field prompts a rich dialogue between gravity, particle physics and cosmology. It is all but certain that at least two cosmic phase transitions have occurred—the electroweak and the QCD phase transitions. The focus of this review will be primarily on phase transitions above such scales, We review different types of phase transitions that can appear in our cosmic history, and their applications and experimental signatures in particular in the context of exciting gravitational waves, which could be potentially be constrained by LIGO/VIRGO, Kagra, LISA, and Decigo.

Keywords: cosmological phase transitions, gravitational waves, cosmology

(Some figures may appear in colour only in the online journal)

1. Introduction

Cosmic phase transitions [1–4] are macroscopic cosmic events so dramatic that they are capable of leaving imprint via non-adiabatic vacuum fluctuations and creation of particles [5–7], formation of defects [4, 8–10], generation of magnetic field [11, 12], generation of baryonic asymmetry [13–17], and the gravitational wave background [18], and yet their properties are determined by particle physics. Indeed all the macroscopic properties such as the order of the phase transition [1, 2], the length of the phase transition [1, 2], the latent heat [20], and the amount of supercooling that occurs [19–21] are all controlled by the quantum properties such as effective mass, effective coupling, finite temperature effects whose masses are near the temperature scale at which the phase transition occurs, for an extensive reviews on these topics, see [17].

Furthermore, the particle physics responsible for a cosmic phase transition can potentially occur at any scale: from the QCD scale \( \sim 10^{-1} \) GeV [22] right up to the GUT scale \( \sim 10^{15} \) GeV [23]. Thus, any cosmic phase transition might shed light on particle physics occurring at scales that are potentially out of reach to both present and future particle colliders. Besides, thermal phase transitions, non-thermal phase transitions can also occur. Usually, they occur after the end of inflation, prompted by the inflationary sector or some hidden sector [5–7]. Such phase transitions do not require initial conditions to be set by thermal initial conditions.

It is likely that at least two transitions have occurred: the QCD transition between a quark gluon plasma and a hadron gas [22], and the electroweak transition [24] in which
electroweak symmetry was spontaneously broken allowing for the standard model (SM) particles to acquire gauge invariant masses [25, 26]. The latter phase transition may be involved in generating baryon asymmetry of the Universe via the electroweak (EW) baryogenesis mechanism [13, 16, 27–29]. Both of these phase transitions can leave an observable signature in the gravitational wave background [30–32] as well as affecting both the cold and hot dark matter backgrounds [33]. Other phase transitions are also possible: examples of which include the spontaneously breaking of gauge symmetry of Grand Unified Theory (GUT) giving rise to inflation [34–36], and the production of dark matter through a strongly first order phase transition [37].

The nature of these phase transitions not only affect any observable relic footprints they might leave, but also their utility. For example, if a relic gravitational background suggests that the electroweak phase transition occurred through bubble nucleation then that could be a sign that the baryon asymmetry of the Universe was produced during this phase transition. First order phase transitions also generate primordial magnetic field during the turbulence phase of the plasma and bubble collision [38], see a very nice review [12], and in some cases they may generate domain walls and strings [4, 5, 39, 40], as it happens in the case of Next to Minimal Supersymmetric standard model (NMSSM) [9, 41].

Furthermore the ingredients behind cosmic phase transitions have many discovery avenues. Observation of proton decay would give strong evidence of grand unification [42, 43] which is an ingredient in a GUT phase transition, discovery of new scalar particles would give more information about the electroweak phase transition or new massive gauge bosons could indicate a dark phase transition. Finally the principles behind cosmic phase transitions can in principle be tested in condensed matter systems which can imitate cosmological situations for a given Lagrangian [44].

2. Basic cosmological overview

Let us briefly summarize the early Universe cosmology in chronological order. How the Universe began remains a profound question, for which we do not have direct experimental evidence yet. Nonetheless, we can speculate based on sound physical arguments and the observations confirmed by the detection of cosmic microwave background (CMB) radiation [45, 46].

2.1. How did the Universe begin?

Einstein’s theory of gravity (GR) is extremely successful in the infrared (IR) matching of all possible observables [47], including the recent discovery of gravitational waves from mergers of two blackholes [48], and binary neutron star mergers [49]. However at short distances and small time scales, i.e. in the ultraviolet (UV), GR has pathologies, besides being a non-renormalizable theory, GR introduces cosmological and blackhole singularities, see [50], and in some cases naked singularities, see [51]. In GR, our Universe has a distinct starting point, a singular spacetime—as long as all the standard energy conditions are always satisfied, i.e. strong, weak, and null energy conditions, see [50]. It is possible to address the cosmological singularity problem without violating the matter energy conditions by weakening the gravitational interaction in the UV. This can happen in ghost free infinite derivative gravity inspired from string field theory [52, 53]. There could be two consequences for such study; one could be a realization of a non-singular bounce [54, 55], and the other scenario would be that Universe could be frozen in time in the UV, such that the Universe becomes conformal as \( t \to 0 \) [56]. Bouncing cosmologies and cosmological density perturbations have been reviewed in this nice review [57, 58]. There is a strong indication that this non-singular initial phase of the Universe has a key role to play towards understanding the subsequent phases of the Universe such as cosmic inflation, horizon, homogeneity and isotropy of the Universe, to create appropriate initial conditions for the Universe. In this review we will not discuss any further the very genesis of the Universe, we will merely assume that Universe is homogeneous and isotropic from the very beginning. We will discuss inflationary cosmology very briefly, but inflationary cosmology has its own limitations, when it comes to explaining the initial condition problem—it cannot solve or address the initial singularity problem [59].

2.2. Cosmic inflation to reheating

2.2.1. Primordial inflation. A phase of primordial inflation addresses some of the key challenges of the hot big bang cosmology, such as the flatness and the horizon problems, i.e. generating the large scale structures on roughly 4000 Mpc scale, and the age of the Universe \( \sim 13.8 \) Gyr [34–36, 60–62]. Cosmic inflation is a very successful paradigm, and we know lot more about inflationary predictions than the pre-inflationary phase of the Universe, such as big bounce or freeze-in phase of the Universe. Within the inflationary paradigm, with the help of Planck satellite we could see only the eight e-folding of inflation around the pivot scale 0.05 Mpc \(^{-1} \) [46].

One of the key predictions of inflation remains that of stretching the long wavelength quantum fluctuations on dark matter on scales larger than the size of the Hubble patch during inflation [63–66], in order to match the current observations, in the temperature anisotropy in CMB, which has been observed quite precisely by number of space based missions, starting from COBE [67], WMAP [45], and now by Planck [46]. The other key predictions for inflation is that by inflating the scale factor of the Universe, it makes the spatial curvature of the Universe flatter and flatter. As a consequence the quantum fluctuations in the scalar field becomes very close to the Gaussian random fluctuations, with almost scale invariant perturbations. Indeed, the epoch of inflation has to come to an end, and this yields a slight departure from the scale invariant perturbations, which we we have inferred from flat-\( \Lambda \)CDM paradigm, where flat signifies the spatial flatness of the Universe, \( \Omega_k \), see [36]. Note that \( \Lambda \) stands for late cosmological constant—the fact that the current Universe appears to be accelerating, as it is evident from dark energy surveys
[68], and CDM stands for the cold dark matter paradigm, which is non-relativistic, potentially non-baryonic form of matter. The CDM is required to form the very first stars and structures in the Universe, effectively with zero equation of state parameter, i.e., pressure-less fluid. There are compelling paradigms other than $\Lambda$CDM, such as modified Newtonian gravity [69], which might as well lead to structure formation [70] and explain some of the outstanding observations of the baryonic physics in the neighborhood of our galaxy, but here in this review we will limit ourselves to CDM paradigm, other variants of CDM are warm dark matter (WDM) [71], or fuzzy dark matter scenarios [72], which can ameliorate some of the problems related to CDM in the structure formation simulations versus observations, see [73].

Cosmic inflation within GR can be explained in a very simple manner. Inflation happens along a flat direction with a non-negligible slope of the potential, given by the order parameter, $\phi$, as an inflaton, and its potential, $V(\phi)$. Inflation could last eternally in future due to stochastic fluctuations of the scalar field in almost deSitter background [74, 75], nevertheless, in a given observable Hubble patch, it has to come to an end. Note that inflation cannot be past eternal unless the singularity problem can be resolved [76]. The current data, from Planck [77], at best can probe the second derivative of the potential and not beyond that. The inflationary predictions [77] are compatible with slow-roll inflation, which assumes that the potential dominates over the kinetic energy $\dot{\phi}^2 \ll V(\phi)$, and $\ddot{\phi} \ll V'(\phi)$, where dots w.r.t physical time, $t$, therefore the Friedmann and the Klein–Gordon equations are approximated as:

$$H^2 \approx \frac{V(\phi)}{3M^2_P}, \quad 3H\dot{\phi} \approx -V'(\phi), \quad (1)$$

where prime denotes derivative with respect to $\phi$. The slow-roll conditions, which parametrize the shape of the potential, are given by:

$$\epsilon(\phi) \equiv \frac{M^2_P}{2} \left( \frac{V'}{V} \right)^2 \ll 1, \quad |\eta(\phi)| \equiv \frac{M^2_P}{2} \left| \frac{V''}{V} \right| \ll 1. \quad (2)$$

Note that the slow-roll conditions are violated when $\epsilon \sim 1$, and $\eta \sim 1$, which marks the end of inflation. The number of e-foldings can be defined between, $t$, and the end of inflation, $t_{\text{end}}$:

$$N \equiv \ln \frac{a(t_{\text{end}})}{a(t)} = \int_{t_{\text{end}}}^{t_{\text{end}}} H dt \approx \frac{1}{M^2_P} \int_{\phi_{\text{end}}}^{\phi} V^{1/2} \phi' d\phi. \quad (3)$$

where $\phi_{\text{end}}$ is defined by $\epsilon(\phi_{\text{end}}) \sim 1$, provided inflation comes to an end via a violation of the slow-roll conditions. The number of e-foldings can be related to the Hubble crossing mode $k = a\dot{\phi}H$ by comparing with the present Hubble length $a_0H_0$. The final result is [78, 79]

$$N(k) = 62 - \ln \frac{k}{a_0H_0} - \ln \left( \frac{10^{16} \text{GeV}}{V_{k}^{1/4}} \right) + \ln \frac{V_{k}^{1/4}}{V_{\text{end}}^{1/4}} - \frac{1}{3} \ln \frac{V_{\text{end}}^{1/4}}{\rho_k^{1/4}}, \quad (4)$$

where the subscript $\text{end}$ (R) refer to the end of inflation (end of reheating). Today’s Hubble length would correspond to $N_R \equiv N(k = a_0H_0)$ number of e-foldings, whose actual value would depend on the equation of state, i.e., $\omega = p/\rho$ ($\rho$ denotes the pressure, $\rho$ denotes the energy density), from the end of inflation to radiation and matter dominated epochs. A high scale inflation with a prompt reheating with relativistic species would yield approximately, $N_0 \approx 50–60$. A significant modification in the number of e-foldings can take place. If the scale of inflation is low, and if the reheat temperature is as low as that of 1 MeV, roughly the temperature before the Big Bang Nucleosynthesis (BBN), for a review [80], the number of e-foldings to explain the horizon and the flatness problem could be as low as $\sim 25$, see [81–85].

For a single field slow-roll inflation there exists a late time attractor behaviour, such that the evolution of a scalar field after sufficient e-foldings become independent of the initial conditions [86, 87]. This particular initial condition is solely related to the homogeneous inflaton and its initial velocity, and has nothing to do with the initial homogeneity and isotropy of the Universe. Inflation as such does not solve these problems, there are obstructions to that within GR due to focusing theorems due to Raychaudhury [88], and Hawking–Penrose singularity theorems [50, 89]. The initial patch of the Universe should be homogeneous on scales larger than the inflating Hubble patch [11], similarly, in order to inflate such a patch, the patch should be already isotropic. In this regard inflation assumes homogeneity and isotropy of spacetime within GR [90, 91].

2.3. Primordial perturbations

2.3.1. Scalar perturbations. The small inhomogeneities in the inflaton field can be recast as, $\phi(\vec{x}, t) = \bar{\phi}(t) + \delta \phi(\vec{x}, t)$, where $\delta \phi \ll \bar{\phi}$, is the inflaton perturbations in the background metric. During inflation $\delta \phi$ are stretched outside the Hubble patch, because $m^2 \sim V' < H^2$. These fluctuations can then be tracked from a sub-Hubble to that of a super-Hubble length scales right when the wave numbers have crossed the Hubble patch, these fluctuations are random Gaussian, and can be given by:

$$\langle |\delta \phi_k|^2 \rangle = \langle (H(t_*)^2/2k^4), \quad (5)$$

where $t_*$ denotes the instant of Hubble crossing. One can define a power spectrum for the perturbed scalar field

$$P_{\phi}(k) = \frac{k^3}{2\pi^2} \langle |\delta \phi_k|^2 \rangle = \left[ \frac{H(t_*)}{2\pi} \right]^2 \left[ \frac{H}{2\pi} \right] \left. \right|_{k = aH} \quad . \quad (6)$$

Note that the phase of $\delta \phi_k$ can be arbitrary, and therefore, inflation has generated a Gaussian perturbation. We can calculate the power spectrum for the metric perturbations, this is what we observe in the CMB, translated into temperature anisotropy. Since, the separation between the background metric and a perturbed metric is not unique, a choice of gauge, or a choice of a particular coordinate system becomes
necessary to simplify the metric perturbations. One particular choice would be to fix the gauge where the non-relativistic limit of the full perturbed Einstein equation can be recast as a Poisson equation with a Newtonian gravitational potential, Φ. The induced metric perturbations can be written in GR as, e.g. [63]:

\[ ds^2 = a^2(\tau) \left[ (1 + 2\Phi)dr^2 - (1 - 2\Psi)d\vec{x}'d\vec{x}'^2 \right], \]  

(7)

and when the spatial part of the energy momentum tensor is diagonal, i.e. \( \delta T^i_j = \delta_j^i \), it follows that \( \Phi = \Psi \). For a critical density Universe, i.e. for a flat Universe, \( \delta_k \equiv \delta_k^i \), the power spectrum is given by:

\[ \delta_k^2 \equiv \frac{4}{9} P_\Phi(k) = \frac{4}{25} \frac{9}{(\frac{H}{\phi})^2} \]  

(8)

where the right hand side can be evaluated at the time of horizon exit \( k = aH \). It is convenient for the observers to express the perturbations in a different gauge, known as the comoving synchronous gauge, where on the comoving hypersurface the energy flux vanishes, and the amplitude is denoted by \( \bar{\zeta} \) [63, 92]. The comoving curvature perturbation \( \bar{\zeta} \) is a conserved quantity for the super-Hubble modes, \( k \to 0 \), and \( \bar{\zeta} = -(5/3)\Phi_k \). Therefore, \( \delta_k \) can also be expressed in terms of the curvature perturbations [93] \( \delta_k = \frac{2}{5} \left( \frac{\dot{\Phi}}{\dot{\phi}} \right)^2 \bar{\zeta} \). The corresponding power spectrum \( \delta_k^2 = (4/25)P_\zeta(k) = (4/25)(H/\dot{\phi})^2(H/2\pi)^2 \). With the help of the slow-roll equation \( 3H\dot{\phi} = -V' \), and the critical density formula \( 3H^2M_p^2 = V \), one obtains

\[ \delta_k^2 \approx \frac{1}{75\pi^2M_p^6} \frac{V^3}{V''} = \frac{1}{150\pi^2M_p^6} \frac{V}{\epsilon}, \quad P_\zeta(k) = \frac{1}{24\pi^2M_p^6} \frac{V}{\epsilon}, \]  

(9)

where we have used the slow-roll parameter \( \epsilon \equiv \langle M_p^2/2 \rangle (V'/V)^2 \). If we assume that the primordial spectrum can be approximated by a power law, see [77]

\[ P_\zeta(k) \approx (3.044 \pm 0.014) \times 10^{-10} \left( \frac{k}{k_0} \right)^{n_s-1}, \]  

(10)

where \( n_s \) is called the spectral index (or spectral tilt), the reference scale is: \( k_0 = 7.5a_0H_0 \sim 0.002 \text{Mpc}^{-1} \), and the error bar on the normalization is given by the characterization of polarization at low and high multipoles, Planck temperature, polarization, and lensing data yields at 68% CL [77]

\[ n_s(k_0) = 0.9649 \pm 0.0042. \]  

(11)

In the slow-roll approximation, this tilt can be expressed in terms of the slow-roll parameters and at first order:

\[ n_s - 1 = -6\epsilon + 2\eta + O(\epsilon^2, \eta^2, \epsilon\eta, \epsilon\xi^2), \]  

(12)

where

\[ \xi^2 \equiv M_p^4 \frac{V'(d^3V/d\phi^3)}{V^2}, \quad \sigma^3 \equiv M_p^4 \frac{V''(d^4V/d\phi^4)}{V^3}. \]  

(13)

The running of these parameters are given by [86]. Since the slow-roll inflation requires that \( \epsilon \ll 1, |\eta| \ll 1 \), therefore naturally predicts small variation in the spectral index within \( \Delta \ln k \approx 1 \) [94]

\[ \frac{dn(k)}{d\ln k} = -16\epsilon\eta + 24\epsilon^2 + 2\xi^2. \]  

(14)

There is no evidence of scale dependence of \( n_s \) has been found by the latest Planck data [77].

2.3.2. Tensor perturbations. Like scalar field induced metric perturbations during inflation, we would also expect pure stochastic gravitational waves [95–99]. For reviews on gravitational waves, see [63, 100]. The gravitational wave perturbations are described by a line element \( ds^2 + 2\delta x^i \), where

\[ ds^2 = a^2(\tau)(dr^2 - d\vec{x}'d\vec{x}'^i), \quad \delta x^i = -a^2(\tau)h_i d\vec{x}'^i. \]  

(15)

The 3-tensor \( h_i \) is symmetric, traceless \( \delta^i_j h_{ij} = 0 \), and divergenceless \( \nabla_i h_{ij} = 0 \) (\( \nabla_i \) is a covariant derivative), and also gauge and conformally invariant. Massless spin 2 gravitons have two transverse degrees of freedom (d.o.f) For the Einstein’s GR, the gravitational wave equation of motion follows that of a massless Klein Gordon equation [63]. Especially, for a flat Universe

\[ \ddot{h}_i + 3H \dot{h}_i + (k^2/\epsilon^2) h_i = 0. \]  

(16)

As any massless field, the gravitational waves also feel the quantum fluctuations in an expanding background. The spectrum mimics that of equation (6)

\[ P_{\text{grav}}(k) = \frac{2}{M_p^2} \left( \frac{H}{2\pi} \right)^2 \left| k = aH \right. \]  

(17)

The corresponding spectral index can be expanded in terms of the slow-roll parameters at first order as

\[ r \equiv \frac{P_{\text{grav}}}{P_\zeta} = 16\epsilon, \quad n_t = \frac{d\ln P_{\text{grav}}(k)}{d\ln k} \approx -2\epsilon. \]  

(18)

Note that the tensor spectral index is negative, in some sense gravitational waves spectrum is solely determined by the the Hubble expansion rate during inflation and the initial vacuum condition. Relaxing the initial vacuum condition may lead to different predictions in the value of tensor-to-scalar ratio, \( r \), see [101, 102]. A classical initial condition can also produce \( r \), albeit the magnitude will be very tiny [101]. So, non detection of primordial gravitational waves does not confirm the quantum nature of gravitons in CMB base experiments. The latest constraint on the tensor to scalar ratio is given by the Planck upper limit 95%CL is \( r < 0.1 \), which is further tightened by BICEP2/Keck Array BK14 data \( r < 0.064 \) [77].

2.4. Reheating phase

There is no dearth of models of inflation which can potentially match the current set of observations in CMB, for a review see [103]. Within particle physics, typically inflation is assumed to be driven by SM gauge singlets, either driven by a single or multiple fields, such as hybrid [104], or infinitely many, assisted inflation [105, 106]. However, well motivated particle physics models are SM driven Higgs inflation [107], which requires unnatural coupling between the SM Higgs
and the Ricci scalar, which leaves the model very similar to the Starobinsky’s model of inflation of $R + \alpha R^2$ [62], after one-loop computation [108]. Note that in the original paper the relative sign difference was negative, and the motif was to obtain a bouncing Universe, but with a ghost in the spin-2 sector. Amongst the well-motivated models of inflation, Starobinsky’s model remains very minimal in content and driven purely within the gravitational sector. Its UV completion has been given in [52, 109, 110]. There are also other well-motivated models of inflation within particle physics, where the inflaton can be recognized by the supersymmetric partners of quarks and leptons, namely the gauge invariant combination of squarks and sleptons carrying the SM gauge charges within minimal supersymmetric standard model (MSSM) [111, 112], for a review of MSSM see [113, 114]. However, there has been no evidence of supersymmetric partners at the LHC, which has constrained the scale of inflation above the 3–4 TeV scale [115]. These models, i.e. MSSM and Higgs inflation, are also known as viable sector models of inflation, because the inflaton directly decays into visible sector d.o.f. All the Yukawas and gauge couplings are well known in these models. In this review we will not delve into model building of inflation any further, as this is not so relevant for this discussions below, but we will now divert to how the inflaton decays and thermalizes the Universe.

Typically, inflation ends via smooth phase transition as we had discussed, i.e. by violation of slow roll conditions. Inflation could also end with tachyonic instability [104], in some cases inflation can be driven by tachyons as well [116], or via tunneling from the inflationary vacuum to the SM type vacuum [36]. Unfortunately, in the Guth’s model the bubble never thermalizes, the bubble wall is still expanding in the deSitter background, and inside the bubble the SM Universe is super cooled. The SM vacuum needs to be thermalized by the collision of 2 or more bubbles, which never takes place if the bubble nucleation rate is smaller than the Hubble expansion rate of the deSitter. Depending on the gauge group, the phase where inflation comes to an end can create topological defects, see [3, 10, 117], such as cosmic strings, or domain walls, etc., however, these are very much model dependent. The phase after inflation leads to reheating and preheating. Non-perturbative preheating can give rise to 1st order phase transitions [5, 7], gravitational waves [118–126], magnetic fields [127], topological defects [6, 8, 9, 39], and non-topological solitons [128, 129].

Topological defects are another consequences of phase transitions, it was Kibble [2, 130] and Zurek [131] who independently postulated the formation of topological defects during cosmic phase transition. The topological defect is also known as a solitonic solution in quantum field theory which are homotopically distinct from the vacuum solution. In topology, two continuous functions if they can be continuously deformed into each other, then such deformations are known as homotopy between the two functions. If not, then they are homotopically distinct functions. The latter produces defects, kinks, cosmic strings, cosmic textures, and also Dirac monopoles, for a review see [4, 10]. Besides, topological solitons, there are also non-topological solutions in any interacting field theory where the boundary conditions at infinity are the same as that of the vacuum state [132], for example Q-ball, a detailed review of Q-balls, see [133, 134].

2.4.1. Perturbative decay and thermalization phase. During inflation, the Universe is cold and devoid of any thermal entropy. It is thus paramount to create a thermal bath, which can at least achieve local thermodynamical equilibrium (LTE), means that the species can be in thermal equilibrium as long as $\Gamma \geq H(t)$, where $\Gamma$ denotes the interaction rate and $H(t)$ is the Hubble expansion rate. Note that $\Gamma$ is solely determined by the particle physics interaction rate at a given energy, temperature, while $H(t)$ is the Hubble expansion rate of the Universe. For the species in LTE, the energy density, $\rho$, and the number density, $n$, for relativistic particles are given by [135]

$$\rho = \begin{cases} \pi^2/30 & (\text{Boson}) \\
(7/8) \pi^2/30 & (\text{Fermion})
\end{cases} T^4, \quad n = \begin{cases} (\zeta/3)/\pi^2 & T^3, \\
(3/4) (\zeta/3)/\pi^2 & T^3,
\end{cases}$$

(19)

where $T$ is the temperature of an ambient bath, shared by all the species present in the bath. Typically, the average energy of every species will be shared $\langle E \rangle \sim \rho^{1/4}$, and $n \sim \rho^{3/4}$ hold, with $\langle E \rangle = (\rho/n) \approx 3T$ being the average particle energy.

On the other hand, right after the inflaton has decayed, the energy density of the Universe is determined by the total decay width, $\Gamma_{\phi}$, of the inflaton to the relativistic species, $\rho \approx 3 (\Gamma_{\phi} M_{P})^2$. The ambient plasma has a thermal entropy, given by: $\langle E \rangle \approx m_{\phi} \gg \rho^{1/4}$. Then, the total number density is roughly given by $n \approx (\rho/m_{\phi}) \ll \rho^{3/4}$. Note that the initial energy density $\rho$ is always bounded below the energy density of the inflaton energy, i.e. $\rho \leq 3H^2 M_{P}^2$. Therefore, the decay products which creates the ambient plasma results in a very dilute plasma, the number density of the decayed products is very tiny, though the energy of the decayed particles can be as large as that of the inflaton mass, i.e. $m_{\phi}$. This suggests that the initial plasma is far from full thermal equilibrium initially [136–143].

Reaching full equilibrium requires re-distribution of the energy among different particles, kinetic equilibrium, as well as increasing the total number of particles, chemical equilibrium. Therefore, both the number-conserving and the number-violating reactions must be taken into account. Kinetic equilibrium can be achieved by $2 \rightarrow 2$ scatterings with gauge boson exchange in the $t$-channel [137, 138]. While the chemical equilibrium is achieved by changing the number of particles in the reheat plasma. It was recognized in [137], see also [138], that the most relevant processes are $2 \rightarrow 3$ scatterings with gauge-boson exchange in the $t$-channel. The latter is the inelastic scattering, when this become efficient, the scattering rate exceeds that of the Hubble expansion rate, and the number of particles also increases very rapidly [144], due to the fact that the produced gauge bosons subsequently participate in similar $2 \rightarrow 3$ scatterings. During this phase, decays of particles can also be considered, but they do not play an important role, they cannot increase the number of particles to
the required level. The full thermal LTE is established shortly after $2 \rightarrow 3$ scatterings become efficient. For this reason, to a very good approximation, one can use the rate for inelastic scatterings as a thermalization rate of the Universe $\Gamma_{\text{th}}$. If the inflaton decay products have SM like gauge interactions, i.e. relatively large gauge interactions, then the Universe reaches full thermal equilibrium quite quickly, the main reason is that the $2 \rightarrow 3$ scatterings with gauge boson exchange in the $t$-channel are indeed very efficient, see [137, 138, 145]. During this phase of thermalization one can produce massive long-lived or stable weakly interacting massive particles (WIMPs), or long lived feebly interacting massive particles (FIMPS) [146–153].

A rough estimate of the reheating temperature can be made. The release of the inflaton energy density into the thermal bath of relativistic particles take place when $H(a) = \sqrt{(1/3M_p^2\rho_D(a_i/a)^{3/2}) = \Gamma_{\phi}}$. The energy density of the thermal bath is determined by the reheating temperature $T_R$, or the temperature of the relativistic bath is given by:

$$T_R = \left( \frac{90}{\pi^2g_*} \right)^{1/4} \sqrt{T_{\text{th}} M_p} = 0.3 \left( \frac{200}{g_*} \right)^{1/4} \sqrt{T_{\text{th}} M_p},$$

where $g_*$ denotes the effective relativistic degrees of freedom in the plasma. However the inflaton might not decay instantaneously. In such a case there might already exist a thermal plasma of some relativistic species at a temperature higher than the reheating temperature already before the end of reheating [135]. If the inflaton decays with a rate $\Gamma_{\phi}$ then the instantaneous plasma temperature is found to be [135]:

$$T_{\text{inst}}(a) \sim \left( g_*^{-1/2} H T_{\text{th}} M_p^2 \right)^{1/4}.$$  

The temperature of the Universe reaches its maximum $T_{\text{max}}$ soon after the inflaton field starts oscillating around the minimum. Once the maximum temperature is reached, then $\rho_{\phi} \sim a^{-3/2}$, and $T \sim a^{-3/8}$ until reheating and thermalization is completely over. Thermalization is achieved when both kinetic and chemical equilibrium are reached. For a successful cosmology one needs to ask how the inflaton energy gets converted into the SM degrees of freedom. For large reheat temperatures, $T_R \sim 10^9$ GeV, the Universe could abundantly create thermal relics of unstable gravitinos with a mass of order 100–1000 GeV, which could spoil the success of BBN [154–159] (for the effects of lighter unstable relics see [160]). For extremely low reheat temperatures, i.e. $T_R \sim O(1–10)$ MeV, it becomes a great challenge to obtain matter-anti-matter asymmetry and the right abundance for the dark matter. Only a few particle physics scenarios can successfully create baryons and dark matter at such a low temperature, see for instance [161].

2.5. Non-thermal phase and reheating

The Universe after inflation could be reheated in a much violent fashion via non-perturbative, non-thermal way. The Universe in this epoch prior to the attaining LTE could be completely out of equilibrium. This could lead to rapid and efficient transfer of inflaton energy, the process is also dubbed as preheating. Indeed, preheating is model dependent, but in a wide class of inflationary models preheating criteria can be satisfied with ease. One of the key ingredients is that the inflaton couples to essentially massless field $\chi$, through interaction term like $\phi^2\chi^2$. The quantum modes of $\chi$ can then be excited during the inflaton oscillations via a parametric resonance [5, 139–142, 162], for a review see [163]. During preheating, fermions can also be excited, but their occupation number can not grow arbitrarily large due to Pauli blocking [164–171]. Also, one can excite the gauge fields which may have applications for cold electroweak baryogenesis [172–174], and magnetic fields as well [127]. The epoch of preheating has been performed on lattice, see [175–177].

2.5.1. Parametric resonance. Let us briefly discuss preheating in the simplest but most general setup. Let us consider the relevant renormalizable couplings between the inflaton $\phi$ and a scalar field $\chi$, for which the potential will be given by:

$$V = \frac{1}{2} m_\phi^2 \phi^2 + \frac{1}{2} m_\chi^2 \chi^2 + \sigma \phi \chi^2 + h^2 \phi^2 \chi^2 + \kappa \chi^4.$$  

where we have considered $\phi$ and $\chi$ to be real, and the kinetic terms are all canonical. Furthermore, $\phi$ is a gauge singlet inflaton. Preheating with non-canonical terms has been studied in [178]. Note that $\sigma$ has a (mass) dimension. The only scalar field in the SM is the Higgs doublet. During the preheating term like $\chi^4$ can be excited, but their occupation number can dominate at the beginning of inflaton oscillations (i.e. when the Hubble expansion rate is $H(t) \approx m_\phi$). Preheating typically occurs within a narrow window for $h \lesssim 10^{-4} < h < 10^{-3}$. The $h^2 \phi^2 \chi^2$ term also yields a quartic self-coupling for the inflaton at a one-loop level which is typically constrained by the temperature anisotropy of the CMB perturbations, i.e. $\kappa \lesssim 10^{-12}$. Neglecting the self interaction for $\chi$ field, the equation of motion for $\chi$ quanta is given by:

$$\dot{\chi}_k + \frac{3}{a} \dot{a} \chi_k + \left( \frac{k^2}{a^2} + m_\phi^2 + 2(\sigma \phi + h^2 \phi^2) \right) \chi_k = 0.$$  

It is assumed that the inflaton oscillations are homogenous, $\phi(t) = \phi(t) \sin(m_{\phi}t)$, where $\phi(t) \approx (M_p/\sqrt{3\pi m_{\phi}})$, for the inflaton mass $m_{\phi}$. The occupation number of the excited $\chi$ is given by:

$$n_k = \frac{\omega_k}{\omega_k} \left( \frac{\omega_k^2}{\omega_k} + |\chi_k|^2 \right) - \frac{1}{2}.$$  

There exists a possibility of a narrow resonance production of $\chi_k \propto \exp(i \Delta_k z)$, where $\Delta_k$ is set by the instability band $\Delta_k^G$ labeled by an integer $n$, and $z = m_{\phi} t$. quanta, see [139–142], when expansion of the Universe and the trilinear interaction are small. The resonance occurs for $k = 0.5 m_{\phi} (1 \pm q/2)$, where $\mu_k$ takes the maximum value $\mu_k = q/2$, where $q = g^2(\phi^2/4m_{\phi}^2)$.

\[ 6 \]
When the expansion rate of the Universe is taken into account, then the evolution of the oscillating inflaton field also modifies to a damped oscillator:

$$\phi(t) \simeq \frac{M_p}{\sqrt{3\pi}} \frac{\cos(mt)}{m},$$

(25)

where $t$ is the physical time. During this period the stochastic resonance becomes important [139], where there are resonance bands which keep shifting from stability to instability bands. The resonant particle production and re-scatterings of interacting quanta lead to the formation of a plasma consisting of both $\phi$ and $\chi$ quanta with typical energies $\sim 10^{-1} \left( hm_\phi M_p \right)^{1/2}$, see [139]. This plasma attains the kinetic equilibrium first, but the full thermal equilibrium, including both kinetic and chemical, is established over a much longer time scale [176, 177].

The occupation number of particles in the preheated plasma is $\gg 1$ (which is opposite to the situation after the perturbative decay). This implies that the number density of particles is larger than its value in full equilibrium, while the average energy of particles is smaller than the equilibrium value. It gives rise to large effective masses for particles which, right after preheating, is similar to their typical momenta [139].

The momentum dependent frequency, $\omega_k(t) = \sqrt{k^2 + m_\chi^2} + 2hqK$ violates the above condition when

$$k^2 + m_\chi^2 \leq (h^2 \phi m_\phi \dot{\phi})^{2/3} - h^2 \dot{\phi}^2.$$  

(27)

The maximal range of momenta for which particle production occurs corresponds to $\phi(t) = \phi_c$, where $\phi_c = 1/2$.

The maximal value of momentum for particles produced at that epoch can be estimated by $k_{\text{max}}^2 + m_\chi^2 \approx h m_\phi \dot{\phi}$. The resonance becomes efficient for $hm_\phi \dot{\phi} \geq 4m_\chi^2$. Thus, the inflaton oscillations may lead to a copious production of superheavy particles with $m_\chi > m$ if the amplitude of the field $\phi$ is large enough, $h \dot{\phi} > 4m_\chi^2/m_\phi$. Besides narrow and broad resonances, there are other variants of preheating, such as instant [180], and tachyonic preheating triggered via tachyonic instability, where at the classical level the zero mode develops exponential enhancement [181, 182].

2.5.2. Fermionic and gauge preheating. The Dirac equation (in conformal time $\eta$, where $d\eta = t/a(t)$) for a fermionic field is given by [165, 166]:

$$\left( \frac{i}{a} \gamma^\mu \partial_\mu + i \frac{3}{2} H_{\eta} - m(\eta) \right) \psi = 0,$$

(28)

where $m(\eta) = m_\psi + h(\phi(\eta))$, where $m_\psi$ is the bare mass of the fermion. $\alpha$ is the scale factor of the Universe, $H = a'/a^2$ is the Hubble rate and $t$ denotes derivative w.r.t. $\eta$. The particle density per physical volume $V \sim a^3$ at time $\eta$ is given by:

$$n(\eta) \equiv \langle 0 \mid N \mid 0 \rangle = \frac{1}{\pi^2 a^3} \int d^k k^2 |\beta_k|^2,$$

(29)

where $\alpha_k, \beta_k$ are the Bugolyubov’s coefficients satisfying: $|\alpha_k|^2 + |\beta_k|^2 = 1$. The occupation number of fermions created is thus given by $n_k = |\beta_k|^2$, and the above condition ensures that the Pauli limit $n_k < 1$ is respected. One important physical quantity is the scaling of the total energy

$$\rho_\psi \propto m_\psi N_\psi \propto q m_\psi^{1/2},$$

(30)

which is linear in $q = h^2 \dot{\phi}^2/m_\phi^2$ [165, 166, 168, 169]. Note that $m_\psi(\eta) \propto \eta^{1/2}$. Note that the SM fermions are chiral, if the inflaton is a SM gauge singlet, then it can only couple via dimension-5 operators, i.e.

$$\frac{\lambda}{M_p} \phi(H\phi)q_R,$$

(31)

where $\lambda \sim O(1)$, $H$ is the SM Higgs doublet and $q_Lq_R$ are the SU(2) doublet and the right handed SM fermions, respectively. As a result, preheating of SM fermions from a gauge singlet inflaton becomes less important due to weak coupling. In [166], it was shown that an inflaton coupling to the right handed neutrino, $h_0 \phi N:\nu$, where $N$ is the right handed neutrino, could induce non-thermal leptogenesis, where the right handed neutrinos were treated as gauge singlets. Similar arguments would hold for the inflaton coupling to the SM gauge bosons, where the inflaton can only couple via non-renormalizable operator, i.e.

$$\frac{\lambda}{M_p} \phi F_{\mu\nu} F^{\mu\nu},$$

(32)

where $\lambda \sim O(1)$. Therefore, exciting the SM gauge bosons and the SM fermions through parametric resonance of a gauge singlet inflaton is a daunting task. Inflaton would rather prefer perturbative decay. The only way one can excite SM fermions and gauge fields copiously, if they are directly excited by the oscillations of the SM Higgs boson. This can happen in low scale electroweak baryogenesis [172–174], or in the context of SM Higgs inflation [183]. During the Higgs oscillations the
SM degrees of freedom can be excited via parametric resonance, instant preheating and also via tachyonic preheating. All three phases of preheating are present. The other notable example is the MSSM inflation [184] where gluons and MSSM fermions are excited via instant preheating.

2.5.3. Fragmentation of the inflaton. An intriguing consequence of inflaton coupling to the fermions is the fragmentation of the inflaton as a condensate [128, 129, 185]. This leads to non-thermal phase where the inflaton condensate can fragment to form non-topological solitons, known as Q-balls. The Q-balls can evaporate from their surface, see for a review [133], therefore suppressing the reheating and thermalization time scale. Let us illustrate this idea by studying a simple scenario of an oscillating complex scalar field around its minimum. Typically, the fermionic loops (assuming that the fermions live in a larger representation than bosons) yields a Logarithmic correction to the inflaton mass. Similar corrections also arises within SUSY, in a gravity mediated scenarios [133]

$$V = m^2 |\Phi|^2 \left[ 1 - K \log \left( \frac{|\Phi|^2}{M^2} \right) \right],$$  \hspace{1cm} (33)

where the value of $K$ is determined by the Yukawa coupling $h$ with $K = -C(h^2/16\pi^2)$, where $C$ is some number. If $K < 0$, the inflaton condensate feels a negative pressure for field values $\phi \ll M$, we find:

$$V(\phi) \simeq \frac{1}{2} m^2 \phi^2 \left( \frac{\phi^2}{2M^2} \right)^K \propto \phi^{2+2K}$$  \hspace{1cm} (34)

where we assume $|K| \ll 1$. The average equation of state

$$\langle p \rangle \simeq \frac{K}{2} \langle \rho \rangle \approx -|K| \langle \rho \rangle,$$  \hspace{1cm} (35)

where $p$ and $\rho$ is a pressure and energy density of the scalar field, respectively. The negative value of $K$ corresponds to the negative pressure, which signals the instability of the condensate. At the level of linear perturbations [128] one can show that the field fluctuations grow exponentially if the following condition is met

$$\frac{k^2}{a^2} \left( \frac{k^2}{a^2} + 2m^2_{1/2} K \right) < 0.$$  \hspace{1cm} (36)

The instability band exists for negative $K$, as expected from the negative pressure arguments [133]. The instability band, $k$, is in the range [128] $0 < k^2/a^2 < k^2_{1/2} \equiv 2m^2_{1/2}/|K|$, where $a$ is the expansion factor of the Universe. The most amplified mode lies in the middle of the band, and the maximum growth rate of the perturbations is determined by $\dot{\alpha} \sim |K|m_{1/2}/2$ [133]. When $\delta \phi / \delta \phi_0 \sim O(1)$, the fluctuations become nonlinear. This is the time when the homogeneous condensate breaks down into Q-balls and anti-Q-balls. Such a phenomena can also yield gravitational waves due to the anisotropic stress created during the process of fragmentation of the inflaton, and this has been studied in [125, 126].

2.6. Radiation, dark matter, and dark energy

After the end of inflation, and the end of reheating/preheating yields the most important phase of the Universe, i.e. known as the radiation domination phase. See table 1 for known transitions in the cosmic history. The exact transition from the reheating phase, as we have seen above, depends on lots of parameters and rather model dependent on a particular nature of the BSM physics. However, that the reheating phase must come to an end before the BBN [80, 186] at temperature of $\sim 1$ MeV, when the hadrons have already formed. After reheating, the Universe is primarily dominated by relativistic species, assuming that they are all in LTE, the Hubble expansion rate is then determined by the ambient temperature,

$$H = \sqrt{\frac{\rho}{3M^2_\gamma}} = 1.66 \times g_\ast^{1/2} T^2 \frac{M^2_\gamma}{P},$$  \hspace{1cm} (37)

where $g_\ast$ is the total number of relativistic degrees of freedom and it is given by

$$g_\ast(T) = \sum_{i=b}^n g_i \left( \frac{T_i}{T} \right)^4 + \frac{7}{8} \sum_{i=I}^n g_i \left( \frac{T_i}{T} \right)^4.$$  \hspace{1cm} (38)

Here, $T_i$ denotes the effective temperature of the species $i$, which has decoupled. During the radiation era when $H = (1/2t)$, one finds

$$\frac{t}{1s} \approx 2.42g_\ast^{-1/2} \left( \frac{1 \text{ MeV}}{T} \right)^2.$$  \hspace{1cm} (39)

The light elements, such as $^2H, ^3He, ^4He$, and $^7Li$ are synthesized during the first few hundred seconds [80, 186]. The abundances depend on the baryon-to-photon ratio [80, 187, 188]

$$\frac{n_p}{n_e} = 273.3036 \delta \rho^2 (1 + 7.6958 \times 10^{-3} Y_p) \left( \frac{2.7258 \frac{M}{T^4}}{1 \text{ MeV}} \right)^3.$$  \hspace{1cm} (40)

where $Y_p = 2(n_p/[1 + (n/p)]) \approx 0.25$, and $n/p$ is the ratio of neutron-proton abundance at the temperature of $\sim 0.1$ MeV [80]. The latest constraint on $\Omega_b = \rho_b / \rho_c$ comes from the Planck data, where $\rho_c$ is the critical energy density of the Universe, i.e. $\rho_c = 3H^2/8\pi G$, where $G$ is the Newton’s constant, and $H_0 = 100 \times h$, and $H_0 = 67.4 \pm 0.5$ Km/s/Mpc at 68%CL, while $\Omega_b h^2 = 0.2224 \pm 0.0001$ at 68%CL [46]. Planck also gives constraint on the relativistic species which matches well with the constraints arising from the BBN, i.e.

$$N_{\text{eff}} = 2.99 \pm 0.17.$$

The radiation domination ends when the non-relativistic matter starts dominating the Universe, the radiation-matter equality happens when

$$1 + z_{\text{eq}} = \frac{\Omega_m}{\Omega_r} = 2.41 \Omega_m h^2 \times 10^3 \sim 3.3 \times 10^3,$$  \hspace{1cm} (41)

where $\Omega_m = \Omega_c + \Omega_b$ is the total matter density. The bound on $\Omega_b h^2 = 0.12 \pm 0.001$ at 68% CL, while $\Omega_m = 0.315 \pm 0.007$ at 68% CL [46]. The temperature of the Universe is roughly given
Table 1. List of key times in the early Universe in terms of temperature, redshift and time. We are assuming the reheating temperature was sufficiently high and that each phase transition occurred in a single step. A large baryon chemical potential in the early universe can also change the time of the QCD phase transition slightly and the temperature at which the electroweak phase transition occurs has some model dependence.

| Phase Transition                | $T$ (GeV) | $t$ (s) |
|--------------------------------|-----------|---------|
| Electroweak transition         | $\sim 20$–200 | $10^{-11}$ |
| QCD transition                 | $10^{-4}$  | $10^2$  |
| Big bang nucleosynthesis       | $\times 10^{-4}$ | $10^{12}$ |
| Recombination                  | $10^{-10}$ | $10^{12}$ |

by $T_{eq} = T_0(1 + z_{eq}) = 6.57 \times \Omega_m h^2 \times 10^2 K \sim 6.2 \times 10^3 K$. The creation of non-relativistic cold dark matter has to happen some where deep inside the radiation epoch. There are number of dark matter candidates, which include both thermal and non-thermal candidates, see reviews [152, 189, 190].

Amongst thermal dark matter, the well-known candidate is hot neutrinos [135, 191]. In fact the neutrinos decouple roughly around when the temperature of the Universe falls below 1 MeV, the interaction rate per neutrino falls below the Hubble expansion rate, after which the neutrino number density is conserved and their momentum falls as $1/a$. The neutrino has large free streaming length, which is detrimental to form large scale structures in the Universe. This is the prime reason they fall out of favour for the heavier cold dark matter candidate. The challenge remains is how to slow them down, lacking of any credible mechanism leads to postulating sterile neutrinos as a possible candidate for warm or cold dark matter, for a review see [192, 193]. From structure formation point of view, above 3 KeV sterile neutrino, which are thermally decoupled from the plasma, can be regarded as virtually cold, see recent analysis [71, 194]. As per as cold dark matter is concerned, there are plethora of models [195], but their creation mechanism remains predominantly thermal decoupling, such as freeze out [135, 196–199] or freeze in [200], or non-thermal processes via decay of some heavy particles, such as decay of the inflaton itself, or part of the inflaton itself [201].

The long wavelength CMB perturbations do not grow during the radiation epoch, but once matter domination starts, the initially induced CMB perturbations get a chance to grow and seed density perturbations in matter sector, which includes DM, baryons and photons to form first structures in the Universe, for a review see [70, 202, 203].

From a particle physics perspective, the major phase occurs very late in the history of the Universe when the Universe seems to be accelerating, for a review, see [204, 205]. The latest data from Placnk constraints the dark energy abundance to be $\Omega_d = 0.684 \pm 0.007$ at 68% CL. [46]. This apparent acceleration can be explained by apparent domination of dark energy in its simplest form, i.e. the cosmological constant in the Einstein Hilbert gravity. Indeed, the challenge lies how to protect the apparently small cosmological constant from radiative corrections, see [206, 207], which remains an outstanding unresolved problem. There are proposals to modify gravity in the infrared, however, without much observational or theoretical motivations in our opinion, for a review see [208].

3. Nature of phase transitions

In this chapter we are mostly concerned with thermal phase transition, preheating, fragmentation were clear examples of non-thermal phase transitions, which we had briefly discussed earlier. We will begin with the nature of phase transitions below.

3.1. First and second order transitions

In quantum field theory a phase transition is typically thought of as a transition between one vacuum state and another. For simplicity let us consider the case where the system is in the absolute ground state at some particular high temperature and as the Universe cools a new ground state becomes energetically favorable. If the ground state evolves continuously then this is what is known as a second order phase transition (or more generally a continuous phase transition). Alternatively if there is a discontinuous change in the ground state of the quantum field theory then this is a first order phase transition. There is also a discontinuity in the entropy during a first order phase transition. As such a first order phase transition releases a large amount of latent heat.

To illustrate the different types of phase transition we give a graphic illustration in figure 1 where the left and right panel demonstrate a second and first order phase transition respectively. In the case of a first order phase transition there is a barrier between a local minimum and the absolute ground state. As such the phase transition occurs through quantum tunneling and initially only occurs in regions of space called bubbles. These bubbles of new phase grow and coalesce in a background of the old phase. As the Universe cools until the new phase replaces the old one completing the phase transition.

In the case of first order phase transitions the size of the discontinuity can be compared to the temperature and in the case where the size of discontinuity is comparable or large compared to the temperature, the transition is referred to as a strongly first order phase transition. We spell this out in more detail in later sections. In general a strongly first order phase transition is of particular interest to particle cosmology as the violent process of bubble nucleation and the subsequent collisions can result in striking primordial gravitational wave signals. Furthermore a strongly first order phase transition is of particular interest for explaining the asymmetry between matter and anti-matter. It is also possible to produce magnetic field and defects during either first and 2nd order phase transitions.

Let us conclude this section with a final note. Our statement about the ground state continuously evolving being associated with a 2nd order phase transition is some what simplified. A crossover transition also exhibits this quality. For a second order phase transition the correlation length goes to infinity.
and some masses as well as the specific heat becomes non-analytic at the critical temperature. Conversely a crossover transition has no non-analytic properties and all correlation lengths remain finite.

### 3.2. Thermodynamical parameters

In this section we will discuss the most important thermodynamic parameters during a first order phase transition. We focus on a strongly first order phase transition since these are the main focus of particle cosmology, be it baryogensis, or gravitational wave production. A first order phase transition proceeds through bubble nucleation [209–211]. We will be interested in calculating the temperature at which these bubbles appear, the velocity at which they expand, the total number of bubbles, the fraction of new phase volume, the latent heat and the speed of the phase transition.

In this section our discussion will be in broad strokes with the exception of the wall velocity which we leave to its own section. To perform calculations one needs to know in more detail the effective potential and action at finite temperature which we delay to later sections. Let us begin by describing qualitatively the process of nucleation. The nucleation of bubbles does not occur immediately after the critical temperature even though the new phase has become energetically favourable. Instead, since the transition occurs through tunneling one must wait until the tunnelling rate is fast compared to the Hubble time. Eventually the Universe cools to the point where there is at least one critical size bubble in the Hubble volume. This temperature is denoted as the nucleation temperature. Finally when the volume fraction of the old phase is negligible the phase transition completes at a temperature $T_f$.

The tunnelling of the field from the false vacuum to true vacuum can be described as a solution to the classical equations of motion for the field. Assuming spherical symmetry the equation of motion for a single scalar field with a potential that is bounded from below is [212]

$$\frac{d^2\phi}{dr^2} + 2 \frac{d\phi}{dr} = \frac{dV}{d\phi}, \quad (42)$$

with boundary conditions $\phi'(0) = 0$, $\phi(0) \sim \phi_{\text{true}}$ and $\phi(\infty) = \phi_{\text{false}}$. The is the tunnelling solution to the classical equations of motion, known as the bounce, is the one where the field starts near the true vacuum and continuously evolves to the new one. It is in general non-trivial to find the bounce solution as naive attempts to find a solution tend to find the static solution—which trivially satisfies the left hand side of the above equation—where the field is in its minimum—satisfying the right hand side. As such many different approaches have been proposed to solve for the bounce [213–217]. An approximate solution that is useful for illustrative purposes is the well known kink solution [215, 218]

$$\phi(r) \approx A(r) = \frac{\phi_0}{2} \left[ 1 - \tanh \left( \frac{r - \delta}{L_w} \right) \right] - D[r, L_w, \delta], \quad (43)$$

where $L_w$ defines the wall width, $\delta$ and offset and $D$ a regulating function to make the derivative vanish at zero. As an example consider the potential

$$V = 0.3\phi^2 - \phi^3 + 0.6\phi^4. \quad (44)$$

The kink solution that approximates the true solution is given by [213]

$$A(r) = \frac{0.937}{2} \left[ 1 - \tanh \left( \frac{r - 3.42}{2.12} \right) \right] - \frac{1}{2} \left[ 0.937 \text{Sech} \left( \frac{1.42}{2.12} \right)^2 \right] e^{-r}. \quad (45)$$

The above compares remarkably well to the true bounce solution as can be shown in figure 2. An alternative Ansatz was proposed in [219] and takes the form

$$A(r) = \frac{A_0}{1 + r e^{\sqrt{2}(x-r)}} \quad (46)$$

with the parameters given in the thick and thin wall limits in the above reference. Since all potentials with two minimum separated by a maximum can be approximated by a quartic function over the region of interested it is useful to generically solve general quartic potentials. Indeed since one can always make rescalings and shifts of the form $\phi \to \phi + a$, $\phi \to b\phi$,

\[ \text{This approximation will not hold if the potential is unbounded or the distance to the true vacuum is large compared to the height of the barrier.} \]
it turns out to be sufficient to solve the class of potentials of the form
\[ V(\alpha) = \frac{-4\alpha + 3}{2} \phi^2 - \phi^3 + \alpha \phi^4, \]
for the parameter range \(\alpha \in [0.5, 0.75]\). Using the Ansatz \[213, 215, 218\]
\[ A(r, \alpha) = \frac{x_0(\alpha)}{2} \left[ 1 - \tanh \left( \frac{r - \delta(\alpha)}{L_\alpha(\alpha)} \right) \right] - \frac{1}{2} \phi(\alpha) \text{Sech} \left( \frac{\delta(\alpha)}{E_\alpha(\alpha)} \right)^2 e^{-r}, \]
where we plot the dependencies of the Ansatz parameters in figure 3. In principle if the path between the true and false vacuum can be approximated by a quartic one can use such an ansatz to approximate the bounce solution. The main source of error in this will be the fact that the true bounce solution follows a curved path in field space when multiple fields are involved.

Note that for \(\alpha \sim 0.5\) one has degenerate minima and this is where the bubble wall is the thinnest. In practice the bubble wall width tends to be in the range 1/20 to 1/200 for phenom-enologically viable phase transitions. Denoting the bounce wall width to be in the range 1/20 to 1/200 for phenomenologically viable phase transitions. Denoting the bounce solution as \(\phi_B\), the decay rate of the old phase to the new phase is controlled by the Euclidean action
\[ S_E(\phi_B) = 4\pi \int_0^\infty dr r^2 \left( \frac{d\phi_B}{dr} \right)^2 + V(\phi_B), \]
where \(V\) is the effective potential at finite temperature which for now we leave unspecified. The lack of angular variables in the integration reflects the fact that we are considering a spherically symmetric solution. The decay rate per unit volume of the effective potential is given by \[1\]
\[ \Gamma(T) \sim T^4 \left( \frac{S_E(\phi_B)}{2\pi T} \right)^{3/2} e^{-\frac{\phi(\phi_B)}{T}}. \]
From the decay rate one can write the differential decay probability for a given temperature \(T\) as \[20\]
\[ \frac{dP}{d\ln T} \sim \Gamma(T) \frac{M_p}{T^2} \left( \frac{t_{\phi}T_0}{T} \right)^3, \]
where \(t_{\phi}\) is the age of the Universe and \(T_0\) is the temperature today.

Using the relationship between temperature and time
\[ T^2t = \sqrt{\frac{45}{16\pi^3}} \frac{M_p}{\sqrt{g_*}}. \]
one can derive an approximate condition for when 1/e volume fraction is in the new phase
\[ \frac{S_E(\phi_B)}{T} = 170 - \log(\frac{T}{\text{GeV}}) - 2 \log(g_*). \]
The temperature which satisfies this equation is known as the nucleation temperature. A more precise way of calculating the nucleation temperature is through calculating the total number of bubbles in a Hubble volume at a given temperature, \(T\), \[20\]
\[ N = \int_0^{T_c/T} H^{-1}(T') V(T') \Gamma(T')dT'. \]
where \(T_c\) is defined in figure 1. The nucleation temperature is defined when the above expression is equal to 1. At a given time the total volume fraction of space in the false vacuum is \[20\]
\[ f_{\text{false}}(t) = e^{-I(t)}, \quad I(t) = \int_0^t d\tau' p(T(t'))V(t', t). \]
Here \(V(t', t)\) is the volume of a bubble formed at time \(t'\) evaluated at time \(t\). If one assumes spherical symmetry one can write \(V(t', t) = 4\pi [r(t) - r(t')]^3/3\). Taking \(r(t) - r(t') \sim c_{\nu}(t - t')\) for a constant wall velocity, \(c_{\nu}\), one can write the fraction of volume in the false vacuum as \[20\]
\[ f_{\text{false}} = \exp \left[ -\frac{4\pi}{3} c_{\nu}^3 \int_0^t dt' p(T(t'))(t - t')^3 \right]. \]
Dropping the time dependence of the temperature, the nucleation probability per unit volume in the above equation is given by \[20, 210\]
\[ p(T) = \frac{\omega}{2\pi} \left( \frac{S_E(\phi_B)}{2\pi T} \right)^{3/2} A(T) e^{-\frac{\phi(\phi_B)}{T^2}}. \]
In the above \(\phi_B\) is the bounce solution to the classical equations of motion, \(\omega\) is the angular frequency of the unstable mode and \(A(T)\) is the fluctuation determinant. The phase transition completes when \(f_{\text{false}}(t)\) becomes negligible. The

**Figure 2.** An example of the bounce solution for a sample potential (left panel). Above is presented the solution calculated numerically compared to the Ansatz given in equation (43). Since the difference is invisible to the naked eye we also plot the residuals (right panel).
duration of the phase transition can then determined by taking the difference between this time and the nucleation time. The speed of the phase transition is often parametrized in terms of the time rate of change of the effective action

$$\beta = -\frac{dS}{dt} \approx T \frac{d(S_E/T)}{dT} \bigg|_{T_N}.$$  

The speed of the phase transition is a parameter which controls the frequency and amplitude of relic gravitational wave backgrounds left by cosmic phase transitions. Other parameters that control the amplitude and frequency of relic gravitational waves are the Latent heat and the bubble wall velocity. The latent heat divided by the radiation energy density is given by

$$\alpha = \frac{\Delta \rho}{\rho_N}.$$  

with

$$\rho_N = \frac{\pi^2 g_* T_N^4}{30}, \quad \Delta \rho = \left[ V - \frac{dV}{dT} T_N \right]_{\text{false}} - \left[ V - \frac{dV}{dT} T_N \right]_{\text{true}}.$$  

3.3. Bubble wall velocity

After a bubble nucleates it expands creating an ever larger region of the new vacuum. The act of expansion leads to more particles in the plasma either acquiring or receiving a mass. Furthermore the equilibrium distributions of particles in the plasma gets perturbed near the bubble wall. These processes costs energy and results in resistance to the bubble expansion. If such friction is large, the bubble may not necessarily go ultra relativistic.

Let us begin with the classical equations of motion for a scalar field interacting with fermions and gauge bosons [220]

$$\Box h - \frac{\partial V_0}{\partial h} - \frac{d m_f^2}{d h} \langle \delta h^2 \rangle - \sum \frac{d m_A^2}{d h} \langle A^2 \rangle + \sum \frac{d m_\psi}{d \phi} \langle \bar{\psi}\gamma_\mu \psi \rangle = 0,$$

where $A$ and $\psi$ are gauge bosons and fermions that acquire a mass when $h$ acquires a vacuum expectation value. Also the angular brackets denotes the vacuum distribution at finite temperature but not necessarily in equilibrium. If the particle distributions are their equilibrium functions then the above are the same equations of motion one solves when finding the bounce solution that describes the profile of a bubble wall during nucleation. We will derive this fact in more detail when we review the fate of the effective potential at finite temperature in a later section. The expansion of the bubble wall however perturbs the plasma away from equilibrium and the energy required to perform such a perturbation resists the expansion of the bubble wall. The box operator contains a curvature...
friction term, $(2/r)dh/dr$, which can be neglected when the bubble expands to a sufficiently large size that we can neglect the curvature of the bubble wall. In this case the box operator becomes $\Box = (1 - v^2)\partial^2_z$ with the bubble moving along in the $z$ direction. Such an approximation underestimates the total friction and therefore will provide a slight overestimate of the bubble wall velocity. The 2nd moments of the fields can be written in terms of their vacuum expectation values (which will be zero) and their distributions

$$\langle \delta h^2 \rangle = \langle \delta h^2 \rangle_{vac} + \int \frac{d^3k}{(2\pi)^3} f_h(k, x),$$

$$\langle A^2 \rangle = \langle A^2 \rangle_{vac} + \int \frac{d^3k}{(2\pi)^3} E_k f_A(k, x),$$

$$\langle \dot{\psi}_L \psi_L \rangle = \langle \dot{\psi}_L \psi_L \rangle_{vac} + \int \frac{d^3k}{(2\pi)^3} E_k f_L(k, x). \quad (62)$$

Writing the distributions as

$$f_X = f_{X0} + \delta f_X,$$

where the $\delta f_X$ piece corresponds to the departure from equilibrium, and noting that

$$\frac{\partial V(h, T)}{\partial h} = \frac{\partial V_0}{\partial h} + \sum_j \frac{\partial m_j}{\partial h} \int \frac{d^3p}{(2\pi)^3} E_p f_j(p, x), \quad (63)$$

we can then rewrite the classical equations of motion such they explicitly contain the wall velocity

$$-(1 - v^2_h)\ddot{h} + \frac{\partial V(h, T)}{\partial h} + \sum_j \frac{\partial m_j}{\partial \phi} \int \frac{d^3p}{(2\pi)^3} 2E_p \delta f_j(p, x) = 0. \quad (64)$$

The sum in the above equation represents the resistance to bubble expansion due to the new phase causing particles in the plasma to acquire a mass and depart from their equilibrium distributions. When the WKB condition of $p_j \gg 1/\lambda_n$ is satisfied [220], the particle distributions satisfy Boltzmann equations [221]

$$\left( \xi + \dot{\xi} + p_i \partial_i \right) f_j = -C[f_j]. \quad (65)$$

We can parametrize the distribution function as

$$f_i = \frac{1}{e^{(E/E_\ast) / T} + x}, \quad (66)$$

where the value of $x = \pm 1$ denotes fermions (+) or bosons (−). The various terms in the Boltzmann equation can be written as

$$\dot{\delta f} = f'_0 \left( \partial_i E + \dot{\delta} \partial_i \right) = f'_0 \left( \frac{1}{2E} \partial_i m_i^2 + \partial_i \delta \right),$$

$$\dot{\xi} \delta f = f'_0 \frac{p_i}{E} \partial_i \delta,$$

$$\dot{p}_i \partial_i \delta f \sim 0. \quad (67)$$

The perturbations can be written as a sum of a perturbation in the chemical potential, the fluctuation in the temperature and the fluctuation in the velocity of each species. That is [220]

$$\delta_i = - \left[ \delta \mu_i + \frac{E}{T} \left( \delta T_i + \delta T_{bg} \right) + p_i \left( \delta v_i + \delta v_{bg} \right) \right]. \quad (68)$$

To solve the Boltzmann equations one takes the moments—$\int d^3p/(2\pi)^3$, $\int (E/T)d^3p/(2\pi)^3$ and $\int (p_i/T)d^3p/$ $(2\pi)^3$—of the Boltzmann equations to acquire a set of linear equations

$$c_{2\delta} \partial_t \mu_i + c_{3\delta} \partial_i (\delta T_i + \delta T_{bg}) + (c_{3\delta}/3) T \partial_i (\delta v_i + \delta v_{bg})$$

$$+ \int \frac{d^3p}{(2\pi)^3} \frac{E}{T} C(f_i) = c_{1\delta} \frac{\partial m_i^2}{2T} \partial_t ^2,$$

$$c_{2\delta} \partial_t \mu_i + c_{3\delta} \partial_i (\delta T_i + \delta T_{bg}) + (c_{3\delta}/3) T \partial_i (\delta v_i + \delta v_{bg})$$

$$+ \int \frac{E}{T} \frac{d^3p}{(2\pi)^3} \frac{E}{T} C(f_i) = c_{1\delta} \frac{\partial m_i^2}{2T} \partial_t,$$

$$c_{2\delta} \partial_t \mu_i + c_{3\delta} \partial_i (\delta T_i + \delta T_{bg}) + (c_{3\delta}/3) T \partial_i (\delta v_i + \delta v_{bg})$$

$$+ \int \frac{p_i}{T} \frac{d^3p}{(2\pi)^3} \frac{E}{T} C(f_i) = 0. \quad (69)$$

where we have defined

$$c_{nx} = \int \frac{d^3p}{(2\pi)^3} \frac{E^{n-2}}{T} f_{nx}(p). \quad (70)$$

One can write the above set of equations in a matrix equation [220, 222, 223]

$$\mathbf{M} \dot{\delta} + \Gamma \delta = \partial m_i^2 \dot{F}. \quad (71)$$

In the above the matrix $\Gamma$ is from the set of relevant collision integrals. If we assume that the derivatives of the fluctuations $\delta$ are small then one can invert the above matrix equation to obtain an expression for the fluctuations $\delta \sim \Gamma^{-1} \partial m_i^2 \dot{F}$. For a single field undergoing a phase transition the equations of motion including these perturbations can be written [220]

$$-\ddot{\phi}'' + \frac{\partial V}{\partial \phi} + \frac{1}{2} T \frac{\partial m_i^2}{\partial \phi} (c_{1\delta} \partial \mu_j + c_{2\delta} \partial T_j + c_{2\delta} T \partial \nu_j) = 0. \quad (72)$$

Substituting the approximation for $\delta$ into our equations of motions and taking the bubble wall velocity to be small gives

$$h'' - \partial_h V = \eta \nu \gamma \frac{\partial^2 \phi}{T} \phi', \quad (73)$$

where $\eta$ is given by

$$\eta = \frac{T}{4} G \Gamma^{-1} F. \quad (74)$$

and isolates the part of the friction that is independent of the wall velocity. Here the term $G$ is a vector that for a particle species $i$ has the form $2N_i m_i / v_c^2$, with $v_c = (c_{1\delta}, c_{2\delta}, c_{3\delta})$, and we give the standard model friction terms in table 2 from [224].

A rule of thumb for whether bubble walls can run away was developed by Bodecker and Moore [225]. Assuming the bubble wall reaches ultra relativistic speeds, the pressure that resists expansion due to particles crossing the wall and changing its mass reaches an asymptotic value independent of the Lorentz factor $\gamma$. In the ultra relativistic regime, one needs to only consider particles crossing the wall from the false vacuum to the true. Since no particles enter the false vacuum phase either through reflecting off the wall all through exiting the expanding bubble, the particle distributions can be assumed in the equilibrium distribution. Specifically the pressure reaches the value [225]
\[ P_{t \rightarrow 1} = \sum_{i} n_i \int \frac{d^3p}{2E(2\pi)^3} f_{0,i}(p)(m_i^2(h_T) - m_i^2(h_F)), \]

where \( h_{0,i} \) are the true and false vacuum respectively and the subscript \( 1 \rightarrow 1 \) refers to the process where particles which cross the wall and acquire a contribution to their mass. If the minimum of the mean field potential

\[ V = V_{|T=0} + \frac{1}{2} (h - h_t)^2 T^2, \]

either does not exist or is higher than the false vacuum of the full thermal potential, the bubble wall cannot run away. This criteria implies that phase transitions involving scalar singlets tend to runaway as they introduce additional expansion pressure without introducing too much additional friction. Even still [224] found there are in fact some cases where the bubble wall in a singlet catalyzed electroweak phase transition can expand subsonically, that is the velocity of the bubble wall is slower than the speed of sound within the plasma.

### 3.4. Multistep phase transitions

Thus far we have been considering the case where the phase history of a system is the simple case where one has one phase at high temperature and a different one at low temperature. Condensed matter systems teach us that things may not be so simple. Indeed there has been much recent interest in the case of multistep phase transitions and their application to baryogenesis [13–17, 226, 227] and gravitational waves [228]. In general there are four distinct cases of interest. The first where a symmetry is not broken in one of the phase transitions (for example a gauge singlet field may tunnel from one vacuum to another). The second case is where the same symmetry is broken in both phase transitions such as a two step electroweak phase transition. One can also break two different symmetries in subsequent transitions such as the case where the standard model is extended by a singlet field with a discrete \( Z_2 \) symmetry where \( V(\phi) = V(-\phi) \). Such a model was considered in [229] in the context of gravitational waves. Finally, there is the case where one has a symmetry at zero temperature is broken at an intermediate temperature (and then possibly restored at high temperature).

For the first case consider the example of a real singlet where the first phase transition proceeds as \( (0, \epsilon) \rightarrow (0, \nu) \) where the first phase transition does not begin exactly at the origin as finite temperature effects generate a linear term that shifts the minimum away from \( v = 0 \). Second, the electroweak phase transition proceeds from \( (0, \nu) \rightarrow (\nu, \nu') \) which can have a tree level barrier between the true and false vacuum catalyzing a strongly first order phase transition. In principle such a scenario could lead to exotic gravitational wave effects. In the context of NMSSM, a phase transition where the singlet changes sharply during the electroweak phase transition can boost the efficiency of baryon production [230]. Detailed phenomenological scans of a real singlet extension to the standard model have shown that a scalar as heavy as 800 GeV can still catalyze a SFOEWPT [21, 231]. Such a scenario would take a 100 TeV collider to fully probe [232]. The requirement that a singlet must be no heavier than a 800 GeV sounds model specific leading to a question as to whether a more complicated scalar sector could push the scale of new physics even higher and still catalyze a SFOEWPT. However the result for the singlet extension of the standard model seems to agree with the effective field theory result which also sets the scale of new physics at 800 GeV [233]. Recent work has stressed the difficulties of using effective field theory during a phase transition so this result should be taken with a grain of salt [218].

Next let us consider the second possibility raised above. The idea of a symmetry being broken through multiple transitions has been mainly of interest in the case of electroweak symmetry breaking. One can break electroweak symmetry in a multistep transition when the standard model is extended by at least an additional scalar \( SU(2)_L \) multiplet. The simplest cast is when one adds a scalar \( SU(2)_L \) multiplet to the standard model along with a scalar gauge singlet to act as a catalyst. A component of the \( SU(2)_L \) triplet can acquire a vacuum expectation value at high temperature before the deepest minima is in the direction of the \( SU(2)_L \) doublet fields. This situation was considered in [226] to catalyze electroweak baryogenesis at a higher scale and is illustrated in figure 4. A key attraction of this scenario is the physics that leads to the electroweak phase transition being strongly first order, a requirement for

![Table 2. Table of friction coefficients for relevant SM processes. Table contents taken from [224].](image-url)
electroweak baryogenesis, can be above the electroweak scale avoiding current bounds.

Finally the last possibility raised above is arguably the most exotic. Weinberg demonstrated by example that counter to our intuition, symmetries can be broken at high temperature [24]. Indeed this is the case in some condensed matter systems such as Rochelle salt which prompts us to consider this scenario. It was shown in [234] that if one extends the scalar sector of the standard model by a colored scalar field, one can indeed generate a scenario where the colored scalar acquires a vacuum expectation value before losing it during the electroweak phase transition. If one also includes a gauge singlet the scale of both the color breaking and electroweak phase transitions can both be multi-TeV. This scenario was recently considered as a mechanism for producing the baryon asymmetry of the Universe [227]. Note that the colored scalar cannot be a stop [235] as if one tunnels into a phase where the stop has a vacuum expectation value, one cannot efficiently tunnel back into the SU(3)C symmetric phase when the Universe cools. In addition to adding extra scalars to the effective potential, one can achieve symmetry breaking at high temperature through modifying the effective potential with large chemical potentials [236–238].

4. Effective potentials at finite temperature

It is possible for a scalar field to acquire a vacuum expectation value. This vacuum expectation value can be space-time dependent so we can treat it like a field. It is however, a classical field rather than a quantum field as it does not have excitations that correspond to particle states. To derive the effective potential the process is to shift the scalar fields by the expectation value such that the expectation value of the scalar fields are zero. The part of the shifted Lagrangean that is purely made up of the classical field we call the effective potential. The global minimum of the effective potential is the vacuum expectation value of the unshifted field. It is common to refer to local minima as false vacua as they can decay to the true vacuum through tunnelling.

Let us consider some examples. In the case of the standard model, the Higgs is complex SU(2)L doublet so it formally contains four scalar quantum fields. All four fields can acquire a vacuum expectation value so we can shift by a SU(2)L doublet of four classical fields which are the vacuum expectation values for each field. However, gauge invariance then allows us to perform a rotation such that only a single classical field is necessary. It is customary to choose the shift to be in the following direction in the internal space

\[ H \rightarrow \left( \frac{\phi_1^0 + i\phi_2^0}{\sqrt{2}}, -\frac{\phi_1^0 - i\phi_2^0}{\sqrt{2}} \right) + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \],

(77)

where \( \nu \) is the vacuum expectation value. For now we have omitted any possible space time dependence of \( \nu \). The vacuum expectation value spontaneously breaks SU(2)L symmetry and its associated Noether current (as well as the associated charge) is no longer conserved. Instead the conserved charge is the one that returns zero when acting on the shifted classical fields. It is easy to see that if the Higgs has a hypercharge of \( Y = 1/2 \), the linear combination \( Q_{EM} = \tau_3 + Y \) satisfies this criteria and is then the conserved charge of the theory with broken symmetry.

Next let us consider the case where there are two Higgs doublets with the same hypercharge. In this case we have a total of eight scalar fields between the two SU(2)L doublets. We can again shift all eight scalar fields with a classical field corresponding to their vacuum expectation values. Once again we can use gauge invariance to render some of the classical fields redundant. However, this time we are still left with five classical fields. If we use our freedom to rotate away all but one classical field on the first Higgs doublet one has

\[ H_1 \rightarrow \left( \frac{\phi_1^0 + i\phi_2^0}{\sqrt{2}}, 0 \right), \]

(78)

\[ H_2 \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^0 + i\phi_2^0 \\ \phi_3^0 + i\phi_4^0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} \nu_1^0 + i\nu_2^0 \\ \nu_3^0 + i\nu_4^0 \end{pmatrix} . \]

(79)

In the above, the vacuum expectation value \( v_1^0 \) violates the same charge as the standard model case. In contrast, \( v_3^0 \) and \( v_4^0 \) violate \( Q_{EM} = \tau_3 \) and \( v_2^0 \) is a CP odd vacuum. In zero temperature equilibrium QFT, the vacuum expectation values are fixed and it is unnecessary to consider their space time dependent behaviour. During a cosmic phase transition however, the vacuum expectation value can evolve with space and time.

4.1. Coleman–Weinberg potential

The process for calculating temperature effects on the evolution of the vacuum follows the same recipe as the process of calculating loop effects albeit with finite temperature ingredients. Specifically, in the real time formalism the propagators are replaced by their finite temperature counterparts, the masses are corrected by a Debye term and the time contour changes (as will be discussed in the forth coming subsection). Therefore we first summarize the effects of zero temperature
loop interactions on the effective potential and, by extension, the vacuum expectation value.

The shifted Lagrangian contains interactions between the classical field and quantum fields. One can therefore consider diagrams such as shown in figure 5. Calculating the case of Higgs self interactions, $\lambda h^4$, with the physical Higgs one has [239]

$$
V_{\text{self},h} = i \sum_{n=0}^{\infty} \int \frac{d^4p}{(2\pi)^4} \frac{1}{2n} \left[ \frac{\lambda v^2/2}{p^2 - m_H^2 + i\epsilon} \right]^n,
$$

(80)

whereas interactions between the classical field and the Goldstone modes at one loop gives a contribution to the effective potential

$$
V_{\text{self,GM}} = i \sum_{n=0}^{\infty} \int \frac{d^4p}{(2\pi)^4} \frac{2}{2n} \left[ \frac{\lambda v^2/12}{p^2 - (m_{GB}^2 + \xi m_W^2) + i\epsilon} \right]^n,
$$

(81)

Figure 5. 1PI diagrams contributing to one loop corrections to the effective potential including scalar contributions (top panel), fermion contributions (central panel) and gauge contributions (bottom panel). Reproduced from [243] with permission.

where refers to the Goldstone mode and $\xi$ is the gauge fixing parameter. One can see explicitly in the above that the one loop correction to the effective potential acquires both gauge and renormalization dependence. Indeed, the one loop corrections even depend on the renormalization scheme. These issues we return to later. For now we can work in the Landau gauge ($\xi = 0$) as this conveniently hides the gauge dependence. We also use dimensional regularization in the $\overline{\text{MS}}$ renormalization scheme. We can categorize the one loop corrections to the effective potential by the virtual state in the interaction the correction corresponds to. The one loop corrections due to scalar (including goldstone boson), gauge boson and scalar interactions are respectively, with respect to the running energy $\mu$

$$
V_{\text{CW}}^S = \frac{m^4}{64\pi^2} \left( \log \frac{m^2}{\mu^2} - \frac{3}{2} \right),
$$

(82)

$$
V_{\text{CW}}^B = \frac{m^4}{64\pi^2} \left( \log \frac{m^2}{\mu^2} - \frac{5}{6} \right).
$$

(83)
\[ V_f^{CW} = -\frac{m_f^4}{64\pi^2} \left( \log \frac{m_f^2}{\mu^2} - \frac{3}{2} \right), \]  

where CW stands for the Coleman–Weinberg potential and \(S, GB\) and \(f\) refer to scalars, gauge bosons and fermions respectively. One then has the total one loop correction to the effective potential as

\[ V_1(T = 0) = \sum_b n_b V_b^{CW} - \sum_f n_f V_f^{CW}, \]  

where \(n_{bf}\) is the number of bosonic/fermionic multiplicity factors \(n_i = 12, n_w = 6\) etc.

### 4.2. Thermal corrections from scalars, fermions and gauge bosons

Calculating the finite temperature corrections to the effective potential essentially means repeating the previous analysis using finite temperature propagators, modifying the masses by temperature dependent ‘Debye’ corrections and modifying how we treat time. There are two commonly used treatments of time at finite temperature that give the same result for the corrections to the effective potential. One involves performing calculations in imaginary time; in this case the time domain is compactified to an interval \(p_0 \in [0, \beta]\). The other approach remains in real time formalism where the time contour is modified as shown in figure 6. This approach is slightly more complicated in its formalism, however a few features are more transparent. Therefore we will briefly summarize the closed time path formalism here.

The reason behind the different contour can be understood as follows. Suppose some system at a time \(t = 0\) is in equilibrium. The density matrix then has the familiar form \([240, 241]\)

\[ \rho_0 = \frac{e^{-\beta H}}{\text{Tr}e^{-\beta H}}. \]  

The time dependent density matrix can then be derived by evolving the equilibrium density matrix with time evolution operators,

\[ \rho(t) = U(t, 0)\rho(0)U(t, 0). \]  

Note the explicit form of the time evolution operators

\[ U(t', t) = T \left( e^{-i J_0} \right). \]  

Here \(T\) refers to time. Both the equilibrium and time dependent density matrices can be expressed in terms of the equilibrium operators

\[ \rho(0) = \frac{U(T - i\beta, T)}{\text{Tr}U(T - i\beta, T)}, \]  

\[ \rho(t) = \frac{U(t, 0)U(T - i\beta, T)U(0, t)}{\text{Tr}U(T - i\beta, T)}. \]  

We can then calculate the time dependent expectation value of any operator, \(\langle A(t) \rangle\) and express it purely in terms of time evolution operators and \(A\) \([240, 241]\)

\[ \langle A(t) \rangle = \text{Tr}[\rho(t)A], \]  

\[ = \frac{\text{Tr}U(T - i\beta, T)U(T', T)U(T', t)AU(t, T)}{\text{Tr}U(T - i\beta, T)U(T, T')U(T', T)}. \]  

Taking \(T \to -\infty\) and considering some \(n\)-point correlator for some source, \(J\), we see that we can interpret the above as taking the time contour given in figure 6. As a result we have four distinct types of propagators depending on where the components of the bilinears are on the time contour. It is convenient to represent these in matrix form. We will restrict ourselves to the scalar case \([240, 241]\)

\[ \Delta(x, y) = \begin{pmatrix} \Delta^{++} & -\Delta^{+-} \\ -\Delta^{-+} & -\Delta^{--} \end{pmatrix}, \]  

\[ = \begin{pmatrix} \langle T [\phi(x)\phi^\dagger(y)] \rangle & \langle T[\phi(x)\phi^\dagger(y)] \rangle \\ -\langle \phi(x)\phi^\dagger(y) \rangle & -\langle T[\phi(x)\phi^\dagger(y)] \rangle \end{pmatrix}. \]  

The four propagators in momentum space we give explicitly

\[ i\Delta^{++}(p) = 2\pi\delta(p^2 - m(T)^2) \left[ \Theta(p_0) f(\bar{p}) + \Theta(-p_0)f(-\bar{p}) \right], \]  

\[ i\Delta^{-+}(p) = 2\pi\delta(p^2 - m(T)^2) \left[ \Theta(p_0) (1 + f(\bar{p})) + \Theta(-p_0)f(-\bar{p}) \right], \]  

\[ i\Delta^{++}(p) = \frac{1}{p^2 - m(T)^2 + i\epsilon} \left[ \Theta(p_0) (1 + f(\bar{p})) + \Theta(-p_0)f(-\bar{p}) \right], \]  

\[ i\Delta^{--}(p) = \frac{-1}{p^2 - m(T)^2 - i\epsilon} \left[ \Theta(p_0) f(\bar{p}) + \Theta(-p_0)f(-\bar{p}) \right]. \]  

These propagators can be essentially derived from unitarity and causality \([242]\). Note that the first two propagators vanish at zero temperature. This is expected as they have no zero

![Figure 6. The closed time path contour. The contour begins above the real line far in the past, evolves to the present before dropping below the real line and evolving back into the past. At T = −∞ the contour evolves perpendicularly to the real line.](image-url)
temperature counterpart. The last two operators (referred to as the time and anti time order propagators) contain a sum of zero temperature and finite temperature pieces. The finite temperature pieces are Boltzmann suppressed when the temperature drops well below the masses so the time and anti-time ordered propagators reduce to their zero temperature counterparts as the temperature goes to zero.

The finite temperature corrections to the effective potential (apart from the Debye corrections to the mass) are produced by recalculating the one loop corrections to the effective potential, but this time with the finite temperature versions of the propagator. As an example let us consider just the one loop corrections due to the top quark, the physical Higgs and the classical field. It is actually easier to calculate the derivative of this term with respect to the mass. In this case we just need to calculate a single bubble diagram [239, 241]

$$\frac{\partial V_1}{\partial m^2(t)} = \frac{1}{2} \int \frac{d^4p}{(2\pi)^4} \Delta^{++}(p).$$  \hfill (99)

Note the appearance of the time orders propagator which, as we have stated, is a sum of the zero temperature and finite temperature pieces. Therefore we recover the zero temperature loop correction but now have an additional finite temperature piece given by [239, 241]

$$\Delta V_b = \frac{T^4}{2\pi^2} n_b J_b \left(\frac{m^2}{T^2}\right),$$  \hfill (100)

$$J_b(z^2) = \int_0^\infty dx x^2 \log \left[1 - e^{\sqrt{x^2 + z^2}}\right],$$  \hfill (101)

where $z = m/T$ has been implicitly defined. Note that the above function is complex for negative arguments. The imaginary parts of the effective potential contains some decay which reflects an instability in the system. This issue we return to later in this section. Performing the same analysis with fermions as virtual states one can then derive the finite temperature contributions to the effective potential due to fermions [239, 241]

$$\Delta V_f = \frac{T^4}{2\pi^2} n_f J_f \left(\frac{m^2}{T^2}\right),$$  \hfill (102)

$$J_f(z^2) = \int_0^\infty dx x^2 \log \left[1 + e^{\sqrt{x^2 + z^2}}\right].$$  \hfill (103)

The total contribution at one loop including finite temperature corrections we can then write as [239, 241]

$$V_1(T) = \sum_b n_b (V_b^{\text{CW}} + \Delta V_b) - \sum_f n_f (V_f^{\text{CW}} + \Delta V_f).$$  \hfill (104)

\[4.2.1\] Debye masses and daisy diagrams. Hard thermal loops, $p \sim T$, can cause perturbation theory to break down at finite temperature. One can delay the break down of perturbation theory by performing a resummation in figure 8 which results in the shift of the pole mass by a temperature dependent Debye term. We will discuss different derivation schemes in this section as well as what limits different corrections become important. The dangerous diagrams can be categorized as in figure 8: daisy diagrams, super daisy diagrams, lollipops and Sunsets. Note that the last two types of diagrams only exist in the case where you have a dimensionful trilinear coupling. We will focus therefore on the first two. Daisy contributions of the form given in figure 7 become important when the mass of a particle is small compared to the temperature [244]. Super-daisy diagrams such as the form given in figure 8 are important when the couplings are large and the masses are small compared to the temperature [243]. Consider the simplest possible model, a $\phi^4$ model.

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi^2 - \frac{m^2}{2} \phi^2 - \lambda \phi^4.$$  \hfill (105)

Diagrams such as the one shown in figure 7 contribute to a thermal correction to the mass. Let us explicitly calculate such a diagram in the imaginary time regime [240]

$$\Delta m^2 = \frac{\lambda}{2\beta} \sum_n \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 + m^2} + k^2 + m^2$$  \hfill (106)

$$= \frac{\lambda}{4} \int \frac{d^4k}{(2\pi)^4} \frac{1}{\sqrt{k^2 + m^2}} + \frac{\lambda}{2} \int \frac{d^4k}{(2\pi)^4} \frac{1}{\sqrt{k^2 + m^2}} e^{\sqrt{k^2} + \sqrt{m^2}} - 1.$$  \hfill (107)

In the limit where the temperature is large compared to the mass one can write the high temperature expansion [240, 245].

$$\Delta m^2_T = \frac{\lambda T^2}{24}.$$  \hfill (108)

This indeed is the typical form used for the Debye mass. If one takes the temperature dependent part of the potential evaluated at the thermally corrected mass, $V_T(m^2 + \Delta m^2)$, and performs a high temperature expansion one encounters a common approximation for the daisy contributions [246]

$$\frac{T^4}{2\pi^2} J_b \left(\frac{m^2 + \Delta m^2}{T^2}\right) \sim \frac{T^4}{2\pi^2} J_b \left(\frac{m^2}{T^2}\right) + V_{\text{Daisy}},$$  \hfill (109)

with

$$V_{\text{Daisy}} = \frac{T^4}{12\pi} \sum \left[ \left(\frac{m^2}{T^2}\right) - \left(\frac{m^2 + \Delta m^2}{T^2}\right)^{3/2} \right].$$  \hfill (110)
The terms $m^2 + \Delta m^2$ are the eigenvalues of $m_{sG}^{ij} + \Pi_{sG}^{ij}$ where $m_{sG}^{ij}$ are the scalar and gauge boson mass matrices respectively and [247]

$$
\Pi_s^{ij} = \frac{T^2}{24} n_k (-1)^{j+i} \partial^{kij} L - \frac{T^2}{48} n_k \partial^{kij} L_Y
$$

$$
\Pi_{sG}^{ab} = \frac{2}{3} T^2 \left( \tilde{n}_H \frac{8}{8} + 5 \right) \frac{1}{n_l} \sum_{m} \delta_{ab} \delta^{ij} \Delta m^2 L
$$

with $n_k$ the appropriate multiplicity factors, $s = (0, 1)$ the spin of the boson and $\tilde{n}_H$ is the number of Higgs that couple to a gauge boson. Finally the derivatives $\partial_i$ are derivatives with respect to field $x$, $L_Y$ is the part of the Lagrangian that contain Yukawa interactions and $L = L - L_Y$. Note that only longitudinal gauge bosons acquire thermal mass corrections. It was recently shown by [243] that the high temperature estimation of the thermal corrections can become a poor approximation. As can be seen in figure 9, the high temperature approximation is quite poor for large values of the quartic coupling. More seriously, the high temperature expansion does not show how decoupling when the mass is large compared to the temperature. A more accurate way of calculating the Debye mass is through the self-consistency relation [243]

$$
m_0^2(\phi) + \Delta m^2 = \frac{\partial^2}{\partial \phi^2} [V_0(\phi) + V_{\text{CW}}(\phi)] + \frac{\partial^2}{\partial \phi^2} V_T(m_0^2(\phi) + \Delta m^2),
$$

$$
\Delta m^2 = \frac{\partial^2}{\partial \phi^2} V_T(m_0^2(\phi) + \Delta m^2).
$$

Furthermore [243] developed a scheme for calculating the superdaisy contributions, by solving the equation [243]

$$
V_{\text{super daisy}} = \int d\phi \frac{dV_T(m_0^2(\phi) + \Delta m^2(\phi)) \ dm^2(\phi)}{dm^2(\phi)}.
$$

4.2.2. Gravitational corrections. Gravitational corrections to false vacuum decay suppress the decay rate slightly. Consider the action of a single scalar field coupled to gravity defined by the Euclidean action [248, 249]

$$
S = \int d^4x \sqrt{g} \left[ \frac{1}{2} \partial_{\mu} h \partial^\mu h + V(h) - \frac{R}{2\kappa} - \frac{R}{2f(h)} \right],
$$

where $\kappa = 8\pi G = 8\pi/M_P^2$ and the metric is given by $ds^2 = dr^2 + \rho(r)^2 d\Omega^2$, where $d\Omega^2$ contains the angular part of the metric. In this metric the Ricci scalar has the simple form

$$
R = \frac{-6(\rho^2 \rho'' + 2 \rho')}{\rho^3},
$$

where the prime is a derivative with respect to the radial coordinate, $r$. The equations of motion are

$$
\frac{df}{d\rho} = \frac{\rho}{2} \left( h'' \rho + \rho h'' \right) = \frac{dV}{d\rho} - \frac{1}{2} \frac{df}{d\rho} R.
$$

One can find approximate solutions to the above equations of motion by expanding in the Planck mass,

$$
h(r) \sim h_0(r) + s h_1(r),
$$

$$
\rho(r) \sim \rho_0 + \kappa \rho_1(r),
$$

where $\kappa = \sqrt{8\pi G}$ and $h_0$ is the bounce solution without gravitational corrections. The change in the action, $S = S_0 + \Delta S$, where $S_0$ corresponds to the action of the bounce has a formidable form [249]

$$
\Delta S = \frac{48\pi^2}{M_P^2} \int dr \left[ r^2 \rho_1 \left( \frac{1}{2} h_0^2 + V(h_0) \right) + (r \rho_1')^2 + 2 \rho_1 \rho_1' + 2 \rho_1 \rho_1'' + 2 \rho_1' + 2 \rho_1'' \right].
$$
One can greatly simplify the above expression by making the rescaling $\rho_1 \rightarrow \rho_1 \times s$ where the value of $s$ is chosen by requiring the derivative with respect to $s$ to vanish at $s = 1$. The change in the effective action then has the remarkably simple form

$$\Delta S = \frac{48 \pi^2}{M_P^2} \int \mathrm{d}r \rho_1'^2 \geq 0.$$  

Note that, as we indicate, this correction to the effective action from gravitational effects is always positive indicating that the tunneling rate in turn is suppressed by gravitational corrections. Also it is useful to note that the first order correction to $\rho$ is independent of the gravitational corrections to the bounce action [249]

$$\rho_1' = \frac{r^2}{6} \left[ \frac{1}{2} h_0^2 - V - \frac{3}{r} f'(h_0) h_0 \right].$$  

The correction is to be quite small for weak scale phase transitions, scaling as $\sim \sigma^6 / (\Lambda^4 M_P^2)$ where $\sigma$ is the vacuum expectation value of the non trivial minimum (assuming the false vacuum is at the origin in field space) and $\Lambda$ is the scale of the potential.

### 4.2.3. Finite number density contributions

Consider a scalar field, $\phi$, with a Global $U(1)$ symmetry which corresponds to a Noether charge $Q$. For a charge density $n = Q/V$ the effect of a non-zero charge density for $T > m$ is [236, 250, 251]

$$V(n,T,\phi) \sim V(\phi,T) - \frac{1}{2} n^2 (M^{-1})_{ab} \phi^b$$

$$\sim \frac{n^2}{2 \lambda (3 |\phi|^2 + T^2 / 2)} + \frac{\lambda}{4} T^2 |\phi|^2 + \frac{\lambda}{4} |\phi|^4,$$

where we have indicated how to generalize to more complicated models in the first line of the above equation. If the number density scales with the temperature cubed then the potential actually grows a minimum at high temperature. In this case the symmetry breaking is not caused by the microphysics of the field’s couplings and mass, but is instead caused by the macroscopic conditions.

### 4.3. Topical theoretical issues

#### 4.3.1. Gauge invariance

The gauge dependence of the effective potential has been the subject of much debate. Some have approached the issue by arguing for the benefits of a particular gauge (usually the Landau gauge because of the simplifications that there is no mixing between longitudinal vector boson and goldstone modes and that the gauge fixing parameter is not renormalized). Others have proposed novel solutions [252, 253]. A couple particularly creative ones involve coupling the source to a composite field [252]. Unfortunately [254] argued that one cannot calculate a finite effective potential this way if there are more than three space time dimensions. Another approach was to demonstrate that a gauge independent effective potential can be derived via a field redefinition [253]. While ingenious, this approach invites the criticism that an effective potential that is sensitive to field redefinitions is not an improvement on one that is sensitive to gauge transformations. Recent work by Schwarz et al [255] argued that the effort to produce a gauge invariant version of the effective action was misguided. They argue that the effective action itself is unphysical as its construction involves a test of how the system responds to an external source, $J$. If $J \neq 0$ this is a non-dynamical background charge density. This charge density does not couple to the gauge bosons which means Ward identities are violated [255]. The $J = 0$ case means that $\phi_0$ is extremal. They then argue that all physical quantities are gauge invariant, demonstrating that one can write the minimum of the potential to two loops if one expresses things in a gauge invariant scale, $\mu_X$. The trick is then to do an expansion in $\hbar$ rather than the usual loop expansion as such an expansion manifestly conserves gauge invariance order by order whereas the usual perturbative expansion fails in this regard. Explicitly one has in a $\hbar$ expansion [255, 256]

$$V = \nu_0 + h \nu_1 + h^2 \nu_2 + \cdots$$  

$$V = \nu^{(\text{LO})} + h \nu^{(\text{NLO})} + h^2 \nu^{(\text{NNLO})}.$$  

One can then express the minimum of the potential, itself a physical quantity, in terms of a physically meaningful scale, $\mu_X$ by grouping terms together in a $\hbar$ expansion as follows [255, 256]

$$V_{\min} = \nu^{(\text{LO})(\mu_X)} + h \nu^{(\text{NLO})(\mu_X)} + \nu^{(\text{NNLO})(\mu_X)}$$

$$V = \nu^{(\text{LO})(\mu_X)} + \nu^{(\text{NLO})(\mu_X)} + \nu^{(\text{NNLO})(\mu_X)} + h \nu^{(\text{LO})(\mu_X)} + h^2 \nu^{(\text{NNLO})(\mu_X)}.$$  

At finite temperature, one is often interested in calculating the order parameter as a measure of the strength of the phase transition. Explicitly it is the ratio of the vev at the non-trivial minima to the temperature at which the potential has degenerate minima, $T_C$. One approach was to expand $T_C$ in a $\hbar$ expansion. Such an approach suffers from infrared divergences. Another approach, is to expand both the minima and the potential evaluated at the minima in a $\hbar$ expansion. Such an approach seems
sufficient for calculating the physical sphaleron energy which is the true quantity controlling whether the yield of any particle produced during a phase transition is washed out.

4.3.2. Model dependence of the order parameter. A popular measure of the strength of the phase transition is through the order parameter \( \phi_C/T_C \) where the critical temperature, \( T_C \), is the temperature at which the minimum is degenerate and \( \phi_C \) is the value of the non-trivial minimum at the critical temperature. The order parameter is not gauge invariant and a gauge invariant treatment is given in [256]. Generally many just use the Landau gauge to calculate the order parameter. In baryogenesis one needs the phase transition to be sufficiently strong such that the phase with broken electroweak symmetry has electroweak sphalerons sufficiently suppressed such that a sufficient percentage of any baryon asymmetry produced through CP violating interactions with the bubble wall is preserved and not washed out [256]. A rule of thumb that gets used in the literature is that the phase transition is sufficiently strong if \( \phi_C/T_C \geq 1 \) [239].

A more precise condition involves calculating how much initial baryon asymmetry is preserved during the transition. However, the sphaleron rate depends on the sphaleron energy and the fluctuation determinant. Both of these are very model dependent and even in the standard model with a variable higgs mass the true condition can range from \( \phi_C/T_C > [0.7, 1.5] \) [256, 257].

4.3.3. Imaginary part of effective potential. At both zero and finite temperature the loop corrections to the effective potential is not real everywhere. This occurs when the mass squared for some values of the classical field the mass squared of the physical Higgs and goldstone bosons can be negative leading to complex logarithms in the Coleman Weinberg potential as well as complex contributions from the thermal functions. Furthermore daisy contributions also can be responsible for imaginary contributions. This leads to two related theoretical issues: first the effective potential is convex by construction and yet a negative mass squared appears to contradict this, second what is the physical interpretation of the imaginary components of the effective potential.

The first theoretical issue is known as the convexity problem where the effective potential is convex by construction and yet we frequently encounter effective potentials which are definitely not convex everywhere. The solution to this problem is found in merely bringing clarity to what it means to say the effective potential is convex by construction [258, 259]. The effective action is derived as the functional Legendre transform of \( W[J(x)] = -\ln(0^+|0^-)_J \). This implies that the effective action is concave and the effective potential is convex. This follows from the definition of a Legendre transform, \( L(p) = \min[xp - f(x)] \), which can be written as, \( xp - f(x) \), where \( x(p) \) satisfies the conditions that \( \partial f(x) = p \) and \( \partial^2 f(x) \leq 0 \). In other words the Legendre transform of \( f \) is concave by definition. Now the effective action is the Legendre transform, \( \Gamma[\phi] = \min[J \phi - W(J)] \), which is concave. When one calculates the effective action by summing 1PI diagrams one finds a non-convex effective potential. \( \Gamma[\phi] \). However one does not require, \( \Gamma[\phi] = \Gamma[\bar{\phi}] \), the latter is the concave envelope of the former (and \( V(\phi) \) is the convex envelope of \( V(\phi) \)). The two are equivalent for a constant background evaluated at the absolute minimum.

The second issue is a little more subtle. It was shown by Weinberg and Wu [260] that the imaginary contributions to the effective potential have the interpretation of decay processes. The decay in question is not the scalar fields decaying into other particles. This can be demonstrated from the fact that even a theory with a single scalar field that has no decay modes still obtains imaginary contributions. The decay also is not the non-perturbative decay of the false vacuum as these imaginary contributions are perturbatively derived. In fact, the decay corresponds to an instability in the system where fluctuations around the classical field become large in a set of uncorrelated domains of size \( (V''(\phi))^{-1/2} \). Within each domain the fluctuations grow exponentially with time and the system becomes unstable. This instability becomes important when the false vacuum is decaying and one needs to expand around a space time varying background that includes regions of negative field curvature. If the imaginary part of the effective potential is large compared to the real part then the system is unstable and the usual process of calculating bubble nucleation may be invalid.

This is best understood in direct analogy with the Schwinger effect in electromagnetism\(^6\). If the electric field in some volume, \( V \), is strong enough, electron anti-electron pairs will be spontaneously produced via interactions between the vacuum and the strong electric field. These electron anti-electron pairs will split, aligning themselves with the background electric field which lowers both the background and the energy of the system. The Schwinger effect can be formally understood in terms of effective actions. Defining the one loop correction to the Lagrangian density as \( L_1 \) one can write the pair creation rate per unit volume and time as [261]

\[
\Gamma = -2L_1 = \frac{1}{4\pi^3} \left( eE \right)^2 \sum_{n=1}^{\infty} \frac{1}{n^2} \exp \left[-\frac{n\pi m^2}{|e|E} \right].
\]

Similarly the imaginary components of the effective potential correspond to spontaneous production of scalar quanta. This percolation serves to drive field to the inflection point lowering the total energy to a point that is higher than the minimum. Note that this lowering of the energy is a purely quantum mechanical effect that is different from a classical roll. Thermal corrections are also responsible for an imaginary component to the finite temperature version of the effective potential. These components arise from the fact that we have calculated the effective potential under the assumption of equilibrium and if there is an imaginary component it means that you have a thermal instability to your equilibrium state. The thermal imaginary components will then proceed to take the system out of equilibrium.

It was recently shown that for the standard model case the imaginary parts of the ring sum term effectively cancel the imaginary parts of the one loop corrections

---

\(^6\) We thank Hiren Patel for clarifying this issue.
guaranteeing the stability in this case. Explicitly one has in the high temperature limit [262]

\[
\text{Im}[V_i(T)] \to \sum_{i \in \{A,GB\}} \Theta(-m_i^2) n_i \left[ \frac{|m_i|^4}{64\pi} + \frac{|m_i|^3 T}{12\pi} \right] + \sum_{i \in \{A,GB\}} \Theta(-m_i^2) n_i \left[ \frac{|m_i|^4}{64\pi} \right],
\]

\[
\text{Im}[V_{\text{Ring}}(T)] \to \sum_{i \in \{A,GB\}} \Theta(-m_i^2) n_i \left[ \frac{|m_i|^4 T}{12\pi} \right].
\]

At high temperature these contributions can cancel for the standard model case (apparently also do at \(m \sim T\) which is not much of a surprise since the high temperature limit holds very well until \(m \sim 2T\)). So for the standard model case at least this presents no issue at least in terms of the stability of the system during a phase transition. However, one should note that the oscillation occurs only in the finite temperature expansion and small imaginary component remains even at the origin when electroweak symmetry is restored.

4.3.4. Back reaction of the soliton. To self consistently calculate the tunneling rate which is relevant for calculating various thermodynamic parameters, one needs to include the correction that is due to the fact that one has a space time varying field configuration. Recent works [263–265] have addressed this issue and suggest the following recipe

1. find the approximate bounce solution to the classical equations of motion
2. Insert the bounce solution into the equation for the Greens function and find the new greens function that solves for the case of a \(\phi^4\) theory with quartic term \((\lambda/4!:\)

\[
(-\Box + \frac{\lambda}{2} \phi^2) G(x,y) = \delta(x-y),
\]

where as before \(G\) is a propagator.
3. Calculate the tadpole corrections \(\Sigma_R\) renormalizing in the homogeneous false vacuum
4. Insert tadpole into the equations of motion

\[
-\Box \phi + \Sigma_R \phi = 0,
\]

where \(\Sigma_R = \lambda S(\phi) + \delta \Sigma\) and \(\delta \Sigma\) is contains all the relevant counter terms. Solve the bounce which now solves this corrected equation of motion.
5. Repeat steps 2–4 until one has convergence.

The corrections were found to be very small in the thin wall regime [264, 265] but are expected to be more relevant when one is beyond the thin wall limit.

5. Examples of phase transitions

5.1. Electroweak phase transition

No other phase transition gains as much attention as the electroweak phase transition [13–17, 20, 21, 223, 247, 256, 267–273]. Reheating models generically tend to predict a reheating temperature high enough to restore electroweak symmetry\(^7\). Furthermore, if electroweak symmetry is broken during the cooling of the Universe after reheating via a strongly first order electroweak phase transition, the baryon asymmetry of the Universe can be generated during the phase transition (at the small cost of diluting thermal relics [275]). Let us begin with the standard model to understand why it does not accommodate a strongly first order phase transition and how to extend the SM to catalyze such a scenario. We will follow the conceptual organization of [266] which categorized the different classes of extensions to the Standard model that can successfully accommodate a strongly first order electroweak phase transition. The Higgs potential in the standard model in the high temperature expansion expressed in the Landau gauge is given by [239]

\[
V = D(T^2 - T_0^2) h^2 - ETH^3 + \frac{\lambda}{4} h^4,
\]

where [239]

\[
D = \frac{96}{3} \frac{(g_1^2 + 3g_2^2 + 4g_1^2 + 8\lambda)}{32},
\]

\[
T_0 = \frac{m_1^2}{4D},
\]

\[
E = \frac{3}{96\pi} \left( 2g_1^2 + (g_1^2 + g_2^2)^{3/2} \right),
\]

where in the above \(g_1\) and \(g_2\) are the standard model gauge boson couplings. The strength of the phase transition is given by [239]

\[
\frac{\phi_c}{T_c} = \frac{2E}{\lambda} = \frac{4E}{m_{h1}^2} \simeq \frac{2300}{m_{h1}^2}.
\]

So for a Higgs mass of 125 GeV one has a very weak first order phase transition with an order parameter \(\sim 0.1\). In reality, lattice simulations indicate that the electroweak transition is a smooth crossover. To boost the strength of the electroweak phase transition there are four model classes to achieve this which are depicted in figure 10. These are

I. Boost the effective \(E\) parameter in equation (132) by a factor of at least 5. This can only be achieved through the introduction of new bosonic degrees of freedom that acquire a part of their mass through electroweak symmetry breaking. Also the total mass of the new boson cannot be too heavy as the cubic term is only manifest when the high temperature expansion is valid—that is when the temperature is large or comparable to the mass. Above such a mass the thermal contribution from such a boson is heavily Boltzmann suppressed. The most celebrated example of such an approach is the light stop scenario [276]. Such a scenario is very efficient as the contribution to \(E\) has a multiplicity factor of 12. However, the light stop scenario is highly constrained as it requires a stop lighter than the SM top quark and it is difficult, but not impossible, for such a light stop to evade detection

\(^7\) One caveat is that it is possible through the introduction of many singlets for electroweak symmetry to not restore at high temperature [274].
Another possibility that is equally efficient and can evade detection is that of folded supersymmetry [278] where the SU(3)_C quantum numbers of the stop is not the standard model colour. It is also in principle possible to boost the value of $E$ through light scalars fields.

IIA. The second scenario attempts to introduce a tree level effective cubic term to provide a barrier between the true and false vacuum during the phase transition (a barrier that can persist at zero temperature). Such an operator is forbidden due to gauge invariance unless there are additional scalar fields [279]. Such a scalar field must have their vev also substantially change during the phase transition.

IIB. In this scenario the barrier between the true and false vacuum is created by the effective quartic changing signs and the vacuum is stabilized by the non-renormalizable sextet term. Such a theory can be an effective theory that is valid up to the cutoff scale $\Lambda$. The scale of new physics needs to be relatively low compared to the standard model—between about 500 and 800 GeV—in order to catalyze a strongly first order electroweak phase transition [233]. If the cutoff is too low then the tunneling probability becomes large compared to the age of the Universe. If the cutoff is too high then the effect of the new physics is too feeble to catalyze a strongly first order electroweak phase transition. Recent work has demonstrated that the dimension eight operators are also important for the electroweak phase transition [280].

Figure 10. Model classes for catalyzing a strongly first order electroweak phase transition. From top left to bottom right the barrier is caused by thermal loops, tree level triscalalar interactions, non-renormalizable operators and Coleman Weinberg corrections, respectively. Reprinted figure with permission from [270], Copyright (2013) by the American Physical Society.

III. Perhaps the least explored option of the four is to induce a large contribution from the Coleman Weinberg potential to catalyze a strongly first order electroweak phase transition [277]. The dependence on the dimension six and dimension eight Wilson Coefficients we show in figure 11.

Figure 11. Values of Wilson coefficients for $c_6/(8f^2)h^6$ and $c_8/(16f^4)h^8$ operators with LISA and future $e^+e^-$ circular collider (FCC-ee) constraints given. Note $\lambda_3$ refers to the usual Higgs coupling and FCC-ee can probe values outside the interval $\lambda_3 \in [0.1, 1.9]$. Reproduced from [284], CC BY 3.0.
transition. For instance in the case where one has a large number of inert scalar singlets (say 12 or more) the contribution to the Coleman Weinberg potential can be large enough to catalyze a strongly first order electroweak phase transition. A more recent paper achieved this with the addition of two fermion fields [281].

On top of these possibilities some more exotic possibilities include having cosmologically varying Yukawa couplings [282] or cosmologically varying the gauge coupling such that a strongly first order EWPT is catalyzed by a QCD transition at a higher scale [283].

With the light stop scenario becoming more constrained, most phenomenology research focuses on IIA and IIB scenarios. Both of these types of phase transition often causes substantial supercooling which also implies a larger gravitational wave signal as we will see in forthcoming sections. For type IIB scenarios the Higgs quartic being negative is a generic consequence of fixing the mass and vacuum expectation value to agree with experiment in the presence of a sizable positive Wilson coefficient for a $H^6$ operator. Some examples of extended scalar sectors that generate this operator are given in table 3. Note from the form of the Wilson coefficients given in this table it is of course not guaranteed it has the needed positive sign and some models more easily accommodate this than others.

For scenarios of type IIA one requires an effective trilinear coupling to provide a barrier between the true and false vacuum. For a second field, $\phi$, the electroweak phase transition proceeds along the field space path $(0, v_{\phi}) \to (v_{\phi}, v'_\phi)$ where in principle $v_{\phi}$ can be negligible. Rotating and shifting to field space coordinates $\varphi = a h + b (\phi + v_{\phi})$ and $\vartheta = - b h + a (\phi + v_{\phi})$ where $a$ and $b$ are chosen so that in the rotated and shifted coordinates the phase transition proceeds as $(0, 0) \to (0, v_{\phi})$. Even though a term of the form $h^3$ is forbidden by gauge invariance, if a trilinear coupling between $h$ and $\phi$ exists then in the rotated coordinates this leads to a term of the form $\varphi^3$. As an example consider the real singlet, $S$, with potential [231]

$$V = -\mu^2 |H|^2 + \lambda |H|^4 + \frac{d_1}{2} |H|^2 S + \frac{d_2}{2} |H|^2 S^2 + b_2 S^2 + b_3 S^3 + b_4 S^4 . \tag{135}$$

After rotating to the coordinates $v_\tau = \varphi(T) \sin \alpha(T)$, $v_\phi/\sqrt{2} = \varphi(T) \cos \alpha(T)$ and ignoring the resulting linear term (which means ignoring the existence of a high temperature singlet vev) one has [231]

$$V(T, \varphi) = D(T^2 - T_0^2) \varphi^2 + E \varphi^3 + \lambda' \varphi^4 , \tag{136}$$

with

$$D = \frac{g_1^4 + 3 g_3^4 + 4 g_1^2 + 8 \lambda}{32} ,$$

$$T_0 = -\frac{\mu^2 \cos^2 \alpha - \frac{b_2}{2} \sin^2 \alpha}{D} ,$$

$$E \sim \frac{d_1}{2} \cos^2 \alpha \sin \alpha + \frac{b_3}{3} \sin^3 \alpha ,$$

$$\lambda' = \lambda \sin^4 \alpha + \frac{a_2^2}{2} \sin^2 \alpha \cos^2 \alpha + \frac{b_4}{4} \sin^4 \alpha . \tag{137}$$

where we remind the reader that $\alpha$ is the angle of the phase transition in field space. In the above we have ignored the small corrections to the effective cubic term due to the gauge bosons. Note that the trilinear couplings $a_1$ and $b_2$ enter directly into the effective cubic term. To generate a large enough effective cubic to catalyze a strongly first order electroweak phase transition one usually has $a_1$ as quite sizeable, $-1000 \text{ GeV} \lesssim a_1 \lesssim -100 \text{ GeV}$. In order to comply with LHC constraints on the zero temperature mixing angle between the singlet and Higgs, one requires that the other portal coupling $a_2$ be large and anti correlated with $a_1$ to suppress the mixing angle. For sub TeV single mass, current constraints on the mixing angle are $|\sin \theta| \lesssim 0.2$ [285, 286] with this bound expected to tighten with future colliders [232, 287].

For both IIA and IIB type phase transitions, one can have a barrier between the true and false vacuum that is so large at zero temperature that the false vacuum decay rate is never fast enough compared to Hubble for the phase transition to proceed. A recent proposal demonstrates that one can have the QCD transition reduce the barrier between the true and false vacuum [288]. Specifically in a Randall Sundrum model the radion potential acquires a contribution from gluon condensates. The contribution is negative and becomes important near the origin thus it removes some of the barrier between true and false vacua. Therefore when the gluons form a condensate electroweak symmetry breaking can occur. Thus the
Electroweak phase transition could occur at a much lower scale than usual. Alternatively, it was recently shown that if the electroweak phase transition occurs in two steps, the scale of the electroweak phase transition can be multi-TeV [289].

5.1. QCD phase transition: an example of fermion condensation. The QCD phase transition generally occurs when the temperature of the Universe is 170 MeV assuming no significant baryon chemical potential in the early Universe. The transition is caused by the temperature evolution of the strong coupling constant \( g_s \). At temperatures above the transition temperature the coupling constant is small enough to treat the system perturbatively and the system is an a phase of quark-gluon plasma. As the Universe cools the strong coupling constant grows and quarks and gluons hadronize into colour neutral objects. All colour multiplets are confined to exist then within colour singlet objects such as baryons. Since perturbation theory breaks down during the QCD phase transition, they are best analyzed through lattice simulations. The phase diagram of QCD is shown in terms of temperature and baryon chemical potential in figure 12 which is taken from [290]. Some intuition can be obtained through the bag model [291], for a review see [292].

Although vanilla cosmology would predict that QCD underwent a crossover transition, there have been some recent proposals to catalyze a strongly first order phase transition. One approach is to delay the electroweak phase transition until after the QCD phase transition such that the number of light quarks is large enough for the transition to be strongly first order [293]. The quark nuggets that form during such a transition are a dark matter candidate [293]. Another approach is to take advantage of the fact that the lepton asymmetry is relatively unconstrained. A large enough lepton asymmetry catalyze the QCD transition [294]. Such a phase transition could leave the signature of observable low frequency gravitational waves [298–300].

5.1.2. An example of a multistep phase transition. In this section we briefly give an example of a multi-step phase transition. We will focus on the case where a zero temperature symmetry, SU(3)c in particular, is broken in an intermediate phase before being restored. This can be achieved either by having a large number density or introducing new colored scalars which acquire a vacuum expectation value at an intermediate temperature. Let us consider the latter case. Consider an effective potential that includes the standard model Higgs coupled to a colored scalar field \( C \), which is an SU(2)\(_L\) and U(1)\(_Y\) singlet but a triplet under SU(3)c:

\[
V = -\mu^2 H|H|^2 - \rho_C|C|^2 + \lambda_{HC}|H|^2|C|^2 + \lambda_H \frac{H^4}{2} + \lambda_C \frac{|C|^4}{2}.
\]

(138)

For simplicity let us only include thermal corrections which are quadratic in temperature and fields. Symmetry is restored in the \( H \) and \( C \) directions above a temperature \( T_{f}^{H} \) where [289]

\[
(T_{f}^{H})^2 = \left(\frac{\lambda H v^2}{\left(\frac{\lambda_H}{4} + \frac{\lambda_C}{4} + \frac{\lambda}{16} + \frac{\lambda_H^2}{16} \frac{1}{T_f^H} + \frac{\lambda_C^2}{16} \frac{1}{T_f^H} \right)} \right).
\]

(139)

\[
(T_{f}^{C})^2 = \left(\frac{\mu_C^2}{\frac{\lambda}{4} + \frac{\lambda_C}{4} + \frac{\lambda}{16} + \frac{\mu_C^2}{16} \frac{1}{T_f^C} + \frac{\lambda_C^2}{16} \frac{1}{T_f^C} \right). \]

(140)

If one has \( T_{f}^{H} < T_{f}^{C} \) then one has a range of temperatures where color can be spontaneously broken but electroweak symmetry is restored. In figure 13 we show in the top panel from [289] the possibility of color breaking and restoration. Note that this scenario is not particularly fine tuned however the mass range of the colored scalars is quite light. The addition of a gauge singlet allows the mass of the colored scalar to be multi-TeV.

5.2. Topological/non-topological defects and solitons

5.2.1. Topological defects. The phase transition can also give rise to topological defects [3, 117, 130, 301–304], for a review see [305]. Let us discuss briefly the microscopic origin of the formation of topological defects. Let us suppose that there is a non-trivial charge for \( \psi \) field under some gauge symmetry \( G \), and then \( \psi \) field obtains a non-vanishing VEV due to phase transition, the symmetry group is broken now; \( G \rightarrow H \). The manifold of all the vacua accessible to \( \psi \) is then given by the quotient group after breaking, i.e. \( M = G/H \). As an example, in the case of an abelian Higgs model, the symmetry breaking pattern is very simple \( U(1) \rightarrow I \), and the manifold of vacua is \( M = U(1) \), corresponding to the circle of constant radius in the complex plane \( |\psi| = \text{constant} \). Therefore, the formation and the type of topological defects depend on the topological properties of \( M \) [117, 130], which is classified by the homotopy groups \( \pi_n \) of order \( n \). Each group \( \pi_n(M) \) is composed of all classes of hyper surfaces of dimension \( n \). If any hyper surface can shrink to a point inside \( M \), then the homotopy group contains only one element and becomes trivial and is simply connected [306]. In the opposite case, if \( M \) is not simply connected (for example during the breaking of a discrete group \( Z_n \rightarrow I \)), uncorrelated regions of the Universe would have different vacua separated by the domain walls [130, 301]. The domain walls with tiny energy scale may yield some interesting cosmological consequences [307, 308], including mild acceleration of cosmic expansion [309]. If their energy scale is high, and if they persist in the late Universe, they would simply cause cosmological disasters over dominating the energy density of the Universe [307]. There are ways to tackle the problem if we change the nature of phase transition from \( G \rightarrow H \) to a smooth adiabatic transition [310]. Note that defects can also be formed in a slow first order phase transitions [311].

The formation of topological defects also depends on the space-time dimensions, a \( d \)-dimensional defect is governed by the non-triviality of the homotopy group [306]:

\[
\pi_{3-d} \neq I. \quad (141)
\]

Any symmetry breaking of the form \( G \rightarrow H \times U(1) \) gives rise to monopole (point-like defects). Since the standard model group contains the \( U(1) \) factor, any GUT group breaking

\[\text{review} \]
down to the standard model gauge group leads to this monopole problem. This formation of unwanted defects was one of the original motivation to introduce a phase of primordial inflation.

There is also a class of unstable topological defects which can form even when the topology is trivial [3, 302]. The electro-weak strings can be formed during the electroweak symmetry breaking which are perturbatively stable for a range of parameters which are not realized in nature. In general, the defects are a priori unstable due to plasma effects.

5.2.2. Non-topological solitons. The phase transition can also yield non-topological solitons, such as \( Q \)-balls. The \( Q \)-ball becomes a generic ground state in interacting scalar fields carrying some conserved global charge [132, 312, 313], whose boundary condition at infinity is the same as that for the vacuum state, unlike in the case of topological solitons such as magnetic monopoles [314, 315]. A detailed review of non-topological solitons can be found in, e.g. [132]. Formation of \( Q \)-balls can be extended to many scalar fields with various \( U(1) \) charges [316, 317], with a non-Abelian symmetries [318], and also with local gauge symmetries [319]. The main difference between global and local \( Q \)-balls is that in the latter case the charge of the stable \( Q \)-ball is bounded from above.

There is a theorem [132, 313], which states that if there exists a range for a field \( \phi \), in a potential \( U(\phi^2) \), which contains an attractive interaction, then a non-topological soliton should exist for

\[
\nu^2 < \omega^2 < m_\phi^2,
\]

where \( U(\phi^2) \to m_\phi^2 \phi^2 \) when \( \phi \to 0 \). The value of \( \omega = \sqrt{k^2 + m_\phi^2} \) determines the frequency of the \( \phi \) quanta in the field space. A necessary condition for the existence of a solitonic solution is \( \nu^2 < m_\phi^2 \), which means that there exists a parabola \( \nu^2 \phi^2 \) tangent to \( U(\phi^2) \) at \( \phi = \pm \phi_0 \), with \( \nu^2 < m_\phi^2 \).

For a sufficiently large \( Q \), the energy of a soliton is then given by

\[
E = |\nu|Q < m_\phi|Q|,
\]

which ensures its stability against decay into plane wave solutions with \( \phi \simeq \phi_0 \) inside the \( Q \)-ball, and \( \phi \simeq 0 \) outside. The global \( U(1) \) symmetry is broken inside and remains unbroken outside.

The \( Q \)-balls can be formed after inflation, as we had discussed earlier, but can be formed at later stages by the dynamics of a scalar field, such as present in supersymmetric theories due to plenty of supersymmetric flat directions, made up of squarks and sleptons, for a review see [133, 320]. The stability of the \( Q \)-balls can contribute to the dark matter abundance [185], see review [133]. During the formation, gravitational waves can also be generated [125, 126].
6. Phase transitions and cosmic signatures

6.1. Gravitational waves

First order cosmological phase transitions proceed via bubble nucleation. While an isolated spherical bubble produces no gravitational waves as such an event has no quadrupole moment, the violent process of bubble collision does [18, 30, 248, 322–334], for a review of gravitational waves, see [335]. Upon collisions of such bubbles, the latent heat will be converted to bulk flow of the plasma, as well as to kinetic energy of the scalar fields. The fraction of energy converted to gravitational waves per decade is

$$\Omega_{GW} = \omega \frac{dE_{GW}}{d\omega} \frac{1}{E_{\text{in}}},$$

(144)

where

$$\frac{dE_{GW}}{d\omega} = 2G\omega^2 \Lambda_{ij,lm}(\hat{k}) \tilde{T}_{ij}(\hat{k}, \omega) \tilde{T}_{lm}(\hat{k}, \omega),$$

(145)

where $\tilde{T}_{ij}$ is the fourier transform of the stress energy tensor and

$$\Lambda_{ij,lm} = \delta_{il} \delta_{jm} - 2k_i k_j \delta_{lm} - \frac{1}{2} \delta_{ij} \delta_{lm} + \frac{1}{2} \delta_{ij} \delta_{kl} + \frac{1}{2} \delta_{lm} \delta_{ki}.$$

(146)

is a projection operator. The contributions to the gravitational wave spectrum can be modeled as a sum of three contributions characterized by a contribution to the stress energy tensor and an efficiency parameter $\kappa$, which denotes the efficiency that the latent heat can be converted into a particular source of gravitational waves. The three contributions to the stress energy tensor are as follows
1. A contribution from the initial collision of scalar field shells. The stress energy tensor contribution is

$$T_{\mu\nu} = \partial_{\mu}\phi\partial_{\nu}\phi - g_{\mu\nu} \left[ \frac{1}{2} \partial_{\rho}\phi\partial^{\rho}\phi - V_{0} \right].$$

(147)

2. The interaction between kinetic shells going at the speed of sound [336]. The stress energy contribution is [337]

$$T_{\mu\nu} = \sum_{i} \int \frac{d^{4}k}{(2\pi)^{3}2E_{i}} k_{\mu}k_{\nu}f_{i}(k),$$

(148)

where $v_{i}$ is the three velocity of the relativistic fluid, $\gamma$ is the Lorentz factor, $\epsilon$ is the energy and $p$ is the pressure.

3. A contribution due to magnetohydrodynamic turbulence [38] which again is prominent after the collision of the scalar shells. This contribution is usually subdominant. The spatial components of this contribution to the stress energy are

$$T_{ij}(x, \eta) = \frac{4\pi}{3} \rho(\eta)v_{i}(x, \eta)v_{j}(x, \eta),$$

(149)

where $v_{i}$ is the turbulent velocity and $\eta$ is the conformal time.

Recent work has suggested the existence of a fourth contribution from quantum fluctuations in bubble wall collisions [338–340]. They considered a double well potential and demonstrated that quantum fluctuations break the SO(2,1) symmetry of bubble wall collisions. One has a parametric instability and wiggles on the wall from the collision and break SO(2,1). The size of this contribution relative to LISA sensitivity is an open problem and we therefore focus on the contributions to the gravitational wave spectrum that are better understood. The total gravitational wave spectrum can then be modeled as a sum of the three contributions [337]

$$\Omega(f)h^{2} = \Omega_{\text{col}}(f)h^{2} + \Omega_{\text{sw}}(f)h^{2} + \Omega_{\text{turb}}(f)h^{2}.$$  

(150)

The change in free energy between during the phase transition gives a limit to how much vacuum energy can be converted into gravitational waves. The efficiency of converting vacuum energy into scalar field gradient energy is denoted by $\kappa_{\text{col}}$ (the first contribution in the above list) controls the efficiency of producing contributions to $\Omega_{\text{col}}$. The efficiency parameter, $\kappa_{\text{col}}$, is typically small, making this contribution sub-dominant. Specifically it be found by calculating the gradient density $\rho_{D} = \frac{1}{2}(\nabla\phi)^{2}$, and potential energy density $\Delta V(\phi)$ for a bounce solution [341]

$$\kappa_{\text{col}} = \frac{2\rho_{D}}{\rho_{\text{vac}} - \rho_{\text{V}}}.$$  

(151)

Ignoring the turbulence contribution, the conservation of energy and momentum gives

$$\partial_{\mu}T_{\mu\nu}^{\text{plasma}} + \partial_{\mu}T_{\mu\nu}^{\text{field}} = 0.$$  

(152)

One can parametrize the plasma contribution as follows [321]

$$T_{\mu\nu}^{\text{plasma}} = w_{u}u_{\mu}u_{\nu} - g_{\mu\nu}p_{u}.$$  

(153)

where $u = (\gamma, \gamma v)$ is the four velocity field of the plasma. If we ignore the field contribution we can calculate the fluid velocity from the equations $\partial_{\nu}T^{\nu} = \partial_{\nu}T^{\text{sw}} = 0$ from which one obtains [321]

$$w_{+}v_{+}^{2} + w_{-}v_{-}^{2} = \gamma_{+}^{2},$$  

(154)

and

$$w_{+}v_{+}^{2} = w_{-}v_{-}^{2} - \gamma_{+}^{2},$$  

(155)

where $\pm$ denotes the symmetric and broken phases respectively. Defining [321]

$$a_{+} \sim \frac{\pi^{2}}{30} \sum_{i} (N_{ri}^{i} + \frac{7}{8} N_{ri}^{\text{f}}),$$  

(156)

one can define an expression for the fluid velocity [321]

$$v_{+} = \frac{1}{1 + a_{+}} \left[ \frac{v_{-}}{2} + \frac{1}{6a_{+}} \pm \sqrt{\left( \frac{v_{-}}{2} + \frac{1}{6a_{+}} \right)^{2} + a_{-}^{2} \left( \frac{2}{3}a_{+} - \frac{1}{3} \right)} \right].$$  

(157)

a detonation has $v_{+} > v_{-}$ and deflagration is $v_{-} < v_{+}$. The latter only exists when $a_{+} < 1/3$. In a detonation the wall moves at supersonic speed and the plasma it expands into is at rest. In contrast, a deflagration has the wall expanding into the perturbed plasma. Simulations show that the efficiency coefficient for a deflagration (wall velocity smaller than the speed of sound) is [321]

$$\kappa_{\text{sw}} = \frac{6.9v_{u}\alpha}{1.36 - 0.037\sqrt{\alpha} + \alpha},$$  

(158)

whereas for detonations (runaway walls) one has

$$\kappa_{\text{sw}} = \frac{0.73 + 0.083\sqrt{\alpha} + \alpha}{\alpha}.$$  

(159)

Here $\alpha$ is the ratio of Latent heat to vacuum energy and $v_{u}$ is the wall velocity. These thermodynamic quantities are defined in section 3.2. Alternatively if one knows the fluid radial velocity profile, $V_{r}(\xi)$ one can explicitly calculate the efficiency as [337]

$$\kappa_{f} = \frac{3}{c_{p}v_{u}^{3}} \int d^{2}\omega(\xi) V_{r}^{2}(\xi) \gamma^{2}\xi^{2}.$$  

(160)

Here $\xi = r/t$ and $\omega$ is the enthalpy. If the bubble wall does not runaway, the sound wave and turbulence terms are expected to dominate. If the bubbles runaway, the collision term becomes more important and in fact dominates for very large $\alpha$.

Apart from the efficiency parameters that define the efficiency of converting energy available to gravitational wave energy, the gravitational wave power spectrum is controlled by the ratio of Latent heat to vacuum energy, the bubble wall velocity and the speed of the phase transition compared to the Hubble rate $\beta/H$. As well as parameters that are numerically derived from analytical fits to numerical simulations.
6.1.1. Collision term. The interaction of the bubbles can be well approximated by the ‘envelope approximation’ [326] which is the combination of two approximations—first that the stress energy tensor is non-zero only in an infinitesimal region at the bubble wall and second that the stress energy tensor vanishes when the bubble overlaps. This contribution becomes most significant when the bubble runs away \( \gamma \to \infty \). This contribution can be derived analytically through calculation the correlation of the stress energy tensor \( \lambda T(x) T(y) \) [325]. Under the envelope approximation the stress energy tensor due to a bubble nucleated at \( x_N = (t_N, x_N) \) is given by

\[
T_{ij} = \rho(x)(x - x_N), (x - x_N),
\]

with the energy density localized around the bubble wall in accordance with the envelope approximation [325]

\[
\rho(x) = \begin{cases} 
\frac{4\pi}{3} \left( \frac{\dot{v}(t)}{\dot{v}\rfloor_{t=0}} \right)^3 & \text{if} \quad |x - x_N| < \dot{v}(t) \\
0 & \text{otherwise.}
\end{cases}
\]

(161)

Here \( r(t) = v(t - t_N) \) and \( \dot{r}(t) = \dot{v}(t) + l_B \) are the interior and exterior edge of the bubble wall respectively and \( \rho_0 \) is the latent heat released by the transition. The nucleation rate is

\[
\text{If the phase transition is sufficiently quick, } \beta / H \gg 1 \text{ one can ignore the expansion of the Universe and write the metric as}
\]

\[
ds^2 = -dt^2 + (\delta_{ij} + 2h_{ij})dx^i dx^j.
\]

(162)

From the equations of motion the tensor perturbations satisfy the following

\[
\ddot{h}_{ij}(t, k) + k^2 h_{ij}(t, k) = 8\pi G \Pi_{ij}(t,k),
\]

(164)

where \( \Pi_{ij} \) is related to the fourier transform of the stress energy tensor via a projection operator

\[
\Pi_{ij}(t,k) = \Lambda_{ij,kl} T_{kl}(t,k).
\]

(165)

The tensor perturbations can be solved in terms of a Greens function.

The total energy of the gravitational waves is given by the oscillation and ensemble average of the correlator [325]

\[
\rho_{GW} = \frac{\langle \ddot{h}_{ij}(t,x) \dot{h}_{ij}(t,x) \rangle}{8\pi G}.
\]

(166)

From which we can derive the gravitational wave spectrum [325]

\[
\Omega_{GW} = \frac{1}{\rho_i} \frac{d\rho_{GW}}{dt} \ln k
\]

(167)

\[
= \frac{2Gk^3}{\pi \rho_{tot}} \int_0^{t_0} dt_i \int_0^{t_0} dt_j \cos(k(t_x - t_j)) \Pi(t_x, t_j, k),
\]

(168)

where \( \rho_i \) is the total energy \( \rho_0 + \rho_{rad} \) and \( \Pi(t_x, t_j, k) \) is the fourier transform of the stress energy tensor correlation function contracted with projection operators [325]

\[
\Pi(t_x, t_j, k) = \int d^3r e^{i \vec{k} \cdot \vec{r}} \Lambda_{ij,kl} \Lambda_{ij,mm}(T_{kl}(t_x, x) T_{mn}(t_j, y)),
\]

(169)

\[ r = x - y. \]

Defining the quantity \( \alpha = \rho_0 / \rho_{rad} \) and using the fact that \( H^2 = 8\pi G \rho_i / \alpha \), we can write

\[
\Omega_{GW} = \kappa^2 \left( \frac{\beta}{H} \right)^{-2} \left( \frac{\alpha}{1 + \alpha} \right)^2 \Delta(k/\beta, \nu_w),
\]

\[
\Delta(k/\beta, \nu_w) = \frac{3\beta^3}{4\pi^2 \kappa^2 \rho_0^2} \int_0^\infty dt_i \int_0^\infty dt_j \cos(k(t_x - t_j)) \Pi(t_x, t_j, k).
\]

(170)

The ratio of the scale factor at the time of transition to the scale factor today is \( a_0/a^* = 8 \times 10^{-16}(100/\text{GeV}) \) [100 GeV/100] which can be used to relate the gravitational wave spectrum at transition to its spectrum today [325]

\[
f = 1.65 \times 10^{-3} \text{Hz} \left( \frac{f}{100 \text{Hz}} \right) \left( \frac{\beta}{H} \right) \left( \frac{T}{100 \text{ GeV}} \right) \left( \frac{\rho}{100} \right)^{1/6} \]

\[ \times \frac{0.35}{1 + 0.069 \nu_w + 0.69 \nu_w^2} \]

\[
\Omega_{GW} h^2 = 1.67 \times 10^{-5} \left( \frac{\nu_w}{100} \right)^{-1} \kappa^2 \left( \frac{\beta}{H} \right) \left( \frac{\alpha}{1 + \alpha} \right)^2 \times \frac{0.48 \nu_w^2}{1 + 5.3 \nu_w^2 + 5.0 \nu_w}. \]

(171)

The dependence on \( \nu_w \) unfortunately comes from numerically fitting. All that remains is an analytical calculation of \( \Delta \). Such a calculation is difficult in practice however one can acquire a closed form solution in terms of integrals of spherical Bessel functions. The asymptotic form can be derived from the asymptotic expansions of the spherical bessel functions and one finds that \( \Delta \sim k^3 \) for \( k/\beta < 1 \) and \( k^{-1} \) for \( k/\beta > 1 \). Numerically fitting to the integral over Bessel functions for \( \nu_w \) close to unity one has for the frequency spectrum, one finds that \( \Delta \) is well approximated by [337]

\[
\Delta = \frac{\Delta_{peak}}{c_1 \left( \frac{f_{\text{peak}}}{f_{\text{sc}}(c_1 + c_2)} \right)^{-3} + (1 - c_1 - c_2) \left( \frac{f_{\text{peak}}}{f_{\text{sc}}(c_1 + c_2)} \right)^{-1} + c_2 \left( \frac{f_{\text{peak}}}{f_{\text{sc}}(c_1 + c_2)} \right)}.
\]

(172)

where fitting yields \( c_1 = 0.064 \) and \( c_2 = 0.48 \). Note that recent work analyzing a vacuum transition (that is, a case where the plasma is ignored) [341] demonstrated that the envelope approximation breaks down right when it starts to become visible and the true spectrum is dampened. This seems to imply that the collision term is always sub-dominant.

6.1.2. Sound waves. The contributions from the plasma flow are much harder to capture in a model. Moreover, recent studies indicate [342] that the plasma flow contributions dominate over the scalar field contributions, since the plasma flow continues to source GWs long after the collisions of the bubbles.

Progress in this area has been largely dominated by large-scale hydrodynamic simulations. Nevertheless, well-motivated simplified models have been developed recently, such as the recent bulk flow model [327] and sound shell model [343]. Such models may describe the physics in regimes in which simulations have limitations [344].

The sound wave contribution is typically larger than the other contributions. Its power spectrum is [342]
where $\nu$ is the turbulent velocity which is a random variable, 
\[
\nu_i(k, \eta) = \frac{v_i(k)}{V_i(k)e^{i\omega_\eta}} \quad \text{for } k < L_B^{-1},
\]
\[
\nu_i(k, \eta) = \frac{V_i(k)e^{i\omega_\eta}}{L_B^{-1} < k < L_K^{-1}},
\]
and $\omega_\eta$ is the frequency associated with an eddy of size $l = 1/k$.

The needed fourier transform of the stress energy tensor is as follows [38, 347–351]
\[
T_i(k, \eta) \equiv \frac{4\pi\rho(\eta)}{3} \int \frac{d^3q}{(2\pi)^3} \bar{v}(q)\nu_i(k-q)e^{i\omega_\eta}e^{i\omega_{k-q}\eta},
\]
and $\rho(\eta)$ is the energy density at conformal time $\eta$. This contribution can only be modeled numerically. Caprini et al [350, 352], noted that when modeling the contribution from turbulence, one needs to take into account that the turbulence continues long after after the phase transition is complete. If the source is long lasting need to take expansion into account. For example for $T_s = 100$ and $\beta/H_* = 100$ one finds the turbulence is not complete until $T \sim 120$ MeV. This causes some amplification. The effect is rather modest however, as the decay time of source (controlled by eddy turnover time) is much smaller than Hubble time. Indeed they found an amplification of a factor of about 2. Taking this into account, simulations show that one can achieve a reasonable fit with a power spectrum governed by our usual thermal parameters. [350, 353]
\[
h^2\Omega_{turb}(f) = 3.35 \times 10^{-4} \left( \frac{\beta}{H} \right)^{-1} \kappa_{turb}^{3/2} \left( \frac{\alpha}{1+\alpha} \right)^{3/2} \left( \frac{100}{g_*} \right)^{1/3} \nu_i S_{turb}(f).
\]

There is as yet no known method for directly calculating the efficiency parameter, however, this contribution is expected to be sub-dominant. The spectrum is [350, 353]
\[
S_{turb} = \frac{(f/f_{turb})^3}{[1 + (f/f_{turb})^{1/3}]} \left( 1 + 8\pi f/h_* \right),
\]
with
\[
h_* = 1.65 \times 10^{-7} \text{ Hz} \left( \frac{T_N}{\text{GeV}} \right) \left( \frac{g_*}{100} \right)^{1/6},
\]
\[
f_{turb} = 27 \times 10^{-7} \text{ Hz} \left( \frac{\nu_i}{H} \right) \left( \frac{T_N}{\text{GeV}} \right) \left( \frac{g_*}{100} \right)^{1/6}.
\]

6.1.4. Detection of gravitational waves from cosmic phase transitions. Of all the thermal parameters only the nucleation temperature depends strongly on the scale. $\beta/H \sim \log[T_s/M_{pl}]$ also depends weakly on the scale but is more strongly influenced by the ratio of scales $\nu/\Lambda$ as is $\alpha$ and the wall velocity. The transition temperature also controls the peak frequency. Therefore the scale of the new physics can be directly linked to the peak frequency. As each gravitational wave detector probes a different frequency, each probes a different scale
At the very low frequency one has pulsar timing arrays which probes phase transitions at the sub GeV scale. Lisa probes the electroweak phase transition and LIGO/VIRGO as well as KAGRA probes a scale of around $10^7$ GeV. KAGRA will soon be online and is expected to break a degeneracy in testing polarization. The precise scale of new physics that is probed depends on the thermal parameters and varies for the soundwave, turbulent and collision contributions to the total spectrum. In figure 16 we show the scales probed for $\alpha = 1$, $v_w = 1$ and $\beta/H^* = 1.3 \log[T_*/M_{Pl}]$ for all three contributions. A phase transition with a peak frequency visible by LIGO/VIRGO can be motivated by vacuum stability, split supersymmetry, a Pati–Salam transition or neutrino masses. Lisa probes both the electroweak phase transition and LIGO/VIRGO whereas pulsar timing arrays can probe supercooled electroweak phase transitions and the QCD phase transition. To probe the scale in between Lisa and LIGO/VIRGO, several other experiments have been proposed including Magis, BBO and Decigo.

More information about the underlying physics that produced a primordial gravitational wave signal can be gleaned from measuring primordial anisotropies that result from a strongly first order phase transition. Work by [401] analysed phase transitions occurring between 1–1000 TeV and demonstrated that we will obtain new anisotropies that can affect the CMB. One can then check to see if it is a dark sector or visible sector phase transition by checking correlations of $\delta \rho/\rho$ with the CMB. If $\delta \rho/\rho$ is uncorrelated with the CMB one knows that the Universe had a dark sector phase transition.

Figure 14. From left to right: depiction of deflagration, detonation and hybrid phase transitions. The green bands denote the sound shell whereas the solid black line denotes the scalar shell. Reproduced from [325]. © 2010 IOP Publishing Ltd and Sissa Medialab srl.
A triumph of modern cosmology is that two different measurements of the baryon to entropy ratio have concordance. The first is through BBN constraints where deuterium in particular is sensitive to the initial ratio of the baryon to entropy density,

\[ Y_B = \frac{n_b - \bar{n}_b}{s} = 7.3 \pm 2.5 \times 10^{-11}. \] (185)

Furthermore Planck measurements of oscillations in the CMB power spectrum give an overlapping estimate of the baryon yield

\[ Y_B = 8.59 \pm 0.11 \times 10^{-11}. \] (186)

This is unlikely to be an initial condition in any cosmology involving inflation. Although there exists, in the authors words, an ‘ugly and inelegant’ exception, inflation tends to wash out any initial baryon asymmetry. To produce a baryon asymmetry in a CPT conserving theory one needs to satisfy three conditions known as the Sakharov conditions:

1. C and CP violation (one or the other is insufficient)
2. Violation of baryon number conservation
3. a departure from thermal equilibrium.

Electroweak baryogenesis [13–17, 241, 411–414] generates this during the electroweak phase transition where topological processes known as sphalerons efficiently produce both baryons and anti-baryons in the symmetric phase. If the electroweak phase transition is strongly first order, CP violating interactions with bubbles of electroweak broken phase biases the sphalerons to produce more baryons than anti baryons. As the bubbles of broken phase expands, some of the net asymmetry is swept up into the low temperature phase and makes up the present asymmetry. If the electroweak phase transition is strongly first order, the initial baryon asymmetry produced during the transition will not be washed out by the very sphalerons which formed them.

The standard model fails to produce a sufficiently large baryon asymmetry. The standard model falls short on two Sakharov conditions, for a Higgs mass of 125 GeV the departure from equilibrium is too weak as the electroweak transition is actually a crossover transition. Furthermore, the CP violation in the CKM matrix is far too feeble to sufficiently bias the electroweak sphalerons. Therefore if electroweak baryogenesis is part of our cosmic history, one needs to extend the standard model to accommodate both Sakharov conditions. The required extensions to the standard model are in principle probable by experiment with both particle colliders and gravitational wave observatories probing the ingredients for a strongly first order electroweak phase transition while searches for permanent electric dipole moments probe sources of CP violation. The fact that electroweak baryogenesis is both testable and minimal makes it one of the most attractive paradigms.

Calculating the baryon asymmetry during a cosmic phase transition is a difficult problem. One usually calculates the

![Figure 15. Gravitational wave spectrum against Lisa one year sensitivity curves. Thermal parameters are $T_\ast = 100$, $\alpha = 1$ and $v_w = 1$. From left to right the top panels have $\beta/H = (1, 10)$ respectively and the bottom panels are $\beta/H = (10, 100)$, respectively. The black line is the total spectrum whereas the blue, green and red lines are the collision, sound wave and turbulence terms, respectively. Reproduced from [33]. © 2016 IOP Publishing Ltd and Sissa Medialab srl.](image)
overall left handed number density produced through CP violating interactions and then assume those processes are fast compared to weak sphaleron processes. In this case one can uncouple the dynamics of the baryon asymmetry production from the dynamics of the production of a chiral asymmetry. In this case the baryon asymmetry is given by

$$DQ \rho_B''(z) - v_w \rho_B''(z) - \Theta(z) R \rho_B = \Theta(-z) \frac{n_F}{2} \Gamma_{wL}(z).$$ (187)

Solving the above equation one finds that the baryon asymmetry is proportional to the sphaleron rate divided by the entropy which is the same order of magnitude of the observed baryon asymmetry. Therefore electroweak baryogenesis naturally produces the correct order of magnitude for the baryon asymmetry. Therefore it is customary to follow one of two approximate treatments: the first a semi-classical treatment using WKB methods [416, 417]. The second is known as the vev-insertion framework where one makes the assumption that the bulk of baryon production occurs immediately outside the bubble wall where vev is small so we can use the degrees of freedom and mass basis of the symmetric phase [415].

The vev insertion paradigm utilizes the closed time path formalism and captures resonance and memory effects which can substantially boost the overall asymmetry and has the advantage that it can be solved analytically [415, 418]. The vev insertion paradigm neglects flavour oscillation effects which can dampen the resonance [419, 420, 420]. Including gradient effects appears to recover some of the resonance [421].

When various approaches to calculating CP violating sources is valid remains an open problem in the field [422].

Since the standard model fails on two accounts to satisfy the Sakharov conditions, it is typical to extend the standard

---

**Figure 16.** Gravitational wave detectors probing the scale of a phase transition. Sound wave (top panel), collision (bottom left) and turbulence (bottom right) contributions to a gravitational wave source with $\alpha = v_w = 1$, $(\kappa_{\text{col}}, \kappa_{\text{turb}}) = (0.3, 0.02)$ and $\beta/H^* = 1.3 \log[T^*/M_{\text{pl}}]$ for phase transitions occurring at a scale $T_n = (10^{-5}, 10^{-2}, 10, 3 \times 10^5)$ GeV respectively against initial sensitivities of LIGO/VIRGO/Virgo [354, 355], LISA [356] and the European Pulsar Timing Array (EPTA) [357]. After integrating over frequency the sensitivity improves by several orders of magnitude [358]. Also see [333].
model by two sectors—one sector which catalyzes the electroweak phase transition, and another which is responsible for CP violating interactions with the bubble wall. If both sectors are heavy compared to the weak scale then one can in principle use an effective field theory approach [218, 423]. More common is to look at the case where the new physics sectors are weak scale themselves. For example, in the MSSM, if one had a stop lighter than the standard model top it could catalyze a strongly first order electroweak phase transition. The CP violation can then occur through stop-Higgs interactions or gaugino–Higgs interactions [415, 424–431]. The existence of colored scalars in the plasma also provide substantial drag on the bubble wall making the wall velocity naturally small which tends to make baryon production more efficient (though also makes the gravitational waves from the electroweak phase transition less visible). Unfortunately the light stop mechanism for catalyzing the electroweak phase transition is in serious conflict with collider constraints [277]. Indeed the EWBG within the MSSM was starting to look unviable even in the early LHC era [432]. Furthermore, EDM limits make both sources of CP violation severely constrained. Extending the MSSM by a gauge singlet (that is the NMSSM), one can catalyze a strongly first order electroweak phase transition with the additional scalar singlet [223, 373, 433–439] and the source of CP violation can be Singlino–Higgsino–Higgs interactions [230, 440]. Alternatively one can extend the MSSM by effective operators that catalyze the CP violation [441]. It is worth commenting that the minimal model of baryogenesis probably requires two additions to the standard Model to be viable—an addition that provides a source of CP violation and a source that catalyses a strong first order electroweak phase transition. Some examples of such minimal models include the standard model with an CPV effective operator and the addition of an effective operator [442] or an additional scalar [443] to catalyze the transition. Alternatively it has been shown that the addition of two additional fermions is sufficient [281].

Within the minimal supersymmetric standard model (MSSM) and 2HDM (Higgs doublet model) using the vev insertion frame work, one finds that the strength of CP violating sources for tree level interactions with the bubble wall are suppressed by a factor of $\Delta \beta \sim 10^{-2}$ where $\tan \beta(z)$ is the space time dependent ratio of the vevs $\langle v_u(z) \rangle / \langle v_d(z) \rangle$ and $\Delta \beta$ is its maximal variation. A study of the NMSSM showed that the addition of a gauge singlet can boost $\Delta \beta$, and therefore the baryon asymmetry, by an order of magnitude [223]. By contrast, if CP violation is a loop effect (for example the term $H_u \bar{f} f (\alpha + \frac{\Delta \beta}{\lambda} |H|^2)$ can contain a relative phase), one no longer has a $\Delta \beta$ suppression but instead supressed by a factor $v^2/\Lambda^2$. Therefore the scale of CPV physics can be reasonably high. Furthermore, tree level CP violating interactions result in a baryon asymmetry that is essentially independent of the bubble wall width in contrast to the case where the CP violation is loop induced where a strong dependence on the bubble wall width results. Finally we note that the electroweak phase transition need not be weak scale. Indeed if the phase transition proceeds through a multistep procedure either through an intermediate transition that breaks another symmetry [227], or through a two step electroweak phase transition [226], the scale of new physics required can be at the multi-TeV level and are best probed by gravitational wave observers and future colliders.

Outside of supersymmetry, Baryogenesis can also be linked with the production of dark matter [435, 436, 443–449] and has been explored in extended scalar sectors [443, 445, 446, 450, 451] and other low scale phase transitions [452]. It has

**Figure 17.** Thermal parameters from a dark Higgs with (right panel) and without (left panel) non-renormalizable operators for various models. In the above $N$ denotes the rank of the group and $N_0$ denotes the number of fermions coupled with unity Yukawa coupling. The plot points are coloured by their effective zero temperature mass. Note that $\xi$ in the above denotes the usual thermal parameter $\alpha$. The dashed contours in the plots correspond to the GW amplitude $\Omega_{gw}$, with $v_\alpha = 0.5$. The upper thicker contour corresponds to the LISA 1-year peak sensitivity [359]. The lower thicker dashed contour corresponds to the LISA for a power-law spectrum (integrated over frequency), taken from [358]. Reproduced from [364]. CC BY 3.0.
also been proposed that baryogenesis occurs spontaneously during the electroweak transition [453]. One can also use CP violation in the lepton sector to produce enough baryon asymmetry [454]. We end this section by noting that even if the baryon asymmetry is produced through leptogenesis, it may still involve a phase transition [455].

Acknowledgments

We would like to thank Mark Hindmarsh, Djuna Croon, Pratika Dayal, Ruth Durrer, Hiren Patel, Michael Ramsey Musolf, Jim Cline and Jonathan Kozaczuk. A M’s research is financially supported by Netherlands Organisation for Scientific Research (NWO) Grant No. 680-91-119. Triumf receives federal funding via a contribution agreement with the National Research Council of Canada. This work was performed in part at Aspen Center for Physics, which is supported by National Science Foundation Grant PHY-1607611.

ORCID iDs

Anupam Mazumdar https://orcid.org/0000-0002-0967-8964

References

[1] Linde A D 1979 Rep. Prog. Phys. 42 389

Figure 18. Left panels: the spectrum due to two simultaneous PTs, given by the red line, leads to a different spectrum than consecutive transitions. This can be seen by comparison with the gray line, which is predicted by the same thermal parameters. Right panels: the spectrum from simultaneous PTs can not be fitted to the spectrum than from a single PT. Reproduced from [231]. CC BY 3.0.
Anupam Mazumdar is a theoretical physicist specializing in cosmology, phase transition, particle physics, and quantum gravity. He has initiated the ideas on multi-field theory of inflation, such as assisted inflation, visible sector inflation such as supersymmetric standard model inflation. He has worked on the ghost-free and singularity-free construction of infinite derivative theories of gravity, which can potentially resolve cosmological and black hole singularities, yielding a non-singular compact object without event horizon in the latter case.

Graham White is a particle cosmologist from Australia. He has worked probes of the early Universe - from the impact of collider constraints to big bang nucleosynthesis and gravitational waves. In particular, he is interested in the question of why there is more matter than anti-matter in the Universe and has authored a book on the subject. To this end he has worked on new testable paradigms as well as ways to constrain leptogenesis.