Contents

Chapter 1

Background to the problem

Chapter 2

The Effective Potential

The effective potential is a tool which was found to be useful for the study of spontaneous symmetry breaking and restoration in the early 1970's. In the present context, it attempts to include radiative corrections to the scalar tree-level potential in the standard model (or extended standard model) by including the loop graphs. We are also able to introduce the effects of a hot vacuum by considering the potential in a heat bath which represents the background of pair production in a hot universe. In this way we can define an effective potential to describe the behaviour of the scalar sector at or near the electroweak phase transition.

As described below, the effective potential can be written in an integral form but the integral is impossible to calculate in general terms. Hence it is common to resort to expansions at high or low temperatures. Obviously it is then important to investigate the validity of these expansions.

It is also possible to improve the basic loop expansion by making corrections to the temperature dependent masses. In fact, these corrections, known as "mass insertions", are essential to the correct understanding of the nature of the phase transition because the resulting diagrams (commonly known as the "daisy diagrams", see fig. XXX) are comparable in size to the diagrams in the leading terms in the high temperature expansion.

We will describe below the definition of the effective potential (V_{eff}), the inclusion of the temperature dependence in V_{eff}, the high temperature expansion of V_{eff} and its validity, and the derivation of the effective potential in the two Higgs-doublet case under study.

Note that since the detailed calculations of the effective potential and its high temperature expansion are well known from work such as the classic paper by Dolan and Jackiw [XX], the details are not reiterated here. However, the important points of the calculations are included for completeness.

2.1 Definition of effective potential

2.2 Temperature-dependence in V_{eff}

2.2.1 High-temperature expansion for V_{eff}

2.2.2 Validity of the high-temperature expansion

2.3 V_{eff} in the two Higgs-doublet standard model

Chapter 3

The Model

3.1 Basic approach to the model

We had several fundamental points to address in our construction of a computational model and our approach to using this model. Firstly, let us consider how we intend to use the model for a search of the scalar sector parameter space and then look at what we will require of the model itself.

3.1.1 How to search the parameter space?

Our basic aim is to investigate the parameter space available to the extended scalar sector in a search for a region, or regions, which can support baryogenesis (by bubble nucleation) at the electroweak phase transition and, in particular, avoid any subsequent washout of baryon number by sphaleron transitions. We wish to consider the most general scalar potential possible, and this contains 9 parameters, as we have already seen. Only one of these parameters is (effectively) known — ie., we know experimentally that the VEV of the scalar fields is

$$v_0 = 246GeV \tag{3.1.1}$$

So, we are left with 8 free parameters, giving us an 8–dimensional parameter space to work in. It will come as no surprise that preliminary investigations of the problem suggested that it was all but impossible to make an analytical search of such a complicated parameter space — ie., we cannot generally derive analytical boundaries in the space which contain the region supporting electroweak baryogenesis. Hence, we determined to develop a random sampling approach to the investigation of the parameter space.

This was initially envisaged in the following way —

i. Choose a point entirely at random in the parameter space

ii. Model the potential at this point.

iii. Does this point support baryogenesis and avoid the washout? If not, then discard the point and continue the search. If it does support baryogenesis, then retain the details of the point and its potential for analysis, and continue the search.

Now, this is clearly a very simplistic view of what we are likely to encounter in such a process. We must take into account a great deal more than is immediately apparent here. For example, what do we mean by "choose at random"? Should we do the obvious thing and take a flat distribution over all of our free parameters? Could the distributions introduce bias into our results and conclusions?

Well, the first point to note is that we must impose the experimental condition given above in eqn(3.1.1). In the "typical" notation of the scalar potential ($-\mu_i^2$, λ_i, μ_R^2 and μ_I^2), v_0 does not appear as a parameter and so we must begin by determining relationships between v_0 and the parameters in the "typical" notation (which we will refer to as the typical parameters or typical parameter set). If we were to simply choose the 9 typical parameters at random then we would never strike on exactly those points for which eqn(3.1.1) is satisfied, so we see the importance of this point.

Next, we note that *a priori* we have no feel for the likely ranges of the parameters. Yes, we know that we must keep the λ's small, but how small? And more problematically, what about the ranges for the dimensionful couplings — there are no smallness arguments available here. We realise that, since we are uncertain of what we might call reasonable ranges, we are also uncertain even of which parts of the parameter space can be considered physically acceptable, with or without supporting a first order phase transition. To combat these problems we were able to develop a "useful" parameter set, obviously equivalent but not identical to the typical parameter set. These useful parameters allowed us to make random choices in ranges which were well understood and hence to *derive* the dimensionful couplings μ_1^2 and μ_2^2 from the chosen parameters.

We also realise that there *are* conditions which we can (indeed must) impose to ensure that the potential is physically acceptable, but that these are not explicit enough to allow us to make appropriate choices of the parameters in every case. So, we must develop a system to check whether or not each chosen point in the parameter space complies with these physical constraints. The conditions themselves have been well known for several years in the standard model scalar sector but must be re-derived, *for the full general form of the potential*, in the two—Higgs—doublet extended scalar sector.

All of the points raised above will be discussed in detail in the remainder of this chapter.

3.1.2 Requirements on the model

In light of how we intend to use the model, it must be able to accept any given parameter set, convert that set into the equivalent typical parameter set, and then provide us with quantitative results on the important properties of the effective potential as they pertain to the phase transition. Such properties will obviously include the critical temperature, the value of $v(T_c)/T_c$, perhaps the depth of the minimum at the critical temperature, and certainly the physical mass spectrum of the scalar sector.

Having already described the definition of the effective potential and the expansions used in evaluating the potential, it is apparent that providing such a model is non-trivial but it is possible. We will provide a brief overview of the working of the model and spend a good deal more time on the specifically numerical techniques used, why they were required, how they were implemented and checked, and some of the problems encountered which were specifically related to the properties of the effective potential.

3.2 Searching the parameter space

3.2.1 Notation

We have opted for the typical notation in the tree level scalar potential, viz.

$$\begin{aligned} V_0 &= -\mu_1^2\Phi_1^\dagger\Phi_1 - \mu_2^2\Phi_2^\dagger\Phi_2 + \lambda_1\left(\Phi_1^\dagger\Phi_1\right)^2 + \lambda_2\left(\Phi_2^\dagger\Phi_2\right)^2 \\ &\quad +\lambda_3\left(\Phi_1^\dagger\Phi_1\right)\left(\Phi_2^\dagger\Phi_2\right) + \lambda_4\left|\Phi_1^\dagger\Phi_2\right| + \frac{\lambda_5}{2}\left\{\left(\Phi_1^\dagger\Phi_2\right)^2 + \left(\Phi_2^\dagger\Phi_1\right)^2\right\} \\ &\quad -\mu_3^2\Phi_1^\dagger\Phi_2 - \mu_3^{2*}\Phi_2^\dagger\Phi_1 \end{aligned} \tag{3.2.2}$$

However, this does not allow us to directly impose the condition in eqn(3.1.1). So, we note that the potential effectively reduces to a function of only 3 variables — those scalar fields which develop non-zero VEV's, as shown below:

$$\langle\Phi_1\rangle = \begin{pmatrix} 0 \\ v_3 \end{pmatrix} \qquad \langle\Phi_2\rangle = \begin{pmatrix} 0 \\ v_7 + iv_8 \end{pmatrix} \tag{3.2.3}$$

Note that the variables v_i have been labelled according to their corresponding scalar fields and will be used to refer specifically to the VEV's but that we will write a simplified version of the potential using the 3 variables σ_i, defined below

$$\Phi_1 = \begin{pmatrix} 0 \\ \sigma_3 \end{pmatrix} \qquad \Phi_2 = \begin{pmatrix} 0 \\ \sigma_7 + i\sigma_8 \end{pmatrix} \tag{3.2.4}$$

We will also often use the notation $(\sigma_1, \sigma_2, \theta)$, where

$$\Phi_1 = \begin{pmatrix} 0 \\ \sigma_1 \end{pmatrix} \qquad \Phi_2 = \begin{pmatrix} 0 \\ \sigma_2 e^{i\theta} \end{pmatrix} \tag{3.2.5}$$

and this obviously has a corresponding notation for the VEV's, (v_1, v_2, θ_0). It should be clear by the context which notation we are using and we will not normally explain which is in use.

3.2.2 Defining the "useful" parameter set

Firstly, consider how we will impose the condition of eqn(3.1.1). We use the fact that $v_1^2 + v_2^2 = v_0^2$ to parameterise the VEV's in the following way:

$$v_1 = v_0 \cos(\beta_0) \qquad v_2 = v_0 \sin(\beta_0) \tag{3.2.6}$$

and

$$v_7 = v_2 \sin(\theta_0) \qquad v_8 = v_2 \cos(\theta_0) \tag{3.2.7}$$

Now, consider the following conditions, derived by setting the derivatives of the potential equal to zero at the minimum and using the parameterisation above.

$$-\mu_1^2 + \lambda_1 v_1^2 + \frac{\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)}{2} v_2^2 - \gamma_1 \frac{v_2}{v_1} = 0 \tag{3.2.8}$$

$$-\mu_2^2 + \lambda_2 v_2^2 + \frac{\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)}{2} v_1^2 - \gamma_1 \frac{v_1}{v_2} = 0 \tag{3.2.9}$$

$$\mu_R^2 \sin(\theta_0) + \mu_I^2 \cos(\theta_0) - \lambda_5 \sin(\theta_0)\cos(\theta_0) v_1 v_2 = 0 \tag{3.2.10}$$

In these conditions we can consider the VEV's as free parameters and hence derive expressions for $\mu_{1,2}^2$. As mentioned in section 3.1.1, this would free us from the problem of determining reasonable ranges for $\mu_{1,2}^2$ and replace this problem with that of choosing an appropriate position for the minimum of the potential. There are 3 parameters in a choice of the minimum — v_0, which must be $246GeV$; β_0, which we can constrain as we shall see shortly; and θ_0, which is more difficult to constrain. However, if we are to solve the conditions above, note that they contain 12 different parameters between them, including v_0, β_0 and θ_0. Since v_0 is fixed, I may choose any 8 other parameters at random and then solve for the remaining 3. There are many possibilities here but we have already stated that we think it might be useful to solve for $\mu_{1,2}^2$ and so we must decide which one other parameter we would prefer to derive rather than choose at random. It seems somewhat inconsistent to derive one of the λ's, say, and choose the 4 others. We also already know that we can use small coupling arguments to place constraints on the λ's, so we are left to consider μ_R^2, μ_I^2 and the angles β_0 and θ_0. We also know that we want to choose the value of β_0 because we can put constraints on it.

And again, to choose the value of one of μ_R^2 and μ_I^2, and derive the other, seems inconsistent: with such a system, could we be sure of a full investigation of the effects of the μ_3^2 term, which surely must be of interest in our general case?

This leaves us with the option of deriving the angle θ_0. As already stated, it is not immediately apparent how we might constrain θ_0 and so it seems reasonable to make this choice. (We will see later, in section 3.2.4 that it is possible to determine a well-defined range which must contain θ_0. This should not be viewed, however, as constraining θ_0 — it is merely constraining the breadth of our search when we attempt to solve for θ_0.) So, our "useful" parameter set consists of

$$\left(\beta_0, \lambda_i, \mu_R^2, \mu_I^2\right) \tag{3.2.11}$$

and, of course, $v_0 = 246GeV$.

3.2.3 Symmetries in the effective potential

There are several symmetries present in our general form of the scalar potential. These symmetries will allow us to restrict the range of our search because it becomes clear that the signs of several parameters are not measurable. So, let us begin with the tree level potential in the following form —

$$\begin{aligned} V_0 \quad = \quad & -\frac{\mu_1^2}{2}\sigma_1^2 - \frac{\mu_2^2}{2}\sigma_2^2 + \frac{\lambda_1}{4}\sigma_1^4 + \frac{\lambda_2}{4}\sigma_2^4 + \frac{1}{4}\left[\lambda_3 + \lambda_4 + \lambda_5\cos(2\theta)\right]\sigma_1^2\sigma_2^2 \\ & -\frac{1}{2}\left[\mu_R^2\cos(\theta) - \mu_I^2\sin(\theta)\right]\sigma_1\sigma_2 \end{aligned} \tag{3.2.12}$$

We will state the symmetries at tree level without proof (since they are fairly clear and simple to check) then discuss how the symmetries extend to higher order terms in the effective potential.

$$\begin{aligned} \sigma_1 \to -\sigma_1, \quad & \sigma_2 \to -\sigma_2 \\ \sigma_i \to -\sigma_i, \quad & \mu_R^2 \to -\mu_R^2, \quad \theta \to -\theta \\ \sigma_i \to -\sigma_i, \quad & \mu_I^2 \to -\mu_I^2, \quad \theta \to \pi - \theta \\ & \mu_R^2 \to -\mu_R^2, \quad \theta \to \pi - \theta \\ & \mu_I^2 \to -\mu_I^2, \quad \theta \to -\theta \end{aligned} \tag{3.2.13}$$

It is not obvious, *a priori*, that these symmetries will also hold in the higher order terms since each of these terms arises from derivatives of all of the lower order terms in the effective potential and taking derivatives will change the order of σ_i^n terms, say. However, since the symmetries hold in the tree level potential then they must also hold in each term of the second derivative matrix (because the shape of the potential must be the same in any two cases related by any of the symmetries above.) Since the symmetries will hold term by term in the curvature matrix, they must also hold in the trace of the matrix and hence we see that the symmetries also hold at the $V_0 + T^2/24\mathrm{Tr}\,[M^2]$ level. We can then follow a similar line of argument to show that these symmetries will hold at the $V_0 + T^2/24\mathrm{Tr}\,[M^2] - T/(12\pi)\mathrm{Tr}\,[M^3]$ level, and so on.

The consequences of these symmetries for the parameters of the model are now clear — we are able to choose the signs of v_1, v_2, μ_R^2 and μ_I^2 without loss of generality. The obvious line to take here is to simply insist that

$$v_1, v_2, \mu_R^2, \mu_I^2 \geq 0 \tag{3.2.14}$$

by definition. Note that this implies we must have β_0 (the phase angle between the VEV's of the first and second doublets) in the first quadrant.

What is not so clear is the resulting effect of the convention in eqn.(3.2.14) on θ_0. We will now explain that, *under this convention*, we must find θ_0 in the fourth quadrant.

3.2.4 The location of θ_0

When we consider eqn.(3.2.12) above, we realise that we can write it as two parts, one of which is dependent on θ. Let us define the θ–dependent part to be $V_0{}'(\theta)$. Then we have

$$V_0{}'(\theta) = \left\{\lambda_5\sigma_1\sigma_2\cos(2\theta) - \mu_R^2\cos(\theta) + \mu_I^2\sin(\theta)\right\}\sigma_1\sigma_2 \tag{3.2.15}$$

Naively, one might suggest that we can minimise the potential by simply minimising both the θ–dependent and θ–independent parts of the potential individually.

However, $V_0{}'(\theta)$ and the θ–independent part of V_0 are not independent of one another, they are linked by (σ_1, σ_2), and hence this naive approach will generally fail. On the other hand, for any given values of (σ_1, σ_2) we can minimise the contribution of $V_0{}'(\theta)$ to the potential by varying θ. Hence it is certainly true that θ_0 is just the value of θ which minimises $V_0{}'(\theta)$ with $(\sigma_1, \sigma_2) = (v_1, v_2)$.

Now, our definition of $V_0{}'(\theta)$ clearly shows that minimisation with respect to θ is exactly equivalent to solving eqn.(3.2.10) above, which is a statement of the condition that the derivative of the tree level potential with respect to θ should be zero at the minimum. This is hardly surprising but is important to our argument here because, assuming that eqn.(3.2.10) has valid solutions [1] then θ_0 always lies at the global minimum of $V_0{}'(\theta)$, with v_1 and v_2 held constant. So, if we can find a range in which $V_0{}'(\theta)$ must achieve its minimum then we have found a range which definitely contains θ_0.

Following this line of argument, it is easily shown that

$$\theta_0 \in \left[-\frac{\pi}{2}, 0\right] \tag{3.2.16}$$

by considering the signs of the terms in eqn.(3.2.15) *with the convention developed in eqn.(3.2.14) of the last section.*

So, we can now define

$$\begin{array}{c} \beta_0 \in \left[0, \dfrac{\pi}{2}\right], \qquad \mu_R^2 \geq 0, \qquad \mu_I^2 \geq 0 \\ \theta_0 \in \left[-\frac{\pi}{2}, 0\right] \end{array} \tag{3.2.17}$$

3.2.5 Determination of θ_0, μ_1^2 and μ_2^2

Clearly, all derivatives of the potential with respect to the scalar fields must be zero at the global minimum. These conditions lead to only 3 independent expressions, eqns.(3.2.8 — 3.2.10), which are repeated here for convenience.

$$-\mu_1^2 + \lambda_1 v_1^2 + \frac{\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)}{2} v_2^2 - \gamma_1 \frac{v_2}{v_1} = 0 \tag{3.2.18}$$

[1] We shall see in the next section that there are *always* valid solutions.

$$-\mu_2^2 + \lambda_2 v_2^2 + \frac{\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)}{2} v_1^2 - \gamma_1 \frac{v_1}{v_2} = 0 \qquad (3.2.19)$$

$$\mu_R^2 \sin(\theta_0) + \mu_I^2 \cos(\theta_0) - \frac{\lambda_5}{2} v_1 v_2 \sin(2\theta_0) = 0 \qquad (3.2.20)$$

where $\gamma_1 = \mu_R^2 \cos(\theta_0) - \mu_I^2 \sin(\theta_0)$.

The equations $3.2.18, 3.2.19$ above both contain two unknowns and so we are forced to solve eqn.(3.2.20) first. Unfortunately, this cannot be done analytically [2] but it is instructive to consider the form of eqn.(3.2.20). So, let us first consider whether or not there exist legitimate solutions to eqn.(3.2.20), bearing in mind that we wish to find solutions in the fourth quadrant, as explained in the last section. We re-write in the following form

$$\lambda_5 v_1 v_2 \sin(\theta_0) - \mu_I^2 = \mu_R^2 \tan(\theta_0) \qquad (3.2.21)$$

Note μ_R^2 is positive, and $\tan(\theta_0)$ varies continuously from $-\infty$ (at $\theta_0 = -\pi/2$) to 0 (at $\theta_0 = 0$) in the fourth quadrant. Defining the LHS of eqn.(3.2.21) to be a function of θ_0, $f(\theta_0)$, then $f(\theta_0)$ will vary continuously, in the fourth quadrant, over a *finite* range $f(-\pi/2)$ to $f(0)$ [3] where $f(0) \leq 0$. Hence, by continuity arguments, there does exist at least one solution to eqn.(3.2.21). If we consider the graph below we see that there are, in fact, always at least two distinct solutions, and at most four, but, more importantly, there is always one *and only one* solution in the fourth quadrant. So, although we must solve eqn.(3.2.20) numerically, we can restrict our search to the fourth quadrant and be confident that the resulting θ_0 is the required value.

Having evaluated θ_0, calculation of μ_1^2 and μ_2^2 is straightforward.

---

[2] Actually, we can solve analytically, by rearranging the expression as a quartic equation in $\cos(\theta)$ and solving. However, the resulting expressions are so complex and long-winded as to be unusable.

[3] Note that $f(-\pi/2)$ might be greater *or* less than $f(0)$, depending on the sign of λ_5.

3.3 Conditions for a physically acceptable potential

It has already been noted in this chapter that we cannot analytically determine the boundaries of those regions of the parameter space which are able to support electroweak baryogenesis and avoid the possibility of subsequent baryon number washout by sphaleron transitions shortly after the electroweak phase transition. (Incidentally, we shall herewith refer to such regions as "regions which support baryogenesis" or "baryogenesis regions.") Hence the need for a numerical search of the parameter space.

However, it must also be noted that there are regions of the parameter space, lying outwith the baryogenesis regions, where the potential becomes completely unphysical — eg., where the potential becomes unbounded below and thereby has no global minimum or stable vacuum. It is obvious that the baryogenesis regions must be wholly contained within regions where the potential *is* physically acceptable (we will refer to these regions as "physically acceptable regions" or "physical regions.") These physical regions will form our basis parameter space (since it is not physically sensible to consider the unphysical regions) and we will be specifically interested in what proportion of the physical region is contained in the baryogenesis region. Then, ideally we would like to perform our parameter search over these physically acceptable regions. Hence we must decide what constitutes a physically acceptable potential.

The simplest approach to deciding what constitutes a physically acceptable potential is to decide when a potential is unphysical. The next few sections will explain the various ways in which the potential can become unphysical and derive the conditions which form the boundaries between the physical and unphysical regions. Of course, such conditions have been considered and derived before in many references. However, they have never been derived in the most general two-Higgs-doublet model until now. [4]

[4] The general conditions have also been derived by Kastening in REF. and we were able to

3.3.1 Stability of the vacuum

When the potential is unbounded below then there is no stable minimum and, hence, no stable vacuum. The fields would tend to roll ever further from the origin and take on infinitely large VEV's, resulting in an infinite energy density in the vacuum. This is, of course, entirely unacceptable. So, under what conditions does the potential take on these unfortunate characteristics?

It is useful to prove the following result at this point —

Let

$$\Phi_1 = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \Phi_2 = \begin{pmatrix} v_3 \\ v_4 \end{pmatrix} \tag{3.3.22}$$

$$\text{So,}\Phi_1^\dagger\Phi_1 = |v_1|^2 + |v_2|^2, \quad \Phi_2^\dagger\Phi_2 = |v_3|^2 + |v_4|^2, \quad \Phi_1^\dagger\Phi_2 = \overline{v_1}v_3 + \overline{v_2}v_4 \tag{3.3.23}$$

Then we have

$$\begin{aligned} & \left(\Phi_1^\dagger\Phi_1\right)\left(\Phi_2^\dagger\Phi_2\right) - \left(\Phi_1^\dagger\Phi_2\right)\left(\Phi_2^\dagger\Phi_1\right) \\ = \; & |v_2|^2|v_3|^2 + |v_1|^2|v_4|^2 - v_1v_4\overline{v_2v_3} - \overline{v_1v_4}v_2v_3 \\ = \; & (v_1v_4 - v_2v_3)\overline{(v_1v_4 - v_2v_3)} \\ = \; & |v_1v_4 - v_2v_3|^2 \qquad \geq 0 \end{aligned} \tag{3.3.24}$$

So, the result is $(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) \geq (\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1)$. This will be put to use in the following series of derivations.

Let us now introduce a notation to be used in deriving the necessary conditions on the parameters of the potential. The notation will be

$$\Phi_1^\dagger\Phi_1 = A, \quad \Phi_2^\dagger\Phi_2 = B, \quad \Phi_1^\dagger\Phi_2 = X + iY \tag{3.3.25}$$

where A, B, X, Y are all real, and the result above leads to the condition $AB \geq (X^2 + Y^2)$.

Well, the behaviour of the potential at large fields depends only on the quartic terms. For the potential to be bounded below requires that the quartic terms are positive at large fields or, equivalently, that the gradient of the quartic terms

use REF. as a check on our own work.

should be positive (negative) at large positive (negative) fields. So, it is necessary that

$$\lambda_1 A^2 + \lambda_2 B^2 + \lambda_3 AB + (\lambda_4 + \lambda_5)X^2 + (\lambda_4 - \lambda_5)Y^2 \quad > \quad 0 \qquad (3.3.26)$$

Firstly, consider the two directions $A = 0$ and $B = 0$. Then we quickly realise the necessary conditions

$$\begin{aligned} \lambda_1 &> 0 \\ \lambda_2 &> 0 \end{aligned} \qquad (3.3.27)$$

These conditions are already familiar from the special cases of two-Higgs-doublet models which have been studied regularly in the literature to date [REFS].

Now, using the result proved above, the minimum value of $\{(\lambda_4 + \lambda_5)X^2 + (\lambda_4 - \lambda_5)Y^2\}$ for any given AB is $\lambda_{min}ab$, where λ_{min} is defined by

$$\lambda_{min} = \min(0, \lambda_4 + \lambda_5, \lambda_4 - \lambda_5) \qquad (3.3.28)$$

So, in order to check that the quartic terms are positive at large fields, it is necessary to minimise the expression

$$\left(\sqrt{\lambda_1}A - \sqrt{\lambda_2}B\right)^2 + \left(\lambda_3 + \lambda_{min} + 2\sqrt{\lambda_1\lambda_2}\right)AB \qquad (3.3.29)$$

However, when $A/B = \sqrt{\lambda_2/\lambda_1}$, then it is necessary that $\lambda_3 + \lambda_{min} + 2\sqrt{\lambda_1\lambda_2} > 0$.

This can be re–written to provide the necessary condition

$$\min(\lambda_3, \lambda_3 + \lambda_4 + \lambda_5, \lambda_3 + \lambda_4 - \lambda_5) + 2\sqrt{\lambda_1\lambda_2} > 0 \qquad (3.3.30)$$

3.3.2 Conservation of E.M.

The true vacuum is the state which provides the spontaneous breaking of the $SU(2)_L \times U(1)_Y$ symmetry which is required in the Higgs' mechanism, but must not break the $SU(1)_em$ symmetry. It is well known [REFS] that this is equivalent to the condition that the VEV's of the two doublets must be parallel, since otherwise the Lagrangian is not invariant under the combination of generators

$Q = T_3 + Y/2$ — ie. we develop a charged VEV. (IS THAT COMMENT TRUE?) So, we wish to derive the conditions on the parameters of the potential such that the global minimum must lie in some direction with the two doublets parallel.

In order to simplify the derivation of the necessary conditions, it is useful to rewrite the tree level potential in the following form —

$$\begin{aligned} V' \;=\; & -\mu_1^2\left\{A - \frac{v_1^2}{2}\right\} - \mu_2^2\left\{B - \frac{v_2^2}{2}\right\} \\ & +\lambda_1\left\{A^2 - \frac{v_1^4}{4}\right\} + \lambda_2\left\{B^2 - \frac{v_2^4}{4}\right\} + \lambda_3\left\{AB - \frac{v_1^2 v_2^2}{4}\right\} \\ & +\lambda_4\left\{X^2 + Y^2 - \frac{v_1^2 v_2^2}{4}\right\} + \frac{\lambda_5}{2}\left\{(X+iY)^2 + (X-iY)^2 - \cos(2\theta_0)\frac{v_1^2 v_2^2}{2}\right\} \\ & -\mu_3^2(X+iY) - \mu_3^{2\star}(X-iY) + \left(\mu_R^2\cos(\theta_0) - \mu_I^2\sin(\theta_0)\right) v_1 v_2 \end{aligned} \tag{3.3.31}$$

where the global minimum is *defined* to lie at the point

$$\Phi_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v_1 \end{pmatrix} \quad \Phi_2 = \frac{\exp(i\theta_0)}{\sqrt{2}}\begin{pmatrix} 0 \\ v_2 \end{pmatrix} \tag{3.3.32}$$

Then clearly $V' \geq 0$ for all fields Φ_1 and Φ_2 and it is possible to derive necessary conditions on the parameters of the potential by considering the sign of the tree level potential at various points away from the global minimum.

Consider the point

$$\Phi_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v_1 \end{pmatrix} \quad \Phi_2 = \frac{\exp(-i\theta_0)}{\sqrt{2}}\begin{pmatrix} 0 \\ -v_2 \end{pmatrix} \tag{3.3.33}$$

All terms in V', except the soft symmetry–breaking term, are zero at this point. The soft breaking term becomes

$$\begin{aligned} & \left\{\mu_R^2\cos(\theta_0) + \mu_I^2\sin(\theta_0)\right\} v_1 v_2 + \left\{\mu_R^2\cos(\theta_0) - \mu_I^2\sin(\theta_0)\right\} v_1 v_2 \\ & \qquad = 2\mu_R^2\cos(\theta_0) v_1 v_2 \end{aligned} \tag{3.3.34}$$

and hence, in our sign convention, it is necessary that

$$cos(\theta_0) \geq 0 \tag{3.3.35}$$

Similarly, by considering the point where $\theta_0 \to -\theta_0$ and the v_i remain unchanged, the soft symmetry-breaking term is the only changed term again. In our sign convention, this leads to the result

$$\begin{array}{rcl} & \{-2\mu_I^2 \sin(\theta_0)\}\, v_1 v_2 & \geq 0 \\ \to & \sin(\theta_0) & \leq 0 \end{array} \qquad (3.3.36)$$

These two results combine to constrain θ_0 to the fourth quadrant, as has already been demonstrated. This proof of these constraints is perhaps more elegant than the previous proof but the first version was a more intuitive approach and I feel that it has been enlightening to see both approaches.

Consider now the case

$$\Phi_1 = \frac{1}{\sqrt{2}}\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{4}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \quad \Phi_2 = \tfrac{\exp(i\theta_0)}{\sqrt{2}} \left(\tfrac{\lambda_1}{\lambda_2}\right)^{\frac{1}{4}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix} \qquad (3.3.37)$$

Then all terms in V' involving $\Phi_i^\dagger \Phi_j$ go to zero and the resulting potential has the form

$$\begin{aligned} V' &= -\frac{\mu_1^2}{2}\left\{\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{2}} v_2^2 - v_1^2\right\} - \frac{\mu_2^2}{2}\left\{\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{2}} v_1^2 - v_2^2\right\} \\ &\quad + \frac{\lambda_1}{4}\left\{\left(\frac{\lambda_2}{\lambda_1}\right) v_2^4 - v_1^4\right\} + \frac{\lambda_2}{4}\left\{\left(\frac{\lambda_1}{\lambda_2}\right) v_1^4 - v_2^4\right\} \\ &= -\frac{\mu_1^2}{2}\left\{\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{2}} v_2^2 - v_1^2\right\} - \frac{\mu_2^2}{2}\left\{\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{2}} v_1^2 - v_2^2\right\} \\ &= -\frac{1}{2}\left\{\lambda_1 v_1^2 + \frac{\Sigma_+}{2} v_2^2 - \gamma_1 \frac{v_2}{v_1}\right\}\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{4}}\left\{\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{4}} v_2^2 - \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{4}} v_1^2\right\} \\ &\quad -\frac{1}{2}\left\{\lambda_2 v_2^2 + \frac{\Sigma_+}{2} v_1^2 - \gamma_1 \frac{v_1}{v_2}\right\}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{4}}\left\{\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{4}} v_1^2 - \left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{4}} v_2^2\right\} \\ &= \frac{1}{2}\left\{\sqrt{\lambda_1 \lambda_2} - \frac{1}{2}\left(\Sigma_+ - \frac{2\gamma_1}{v_1 v_2}\right)\right\}\left\{\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{4}} v_1^2 - \left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{4}} v_2^2\right\}^2 \\ &\geq 0 \end{aligned} \qquad (3.3.38)$$

Hence, the necessary condition is

$$\left(\Sigma_+ - \frac{2\gamma_1}{v_1 v_2}\right) \leq 2\sqrt{\lambda_1 \lambda_2} \qquad (3.3.39)$$

The last necessary condition required is most simply derived by considering the physical mass spectrum.

3.3.3 Reality of the physical masses

Since we have defined that we are working at the global minimum then it must be true that the curvature matrix, evaluated at the minimum, is semi-positive definite — ie. the mass spectrum has only non-negative masses. We have 8 scalar fields and hence the mass (ie. curvature) matrix is an 8×8 matrix but it can be arranged in a block diagonal form as shown below —

$$M^2 = \begin{pmatrix} M^2_{1256} & O_4 \\ O_4 & M^2_{3478} \end{pmatrix} \tag{3.3.40}$$

where the submatrices M^2_{1256} and M^2_{3478} are formed from the bases $\{\phi_1, \phi_2, \phi_5, \phi_6\}$ and $\{\phi_3, \phi_4, \phi_7, \phi_8\}$ respectively, and have the forms

$$M^2_{1256} = \begin{pmatrix} a_{1256} & 0 & c_{1256} & d_{1256} \\ 0 & a_{1256} & -d_{1256} & c_{1256} \\ c_{1256} & -d_{1256} & b_{1256} & 0 \\ d_{1256} & c_{1256} & 0 & b_{1256} \end{pmatrix} \tag{3.3.41}$$

$$M^2_{3478} = \begin{pmatrix} a_{3478} & e_{3478} & f_{3478} & g_{3478} \\ e_{3478} & b_{3478} & -g_{3478} & h_{3478} \\ f_{3478} & -g_{3478} & c_{3478} & j_{3478} \\ g_{3478} & h_{3478} & j_{3478} & d_{3478} \end{pmatrix} \tag{3.3.42}$$

For completeness, the full form of these matrices can be found in Appendix 7.

We can easily diagonalise the matrix M^2_{1256}, as seen in chapter XX, which leads to the eigenvalues

$$\begin{aligned} M^2_{\chi^\pm} &= 0 \\ M^2_{+} &= -\frac{v_0^2}{2}\left[\lambda_4 + \lambda_5\cos(2\theta_0) - \frac{\mu_R^2\cos(\theta_0) - \mu_I^2\sin(\theta_0)}{v_1 v_2}\right] \end{aligned} \tag{3.3.43}$$

where $M^2_{\chi^\pm}$ correspond to the would-be Goldstone bosons $\chi^\pm$ and M^2_+ corresponds to the charged scalars $H^\pm$.

Since M^2_+ must be positive, then we have the condition

$$\lambda_4 + \lambda_5\cos(2\theta_0) \leq \frac{2\mu_R^2\cos(\theta_0) - \mu_I^2\sin(\theta_0)}{v_1 v_2}$$

$$\rightarrow \qquad (\lambda_4 + \lambda_5 \cos(2\theta_0)) \leq \tfrac{2\gamma_1}{v_1 v_2} \tag{3.3.44}$$

Since μ_R^2, μ_I^2, v_1 and v_2 are all positive, and θ_0 lies in the fourth quadrant, then the RHS of this inequality is certainly positive. Mathematically, the value of γ_1 can be placed in the range

$$\min\left(\mu_R^2, \mu_I^2\right) \leq \gamma_1 \leq 1.41\max\left(\mu_R^2, \mu_I^2\right) \tag{3.3.45}$$

but it should be remembered that the value of θ_0 is determined by the physical requirements of the model and not simple mathematical possibilities — eg., if $\mu_I^2 = 0$, then $\theta_0 = 2\pi$ and the lower limit given above can then appear very loose. Nevertheless, these limits are useful in understanding the effect of the constraint in eqn.(3.3.44). Since μ_R^2 and μ_I^2 are restricted to be less than $22.5 \times 10^3 GeV^2$ in our computations, then γ_1 is certainly always less than about $30 \times 10^3 GeV^2$. For given values of μ_R^2 and μ_I^2, the minimal values of γ_1 occur for θ_0 around $3\pi/2$ or 2π.

The largest possible value of $(v_1 v_2)$ is $v_0^2/2$, which is of the order $30 \times 10^3 GeV^2$. This maxmimum value occurs for values of β_0 around $\pi/4$. The minimum value of $(v_1 v_2)$ in our allowed range lies at $\beta_0 = \pi/2$, which would result in the RHS of (3.3.44) becoming very large in any case (except, of course, in the case of $\mu_3^2 = 0$, where θ_0 is no longer a required parameter of the model.)

Hence, at large values of β_0, and in cases of large $|\mu_3^2|$ with θ_0 lying in the interior of the fourth quadrant, the RHS of the condition will tend to be larger than the upper limit of the LHS of the inequality. In these circumstances, the condition is clearly very weak. However, for smaller $|\mu_3^2|$, and especially with β_0 around the region of $\pi/4$, the condition can become quite tight.

3.3.4 Necessary and sufficient conditions for a physical potential

In all, 7 necessary conditions have been listed — eqns (3.3.27) (two separate conditions), (3.3.30),(3.3.35),(3.3.36),(3.3.39), and (3.3.44). Kastening [REF] has shown that these same conditions are also *sufficient* for the existence of a stable vacuum with the correct symmetry-breaking properties. (The vacuum stability requirement is equivalent to requiring that the potential is bounded below. The correct symmetry-breaking pattern is equivalent to insisting that $V' \geq 0$ for all fields, since the minimum of V' was defined to lie in a direction with parallel doublets.)

In determining whether a particular point in the parameter space lies within the physical region it is essential to check that it satisfies each of these conditions.

The two conditions on λ_1 and λ_2 are implemented in the model by simply restricting them both to the range $(0, 1]$ at the time of choosing their values.

The two conditions which are used to place θ_0 in the fourth quadrant are imposed by restricting the search for solutions for θ_0 (see chapter XX) to the fourth quadrant.

The other three conditions, which are all considerably more complicated in their form than the four just mentioned above, must be checked explicitly by the model. In fact, condition (3.3.44) is tested numerically by checking that all four physical mass-squared values are positive — this overkill provided a safety net in case of numerical or coding problems. The last two conditions are simply checked at appropriate points in the model by evaluation of the expressions in the inequalities.

It is important to note that the symmetry-breaking conditions have been derived at small fields but the vacuum stability conditions depend on behaviour at large fields. Hence, conditions (3.3.27) and (3.3.30) might be derived using the renormalisation group equation to provide "running" parameters, ensuring that the vacuum is stable at least up to the scale of new physics — the SUSY scale, say. However, such fine-tuning here would be unjustified considering the uncertainties which exist in the numerical values of the limits which should be imposed on $m(T_c)/T_c$ and $v(T_c)/T_c$ [REFS].

3.4 Calculations at the critical temperature

Up to this point in the choice of parameters and imposition of constraints we have implicitly been working only at tree level since we have derived constraints from knowledge of the physical vacuum in the present universe, essentially at zero temperature. However, it is the properties of the scalar potential near the critical temperature which are required to determine the nature of the EWPT and whether the potential can support baryogenesis.

Clearly, the critical temperature, its definition and calculation, are of major consequence to the determination of the nature of the potential and the resulting EWPT. Hence, a chapter has been devoted to the description of this part of the work. At this point the reader might feel it worthwhile to consider the chapter on the critical temperature before continuing here. However, the details of the definition and calculation of the critical temperature are not required to follow the remainder of this chapter — we will describe the calculations made *at* the critical temperature without having to describe the critical temperature itself.

3.4.1 Minimisation of the scalar potential at $T = T_c$

Recall that we were unable to diagonalise the tree level mass matrix because, effectively, we could not diagonalise the submatrix formed from the neutral scalar fields. Hence, we are also unable to diagonalise the mass matrix with the first loop correction term in the potential, ie. the "T^2" term. So, we are forced to resort to numerical evaluation of the mass eigenvalues at the T^2 level in order to calculate the $\mathrm{Tr}\,[M^3]$ term. Obviously, then, we have no choice but to perform a numerical minimisation of the potential since our evaluation is not analytical.

One of the most notorious problems in numerical minimisation is finding a *global* minimum — how can one be sure that the minimum one finds is not simply local? We face this problem now since we must find the global minimum of the scalar effective potential at the critical temperature. We describe below the procedure developed to perform this minimisation and the major problem we had to overcome before the procedure could be considered reliable — the discussion of this problem helps to highlight a source of error in the evaluation of the potential.

3.4.2 The numerical minimisation routine

3.4.3 Finding the global minimum

The numerical minimisation routine requires a first estimate of the position of the minimum, from which it follows the route of fastest descent in an attempt to find the minimum. Such an initial point is not necessarily an estimate of the position of the minimum. We might try instead to provide some point from which we feel confident that the route of fastest descent will lead to the minimum — eg. in the case of minimisation at zero temperature (or, equivalently, at tree level) we know that the minimum lies in the quadrant with $v_1, v_2 > 0$ and $\theta_0 \in [3\pi/2, 2\pi]$, and hence we might choose some point lying in the direction $v_1 = v_2 \equiv v$ with $\theta_0 = 7\pi/8$ and set $v = v_0/\sqrt{2}$. Since this point lies in "the centre" of the quadrant where we know the minimum must lie, we might feel confident that the route of fastest descent from this point will lead to the minimum although we have not attempted to estimate where the minimum lies within the quadrant.

However, it is very difficult to see how we might choose such a point beyond tree level due to the complexity of the problem, and with T_c typically of order $100GeV$ it seems unlikely that we could rely on the position of the minimum at zero T as a good estimate for that at T_c. Indeed, we also derived the constraints on θ_0 at zero T and cannot guarantee that θ_{T_c} is similarly constrained. So, we developed a procedure involving multiple minimisations, each started from a different inital point which is chosen at random. We then take the best minimum found by these multiple attempts as our best estimate of the global minimum.

Minimisation of the potential is a numerically intensive calculation because of the need for repeated evaluations of the potential, each of which involves numerical diagonalisation problems. Hence, it is desirable to perform as few minimisations as possible in the search for the global minimum.

To this end, I undertook a simple study of the number of randomly chosen initial points required to achieve the best approximation to the global minimum. Consider a given point in the parameter space of the model. I was interested in estimating how few initial points are required in the minimisation procedure in order to reliably find the global minimum for the potential described at this point in the parameter space. So, take 20 initial points, called $\underline{p}_i$ say, and evaluate the resulting minima, V_i^{min}, defining the best minimum found as the global minimum. Then there are 20 possible subsets of the $\underline{p}_i$ of the form $S(n,l) = \left\{\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_n\right\}$ with $n \leq 20$. These correspond to the sets of initial points which would be used if we only chose n initial points. Define $V^{min}(n,l)$ as the best minimum found from $S(n,l)$. Then, by comparing the $V^{min}(n,l)$ with the global minimum it is possible to optimise the choice of n, the number of initial points, by determining the smallest value of n which returns the global minimum with acceptably few failures.

This analysis was performed for several hundred randomly chosen points in the parameter space and led to the following results.

TABLE showing n=5 as good choice.

As demonstrated in the table above, the minimisation procedure can be optimised for $n = 5$ with only $0.xx\%$ failure rate. As we have already pointed out, the high temperature expansion approximations we are using are good to about 5% at our upper limits on $m(T_c)/T_c$ and hence it can be considered reasonable to accept an uncertainty of $0.xx\%$ in the number of accurately minimised potentials in our parameter space search.

3.4.4 Problems with minimisation

A subtle problem was encountered in the original attempts to minimise the effective potential. The numerical minimisation procedure was found to return a point as its best estimate of the minimum but claim that it had failed to meet all of the conditions for a minimum even though it could find no lower points. This led me to make numerical investigations of the potential at several points in the parameter space, using the computer model to evaluate the potential on a grid of points in $(\sigma_1, \sigma_2, \theta)$-space. These investigations were centred on the "failed" points returned by the numerical minimisation procedure in order to determine why the procedure was failing at these points yet could find no lower points. The resulting figures demonstrated very clearly that the minimisation procedure had indeed found a minimum and it became rapidly clear that there was a lesson to be learned herein.

Evaluating the potentials on ever finer grids it eventually became apparent that there were, in fact, several local minima located in the regions being investigated. However, this fact could not be reconciled with the form of the potential. The local minima were found to lie in a very small region around the point originally returned by the minimsation procedure — the width of the region being several orders of magnitude smaller than the values of the fields in the region. This implies that the potential should have a high frequency component in its form but that is clearly not the case since it is a multi-dimensional, fourth order polynomial. So, it appeared that some uncertainty in the evaluation of the potential must be the source of the apparent rapid variation behaviour.

Since the evaluation of the potential derives from a largely analytical form, there appear to be only two possible sources of this uncertainty. The first is the numerical precision of the machine itself. However, since the calculations involved in the evaluation of the potential are mostly simple evaluations of polynomials, and were performed in double precision FORTRAN, we can rely on them to xx significant figures and the variations observed in the potential were at a level many orders of magnitude larger than this. This leads us to the second possible source of error — the evaluation of the $\mathrm{Tr}\,[M^3]$ term. This relies on numerical estimation of the eigenvalues of the curvature matrix at the point of evaluation. However, the procedure used to find these eigenvalues is known to converge to the eigenvalues slowly (only linearly, in fact) and to converge even more slowly when we have repeated eigenvalues OR ZERO EIGENVALUES (NOT SURE ABOUT

THAT ONE. CHECK IT!) Of course, we have both repeated eigenvalues and even repeated zero eigenvalues in the curvature matrix evaluated at the global minimum. Hence, the evaluation of the $\mathrm{Tr}\,[M^3]$ term is likely to be considerably less accurate than expected from arguments on machine precision.

It is a feature of minimisation routines in general that they must be given a series of conditions on which to base the decision to halt their searches. These conditions compare the value of the function and its (estimated) derivatives at the test-point and a group of "nearby" points. When the function is found to be lower on the test-point than on its neighbours, and the derivatives also have appropriate values (in particular, the matrix of second derivatives is positive definite,) AND the "nearby" points are within a stated distance of the test-point, then the routine will halt its search. When we have a well-behaved function, which can be evaluated to accuracies close to the machine precision, then the tolerances set on the distances between the test-point and its neighbours can be set very small. However, when uncertainties in the evaluations are much larger than the machine precision then the tolerances need to be much larger than we might at first expect. In the case in hand, it was apparent that we should try relaxing the test conditions for the minimum by increasing the tolerances.

Tests of the minimisation procedure with larger tolerances proved very successful — the procedure no longer complained of apparent failure and the minima returned were the same as those found before. However, it became apparent that the optimal values for the tolerances actually depend on the values of μ_R^2 and μ_I^2 and so the final version of the minimisation procedure adapts the tolerances according to the values of μ_R^2 and μ_I^2.

3.4.5 Determination of the nature of the EWPT

Having minimised the effective potential at the critical temperature we are now in a position to determine the nature of the EWPT. With the global minimum found to lie at $(v_1(T_c), v_2(T_c), \theta(T_c))$ then, as described in chapter XX, the condition for a first order phase transition is

$$\frac{v(T_c)}{T_c} > 1 \tag{3.4.46}$$

where $v(T_c)^2 = v_1(T_c)^2 + v_2(T_c)^2$. This is easily checked at this point. However, we have also pointed out in chapter XX that our expansions are only valid in the high temperature limit which corresponds to the following condition on the scalar masses and the top mass —

$$\frac{m_i(T_c)}{T_c} < 1.6 \quad \forall \quad i = 1, \ldots, 8; \mathrm{top} \tag{3.4.47}$$

(The expansions are good to better than 10% up to this limit. CHECK THESE NUMBERS IN PAPER FROM SLAC.)

The $m_i(T_c)$ are calculated numerically as the eigenvalues of the curvature matrix at the global minimum and hence, as explained in the previous section, the values do contain uncertainties due to the inherently slow convergence of the XXXXXX method used to solve the eigenvalue problem CHECK THE XXXXX!. However, these uncertainties are at least one or two orders of magnitude smaller than the eigenvalues (CAN I PROVE THIS?), except in the case of very small eigenvalues which are, in any case, going to meet the condition of eqn(3.4.47). So, these uncertainties will not have any noticeable effect on the number cases passing or failing on condition 3.4.47.

3.5 Summary of the modelling procedure

We can now provide a summary of how the parameter space search is performed.

1 — Choose a point in the parameter space — these points are chosen on a grid in the (μ_R^2, μ_I^2)-plane. The grid is a 15×15 square, with spacing of $1.5 \times 10^3\mathrm{GeV}^2$ in both directions, giving a range of $0 \to 22.5 \times 10^3\mathrm{GeV}^2$ in both μ_R^2 and μ_I^2. A set of n points is examined at each point on the grid. The points are chosen in the form of the "useful" parameter set, described above in section XX. They are then converted to the typical parameter set, as described in section YY.

2 — Check that the chosen point lies in the physical region — we only wish to consider points which lead to physically acceptable potentials. Hence we check that the chosen point meets the criteria which define the physically acceptable region of the parameter space.

NB! If the chosen point is found to lie in the unphysical region then it is immediately abandoned and **is not included** in the count of the number of points examined. This means that the proportion of the **physical** region which also supports baryogenesis is given by the ratio of the number of points found to support baryogenesis to the total number of points examined.

3 — Determine the critical temperature — as described in chapter XX.

4 — Minimise the effective potential at the critical temperature — described above in section xx.

5 — Check the validity of the high temperature expansion — equivalent to imposing the condition 3.4.47. If the high temperature is found to be invalid then we abandon the point but we **do** include it in the count of the number of points examined since it does lie in the physical region.

6 — Determine the nature of the phase transition — the global minimum at $T = T_c$ must meet the condition 3.4.46 in order for the phase transition to be first order.

Points which are found to be physical, to validate the use of the high temperature expansion and to provide a first order phase transition are recorded for later analysis.

We have already defined what we mean by the unphysical, the physical and the baryogenesis regions of the parameter space. Note, however, that there is, in fact, a fourth type of region — the region, contained in the physical region, where the high temperature expansion is invalid. This region, which we will refer to as the "invalid region", cannot be studied by our methods and so it is unknown what proportion of this region might support baryogenesis. Points which we find to lie in the invalid region **are** counted, as explained in part 5 above, and hence we must be clear that the proportion of the physical region which we find to support baryogenesis is actually a lower limit. We also provide a study, in our results, of the proportion of the physical region which lies in the invalid region.

Chapter 4

The Critical Temperature

4.1 Definition of the Critical Temperature

How should we define the critical temperature? Well, clearly this is the temperature at which the phase transition occurs. However, it is not clear what we mean by "the phase transition occurs." For example, how long does the phase transition take? If it is over very quickly then we can safely define a temperature at which the transition is said to have occured. On the other hand, if the phase transition occurs over a relatively long period of time then we might find a *range* of temperatures over which the transition can be said to occur. What does become clear is that we must *define* the temperature at which we consider the phase transition to occur and justify such a definition.

Let us continue by discussing what we might consider as possible definitions of the critical temperature.

i. the phase transition might be said to occur at the temperature (T_1) at which the global minimum of the scalar potential (or the true, symmetric vacuum, if you will) becomes degenerate with the secondary, local minimum (or false, non-symmetric vacuum.) Since,in general, we expect to find a potential barrier between these two minima, this would require the fields to tunnel through the barrier, resulting in a suppressed transition rate and hence a slow phase transition.

However, let us note that, at any temperature above this T_1, even though it might be possible for the fields to tunnel through the potential barrier, it will always be energetically more favourable for the reverse transition to occur and the fields will not condense into the non-symmetric vacuum. Hence, T_1 is certainly the maximum possible value to be considered for the critical temperature.

ii. At some temperature $T_2 < T_1$, the free energy available when the scalar fields tunnel through the potential barrier between the origin and the new, true vacuum, is large enough to overcome the surface energy of the bubbles of new vacuum which are forming. This allows these bubbles to expand indefinitely and this is where we would most likely define the phase transition to occur, or at least to start. However, a detailed definition of this temperature would require a detailed study of the dynamics of the bubbles of new vacuum and this is, in itself, a problem which has been extensively studied outwith the present problem in hand. Certainly, such a study lies beyond the scope of our aims here and so we go on to discuss the third possibility.

iii. We are specifically interested in determining whether the standard model with two Higgs doublets can avoid the wash-out of a baryon asymmetry by sphaleron transitions immediately AFTER the phase transition has occured. Now, although, as described above, the exact value of T_2 would be an extremely complex definition involving far more than a study of the properties of the scalar potential alone, we can be sure that $T_2 > T_3$, where T_3 is the temperature at which the potential barrier between the origin and the true vacuum disappears. A moment's thought reveals that this is simply the temperature at which the origin is no longer a local minimum and becomes a saddle point instead. So, we can define T_3 as the highest temperature at which the curvature matrix of the scalar potential will have a zero eigenvalue — ie., the temperature at which the origin becomes an unstable point for the fields and they can "roll away", in a classical fashion, to the true vacuum. Since there is now nothing to prevent the phase transition occurring, it is clear that T_2 must indeed be greater than T_3, otherwise the phase transition will be second order.

These three situations are represented in *Fig 1*. We must decide on one of these definitions as our critical temperature. Certainly it will be true that a first order phase transition will start at some $T_2 > T_3$ and will finish (if it has not already done so) at T_3, since there is no longer any suppression of the transition at this point.

In our calculations we will use the definition

$$T_c = T_3, \tag{4.1.1}$$

in keeping with the common practice in the literature, and for the following reasons. We wish to study the conditions which exist immediately after the phase transition has occurred. It is certainly true that the phase transition has finished by the time the origin "turns over", as described in (*iii*) above. Any particular choice of T_2, being dependent on the dynamics of the phase transition, is a model-dependent value while T_3 is purely a function of the parameters of the scalar sector. In this way, T_3 tells us something about the conditions after the phase transition which is fundamental to the particular scalar sector under consideration. Also, it is fortunate that T_3 can be derived analytically, as we will show shortly.

We obtain a measure of how strongly first order the phase transition *was* by considering the quantity $v_0(T_c)/T_c$, where $v_0(T_c)$ is the magnitude of the scalar field VEV at T_c (ie., the distance from the origin of the global minimum in the scalar effective potential at T_c.) At this point we turn to the dynamics of the sphaleron transition to consider what we must demand of $v_0(T_c)/T_c$. This has been studied extensively in the literature and we draw the following limit from such studies:

$$\frac{v_0(T_c)}{T_c} \geq 1.0 \tag{4.1.2}$$

THIS WILL BE EXPANDED UPON.

4.2 Derivation of the critical temperature

T_c is the highest temperature at which the curvature of the scalar potential becomes negative in some direction *at the origin,* ie. at which the curvature matrix, evaluated at the origin, becomes non-positive-definite. Hence, we must evaluate the curvature matrix and note that it is considerably simplified at the origin as compared to the general form, allowing us to diagonalise the matrix which we cannot do for the general form. Having determined the diagonal form, or essentially determined the eigenvalues, we can go on to evaluate T_c.

We begin by defining the curvature matrix in question. We will call the matrix C_0, and define it as

$$[C_0]_{ij} = \frac{\partial^2}{\partial\phi_i \partial\phi_j} \left\{ V_0 + \frac{T^2}{24} \times \mathrm{Tr}\left[M^2(\phi, T)\right] \right\} \Big|_{origin} \tag{4.2.3}$$

Note that we only include terms up to the T^2 term. This approach was first suggested by *DOLAN AND JACKIW? NOT SURE*, and neglects the $\mathrm{Tr}\,[M^3]$ term because of the possibility of introducing imaginary terms into the critical temperature. [THIS HAS TO BE EXPANDED UPON.]

Then C_0 has the form

$$C_0 = \begin{pmatrix} A & 0 & 0 & 0 & -\gamma_1 & \gamma_2 & 0 & 0 \\ 0 & A & 0 & 0 & -\gamma_2 & -\gamma_1 & 0 & 0 \\ 0 & 0 & A & 0 & 0 & 0 & -\gamma_1 & \gamma_2 \\ 0 & 0 & 0 & A & 0 & 0 & -\gamma_2 & -\gamma_1 \\ -\gamma_1 & -\gamma_2 & 0 & 0 & B & 0 & 0 & 0 \\ \gamma_2 & -\gamma_1 & 0 & 0 & 0 & B & 0 & 0 \\ 0 & 0 & -\gamma_1 & -\gamma_2 & 0 & 0 & B & 0 \\ 0 & 0 & \gamma_2 & -\gamma_1 & 0 & 0 & 0 & B \end{pmatrix} \tag{4.2.4}$$

where

$$\begin{aligned} A &= -\mu_1^2 + \frac{T^2}{12}\left(6\lambda_1 + 2\lambda_3 + \lambda_4\right) + \frac{T^2}{16}\left(3g^2 + g'^2 + 4g_t^2\right) \\ B &= -\mu_2^2 + \frac{T^2}{12}\left(6\lambda_2 + 2\lambda_3 + \lambda_4\right) + \frac{T^2}{16}\left(3g^2 + g'^2 + 4g_b^2 + \frac{4}{3}g_\tau^2\right) \\ \gamma_1 &= \mu_R^2\cos(\theta) - \mu_I^2\sin(\theta) \\ \gamma_2 &= \mu_R^2\sin(\theta) + \mu_I^2\cos(\theta) \end{aligned} \tag{4.2.5}$$

This can be rotated to the form

$$C_0 = \begin{pmatrix} S & 0_4 \\ 0_4 & S \end{pmatrix} \tag{4.2.6}$$

where S is the matrix

$$S = \begin{pmatrix} A & -\gamma_1 & 0 & \gamma_2 \\ -\gamma_1 & B & -\gamma_2 & 0 \\ 0 & -\gamma_2 & A & -\gamma_1 \\ \gamma_2 & 0 & -\gamma_1 & B \end{pmatrix} \tag{4.2.7}$$

Clearly then, C_0 has *at most* 4 distinct eigenvalues, which are just the eigenvalues of S. We derive these eigenvalues by considering the characteristic equation of S, namely $\det[S - xI_4] = 0$. This gives

$$\left(x^2 - (A+B)x - \left(\gamma_1^2 + \gamma_2^2\right) + AB\right)^2 = 0 \tag{4.2.8}$$

and so we obtain

$$x = \frac{(A+B) \pm \sqrt{(A-B)^2 - 4\left(\gamma_1^2 + \gamma_2^2\right)}}{2} \tag{4.2.9}$$

However, this form is not conducive to determining the critical temperature. Recall that we are interested in finding the highest temperature at which one of these eigenvalues is zero. Clearly, the smaller eigenvalue will go to zero first (ie., the one with the negative square root term.) Hence we can proceed by writing

$$\begin{aligned} (A+B)^2 = & \quad (A-B)^2 - 4\left(\gamma_1^2 + \gamma_2^2\right) \\ \Rightarrow \quad AB + (\mu_R^2 + \mu_I^2) & \quad = 0 \end{aligned} \tag{4.2.10}$$

If we re-write A in the form $A = -\mu_1^2 + A'T^2$, and similarly for B, then eqn(4.2.10) reduces to

$$A'B'T^4 - \left(\mu_1^2 B' + \mu_2^2 A'\right) T^2 + \left(\mu_R^2 + \mu_I^2\right) + \mu_1^2\mu_2^2 = 0 \tag{4.2.11}$$

which leads us to

$$T^2 = \frac{(\mu_1^2 B' + \mu_2^2 A') \pm \sqrt{\left(\mu_1^2 B' - \mu_2^2 A'\right)^2 + 4A'B'\left(\mu_R^2 + \mu_I^2\right)}}{2A'B'} \tag{4.2.12}$$

Hence, we have four distinct possibilities for the critical temperature. However, two of them will be negative (assuming T^2 above is positive, which we will discuss in a moment), which we cannot consider physical. So, our definition of T_c will be the positive temperature such that

$$T_c^2 = \max\left\{\frac{(\mu_1^2 B' + \mu_2^2 A') \pm \sqrt{(\mu_1^2 B' - \mu_2^2 A')^2 + 4A'B'\left|\mu_3^2\right|^2}}{2A'B'}\right\} \tag{4.2.13}$$

Note that we must ensure that T_c is physical by considering the possibility of T^2 above becoming negative, leading to a purely imaginary temperature, or the term inside the square root becoming negative, leading to a complex temperature. It should be clearly understood, however, that these constraints say nothing about whether or not the parameter set in question describes a physically realistic potential, only whether the origin is always (or never) a minimum of the potential. If T_c is not real, and the origin is always at least a local minimum, then it is possible that the phase transition might still proceed by tunnelling and bubble nucleation, and it would certainly be first order. The case where the origin is never a minimum, at any temperature, is unphysical since it implies that the potential is unbounded below. (CHECK THIS!)

We discuss these constraints further in the description of our computer modelling techniques.

Let us complete this section with a comment on the critical temperature in the special cases where we have μ_I^2 and/or μ_R^2 zero. When $\mu_I^2 = 0$, then the critical temperature will look almost exactly the same as the definition above, with $\left|\mu_3^2\right|^2$ replaced by μ_R^2. When μ_R^2 is also zero, then the solutions for critical temperatures reduce drastically to the forms $-\mu_{1[2]}^2 + A'[B']T^2 = 0$ but note that there are still four mathematical solutions for T, it is still possible to have unphysical critical temperatures and the matrix C_0 would still have two distinct eigenvalues. The point which I am trying to demonstrate is that, although the introduction of the μ_3^2 term, especially with μ_3^2 complex, leads to a great deal more complexity in the study of the effective potential, and also makes the T_c calculation appear considerably more daunting at first glance, it has a barely noticeable effect on the complexity of the critical temperature calculations. This can be put down to the remarkable properties of the matrix S which we defined above and we will discuss these properties in considerable detail in the chapters on corrections to the critical temperature.

Chapter 5

Results

Our aim in constructing the two Higgs-doublet standard model was to investigate the possibility of supporting baryogenesis in the electroweak scenario and, in so doing, to circumvent the problem of disagreement between theoretical upper limits and experimental lower limits on the Higgs boson mass. Hence, we are primarily interested in the mass spectrum of the model in the baryogenesis region of the parameter space. However, having found a viable region of the parameter space for baryogenesis, it is also desirable to discuss the structure of the baryogenesis region with a view to understanding how the extension to the standard model has allowed the extra freedom to circumvent the mass problem.

Since the constraints which form the boundaries of the baryogenesis region, namely the conditions on $m(T_c)/T_c$ and $v(T_c)/T_c$, have proved impenetrable in an analytical investigation of the parameter space, it proves useful to compare the population of the physical region with that of the baryogenesis region. This allows the effects of these constraints to be measured and compared by a numerical analysis of the data. To this end, a short study of the physical region was undertaken.

5.1 A study of the physical region

In order to produce data in the physical region it is simply necessary to run the computational model with no cuts on $m(T_c)/T_c$ and $v(T_c)/T_c$. These constraints were developed purely from considerations of the validity of the high-temperature expansion and the requirement of a first order phase transition, and as such they play no part in determining whether the model considers a particular case to be physical or not. So, by simply dropping these constraints we build a knowledge of the population of the entire physical region (within the limits imposed on $|\mu_3^2|$, the λ_i and β_0) as opposed to the baryogenesis region.

A set of data was calculated in this manner — 50 *physical* points were modelled at each point on the 16×16–μ_3^2 grid and the resulting data were analysed in the same fashion as those data computed for the baryogenesis region. Where it proves useful or revealing, this "physical data set" will be compared with the "baryogenesis data set".

The major strength of this approach is in revealing the effects of the $m(T_c)/T_c$ and $v(T_c)/T_c$ constraints. It is important to understand how the effects of these two constraints differ from one another since the condition on $v(T_c)/T_c$ is truly physical — it has been developed by considering the order of the phase transition — while the condition on $m(T_c)/T_c$ represents only a limit in our techniques. Hence, behaviour controlled by the $v(T_c)/T_c$ constraint can be considered as the true physical nature of the model, while behaviour governed by the $m(T_c)/T_c$ condition is to be viewed as uncertain.

It is interesting to compare the effects of the various conditions on vacuum stability and the symmetry-breaking pattern of the vacuum state, which form the boundaries of the physical region (see chapter X). The relative importance of these constraints is simply studied by considering the number of parameter points which fail under each constraint and this study can be used to explain the observed distributions of the coupling constants in the physical data-set. The graphs XX-YY show these distributions.

Let us turn first to the distributions of λ_1 and λ_2. Recall, firstly, that both of these couplings are constrained to be positive — hence the sudden cut-off in frequency as the couplings become negative. The observed steady increase in frequency of physical cases as both λ_1 and λ_2 increase can be attributed to conditions (3.3.30) and (3.3.39), namely

$$\begin{aligned} \min(\lambda_3, \lambda_3+\lambda_4+\lambda_5, \lambda_3+\lambda_4-\lambda_5) &> -2\sqrt{\lambda_1\lambda_2} \\ \left(\Sigma_+ - \frac{2\gamma_1}{v_1v_2}\right) &\leq 2\sqrt{\lambda_1\lambda_2} \end{aligned} \tag{5.1.1}$$

Since $(\lambda_3+\lambda_4\pm\lambda_5)$ lies in the range $[-3,3]$, smaller values of $2\sqrt{\lambda_1\lambda_2}$ will more often be cut than the larger values, tending to suppress the frequency of smaller values of λ_1 and λ_2.

Similarly, since $\gamma_1/(v_1v_2) > 0$, and Σ_+ lies in the range $[-3,3]$, the left-hand side of the second condition above lies in a range $[-x,3]$, where $-x < -3$. This means that the second condition will tend to be easily satisfied but certainly more easily satisfied at larger λ_1 and λ_2. This also contributes to the decreased frequencies at small λ_1 and λ_2.

Also, since both constraints are symmetrical in λ_1 and λ_2, their effect on the respective distributions will be identical. This will result in similar distributions, as observed.

Comparing the distributions of λ_3 and λ_4, it is clearly seen that λ_3 tends to rise more rapidly but is notably less frequent at the lower end of the range. To explain these distributions, turn to condition (3.3.44) which ensures that the charged mass-squared is positive — ie.,

$$\lambda_4 + \lambda_5\cos(2\theta_0) - \frac{2\gamma_1}{v_1v_2} \leq 0 \tag{5.1.2}$$

Hence, the LHS of condition (3.3.39), which has been reproduced above, can be constrained to be less than 1. This is easily seen by re-writing the expression in the form

$$\lambda_3 + \left(\lambda_4 + \lambda_5\cos(2\theta_0) - \frac{2\gamma_1}{v_1v_2}\right) \tag{5.1.3}$$

and noting that the term in brackets was proved negative above and that λ_3 lies in the range $[-1, 1]$.

Since the maximum value of the RHS of (3.3.39) is 2, and we have constrained the LHS to be less than 1 by using condition (3.3.44), it is to be expected that (3.3.39) should be seen to fail on relatively few cases after the imposition of (3.3.44). In the computational model, (3.3.44) was indeed imposed before (3.3.39) and Table X below demonstrates that, as expected, there were very few cases found to fail on (3.3.44) after having passed on (3.3.39).

Having seen that condition (3.3.39) plays a minor role in shaping the physical region in the computational model, the suppression of cases with smaller λ_3 values should be put down to the effects of condition (3.3.30), which can be seen above. In the same way that this constraint depressed the frequency of smaller values of λ_1 and λ_2, it will also suppress smaller values of λ_3 and λ_4.

Lastly, note that the distribution of λ_5 is essentially symmetrical about zero. Considering conditions (3.3.30) and (3.3.44), the possibility of a symmetrical behaviour in the distribution of λ_5 becomes apparent. In (3.3.30), for example, the "minimum function" term can be written as $\lambda_3 + \min(0, \lambda_4 - |\lambda_5|)$. The inclusion of the modulus brackets around λ_5 make it clear that the behaviour of λ_5 in this expression will not depend on its sign — ie., will be symmetrical about zero. In (3.3.44), λ_5 is paired with $\cos(2\theta_0)$. With θ_0 in the fourth quadrant then $\cos(2\theta_0)$ lies in the range $[-1, 1]$. Hence, the sign of the λ_5 term can once again be positive or negative, depending on the value of θ_0, and so the effect on the distribution of λ_5 is to make it symmetrical about zero. Condition (3.3.44) will not contribute a symmetrical factor to the distribution of λ_5 but, as already seen, it does not play such an important part in the formation of the physical region and so its effect will be secondary in the distribution graphs.

The conditions on λ_1, λ_2 and θ_0 are imposed at the time of choosing the parameters and, hence, these conditions are never tested as such. Therefore they are never seen to fail. However, a comparison of the failure rates on tests of the other conditions demonstrate their relative importance in shaping the physical region and/or the "overlap" in their effects.

Tables X and Y give the failure rates for the conditions (3.3.30) and (3.3.39) — condition (3.3.44), ensuring a real, positive charged mass, is NEVER seen to fail.

TABLE

However, assuming that there is an overlap in the effects of the individual conditions, these failure rates are likely to depend on the order in which the conditions are tested.

5.2 The size of the baryogenesis region

Let us begin by discussing the size of the baryogenesis region, in particular as a proportion of the size of the physical region, and also discuss how the the physical region is shaped by the constraints derived to ensure the potential is physical.

As described in chapter XX, we investigated the phase transition by choosing random points in the parameter space, making the choice in the form of the "useful" parameter set $(\mu_R^2, \mu_I^2, \beta, \lambda_i)$. In order to explore the effect of the soft symmetry-breaking term in the potential, μ_R^2 and μ_I^2 were **not** chosen at random but were chosen on a 16×16 grid of points in the μ_R^2, μ_I^2-plane. At each point on the μ_R^2, μ_I^2-grid (extending from $0 \rightarrow 22.5 \times 10^3$GeV in steps of 1.5×10^3GeV,) 10^3 points of the form (β, λ_i) **which lay in the physical region** were chosen and the properties of their corresponding potentials were evaluated.

Number of physical points checked	Number of points found in baryogenesis region	%-age of physical space occupied by baryogenesis region	
$16 \times 16 \times 10^3 = 25.6 \times 10^4$	13117	$5.1 \pm 0.9\%$	(5.2.4)

As can be seen in table XX, of over one quarter million potentials which were checked in the physical region, **13117** points were found to lie in the baryogenesis region — so, the baryogenesis region occupies $5.1 \pm 0.9\%$ of the physical region, with the uncertainty in the number of points in the baryogenesis region calculated simply as $\pm\sqrt{n}$.

Consider now the frequency of baryogenesis-supporting cases as a function of μ_3^2, presented in a 3-dimensional histogram form (which we will refer to as "lego-plots", for obvious reasons) in graph XX. This clearly shows that the introduction of the soft symmetry-breaking term has enormously increased the available parameter space to support baryogenesis. This, in itself, is not enough to ensure that the model overcomes the disagreement between theory and experiment: that relies on the mass distributions, which will be discussed in due course. However, it immediately shows that this extended Higgs' sector is allowing extra freedom in the choice of parameters, suggesting that the possibility of sidestepping the mass problems is at least more likely in this model than in the Standard Model.

Naturally, it is interesting to consider the shape of this distribution of baryogenesis-supporting cases in more detail. It is important to explain how the soft symmetry-breaking term brings about the increase in frequency at small $|\mu_3^2|$, then lets the frequency fall away at larger values. In order to shed some light on this behaviour, it is useful at this point to introduce a set of data produced by working on the whole of the physical region, as opposed to just the baryogenesis region.

5.3 The scalar mass spectrum

It is possible to gain a limited analytical understanding of the scalar mass spectrum by considering the diagonalisation of the zero-temperature mass matrix. However, even if this diagonalisation was possible in full in the most general case (which we shall see shortly it is not), we would still be limited in what we could learn, in an analytical, quantitative fashion, about the mass spectrum. This is because the detailed shape of the spectrum derives from the effects of the constraints on the ratios $m(T_c)/T_c$ and $v(T_c)/T_c$, which are imposed at the critical temperature. We have already seen that it is not feasible even to derive the values of these ratios in individual cases, let alone attempt a general study of the effects of the relevant constraints.

However, we can make certain observations on the *likely* effect of the soft symmetry-breaking term by considering the diagonalisation of the zero temperature mass matrix in the CP-invariant case where $\mu_R^2 \neq 0$ and $\mu_i^2 = 0$. Hence let us briefly consider the diagonalisation of the zero temperature mass matrix.

Let us begin by noting once more that we can reduce the mass matrix to a block diagonal form, where the main-diagonal blocks relate to the charged and neutral sectors respectively. Since the off-diagonal blocks are zero matrices, we can reduce the diagonalisation problem to two 4×4 diagonalisation problems. So, let us look first at the charged sector.

5.3.1 The charged sector masses

This sector diagonalises readily. The full form of the matrix can be found in the appendix XX, but the general structure is

$$M_{1256}^2 = \begin{pmatrix} a & c & 0 & d \\ c & b & -d & 0 \\ 0 & -d & a & c \\ d & 0 & c & b \end{pmatrix} \tag{5.3.5}$$

We form the characteristic equation $\det\left[M_{1256}^2 - m^2 I_4\right] = 0$ to give

$$\left(m^4 - (a+b)m^2 + ab - (c^2 + d^2)\right)^2 = 0 \tag{5.3.6}$$

which in turn returns the solutions for the squared masses, m^2, as

$$m^2 = \frac{a+b}{2} \pm \frac{\sqrt{(a-b)^2 + 4(c^2 + d^2)}}{2} \tag{5.3.7}$$

Having substituted from the full form of the matrix we arrive at the masses

$$\begin{aligned} M_{\chi^\pm}^2 &= 0 \\ M_{H^\pm}^2 &= -\frac{v^2}{2}\left[\lambda_4 + \lambda_5 \cos(2\theta_0) - \frac{2\left(\mu_R^2 \cos(\theta_0) - \mu_I^2 \sin(\theta_0)\right)}{v_1 v_2}\right] \end{aligned} \tag{5.3.8}$$

where the $\chi^\pm$ are the would-be charged Goldstone bosons, “swallowed” by the $W^\pm$ bosons, and the $H^\pm$ are the charged Higgs bosons.

Since we know that we have constrained θ_0 to lie in the fourth quadrant, and $\mu_R^2, \mu_I^2, v_1, v_2 > 0$, we can see that the term $-2\left(\mu_R^2 \cos(\theta_0) - \mu_I^2 \sin(\theta_0)\right)/v_1 v_2$ will be positive and hence we see that the charged scalar mass will tend to increase with $|\mu_3^2|$.

Before discussing this further, consider now the neutral sector.

5.3.2 The neutral sector masses

The full form of the matrix can be found in appendix XX. In this most general form, diagonalisation leads to one would-be Goldstone boson and three massive particles with masses whose forms are impenetrable. However, if we take the simpler, CP–invariant case with $\mu_i^2 = 0$, then the matrix has a structure

$$M_{3478}^2 = \begin{pmatrix} a & 0 & e & 0 \\ 0 & b & 0 & f \\ e & 0 & c & 0 \\ 0 & f & 0 & d \end{pmatrix} \tag{5.3.9}$$

The eigenvalues of this form are readily evaluated following the same procedure as described above. The resulting masses are

$$\left(\mu_R^2 - \lambda_5 v_1 v_2\right) \frac{v^2}{v_1 v_2} \text{and}$$

$$\lambda_1 v_1^2 + \lambda_2 v_2^2 + \mu_R^2 \frac{v^2}{2 v_1 v_2} \pm \sqrt{\left(\lambda_1 v_1^2 - \lambda_2 v_2^2 + \mu_R^2 \frac{(v_2^2 - v_1^2)}{2 v_1 v_2}\right)^2 + \left(\Sigma_+ v_1 v_2 - \mu_R^2\right)^2} \tag{5.3.10}$$

and, of course, $M_{\chi^0}^2 = 0$, the would-be Goldstone boson "swallowed" by the W^0 boson.

Considering the first expression given above we see that the introduction of the soft symmetry-breaking term has resulted in an increase in the mass. Consider now the second expressions. Clearly, the expression of the form $A + \sqrt{B}$ represents the larger mass. It is not so immediately obvious what effect has resulted but, noting that the μ_R^4 terms inside the square root have positive co-efficients, we can conclude that large values of μ_R^2 will certainly result in an increase in the larger mass. At lower μ_R^2, the behaviour of this larger mass will be dependent on the values of $\lambda_1, \lambda_2, \Sigma_+, v_1$ and v_2. However, we should expect that, as μ_R^2 increases and the μ_R^4 terms become more dominant, the mass will tend to increase.

It is more difficult to estimate what will happen in the case of the lower mass in the second expressions. The square root term will tend to decrease the mass as μ_R^2 increases but the term $\mu_R^2 v^2/(2v_1v_2)$, which lies outwith the square root, will tend to increase. Hence we might expect that, whether the observed behaviour is a rise or fall in mass, it will tend to be slower than the rise observed in the other masses.

Although these expressions, and the suggestions drawn from them, were developed in a simplified case, I suggest that we might expect the effect of introducing a real-valued soft symmetry-breaking coupling to mirror the effects of a fully complex-valued coupling. Hence, combining the information on the charged and neutral sectors,it seems likely that we can expect the neutral masses to increase with $|\mu_3^2|$.

However, in the baryogenesis region, the population is biased by the constraints on $m(T_c)/T_c$ and $v(T_c)/T_c$. Hence, *a priori*, we cannot assume that this expected behaviour will be exhibited in the baryogenesis region. Even the population of the physical region is biased by the constraints imposed to ensure a physical potential. However, the physical region is considerably larger than the baryogenesis region and so the population will be more representative of the parameter space as a whole. Hence, we can expect this behaviour to be exhibited over the physical region. This is an important point — since the constraints which form the boundaries of the baryogenesis region are so opaque in an analytical approach, it becomes almost impossible to say anything analytically about the baryogenesis region. However, we can develop a deeper insight into the physical region as a whole and then, by consideration of the numerical results, form a better understanding of how the constraints behave and shape the baryogenesis region.

Hence, it is often constructive to compare the results derived in the baryogenesis region with the same results derived over the whole physical region. To this end, a set of data was produced without using the constraints on $m(T_c)/T_c$ and

$v(T_c)/T_c$. This has the effect of extending the recorded data to fill the physical region.

Presented below are data on the effect of the soft symmetry-breaking term, drawing comparisons between the masses in the physical region and those found in the baryogenesis region. Following this, the observed scalar mass spectra are discussed, again drawing comparisons between the baryogenesis and physical regions.

5.4 The effect of non-zero μ_3^2 on scalar masses

As described already, it is to be expected, in the physical region at least, that the masses of the scalars should tend to increase with $|\mu_3^2|$. In order to investigate this prediction, consider the mean mass as a function of μ_3^2, as seen in graphs XX-YY.

Note firstly that all four physical masses do increase steadily with $|\mu_3^2|$. Now, the effects of the constraints on $m(T_c)/T_c$ and $v(T_c)/T_c$ will certainly be to reduce the overall number of cases appearing in these graphs. However, the way in which the constraints might change the numbers appearing in each bin, and hence change the shape of the graphs, is not so predictable.

5.5 Can we predict the observed behaviour?

Left without a thorough analytical approach to evaluate or manipulate $m(T_c)/T_c$ and $v(T_c)/T_c$, it is not possible to predict with any authority the shape of the scalar mass spectrum or the behaviour of the masses with respect to the soft symmetry-breaking term. However, with the benefit of hindsight, it is possible to show that the observed behaviour is consistent with the behaviour of $m(T_c)$, $v(T_c)$ and T_c.

Considering T_c first - the expression for T_c is (see chapter XX)

$$T_c^2 = \max\left\{ \frac{(\mu_1^2 B' + \mu_2^2 A') \pm \sqrt{(\mu_1^2 B' - \mu_2^2 A')^2 + 4A'B' \left|\mu_3^2\right|^2}}{2A'B'} \right\} \tag{5.5.11}$$

where

$$
\begin{aligned}
A' &= \frac{1}{12}\left(6\lambda_1 + 2\lambda_3 + \lambda_4\right) + \frac{1}{16}\left(3g^2 + g'^2 + 4g_t^2\right) \\
B' &= \frac{1}{12}\left(6\lambda_2 + 2\lambda_3 + \lambda_4\right) + \frac{1}{16}\left(3g^2 + g'^2 + 4g_b^2 + \frac{4}{3}g_\tau^2\right)
\end{aligned}
\qquad (5.5.12)
$$

A few moments consideration of T_c reveals that, to give maximum T_c, the sign of the square root term must be the same as the sign of $A'B'$. Hence, when $A'B'$ is positive, the effect of μ_3^2 is to increase T_c, and when $A'B'$ is negative then T_c is decreased. However, since the λ_i are all typically of the same order, and λ_1, λ_2 are always positive, then $(6\lambda_{1[2]} + 2\lambda_3 + \lambda_4)$ will tend to be positive, as are all other terms in A' and B'. So, $A'B'$ will also tend to be positive, resulting in an increase in T_c with $|\mu_3^2|$. This is indeed the observed behaviour in the physical region, as depicted in graph XX.

However, as T_c increases, we require a corresponding increase in $v(T_c)$ if we are to maintain a first order phase transition. As already stated, the only realistic approach to a study of how $v(T_c)$ behaves is purely numerical and graph XX depicts the mean value of $v = \sqrt{v_1^2 + v_2^2}$ with respect to μ_3^2. The graph clearly demonstrates that $v(T_c)$ does indeed increase with $|\mu_3^2|$.

When considering how $m(T_c)$ behaves, it is important to bear in mind that the physical masses are essentially unrelated to the values $m_i(T_c)$. By this we mean that it is not clear, in any given case, whether the masses are likely to increase or decrease with T because the masses must always be evaluated at the minimum of the potential and we have only a numerically derived knowledge of the position of the minimum. So, an analytical investigation of how $m(T)$ varies with T becomes effectively impossible and it is entirely unclear how $m(T_c)$ will behave. Hence, turning to the numerical data, graphs XX-YY illustrate that the mean values of the mass eigenvalues at T_c increase with $|\mu_3^2|$ *in the physical region.* Note also that the mean values of the eigenvalues at T_c are GREATER/LESS than the physical scalar masses.

When it comes to discussing the results above in the restricted space of the baryogenesis region, we can use the results in the physical region to form an understanding of what can be expected. HAVE TO PRODUCE GRAPHS OF $m_i(T_c)$ BEFORE CONTINUING THIS.

5.6 The scalar mass spectrum

It has already been demonstrated that increases in $|\mu_3^2|$ lead to increases in the scalar masses. However, we have also seen that there are corresponding increases in $m_i(T_c)$ (CHECK THIS!) which reduce the population of the baryogenesis region at larger $|\mu_3^2|$. Hence, the distributions of the physical masses are likely to be shaped at the higher end by the constraint on $m(T_c)/T_c$, while the lower end should be dominated by the cuts on $v(T_c)/T_c$.

Graphs XX-YY show how the scalar mass distributions are constructed. These were plotted using the data set which did not make cuts using the conditions on $m(T_c)/T_c$ and $v(T_c)/T_c$, and hence represent the situation over the physical region as a whole. For each different scalar mass there are five graphs given - these are

i. The mass distribution formed using all of the cases found in the data set - ie., the mass distribution of the physical scalar in the physical region as a whole.

ii. The mass distribution of those cases which would be cut from the data by the constraint on $v(T_c)/T_c$ *only.*

iii. The mass distribution of those cases which would be cut from the data by the constraint on $m(T_c)/T_c$ *only.*

iv. The mass distribution of those cases which would be cut from the data by *both* constraints simultaneously.

v. The mass distribution of those cases which meet both constraints - ie., the mass distribution of the physical scalar in the baryogenesis region.

For each of the four physical particles, part (v) above is reproduced in the graphs XX-YY using the full baryogenesis data set, where the frequencies involved are considerably higher, improving the statistical standing of the distributions.

There are several points to be noted from graphs XX-YY (IE. THE ONES

SHOWING THE CONSTRUCTION OF THE DISTRIBUTIONS). Let us first compare the distributions as seen in the physical region and the baryogenesis region. Obviously, since the baryogenesis region occupies only about 5% of the volume of the physical region, the frequency of masses observed in the baryogenesis region is considerably smaller than in the physical region. However, graph XX demonstrates the relative shapes of the distributions for the heaviest neutral mass by normalising the maximum height of both graphs to unity. This clearly shows that the conditions for baryogenesis have left the shape of the distribution essentially unchanged but have removed a small "head" (ie., those cases with small masses) from the distribution and, much more noticeably, cut off a long and fairly well-populated tail. This is repeated for the other three scalar masses but has not been reproduced graphically.

Secondly, consider in more detail how the head and tail of the distribution, as seen in the physical region, have been removed. The graphs XX-YY (CONSTRUCTION TIME AGAIN!) show clearly that the constraint on $v(T_c)/T_c$ dominates the overall reduction in observed frequencies. However, it is at lower masses that this constraint has (IS THIS BECAUSE HIGHER PHYSICAL MASSES LEAD USUALLY TO HIGHER $m_i(T_c)$?) its greatest effect and is almost entirely responsible for removing the head of the distribution. At larger masses, just beyond the modal mass value, the condition on $m(T_c)/T_c$ comes to dominate and is almost entirely responsible for the removal of the far reaches of the distribution's tail. However, in the more heavily populated part of the tail, both constraints play a very important part in shaping the distribution. (CAN I TIE THIS IN WITH THE PINCHING OBSERVED IN THE MASS RANGES AND SO ON? BOTH IMPORTANT AT HIGH MASS AND HIGH MASSES COME TYPICALLY AT HIGH μ_3^2?)

Hence, bearing in mind that the $m(T_c)/T_c$ constraint is a limitation of the high–T expansion, it is possible that much of the tail observed in the physical region might support baryogenesis in a non–perturbative study of the phase transition. Note, however, that the physical region does display

Chapter 6

A Correction to the Critical Temperature

Until now we have determined T_c by finding the largest temperature at which one of the eigenvalues of the matrix

$$\begin{aligned} C_{ij}^2 &= \frac{\partial^2 V_2}{\partial \phi_i \partial \phi_j} \\ &= \frac{\partial^2}{\partial \phi_i \partial \phi_j} \left\{ V_0 + \frac{T^2}{24} \left(\mathrm{Tr} \left[M_0^2(\phi) \right] + V_g + V_t \right) \right\} \end{aligned} \tag{6.0.1}$$

goes to zero *when evaluated at the origin.* This procedure ignores the term proportional to $\mathrm{Tr}\,[M^3]$, as discussed earlier, and this has been the typical approach to evaluation of T_c in the literature to date. However, it has been suggested lately, as already mentioned, that lattice calculations point towards a poor evaluation of T_c as one of the main drawbacks of the perturbative approach. So, we suggest including the next term in the series when making the evaluation of T_c. We also suggest that this is only sensible, since it is exactly this term which provides the possibilty of first order behaviour and, by neglecting it in T_c, it seems at least possible that we are neglecting an important factor in T_c.

If we want to include the term $T/12\pi \mathrm{Tr}\,[M^3]$ in our evaluation of T_c, then we must determine a correction to the matrix C^2 defined above. We denote this by

C'^2 and define it thus

$$C'_{ij}{}^2 = \frac{-T}{12\pi}\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left\{\mathrm{Tr}\left[M^3(\phi,T)\right]\right\} \tag{6.0.2}$$

Now, since we have already pointed out that $\mathrm{Tr}\,[M^3(\phi,T)]$ cannot be calculated explicitly, we have to proceed in the following manner. We may write $\mathrm{Tr}\,[M^3(\phi,T)] = \sum_{i=1}^{8}\eta_i^3(\phi,T)$ where $\eta_i^2(\phi,T)$ are the eigenvalues of $M^2(\phi,T)$. We may then re-write this in the form $\mathrm{Tr}\,[M^3(\phi,T)] = \sum_{i=1}^{8}[\eta_i^2(\phi,T)]^{\frac{3}{2}}$. We now proceed to differentiate this form

$$\begin{aligned}\frac{\partial}{\partial\phi_i}\left(\mathrm{Tr}\left[M^3(\phi,T)\right]\right) &= \frac{\partial}{\partial\phi_i}\left\{\sum_{k=1}^{8}\left[\eta_k^2(\phi,T)\right]^{\frac{3}{2}}\right\} \\ &= \sum_{k=1}^{8}\frac{3}{2}\left(\eta_k^2(\phi,T)\right)^{\frac{1}{2}}\frac{\partial}{\partial\phi_i}\left(\eta_k^2(\phi,T)\right) \end{aligned}\tag{6.0.3}$$

and

$$\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left(\mathrm{Tr}\left[M^3(\phi,T)\right]\right) = \tag{6.0.4}$$

$$\sum_{k=1}^{8}\left\{\frac{3}{4}\left(\eta_k^2(\phi,T)\right)^{-\frac{1}{2}}\frac{\partial}{\partial\phi_i}\left(\eta_k^2(\phi,T)\right)\frac{\partial}{\partial\phi_j}\left(\eta_k^2(\phi,T)\right) + \frac{3}{2}\left(\eta_k^2(\phi,T)\right)^{\frac{1}{2}}\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left(\eta_k^2(\phi,T)\right)\right\} \tag{6.0.5}$$

At this point we make two observations.Firstly, note that all terms in the elements of $M^2(\phi,T)$ are of even power in the fields ϕ_i. Hence, all terms in the first derivatives of $M^2(\phi,T)$ and $\eta_i^2(\phi,T)$ must be odd powered in ϕ_i. Secondly, we note that we are only interested in evaluating the correction matrix C'^2 *at the origin*, where all scalar fields are zero, and hence we can ignore the terms with factors of $\frac{\partial}{\partial\phi_j}\left(\eta_k^2(\phi,T)\right)$.

So, our correction matrix will effectively reduce to

$$C'_{ij}{}^2 = -\frac{T}{12\pi}\sum_{k=1}^{8}\left\{\frac{3}{2}\left(\eta_k^2(\phi,T)\right)^{\frac{1}{2}}\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left(\eta_k^2(\phi,T)\right)\right\} \tag{6.0.6}$$

We still have a problem — we do not have explicit expressions for the eigenvalues $\eta_k^2(\phi,T)$ since we cannot diagonalise the matrix $M^2(\phi,T)$ in general (at least, not such that the expressions we obtain are realistically useable.) However, let us continue by noting that we only require the expression $\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left(\eta_k^2(\phi,T)\right)$.

Let us point out that we *can* diagonalise $M^2(\phi,T)$ *at the origin*, and that this is just our definition of the matrix C^2. Let $\rho_k^2(T)$ be the eigenvalues of C^2. Then

$$\det\left[C^2 - \mathrm{diag}\left(\rho_k^2(T)\right)\right] = 0 \tag{6.0.7}$$

by definition. Note that this expression contains no field-dependence (since it is evaluated at the origin) but, if we were to include the field-dependent parts of the curvature matrix ($F^2(\phi,T)$, say) and the eigenvalues ($\epsilon_k(\phi,T)$, say), then we would have

$$\det\left[C^2 + F^2(\phi,T) - \mathrm{diag}\left(\rho_k^2(T) + \epsilon_k(\phi,T)\right)\right] = 0 \tag{6.0.8}$$

Please take note of the notation here — we might strictly have written $(\rho_k^2(T) + \epsilon_k^2(\phi,T))$ above but we are going to work extensively with powers of $\epsilon_k(\phi,T)$ and have chosen the notation for clarity.

We know that $F^2(\phi,T)$ is simply defined as

$$F_{ij}^2(\phi,T) = M_{ij}^2(\phi,T) - C_{ij}^2(T) \tag{6.0.9}$$

and we have expressions for $C_{ij}^2(T)$ and $\rho_k^2(T)$ for all i,j,k in $1,\ldots,8$. Hence, in principle, we can evaluate $\epsilon_k(\phi,T)$. However, we already know that this will be unworkable and so we suggest making an expansion of the determinant above, with appropriate small value approximations.

6.1 Expansion of determinant of the curvature matrix

Now,

$$\begin{aligned} & \det\left[C^2 + F^2(\phi,T) - \mathrm{diag}\left(\rho_k^2(T) + \epsilon_k(\phi,T)\right)\right] \\ = \quad & \det\left[C^2 - \mathrm{diag}\left(\rho_k^2(T)\right)\right] + A_{ij}(T)\Psi_{ij} + B_k(T)\epsilon_k(\phi,T) + \ldots \end{aligned}$$

$$
\begin{aligned}
&= \quad A_{ij}(T)\Psi_{ij} + B_k(T)\epsilon_k(\phi, T) + \ldots \\
&\equiv \quad 0
\end{aligned} \tag{6.1.10}
$$

where we are using the notation of summation over the indices i, j, and Ψ_{ij} is simply the second order field-dependent part of F^2_{ij}.

Our ultimate aim is to evaluate $\epsilon_k(\phi, T)$ at the origin, so we can consider the ϕ_i to be small. If we also assume that $\epsilon_k(\phi = 0, T)$ is small compared to $\rho^2_k(T)$, then we can make small value approximations and say

$$
\begin{aligned}
A_{ij}(T)\Psi_{ij} + B_k(T)\epsilon_k(\phi, T) &\approx 0 \\
\Rightarrow \epsilon_k(\phi, T) &= -\frac{A_{ij}(T)}{B_k(T)}\Psi_{ij}
\end{aligned} \tag{6.1.11}
$$

So, it only remains to evaluate $A_{ij}(T)\Psi_{ij}$ — the sum of terms, in the determinant above, which are second order in ϕ_i — and $B_k(T)$ — the co-efficient of $\epsilon_k(\phi, T)$ in the determinant. We will shortly see that the problem is not actually quite this simple but is still tractable, so let us continue with this line of argument.

Consider the matrix $C^2_0(T) = C^2(T) - \mathrm{diag}(\rho^2_k(T))$ in block diagonal form as explained previously in our discussion of T_c.

$$
C^2_0(T) \equiv C^2(T) - \mathrm{diag}\left(\rho^2_k(T)\right) = \begin{pmatrix} S & O_4 \\ O_4 & S \end{pmatrix} \tag{6.1.12}
$$

where

$$
S = \begin{pmatrix}
A - \rho^2_k(T) & -\gamma_1 & 0 & \gamma_2 \\
-\gamma_1 & B - \rho^2_k(T) & -\gamma_2 & 0 \\
0 & -\gamma_2 & A - \rho^2_k(T) & -\gamma_1 \\
\gamma_2 & 0 & -\gamma_1 & B - \rho^2_k(T)
\end{pmatrix} \tag{6.1.13}
$$

We will obtain the second order field-dependent terms by simply evaluating the determinants of the matrices

$$
\left[X^2_{ij}\right]_{kl} = \left[C^2_0(T)\right]_{kl} + \Psi_{ij}\delta_{ik}\delta_{jl} \tag{6.1.14}
$$

(Please note that X^2_{ij} is a matrix, not an element of a matrix. The indices i, j label the matrix, not its elements.)

In principle, this will lead to 64 terms. However, a short discussion of the properties of $C_0^2(T)$ will drastically reduce this number. Firstly, let us note that the only terms in $\det[X_{ij}^2]$ which are proportional to Ψ_{kl} are just those from the k,l–th signed co-factor of X_{ij}^2, which is exactly equivalent to the k,l–th signed co-factor of C_0^2. So, we can proceed by determining the signed co-factors of C_0^2. Secondly, recall that the determinant of a matrix can be defined as the sum of all possible terms (with appropriate signs) which can be formed as the product of one, and only one, element from each row and each column of the matrix. This definition will prove useful in heavily reducing our workload.

We begin by considering the k,l–th co-factors, with the k,l–th element lying in one of the off-diagonal zero-blocks. Having removed the k–th row and l–th column, we are left with a matrix which no longer has block diagonal form, and it is impossible to choose one, and only one, element from each row and each column without choosing at least one zero. Hence, every term in these co-factors will be zero and so each of these co-factors are also zero. This accounts for 32 of the possible 64 terms mentioned above!

When we consider the k,l–th co-factor, with the k,l–th element lying in one of the main diagonal S–blocks, then we are left with a block diagonal form after removing the k–th row and l–th column. Note, however, that we can make use of the symmetry in $C_0^2(T)$ here. For example, the $(1,1)$–co-factor will be equivalent to the $(5,5)$–co- factor, and so on. Also, the $(1,1)$–co-factor of C_0^2 will simply be the $(1,1)$–co-factor of S multiplied by the determinant of S, and so on. So, we should evaluate the determinant of S and all of its signed co-factors.

The determinant of S is

$$\det[S] = \left(AB - (A+B)\rho_k^2(T) + \rho_k^4(T) + \gamma_1^2 + \gamma_2^2\right)^2 \tag{6.1.15}$$

and the matrix of co-factors is

$$S_c = \begin{pmatrix} B-\rho_k^2(T) & \gamma_1 & 0 & -\gamma_2 \\ \gamma_1 & A-\rho_k^2(T) & \gamma_2 & 0 \\ 0 & \gamma_2 & B-\rho_k^2(T) & \gamma_1 \\ -\gamma_2 & 0 & \gamma_1 & A-\rho_k^2(T) \end{pmatrix} \times \left(AB-(A+B)\rho_k^2(T)+\rho_k^4(T)+\gamma_1^2+\gamma_2^2\right) \qquad (6.1.16)$$

where $[S_c]_{ij}$ is the i,j–th co-factor of S.

Now, the determinant of C_0^2 will clearly be the square of the determinant of S. But, *by definition*, the determinant of C_0^2 is zero. Hence, we conclude that the determinant of S is also zero and then

$$\left(AB-(A+B)\rho_k^2(T)+\rho_k^4(T)+\gamma_1^2+\gamma_2^2\right) = 0 \qquad (6.1.17)$$

This is the factor which appears in all of the entries of S_c and so we see that ALL of the co-factors of $C_0^2(T)$ *and* S are zero. In the notation of eq.(6.1.10), this indicates that $A_{ij}(T)=0$ for all i,j. What about $B_k(T)$? Well, the co-efficient of $\epsilon_k(\phi,T)$ is just a special case of the work we have done. The terms proportional to $\epsilon_k(\phi,T)$ in the expansion in eq.(6.1.10) are just those from the determinants of the matrices

$$\left[Y_k^2\right]_{ij} = \left[C_0^2(T)\right]_{ij} - \epsilon_k(\phi,T)\delta_{ki}\delta_{kj} \qquad (6.1.18)$$

and so Y_k^2 will behave like X_{kk}^2, and the co-efficient of $\epsilon_k(\phi,T)$ will be the sum of the diagonal co-factors of $C_0^2(T)$. Hence, $B_k(T)=0$ also!

6.2 The expansion to second order

Where does this leave our approach to approximating $\epsilon_k(\phi,T)$? Well, these results on $A_{ij}(T)$ and $B_k(T)$ mean that there are no first order corrections to the determinant of the curvature matrix evaluated at the origin but it does not imply that $\epsilon_k(\phi,T)$ is zero — simply that we must go to higher orders in our expansion of the determinant in order to calculate $\epsilon_k(\phi,T)$. We will now go on to show that we must, in fact, expand the determinant to fourth order in small values before we find non-zero co-efficients. In order to clarify the arguments that we are about to go through, we will develop a notation for working with sub-matrices of $C_0^2(T)$ which will allow us to discuss co-factors and terms within co-factors, and so on. We will develop this notation by considering the co-efficients of the terms in the determinant which are fourth order in the fields or, equivalently, second order in Ψ_{ij}.

What will the co-efficient of $\Psi_{ij}\Psi_{kl}$ be? Well, a little thought reveals that it is just the (appropriately signed) determinant of the submatrix $[\overline{C_0^2(T)}]_{\{ik/jl\}}$ formed by removing the i–th and k–th rows, and j–th and l–th columns of $C_0^2(T)$. We will simplify the notation for the submatrix to $[\overline{C^2}]_{\{ik/jl\}}$ and will later extend it to consider the co-efficients of terms like $\Psi_{ij}\Psi_{kl}\Psi_{mn}$ by writing $[\overline{C^2}]_{\{ikm/jln\}}$.

So, we want to know the determinants of all possible matrices $[\overline{C^2}]_{\{ik/jl\}}$. Note, first of all, that we cannot have $i = k$ or $j = l$ because the determinant never contains terms with two factors from the same row or column. Having noted this, we can go on to show that ALL of these determinants are zero! Again, we can resort to simple arguments, as opposed to direct calculation, to prove this. There are only 5 situations to consider —

1. Both (i, j) and (k, l) are in the same S–block. Then $[\overline{C^2}]_{\{ik/jl\}}$ has a block diagonal form, with a complete S–block as one of the main diagonal blocks. Hence the determinant will contain a factor of $\det[S]$ which is identically zero. So, all determinants of this form are zero.

2. (i, j) and (k, l) are in different S–blocks. Then $[\overline{C^2}]_{\{ik/jl\}}$ has a block diagonal form again, but now the two blocks on the main diagonal are both submatrices of S and the determinant will be a product of two of the co-factors of S. Since these co-factors are all zero, the determinant will also be zero. So, all determinants of this form are zero.

3. Both (i, j) and (k, l) are in the same zero-block. Then $[\overline{C^2}]_{\{ik/jl\}}$ does not have a block diagonal form and it is impossible to choose one, and only one, element from each row and column of the $[\overline{C^2}]_{\{ik/jl\}}$ without choosing one to be

zero. Hence the determinant will always be zero.

4. (i,j) and (k,l) are in different zero-blocks. Then $[\overline{C^2}]_{\{ik/jl\}}$ has a block diagonal structure exactly like that in case(2), and so the determinant is always zero.

5. (i,j) and (k,l) can lie one in each type of block; ie. (i,j) in one of the S-blocks and (k,l) in one of the zero-blocks, or vice-versa. Then $[\overline{C^2}]_{\{ik/jl\}}$ will always have a structure which is not block diagonal and we have the same situation as in case(3). So, the determinant is always zero.

This covers all possibilities and so we see that all co-efficients of $\Psi_{ij}\Psi_{kl}$ are zero. Again, the co-efficient of $\epsilon_k^2(\phi,T)$ is just a special case of the results we have just given and so there is also no co-efficient of $\epsilon_k^2(\phi,T)$. However, the case of the co-efficient of $\epsilon_k(\phi,T)\Psi_{ij}$ deserves comment. Recall that $\epsilon_k(\phi,T)$ only ever appears in the diagonal terms of the curvature matrix. If Ψ_{ij} appears in the same diagonal term (ie. $i=j=k$) then clearly we can never form a term proportional to $\epsilon_k(\phi,T)\Psi_{kk}$ in the determinant and so we have no co-efficient for such terms. Every other possible placement of $\epsilon_k(\phi,T)$ and Ψ_{ij} is exactly equivalent to one of the cases described in cases(1–5) above and again we will have no co-efficient for these terms.

We conclude that there are no non-zero terms in the expansion of eq(6.1.10) which are of second order in small values. We proceed to the third order terms.

6.3 The expansion to third order

We now consider the co-efficients of terms like $\Psi_{ij}\Psi_{kl}\Psi_{mn}$. Extending the ideas we have already seen, these co-efficients will simply be the determinants of the matrices $[\overline{C^2}]_{\{ikm/jln\}}$. We claim now that all of these determinants are zero and will prove this in a manner similar to the proof given in the last section. Notice now that we cannot allow any pair of i,k,m to be equal, nor any pair of j,l,n.

1. We might have all of (i,j), (k,l) and (m,n) in the same S-block. Then $[\overline{C^2}]_{\{ikm/jln\}}$ has a block diagonal structure with S as one of the diagonal blocks.

This introduces a factor det$[S]$ into the determinant, which is identically zero, and hence all determinants of this form are zero.

2. We might have two of (i,j), (k,l) and (m,n) in the same S-block and the third in the other S-block. Then $[\overline{C^2}]_{\{ikm/jln\}}$ has a block diagonal structure with a submatrix of S as one of the diagonal blocks. This introduces one of the co-factors of S as a factor in the determinant and these co-factors are all zero. Hence, all determinants of this form are zero.

3. We might have all of (i,j), (k,l) and (m,n) in the same zero-block. Then $[\overline{C^2}]_{\{ikm/jln\}}$ does not have a block diagonal form and it is impossible to choose one, and only one, element from each row and each column of $[\overline{C^2}]_{\{ikm/jln\}}$ without choosing one of them to be zero. Hence, all determinants of this form will be zero.

4. We might have two of (i,j), (k,l) and (m,n) in the same zero-block and the third in the other zero-block. Then $[\overline{C^2}]_{\{ikm/jln\}}$ does not have a block diagonal form and the determinant will be zero by the same line of argument used in case(3). Hence, all determinants of this form will be zero.

5. We might have two of (i,j), (k,l) and (m,n) in the same S-block and the third in a zero-block. Then $[\overline{C^2}]_{\{ikm/jln\}}$ does not have a block diagonal form and the determinant will be zero, as in cases(3,4). Hence, all determinants of this form will be zero.

6. We might have two of (i,j), (k,l) and (m,n) in the same zero-block and the third in an S-block. Then $[\overline{C^2}]_{\{ikm/jln\}}$ does not have a block diagonal form and the determinant will be zero, as in cases(3–5). Hence, all determinants of this form will be zero.

7. We might have two of (i,j), (k,l) and (m,n) in different S-blocks, (ie. one in each S-block), and the third in a zero-block. Then $[\overline{C^2}]_{\{ikm/ljn\}}$ does not have a block diagonal form and the determinant will be zero, as in cases(3–6). Hence, all determinants of this form will be zero.

8. We might have two of (i,j), (k,l) and (m,n) in different zero-blocks, (ie. one in each zero-block), and the third in an S-block. Then $[\overline{C^2}]_{\{ikm/ljn\}}$ does not have a block diagonal form and the determinant will be zero, as in cases(3–7).

Hence, all determinants of this form will be zero.

Consideration of the co-efficients of $\epsilon_k^3(\phi,T)$, $\epsilon_k^2(\phi,T)\Psi_{ij}$ and $\epsilon_k(\phi,T)\Psi_{ij}\Psi_{kl}$ follow on directly from cases(1–8) and the discussion at the end of the last section. Again, all of these co-efficients are identically zero.

So, we have proved, as pointed out in section 6.2, that we must go to fourth order in small values, in our expansion of the determinant of the curvature matrix at the origin, before we are able to evaluate $\epsilon_k(\phi,T)$.

6.4 The expansion to fourth order

We come now to the fourth order terms like $\Psi_{ij}\Psi_{kl}\Psi_{mn}\Psi_{pq}$. The co-efficients of these terms are not all zero, although most of them are. However, there are so many non-zero terms that this would be effectively impossible to handle if we did not make the following observation. We are ultimately interested in evaluating the terms $\frac{\partial^2}{\partial\phi_i\partial\phi_j}(\epsilon_k(\phi,T))$ *at the origin.* So, we are only interested in terms in $\epsilon_k(\phi,T)$ which are of second order in the scalar fields. In our fourth order expansion of the curvature matrix at the origin we have,formally

$$\epsilon_k^4 + \epsilon_k^3\Psi_{ij} + \epsilon_k^2\Psi_{ij}^2 + \epsilon_k\Psi_{ij}^3 + \Psi_{ij}^4 = 0 \tag{6.4.19}$$

Since we will evaluate the second derivatives of $\epsilon_k(\phi,T)$ at the origin, we can be sure that the Ψ_{ij} are as small as we like

Chapter 7

Appendix : The matrices

$$
\begin{aligned}
M^2_{1256} &= \begin{pmatrix} a_{1256} & 0 & c_{1256} & d_{1256} \\ 0 & a_{1256} & -d_{1256} & c_{1256} \\ c_{1256} & -d_{1256} & b_{1256} & 0 \\ d_{1256} & c_{1256} & 0 & b_{1256} \end{pmatrix} \\
a_{1256} &= -\mu_1^2 + \lambda_1 v_1^2 + \lambda_3/2 v_2^2 \\
b_{1256} &= -\mu_2^2 + \lambda_2 v_2^2 + \lambda_3/2 v_1^2 \\
c_{1256} &= \lambda_4/2 v_1 v_2 + \lambda_5/2\cos(2\theta_0) v_1 v_2 - \mu_R^2 \cos(\theta_0) + \mu_I^2 \sin(\theta_0) \\
d_{1256} &= \mu_R^2 \sin(\theta_0) + \mu_I^2 \cos(\theta_0) - \lambda_5 \sin(2\theta_0) v_1 v_2
\end{aligned} \tag{7.0.1}
$$

$$
\begin{aligned}
M^2_{3478} &= \begin{pmatrix} a_{3478} & e_{3478} & f_{3478} & g_{3478} \\ e_{3478} & b_{3478} & -g_{3478} & h_{3478} \\ f_{3478} & -g_{3478} & c_{3478} & j_{3478} \\ g_{3478} & h_{3478} & j_{3478} & d_{3478} \end{pmatrix} \\
a_{3478} &= -\mu_1^2 + 3\lambda_1 v_1^2 + \frac{1}{2}\left(\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)\right) v_2^2 \\
b_{3478} &= -\mu_1^2 + \lambda_1 v_1^2 + \frac{1}{2}\left(\lambda_3 + \lambda_4 - \lambda_5 \cos(2\theta_0)\right) v_2^2 \\
c_{3478} &= -\mu_2^2 + 3\lambda_2 v_2^2 + \frac{1}{2}\left(\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)\right) v_1^2 \\
d_{3478} &= -\mu_2^2 + \lambda_2 v_2^2 + \frac{1}{2}\left(\lambda_3 + \lambda_4 - \lambda_5 \cos(2\theta_0)\right) v_1^2 \\
e_{3478} &= \lambda_5/2 \sin(2\theta_0) v_2^2 \\
f_{3478} &= \left(\lambda_3 + \lambda_4 + \lambda_5 \cos(2\theta_0)\right) v_1 v_2 - \mu_R^2 \cos(\theta_0) + \mu_I^2 \sin(\theta_0) \\
g_{3478} &= -\lambda_5 \sin(2\theta_0) v_1 v_2 + \mu_R^2 \sin(\theta_0) + \mu_I^2 \cos(\theta_0)
\end{aligned}
$$

$$
\begin{aligned}
h_{3478} &= \lambda_5 \cos(2\theta_0) v_1 v_2 - \mu_R^2 \cos(\theta_0) + \mu_I^2 \sin(\theta_0) \\
j_{3478} &= -\lambda_5/2 \sin(2\theta_0) v_1^2
\end{aligned}
\tag{7.0.2}
$$