Radboud Universiteit

Physics and Astronomy

Department of High Energy Physics

Alternative models matched to the data of GW150914

Author:
M.J.J. Magnée (s1053603)

Supervised by:
Béatrice Bonga
Badri Krishnan

June 24, 2022

Abstract

The gravitational wave event GW150914 was the first event ever measured by a gravitational wave detector. It is found that this gravitational wave signal is caused by the inspiraling of two black holes orbiting around each other. It is a very loud event, in the sense that the signal can be easily distinguished from the background noise. This makes the data from the event therefore easy to work with, for example to test if totally different sources, with different frequency behaviour in time, could fit the data as well as the original binary model does. So, this thesis aims to invent astrophysically likely or unlikely models and to confront them with the data obtained from the GW150914 event. For each different model, the gravitational wave energy flux is linked to the mechanical energy loss, from which a relationship can be obtained between the frequency and the time. This obtained relationship is fitted to the data obtained from the LIGO detectors in Hanford and Livingston. This is done first for the original binary model, after which 2 different models are studied: a binary system with slightly time varying mass, and a rotating rigid spirally shaped mass distribution.

Contents

Introduction **6**

I Theoretical Background 7

1 Theoretical concepts of gravitational waves **8**
1.1 Linearized Gravity . . . 8
1.2 gauge transformations . . . 8
1.3 quadrupole formula . . . 9
1.4 Conversion between wave frame and source frame . . . 9
1.5 Measuring the gravitational wave amplitude at a detector . . . 10
1.6 Energy of a gravitational wave . . . 13

2 Model of inspiraling binary: original derivation **15**
2.1 Introduction . . . 15
2.2 Setup of the system . . . 15
2.3 Energy change of the system . . . 16
2.4 The gravitational wave frequency as a function of time . . . 17
2.5 Deriving the amplitude as function of time . . . 17
2.5.1 Relating the amplitude of a gravitational wave to the physical distance . . . 17

3 Derivation of alternative models **19**
3.1 Binary system with time dependent mass . . . 19
3.2 Rotating rigid spirally shaped mass distribution . . . 19
3.2.1 Form of the mass distribution . . . 19
3.2.2 Inertia and quadrupole moment tensor in restframe . . . 20
3.2.3 Inertia and quadrupole moment tensor in the (constant) source frame . . . 21
3.2.4 Energy loss due to gravitational wave emission . . . 22
3.2.5 Kinetic energy of the rotating spiral . . . 22
3.2.6 Potential energy of the rotating spiral . . . 23
3.2.7 Energy balance; obtaining the frequency as function of time . . . 23

4 Processing Gravitational Wave Data **25**
4.1 The used gravitational wave data sets . . . 25
4.1.1 Data set from Livingston and Hanford . . . 25
4.2 Whitening the signal . . . 26
4.2.1 Determining the Amplitude Spectral Density . . . 26
4.2.2 The whitening process . . . 27
4.3 Frequency estimation . . . 27
4.3.1 Combining both data sets from Livingston and Hanford in one frequency plot . . . 27
4.3.2 Obtaining the frequency as function of time in GW data set . . . 29
4.4 Amplitude estimation . . . 29
4.4.1 Finding the maxima of a signal . . . 29
4.4.2 Determining the correct value of the amplitude as function of time . . . 29
4.5 Extracting uncertainties . . . 30

4.5.1 Uncertainties in the data points . . . 30
4.5.2 Uncertainties in the period of the signal . . . 31
4.5.3 Uncertainty in the frequency . . . 32
4.5.4 Uncertainties in the amplitude of the signal . . . 32

5 Bayesian Statistics: Probability Density Function for fit parameters 33
5.1 Finding the best fit: least squares . . . 33
5.1.1 Bayes' rule . . . 33
5.1.2 Prior . . . 33
5.1.3 Likelihood function . . . 35
5.1.4 (Log)posterior . . . 35
5.1.5 Uncertainties in least squares fit . . . 36
5.1.6 Marginal probability density . . . 36

II Results 37

6 Results for the original model: Binary 38
6.1 Determining the Mass and the coalescence time . . . 38
6.1.1 Fit function . . . 38
6.1.2 Relating the fit parameters to the model parameters . . . 38
6.1.3 Prior for model and fit parameters . . . 39
6.1.4 Plot of the data and fit: uniform prior for the fit parameters . . . 40
6.1.5 Posterior for fit parameters; uniform prior for fit parameters . . . 40
6.1.6 Posterior for fit parameters; non uniform prior for fit parameters . . . 41
6.1.7 The values of the model parameters . . . 42
6.2 Estimate the power in the amplitude time plot . . . 42
6.2.1 Estimation of the power for the data from Livingston . . . 42
6.2.2 Estimation of the power for the data from Hanford . . . 43

7 Results for the binary with Time varying mass 45
7.1 Fit function . . . 45
7.2 Relating model parameters to fit parameters . . . 45
7.3 Appropriate prior for the fit and model parameters . . . 46
7.3.1 Uniform prior for the model parameters . . . 46
7.3.2 Uniform prior for the fit parameters . . . 47
7.4 Plot and fit of the data . . . 47
7.5 Posterior for fit parameters . . . 48
7.5.1 Uniform prior for the fit parameters . . . 48
7.5.2 Non uniform prior for the fit parameters . . . 49

8 Results for the spirally shaped mass 51
8.1 Fit function used . . . 51
8.2 Relating model parameters to fit parameters . . . 51
8.3 Prior for the model and fit parameters . . . 51
8.3.1 Non uniform prior for the fit parameters . . . 51
8.3.2 Uniform prior for the fit parameters . . . 52
8.4 Plot and fit of the data: uniform prior for fit parameters . . . 53
8.5 Posterior probability density function for the fit parameters . . . 53
8.5.1 Uniform prior for the fit parameters . . . 53
8.5.2 Non uniform prior for the fit parameters . . . 54

III Conclusion **56**

9 Conclusion **57**
9.1 The original model . . . 57
9.2 The original model with time dependent chirp mass . . . 57
9.3 Spiral model . . . 58

10 Discussion **59**

Bibliography **60**

A Probability Identities **61**
A.1 Coordinate Transformations . . . 61
A.2 Uncertainties . . . 61

B Derivations **62**
B.1 Transforming a tensor to the Transverse Traceless gauge . . . 62
B.2 Quadrupole and Inertia tensor . . . 62
B.3 Combining a sine and cosine wave . . . 63

C Determining the effective physical distance **64**
C.1 Determining the effective physical distance for Livingston . . . 64
C.1.1 Fit function . . . 64
C.1.2 Relating fit parameters to model parameters . . . 64
C.1.3 Prior for model and fit parameters . . . 64
C.1.4 Plot and fit of the data . . . 65
C.1.5 The values of the model and fit parameters . . . 66
C.1.6 Likelihood function for the fit parameters; uniform model parameter prior . . . 66
C.2 Determining the effective physical distance for Hanford . . . 67
C.2.1 Fit function . . . 67
C.2.2 Relating fit parameters to model parameters . . . 67
C.2.3 Prior for model and fit parameters . . . 67
C.2.4 Plot and fit of the data . . . 67
C.2.5 The values of the model and fit parameters . . . 67
C.2.6 Likelihood function for the fit parameters; uniform model parameter prior . . . 68

Introduction

Theoretical gravitational wave research is done for around one hundred years by now, with Dr. Albert Einstein himself at the foundations of it. He published his theory of General Relativity around 1915, whereas he proceeded working on his theory after his publication. He could explain several features of nature, that could not be explained with the laws of Newton, such as gravitational waves. Albert Einstein predicted this feature, although there were no instruments present at that time to test his prediction experimentally. Nowadays, there are multiple detectors with the ability to detect gravitational wave signals, such as the Laser Interferometer Gravitational-Wave Observatory, abbreviated as LIGO. It is worth mentioning briefly how the current state of the art gravitational wave detectors are founded from the beginning and what the future plans of gravitational wave detection will be.

In the 1960's, researchers Mikhail Gertsenshtein and Vladislav Pustovoit start thinking to built a light interferometer in Moscow Russia, whereas several years later Joseph Weber and Rainer Weiss in America came up with a likewise idea. Both ideas arose in order to test Einstein's prediction: the existence of gravitational waves. But a real working detector was not built at that time. In 1990, the first funding for the LIGO observatory was given, and in 1992 the locations are chosen to build a detector around Livingston and one around Hanford, United States of America. These detectors belong to the first generation of gravitational wave detectors. From 1994 up until 1998 LIGO was built at these two locations, whereas the installation was done from 1999 up until 2002. The initial LIGO observations start at around 2002, but until 2010 there were no detections of gravitational waves at all. At the year 2004, the project 'enhanced LIGO' started, after which a multi-year shutdown took place to improve LIGO even more, called 'advanced LIGO'. Mid September of 2015, the 14th of September to be exactly, advanced LIGO start doing its first science observations, with a sensitivity of four times the sensitivity the initial LIGO detectors had. Furthermore, there exist another detector called VIRGO, which is built around the year 2000. This is also a Michelson Interferometer, just like LIGO. This detector is located near Pisa in Italy. All these three current existing detectors together form the LIGO-Virgo Network (abbreviated as LVN), which will be expanded in the future by the addition of several detectors to be built. For Example, in 2020, a fourth gravitational wave detector became operational (see Ref. [9]). This detector is located in Kamioka Observatory in Gifu Prefecture, Japan. This project is called Kamioka Gravitational Wave Detector, abbreviated as KAGRA. It has recently joined the LVN (in March 2022), such that the LVN network currently consists of 4 detectors. In the future, one of the projects that will join is called LIGO india (Ref. [7]), which is currently being built in Aundh, India. First, the LIGO detectors in Hanford and Livingston will be updated to the so called Advanced+ version of the LIGO advanced project, after which they will become operational in this updated version in 2026. Soon after this, the detector in Aundh India will also become operational in the updated version (Advanced+ version). These belong to the second generation of gravitational wave detectors. Another detector, called the Einstein Telescope, is planned to be built in Europe. This will become a gravitational wave detector of the third generation. The sensitivity of this detector will be an order of magnitude larger than the sensitivity of the second generation gravitational wave detectors, whereas it will also have a wider accessible frequency band. With this third generation detector, much more accurate measurements can be done, and also gravitational wave signals from objects with larger masses can be measured, due to the improved sensitivity (Ref. [6]).

As is already mentioned, on September 14th in 2015, the first gravitational wave event was measured, given the name GW150914. Both signals measured at the detectors in Livingston and Hanford show a very strong correlation, that sticks out above the background noise, such that it must be a gravitational wave signal. The researchers of the LIGO collaboration paper Ref. [1] applied a simplified approach

to estimate the parameters of this model. The actual parameter estimation is more complex, but this simplified approach works really well. It is determined that this signal comes from the merging of two 30 solar mass black holes that orbited around each other, where their separation distance becomes smaller as time evolved. The event lasted roughly $0.2s$ and the frequency rose up from 35 to 250Hz within 8 cycles. This signal is in the audible range, and can be resembled with the 'chirp' of a bird. The mass parameter $\mathcal{M}$ that shows up in the calculations, is therefore named the 'chirp mass'. This parameter is a certain function of both the masses in the binary system, which has a very specific form due to the fact that the parameters with the dimension 'mass' are showing up and are collected in this form. More insights about the chirp mass can be found in the first chapters of the first part of this Thesis.
After this first event, many more events were measured with these LIGO detectors, but this first signal is very unique in the sense that it is a loud signal. The wave pattern can be recognised very easily in the data (after a whitening process), whereas this is hard for the signals detected from other events. This makes this event very approachable to work with. The frequency increase as a function of time can be estimated, where also the amplitude increase in time can be estimated. In general, the frequency can not be estimated very well for a usual gravitational wave signal obtained by a LIGO detector, because the signal can be recognised as easily as it can be done for this signal. This means that the zero crossings of the wave signal cannot be determined very accurately. Nevertheless, for this signal the estimations of the frequency *can* be done, because this gravitational wave signal is a loud signal. After a whitening process, the signal can be clearly distinguished from the background noise. The wave pattern can be seen very easily in the whitened signal (i.e. the signal really looks like a sinusoidal curve with an increasing frequency), such that zero crossings of the signal can be determined very precisely. From this, a frequency estimation can be done by using the time difference between successive zero crossings. For other signals, the wave signal does not look much like a sinusoidal curve, due to the large influence of the background noise. Also the power law between the gravitational wave strain (amplitude) and the time is estimated from the data. This is done by looking at the maxima of the signal, from which the amplitude increase can be estimated. In this relation, the effective distance can be estimated, but for this a very accurate value of the order of magnitude of the gravitational wave amplitude is needed.

This master thesis aims to invent crazy, (astro) physically likely or unlikely models that produce gravitational waves and to confront them with the data from the very first gravitational wave event GW150914. The basic theory here that is used, is the original Theory of General Relativity published by Albert Einstein around 1915. This thesis does not describe alternative theories for gravity, but only covers gravitational wave sources within the (linearized) General Theory of Relativity. For each model, a theoretical derivation is done, from which a relationship is found between the frequency of the gravitational wave and the time. This relation is then fit through the data of the event GW150914, in order to check if the crazy model in question could produce gravitational waves in the same way as the binary black hole merger of the original event did. For each model, only the mass, time of merging (the coalescence time) and in one case the mass rate of change are to be determined. The most probable values that are found for these parameters and their probability density functions are used to determine if a model can be fit well to the data (it is physically likely) or if a model does not fit the data really well (it is physically improbable). So, the research question is: Could other models match the data from the gravitational wave event GW150914 as well as the original model of the inspiraling binary system does?

In the first part, the basic concepts of gravitational wave theory are derived and explained. First, the original model is derived. This is done in order to get familiar with processing gravitational wave data and to determine, in the same way as is done in Ref. [1], a value for the physical parameters called the chirp mass and the coalescence time. Also the power law between the dimensionless gravitational wave strain and the time is estimated for both the data from Hanford and from Livingston, which was not done by Ref. [1]. After this model, an inspiraling binary system is considered, but this time the chirp mass is made time dependent, such that a third unknown joins the other two. As a third model, a spirally shaped rotating rigid mass distribution is theoretically derived, which as again two unknown parameters. After this, it is explained how the gravitational wave data is processed with Python and also the Bayesian statistics used to match the theoretical models to the data is explained. Thereafter, all three models are matched to the data in the second part, after which the most probable values and likelihood function as well as marginal probability density functions are shown for the fit parameters. It is determined how likely these values for the fit and model parameters are for the chosen model. An elaborated conclusion

and discussion about the models and their obtained values can be found in the third part of this thesis.

Part I

Theoretical Background

Chapter 1

Theoretical concepts of gravitational waves

1.1 Linearized Gravity

In the General theory of Relativity, spacetime curvature is related to the mass and energy. In ten equations, Einstein established how matter tells spacetime how to curve, and how spacetime tells matter how to move (Ref. [4]):

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{1.1}$$

with $R_{\mu\nu}$ the Ricci tensor, R the Ricci scalar, $g_{\mu\nu}$ the metric tensor of the spacetime and $T_{\mu\nu}$ the stress energy tensor. The last tensor describes the matter/energy distribution present in a spacetime. The left hand side together is defined as the *Einstein tensor*: $G_{\mu\nu}$. These so called Einstein Equations are non-linear and second order differential equations, which makes them hard to solve. But in physics, there is one very powerful tool: approximation by perturbation. In this method, a spacetime is assumed to be a background spacetime, with a small perturbation caused by a certain source $\mathcal{S}$ (whose energy/momentum distribution is described by $T_{\mu\nu}$). The ten Einstein Equations are linearized in the sense that for the metric, the following ansatz is filled in into Equation 1.1:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}, \text{ with } |h_{\mu\nu}| \ll 1. \tag{1.2}$$

One can see here that the metric is written as the background Minkowski metric ($\eta_{\mu\nu}$) with a small perturbation ($h_{\mu\nu}$).

1.2 gauge transformations

In the linearized theory of Gravity, there are certain gauge transformations that can be chosen that simplify calculations, but do not change the physics (hence it is a gauge transformation). Because there are ten linear independent Einstein Equations, there are ten independent components of the perturbation (both due to the fact that all the tensors in Equation 1.1 are symmetric tensors). With the help of the gauge freedom, eight of the ten components can be fixed. There are several possibilities, but for the purpose of this thesis, the Lorenz gauge together with the transverse traceless gauge are used. In each inertial system one considers, these gauge transformations can be done. First, the Lorentz gauge is applied. The linearized Einstein Equations take then the form:

$$\Box\bar{h}_{\mu\nu} = \frac{16\pi G}{c^4}T_{\mu\nu}, \tag{1.3}$$

where $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$ is the so called *trace reversed metric perturbation* and the h is the trace of the metric perturbation in Minkowski space. When knowing this 'h bar', one can easily find the metric perturbation by using the reversed trace property. This gauge fixing takes away four out of the total

eight gauge degrees of freedom.

There is still additional gauge freedom left. This is called the *transverse traceless gauge*, and it is defined such that the radiative part of the trace reversed metric perturbation satisfies for a non-stationary source (i.e. depending on time):

$$\bar{h}_{0\mu} \equiv 0 \text{ and } \bar{h} \equiv 0. \tag{1.4}$$

In the transverse traceless gauge, the metric perturbation is traceless, such that holds $\bar{h}_{\mu\nu} = h_{\mu\nu}$. Furthermore, each temporal component of the perturbation is 0, i.e. $h_{ti} = h_{tt} = 0$, such that only the spatial part of the metric perturbation needs to be taken into account. These gauge transformations can take the radiative part of the metric perturbation to:

$$h_{\mu\nu} = h^{(+)}e^{(+)}_{\mu\nu} + h^{(\times)}e^{(\times)}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h^{+} & h^{\times} & 0 \\ 0 & h^{\times} & -h^{+} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.5}$$

In each frame, the metric perturbation can be transformed into the form of Equation 1.5 due to the additional gauge freedom. This form of $h_{\mu\nu}$ is *extremely helpful and handy in the frame where a gravitational wave is propagating in the z-direction*, because this frame will enter the detector. In this gauge, the distinction can be made between + and × polarization, which is useful in the frame of the propagating wave. This is due to the fact that a detector is assumed to be able to distinguish between these 2 polarizations. It is important to note that in general the metric perturbation is not yet in this form in the frame where the wave is travelling in the z-direction. In Appendix B.1 it can be seen how to transform a general tensor into the transverse traceless gauge in a general frame of reference.

1.3 quadrupole formula

The solution of Equation 1.3 is expressed in terms of the mass quadrupole moment tensor, for convenience defined here as:

$$Q_{ij} \equiv \int \rho(\vec{x}) \left[x_i x_j - \frac{1}{3}|\vec{x}|^2 \delta_{ij} \right] \mathrm{d}^3x, \tag{1.6}$$

This tensor is defined as a traceless tensor, which stays traceless when a coordinate transformation is applied between two (with respect to each other) stationary inertial frames. This form appears in the derivation of the so called *quadrupole formula*. The quadrupole formula was discovered by Albert Einstein himself in 1918. He linked the metric perturbation to the change in time of the mass quadrupole moment. He did this using the conservation of energy-momentum of the source and the divergence theorem of Gauss. The quadrupole formula in all its glory is given by:

$$h_{ij}(t, r) = \frac{2G}{c^4 r} \cdot \ddot{Q}_{ij}(t - r/c). \tag{1.7}$$

In this equation, t is the time measured on a clock at a distance r from the source and the dubble dot above Q_{ij} represents the second time derivative. This is known as the famous quadrupole formula. This is only valid for compact sources causing a small perturbation on the background spacetime. Compact sources are sources that have a small diameter compared to the distance between the source and the point where the gravitational wave is regarded, i.e. $d_{\mathcal{S}} \ll r$. Because the quadrupole moment tensor Q_{ij} appearing in Equation 1.7 is defined to be traceless (and stays traceless in each frame that is used here), the h_{ij} in Equation 1.7 is immediately traceless too (such that the bar above the h_{ij} is not needed, because $h_{ij} = \bar{h}_{ij}$ if the first one is traceless). This gives that the bar can be left out in the quadrupole formula (Equation 1.7) Keep in mind that the h_{ij} in the quadrupole formula is already traceless, but not necessarily transverse yet. To make it transverse, see Appendix B.1 how to take a general tensor to the transverse traceless gauge.

1.4 Conversion between wave frame and source frame

In this thesis, only sources are considered that rotate in one plane around a constant rotational axis. For convenience, a frame is introduced where z-axis is aligned with the rotational axis, and thus in which the

mass distribution rotates in the xy-plane. The mass quadrupole moment tensor is calculated the easiest way in the frame in which the mass distribution is rotating around its own z-axis due to rotational symmetries, whereas the most convenient way to express the metric perturbation is in the TT gauge in the frame in which the wave propagates in its z-direction. Via 2 constant coordinate rotations around coordinate axes, the frame in which the mass distribution is rotating around its z-axis can be linked to the frame of the wave. See Figure 1.1 for a visualization of the needed coordinate rotations and the relation between the source frame and the wave frame.

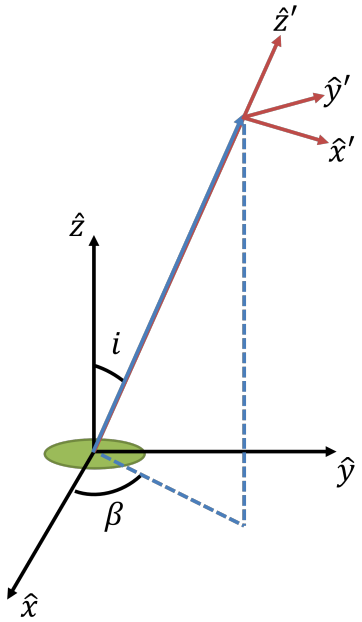


Figure 1.1: Relation between the source frame (black coordinate system) and the wave frame (red coordinate system). The source is represented as the green circle, which is small compared to the distance between source and detector. It is assumed to be rotating around the $\hat{z}$ axis in the xy-plane. The 2 coordinate rotations are denoted with β and i, where i is the inclination angle. The propagation vector of the wave is aligned with the usual spherical unit vector $\hat{r}$ defined in the source frame.

Because in this thesis the used mass distributions are assumed to stay within the xy-plane and rotate around the z-axis in the source frame, the coordinate rotation in the β direction is therefore in fact redundant; it operates as an offset in this direction, because the mass distribution is rotating in this direction. After these rotations, the second time derivative is taken from the obtained quadrupole moment tensor. Then it is evaluated at the retarded time, and as a last step the obtained tensor is taken to the transverse traceless form using the steps in Appendix B.1. In this form, the h^+ and $h^\times$ can be extracted. Further on in this thesis, the quadrupole moment and from this the metric perturbation are calculated for each different model.

In the wave frame, the $\hat{x}'$ and $\hat{y}'$ axis can be oriented differently around their $\hat{z}'$-axis for each situation, as long as both axes are perpendicular to each other and as long as they form a right handed coordinate system. This rotation is elaborated on later and is denoted as the polarization angle ψ.

1.5 Measuring the gravitational wave amplitude at a detector

It is important to note that the distance travelled by the wave $|\vec{x}| \equiv r$ is much larger than the radius of the Earth, i.e. $r \gg R_e$, where R_e is the radius of the Earth. This means that the waves passing the Earth (and thus the waves entering the detector) can be seen as plane waves. The metric perturbation

tensor can then be decomposed into plane waves as a Fourier series as follows:

$$h^{ij}(t,\vec{x}) = \sum_{A=+,\times} e_A^{ij} \int_{-\infty}^{\infty} \mathrm{d}f\ \tilde{h}_A(f) e^{-2\pi i f(t-\hat{n}\cdot\vec{x}/c)} \approx \sum_{A=+,\times} e_A^{ij} \int_{-\infty}^{\infty} \mathrm{d}f\ \tilde{h}_A(f) e^{-2\pi i f t}, \tag{1.8}$$

with e_A^{ij} the $+$ or $\times$ polarization tensors of Equation 1.5. In the last step, the fact is used that the wave length λ of the gravitational wave is much larger than the size of the detector d, i.e. $d \ll \lambda$. Therefore, the wave is assumed to be constant in space at sizes of the order of the size of the detector d (i.e. h_{ij} does not have a spatial dependence at the detector). The choice of the origin of the coordinate system in which this wave is described can therefore be chosen arbitrarily. The origin is chosen to be located at the detector, such that $x = 0$. The output of the detector is a scalar, whereas h^{ij} is a tensor. The output of the detector will have the form:

$$h(t) = \sum_{A=+,\times} \mathcal{D}^{ij} e_{ij}^A h_A(t) = \sum_{A=+,\times} F_A(\theta,\phi,\psi) h_A(t), \tag{1.9}$$

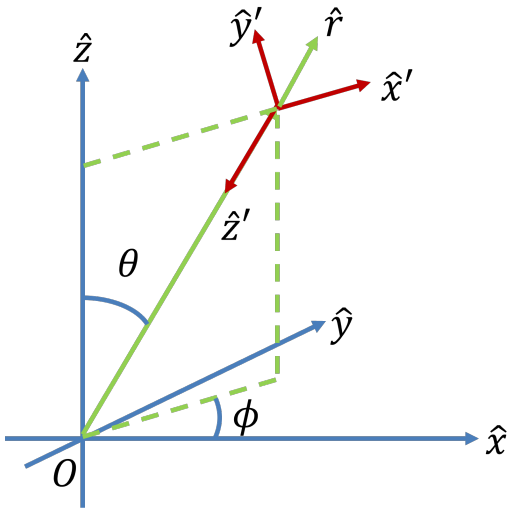
where $\mathcal{D}^{ij}$ is the detector tensor translating the metric perturbation h^{ij} to a scalar gravitational wave strain h_A. Here A can be either $+$ or $\times$. This detector tensor is for a detector with one arm along the x-axis and the other arm along the y-axis of the detector frame (such as the detectors in Livingston and Hanford) defined as:

$$\mathcal{D}_{ij} = \frac{1}{2}(\hat{x}_i\hat{x}_j - \hat{y}_i\hat{y}_j). \tag{1.10}$$

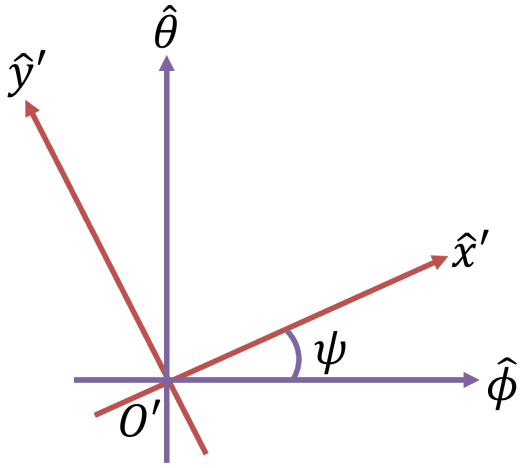
The minus sign is due to the fact that in the detector, one needs to calculate (and measure) the path difference between the arms along the x- and y-axis. One wants to measure a symmetric and trace free tensor (this is how the metric perturbation is constructed using the transverse traceless gauge), such that the detector tensor also shares these properties. The input for the detector can therefore be written as:

$$h_{\mathcal{D}}(t) = \frac{1}{2}(\hat{x}_i\hat{x}_j - \hat{y}_i\hat{y}_j)e_+^{ij}h_+^{\mathcal{D}} + \frac{1}{2}(\hat{x}_i\hat{x}_j - \hat{y}_i\hat{y}_j)e_\times^{ij}h_\times^{\mathcal{D}} = \frac{1}{2}\left[h_+^{xx} - h_+^{yy}\right]\Big|_{\mathcal{S}_{\mathcal{D}}}, \tag{1.11}$$

where the subscript $\mathcal{D}$ denotes that this quantity is calculated in the frame of the detector. Both h_+^{xx} and h_+^{yy} must be in the frame of the detector. There is thus one transformation needed in order to obtain the gravitational wave amplitude seen in the frame of the detector as function of the gravitational wave amplitudes in the frame of the wave. This is the transformation described in Figure 1.2 below:



(a) Figure that visualizes the coordinate rotations from the frame of the wave (red frame) to the frame of the detector (blue frame). There is a rotation needed with an angle ϕ around the $\hat{z}$-axis, and a rotation with an angle θ around the $\hat{y}$ axis. After this, the system needs to be rotated around the new obtained $\hat{z}$-axis with an angle ψ. The propagation vector of the wave $\hat{k}$ is in the frame of the source aligned with the $\hat{z}'$ axis, but is aligned with the spherical unit vector -$\hat{r}$ defined in the detector frame.



(b) Figure that visualizes the polarization angle ψ. The $\hat{\theta}$ and $\hat{\phi}$ axes are the in the detector frame defined spherical unit vectors in the θ and ϕ-direction respectively. The polarization angle ψ is defined as the angle between the $\hat{x}'$ unit vector in the wave frame and the $\hat{\phi}$ vector defined as the spherical unit vector in the ϕ-direction in the detector frame (defined in the usual way).

Figure 1.2: Coordinate transformation between wave frame and detector frame. The detector frame is denoted with the blue coordinate system, whereas the wave frame is denoted with the red coordinate frame. The $\hat{\theta}$ and $\hat{\phi}$ are the usual spherical unit vectors defined in the detector frame.

From source to detector, this means that a rotation with an angle ψ around the $\hat{z}'$ axis is needed first, to align the $\hat{x}'$-axis with the $\hat{\phi}$ unit vector (and to align the $\hat{y}'$-axis with the $\hat{\theta}$ unit vector). This rotation takes the polarization amplitudes in the wave frame to:

$$h_+ \to h_+ \cos(2\psi) + h_\times \sin(2\psi) \tag{1.12a}$$
$$h_\times \to h_\times \cos(2\psi) - h_+ \sin(2\psi), \tag{1.12b}$$

With this rotation, the $\hat{x}'$ axis is aligned with the $\hat{\phi}$ unit vector. Now, a rotation around the z-axis with an angle ϕ is needed, followed by a rotation around the y'-axis with an angle θ (keep in mind that the $\hat{z}'$ axis is oriented in the negative $\hat{r}$-direction defined in the detector frame). Applying these rotations, one obtains a relationship between the metric perturbation tensor in the frame of the detector as function of the position in the sky (θ, ϕ), as well as the $+$ and $\times$ polarization amplitudes in the frame of the wave. Taking the rotation in Equation 1.12 into account, the gravitational wave strain detected at the detector becomes also a function of the polarization angle ψ. As can be seen from Equation 1.11, only the difference between h_+^{xx} and h_+^{yy} in the detector frame needs to be calculated. With the calculation rules for rotations around coordinate axes as is visualized in Figures 1.2 and 1.1, the following is obtained for this difference:

$$\begin{aligned} h_{\mathcal{D}}(t) =& h_+ \underbrace{\left[\cos(2\psi)\cdot\left(\frac{1+\cos^2(\theta)}{2}\right)\cos(2\phi) - \sin(2\psi)\cdot\cos(\theta)\cdot\sin(2\phi)\right]}_{F_+(\theta,\phi,\psi)} \\ &+ h_\times \underbrace{\left[\cos(2\psi)\cdot\cos(\theta)\cdot\sin(2\phi) + \sin(2\psi)\cdot\left(\frac{1+\cos^2(\theta)}{2}\right)\cos(2\phi)\right]}_{F_\times(\theta,\phi,\psi)}, \end{aligned} \tag{1.13}$$

where θ and ϕ represent the position in the sky with respect to the detector and ψ the polarization angle of the wave. Here F_+ and $F_\times$ are antenna response functions for the $+$-polarization and $\times$-polarization

respectively. These antenna response functions are different for each different detector, because each detector is oriented in a different way with respect to the wave (i.e. for each detector, the θ and ϕ may be different). If the quadrupole moment tensor is calculated in the source frame, transformed to the wave frame and put into the TT gauge form, Equation 1.7 can be used to determine the h_+ and $h_\times$ in the wave frame. With these, one can calculate the wave strain $h_\mathcal{D}$ that is measured by the detector, if one knows the angles θ and ϕ at which the wave is coming in. When more and more detectors are used, the estimation of θ and ϕ for each detector can be made more and more precise.

In this thesis, it is enough to work with a polarization angle $\psi = 0$, but for completeness it is taken into account in this derivation. If $\psi = 0$ (see Ref: [5]), the expressions for the antenna response functions in Equation 1.13 reduce to the expressions of the antenna response functions below:

$$h_\mathcal{D}(t) = \frac{1}{2}\left(h_+^{xx} - h_+^{yy}\right)\Big|_\mathcal{D} = h_+ \cdot \underbrace{\left(\frac{1+\cos^2(\theta)}{2}\right)\cos(2\phi)}_{\equiv F_+(\theta,\phi,\psi=0)} + h_\times \cdot \underbrace{\cos(\theta)\sin(2\phi)}_{\equiv F_\times(\theta,\phi,\psi=0)}. \tag{1.14}$$

For longer lasting signals, the rotational motion of the Earth needs to be taken into account, because the detector is then moving with respect to the gravitational wave signal. This can be done via a transformation to the Geogentric frame of the Earth. The signal that is used here last approximately 0.2 seconds, such that the rotational influence is really small, but for longer signals, this needs to be taken into account.

1.6 Energy of a gravitational wave

The Einstein Equations are linearized to obtain an equation for the small perturbation of the metric $h_{\mu\nu}$ caused by the *source only*, i.e. the energy momentum tensor $T_{\mu\nu}$ only contains energy and momentum of the source. This is due to the fact that only compact sources are considered, such that the energy momentum tensor $T_{\mu\nu}$ is of the same order as the metric perturbation, i.e. $|T_{\mu\nu}| << 1$. If the Einstein equations are expanded up to second order on both sides in Equation 1.1, the first order terms again result in Equation 1.3, whereas the second order on the left hand side is related to the second order part of the energy momentum tensor. As is already explained, the source only gives a first order term in the energy-momentum tensor, such that can be concluded that the second order term of the energy-momentum tensor must be the energy-momentum of the gravitational wave itself because this phenomenon also contains energy and momentum, which is much smaller compared to the energy and momentum of the source. This second order part of the energy momentum tensor is denoted with $t_{\mu\nu}$.

Energy in general relativity is in first place not really meaningful, because one can transform at each spacetime point to a freely falling frame to undo all gravitational effects. To make a meaningful definition of energy, one should average over the second order terms of the Einstein tensor (Ref. [4]):

$$\left\langle G_{\mu\nu}^{(2)} \right\rangle = \frac{8\pi G}{c^4} t_{\mu\nu}. \tag{1.15}$$

Expanding the metric to a second order, collecting all second order terms and time averaging the result, gives one the tools to determine the energy flux of a gravitational wave. In the end, this leads to the energy per unit time per solid angle:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{r^2c^3}{32\pi G}\left\langle \dot{h}_{ij}^{TT}\dot{h}_{TT}^{ij} \right\rangle = \frac{r^2c^3}{16\pi G}\left\langle \dot{h}_+^2 + \dot{h}_\times^2 \right\rangle, \tag{1.16}$$

with the metric perturbation tensor and time derivative of the quadrupole moment tensor in transverse traceless gauge. In transverse traceless gauge, these polarization amplitudes are defined, such that the relation in Equation 1.17 only holds in TT gauge. It does not matter in which frame the $h_{\mu\nu}$ is changed to the transverse traceless gauge (as is explained before), but this is most convenient in the frame in which the wave is travelling in the $\hat{z}$-direction. The i and j here only run over the spatial part of the metric perturbation tensor, because each (partly) temporal component is 0. To obtain a differential equation

for the frequency evolution in time, one integrates Equation 1.16 over the angles β and i in spherical coordinates to obtain the total radiated energy per unit time:

$$\frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}t} = \int_0^{2\pi} \int_0^{\pi} \frac{\mathrm{d}P}{\mathrm{d}\Omega}(i, \beta) \cdot \sin(i) \mathrm{d}i \mathrm{d}\beta \tag{1.17}$$

Here i and β are chosen to represent the usual θ and ϕ spherical angles, but are named different because θ and ϕ are already reserved for the position of the source in the sky with respect to the detector.

After obtaining the energy of the gravitational wave per unit of time, one equates this to the mechanical energy *loss* per unit time of the source. From this, an equation can be derived for the frequency as function of time. This is done for the different models later on.

Chapter 2

Model of inspiraling binary: original derivation

2.1 Introduction

This thesis aims to find alternative models that are confronted with the data from event GW150914, but first the theoretical frequency-time relation of the original model is derived, after which this model is matched to the data. This is done in the same way as the researchers of the LIGO collaboration paper Ref. [1] did in a simplified way as the original parameter estimation. This thesis does not replace the work done by the researchers of this article, but it is a good way to get familiar with deriving theoretical models and with processing gravitational wave data in a simplified version.

2.2 Setup of the system

Assume two black holes orbiting around each other. One black hole may have a larger mass than the other (without loss of generality one can define one to have a larger mass). The orbit they follow is assumed to be small and circular, and they are assumed to cause a small perturbation on the Minkowski background metric. Together with the fact that the source is far away from Earth (i.e. the radial distance between Earth and source denoted with r is much larger than the source diameter d_S), this also means that this source can be assumed to be a compact source, such that the quadrupole Formula in Equation 1.7 (valid for compact sources) can be used. This source is thus assumed to cause only a small perturbation, which means that this source can be treated Newtonian. It satisfies therefore the laws of Kepler. As can be found in mechanics textbooks like Ref. [4], the 2 body problem is split into 2 separate problems: one equation of motion for the centre of mass, and one equation of motion for the reduces mass μ which is defined as: $\mu \equiv (M_1 M_2)/M$, with M the total combined mass. The orbit is circular with radius a, which means that the law of Kepler takes the following form for this binary system:

$$a^3 = \frac{MG}{\omega^2} \tag{2.1}$$

See Figure 2.1 to see how the mass distribution looks like at a time t. The orbit is in the xy-plane without loss of generality, where the z-axis is pointing out of the paper. The masses stay in this plane due to the conservation of angular momentum; angular momentum cannot be changed significantly by gravitational waves at the time scales of the merging process (Ref. [1]).

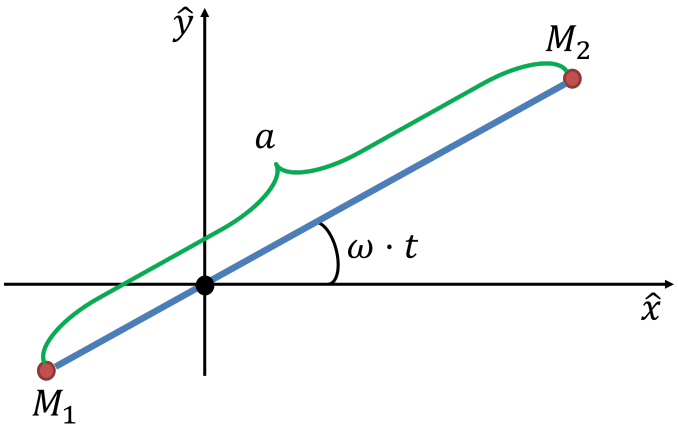


Figure 2.1: Geometry of the inspiraling binary system. Without loss of generality mass M_1 is assumed to be the larger mass and M_2 is assumed to be the smaller mass. The radius a and orientation at time t can be seen in the figure. The orientation at time $t = 0$ (the starting configuration) can be chosen arbitrarily by fixing the coordinate system initially at a different angle.

The quadrupole moment tensor in the frame of the source is calculated to be:

$$Q_{ij} = \frac{\mu a^2}{2} \cdot \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{pmatrix} + \frac{\mu a^2}{2} \cdot \begin{pmatrix} \cos(2\omega t) & \sin(2\omega t) & 0 \\ \sin(2\omega t) & -\cos(2\omega t) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.2}$$

The coordinate transformation described in Section 1.4 is a constant, time independent coordinate transformation. It therefore does not matter if the time derivative is taken first or if this coordinate transformation is done first. The second time derivative of the quadrupole moment tensor of Equation 2.2 is taken, and the obtained tensor is taken into the transverse traceless (TT) gauge, after which the metric perturbation in TT gauge is calculated via the quadrupole formula of Equation 1.7. The following form of the polarization amplitudes is obtained:

$$h_+ = -\frac{4\mu G a^2 \omega^2}{c^4 r} \cdot \left[\frac{1 + \cos^2(i)}{2}\right] \cdot \cos(2\omega(t - r/c)) \tag{2.3a}$$

$$h_\times = -\frac{4\mu G a^2 \omega^2}{c^4 r} \cdot \cos(i) \cdot \sin(2\omega(t - r/c)). \tag{2.3b}$$

With these polarization amplitudes, the following is obtained for the gravitational wave power per solid angle (using Equation 1.16, see Ref. [4]):

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{\mu^2 G a^4 \omega^3}{2\pi c^5} \cdot \left[1 + 6\cos^2(i) + \cos^4(i)\right], \tag{2.4}$$

such that the total radiated energy per unit of time equals (use Equation 1.17):

$$\frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}t} = \int_0^{2\pi} \int_0^{\pi} \frac{\mathrm{d}P}{\mathrm{d}\Omega} \cdot \sin(i) \mathrm{d}i \mathrm{d}\beta = \frac{32 G \mu^2 a^4 \omega^6}{5c^5} \tag{2.5}$$

Keep in mind that in this derivation, the angular frequency is only slightly changing in time on time scales of one single rotation T, but it does change significantly on larger time scales $\tau \gg T$, time scales of multiple rotations.

2.3 Energy change of the system

As we know from Keplers law for circular orbits, the potential energy is minus 2 times the kinetic energy, such that the total mechanical energy of two black holes orbiting around each other in circular orbits

with radius a and rotational angular frequency ω equals:

$$E_{\text{mech}} = -\frac{1}{2}\frac{GM\mu}{a}. \tag{2.6}$$

As this binary black hole system shrinks, the radius obviously decreases. Because of Keplers law (which is in fact conservation of angular momentum), this means that the binary spins up. With the help of Equation 2.1, the increase of the radius per unit of time (it is decreasing, so the increase is negative) can be linked to the increase in frequency per unit of time. The total increase in mechanical energy as a function of the time derivative of the frequency then becomes:

$$\frac{\mathrm{d}E_{\text{mech}}}{\mathrm{d}t} = \frac{GM\mu}{a^2}\cdot\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{GM\mu}{a^2}\cdot -\frac{2}{3}\frac{MG}{a^2\omega^3} = -\frac{2}{3}\cdot\frac{M^{2/3}G^{2/3}\mu}{\omega^{1/3}}\cdot\dot{\omega} \tag{2.7}$$

Because Equation 2.7 is the increase in mechanical energy and Equation 2.5 represents the energy that is *taken away* by the gravitational wave, this means there needs to be a minus sign in front of Equation 2.7 to represents the mechanical energy loss. Equating these gives the following differential equation:

$$\dot{\omega} = \left(\frac{96}{5}\right)\cdot\frac{\omega^{11/3}}{c^5}\cdot(G\mathcal{M})^{5/3}, \tag{2.8}$$

where $\mathcal{M}$ is the Chirp mass defined as:

$$\mathcal{M} \equiv \frac{(M_1M_2)^{3/5}}{(M_1+M_2)^{1/5}} = \mu^{3/5}\cdot(M_1+M_2)^{2/5}. \tag{2.9}$$

The chirp mass $\mathcal{M}$ is the effective mass of the binary system. To compare: In Celestial mechanics, the reduced mass μ is the effective mass. In the two body problem in Celestial mechanics, the problem of two bodies orbiting around each other is reduced to a problem where a single mass μ is orbiting around an object that has a mass equal to the sum of the 2 masses. In this Inspiraling Binary problem, there is effectively a single object with a mass $\mathcal{M}$ that is orbiting another object standing still. This system then has a quadrupole moment tensor of Equation 2.2, where the orbital radius a is decreasing.

2.4 The gravitational wave frequency as a function of time

In the quadrupole moment of Equation 2.2, it can be seen that the angular frequency of the gravitational wave equals twice the rotational angular frequency (this holds as well for the normal frequency f). This means that the following ω in Equation 2.8 needs to be substituted by the following to obtain a differential equation for the frequency of the gravitational wave:

$$\omega = \frac{\omega_{\text{GW}}}{2} = \frac{2\pi f_{\text{GW}}}{2} = \frac{f_{\text{GW}}}{\pi}. \tag{2.10}$$

The solution of the differential equation for f_{GW} is the evolution of the gravitational wave frequency in time (Ref. [1]):

$$f_{\text{GW}}^{-8/3} = \frac{(8\pi)^{8/3}}{5}\cdot\left(\frac{G\mathcal{M}}{c^3}\right)^{5/3}\cdot(t_0-t). \tag{2.11}$$

In this formula, $\mathcal{M}$ is the chirp mass defined in Equation 2.9 and t_0 is the time of coalescence (the time when the binary is merged into one massive black hole). For convenience, the mass will be measured in solar masses. This is due to the fact that the numbers are small compared to SI units of mass.

2.5 Deriving the amplitude as function of time

2.5.1 Relating the amplitude of a gravitational wave to the physical distance

To derive how the amplitude as function of time behaves, Ref. [2] derived a relation between the gravitational wave amplitude $h_{\mathcal{D}}$ as function of the time t. Mind that there is a minus sign in front, but this does not matter since the absolute value is taken to study the behaviour of the amplitude in time. The

relation between $h_{\mathcal{D}}$ and the time is found by using Equation 1.11 and Equation 2.11, together with the rule to combine a cosine and a sine described in Appendix B.3:

$$h_{\mathcal{D}}(t) = -\left(\frac{G\mathcal{M}}{c^2 D_{\text{eff}}}\right) \cdot \left(\frac{t_0 - t}{5G\mathcal{M}/c^3}\right)^{-1/4} \cdot \cos(2\omega(t - r/c) + 2\phi_0), \tag{2.12}$$

where t_0 is the coalescence time of arrival of the signal at the detector, ϕ_0 is a phase (not of interest here, but can be tuned using the β rotation described in Section 1.5) and D_{eff} the effective physical distance defined as:

$$D_{\text{eff}} \equiv D\left[F_+^2 \cdot \left(\frac{1 + \cos^2(i)}{2}\right)^2 + F_\times^2 \cdot \cos^2(i)\right]^{-1/2}. \tag{2.13}$$

Here F_+ and $F_\times$ are the antenna response functions for the +- and ×-polarizations respectively (Ref. [5]). These are functions that are characteristics of the detector used. See Section 1.5 for the definitions of the antenna response functions F_+ and $F_\times$ and their θ and ϕ dependence.
In Equation 2.13 there is also an angle of inclination i, which is the angle between the normal vector to the orbital plane and the line connecting the origin of the source and the origin of the detector. This has nothing to do with the position of the source in the sky, but it is the inclination angle with respect to the line connecting detector and source.

Due to the whitening process, the gravitational wave strain does not have the correct order of magnitude. An attempt is done to estimate the effective distance the binary system should have according to these data, which can be seen in Appendix C. The estimation of the correct amplification factor (and thus the correct order of magnitude of the gravitational wave signal) is a bit difficult, and goes beyond the scope of this thesis. This gives that the value obtained for the effective distance is not a realistic value for this binary system. Nevertheless, it was useful to attempt the estimation of the effective distance in order to learn how to process gravitational wave data from the signal even more. It was also educational to study a way of obtaining this amplification factor. But, this means for this thesis that it is only relevant to check the $-1/4$ behaviour of the gravitational wave strain as function of the time in Equation 2.12. This will be elaborated on in chapter 6.

Chapter 3

Derivation of alternative models

3.1 Binary system with time dependent mass

This alternative model is a model that probes beyond the Theory of General Relativity: the mass is changing, which is not likely to happen within General Relativity. This model is used to test whether such time dependent masses could be fitted through the data of the event GW150914. The fit function for the inspiraling binary with time dependent mass consists of a very important concept: The mass is assumed to change linearly in time. This can be expressed as follows:

$$\mathcal{M} = \mathcal{M}_0 + \dot{\mathcal{M}}_0 \cdot (t - t_0) \tag{3.1}$$

Of course, more higher order terms can be taken into account, such as second or third order derivatives. These higher order terms are less constrained, such that expanding up to the first order is appropriate for this thesis.

Due to the fact that the mass is assumed to change only linearly, this means that a first order Taylor expansion of the Chirp mass to the power 5/3 is required. A first order is sufficient due to the infinitesimal condition. Taylor expanding the chirp mass to the power 5/3 around t_0 gives:

$$\mathcal{M}^{5/3} = (\mathcal{M}_0)^{5/3} + \frac{5}{3} \cdot \dot{\mathcal{M}}_0 \cdot \mathcal{M}_0^{2/3} \cdot (t - t_0) \tag{3.2}$$

With these tools, the relation in Equation 2.11 now becomes:

$$\begin{aligned} f_{\mathrm{GW}}^{-8/3} =& -\frac{(8\pi)^{8/3}}{5} \cdot \left(\frac{G}{c^3}\right)^{5/3} \cdot \left[\mathcal{M}_0^{5/3} + \frac{5}{3} \cdot \dot{\mathcal{M}}_0 \cdot \mathcal{M}_0^{2/3} \cdot (t - t_0)\right] \cdot (t - t_0) \\ =& \frac{(8\pi)^{8/3}}{5} \cdot \left(\frac{G}{c^3}\right)^{5/3} \cdot \left[-\frac{5}{3} \cdot \dot{\mathcal{M}}_0 \cdot \mathcal{M}_0^{2/3} \cdot (t - t_0)^2 - \mathcal{M}_0^{5/3} \cdot (t - t_0)\right]. \end{aligned} \tag{3.3}$$

To avoid calculations and results with large numbers, the masses are expressed in solar masses $M_\odot$ instead of 'kilograms'.

3.2 Rotating rigid spirally shaped mass distribution

The previous crazy model is an extension of the original model, which gives that the fit function for this model can be derived really easily. The model described in this section is a total newly introduced model, which has nothing to do with the original model. The derivation of the equation that describes the frequency as function of time is therefore a bit longer than the derivation for the previous crazy model.

3.2.1 Form of the mass distribution

Assume a mass distribution that has been formed in a spiral shape, which has a total mass M being constant in time. See Figure 3.1. The spiral has formula:

$$r(\phi) = \alpha\phi \text{ with } \phi \in [0, 2\pi]. \tag{3.4}$$

The parameter α contains information about the size of the spiral. Because we know that the distribution is shaped like a spiral, the size of the spiral should be larger than its Schwarzschild radius. It is a bit strange to mention a 'Schwarzschild radius' for an object that is not a sphere, but this radius is the boundary that becomes an event horizon if all of its total mass is contained within a sphere of that radius. Here it is assumed that the size of the spiral distribution is larger than or equal to the Schwarzschild radius that appears if the total mass was collected in a spherical distribution with that radius, i.e.

$$r_0 = r(\phi = 2\pi) - r(\phi = 0) = \alpha \cdot 2\pi \geq \frac{2MG}{c^2} \Rightarrow \alpha \geq \frac{MG}{\pi c^2}. \tag{3.5}$$

For this thesis, it is sufficient to work out the spiral model for only one value of α. For convenience, the minimal value is chosen as value for α: $\alpha = MG/(\pi c^2)$.

Assumed is also that the mass density is constant in space and time. It is a one dimensional distribution, denoted with λ, such that $\mathrm{d}m = \lambda \mathrm{d}l$ holds, where $\mathrm{d}m$ represents an infinitesimal mass part. The total mass expressed in terms of the mass density can be calculated by integrating over all pieces $\mathrm{d}m$ as:

$$M = \int \mathrm{d}m = \int_0^l \lambda \mathrm{d}l = \lambda\alpha \underbrace{\int_0^{2\pi} \sqrt{1+\phi^2}\mathrm{d}\phi}_{\equiv\Phi} = \lambda \cdot l = \Phi\lambda\alpha, \tag{3.6}$$

with l the length of the spiral, calculated via the integral above in the system co-rotating with the spiral as:

$$l = \int_0^{2\pi} \alpha \cdot \sqrt{1+\phi^2}\mathrm{d}\phi. \tag{3.7}$$

Because the shape of the spiral is assumed to be constant and because the limits of the spiral (the limits in the integral) are assumed to stay constant in time, the following integral is a constant integral denoted with Φ. It is very useful to express results in this constant:

$$\Phi = \int_0^{2\pi} \sqrt{1+\phi^2}\mathrm{d}\phi \approx 21.2563. \tag{3.8}$$

This means that the following relation between l and Φ is obtained, because α is a constant: $l = \alpha\Phi$.

What is important here is that ω is only a slowly varying function in time. This means that it can be assumed to be constant on timescales T in the order of the period of one rotation, but is changing on timescales $\tau \gg T$. These are timescales of many rotations. This is an assumption that is also used in the derivation of the original model.

Because the spiral is emitting gravitational waves and the body is rigid, the bodies energy should decrease when emitting waves, which means it slows down such that the emitted frequency drops as time moves forward. But this contradicts the behaviour of the data; from the data we can conclude that the frequency rises as time increases. This means that the mass needs to be negative in order to create a system that behaves like the fashion of the data. Although this model is astrophysically very unlikely, it could be a step towards describing negative mass distributions. But it could be possible that within the uncertainty, a positive mass is possible. This contradiction on forehand is useful to be taken into account, but the results will determine to what extend this model will be ruled out or not.

3.2.2 Inertia and quadrupole moment tensor in restframe

The inertia tensor is calculated in the frame co-rotating with the mass distribution to be:

$$I = \lambda\alpha^3 \begin{pmatrix} J & R & 0 \\ R & P-J & 0 \\ 0 & 0 & P \end{pmatrix}, \tag{3.9}$$

with

$$J = \int_0^{2\pi} \phi^2 \sin^2(\phi)\sqrt{1+\phi^2}\mathrm{d}\phi \approx 184.80 \tag{3.10a}$$

$$R = \int_0^{2\pi} \phi^2 \sin(2\phi)\sqrt{1+\phi^2}\mathrm{d}\phi \approx 120.93 \tag{3.10b}$$

$$P = \int_0^{2\pi} \phi^2 \sqrt{1+\phi^2}\mathrm{d}\phi \approx 399.22 \tag{3.10c}$$

These integrals appear when the inertia tensor is calculated via its definition in Appendix B.2. All of these three functions are constants in space and time in the rest frame of the spiral (co-rotating frame). The only things that can change is the angular velocity ω.

A coordinate transformation is needed to calculate it for this spiral shaped mass distribution, but this is very straight forward: a transformation is done to a coordinate system that has the z-axis aligned with the source frame, and the x-axis tangent to each point of the spiral and the y-axis perpendicular to the tangent line at each point and perpendicular to the z-axis. See Figure 3.1 for clarification.

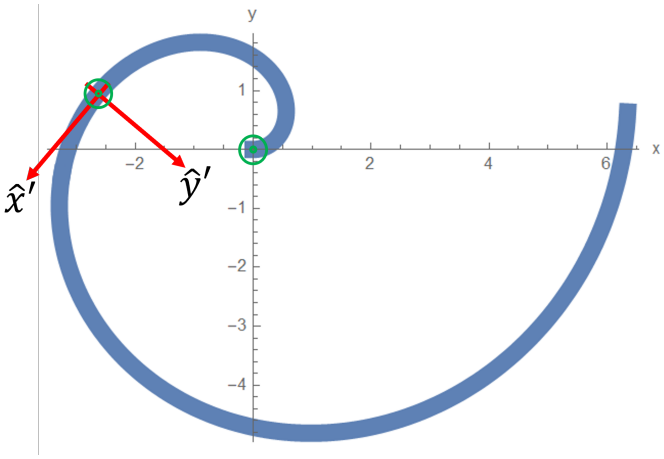


Figure 3.1: Geometry of the rigid spiral with a coordinate system introduced at each point of the spiral to simplify the calculation of the inertia tensor and the total length of the spiral. The green circle represents the z-axis pointing out of the paper in both the stationary (located at the origin of the given coordinate system) as well as in the coordinate systems along the spiral. The spiral is assumed to be rotating around the green dot located at the origin.

3.2.3 Inertia and quadrupole moment tensor in the (constant) source frame

The inertia tensor takes the following form in the source frame:

$$I_S = \frac{\lambda\alpha^3}{2} \cdot \left[\begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & 2P \end{pmatrix} + \begin{pmatrix} 2J-P & 2R & 0 \\ 2R & -(2J-P) & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \cos(2\omega t) + \begin{pmatrix} -2R & (2J-P) & 0 \\ (2J-P) & 2R & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \sin(2\omega t) \right]. \tag{3.11}$$

The quadrupole moment tensor is then calculated in the source frame with the help of the relation given in Appendix B.2 as:

$$Q = \frac{\lambda\alpha^3}{2} \cdot \left[\begin{pmatrix} \frac{P}{3} & 0 & 0 \\ 0 & \frac{P}{3} & 0 \\ 0 & 0 & -\frac{2P}{3} \end{pmatrix} - \begin{pmatrix} 2J-P & 2R & 0 \\ 2R & -(2J-P) & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \cos(2\omega t) - \begin{pmatrix} -2R & (2J-P) & 0 \\ (2J-P) & 2R & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \sin(2\omega t) \right]. \tag{3.12}$$

The time derivative of this tensor is now calculated. This can be done in this step of the calculation because the frame of the wave is not moving with respect to the stationary system in which the spiral is

rotating, as is explained before. After this time derivative, the tensor is taken to the wave frame. The form the tensor takes in Traceless gauge in the wave frame is:

$$\begin{aligned}\ddot{Q} =& -4\lambda\alpha^3\omega^2 \cdot \begin{pmatrix} \left(\frac{1+\cos^2(i)}{2}\right)(2J-P) & 2R\cos(i) & 0 \\ 2R\cos(i) & -\left(\frac{1+\cos^2(i)}{2}\right)(2J-P) & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \cos(2\omega(t-r/c)) \\ &+ 4\lambda\alpha^3\omega^2 \cdot \begin{pmatrix} 2R\left(\frac{1+\cos^2(i)}{2}\right) & -(2J-P)\cos(i) & 0 \\ -(2J-P)\cos(i) & -2R\left(\frac{1+\cos^2(i)}{2}\right) & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \sin(2\omega(t-r/c)),\end{aligned} \tag{3.13}$$

such that the metric perturbation h_{ij} can be calculated via the quadrupole moment formula in Equation 1.7. It turns out that the quadrupole moment tensor is already transverse (and traceless by definition)! This means that the quadrupole moment tensor is already in transverse traceless gauge, which means also that automatically h_{ij} is in transverse traceless gauge. The polarization amplitudes h_+ and $h_\times$ are then calculated via Equation 1.5. This gives for these amplitudes:

$$h_+ = -\frac{4G\lambda\alpha^3\omega^2}{c^4 r} \cdot \left[\frac{1+\cos^2(i)}{2}\right] \cdot [(2J-P)\cos(2\omega(t-r/c)) + 2R\sin(2\omega(t-r/c))] \tag{3.14a}$$

$$h_\times = -\frac{4G\lambda\alpha^3\omega^2}{c^4 r} \cdot \cos(i) \cdot [2R\cos(2\omega(t-r/c)) - (2J-P)\sin(2\omega(t-r/c))] \tag{3.14b}$$

3.2.4 Energy loss due to gravitational wave emission

With the help of Equation 1.17, the following relation for the energy carried away by a gravitational wave is obtained:

$$\frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}t} = \frac{r^2c^3}{32\pi G}\int \left\langle \dot{h}_{ij}^{TT}\dot{h}_{TT}^{ij}\right\rangle_T \mathrm{d}\Omega = \frac{r^2c^3}{16\pi G}\int_0^\pi\int_0^{2\pi} \left\langle \dot{h}_+^2 + \dot{h}_\times^2\right\rangle_T \sin(\theta)\mathrm{d}\phi\mathrm{d}\theta, \tag{3.15}$$

where $\langle ...\rangle_T$ denote time averaging over one rotation period T. For this purpose, it means that terms containing the product of a sine and cosine average to 0, whereas terms containing the square of a cosine or sine average to 1/2. This means that the integrand takes the form:

$$\begin{aligned}\left\langle \dot{h}_+^2 + \dot{h}_\times^2\right\rangle_T &= \left(\frac{8G\lambda\alpha^3\omega^3}{c^4 r}\right)^2 \left[\frac{(1+\cos^2(\theta))^2}{4}\left[(2J-P)^2 + (2R)^2\right] + \cos^2(\theta)\left[(2R)^2 + (2J-P)^2\right]\right] \\ &= \left(\frac{4G\lambda\alpha^3\omega^3}{c^4 r}\right)^2 \left[(2J-P)^2 + (2R)^2\right]\left(\frac{1}{4} + \frac{3}{2}\cos^2(\theta) + \frac{1}{4}\cos^4(\theta)\right),\end{aligned} \tag{3.16}$$

such that the total energy loss per unit of time equals:

$$\frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}t} = \frac{r^2c^3}{16\pi G} \cdot \left(\frac{4G\lambda\alpha^3\omega^3}{c^4 r}\right)^2 \left[(2J-P)^2 + (2R)^2\right] \cdot \frac{16\pi}{5} = \frac{16M^2G\alpha^4\omega^6}{5\Phi^2c^5} \cdot S, \tag{3.17}$$

where $S \equiv \left[(2J-P)^2 + (2R)^2\right] = 5.9372 \cdot 10^4$ is a dimensionless constant.

3.2.5 Kinetic energy of the rotating spiral

Because the source frame as well as the wave frame are non moving, the kinetic energy of this spiral does not depend on the frame in which it is calculated (wave or source frame). The kinetic energy of the rotating spiral is easily calculated in the frame of the source:

$$E_k = \frac{1}{2}I_{33}\omega^2 = \frac{\lambda\alpha^3}{2}P\omega^2 = \frac{M\alpha^2}{2\Phi}P\omega^2 \tag{3.18}$$

The change of kinetic energy per unit time is calculated to be:

$$\frac{\mathrm{d}E_k}{\mathrm{d}t} = \frac{M\alpha^2}{\Phi} \cdot P \cdot \omega \cdot \dot{\omega} \tag{3.19}$$

3.2.6 Potential energy of the rotating spiral

The potential energy of the spiral is calculated using the assumption that the mass only feels the potential energy along the line of the spiral, such that the potential energy for a mass at angle ϕ_1 and distance the origin $r_1 = \alpha\phi_1$ does not feel for example the gravitational force from a mass at angle $\phi_1 + \pi$ and distance to the origin $r_2 = \alpha\phi_1 + 2\pi\alpha$ (opposite masses). This assumption is justified by the fact that the distance $(r_2 - r_1)$ is large, because α is large enough (α equals the Schwarzschild radius). This means that the spiral can be "unrolled", which means that it can be seen as a straight line. The calculation procedure is as follows:

1. The line will be build up like a string of beads, where each bead is a bit of mass. This bit of mass is $\mathrm{d}m = \lambda \mathrm{d}l$.
2. This bit of mass needs to be stringed along the line. This means that it needs to be guided along the potential energy of the mass already present. This mass is given by: $M(s) = \int_0^s \lambda \mathrm{d}s' = \lambda s$, with s the length of the string with the mass until then.
3. The energy *stored* in this system is therefore *minus* the energy that need to be overcome. This is the energy a person uses to push this bead along the wire; $\mathrm{d}W_{\text{person}} = -\mathrm{d}W_{\text{grav}}$.
4. The energy needed to push the bead $\mathrm{d}m$ from ∞ to a distance s from the origin, with the potential energy of the mass already present is given by: $\mathrm{d}W_{\text{grav}} = \frac{GM(s)\mathrm{d}m}{s}$.

The total potential energy stored in the system is calculated to be:

$$E_p = W_{\text{person}} = \int \mathrm{d}W_{\text{system}} = -G\int_0^l \frac{\lambda s \cdot \lambda \mathrm{d}l}{s} = -G\lambda^2 \int_0^l \mathrm{d}l = -G\lambda^2 l, \tag{3.20}$$

with again $l = \alpha\Phi$. It is better to express the potential energy in terms of the total mass M and the constant integral Φ:

$$E_p = \frac{GM^2}{\Phi} \cdot \frac{1}{\alpha} = \frac{GM^2}{l} \tag{3.21}$$

The change of the potential per unit time is calculated to be 0, because it does not depend on ω, which is the only parameter that changes in time.

3.2.7 Energy balance; obtaining the frequency as function of time

The total change in mechanical energy per unit of time equals minus the energy per unit time that the gravitational waves take away. There is a minus sign here, because gravitational waves *take away* energy, whereas it is compared to the *gain* in mechanical energy. The change in mechanical energy per unit time is calculated to be:

$$\frac{\mathrm{d}E_{\text{mech}}}{\mathrm{d}t} = \frac{\mathrm{d}E_k}{\mathrm{d}t} + \frac{\mathrm{d}E_p}{\mathrm{d}t} = \frac{M\alpha^2}{\Phi} \cdot P \cdot \omega \cdot \dot{\omega} \tag{3.22}$$

Equating this to minus the energy per unit time of the gravitational waves gives:

$$\frac{16M^2G\alpha^4\omega^6}{5\Phi^2c^5} \cdot S = -\frac{MP\alpha^2}{\Phi} \cdot \omega \cdot \dot{\omega} \tag{3.23}$$

Simplifying the result gives:

$$\frac{\dot{\omega}}{\omega^5} = -\frac{16MG\alpha^2 S}{5P\Phi c^5} = -\frac{16}{5\pi^2 c^3} \cdot \frac{S}{\Phi \cdot P} \cdot \left(\frac{MG}{c^2}\right)^3 = -C_0 \cdot \left(\frac{MG}{c^2}\right)^3, \tag{3.24}$$

where $C_0 \equiv 16S/\left(5\pi^2 \cdot c^3 \cdot \Phi \cdot P\right)$. In the one but last step the minimal value of α that it chosen to work with is substituted. It is chosen to work with the mass in solar masses, because this gives smaller numbers compared to 'kilograms'. With the mass given in kilograms here, Equation 3.24 takes the form:

$$\frac{\dot{\omega}}{\omega^5} = -C_0 \cdot \left(\frac{MG}{c^2}\right)^3. \tag{3.25}$$

This Equation has multiple solutions, but the only physical solution (real and positive) equals:

$$\omega(t) = \frac{1}{\sqrt{2}} \cdot \left[-C_1 + C_0 \cdot M'^3 \cdot (t - t_0)\right]^{-1/4} \tag{3.26}$$

with C_1 an integration constant, to be determined by the boundary condition $\omega_0 = \omega(t = t_0)$. It turns out that this $C_1 = -1/(4\omega_0^4)$. From the gravitational wave signal, it can be seen from the data that the period shrinks to 0 if the signal gets closer to t_0, such that the frequency becomes infinite as the time t approaches t_0. This means that $\omega_0 \to \infty$ such that $C_1 = 0$. The gravitational wave angular frequency is exactly twice the rotational angular frequency, as can be concluded from the derived Equation 3.14. One obtains the following linear relation in time:

$$f_{GW}^{-4} = (2 \cdot 2\pi\omega(t))^{-4} = \frac{1}{64\pi^4} \cdot C_0 \cdot \left(\frac{MG}{c^2}\right)^3 \cdot (t - t_0) \tag{3.27}$$

The integration constant is chosen to be t_0, in the same fashion as is done for the original model. For a mass distribution that is slowing down, this integration constant cannot be named the coalescence time, since nothing coalesces. This offset time can be interpreted as the time at which the spiral stops rotating.

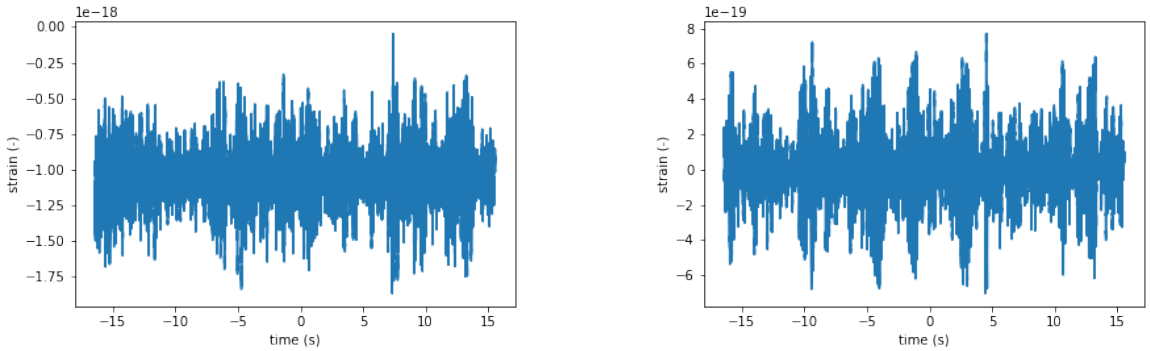
Chapter 4

Processing Gravitational Wave Data

4.1 The used gravitational wave data sets

4.1.1 Data set from Livingston and Hanford

From the detector in Livingston and from that in Hanford, the obtained data sets are very similar. Both sets consist of 4096 samples per second (i.e. the sampling frequency $f_s = 4096Hz$), and the length of the measured window is $32s$ for both detectors. The signals obtained at Livingston and Handford are plotted in Figure 4.1 a and b respectively:



(a) Livingston

(b) Hanford

Figure 4.1: Original signal obtained at the LIGO detectors

Both signals are built up in the sense that the inspiraling phase lies around $t \approx 0$, such that the entry that lies closest to $t = 0$ is found by Python. Around that entry the signal is plotted to find what range of entries is appropriate. Because both detectors are located at a different position on Earth, there could be a time difference between both signals. But this time difference needs to be small, because the distance is relatively small and the speed of light is very large. See Figure 4.2.

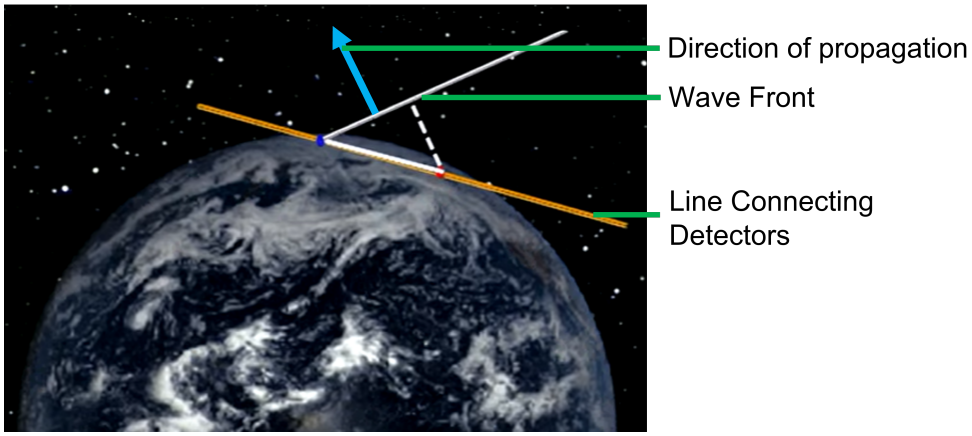
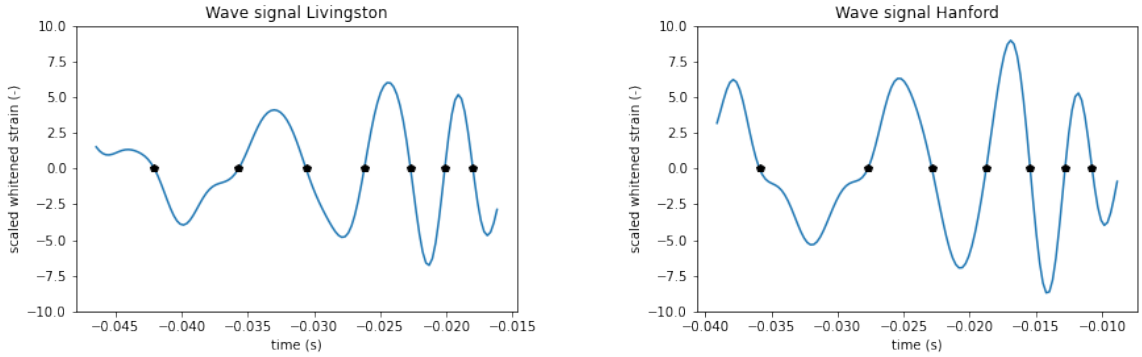


Figure 4.2: The detector in Livingston (Red) and Hanford (Blue) are located at different positions on Earth. The wave may be propagating with an angle with respect to the line connecting both detectors (See Ref. [3]).

The maximum time difference (i.e. when the wave propagates in the direction of the line connecting the detectors) is around $10ms$. From this it can be concluded that the excitement still happens around $t \approx 0$ with such a small time difference. The range in which the signal is present is determined by plotting several ranges around the point where $t = 0$, as is explained. Eventually, the obtained plot for both detectors takes the form as in Figure 4.3 below:



(a) Livingston

(b) Hanford

Figure 4.3: Whitened signal obtained for the LIGO detectors. The black dots are the zero crossings used to estimate the period of the signal at certain time stamps.

4.2 Whitening the signal

4.2.1 Determining the Amplitude Spectral Density

With the help of the option 'ASD' in Python (which stands for 'amplitude spectral density'), an amplitude spectral density (ASD) can be made. An amplitude spectral density is in fact the root of the power spectral density (PSD). A PSD describes the distribution of power into frequency components present in the signal. One can determine 'how much' a certain frequency is present in a signal. The sample frequency f_s is of course equal to the reciprocal of the time difference between 2 successive data points ($f_s = 4096Hz$. This is due to the fact that the signal is already sampled, such that the time stamps where the data points are located are already chosen when this signal was sampled. Together with Python, an amplitude spectral density (ASD) is made from the PSD. This is done for the data measured by the detector in Hanford, as well as for the data obtained by the detector in Livingston. See Figure 4.4

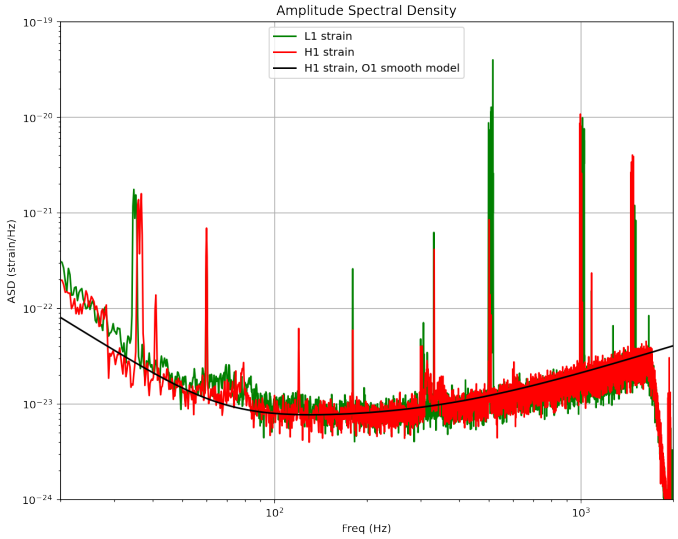


Figure 4.4: Amplitude Spectral Density for the signals obtained from the detectors at Livingston (Green) and Hanford (Red). With the black line a smooth model is plotted for the data obtained from the detector in Hanford.

4.2.2 The whitening process

Whitening a signal is a relatively easy operation. What is basically done when whitening a signal is: first transform the signal to the frequency domain by a fast Fourier transform, multiply the data set given in random variables with a so called 'whitening matrix' to obtain a data set in new variables that has a diagonal covariance matrix (i.e. the variables are uncorrelated), and then transform it back to the time domain. What is obtained then, is a signal which has an amplified amplitude. For the frequency this is no problem, because the frequency is estimated by looking at successive zero crossings of the signal, which are unaltered. This will be elaborated on in Section 4.3. For calculations where the real amplitude is needed for, this amplification factor needs to be taken into account. This will be discussed broadly in Section 4.4.2.

4.3 Frequency estimation

4.3.1 Combining both data sets from Livingston and Hanford in one frequency plot

To combine the data from Hanford and Livingston, the time axis of both signals need to be aligned. In Ref. [1], the signals are aligned as much as possible. In this article, the researchers inverted one of the signals and shifted this one by 6.9ms. In this case, the overlap of both signals is maximal. This needs to be done, because now both frequency estimates align on the time scale. If the signals were not aligned, one would obtain 2 different data sets that cannot be combined; the time shift would then be present in the frequency estimation plot. The overlap of the signals is only used to determine the time shift, and not for determining the frequency. This time shift is then used to shift the signal from the detector in Livingston by $0.0069s$. Both signals are then used separately to estimate the frequency at certain time stamps. In the end, the frequency plots of both signals are combined into one frequency plot. See Figure 4.5 how the maximal overlap looks like.

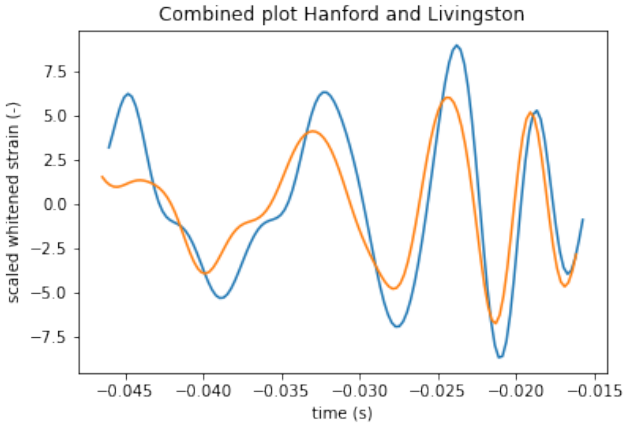


Figure 4.5: The signals from the detector in Livingston and Hanford plotted in one figure, where the signal obtained from the detector in Livingston is time shifted in such a way that it creates maximum alignment (i.e. overlap) with the signal obtained with the detector in Hanford.

The frequencies are determined by successive zero crossings of the yellow curve and the blue curve separately in Figure 4.5. As can be seen from the figure above, both gravitational wave signals agree with each other about the frequency increase in time; the zero-crossings of both signals are almost the same and have the same successive tendency (i.e. have roughly the same increase rate). The combined frequency-time plot takes the form as in Figure 4.6:

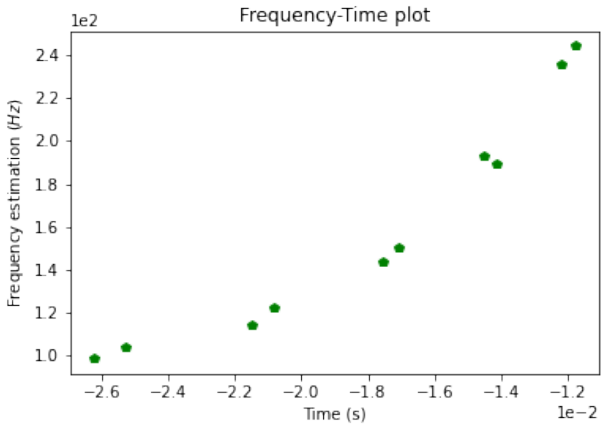


Figure 4.6: Frequency estimation for the signal during the inspiraling phase, where both the signal measured in Livingston as well as the signal measured in Hanford is taken into account.

Points with a too large uncertainty and points that lie within the ring-down phase are left out. The points that are assumed to satisfy both conditions lie in between $t = -0.027$ and $t = -0.011$ (i.e. the range in Figure 4.6 above). The uncertainty becomes to large for points at $t < -0.027$, because the amplitude has roughly the same value as the noise. Therefore, at these time stamps the signal is hardly distinguishable from the noise. As time evolves (and thus the amplitude increases), the signal becomes more and more distinguishable from the noise, until it eventually disappears. The signal dies out as the system comes in the so called 'ringdown-phase', where the mass distribution evolves to a sphere such that the quadrupole moment becomes 0. The models that are derived only hold in the inspiral phase, such that only points estimating the frequency that lie *in the time range of the inspiraling phase* are taken into account. This means that points for $t > -0.011$ are left out. For both signals, only five points are selected to which data is fitted. In total, there are ten points to estimate the frequency as a function of time with. The points that are taken into account are plotted in Figure 4.6.

4.3.2 Obtaining the frequency as function of time in GW data set

In a gravitational wave signal, the frequency is in general time dependent. It is determined by looking at the time difference between successive zero crossings of the signal. Because the signal is built up as a sinusoid curve (with decreasing period), the period of the signal can be estimated by successive zero crossings of the signal. The time of zero crossing i is denoted with Z_i This time difference is assumed to represent half of the period of the signal at time τ_i, where τ_i is chosen to be exactly in the middle of the successive zero crossings i and $i+1$. Here i is an index for the zero crossings. The frequency f_i at time τ_i are therefore estimated as:

$$\tau_i = \frac{Z_i + Z_{i+1}}{2} \tag{4.1a}$$

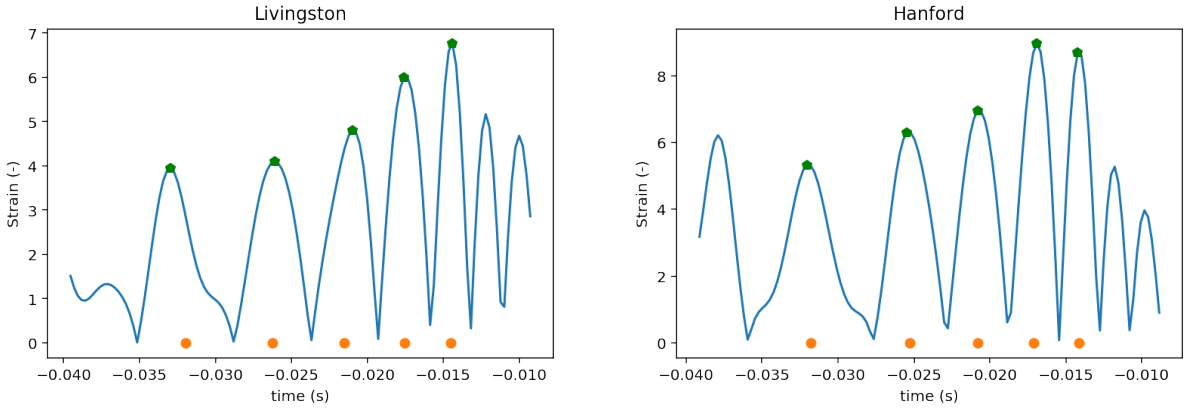
$$f_i = \frac{1}{2 \cdot (Z_{i+1} - Z_i)}, \tag{4.1b}$$

During the rest of the thesis, i will run from 1 up to the number of zero crossings taken into account for a signal.

4.4 Amplitude estimation

4.4.1 Finding the maxima of a signal

One has now the whitened (and amplified) signal. If one wants to look at the evolution of the amplitude of a signal, the values of the extreme values are very useful. From this, one can determine how the amplitude behaves as a function of time, because for a signal, the amplitude is defined as the maximal deviation from the equilibrium position. The signal has positive and negative deviation (because it is oscillating). In order to take both the maxima and the minima into account, the absolute value of the signal is taken, and with help of Python, all the maxima are found (the minima are turned into maxima when taking the absolute value). The result is shown in Figure 4.7 below for both the detector located at Hanford as well as for the detector at Livingston.



(a) label 1

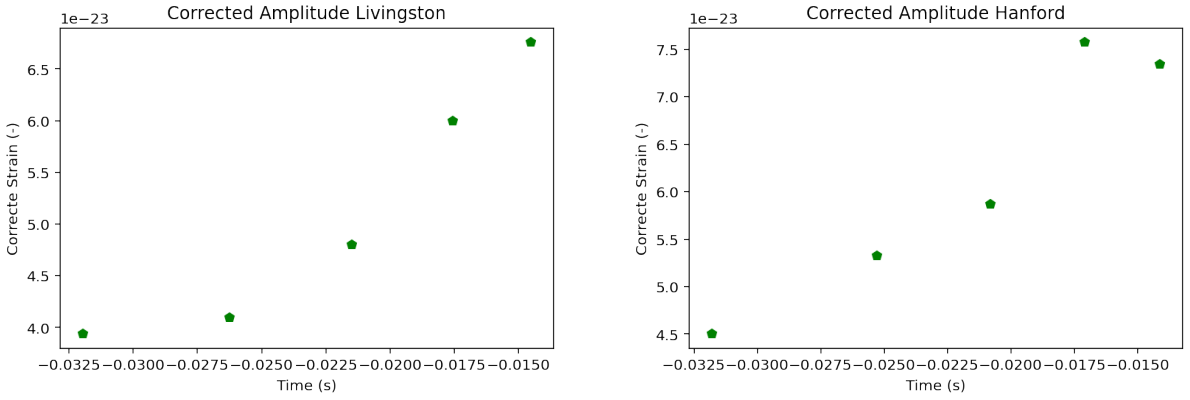
(b) label 2

Figure 4.7: 2 Figures side by side

4.4.2 Determining the correct value of the amplitude as function of time

In order to determine the real values of the amplitude, another normalization factor needs to be used. But not the whole signal is multiplied with this normalization factor, because we have now found the times at which the amplitude has a maximum value and the value of that whitened amplitude. We also have determined the frequency estimation at approximately these time stamps. This is due to the fact that an extreme value of an oscillating signal is halfway in between two successive zero crossings, and the frequency estimations are also assumed to be at times half way in between two successive zero crossings.

Of course, it is possible that the time stamps of the maxima are not exactly equal tot he time stamps of the frequency estimations. But in order to find the amplification factor, the value of the frequency is needed (because the amplification factor is taken from the ASD). The difference between these time stamps is therefore neglected (because this difference is very small). At the time stamp of each maximum, the frequency is estimated. At these values of the frequencies, the values of the ASD are obtained. These values are averaged, from which the result is the estimation of the amplification factor. The correct value of the amplitude is plotted against the time for both the detectors located at Livingston and Hanford in Figure 4.8 below. Here only the value of the peaks are shown.



(a) label 1

(b) label 2

Figure 4.8: 2 Figures side by side

Because this method is a bit inaccurate due to the fact that the 'undo-procedure' is not exactly the reverse operation of the whitening process, the (effective) distance estimated with this method is also inaccurate, and may be incorrect. Due to this fact, the power law between the gravitational wave strain and the time in Equation 2.12 is estimated for both the signal obtained in Hanford as well as for the signal obtained in Livingston. The amplification factor of the gravitational wave signal may be incorrect, but this only causes a different prefactor in Equation 2.12. The same 1/4 power law remains, such that this can be determined very accurately.

4.5 Extracting uncertainties

4.5.1 Uncertainties in the data points

Of course are the obtained values for the frequencies not precisely correct, but they have uncertainties. A zero crossing could be a bit more to the left or to the right. In order to take this into account, it is assumed that a zero crossing could be at most lie at the location of the points where the signal takes half the value of the nearest absolute maximum (to the left or to the right). The time t_i where the signal has a local maximum is denoted with M_i. These maxima both contain maxima and minima in the signal, but they are in the figure all maxima due to the absolute value. The point in time where the signal takes half the value of its maximum is denoted with H_i. There are twice as much half-maxima than maxima, as is clear. This means that zero crossing Z_i can be at most located to the right at the time where the signal takes half of the value of the nearest absolute maximum to the right (at H_{2i}), and can be at most located to the left at the time where the signal takes half of the value of the nearest absolute maximum to the left (at H_{2i-1}). This means that zero crossing Z_3 is located in between half-maximum H_5 and H_6. See Figure 4.9 for clarification.

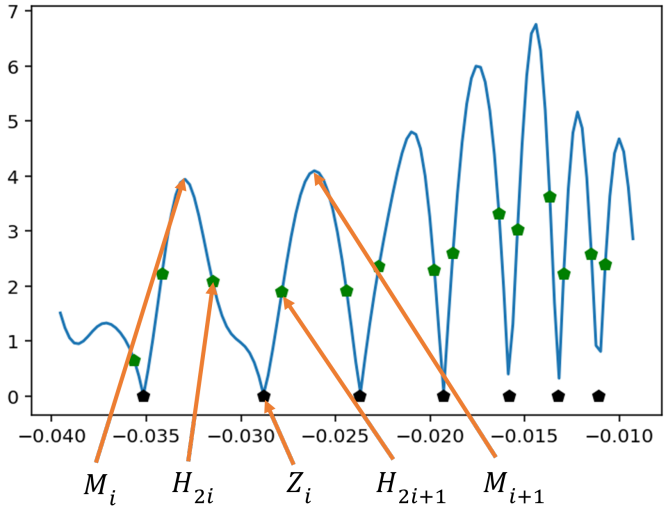


Figure 4.9: Uncertainty estimation of the zero crossings. Here the different points used in the derivation are indicated below with an orange arrow. A black dot (zero crossing) could have a value in between the previous and following green dot (half maxima).

The signal is well organized, in the sense that zero crossing Z_i is wedged in between absolute extreme value M_{i-1} and absolute extreme value M_i, as can also be seen in Figure 4.9. For the data coming from the detector at Hanford the exact same thing is done. This signal is also well organized in the same sense. As is already said, points that satisfy $-0.030 < t < -0.012$ are taken into account.

4.5.2 Uncertainties in the period of the signal

The maximal and minimal values of a zero crossing point Z_i is now clear, but the maximal and minimal value of the period of the signal need to be calculated in order to determine the uncertainty in the frequency. The maximal value of the period is calculated using the maximal time difference between 2 successive zero crossings, i.e. the minimum (i.e. left most) value of zero crossing i (at H_{2i-1}) and the maximum (i.e. right most) value of zero crossing $i+1$ (at $H_{2(i+1)} = H_{2i+2}$):

$$T_i^{\max} = Z_{i+1}^{\max} - Z_i^{\min} = H_{2i+2} - H_{2i-1}, \tag{4.2}$$

where $Z_i^{\max}$ is the value most to the right that zero crossing i could have. The minimal value (value most to the left) of period i is determined in the same way by calculating the minimal time difference between successive zero crossings:

$$T_i^{\min} = Z_{i+1}^{\min} - Z_i^{\max} = H_{2i+1} - H_{2i}. \tag{4.3}$$

We have now that:

$$T_i^{\min} \leq T_i^T \leq T_i^{\max}. \tag{4.4}$$

The true value T_i^T of the period is in between these boundaries, with 68% confidence, such that the uncertainty is estimated as follows: The uncertainty from above is estimated by the difference between the maximum value $\leq T_i^{\max}$ and the calculated value T_i, whereas the uncertainty from below is estimated by the difference between the calculated value T_i and the minimum value $T_i^{\min}$. In the end, the uncertainty is estimated as the mean of the uncertainty from above and below:

$$\sigma_i^+ \equiv T_i^{\max} - T_i \tag{4.5a}$$

$$\sigma_i^- \equiv T_i - T_i^{\min} \tag{4.5b}$$

$$\sigma_i^T \equiv \frac{\sigma_i^+ + \sigma_i^-}{2} = \frac{T_i^{\max} - T_i^{\min}}{2} \tag{4.5c}$$

4.5.3 Uncertainty in the frequency

The uncertainty in a period is calculated above. The uncertainty in frequency f_i at time τ_i is calculate via Equation A.6 in Appendix A.2 to be:

$$\sigma_i^f = f_i^2 \cdot \sigma_i^T . \tag{4.6}$$

The frequency as function of time, including uncertainties at all the data points is visualized in Figure 4.10 below:

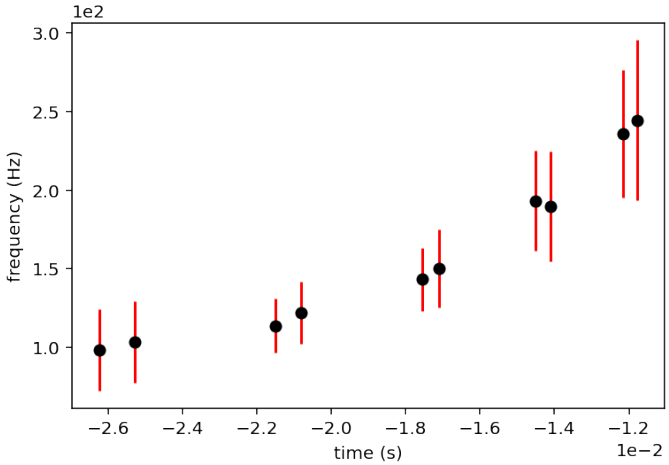
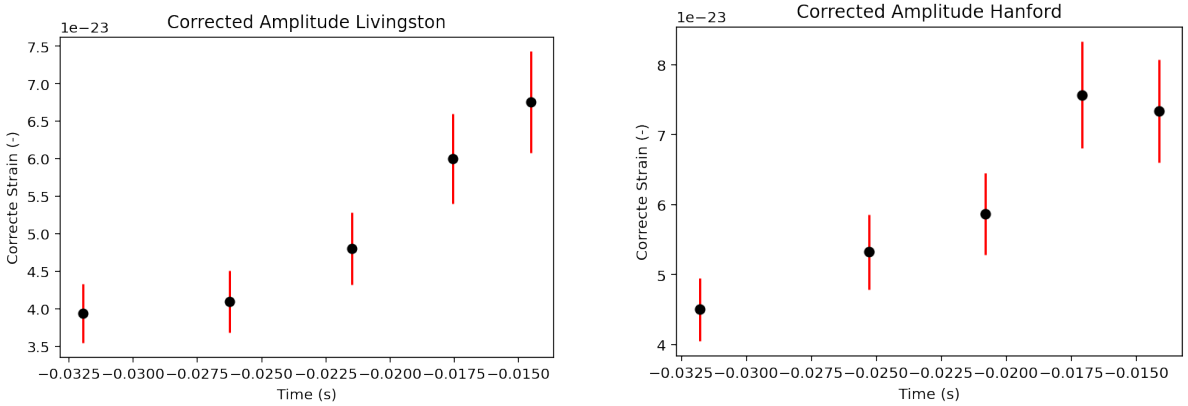


Figure 4.10: Frequency estimation plotted against time. The red bars denote the error bars at each data point. These are determined using the steps described in the previous section.

4.5.4 Uncertainties in the amplitude of the signal

In order to estimate the uncertainty in the value of the amplitude of the signal, it is assumed that the real value is deviated at most 10% from the estimated value, i.e. 10% above and 10% below the estimated value. The error bar is therefore dependent on the value of the amplitude. See Figure 4.11 below:



(a) Livingston

(b) Hanford

Figure 4.11: Corrected strain $h(t)$ plotted against the time t. The red bars denote the error bar at each data point, which is assumed to reach to 0.9 times its value below, and 1.1 times its value above.

Chapter 5

Bayesian Statistics: Probability Density Function for fit parameters

5.1 Finding the best fit: least squares

5.1.1 Bayes' rule

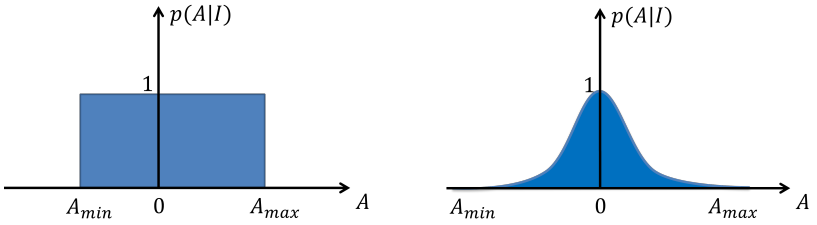
In this thesis, linear and quadratic fit functions are used to fit through data points. To determine the most optimal parameters, the square of the distance between the fit function and the data points needs to be minimized (for certain values of the fit parameters). This is called the principle of least squares. For notational simplicity, the data points are collected into the vector $\mathbf{D}$, and the model parameters are collected into the vector $\mathbf{X}$ (not necessarily of the same size). There is also background information present, because this is indispensable in an experiment. In processing these data, the use of Bayes' rule comes in handy. Bayes' rule tells us that the posterior probability density function for the parameters X_i given data D_i can be written as a product of a prior and a likelihood function in the following fashion (Ref. [8]):

$$\underbrace{p(\mathbf{X}|\mathbf{D},I)}_{\text{posterior}} \propto \underbrace{p(\mathbf{D}|\mathbf{X},I)}_{\text{likelihood}} \times \underbrace{p(\mathbf{X}|I)}_{\text{prior}}. \tag{5.1}$$

In any case, the prior and likelihood function are probability densities on their own, such that in principle they should be normalized separately. But because they are multiplied with each other to form the posterior probability density function, both the prior and likelihood function do not necessarily need to be normalized (but normalizing them separately is of course much properer). The overall proportionality constant can be determined by normalizing the posterior probability density function. When one wants to know a real probability, one needs to determine this factor. If one only wants to see the behaviour of the posterior, this factor can be omitted.

5.1.2 Prior

A prior is a probability density function for the theory parameters $\mathbf{X}$ given the background information I. This probability density function is based on *prior knowlegde* about the parameters; it is a choice, a decision about what possible values one could obtain for the unknown parameters. A prior is chosen to be nonzero within a certain range of a parameter and zero outside of that range. It is much more elegant to normalize the prior, such that a uniform prior takes a value that is the recripocal of the range. For calculational convenience, the uniform prior is chosen to have a value of 1 within the range. The overall normalization factor can be absorbed in the overall normalization constant of the posterior. Mostly, a prior is taken to be uniform (if each value is equally likely). Another common choice for a prior is a Gaussian prior, for a parameter that has a very large probability to have a certain value and very small probability that the value deviates much from this value. In Figure 5.1 below, one finds 2 examples of priors for a parameter 'A'.



(a) Uniform prior (b) Gaussian prior

Figure 5.1: Two different priors for a parameter 'A' that could be used.

In this thesis, the model parameters, such as the mass or the coalescence time, are not the same as the fit parameters, due to the fact that Python is not able to work with the large numbers of the model parameters. The choice is made to use fit functions in an as easy form as possible. There is no prior knowlegde present about the *fit* parameters, but there is only information present about the *model* parameters, which are related to the *fit* parameters. A conversion from prior of the model parameters to the prior of the fit parameters is done by a coordinate transformation described in Appendix A.1. Due to this, the prior of the fit parameters will no longer be 'just' a constant, but will be a function of the fit parameters f($\mathbf{X}$). The boundaries also change, such that the relation between model and fit parameter needs to be used to determine the boundaries for the prior for the fit parameters (from the boundaries of the model parameters). This means that the prior takes the form as in Equation 5.2 below:

$$p(\mathbf{X}|I) = \begin{cases} f(\mathbf{X}) & \mathbf{X}_{min} \leq \mathbf{X} \leq \mathbf{X}_{max} \\ 0 & \text{else,} \end{cases} \Rightarrow \log p(\mathbf{X}|I) = \begin{cases} \log(f(\mathbf{X})) & \mathbf{X}_{min} \leq \mathbf{X} \leq \mathbf{X}_{max} \\ -\infty & \text{else,} \end{cases} \tag{5.2}$$

where the minimal and maximal values depend on the most probable value of the parameter in the used model.

But, the least square method is valid as long as the prior for the variables for which the least squares are calculated is assumed to be constant. Previously, it can be seen that a constant prior is used for the model parameters, from which a parameter dependent prior is found for the fit parameters. But here it is done the other way around: First a constant prior for the fit parameters is used to determine the most likely values for the fit parameters using the least χ^2 fit method. This fit parameter prior is transformed to a model parameter prior (with the inverse transformation), to see how the prior for the model parameters would behave if the prior for the fit parameters needs to be constant. Then the posterior and marginal probability density function with constant fit parameter prior are determined (such that the distributions for the most likely χ^2-values can be seen). As a last step, the non uniform prior for the fit parameter (determined from the constant prior for the model parameters) is determined, to see how the posterior and marginal probability density functions look like when this prior is used. The uniform prior for the fit parameters is then, with an inverse coordinate transformation, transformed into the prior for the model parameters. This is done in order to check what the prior for the model parameters would be, in order to obtain a uniform prior for the fit parameters. From this, a conclusion can be drawn about the prior for both the fit and model parameters. This can be done, because a prior is a guess based on prior knowledge, but it is a choice that is made by the data analyst: the prior is in fact an arbitrary choice. Anyway, the uniform prior and logarithm of it for the fit parameters take the form:

$$p(\mathbf{X}, I) = \begin{cases} \mathcal{N} & \text{if } \mathbf{X}_{min} \leq \mathbf{X} \leq \mathbf{X}_{max} \\ 0 & \text{else.} \end{cases} \tag{5.3}$$

The constant factor of the prior $\mathcal{N}$ can be chosen to be 1, because it easily becomes 0 when taking the logarithm. It could be any other constant of course, but this constant can be absorbed into the overall proportionality constant in the posterior probability density function. The normalization constant $\mathcal{N}$ (and also this same symbol) is used in the result part of this thesis for different priors, but it can have a different value for each different prior. For each normalization constant, the same symbol $\mathcal{N}$ is therefore used multiple times.

5.1.3 Likelihood function

The likelihood function is the probability density function for the data points *given that the parameters are true*. To derive a probability density function for the data points, one starts with the noise. Assumed is that the noise behaves Gaussian, which means that it has a mean value of 0 and a standard deviation of σ_i, where i runs from 1 up to the number of data points. The probability density function for the noise (i.e. the probability density function to obtain a certain value for the noise) is therefore easily found to be:

$$p(n_i|I) \propto \frac{1}{\sigma_i} \exp\left(-\frac{n_i^2}{2\sigma_i^2}\right) \tag{5.4}$$

The noise at each data point is furthermore assumed to be independent of the noise at the other ones, such that the probability density for the whole noise set (i.e. at all data points together) can be written as:

$$p(\{n_i\}|I) = \prod_{i=1}^{N} p(n_i) \propto \prod_{i=1}^{N} \left(\frac{1}{\sigma_i} \cdot \exp\left[-\frac{n_i^2}{2\sigma_i^2}\right]\right). \tag{5.5}$$

From this, the likelihood for the data can be found. The difference between the value of a data point and the value that is predicted by the model is assumed to be caused only by the noise, i.e. $n_i = D_i - y_i$, where D_i represents data point i and y_i represents the predicted value of data point i. It is worth mentioning that the standard deviation of the data points is equal to the standard deviation of the noise, because only the noise is causing a deviation from the predicted model. Using that the noise is equal to the difference of the measured value and the value the model predicts, the likelihood function for the data points is found to be:

$$p(\mathbf{D}|\mathbf{X}, I) \propto \left[\prod_{i=1}^{N} \left(\frac{1}{\sigma_i}\right)\right] \cdot \exp\left[\sum_{i=1}^{N} -\frac{1}{2} \cdot \left(\frac{D_i - y_i}{\sigma_i}\right)^2\right] \tag{5.6}$$

It is better to work with the logarithm of the likelihood, which is calculated as:

$$\log p(\mathbf{D}|\mathbf{X}, I) \equiv \log(p(\{n_i\}|\mathbf{X}, I)) \propto -\frac{1}{2} \sum_{i=1}^{N} \left[\log\left(\sigma_i^2\right) + \left(\frac{D_i - y_i}{\sigma_i}\right)^2\right]. \tag{5.7}$$

5.1.4 (Log)posterior

According to the rule of Bayes, the posterior is obtained by multiplying the prior with the likelihood function. From the calculation rules of logarithms, it can be concluded that the logarithm of the posterior can be obtained by summing the logarithm of the prior and the logarithm of the likelihood function:

$$\begin{aligned} \log p(\mathbf{X}|\mathbf{D}, I) &\propto \log p(\mathbf{X}|I) + \log p(\mathbf{D}|\mathbf{X}, I) \\ &\propto \begin{cases} -\frac{1}{2} \sum_{i=1}^{N} \left[\log\left(\sigma_i^2\right) + \left(\frac{D_i - y_i}{\sigma_i}\right)^2\right] & \text{if } \mathbf{X}_{min} \leq \mathbf{X} \leq \mathbf{X}_{max} \\ -\infty \text{ if else} \end{cases} \end{aligned} \tag{5.8}$$

The total posterior can be found by exponentiating the logposterior:

$$p(\mathbf{X}|\mathbf{D}, I) \propto \left[\prod_{i=1}^{N} \frac{1}{\sigma_i}\right] \cdot \exp\left[-\frac{1}{2}\chi^2\right], \tag{5.9}$$

with

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{D_i - y_i}{\sigma_i}\right)^2 \tag{5.10}$$

There are certain values of the parameters that give a maximum in the likelihood function (which are the values that are thus most likely to be the true values of the parameters). This maximum of the posterior should be shaped like a peak, such that this peak contains the most probable values of the parameters. This corresponds in this case by minimizing the value of χ^2, which is called the *residual sum of squares*. With the values of the parameters that minimize the χ^2, a fit function can be plotted through the data. This method is therefore called the *least squares method*, such that the residual sum of squares is minimized.

5.1.5 Uncertainties in least squares fit

Of course the values of the fit parameters that maximize the likelihood function have a certain uncertainty. This uncertainty is linked to the width of the posterior function. For this, the Hessian is used, which is a second order derivative matrix defined as:

$$H_{ij} \equiv \frac{\partial^2 \left(\chi^2\right)}{\partial X_i \partial X_j}. \tag{5.11}$$

The covariance matrix is then related to the inverse of the Hessian (see Ref. [8]):

$$\sigma_{ij}^2 = 2 \left[H^{-1}\right]_{ij}, \tag{5.12}$$

where i and j are the indeces of the fit parameters. This means that both of them run from 1 up to the amount of fit parameters used in the model. The uncertainties in the fit parameters are given by the square root of the diagonal elements of the covariance matrix given in Equation 5.12. Due to the definition of this theory, the square root of the i-th diagonal element of the covariance matrix in Equation 5.12 corresponds to the uncertainty in the i-th fit parameter (i.e. the i-th entry of the vector $\mathbf{X}$).

5.1.6 Marginal probability density

It is also important to determine the marginal probability densities, i.e. the probability density of one variable irrespective of the outcome of the others. This so called marginal probability density is obtained by integrating out the other variables and keep the variable for which the marginal probability is determined X_i as a variable:

$$p(X_i|\mathbf{D}, I) \propto \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(X_1, ..., X_N|\mathbf{D}, I) \mathrm{d}^{N-1} X_{j \neq i}. \tag{5.13}$$

This is a one dimensional probability density function, in the sense that the other variables have been marginalized out; an integral is taken over all possible values, such that all these possible values of the parameters are taken into account. All that is left then is a marginal probability density function for one variable that is not integrated out. The infinitesimal in the integral ($\mathrm{d}^{N-1} X_{j \neq i}$) represents an $(N-1)$-dimensional volume element, not containing the variable for which the marginal probability density is determined.

Part II

Results

Chapter 6

Results for the original model: Binary

The theoretical derivation of this binary model is done in Chapter 2, such that this model of the frequency as function of time can be fitted to the obtained data from the event GW150914. Fitting the original model to the data gives better insights in the way gravitational wave data are processed, because the same simplified method is used as is used in Ref. [1]. The frequency dependence is derived, from which an estimation of the chirp mass and the coalescence time are determined. These are compared to the values obtained in Ref. [1], such that the method of working with gravitational wave data is used once for a model this is already done for. After this, the same method is used for the crazy models. For this original model, also a power law estimation is done between the gravitational wave strain and the time. The power law is derived theoretically and fitted through the data as well. This is not done by Ref. [1], but this is an extension to the research described in this paper. From this, again the coalescence time can be estimated, but also an attempt is done for the estimation of the effective distance between the detector and the source for both the detectors located in Livingston and Hanford.

6.1 Determining the Mass and the coalescence time

6.1.1 Fit function

From the form of the gravitational wave in Equation 2.11, it can be seen that if the gravitational wave frequency is raised to the power $-8/3$ and plotted against the time t, a linear function of the following form is a good candidate to fit through the data:

$$y = -A \cdot (x - B), \tag{6.1}$$

where y corresponds to $f_{\mathrm{GW}}^{-8/3}$ in $\mathrm{Hz}^{-8/3}$ and x corresponds to the time t in s.

6.1.2 Relating the fit parameters to the model parameters

The fit parameters A and B are related to the model parameters $\mathcal{M}$ and t_0 as follows (by comparing Equation 6.1 with Equation 2.11):

$$A = \frac{(8\pi)^{8/3}}{5} \cdot \left(\frac{G}{c^3}\right)^{5/3} \cdot \mathcal{M}^{5/3} \tag{6.2a}$$

$$B = t_0 \tag{6.2b}$$

These relations are easily inverted to give:

$$\mathcal{M} = A^{3/5} \cdot \left(\frac{c^3}{G}\right) \cdot \left(\frac{5}{(8\pi)^{8/3}}\right)^{3/5} \tag{6.3a}$$

$$t_0 = B. \tag{6.3b}$$

The determinant of the Jacobian needed for the coordinate transform is in fact one dimensional, because B is identically t_0. It is calculated to be:

$$\det(J) = \frac{\partial \mathcal{M}}{\partial A} = \frac{3}{5} \cdot \left(\frac{c^3}{G}\right) \cdot \left(\frac{5}{(8\pi)^{8/3}}\right)^{3/5} \cdot A^{-2/5}. \tag{6.4}$$

6.1.3 Prior for model and fit parameters

Uniform prior for model parameters; non uniform prior for fit parameters

As it is already said, the fit parameters and model parameters are in this case related non linearly, which means that the prior for the fit parameters becomes a function of these parameters (instead of being constant). This will be discussed further on more broadly. For the model parameters, the prior for the chirp mass is chosen to be uniform (it has a constant value) between 0 and 100 solar masses and 0 outside of that. This is chosen this way to have a wide range of possible realistic values, such that the most probable value lies within this range. For the coalescence time, the prior does also have a constant value in between -0.1 and 0.1, because from the plot it can be seen that the time at which the signal is present (and where the coalescence time needs to be) certainly lies in between these values. Because the chirp mass and the coalescence time are independent from one another, the combined prior is just the product of the separate priors:

$$p(\mathcal{M}, t_0|I) = p(\mathcal{M}|I) \times p(t_0|I) = \begin{cases} \mathcal{N} & \text{if } 0 < \mathcal{M} \leq 100 M_\odot \ \& \ -0.1 \leq t_0 \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{6.5}$$

Strictly speaking, the prior above needs to be normalized, i.e. $\mathcal{N}$ should have the value $\mathcal{N} = [(100 M_\odot \cdot 0.2)]^{-1}$, such that the volume under the two dimensional prior is 1. But as is explained in Chapter 5, this normalization constant $\mathcal{N}$ can be absorbed into the overall normalization constant of the posterior (such that for convenience the prior is chosen to have value 1). With the coordinate transformation, the prior for the fit parameters becomes:

$$p(A, B) = p(\mathcal{M}, t_0) \cdot \frac{\mathrm{d}\mathcal{M}}{\mathrm{d}A} \propto \begin{cases} A^{-2/5} & \text{if } 0 \leq A \leq 3.33 \ \& \ -0.1 \leq B \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{6.6}$$

Constant prior for fit parameters

With the prior in Equation 6.6, an estimate of the posterior probability density function can be made for the fit parameters, but the most probable value is not determined using this prior (such that the most probable value determined with the χ^2 method does not correspond to the most probable value in these posterior and marginal probability density functions). In the derivation for the least χ^2 method, a constant prior is needed, such that only the likelihood function is depending on the model parameters. Now the same range for the model parameters is used (because the range of possible values does not change), but the function value of the prior is taken to be constant (and A and B are assumed to be independent):

$$p(A, B) = \begin{cases} \mathcal{N} & \text{if } 0 \leq A \leq 3.33 \wedge -0.1 \leq B \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{6.7}$$

Again, this prior is normalized if $\mathcal{N}$ takes the uniform value $\mathcal{N} = [3.33135 \cdot 0.2]^{-1}$. This normalization factor is absorbed into the overall normalization constant appearing in the posterior, such that the prior takes the value 1. With a coordinate transform between model and fit parameters, the prior for the model parameters becomes:

$$p(\mathcal{M}, t_0) \propto \begin{cases} \mathcal{M}^{2/3} & \text{if } 0 \leq \mathcal{M} \leq 100 M_\odot \wedge -0.1 \leq t_0 \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{6.8}$$

This prior is not uniform, but it is nowhere sharply peaked and it is broad enough. Because a prior is a choice, it is legit estimation to take a constant prior for the fit parameters. Then a uniform prior for the fit parameters can be used, such that the posterior and marginal probability density functions belong to the values and uncertainties of the most probable values determined with the least χ^2 method. Both priors are used in the calculations.

6.1.4 Plot of the data and fit: uniform prior for the fit parameters

The data and uncertainties are obtained via the procedure described in Section 4.5. When the frequency estimated from the data of both detectors combined is raised to the power $-8/3$ and is plotted against the time, one obtains the following fashion with least squares fit:

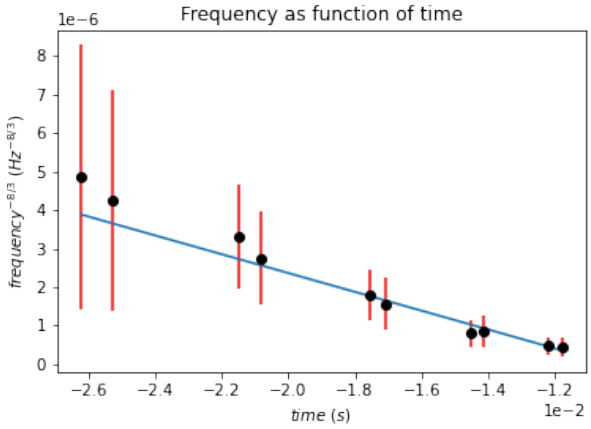


Figure 6.1: Frequency estimation raised to the power $-8/3$ plotted against the time t for the data of the detector in Livingston and Hanford combined. The red bars denote the error bars, whereas the blue line is the least squares fit line.

The fit parameters that minimize the residue are obtained with Python to be:

$$A = (0.24 \pm 0.02) \cdot 10^{-3} Hz^{-5/3} \tag{6.9a}$$

$$B = (-0.0118 \pm 0.0002) s \tag{6.9b}$$

These values and uncertainties are calculated using the least χ^2 approach (including the covariance matrix) described in Chapter 5.

6.1.5 Posterior for fit parameters; uniform prior for fit parameters

The marginal and posterior probability density functions calculated with the uniform prior for the fit parameters belong to the most likely values for the fit parameters obtained with the least χ^2 method. These probability density functions describe how the most likely values of A and B are distributed. The prior of Equation 6.7 is used. See Figure 6.2 below.

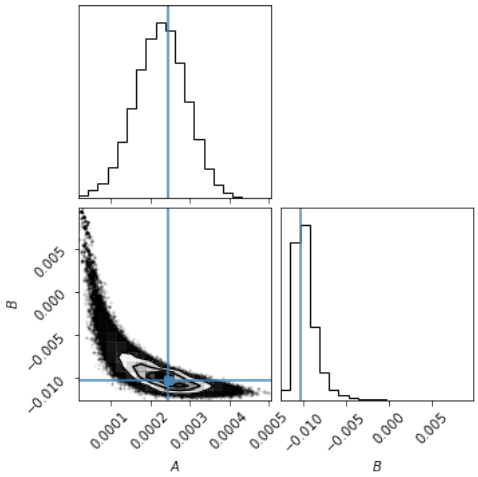


Figure 6.2: Likelihood function and marginal probability density functions for the fit parameters for the frequency evolution in time. The values that maximize the likelihood function are denoted with the blue lines, and the point where the maximum of the likelihood function is located is denoted with a blue dot. A uniform prior for the fit parameters is used here, such that the blue lines correspond to the maximum values in the marginal probability density functions.

6.1.6 Posterior for fit parameters; non uniform prior for fit parameters

With the recipe in Section 5.1.6, one can determine the form of the (marginal) likelihood function for the fit parameters using the prior specified before. To determine the pdf for A and B and the marginal pdf for both of them separately, the Python Corner package is used. With the prior of Equation 6.6, the following posterior and marginal probability density function are found for A and B:

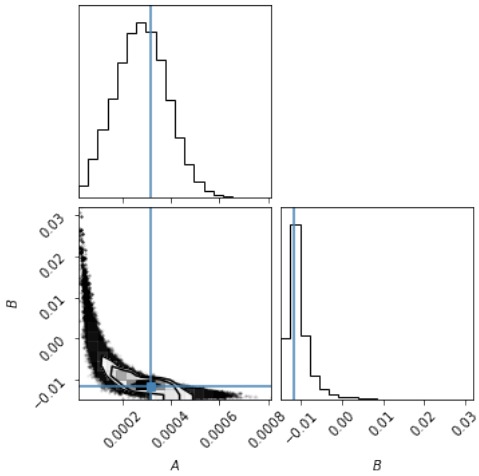


Figure 6.3: Likelihood function and marginal probability density functions for the fit parameters for the frequency evolution in time. A non uniform prior for the fit parameters is used here. The values that maximize the likelihood function are denoted with the blue lines, and the point where the maximum of the likelihood function is located is denoted with a blue dot.

The value of A and B that minimizes the residu is denoted with a blue line in the marginal pdf and with a blue dot in the likelihood function. Keep in mind that these values are the most probable values obtained by maximizing the posterior probability density function when the prior in Equation 6.7 is used,

and are not the most probable values belonging to the distributions shown in Figure 6.3.

6.1.7 The values of the model parameters

The most likely values of the model parameters are determined via the obtained relations in Equation 6.3, where the uncertainties are calculated using Appendix A.2:

$$\mathcal{M} = (20.6 \pm 1.5) \cdot M_{\odot} \tag{6.10a}$$

$$t_0 = (-0.0103 \pm 0.0003)\text{s} \tag{6.10b}$$

The values of the uncertainties in the chirp mass $\mathcal{M}$ and the coalescence time t_0 are 68%-interval uncertainties. This means that these uncertainties are 1 standard deviation in the accessory parameter.

6.2 Estimate the power in the amplitude time plot

Determining the correct order of magnitude of the amplitude is a tedious task, which goes beyond the scope of this thesis. Despite this, an attempt is done to estimate the effective distance in Appendix C. For this thesis, it is enough to estimate the power that pops up in the relation between the (dimensionless) gravitational wave strain $h(t)$ and the time t. The relation takes the form of Equation 2.12. Because the gravitational wave strain is dimensionless, its logarithm may be taken without changes:

$$\log(h(t)) = \log(C_3) - \frac{1}{4}\log(t_0 - t), \tag{6.11}$$

where

$$C_3 \equiv \left(\frac{D_{\text{eff}}^4 \cdot c^{11}}{5M^5G^5}\right)^{-1/4}. \tag{6.12}$$

If one defines now $x \equiv \log(t_0 - t)$ and $y = \log(h)$, the most appropriate fit function to determine the power to which the time has to be raised in order to obtain a linear relationship between the gravitational wave strain and the time equals:

$$y = n \cdot x + b, \tag{6.13}$$

where n is the interesting parameter to be estimated. The factor should equal the value $-1/4$, but this will be estimated with a least square fit for the data from Hanford and Livingston separately. In the fit function of Equation 6.13, the value for t_0 obtained in the frequency estimation will be used here. This value equals $t_0 = -0.0103s$

6.2.1 Estimation of the power for the data from Livingston

If the logarithm of the gravitational wave strain is now plotted against the logarithm of $t_0 - t$, the following is obtained (See Figure 6.4):

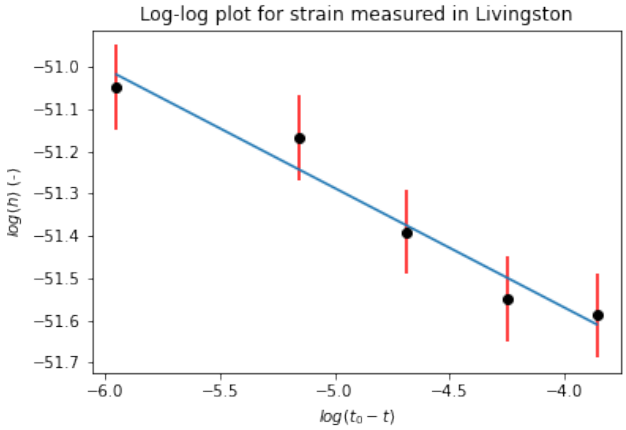


Figure 6.4: The logarithm of the gravitational wave strain plotted against the logarithm of the time for the data from Livingston. The error bars are denoted with red bars and the blue line is the least squares fit line.

The value of the power n is estimated to be:

$$n = -0.28 \pm 0.04. \tag{6.14}$$

This uncertainty is again one standard deviation, calculated with the rules in Appendix A.2. As can be seen, this result is good in accordance with the expected value of $-1/4$ in Equation 2.12.

6.2.2 Estimation of the power for the data from Hanford

For the data from Hanford, the exact same is done as for the data from Livingston. The following figure is obtained when plotting the logarithm of the gravitational wave strain from Hanford against the logarithm of the time. See Figure 6.5 below.

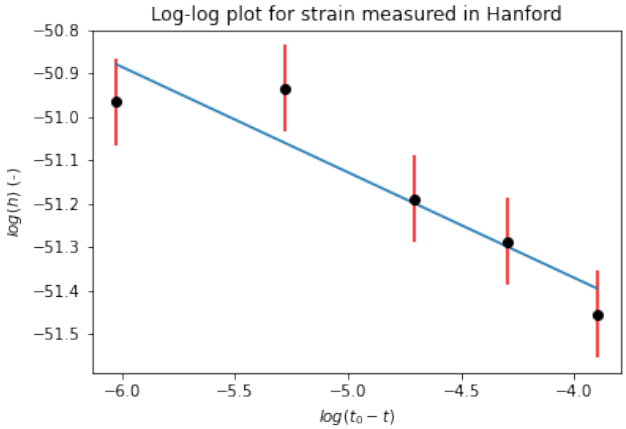


Figure 6.5: The logarithm of the gravitational wave strain plotted against the logarithm of the time for the data from Hanford. The error bars are denoted with red bars and the blue line is the least squares fit line.

The value of the power obtained from the data from Hanford equals:

$$n = -0.24 \pm 0.04 \tag{6.15}$$

This uncertainty is of course also one standard deviation and is calculated via the rules in Appendix A.2. As can be seen, both the value of the power obtained for the data from Livingston and from the data

from Hanford are in accordance with the theoretical derivation (Equation 2.12). The uncertainties for both the result of Hanford as well as for the result of Livingston are really small, which makes both the values reliable.

Chapter 7

Results for the binary with Time varying mass

Now the results are shown for the binary system with time dependent mass, derived in Chapter 3. Although this model is (astro) physically unlikely, it is worth studying the physics and behaviour of this system. The results are shown in the same way as was done for the original model.

7.1 Fit function

In order to determine the form of the fit function, the fashion of Equation 3.3 is carefully considered. This means that a fit function of the following form is appropriate:

$$y = A \cdot x^2 + B \cdot x + C \tag{7.1}$$

Here y represents the frequency raised to the power $-8/3$ in $Hz^{-8/3}$ and t the time in s. This is a fit function where the model parameters and fit parameters are not linearly linked to one another, such that a slightly tedious coordinate transform is necessary.

7.2 Relating model parameters to fit parameters

In the fit function, the three parameters A, B, and C are used as in Equation 7.1. The fit parameters and model parameters are related to one another as follows (as can be concluded by comparing 3.3 with 7.1):

$$A = \frac{(8\pi)^{8/3}}{5} \cdot \left(\frac{G}{c^3}\right)^{5/3} \cdot \left[\frac{5}{3} \cdot \dot{\mathcal{M}}_0 \cdot \mathcal{M}_0^{2/3}\right] \tag{7.2a}$$

$$B = \frac{(8\pi)^{8/3}}{5} \cdot \left(\frac{G}{c^3}\right)^{5/3} \cdot \left[\frac{10}{3}\dot{\mathcal{M}}_0\mathcal{M}_0^{2/3}t_0 - \mathcal{M}_0^{5/3}\right] \tag{7.2b}$$

$$C = \frac{(8\pi)^{8/3}}{5} \cdot \left(\frac{G}{c^3}\right)^{5/3} \cdot \left[\mathcal{M}_0^{5/3}t_0 - \frac{5}{3}\dot{\mathcal{M}}_0\mathcal{M}_0^{2/3}t_0^2\right] \tag{7.2c}$$

These relations are solved for $\mathcal{M}_0$, $\dot{\mathcal{M}}_0$ and t_0 to give the only physical solution (i.e. all parameters are real and the mass $\mathcal{M}_0$ is positive):

$$\mathcal{M}_0 = \left(\frac{125}{(8\pi)^8}\right)^{1/5} \cdot \frac{c^3}{G} \cdot D^{3/10} \tag{7.3a}$$

$$\dot{\mathcal{M}}_0 = -\frac{3}{5} \cdot \left(\frac{125}{(8\pi)^8}\right)^{1/5} \cdot \frac{c^3}{G} \cdot A \cdot D^{-1/5} \tag{7.3b}$$

$$t_0 = -\frac{B + \sqrt{D}}{2A}, \tag{7.3c}$$

Where $D \equiv B^2 - 4 \cdot A \cdot C$. This quantity appears in all three solutions, such that it is given its own name for convenience. With these relationships, the prior for the model parameters can be converted to the prior for the fit parameters, using a coordinate transform.

7.3 Appropriate prior for the fit and model parameters

7.3.1 Uniform prior for the model parameters

The prior for the chirp mass of the binary system is chosen to be 1 between 0 and 100 solar masses, because the with this model estimated value of the mass is assumed to be somewhere in the vicinity of the value obtained with the original model (also due to the fact that the mass change is assumed to be small), and 0 outside of that. The value of the chirp mass is already estimated for a model with time independent mass, so we might expect that the value this model predicts differs not that much from the value of the original model, but we are not sure about this.
For the chirp mass loss per unit of time, the prior is chosen to be Gaussian around $\dot{M}_0 = 0$. This is chosen this way, because if one does believe in General Relativity, this factor should be 0 (because the original model does already fit very well to the data from the event GW150914). So it is most likely that the value for $\dot{M}_0$ equals 0, but it could be possible that it is slightly greater or smaller than 0 in this model. To include the most exotic values of the mass loss, the width σ of the prior is chosen to be $\sigma = 250 M_\odot$, and the range the mass loss is assumed to lie within is chosen to be: $|\dot{\mathcal{M}}_0| \leq 3\sigma$.
For the coalescence time, a the exact same prior is used as for the original model. This time needs to have the same value, because it is an intrinsic property of the binary (the binary does not merge at a later time when using a quadratic fit instead of a linear fit). All three parameters are independent of one another, such that the combined prior is just the product of the three separate priors. The prior takes the form:

$$p(M_0, \dot{M}_0, t_0|I) = \begin{cases} \mathcal{N} \cdot \exp\left(-\frac{(\dot{\mathcal{M}}_0)^2}{2\sigma^2}\right) & \text{if } -3\sigma \leq \dot{\mathcal{M}}_0 \leq 3\sigma \text{ and } 0 < \mathcal{M}_0 \leq 100 M_\odot \ \& \ -0.1 \leq t_0 \leq 0.1 \\ 0 \text{ else,} \end{cases} \tag{7.4}$$

Where $\mathcal{N}$ is a normalization factor. Again, it is much proper to normalize a prior (as is explained that it is a probability density function on its own). The normalization factor $\mathcal{N}$ takes the form:

$$\mathcal{N} = \frac{1}{[2.49986 \cdot \sigma] \cdot [100 M_\odot] \cdot [0.2]}, \tag{7.5}$$

where the prior is in the $\dot{\mathcal{M}}_0$ integrated between $-3\sigma \leq \dot{\mathcal{M}}_0 \leq 3\sigma$, which does not deviate much from an integration from $-\infty$ to $+\infty$.

The prior in Equation 7.4 is the prior for the physical parameters $\mathcal{M}_0$, $\dot{\mathcal{M}}_0$ and t_0. A coordinate transformation from these parameters to A, B and C gives the prior in the fit parameters. This is thus a three dimensional transformation. The relations in Equation 7.2 are substituted, and the model prior is multiplied with the determinant of the Jacobian:

$$\det(J) \equiv \det\left(\frac{\partial(\mathcal{M}, \dot{\mathcal{M}}_0, t_0)}{\partial(A, B, C)}\right) = -\frac{2.82679 \cdot 10^{-10} c^{15}}{G^5 \mathcal{M}_0^3} = -1.33959 \cdot 10^{67} \cdot D^{-9/10} \tag{7.6}$$

The proportionality constant in the Jacobian can be left out, because this can be absorbed into the overall normalization constant of the posterior probability density function. The prior for the fit parameters then takes the form:

$$p(A, B, C|I) \propto \begin{cases} \frac{1}{(B^2 - 4 \cdot A \cdot C)^{9/10}} \cdot e^{\left(-\frac{5.99 \cdot A^2}{\left[B^2 - 4 \cdot A \cdot C\right]^{2/5}}\right)} & \text{if } |A| \leq 0.042 \wedge -0.012 \leq B \leq 5.01 \cdot 10^{-3} \wedge |C| \leq 7.52 \cdot 10^{-4} \\ 0 & \text{else,} \end{cases} \tag{7.7}$$

where the boundaries for A, B and C are determined via Equation 7.2, by looking at the maximum and minimum.

7.3.2 Uniform prior for the fit parameters

If the prior for the fit parameters is chosen to be uniform (with which the least χ-squared fit values are calculated), the boundaries of the fit and model parameters do not change; the relation between the model and fit parameters do not change. The parameters A, B and C are assumed to be independent. The prior for the fit parameters takes the following form:

$$p(A,B,C) = \begin{cases} \mathcal{N} & \text{if } |A| \leq 0.042 \wedge -0.012 \leq B \leq 5.01 \cdot 10^{-3} \wedge |C| \leq 7.52 \cdot 10^{-4} \\ 0 & \text{else.} \end{cases} \tag{7.8}$$

For convenience the normalization factor is chosen to be 1, but it can be chosen to be the reciprocal of the three dimensional volume of the fit parameter space in order for the prior to be normalized:

$$\mathcal{N} = \frac{1}{[0.042] \cdot [0.01701] \cdot [7.52 \cdot 10^{-4}]}. \tag{7.9}$$

With the prior in Equation 7.8, and the help of the Jacobian in Equation 7.6, the prior for the model parameters is then determined to be:

$$p(\mathcal{M}_0, \dot{\mathcal{M}}_0, t_0) \propto \begin{cases} \mathcal{M}_0^3 & \text{if } 0 < \mathcal{M}_0 \leq M_\odot \wedge |\dot{\mathcal{M}}_0| \leq 3\sigma \wedge |t_0| \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{7.10}$$

This prior is not uniform, but again is nowhere sharply peaked. A prior is a choice, and together with the fact that the prior is nowhere sharply peaked, also a constant prior should work for both the model parameters and the fit parameters. The uniform prior for the fit parameters is used for calculating the most likely values obtained with the least χ^2 method described in Chapter 5. Here the sharply peaked prior for the chirp mass loss $\dot{\mathcal{M}}_0$ is lost, but a prior is a choice. It is expected that the value of the chirp mass loss will be determined to be 0 (hence a Gaussian prior around 0), but also a constant prior can be used (to make every value equally likely).

7.4 Plot and fit of the data

The quadratic fit function that maximizes the likelihood function (and with constant prior thus maximizes the posterior probability density) is fitted through the data points, as can be seen in Figure 7.1 below. Here the least χ-squared method is used, such that the uniform prior for the fit parameters of Equation 7.8 is used here. See Figure 7.1 below:

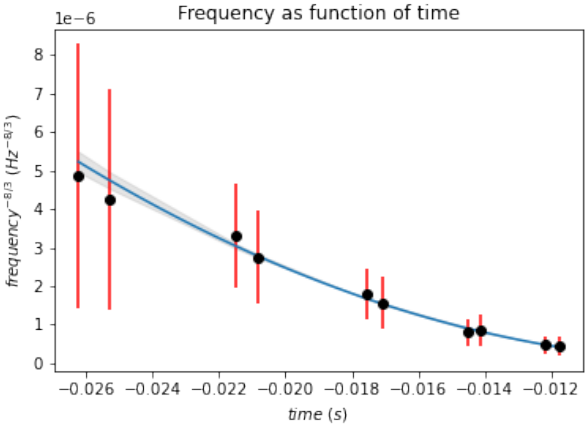


Figure 7.1: Quadratic fit function fitted through the data from the event GW150914. The red bars denote the error bars at each data point, and the blue line represents the least squares quadratic fit line. The grey shaded area is the area of one standard deviation for the parameters.

For the fit parameters, the following values (including uncertainties) maximize the likelihood function:

$$A = (1.3 \pm 0.3) \cdot 10^{-2} s^{2/3} \tag{7.11a}$$

$$B = (1.7 \pm 0.8) \cdot 10^{-4} s^{5/3} \tag{7.11b}$$

$$C = (6.0 \pm 6.3) \cdot 10^{-7} s^{8/3} \tag{7.11c}$$

In the relations between the model and fit parameters, one can see that there is a factor $B^2 - 4AC$, which needs to be positive in order for the model parameters to be real. The uncertainties of A and B are small compared to their value, whereas the value of C has a very large uncertainty. The value of C can therefore be treat a bit different from the values of A and B (because A and B are determined more accurately and are therefore more reliable). In order to satisfy the condition that $B^2 - 4AC > 0$ holds, the following upper bound for C is chosen:

$$C < \frac{B^2}{4A} = 1.24785 \cdot 10^{-7} s^{8/3} \tag{7.12}$$

Here the minimal value of B and maximal value of A is used. This is done because now the most extreme values of A and B that follow from Equation 7.11 are taken into account. This ensures us that the solution is always physical. It turns out that the mass loss rate has a very small imaginary part, such that it can be regarded as a small founding failure. The value of C that is chosen to work it lies in between the possible range of values ($-0.313 \cdot 10^{-7} \leq C \leq 1.24785 \cdot 10^{-7}$ equals: $C = (4.67 \pm 7.80) \cdot 10^{-8}$, where the uncertainty is determined to be half the obtained possible range for C.

For the coalescence time, the mass and the mass change per unit time, this means that the following values maximize the Likelihood function (and minimize the sum of the squared distances between model and data point):

$$\mathcal{M}_0 = (16.5 \pm 8.0) M_\odot \tag{7.13a}$$

$$\dot{\mathcal{M}}_0 = (-8.0 \pm 3.0) \cdot 10^2 M_\odot / s \tag{7.13b}$$

$$t_0 = (-0.013 \pm 0.008) s \tag{7.13c}$$

The uncertainties are calculated using 68%-intervals described in Appendix A.2, such that the values of the uncertainties are 1 standard deviation of the accessory variable. As can be seen, the values of the chirp mass $\mathcal{M}_0$ and the coalescence time t_0 are really in accordance with the values estimated with the original model.

7.5 Posterior for fit parameters

7.5.1 Uniform prior for the fit parameters

The posterior and marginal probability density function for the fit parameters, with most probable values given in the previous section, are calculated using the prior in Equation 7.8. It is determined using the procedure described in Chapter 5 and the Python "Corner" package. This gives the following, as can be seen in Figure 7.2

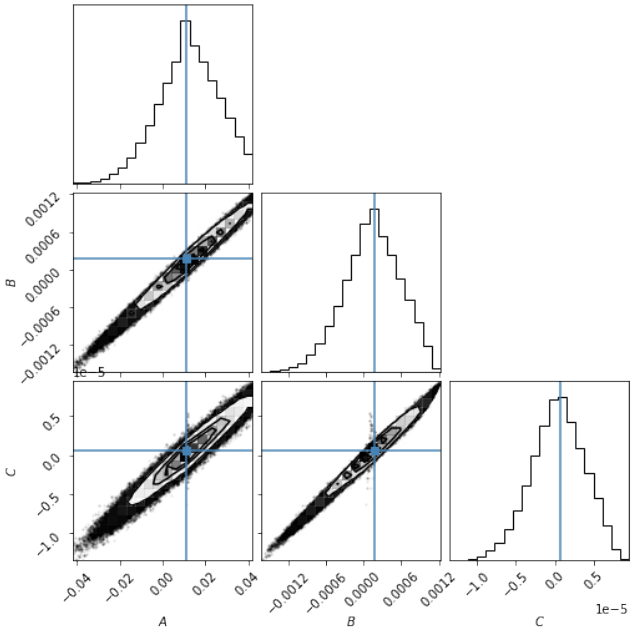


Figure 7.2: Posterior and marginal probability density functions for the fit parameters of the quadratic model. A uniform prior for the fit parameters is used. The values that maximize the likelihood function are denoted with blue lines and blue dots.

7.5.2 Non uniform prior for the fit parameters

Using the non uniform prior for the fit parameters in Equation 7.7, again with the Python "Corner" package a posterior is obtained using this prior. The following posterior function and marginal probability density functions are found and plotted in Figure 7.3:

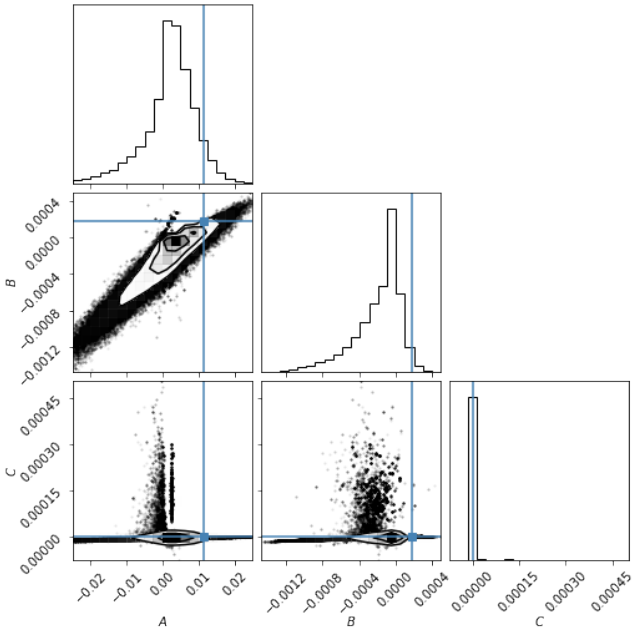


Figure 7.3: Posterior and marginal probability density functions for the fit parameters of the quadratic model. A non uniform prior for the fit parameters is used. The values that maximize the likelihood function are denoted with blue lines and blue dots.

As can be seen, the probability density function with uniform prior for the fit parameters gives a more accurate and reliable result.

Chapter 8

Results for the spirally shaped mass

The last model that is considered is the spirally shaped mass distribution. The theoretical concepts are derived in Chapter 3. The results are again shown in the same way as for the previous models. This model is (astro) physically very unlikely to exist, but due to the unknown physical properties of this model, this model could perhaps match the data really well.

8.1 Fit function used

As is already mentioned, the parameter α determines the size of the spiral. To take away one of the three degrees of freedom, the α is fixed such that it gives the spiral the size of its own Schwarzschild radius. The fit function in terms of the fit parameters is for this model given by (which can be concluded from Equation 3.27):

$$y = m \cdot (x - b), \tag{8.1}$$

where y corresponds to f_{GW}^{-4} in Hz^{-4} and x corresponds to the time in s. For convenience during programming with Python, the same fit function is used for this model as is used for the original binary model, except for the minus sign in front.

8.2 Relating model parameters to fit parameters

To obtain a relationship between the model and fit parameters, Equation 8.1 is compared to Equation 3.27 to obtain the following relationships:

$$m = \frac{1}{20} \cdot \frac{1}{\Phi \pi^6 c^3} \cdot \frac{S}{P} \cdot M^3 \tag{8.2a}$$

$$b = t_0, \tag{8.2b}$$

such that the inverse relations are given by:

$$M = \left(20\Phi\pi^6 c^3 \cdot \frac{P}{S}\right)^{1/3} \cdot m^{1/3} = 4.199 \cdot 10^9 \cdot m^{1/3} \tag{8.3a}$$

$$t_0 = b. \tag{8.3b}$$

The determinant of the Jacobian of the coordinate transformation is calculated to be:

$$\det(J) = \left|\frac{\partial(M, t_0)}{\partial(m, b)}\right| = \frac{1}{3} \cdot \left(20\Phi\pi^6 c^3 \cdot \frac{P}{S}\right)^{1/3} \cdot m^{-2/3}. \tag{8.4}$$

8.3 Prior for the model and fit parameters

8.3.1 Non uniform prior for the fit parameters

For the purpose of working with a prior based on the prior knowledge of the behaviour of the model parameters, the prior for the model parameters is chosen to be uniform. All the values of the model

parameters are equally likely. As is already explained before, only 2 parameters are used in the fit function. This is due to the fact that the value of α is fixed (one value is chosen). For the mass and coalescence time, a constant prior is chosen between a certain range. This range is chosen for the mass to extend from 0 to 10^5 solar masses, and for the coalescence time the same range is chosen: $-0.1 \leq t_0 \leq 0.1$. The range in the mass is chosen this way because this is a very crazy model, which has, with all the other constants together, a smaller pre-factor in the amplitude, such that the mass needs to be much larger to match the data. The range for the coalescence time is chosen the same as for the other models, because from the data it can be concluded that the coalescence time should be certainly within this range. The mass and coalescence time are assumed to be independent, such that the prior becomes:

$$p(M, t_0|I) = \begin{cases} \mathcal{N} & \text{if } 0 \leq M \leq 10^5 M_\odot \text{ and } -0.1 \leq t_0 \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{8.5}$$

The normalization factor $\mathcal{N}$ can be chosen such that the prior is normalized (which gives a really elegant probability density function). Then the value of $\mathcal{N}$ would be:

$$\mathcal{N} = \frac{1}{[10^5 \cdot M_\odot] \cdot [0.2]}. \tag{8.6}$$

This normalization constant can, as is already explained, also be absorbed into the overall normalization constant of the posterior probability density function, such that for convenience the $\mathcal{N}$ can be chosen to be 1. With the determinant of the Jacobian in Equation 8.4, the prior for the fit parameters becomes, using the recipe described in Appendix A.1:

$$p(m, b|I) \propto \begin{cases} m^{-2/3} & \text{if } 0 \leq m \leq 4.4 \cdot 10^{-5} \text{ and } -0.1 \leq b \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{8.7}$$

8.3.2 Uniform prior for the fit parameters

Again, for the least squares approximation, a uniform prior for the fit parameters is needed. This takes the form as in Equation 8.8 below:

$$p(m, b|I) = \begin{cases} \mathcal{N} & \text{if } 0 < m \leq 4.4 \cdot 10^{-5} \text{ and } -0.1 \leq b \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{8.8}$$

The normalization factor $\mathcal{N}$ is chosen to be 1 for calculational convenience in Python, but as is already mentioned. A prior is a probability density function on its own, such that normalizing this function is much more elegant. This constraints the normalization factor to take the value:

$$\mathcal{N} = \frac{1}{[4.4 \cdot 10^{-5}] \cdot [0.2]} \tag{8.9}$$

With the prior in Equation 8.8 and the Jacobian of the coordinate transform in Equation 8.4, the prior for the model parameters is calculated to be:

$$p(M, t_0) \propto \begin{cases} M^{-2} & \text{if } 0 < M \leq 10^5 M_\odot \\ 0 & \text{else.} \end{cases} \tag{8.10}$$

This is a non uniform prior, but it is also not sharply peaked somewhere (except at $M = 0$, but this value is excluded from the prior). This means that a constant prior is a good choice (because a prior can be determined arbitrarily; it is a choice). Normalizing the prior in Equation 8.10 is a bit difficult (due to the vertical asymptote at $M = 0$). This can be solved by taking a value that is slightly above $M = 0$ (for which a smaller or equal sign can be used, such that one has a lower bound that is included).
The result is again shown using both priors; a non uniform prior for the fit parameters determined from a uniform prior for the model parameters (Equation 8.7 is used to determine the posterior) and a uniform prior for the fit parameters (Equation 8.8 is used to determine the posterior).

8.4 Plot and fit of the data: uniform prior for fit parameters

Using the least χ^2-method described in Section 5, the most likely fit parameters are determined and with them the least square fit line is fitted. The uniform prior of Equation 8.8 is used. See Figure 8.1 below:

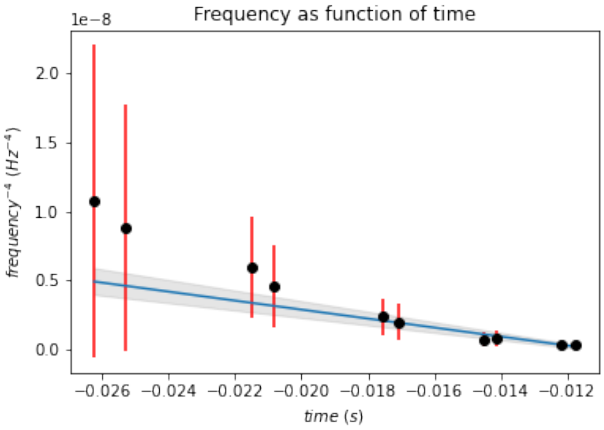


Figure 8.1: Caption

The values for m and b that minimize the residue are determined to be:

$$m = (-3.3 \pm 0.6) \cdot 10^{-7} s^3 \tag{8.11a}$$
$$b = (-0.0112 \pm 0.0003) s \tag{8.11b}$$

The best fit value for m, the value that maximizes the likelihood function, is a negative value. From this we can conclude, by comparing this with Equation 8.1, that the prefactor in front of $(t - t_0)$, needs to be negative. All constants in front of $(t - t_0)$ in Equation 3.27 are positive quantities (which means C_0 is a positive constant), whereas the only thing not yet known is the mass M. The only way that the negative value of m can be matched with the prefactor, is that M needs to be a negative number. This immediately rules out this theoretical model, because only a real positive mass is a physically realistic mass. This model can be excluded from potential sources that create these kind of gravitational wave signals. This is due to the fact that the data predict an increasing frequency as function of time (for times $t < t_0$), whereas this model has a decreasing frequency fashion. It was important to check if maybe, within the uncertainty, m could have a positive value, such that this model would allow a positive mass. But, as can be seen, the fit parameter m can only have a negative value within the uncertainty. The value of the mass is calculated to show what happens. With these values, the Mass and Coalescence time are determined to be:

$$M = (-2.0 \pm 0.1) \cdot 10^4 M_\odot \tag{8.12a}$$
$$t_0 = (-0.0112 \pm 0.0003) s. \tag{8.12b}$$

8.5 Posterior probability density function for the fit parameters

8.5.1 Uniform prior for the fit parameters

The values that minimize the χ^2 are determined using the uniform prior of Equation 8.8. The distribution of the fit parameters using this prior can be seen in Figure 8.2 below:

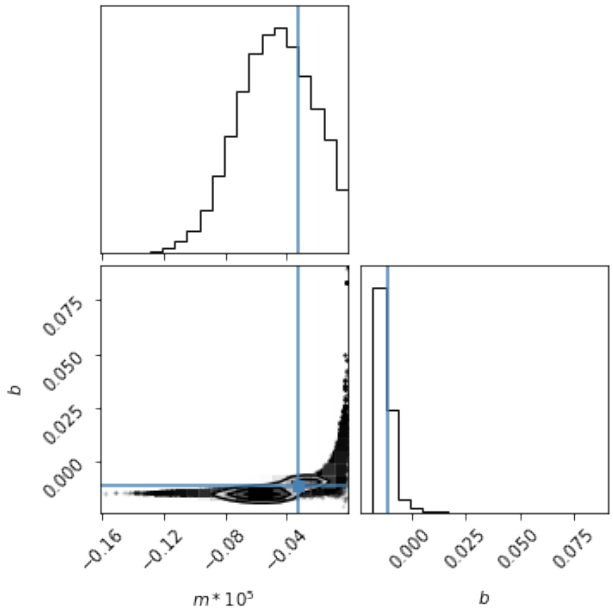


Figure 8.2: Likelihood function and marginal probability density functions for the fit parameters of the spiral model. A uniform prior for the fit parameters is used.

8.5.2 Non uniform prior for the fit parameters

As is done for the previous models, a posterior probability density function and marginal probability density functions for the fit parameters are plotted using the non uniform prior of Equation 8.7. See Figure 8.3 below:

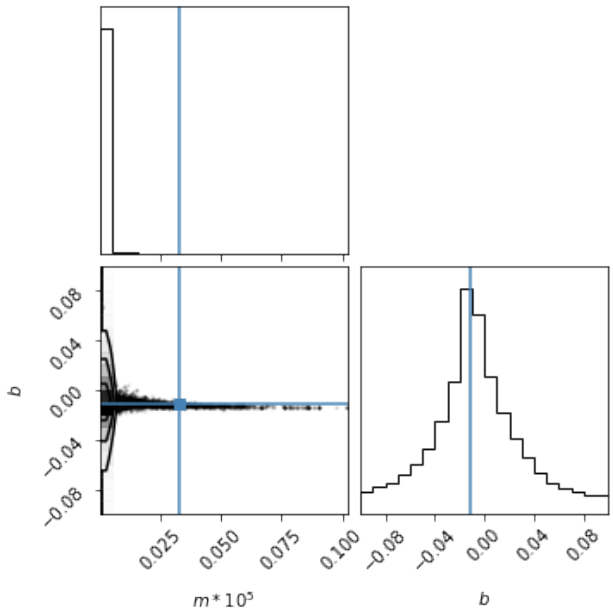


Figure 8.3: Likelihood function and marginal probability density functions for the fit parameters of the spiral model. A non uniform prior for the fit parameters is used.

As can be seen, the non uniform prior for the fit parameters does not work as well as the constant prior for the fit parameters. The uniform prior is more accurate, as can be seen by comparing Figure 8.3 with Figure 8.2. This is due to the fact that the posterior takes a much less complex form if a constant prior is used for the fit parameters (such that the most likely values are calculated more easily).

Part III

Conclusion

Chapter 9

Conclusion

In the previous part, the theoretically derived crazy models are fitted to the data from the LIGO detectors in Livingston and Hanford. Each model will be evaluated separately. The physically likeliness will be regarded, as well as the match with the data is discussed.

9.1 The original model

The original model consists of two black holes orbiting around each other. Their separation distance decreases, whereas their orbital frequency (and with that the gravitational wave frequency) increases due to Keplers laws. This model is of course physically very likely, because this model is made for a binary system that is being observed. This model belongs to the physics that created these gravitational waves.

The coalescence time matches with the gravitational wave signal, because in the plot of the signal one can see that the signal has a rapidly decreasing frequency and amplitude after the estimated coalescence time t_0. This value is (within the uncertainty) a very good estimation of the real time of merging.

The estimated value of the chirp mass is in the same order of magnitude as the one estimated by Ref. [1]. In this article, a simplified version of the determining of the chirp mass is presented. In this thesis, the same method is used as is described in this article. With this method, roughly the same value for the chirp mass is obtained, but due to the fact that an even more simplified version is used in this thesis, the value deviates a bit. In Ref. [1], there is done more research on the maximum value of the chirp mass and on the value of the masses of the individual black holes, but this was not the aim of this thesis. Only an estimation of the chirp mass was needed, because then different models can be compared.

9.2 The original model with time dependent chirp mass

In this model, the same physics is assumed to hold, except for a linear chirp mass change in time. This linear assumption is needed in order for the derivation to be still valid.

First, the estimation of the chirp mass is considered. The most likely value of the chirp mass that is determined with the original model, lies really well within the uncertainty determined with this model. Despite that, the most probable value of the chirp mass is slightly smaller with this model than is determined with the original model, whereas the uncertainty of this value determined with this model is smaller than the most likely value! This means that the most likely value is really reliable, which gives a reason to accept the model as possible crazy model that matches the data.

For the coalescence time, the most probable value determined with this model matches the most probable value determined with the original model very well. Also the uncertainty is smaller than the value of the coalescence time, which makes this value reliable.

The mass loss is unfortunately large compared to the total mass, which is physically not possible. This

makes this model unlikely to happen.

The values obtained with this model for the chirp mass $\mathcal{M}'$, the constant chirp mass change per unit of time $\dot{\mathcal{M}}'_0$ and the coalescence time t_0 all have a relatively small uncertainty, which makes them very reliable. This model cannot be immediately ruled out from the possible models that could be matched to the data obtained at Livingston and Hanford from the gravitational wave event GW150914; it matches the data from LIGO better than expected. But, this model is physically unlikely to happen because the mass loss is really large compared to the total mass.

9.3 Spiral model

For the spiral model, the possible value of its mass M and an offset time t_0 can be determined. In the derivation of this model, these parameters pop up. The offset time is a bit a strange parameter in this model, because the spiral only loses energy by slowing down (and there is no coalescence or merging in this model), such that this offset time can be interpreted as the time at which the spiral 'magically' stops rotating.

The value of this offset time really matches the signal, as it really agrees with the value of the coalescence time estimated with the original model. In the signal, one can see that the frequency becomes infinite at a certain time t_0 (i.e. the coalescence time in the original model). This time stamp is a property of the signal, which means that the same value should be obtained. The most probable value for this offset time calculated for the spiral model corresponds really well to the most likely value for the coalescence time estimated with the original model. The range of possible values is the same, i.e. the uncertainty is roughly the same for this model compared to the original model.

The most probable value for the mass of the spirally shaped mass distribution is a totally different story. The mass that is most likely has a negative sign, which is quite logical. The binary system is losing energy (i.e. getting from negative values close to zero to negative numbers farther away from zero) as it spins up. If the spiral would have a positive mass, it would emit gravitational waves as it slows down, but the energy goes from negative numbers far away from zero to negative numbers close to zero. This is opposite to the binary system, such that a positive mass could be ruled out. It could be possible that within the uncertainty a positive mass belonged to the possible values, but it turns out that only negative values within the uncertainty are possible. A negative mass causes the energy to follow the same fashion as the energy does for the binary system. Therefore the most probable value for the total mass must be a negative value, as can be concluded. The fact that the mass is negative shows that the model is completely (astro) physically unlikely, because negative masses are only a theoretical concept in current physics. Despite this, the (negative) value that is estimated is very reliable, because the relative uncertainty (i.e. the uncertainty divided by the most probable value) is small. But, it is only reliable if the concept of negative mass would be embraced in physics, if for example experimental results or astrophysical relations with negative masses are obtained. The mass is roughly 10^4 times as large as the mass of our own sun, which is a really common value for a mass in astrophysics. This object is assumed to have a size larger than its Schwarzschild radius, because the form of the distribution (spirally shape) is known. If such a model would exist, one should know the form of the distribution before the physics can be done. Otherwise it is just gambling about the shape of the distribution by looking how the frequency relates to the time.

Chapter 10

Discussion

In the previous chapter, the conclusion about the crazy models is taken: it is decided if they are (astro) physically likely or if they are improbable. This is done on the basis of the values of the model parameters; the (chirp) mass, coalescence time or chirp mass change per unit of time are determined and are compared to physically realistic values and their reliability is checked, i.e. how do their uncertainties behave. In this master thesis, the data is modeled in a really basic way with Python, in an even more simplified way as the researchers of Ref. [1] did. This means that the estimations and processes such as whitening could be done more precise when an extensive Python code is written and used. For the purpose of this thesis, it was enough to work with the basic Python code, written by myself. This is justified by the fact that only the procedure of processing gravitational wave data and the physical relevance of the models needed to be studied.

For further research, more crazy models can be thought of. For these newly invented models a theoretical derivation is needed in order for the data to be matched, in order to obtain a relation for the frequency as function of the time. For these models, again a likelihood function can be made, and their physically relevance and physical likeliness can be checked.

Furthermore, one could extend the crazy models discussed in this thesis as follows for example:

The spiral model could for example be extended by making its size time dependent. The angular momentum could be taken constant, such that the rotational angular frequency ω can be linked to the size factor α. This gives a spiral shaped rotating mass distribution, that can collapse and that eventually forms a perfect sphere. This gives an interpretation to the integrating constant t_0 that appears in the equation that describes the frequency as function of time. Then the time t_0 could be interpreted as the time of collapsing. This kind of spirally shaped rotating distribution was tried by me, but the equation that needed to be solved in order to find a relation between the frequency and the time could not be expressed in elementary functions. The solution needs to be approximated or estimated numerically, which is beyond the scope of this Master Thesis.

Also the current spirally shaped distribution can be extended by taking multiple values of α into account; it can be studied if different values of α fit the model even better or worse, or maybe give a positive mass as outcome.

The binary model with time dependent mass could for example be extended by not assuming a linear mass change, but taking higher order terms into account. One should start from elementary Newtonian mechanics, where the mass rate of change is taken into account as a so called 'thrust' force. This gives a totally different derivation, which makes the problem more complicated. This model I tried to derive myself, but it turned out to be a problem too complicated to solve within the Master Thesis. The choice is made to make the mass only linear time dependent, to learn how to work with gravitational wave data and different fit functions, rather than to do heavy time consuming hard elementary Newtonian physics.

Bibliography

[1] B. P. Abbott et al. "The basic physics of the binary black hole merger GW150914". In: *Annalen der Physik* 529.1–2 (Oct. 2016), p. 1600209. ISSN: 1521-3889. DOI: 10.1002/andp.201600209. URL: http://dx.doi.org/10.1002/andp.201600209.

[2] Bruce Allen et al. "FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries". In: *Physical Review D* 85.12 (June 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.85.122006. URL: http://dx.doi.org/10.1103/PhysRevD.85.122006.

[3] David Butler. *Classroom Aid - GW150914 Sky Location*. Youtube. 2016. URL: https://www.youtube.com/watch?v=E0SEAdcF6WU.

[4] M. P. Hobson, G. P. Efstathiou, and A. N. Lasenby. *General Relativity: An Introduction for Physicists*. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511790904.

[5] Michele Maggiore. *Gravitational waves. Volume 1, Theory and experiments Michele Maggiore*. eng. Oxford: Oxford University Press, 2007.

[6] Michele Maggiore et al. "Science case for the Einstein telescope". In: *Journal of Cosmology and Astroparticle Physics* 2020.03 (Mar. 2020), pp. 050–050. DOI: 10.1088/1475-7516/2020/03/050. URL: https://doi.org/10.1088%2F1475-7516%2F2020%2F03%2F050.

[7] M Saleem et al. "The science case for LIGO-India". In: *Classical and Quantum Gravity* 39.2 (Dec. 2021), p. 025004. DOI: 10.1088/1361-6382/ac3b99. URL: https://doi.org/10.1088%2F1361-6382%2Fac3b99.

[8] D. S. Sivia and J. Skilling. *Data Analysis - A Bayesian Tutorial*. 2nd. Oxford Science Publications. Oxford University Press, 2006.

[9] The LIGO Scientific Collaboration et al. *First joint observation by the underground gravitational-wave detector, KAGRA, with GEO600*. 2022. DOI: 10.48550/ARXIV.2203.01270. URL: https://arxiv.org/abs/2203.01270.

Appendix A

Probability Identities

A.1 Coordinate Transformations

Assume there are N variables $\{X_i\}$, $i \in [1, N]$. A probability density function is given by:

$$p(\{X_i\}) = f(X_1, \ldots, X_N), \tag{A.1}$$

where the probability density function is a function of all variables. One can also express the probability density function in terms of some other N variables $\{Y_i\}$, where the probability density function for the variables $\{Y_i\}$ take the form:

$$p(\{Y_i\}) = g(Y_1, \ldots Y_N). \tag{A.2}$$

The variables $\{Y_i\}$ are related to the variables $\{X_i\}$ as follows:

$$X_1 = h_1(Y_1, \ldots, Y_N) \tag{A.3a}$$

$$\vdots \tag{A.3b}$$

$$X_N = h_N(Y_1, \ldots, Y_N), \tag{A.3c}$$

where these h_i are the functional relations between the 2 sets. A furthermore necessary condition is that the determinant of the Jacobian of this transformation does not vanish (i.e. it does have an inverse). The Jacobian J is given by:

$$J \equiv \frac{\partial X_i}{\partial Y_j}, \text{ with } \det(J) \neq 0. \tag{A.4}$$

A true probability for the first set ($\{X_i\}$) that a point $(X_1, ..., X_N)$ lies within the infinitesimal volume element bounded by the values $X_1, ..., X_N$ and $X_1 + \mathrm{d}X_1, ..., X_N + \mathrm{d}X_N$ should equal the probability for the second set ($\{Y_i\}$) that a point $(Y_1, ..., Y_N)$ lies within the infinitesimal volume element bounded by the values $Y_1, ..., Y_N$ and $Y_1 + \mathrm{d}Y_1, ..., Y_N + \mathrm{d}Y_N$. The conversion factor equals the Jacobian:

$$g(\{Y_i\}) = f(\{X_i\}) \cdot \left| \frac{\partial X_i}{\partial Y_j} \right| = f(\{X_i\}) \cdot |\det(J)| \tag{A.5}$$

A.2 Uncertainties

The uncertainties in all the quantities in this thesis are calculated via the rules of 68%-intervals. The symbol S_X is used for the 68% uncertainty in variable X:

For a quantity Y depending on N other quantities $\{X_i\}$ (with $i \in [1, N]$) with uncertainties S_i, i.e. $Y = Y(\{X_i\})$, the uncertainty in variable Y is calculated to be:

$$S_Y = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial Y}{\partial X_i} \right)^2 \cdot S_i^2}. \tag{A.6}$$

Appendix B

Derivations

B.1 Transforming a tensor to the Transverse Traceless gauge

As is already explained, the metric perturbation is yet trace less by the definition of the quadrupole moment tensor, but not yet transverse. Only the spatial part is possibly nonzero, whereas all temporal components are zero. The metric perturbation is made transverse by projecting it onto the transverse plane. In this thesis, the Transverse trace less gauge is applied in the frame where the wave is propagating in the z-direction. This means that the wave has unit wave vector $\hat{k} = (0, 0, 1)$ in this frame. The projection tensor takes the following form for this wave vector:

$$P_{ij} = \delta_{ij} - k_i k_j, \tag{B.1}$$

which in matrix notation looks like:

$$P_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{B.2}$$

Because we are working with tensors with 2 indices, each index 'needs to be projected'. Take a general yet trace less tensor A_{ij}. The transverse projected tensor A^T_{ij} takes the form:

$$A^T_{ij} = P^k_i P^l_j A_{kl}. \tag{B.3}$$

The problem now is that this tensor A^T_{ij} is not necessarily trace less anymore. This is made trace less again by subtracting half the trace from each diagonal element (because the spatial row and column containing a z-entry are made 0 due to the transverse property):

$$A^{TT}_{ij} = \left(P^k_i P^l_j - \frac{1}{2} P_{ij} P^{kl} \right) A_{kl}. \tag{B.4}$$

B.2 Quadrupole and Inertia tensor

The definition for the inertia tensor is given by:

$$I_{ij} \equiv \int \rho(\vec{x}) \cdot \left[|\vec{x}|^2 \delta_{ij} - x_i x_j \right] \mathrm{d}^3 x, \tag{B.5}$$

where $\vec{x}$ is the position vector that points from the origin to a bit of mass $\mathrm{d}m = \rho(\vec{x})\mathrm{d}^3 x$. The Quadrupole moment tensor is closely related to the inertia tensor. The Quadrupole moment tensor is defined as:

$$Q_{ij} \equiv \int \rho(\vec{x}) \cdot \left[x_i x_j - \frac{1}{3} |\vec{x}|^2 \delta_{ij} \right] \mathrm{d}^3 x. \tag{B.6}$$

As can be seen, the quadrupole moment tensor is by definition traceless, whereas this does not hold for the inertia tensor. The quadrupole moment tensor is related to the inertia tensor via:

$$Q_{ij} = -I_{ij} + \frac{1}{3} \cdot \mathrm{Tr}(I) \cdot Id(3), \tag{B.7}$$

where $Id(3)$ represents the 3×3 identity matrix.

B.3 Combining a sine and cosine wave

Assume a wave form takes the following form:

$$w(t) = A \cdot \cos(C \cdot t) + B \cdot \sin(C \cdot t). \tag{B.8}$$

The prefactors A and B can without loss of generality be written as:

$$A = P \cdot \cos(Q) \tag{B.9a}$$
$$B = -P \cdot \sin(Q), \tag{B.9b}$$

due to the fact that there are still two degrees of freedom (P and Q), and because there exist a unique solution for P and Q in terms of A and B. When using the addition rule for cosines, on obtains for the wave signal:

$$w(t) = P \cdot \cos(Q) \cdot \cos(B \cdot t) - P \cdot \sin(Q) \cdot \sin(B \cdot t) = P \cdot \cos(B \cdot t + Q), \tag{B.10}$$

where now

$$P = \sqrt{A^2 + B^2} \tag{B.11a}$$
$$Q = \arctan\left(-\frac{B}{A}\right). \tag{B.11b}$$

Appendix C

Determining the effective physical distance

C.1 Determining the effective physical distance for Livingston

C.1.1 Fit function

The fit function should be of the form of the amplitude in the model, i.e. Equation 2.12. The function used to determine the model parameters is of the following form:

$$y = -A \cdot (x - B), \tag{C.1}$$

where y corresponds to the gravitational wave amplitude (strain) raised to the power -4 (dimensionless) and x corresponds to the time in s.

C.1.2 Relating fit parameters to model parameters

The parameters in Equation C.1 (A and B) are related to the model parameters D_{eff} and t_0 as follows:

$$A = \frac{c^{11}}{5G^5\mathcal{M}^5} \cdot D_{\text{eff}}^4 \tag{C.2a}$$

$$B = t_0, \tag{C.2b}$$

which are easily inverted to give:

$$D_{\text{eff}} = \left(\frac{5}{c}\right)^{1/4} \cdot \left(\frac{\mathcal{M}G}{c^2}\right)^{5/4} \cdot A^{1/4} \tag{C.3a}$$

$$t_0 = B. \tag{C.3b}$$

The Jacobian is calculated to be:

$$\frac{\partial D_{\text{eff}}}{\partial A} = \frac{1}{4} \cdot \left(\frac{5}{c}\right)^{1/4} \cdot \mathcal{M}^{5/4} \cdot A^{-3/4}. \tag{C.4}$$

C.1.3 Prior for model and fit parameters

Uniform prior for the model parameters

For the fit parameters, the prior is determined as follows: The coalescence time t_0 should be the same as for the frequency plot, because both the differential equation for the amplitude evolution as well as for the frequency evolution are integrated between time t and the coalescence time. It is a time stamp at which the frequency should be 0 and the amplitude should be infinite, according to the theory of the inspiraling binary and the characteristic of these gravitational waves. For the distance, the upper bound

for the prior is taken to be the order of magnitude of 10 times the farthest star discovered at Earth, just to be sure the binary is possibly located farther away than the farthest star. The lower bound can be chosen to be 0 for simplicity:

$$p(D_{\text{eff}}, t_0|I) = \begin{cases} \mathcal{N} & \text{if } 0 \leq D_{\text{eff}} \leq 10^{27} \ \& \ -0.1 \leq t_0 \leq 0.1 \\ 0 & \text{else.} \end{cases} \tag{C.5}$$

The normalization factor $\mathcal{N}$ can be chosen to be the reciprocal of the range of the parameters, but for calculational convenience this is chosen to be 1. Because the fit parameter B is exactly equal to the model parameter t_0, no coordinate transform is needed for this parameter. The coordinate transform $D_{\text{eff}} \to A$ needs a Jacobian, where only the variable A appearing here is kept; The overall proportionality constant will be omitted, because this can be corrected with the overall normalization constant for the likelihood function. The prior for the fit parameters takes the form:

$$p(A, B|I) = p(D_{\text{eff}}, t_0|I) \cdot \frac{\partial D_{\text{eff}}}{\partial m} \propto \begin{cases} A^{-3/4} & \text{if } 0 \leq A \leq 1.04225 \cdot 10^{93} \ \& \ -0.1 \leq B \leq 0.1 \\ 0 & \text{else,} \end{cases} \tag{C.6}$$

where the boundaries are calculated via the relation between the model parameters and the fit parameters in Equation C.2.

Uniform prior for the fit parameters

As is done for each model, the least χ^2-method is valid as long as a uniform prior for the fit parameters is used. For this model, this takes the form:

$$p(A, B|I) = \begin{cases} \mathcal{N} & \text{if } 0 < A \leq 1.04225 \cdot 10^{93} \wedge -0.1 \leq B \leq 0.1 \\ 0 & \text{else,} \end{cases} \tag{C.7}$$

where the normalization factor $\mathcal{N}$ can be chosen to be the reciprocal of the (two dimensional) range of the parameters:

$$\mathcal{N} = \frac{1}{[1.04225 \cdot 10^{93}] \cdot [0.2]}. \tag{C.8}$$

In this way the prior is defined in a much proper way, but for calculational convenience (as it is normalized), the normalization constant is chosen to be 1.

C.1.4 Plot and fit of the data

To determine the values of the fit parameter that gives the smallest residue, a least squares fit is done with Python. The result is shown in Figure C.1 below:

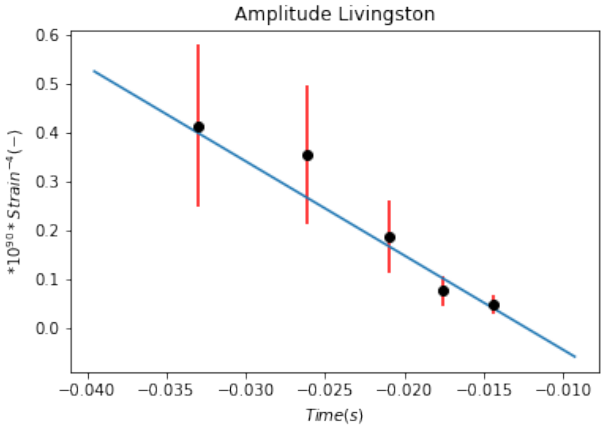


Figure C.1: Dimensionless strain measured at the detector in Livingston plotted against the time. The red bars denote the error bars, whereas the blue line represents the least squares fit line. Here a uniform prior for the fit parameters is used.

In order to determine the value of the parameter A as precise as possible, the y coordinate of the data is multiplied with a factor of 10^{-90}. This causes the y coordinate of the data to be in the order of 10^{-1}. By looking at the fit function in Equation C.1, it can be seen that the parameter A is scaled by the same factor, such that this parameter has the same order of magnitude as the data. The parameter B is unchanged, because all the multiplication is stored in this parameter A due to this specific form of the fit function.

C.1.5 The values of the model and fit parameters

The fit parameters that minimize the residue are determined with the least square method to have the following values for the detector located at Livingston:

$$A = (19.3 \pm 3.8) \cdot 10^{90} s^{-1} \tag{C.9a}$$

$$B = -0.013 \pm 0.003\text{s}, \tag{C.9b}$$

where the uncertainties are also dermined by the Python least square package. The most likely values of the model parameters are calculated from the values of the fit parameters to be:

$$t_0 = -0.013 \pm 0.003\text{s} \tag{C.10a}$$

$$D_{\text{eff}} = (3.0 \pm 0.1) \cdot 10^{26} m \tag{C.10b}$$

where the uncertainties are calculated by the 68% interval calculation rules described in Appendix A.2.

C.1.6 Likelihood function for the fit parameters; uniform model parameter prior

Again, the Python Corner package is used to calculate a contour plot of the likelihood function for A and B, and to determine the marginal probability density functions for both of them separately. The values determined with the least square method are plotted as the blue lines. The result is shown in Figure C.2 below:

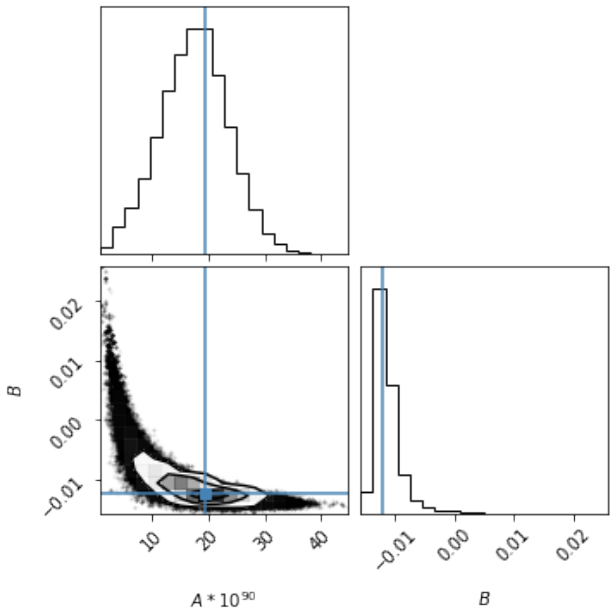


Figure C.2: Likelihood function and Marginal probability density function for the fit parameters for the detector in Livingston. The blue dots represent the most likely values of the fit parameters. A non uniform prior for the fit parameters is used.

C.2 Determining the effective physical distance for Hanford

C.2.1 Fit function

The fit function is of course exactly the same as for the determination of the effective physical distance for Livingston. See Equation C.1.

C.2.2 Relating fit parameters to model parameters

Because the same fit function and the same theoretical model are used, the relation between fit and model parameters is exactly equal to the case for Livingston, as well as the form of the Jacobian.

C.2.3 Prior for model and fit parameters

Also the same prior for the model and fit parameters is used. Again, the reason for this is the same fit function and the same parameters that are determined (D_{eff} and t_0).

C.2.4 Plot and fit of the data

For the detector in Hanford, the data slightly differs from the data measured by the detector in Livingston. The 1/4 behaviour should be present, but the effective distance is slightly different. This is due to the fact that the effective distance depends on the real physical distance from source to detector and it depends on the antenna response factors. These factors have different values for each detector, because each detector is located on a different position on Earth and may be oriented in a certain way. These differences between detectors are stored in these antenna response factors. The strain to the power -4 is plotted against the time can be seen in Figure C.3 below.

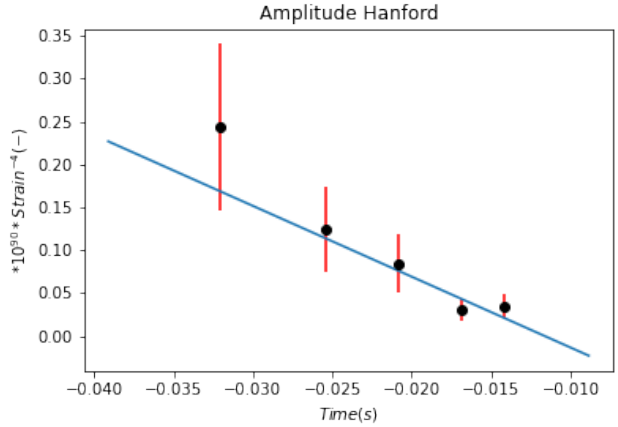


Figure C.3: Dimensionless strain measured at the detector in Hanford plotted against the time. The red bars denote the error bars, whereas the blue line represents the least squares fit line. Here a uniform prior for the fit parameters is used.

C.2.5 The values of the model and fit parameters

With the least squares package of Python, the values for the fit parameters A and B are determined to be:

$$A = (8.2 \pm 3.0) \cdot 10^{90} s^{-1} \tag{C.11a}$$

$$B = (-0.011 \pm 0.002) s. \tag{C.11b}$$

With these values of the fit parameters, the values for the model parameters that maximize the likelihood function are calculated via Equation C.3:

$$D_{\text{eff}} = (2.4 \pm 0.2) \cdot 10^{26} m \tag{C.12a}$$

$$t_0 = (-0.012 \pm 0.002) s \tag{C.12b}$$

C.2.6 Likelihood function for the fit parameters; uniform model parameter prior

With help of the Python 'corner package', the likelihood function as well as the marginal probability density function are determined. This is done via the recipes in Chapter 5. See Figure C.4 below:

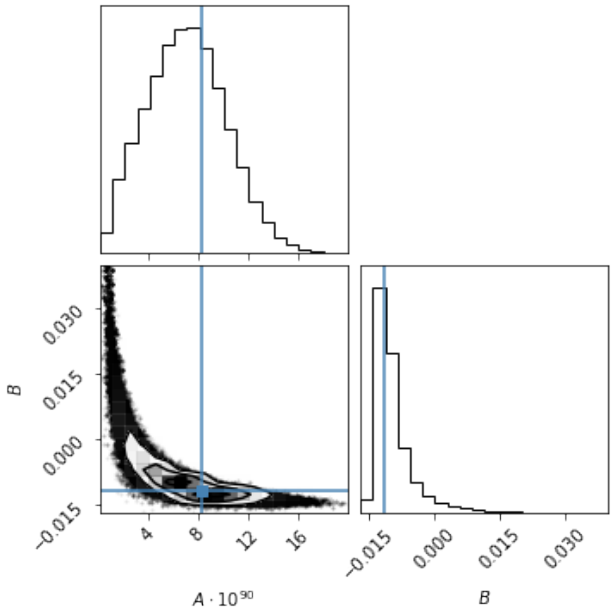


Figure C.4: Likelihood function and Marginal probability density function for the fit parameters for the detector in Hanford. The blue dots represent the most likely values of the fit parameters. A non uniform prior for the fit parameters is used.