

Generalization of the Fast Chirp Transform Algorithm

A Report of the FCT Group

F.A. Jenet, Philip Charlton, J. Edlund, L. Wen,
T.D. Creighton, T.A. Prince, and Massimo Tinto

(Dated: 15 May 2003)

Abstract

The Fast-Chirp Transform (FCT) algorithm, derived by Jenet and Prince [1] for detecting sinusoidal signals of monotonically increasing phase functions of time, is here generalized to encompass signals of arbitrary phase functions. One other important aspect of the FCT algorithm here derived is that the “grid size” in the space of parameters characterizing the phase function itself can be rescaled at will. This feature is essential for implementing hierarchical searches of signals characterized by a set of parameters whose values are unknown. After showing through an example that the computational cost required to perform a search with the FCT is essentially equal to the time required to perform an equal number of fast Fourier transforms, we analyze as an example of its applications the detection of first-Post-Newtonian gravitational wave signals emitted during the coalescence of binary systems containing neutron stars. We find that the recovered signal-to-noise ratio (SNR) in this case is the same as that achievable by implementing Matched Filtering.

PACS numbers: 04.80.Nn, 95.55.Ym, and 07.60.Ly

I. INTRODUCTION

Kilometer-size ground-based interferometric detectors of gravitational radiation are in the process of becoming fully operational at several laboratories around the world. In the next few years, from locations in the United States of America, Italy, Germany, and Japan, these instruments will attempt to detect and study, in the kilohertz frequency band, gravitational waves emitted by astrophysical sources such as spinning neutron stars, supernovae, and coalescing binary systems [2], [3], [4], [5].

Among the most likely sources of gravitational radiation that ground-based interferometer detectors of gravitational waves will search for are coalescing binary systems containing neutron stars and/or black holes. The gravitational waves emitted by these sources have a unique signature that enables them to be extracted from the wide-band data from these detectors by digital filtering techniques [6]. This signature is their accelerating sweep upwards in frequency as the binary orbit decays because of energy loss due to emission of gravitational radiation. Coalescing binaries have the advantage over other sources in signal-to-noise ratio (SNR) by a factor that depends on the square-root of the ratio between the corresponding number of cycles in the wave trains [7].

The standard technique used for extracting these “chirps” from the noisy data is called Matched Filtering (MF). In the presence of colored noise, represented by a random process $n(t)$, the noise-weighted inner product between two real functions u and v is defined as

$$\langle u, v \rangle := 2 \operatorname{Re} \int_0^\infty \frac{\tilde{u}(f) \tilde{v}^*(f)}{S(f)} df, \quad (1)$$

where the symbol $\tilde{}$ over u and v denotes their Fourier transform, $S(f)$ is the one-sided power spectral density of the noise $n(t)$, and the $*$ represents complex conjugation. From this definition, the expression for the SNR, ρ^2 can be written as follows [6]

$$\rho^2 := \frac{\langle u, h \rangle^2}{\operatorname{Var} \langle n, h \rangle}, \quad (2)$$

where $h(t)$ is a gravitational wave signal in the data stream $u(t)$ recorded by the detector. By analyzing the statistics of the SNR ρ^2 , it is possible to statistically infer the presence of a gravitational wave signal in the data. This operation of course needs to be performed over the entire bank of templates over which the SNR statistics is built upon, since a gravitational wave signal is in principle determined by a (finite) set of continuous parameters.

Let us suppose that $h(t; \hat{\tau})$ is a family of signal templates parameterised by $\hat{\tau} \equiv \{\tau_0, \tau_1, \dots, \tau_P\}$, with $h(t; \hat{\tau}) = 0$ for $t < 0$ and $t > T$. A signal with the same shape but commencing at time t_0 is given by $h(t - t_0; \hat{\tau})$, thus the expression for the SNR as a function of arrival time t_0 becomes

$$\rho^2(t_0, \hat{\tau}) = \frac{1}{\sigma^2(\hat{\tau})} \text{Re} \int_0^\infty \frac{\tilde{u}(f) \tilde{h}^*(f; \hat{\tau}) e^{-i2\pi t_0 f}}{S(f)} df \quad (3)$$

where $\sigma^2(\hat{\tau}) = \text{Var} \langle n, h(\hat{\tau}) \rangle$. Rigorous procedures for discretizing the set of continuous parameters characterizing gravitational wave signals from binary systems have been presented elsewhere in the literature [8], [9], and we refer the interested reader to those publications. Here we focus instead on a different algorithm for calculating the expression of the SNR over the bank of templates. This algorithm, which has been called the Fast Chirp Transform (FCT) [1], provides an alternative way of calculating the expression of the SNR given by equation (3). This paper will focus on aspects of the FCT algorithm not previously considered in the literature, and it will be organized as follows.

In section II, after presenting a brief summary of the FCT algorithm [1], we give an alternative derivation of it, one that does not rely on any assumptions about the behavior of the phase function. We use the 2-dimensional case as the simplest case for deriving the FCT for sinusoidal signals of arbitrary phase function, and use it for estimating the error introduced in computing the expression of the SNR by implementing the FCT algorithm. We then address the determination of the grid-size within the parameter space the FCT operates on, and show that it can be rescaled at will. Hierarchical searches of chirping gravitational wave signals present in the data from interferometric detectors could therefore be performed with the FCT. After generalizing the FCT algorithm to n -dimensions, we estimate its computational cost and find it to be equal to that of the Fast Fourier Transform algorithm when 10^6 of more data points are processed. In Section III we then estimate the SNR degradation introduced by various inherent approximations present within the FCT algorithm and find it to be negligibly small. As an example application we show how the FCT can be used for searching first Post-Newtonian gravitational wave signals from coalescing binary systems.

II. THE FAST CHIRP TRANSFORM

The previous paper, [1], introduced the computational technique called the “Fast Chirp Transform” (FCT). The FCT was developed to analyze signals of deterministically varying frequency (also known as “chirping signals”) by efficiently evaluating sums of the following form:

$$Y(k_0, k_1, \dots, k_n) = \sum_{j=0}^{N_0-1} x_j e^{i\Phi(j, k_0, k_1, \dots, k_n)}, \quad (4)$$

with N_0 being the number of data points. Hence, the FCT convolves the data, x_j , with a chirping waveform. The waveform is determined by specifying the phase function, $\Phi(j, k_0, k_1, \dots, k_n)$, which was assumed to be of the form:

$$\Phi(j, k_0, k_1, \dots, k_n) = 2\pi \sum_{l=1}^{n-1} k_l \phi_l(j) + \frac{2\pi k_0 j}{N_0}. \quad (5)$$

A nice property of the FCT algorithm is that one only needs to supply the phase function determining the frequency evolution of the signal, without defining and calculating a set of templates. Since it has been shown [1] that the FCT algorithm is computationally as fast as matched filtering, it should be regarded as a “black box” analysis tool, analogous to the role played by the FFT for the analysis of signals of constant frequency. The user need only input the phase function and the data.

The previous derivation of the FCT divided the data into varying length data segments, where the length of each data segment was determined by the condition that each of the non-linear components of the phase function varied by less than π over the duration of that data segment. Hence, the non-linear phase terms could be treated as constant within each segment. The sum in equation (4) was transformed into n nested sums, with each partial sum evaluated using the Fast Fourier Transform.

The accuracy of the FCT is determined by the validity of the assumption that the non-linear phase terms are constant within each data segment. As the values of the conjugate parameters, k_1, \dots, k_n , increase, this assumption becomes less and less accurate. Hence, the accuracy of the FCT will roll-off with increasing conjugate parameter value, much in the same way the accuracy of the FFT rolls-off with increasing frequency.

Without loss of generality, we will restrict the range of the phase functions, $\phi_i(j)$, to the interval $[0, 1]$, and they will be normalized to have a maximum value of 1. The 2-D transform will be presented first, followed by a generalization to higher dimensions.

A. The 2-D transform

The FCT will first be derived for the 2-dimensional case. Consider the continuous form of equation (4):

$$Y(\omega_0, \omega_1) = \int x(t) e^{i\omega_1 \phi_1(t) + i\omega_0 t} dt. \quad (6)$$

This integral may be written as a two dimensional integral using a delta function:

$$Y(\omega_0, \omega_1) = \int \int x(t) \delta(t' - \phi_1(t)) e^{i\omega_1 t' + i\omega_0 t} dt dt'. \quad (7)$$

The above double integral is simply a Fourier transform of a two dimensional distribution, $F(x(t), t, t')$, given by

$$F(x(t), t, t') = x(t) \delta(t' - \phi_1(t)). \quad (8)$$

$F(x(t), t, t')$ is called an “embedding” of the one dimensional function $x(t)$ in 2-dimensions. The two dimensional distribution may be viewed as a 2-D field of numbers that are zero everywhere except on a one dimensional trajectory defined by:

$$\begin{pmatrix} t \\ t' \end{pmatrix} = \begin{pmatrix} t \\ \phi_1(t) \end{pmatrix} \quad (9)$$

Figure ?? shows a graphical representation of this embedding for the case of $\phi_1(t) = t^2$ and $x(t) = 1$. Given any arbitrary phase function, an appropriate embedding can be constructed that will transform equation (6) into a 2-D Fourier transform.

A similar approach may be used for the 2-D discrete transform:

$$Y(k_0, k_1) = \sum_{j=0}^{N_0-1} x_j e^{i2\pi k_1 \phi_1(j) + i\frac{2\pi k_0 j}{N_0}}. \quad (10)$$

One main difference between the continuous and discrete cases is that for the continuous case, the 2-D Fourier transform of the embedded function is exactly equal to the original integral whereas for the discrete case, the 2-D FCT will only be an approximation to the original sum under consideration. Taking a similar approach to that followed in the continuous case, the general 2-parameter chirp transform for discrete data is defined as

$$C(k_0, k_1) = \sum_{j_1=0}^{N_1-1} \sum_{j_0=0}^{N_0-1} F_d(x_{j_0}, j_0, j_1) e^{i\frac{2\pi j_0 k_0}{N_0} + i\frac{2\pi j_1 k_1}{N_1}} \quad (11)$$

where $F_d(x_{j_0}, j_0, j_1)$ is a discrete embedding of the function x_{j_0} in the 2-dimensional discrete plane. It can be shown that $C(k_0, k_1)$ is equivalent to the FCT derived by Jenet and Prince [1] when

$$F_d(x_{j_0}, j_0, j_1) = x_{j_0} H\left(\frac{1}{2} - |j_1 - N_1 \phi_1(j_0)|\right) \quad (12)$$

where $H(x)$ is the Heaviside step function defined as

$$H(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0. \end{cases} \quad (13)$$

Although other discrete embeddings may be chosen, which may have better properties for specific applications, we will not explore them further in this paper.

Equation (11) shows that the 2-D discrete FCT may be evaluated using a 2-D FFT of the embedding F_d . For any given phase function, $\phi(t)$, the embedding defined by equation (12) gives a well defined way of packing the data into a 2-D array to prepare it for transformation by an FFT algorithm. Figure ??b shows a $N_0 \times N_1$ grid where the ordinate represents j_0 and the abscissa represents j_1 . The darkened squares represent the regions where F_d is non-zero. A phase function of the form $\phi(t) = t(t - t_0)$ was used in this example. In §II B we show that $C(k_0, k_1) \approx Y(k_0, k_1)$ thus validating the use of the FCT to calculate sums of the form (10).

B. Accuracy of the approximation

In this section we show that $C(k_0, k_1)$ is an approximation to the exact chirp transform $Y(k_0, k_1)$ when equation (12) defines the embedding. For this embedding, $F_d(x_{j_0}, j_0, j_1)$ is non-zero only when

$$|j_1 - N_1 \phi_1(j_0)| < \frac{1}{2}. \quad (14)$$

Hence, for a given integer j_0 , there is only one integer j_1 that satisfies the above constraint. Thus we may write

$$j_1 = N_1 \phi_1(j_0) + \Delta(j_0) \quad (15)$$

where $|\Delta(j_0)| < 1/2$. Since F_d is non-zero for only one value of j_1 given a value of j_0 , the j_1 sum in equation (11) becomes

$$C(k_0, k_1) = \sum_{j_0=0}^{N_0-1} x_{j_0} e^{i \frac{2\pi k_0 j_0}{N_0} + i 2\pi k_1 \phi_1(j_0) + i \frac{2\pi \Delta(j_0) k_1}{N_1}}. \quad (16)$$

Since $|\Delta(j_0)| < 1/2$, the difference between the terms in the exponent (i.e. the phase) of the above expression and that of equation (10) is always less than π (this is the same constraint used for the original derivation of the FCT in [1]). We conclude that $C(k_0, k_1)$ is approximately equal to $Y(k_0, k_1)$.

To quantify the approximation we will determine a bound on the difference $|Y(k_0, k_1) - C(k_0, k_1)|$. From Eqs. (10) and (16) we see that

$$\begin{aligned} |Y(k_0, k_1) - C(k_0, k_1)| &= \left| \sum_{j=0}^{N_0-1} x_j e^{i2\pi(k_0 j/N_0 + k_1 \phi_1(j))} (1 - e^{i2\pi\Delta(j)k_1/N_1}) \right| \\ &\leq \sum_{j=0}^{N_0-1} |x_j| |1 - e^{i2\pi\Delta(j)k_1/N_1}|. \end{aligned} \quad (17)$$

Now for any real x ,

$$|1 - e^{ix}|^2 = 2(1 - \cos x) \quad (18)$$

and since $\cos x \geq 1 - x^2/2$ for all x we have the inequality

$$|1 - e^{ix}| \leq \begin{cases} |x| & -2 \leq x \leq 2 \\ 2 & \text{otherwise.} \end{cases} \quad (19)$$

As $|\Delta(j)| \leq 1/2$ for all j it follows that

$$|Y(k_0, k_1) - C(k_0, k_1)| \leq \frac{\pi k_1}{N_1} \sum_{j=0}^{N_0-1} |x_j| \quad \text{for } k_1 \leq 2N_1/\pi \quad (20)$$

$$\leq 2 \sum_{j=0}^{N_0-1} |x_j| \quad \text{for } k_1 > 2N_1/\pi. \quad (21)$$

Note that the inequality (21) would still hold even if $\Delta(j)$ in (17) were replaced with an arbitrary phase difference, so $C(k_0, k_1)$ cannot be regarded as a good approximation to $Y(k_0, k_1)$ once k_1 exceeds $2N_1/\pi$.

C. Parameter spacing

The derivation of the FCT in §II A assumes the range of the phase function ϕ_1 to be within the interval $[0, 1]$, and each value $\phi_1(j)$ is replaced by a nearby value j_1/N_1 , with j_1 satisfying equation (14). For the FCT to be on the same footing as the FFT, the chirp parameter k_1

must be an appropriate integer ie. $k_1 = 0, 1, \dots, N_1 - 1$. In practical applications a different spacing for the chirp parameter is often required. In the case of the Fourier transform, the natural size of each frequency bin is $1/T$ Hz, where T is the length of the data in seconds. A finer spacing of frequency bins can be obtained, for example, by extending the data with zeroes (“oversampling”), while a coarser spacing can be obtained by either decimating or averaging over adjacent frequency bins.

In a similar fashion, by rescaling the phase function it is straightforward to adapt the FCT for a given resolution of chirp parameters. Let us assume that the phase function is bounded and non-negative, but otherwise unrestricted. Suppose that the desired spacing of the chirp parameter is $\Delta\lambda$, so that

$$\lambda_{k_1} = k_1\Delta\lambda \quad \text{for } k_1 = 0, 1, \dots, N_1 - 1. \quad (22)$$

To calculate the FCT for this spacing and phase function, we simply replace the combination $\lambda\phi_1$ by $k_1\hat{\phi}_1$ where $\hat{\phi}_1 = (\Delta\lambda/M)\phi_1$, and calculate the FCT with respect to $\hat{\phi}_1$. The additional scaling factor M is chosen so as to be the smallest positive integer which ensures that $\max \hat{\phi}_1 \leq 1$. Values of the FCT at the spacing $\Delta\lambda$ are obtained by only calculating every M th element in the k_1 direction – that is, $C(k_0, k_1M)$ is the FCT value corresponding to chirp parameter $k_1\Delta\lambda$.

In cases where M turns out to be 1 it is easy to see why this is analogous to oversampling. This will happen when $\Delta\lambda$ and ϕ_1 are such that $\max \Delta\lambda\phi_1$ is already less than 1. In that case the discretisation of $\hat{\phi}_1$ (as in Figure ??) could only have non-zero values for $0 \leq k_1 < N_1 \max \hat{\phi}_1$ and all zeroes when $k_1 \geq N_1 \max \hat{\phi}_1$, exactly as if oversampling had been applied in the k_1 direction.

D. The n -dimensional transform

In the previous section, the concept of embedding a one dimensional function into a two dimensional space was introduced in order to transform the integral to be evaluated into a Fourier transform. The concept may be easily generalized into n -dimensions. For the continuous case, the integral:

$$Y(\omega_0, \omega_1, \dots, \omega_n) = \int x(t) e^{i\omega_n\phi_n(t) + \dots + i\omega_1\phi_1(t) + i\omega_0 t} \quad (23)$$

may be transformed into an n -dimensional Fourier transform of the form:

$$Y(\omega_0, \omega_1, \dots, \omega_n) = \int F(x(t), t, t_1, \dots, t_n) e^{i\omega_n t_n + \dots + i\omega_1 t_1 + i\omega_0 t} \quad (24)$$

where F is an n -dimensional embedding defined by:

$$F(x(t), t, t_1, \dots, t_n) = x(t) \delta(t_1 - \phi_1(t)) \times \delta(t_2 - \phi_2(t)) \times \dots \times \delta(t_n - \phi_n(t)). \quad (25)$$

Analogous to the 2-D case, the n -dimensional embedding may be viewed as a n -dimensional field that is zero everywhere except on a curve defined by:

$$\begin{pmatrix} t \\ t_1 \\ \dots \\ t_n \end{pmatrix} = \begin{pmatrix} t \\ \phi_1(t) \\ \dots \\ \phi_n(t) \end{pmatrix}. \quad (26)$$

The discrete case may be constructed just as easily using a discrete embedding of the form:

$$F_d(x(j_0), j_0, j_1, \dots, j_n) = x_{j_0} H\left(\frac{1}{2} - |j_1 - N_1 \phi_1(j_0)|\right) \times \dots \times H\left(\frac{1}{2} - |j_n - N_n \phi_n(j_0)|\right). \quad (27)$$

Hence, the n -dimensional FCT is a discrete Fourier transform of an n -dimensional array of numbers defined by the above discrete embedding. The set of points defined by F_d is mostly zero except for a small subset that lie near the 1-dimensional trajectory defined by equation (26).

E. Algorithm complexity and parallel implementation

Several implementations of the general FCT algorithm have been developed. The most direct implementation is to simply pack the data into an n -dimensional array as in equation (27) and perform a full n -dimensional FFT of $F_d(x(j_0), j_0, j_1, \dots, j_n)$. This requires $O(N_0 \times \dots \times N_n \log_2 N_0 \times \dots \times N_n)$ operations. However, the fact that the data is arranged so that it follows a one-dimensional “track” allows for a considerable computational saving. Because of the FCT packing scheme, there is only a single non-zero element in each $(n - 1)$ -dimensional matrix obtained by fixing j_0 . The sparsity of this matrix allows the FCT to be calculated

in $O(N_1 \times \dots \times N_n)$ operations for each j_0 , followed by a 1-dimensional FFT for each value of j_1, j_2, \dots, j_n . This reduces the dominant cost of the FCT to be $O(N_0 \times \dots \times N_n \log_2 N_0)$ operations.

An implementation of this method has been developed for the LIGO Analysis Library (LAL) [10]. In this implementation, the FCT is evaluated for a given range of chirp parameters in a single function. Figure ?? shows, as an example, the computational time required to evaluate a 2-D FCT as a function of N_0 (N_1 has been fixed to be equal to 32) relative to the time needed to calculate 32 fast Fourier transforms of length N_0 . Since this ratio approaches unity for a large number of data samples, N_0 , the graph confirms that the performance of our FCT implementation is dominated by a term of $O(N_1 N_0 \log_2 N_0)$, and shows that the time to perform the FCT is essentially the time required to perform fast Fourier transforms. Note that the deviation from this complexity at smaller values of N_0 is due to the fact that the FCT algorithm complexity has additional terms of $O(N_1 N_0)$. The tests were performed on a 2.2 GHz Intel Xeon machine with 2GB of memory running Linux RedHat 9.0. The FCT implementation uses FFTW [11] both internally and for the purposes of this comparison.

While the function itself has not been parallelized, it is a simple matter to parallelize a large FCT calculation by breaking up the space of parameters into smaller regions and evaluate the FCTs of each of them separately on the nodes of a parallel machine. This approach has been implemented in the LIGO Data Analysis System (LDAS) [10] using a Beowulf cluster.

F. The point-spread of the chirp transform

In this section we derive the approximate response of the FCT to a chirping signal of the following form

$$S(x) = \exp \left\{ -i\beta_0 x - i \sum_{k=1}^{n-1} \beta_k \phi_k(x) \right\} \quad (28)$$

where the β_k 's are the chirp parameters, n is the total number of chirp parameters, $\phi_k(x)$ is the phase function corresponding to the k th chirp parameter, and x is a continuous parameter which typically represents time or frequency. The chirp transform of $S(x)$ is

defined as follows:

$$H(\tau_0, \dots, \tau_{n-1}) = \int_{X_{min}}^{X_{max}} S(x) \exp \left\{ i\tau_0 x + i \sum_{k=1}^{n-1} \tau_k \phi_k(x) \right\} dx. \quad (29)$$

$H(\tau_0, \dots, \tau_{n-1})$ is peaked at $\tau_k = \beta_k$ for all k . From now on, the chirp parameters will be denoted as $\hat{\tau}$ and/or $\hat{\beta}$ when referred to collectively. The goal of this section is to show how the chirp transform “power”, $|H(\hat{\tau})|^2$ behaves when $\hat{\tau} \neq \hat{\beta}$. Figure ?? shows a grey-scale plot and a surface plot of a two dimensional chirp transform, $|H(\tau_0, \tau_1)|^2$, with $\phi_1(x) = x^{2.5}$. τ_0 and τ_1 are in units of $1/X_{max}$ and $1/X_{max}^2$ respectively. In these plots, the origin corresponds to the location where $\hat{\tau} = \hat{\beta}$.

These plots show that most of the power is located within a restricted “bow-tie” region of the (τ_0, τ_1) plane. Within this region, the amplitude varies as a strong function of $\hat{\tau}$. In the next section, it is shown that the bounded region is a result of the finite limits of integration in equation (29). Analytical expressions for the boundaries of the bow-tie region are given for the case of the two dimensional chirp transform. In §2.8, an approximate expression for $|H(\hat{\tau})|^2$ is derived for the general n -dimensional chirp transform.

G. Power boundaries

Understanding the response of the chirp transform to a pure chirp as defined in the introduction may be simplified by noting that the behavior of $H(\hat{\tau})$ does not depend on $\hat{\beta}$. Varying $\hat{\beta}$ only changes the location of the maximum. Hence, $\hat{\beta}$ is set to 0 for the remaining discussion. In this case,

$$H(\hat{\tau}) = \int_{X_{min}}^{X_{max}} \exp \left\{ i\tau_0 x + i \sum_{k=1}^{n-1} \tau_k \phi_k(x) \right\} dx. \quad (30)$$

This integral happens to be the finite time Fourier transform of the following signal:

$$F(x) = \exp \{ i\Phi(\tau_1, \dots, \tau_{n-1}, x) \} \quad (31)$$

$$\Phi(\tau_1, \dots, \tau_{n-1}, x) = \sum_{k=1}^{n-1} \tau_k \phi_k(x). \quad (32)$$

$F(x)$ is a signal with unit amplitude and varying frequency. The instantaneous frequency, $f(x)$, is given by

$$f(x) = -\Phi'(\tau_1, \dots, \tau_{n-1}, x) \quad (33)$$

where the prime denotes the derivative with respect to x . For this discussion it is assumed that $\Phi'(\tau_1, \dots, \tau_{n-1}, x)$ has a single minimum and a single maximum over the interval $[X_{\min}, X_{\max}]$. Thus, the instantaneous frequency will sweep from $-\Phi'_{\max}$ to $-\Phi'_{\min}$ over the interval. Since τ_0 corresponds to the frequency parameter in equation (30), most of the power is expected to lie within the region defined by $\tau_0^{\min} \leq \tau_0 \leq \tau_0^{\max}$ where

$$\tau_0^{\min} = -\Phi'(\tau_1, \dots, \tau_{n-1}, x)_{\min} \quad (34)$$

$$\tau_0^{\max} = -\Phi'(\tau_1, \dots, \tau_{n-1}, x)_{\max}. \quad (35)$$

For the case of a two dimensional chirp transform, the following two boundary lines are obtained:

$$\tau_0^{\min} = -\tau_1 \phi'_{1\min} \quad (36)$$

$$\tau_0^{\max} = -\tau_1 \phi'_{1\max}. \quad (37)$$

The top panel in figure ?? shows the same grey-scale plot of $|H(\tau_0, \tau_1)|^2$ as in figure ?? but with the analytical boundaries plotted as well. This figure shows that most of the power does lie within the region defined above.

H. Amplitude variations

An approximate expression for $|H(\hat{\tau})|^2$ is derived in this subsection for the general n -dimensional chirp transform when $\hat{\tau} \neq 0$. It was found in the previous section that the majority of the power is located within a set of surfaces defined by equations (34) and (35). In order to calculate an approximate form for $|H|^2$, the limits of integration in equation (30) will be extended to infinity, $[-\infty, \infty]$, and the integral will be evaluated using the stationary phase approximation (see the Appendix of [1]). Using this prescription, the approximate power, $P(\hat{\tau})$, will be given by

$$P(\hat{\tau}) = \left| \frac{\pi/2}{d^2\Phi(\hat{\tau}, x)/dx^2} \right|, \quad (38)$$

when $\hat{\tau}$ is within the boundary surface defined by equations (34) and (35). Outside of this region, $P(\hat{\tau})$ is set to zero. The variable x in equation (38) may be eliminated using the following relationship:

$$\tau_0 = -\Phi'(\tau_1, \dots, \tau_{n-1}, x). \quad (39)$$

Note that equation (38) is valid only when $d^2\Phi/dx^2$ is non-zero. For the case when $\hat{\tau} \neq 0$ and $d^2\Phi/dx^2 = 0$, higher order derivatives must be used. The necessary modifications are straight forward and will not be discussed further here.

For the case of a two dimensional chirp transform with $\phi_1(x) = x^\eta$, it can be shown using equations (38) and (39) that

$$P(\tau_0, \tau_1) = \left| \frac{A}{\eta(\eta-1)\tau_1^{\frac{1}{\eta-1}} \frac{\tau_0}{\eta} \frac{\eta-2}{\eta-1}} \right| \quad (40)$$

The bottom panel of figure ?? shows a surface plot of $|H(\tau_0, \tau_1)|^2/P(\tau_0, \tau_1)$ for the case of $\phi_1(x) = x^{2.5}$. Notice that the large scale power variations have been removed, leaving only the small scale fluctuations due to finite interval “ring-down” and interference.

III. OPTIMAL FILTERING AND THE FCT

Consider a pair of real vectors u, v of length $M = 2N_0$. The Euclidean inner product $\langle u, v \rangle = \sum_{m=0}^{M-1} u_m v_m$ can be written in terms of the DFTs \tilde{u}, \tilde{v} as

$$\begin{aligned} \langle u, v \rangle &= \frac{1}{2N_0} \sum_{k=0}^{2N_0-1} \tilde{u}_k^* \tilde{v}_k \\ &= \frac{1}{N_0} \text{Re} \sum_{k=0}^{N_0-1} \tilde{u}_k^* \tilde{v}_k + \frac{1}{2N_0} (\tilde{u}_{N_0}^* \tilde{v}_{N_0} - \tilde{u}_0^* \tilde{v}_0). \end{aligned} \quad (41)$$

For detection of a signal in discretely-sampled noisy data, it is usual to employ a *noise-weighted* inner product. By analogy with (41), the inner product weighted by the power spectral density S_n of Gaussian noise n is defined by

$$\langle u, v \rangle = \frac{1}{N_0} \text{Re} \sum_{k=0}^{N_0-1} \frac{\tilde{u}_k^* \tilde{v}_k}{S_n}, \quad (42)$$

where we have assumed that the data has been pre-processed so that the DC and Nyquist components of u and v are negligible.

Suppose that h is a “template” for a signal we want to detect, normalized with respect to $\langle \cdot, \cdot \rangle$. If $u = n + \mathcal{A}h$ represents a data stream consisting of noise n and additive signal h with amplitude \mathcal{A} , the SNR, ρ^2 , is given by

$$\begin{aligned} \rho^2 &= \frac{\langle u, h \rangle^2}{\text{Var} \langle n, h \rangle} \\ &= \frac{\langle n + \mathcal{A}h, h \rangle^2}{\text{Var} \langle n, h \rangle} \end{aligned} \quad (43)$$

where the denominator is obtained over many realizations of n . Thus the expectation of ρ is

$$\begin{aligned} E[\rho^2] &= \frac{E[\langle n, h \rangle^2 + 2\mathcal{A}\langle n, h \rangle \langle h, h \rangle + \mathcal{A}^2 \langle h, h \rangle^2]}{\text{Var} \langle n, h \rangle} \\ &= 1 + \frac{\mathcal{A}^2}{\text{Var} \langle n, h \rangle}. \end{aligned} \quad (44)$$

Now if the second h in (43) is replaced by a different normalized template h' , the expectation obtained would instead be $1 + \mathcal{A}^2 \langle h, h' \rangle^2 / \text{Var} \langle n, h \rangle$. Since $\langle h, h' \rangle^2 \leq 1$ we see that using a mismatched template results in a reduction of the average SNR obtained. For signals present with large SNR, we can interpret $\langle h, h' \rangle^2$ as the fractional loss of SNR due to using the mismatched template h' to detect h .

The expression (43) for the SNR assumes that the template h always starts at the same time. In practise, however, the start-time of a signal is usually unknown and we need to calculate the inner product for all possible start-times of h – that is, we need to calculate the correlation of h and u as a function of time offset. In the frequency domain this can be achieved efficiently using the fast Fourier transform.

Let h^k denote the vector obtained by shifting h to the right by k samples, that is,

$$h_m^k = h_{m-k}, \quad m = 0, 1, 2, \dots, M - 1. \quad (45)$$

Then the components of the Fourier transforms of h^k and h are related via

$$\tilde{h}_m^k = \tilde{h}_m e^{i2\pi km/M}. \quad (46)$$

Then we see that the inner product for each offset is given by

$$\langle u, h^{2k} \rangle = \frac{1}{N_0} \text{Re} \sum_{m=0}^{N_0-1} \frac{\tilde{u}_m^* \tilde{h}_m}{S_n} e^{i2\pi km/N_0} \quad (47)$$

which can be calculated using an inverse FFT of length N_0 . If we now suppose that the template \tilde{h} has the form of a chirp, that is,

$$\tilde{h}_m(\hat{k}) = A_m(\hat{k}) e^{i2\pi \sum_{j=1}^P k_j \phi_j(m)} \quad (48)$$

then the summation in (47) (and hence the SNR) can be calculated via a P -dimensional FCT, where the vector being transformed is

$$\tilde{u}'_m = \frac{\tilde{u}_m^* A_m(\hat{k})}{S_n}. \quad (49)$$

A. Loss of SNR due to FCT approximations

One way of viewing the FCT is that we are replacing an “exact” template (48) with an approximate template, where the approximation is achieved by substituting a nearby rational number for the actual value of the phase at each point. With that in mind we can quantify the loss of SNR incurred by using the FCT instead of an exact template.

First consider the simplest case of a signal with one chirp parameter, so that the exact template is of the form (48) with $P = 1$. Examining Eq. (16), the corresponding FCT template would be

$$\tilde{h}_j^{\text{FCT}}(k_1) = A_j(k_1)e^{i2\pi k_1(\phi_1(j)+\Delta(j)/N_1)}. \quad (50)$$

Note that neither template is normalized, although it is clear that both templates have the same norm.

If a signal were present with precisely the chirp parameter k_1 , we would recover the maximum achievable SNR using the MF template. The FCT template, being slightly mismatched, will recover less of the SNR. The inner product of the two templates has a lower bound given by

$$\begin{aligned} \langle h(k_1), h^{\text{FCT}}(k_1) \rangle &= \frac{1}{N_0} \text{Re} \sum_{m=0}^{N_0-1} \frac{A_m^2(k_1)}{S_n} e^{i2\pi k_1 \Delta(m)/N_1} \\ &= \frac{1}{N_0} \sum_{m=0}^{N_0-1} \frac{A_m^2(k_1)}{S_n} \cos(2\pi k_1 \Delta(m)/N_1) \\ &\geq \frac{1}{N_0} \sum_{m=0}^{N_0-1} \frac{A_m^2(k_1)}{S_n} \left(1 - \frac{1}{2} \left[\frac{2\pi k_1 \Delta(m)}{N_1} \right]^2 \right) \\ &\geq \left(1 - \frac{\pi^2}{2} \left[\frac{k_1}{N_1} \right]^2 \right) \frac{1}{N_0} \sum_{m=0}^{N_0-1} \frac{A_m^2(k_1)}{S_n} \\ &= \left(1 - \frac{\pi^2}{2} \left[\frac{k_1}{N_1} \right]^2 \right) \|h(k_1)\|^2. \end{aligned} \quad (51)$$

This argument can be easily applied to an FCT with P phase functions to show that the fraction of SNR recovered is at least

$$\left(1 - \frac{\pi^2}{2} \left[\frac{k_1}{N_1} + \frac{k_2}{N_2} + \dots + \frac{k_P}{N_P} \right]^2 \right)^2. \quad (52)$$

B. Detection of binary inspiral chirps

To illustrate the use of the FCT in gravitational wave detection, we present an example using the post-Newtonian expansion for gravitational waves emitted during the binary inspiral of neutron stars and black holes. In laser interferometric detection of such signals, the SNR is obtained from (2), where $u(t)$ represents the output data stream of the interferometer, $h(t)$ is a template for a theoretical binary inspiral waveform, and $S(f)$ is the average measured power spectral density of the interferometer output [7], [12].

We will consider only low-mass binaries, since in this case the inspiral waveform is adequately determined by the masses of the two companions. At high masses, the spin of each body must also be taken into account. It is usually more convenient to parameterise the waveforms via the ‘‘chirp times’’ τ_0 and τ_1 rather than the two masses. Given m_1 and m_2 in units of solar mass \mathcal{M}_\odot , the chirp times are given by the following expressions

$$\tau_0 = \frac{5}{256}(\mathcal{M}T_\odot)^{-5/3}\eta^{-1}(\pi f_0)^{-8/3}, \quad (53)$$

$$\tau_1 = \frac{5}{192}(\mathcal{M}T_\odot)^{-1} \left(\frac{743}{336\eta} + \frac{11}{4} \right) (\pi f_0)^{-2} \quad (54)$$

where $\mathcal{M} = m_1 + m_2$ is the total mass, $\eta = m_1 m_2 / \mathcal{M}^2$ is the ‘‘reduced’’ mass, f_0 is a reference frequency and the constant $T_\odot = G\mathcal{M}_\odot / c^3$ has a value of approximately 4.925×10^{-6} s. The chirp times represent the Newtonian and first post-Newtonian contributions to the time it takes the instantaneous frequency to evolve from f_0 to infinity [9]. We choose f_0 to be the frequency at which the noise power ceases to be dominated by the low-frequency seismic component. For LIGO I this value is 40 Hz.

In order to derive an analytic expression of the Fourier transform of the templates, it has been shown [13] that by applying the stationary-phase approximation (SPA) [13] technique the resulting expression is highly accurate. To first post-Newtonian order and neglecting extrinsic signal parameters (such as orientation of the source with respect to the detector) the frequency-domain templates have the form

$$\tilde{h}(f; \tau_0, \tau_1) = f^{-7/6} e^{i2\pi\Phi(f; \tau_0, \tau_1)} \quad (55)$$

$$\Phi(f; \tau_0, \tau_1) = \frac{3}{5}\tau_0 f_0 \left(\frac{f}{f_0} \right)^{-5/3} + \tau_1 f_0 \left(\frac{f}{f_0} \right)^{-1}. \quad (56)$$

Thus $\tilde{h}(f; \tau_0, \tau_1)$ is of the chirp form (5).

A 2-parameter FCT can be used to calculate the matched filter output (3) for this waveform. We will discretise the space of parameters so that $\tau_0(k_0) = k_0\Delta\tau_0$ and $\tau_1(k_1) = k_1\Delta\tau_1$ and set the FCT phase functions to

$$\phi_0(f) = \frac{3}{5}\Delta\tau_0 f_0 \left(\frac{f}{f_0}\right)^{-\frac{5}{3}}, \quad f \geq f_0, \text{ otherwise } 0 \quad (57)$$

$$\phi_1(f) = \Delta\tau_1 f_0 \left(\frac{f}{f_0}\right)^{-1}, \quad f \geq f_0, \text{ otherwise } 0 \quad (58)$$

so that

$$\Phi(f; k_0, k_1) = k_0\phi_0(f) + k_1\phi_1(f). \quad (59)$$

Since seismic noise dominates the spectrum below f_0 , there is no loss in setting the phase functions and amplitude to zero below this value.

There are several considerations in choosing the parameter spacings $\Delta\tau_0$ and $\Delta\tau_1$. As discussed in §III, there is a loss in recovered SNR if a signal has different parameters than the template used to detect it. We can only use a finite grid of templates, yet the signals we wish to detect may be parameterised by any physically valid values of the continuous variables τ_0 and τ_1 . Accordingly, we must make the grid spacing fine enough so that we only lose an acceptably small fraction of SNR for off-grid signals. In addition to the losses due to discretising the parameter space, we must also take into account the additional loss of SNR due to the FCT approximation itself. As seen from (52), this is determined by the number of bins (N_i) in each direction and the largest value of each parameter (k_i) we wish to search over.

For sake of concreteness, we will choose our grid and FCT parameters so that we will recover at least 95% of any signal in the domain of interest. The low-mass regime (1.4 to 20 solar masses) for binary inspirals corresponds to ranges $\tau_0 = 0.3$ to 25 and $\tau_1 = 0.097$ to 1.39 in the chirp parameters. Figure ?? shows the corresponding region of the (τ_0, τ_1) -plane. It can be shown that matched filtering using a uniform grid of templates with $\Delta\tau_0 = 0.02$ and $\Delta\tau_1 = 0.003$ will recover at least 96% of any signal in this range. With this grid spacing the maximum values for the chirp indices are $k_1^{\max} = 1250$ and $k_2^{\max} = 464$. We will choose $N_1 = 78442$ and $N_2 = 29118$ so that they are in approximately the same ratio as k_1^{\max} and k_2^{\max} and so that at worst the FCT recovers 99% of the SNR obtained via matched filtering. Thus, over the whole parameter space the FCT is guaranteed to recover slightly more than

95% of the available SNR for any signal in the specified range. Figure ?? shows the actual overlap of the FCT templates with the matched filter templates as a function of distance from the origin in the the (τ_0, τ_1) -plane. The lower curve represents the lower bound for the overlap obtained from (52).

Acknowledgement

This research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

-
- [1] F.A. Jenet, and T.A. Prince, *Detection of variable frequency signals using a fast chirp transform* *Phys. Rev. D* **62** 122001 (2000)
 - [2] A. Abramovici *et al.*, *Science*, **256**, 325 (1992).
 - [3] B. Caron *et al.*, *Class. Quantum Grav.*, **14**, 1461 (1997).
 - [4] H. Lück *et al.*, *Class. Quantum Grav.*, **14**, 1471 (1997).
 - [5] M. Ando *et al.*, *Phys. Rev. Lett.*, **86**, 3950 (2001).
 - [6] C. Helstrom, *Statistical Theory of Signal Detection*, Pergamon Press, Oxford (1968).
 - [7] K.S. Thorne, In: *300 Years of Gravitation*, eds. Werner Israel, and Stephen W. Hawking, pag. 330, Cambridge University Press, Cambridge (UK) (1987)
 - [8] B.J. Owen, *Phys. Rev. D*, **53**, 6749, (1996).
 - [9] B.J. Owen and B.S. Sathyaprakash, *Phys. Rev. D* **60** 022002 (1999)
 - [10] LIGO/LSC Algorithm Library www.lsc-group.phys.uwm.edu/la1
 - [11] M. Frigo and S.G. Johnson “FFTW: An Adaptive Software Architecture for the FFT”, In: *Proceedings of the 1998 ICASSP Conference*.
 - [12] B. Schutz, in *The Detection of Gravitational Waves*, edited by D. Blair *Cambridge University Press*, Cambridge, England, pp. 406 (1991).
 - [13] S. Droz, D. Knapp, E. Poisson, and B. Owen, *Phys. Rev. D*, **59**, 124016 (1999).