Filtering Seismic Timeseries Data in Passive-source Seismic-Processing (PsSp)

Implementation details of the filters provided for isolating seismic signals

Alexander R. Blanchette*

4 July 2023

Because the goal of this document is to be as useful and accessible as possible it is a *living document*. By that I mean that it will change and grow over time in order to ensure the quality of the content. As such, I don't ever truly plan on it being *finished*.

*arbCoding@gmail.com

Abstract

Seismic data—whether it is ground motion (displacement), velocity, or acceleration—is recorded as a function of time (timeseries). Seismic timeseries recordings contain information from a plethora of sources—natural and anthropogenic. Typically, an analysis task is focused on processing/analyzing signals from a specific target *source*[a]. Generally, seismometers record and are sensitive to a frequency range that is significantly larger than that which any given analytical method/goal requires. Filtering is a critically important tool for isolating signals of interest. In this document I give a brief overview of timeseries filtering in general and provide the specific details of the filters implemented in Passive-source Seismic-processing (PsSp).

Plain Language Summary

Filters are an important step in preparing seismic data to be analyzed. In this document I will hint at some of the basic theory of seismic filtering with most of the focus being placed on the specific filters implemented in PsSp. This is intended to serve as a basic guide on the what/how of filtering seismic data without going through any explicit mathematical derivations; instead I hope this will provide a more intuitive understanding to the reader.

[a] *Source* is being used in the generic sense here; it could refer to the nucleation point of the seismic signals (earthquakes, explosions, etc.) or to structural sources (reflections, refractions, conversions).

Contents

1 Introduction

Filtering of time-dependent recordings (timeseries data) is a task common among a wide range of different scientific/engineering disciplines. While great advances have been made to analyze as much of the seismic timeseries as possible—such as the extraction of approximate Green's functions from the ambient seismic noise field—virtually all seismic analysis first requires the seismic timeseries to be filtered in order to isolate the portion of the recorded data that is most compatible with the intended analysis. In the fairly distant past this was less important due to the limitations of: 1) the available seismographic instruments, and 2) the available computational infrastructure. At present, relatively inexpensive seismometers can record broadband seismic timeseries with the flat-sensitivity range of a high-quality instrument being between a period of 1000 seconds (s) at the low end (a frequency of 0.001 Hz) up to a high of 100 Hz. Such a large dynamic range of frequencies—containing signals from numerous stationary and non-stationary sources—results in a rather complicated timeseries. This excess information convolutes the analysis of the signal(s) of interest. Filtering is the solution to this problem; it allows the analyst to isolate the signal(s) of interest by removing the *noise*[1] from the recorded timeseries. Fortunately—thanks to the many technological advancements over the last few decades—personal computers are sufficiently capable of handling most modern seismic workflows (all except particularly large or high throughput workflows that require more specialized computing infrastructure).

Filtering is a tremendously wide topic of study itself, with a rich history across many otherwise disparate fields of study, well beyond the scope of this simple document. Therefore, I will restrict myself to—at most—provide a brief discussion on the fundamentals of filtering, while leaving many of the details up for the intrepid reader to research on their own. Afterward, I will provide a detailed description of the implementation of filters specific to PsSp. This is not meant to provide a detailed derivation from first principles of the filters employed, but is instead intended to serve as a pragmatic guide on their internal implementation. This is intended as much as a guide to any reader as it is to myself.

2 Background

Filtering of timeseries data is necessary across numerous disciplines (e.g. seismology, electrical engineering, audio engineering, etc.). Due to its prevalence across a diverse range of disciplines, it has been studied extensively over time—with many of the major theoretical advances having been discovered more than 100 years ago (Joseph Fourier published on his series expansion in 1822[2]). In fact, the Butterworth filters employed here were first published by Stephan Butterworth in 1930[1]. Of course, there has been significant progress since, both theoretical and practical. While much could be—and has been—written on the history of filtering signals, I think that it is sufficient here to simply say that it is a quite mature field of study and leave the historical developments to the reader to research on their own, if they so desire.

[1] *Noise* is a bit of a misunderstood term in seismology, especially with the advances in the last few decades in ambient noise seismology. The word 'noise' is used here to mean 'any signal that is not of interest for a specific analysis' and does not provide any suggestion to the source of the noise.

2.1 Mathematical Background

While I do intend for this document to focus more on the development of the reader's intuitive understanding of the filtering process, it would be a significant disservice to totally forgoe all mathemtical formalism. In fact, the equations themselves provide a great deal of insight into inner workings of the filters. In order to provide a sufficiently comprehensive introduction to the application of filters I will need to first introduce a few mathematical concepts. In particular, I need to discuss how it is that we take data that is measured over time and *transform*[2] it to the frequency (spectral) domain (via the *Fourier Transform*) and then back to the time domain (via the *Inverse Fourier Transform*). I must also discuss how we take a filter that is defined in the Lapace (Complex-Frequency) domain and apply it to a signal in the (non-complex) frequency domain.

2.1.1 The Temporal (Time) Domain

Observations are always made as a function of time. Whether the process is constant or dynamic, the very nature of measurement is time-dependent itself. We can, at best, make the rather artifical assumption of constancy and pretend that time doesn't exist—but we can never truly remove the temporal component from our measurements. Seismographic stations record ground motion as a function of time. The distinction between measurements of displacement, velocity, and acceleration is simply a matter of differentiation $x \to v \to a$ or integration $a \to v \to x$. The distinction between ground motion and the actual instrumental measurement is related the internal instrument response[3].

[2]*Transform* is commonly used in mathematics to refer to a linear change of basis. That is, to change the meaning of the dependent-variable in a well-behaved and reversable manner.

[3]Instrument response removal can be thought of as a form of filtering, but it will be a topic of a separate document.

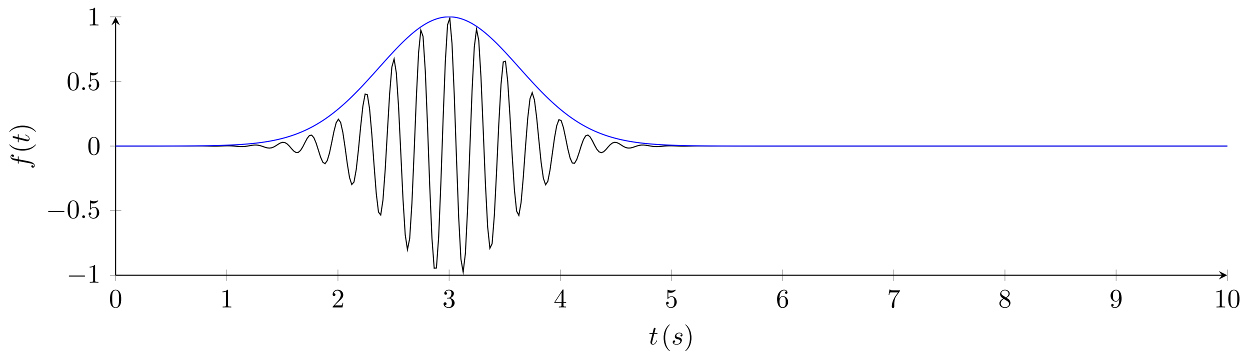

Figure 1: Wave-packet (black) and its corresponding envelope function (blue).

2.1.2 The Spectral (Frequency) Domain

2.1.3 The Laplace (Complex-Frequency) Domain

3 The Butterworth Filter

3.1 Lowpass

$$H(s) = \frac{G_0}{B_n(a)}; a = \frac{s}{w_c} \tag{1}$$

3.2 Highpass

$$B_n(s) = \sum_{k=0}^{n} a_k s^k \tag{2}$$

3.3 Bandpass

3.4 Bandreject

$$a_{k+1} = a_k \frac{cos\,(k\gamma)}{sin((k+1)\gamma)}; a_0 = 1; \gamma = \frac{\pi}{2n}; a_k = a_{n-k} \tag{3}$$

References

[1] Stephen Butterworth et al. “On the theory of filter amplifiers.” In: *Wireless Engineer* 7.6 (1930), pp. 536–541 (cit. on p. 4).

[2] JBJ Fourier. “Théorie analytique de la chaleur: Paris.” In: *Académie des Sciences* 3 (1822) (cit. on p. 4).