

Nonlinear time-warping made simple: A step-by-step tutorial on underwater acoustic modal separation with a single hydrophone

Julien Bonnel, Aaron Thode, Dana Wright, and Ross Chapman

Citation: [The Journal of the Acoustical Society of America](#) **147**, 1897 (2020); doi: 10.1121/10.0000937

View online: <https://doi.org/10.1121/10.0000937>

View Table of Contents: <https://asa.scitation.org/toc/jas/147/3>

Published by the [Acoustical Society of America](#)

ARTICLES YOU MAY BE INTERESTED IN

[Measurement and modeling of sound propagation over continental slope in the South China Sea](#)

[The Journal of the Acoustical Society of America](#) **147**, EL209 (2020); <https://doi.org/10.1121/10.0000801>

[Superposition method for modelling boundaries between media in viscoelastic finite difference time domain simulations](#)

[The Journal of the Acoustical Society of America](#) **146**, 4382 (2019); <https://doi.org/10.1121/1.5139221>

[Source localization in the deep ocean using a convolutional neural network](#)

[The Journal of the Acoustical Society of America](#) **147**, EL314 (2020); <https://doi.org/10.1121/10.0001020>

[Seasonally-invariant head wave speed extracted from ocean noise cross-correlation](#)

[The Journal of the Acoustical Society of America](#) **147**, EL241 (2020); <https://doi.org/10.1121/10.0000877>

[Machine learning in acoustics: Theory and applications](#)

[The Journal of the Acoustical Society of America](#) **146**, 3590 (2019); <https://doi.org/10.1121/1.5133944>

[Analysis of the ray-based blind deconvolution algorithm for shipping sources](#)

[The Journal of the Acoustical Society of America](#) **147**, 1927 (2020); <https://doi.org/10.1121/10.0000919>



Nonlinear time-warping made simple: A step-by-step tutorial on underwater acoustic modal separation with a single hydrophone

Julien Bonnel,^{1,a)} Aaron Thode,² Dana Wright,³ and Ross Chapman⁴

¹Applied Ocean Physics and Engineering Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

²Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA

³University of Washington and National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center Marine Mammal Lab, Seattle, Washington 98105, USA

⁴School of Earth and Ocean Sciences, University of Victoria, Victoria, Canada

ABSTRACT:

Classical ocean acoustic experiments involve the use of synchronized arrays of sensors. However, the need to cover large areas and/or the use of small robotic platforms has evoked interest in single-hydrophone processing methods for localizing a source or characterizing the propagation environment. One such processing method is “warping,” a non-linear, physics-based signal processing tool dedicated to decomposing multipath features of low-frequency transient signals (frequency $f < 500$ Hz), after their propagation through shallow water (depth $D < 200$ m) and their reception on a distant single hydrophone (range $r > 1$ km). Since its introduction to the underwater acoustics community in 2010, warping has been adopted in the ocean acoustics literature, mostly as a pre-processing method for single receiver geoacoustic inversion. Warping also has potential applications in other specialties, including bioacoustics; however, the technique can be daunting to many potential users unfamiliar with its intricacies. Consequently, this tutorial article covers basic warping theory, presents simulation examples, and provides practical experimental strategies. Accompanying supplementary material provides MATLAB code and simulated and experimental datasets for easy implementation of warping on both impulsive and frequency-modulated signals from both biotic and man-made sources. This combined material should provide interested readers with user-friendly resources for implementing warping methods into their own research. © 2020 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0000937>

(Received 5 September 2019; revised 30 January 2020; accepted 3 March 2020; published online 25 March 2020)

[Editor: John A. Colosi]

Pages: 1897–1926

TABLE OF CONTENTS

I. INTRODUCTION.....	1898	1. Warping.....	1907
A. Context.....	1898	2. Inverse warping.....	1908
B. Outline of the tutorial.....	1899	B. Example.....	1908
II. UNDERSTANDING THE RECEIVED SIGNAL.....	1900	V. PRACTICAL CONSIDERATIONS WHEN	
A. Single receiver context.....	1900	WARPING.....	1908
B. Modal propagation.....	1901	A. Warping parameters.....	1908
C. A simplified coastal environment.....	1902	B. Impulsive signals and time origin	
D. Time-frequency representations.....	1903	identification.....	1908
E. Spectrogram.....	1904	C. Non-impulsive signals and source	
III. WARPING: THEORY.....	1904	deconvolution.....	1910
A. Understanding warping.....	1904	D. Frequency modulated signals and phase	
B. Mathematical definition of warping.....	1906	compensation.....	1911
C. Dispersion based warping.....	1906	VI. APPLICATIONS.....	1913
1. Warping model: The isovelocity		A. Modal filtering and dispersion curve	
waveguide.....	1906	estimation.....	1913
2. Warping function.....	1907	1. Warping.....	1913
IV. WARPING: ALGORITHM.....	1907	2. Time-frequency filtering of warped	
A. Numerical implementation of warping.....	1907	modes.....	1913
		3. Isolating modes in the original time	
		domain.....	1914
		4. Dispersion curve estimation.....	1914
		5. Further technical details.....	1914

^{a)}Electronic mail: jbonnel@whoi.edu, ORCID: 0000-0001-5142-3159.

B. Source localization	1915
VII. EXPERIMENTAL EXAMPLES	1915
A. Right whale gunshot	1916
B. Combustive sound source	1917
C. Bowhead whale upsweep	1918
VIII. CONCLUSION	1919

LIST OF NOTATIONS

The notations are organized into five subsections. In each subsection, notations are ordered alphabetically.

General notations

f	Frequency
f_s	Sampling frequency
j	Unit imaginary number
λ	Wavelength
t	Time
dt	Time shift

Environment and experiment geometry

c_w	Water sound speed
c_b	Basement (seabed) sound speed
D	Water depth
r	Source/receiver range
ρ_w	Water density
ρ_b	Basement density
z_s	Source depth
z_r	Receiver depth

Signal related notations

g	Green's function or impulse response
ϕ_s	Phase of the source
s	Source signal
t_s	Source group delay
y	Received signal
y_{deconv}	Received signal after source deconvolution
y_{pc}	Received signal after phase compensation
y_u	Unwarped signal
y_w	Warped signal

Warping related notations

\mathcal{D}_f	Frequency domain of the original signal
\mathcal{D}_f^h	Frequency domain of the warped signal
\mathcal{D}_t	Time domain of the original signal
\mathcal{D}_t^h	Time domain of the warped signal
f_{min}	Minimal frequency of the original signal
f_{min}^h	Minimal frequency of the warped signal
f_{max}	Maximal frequency of the original signal
f_{max}^h	Maximal frequency of the warped signal
f_s^h	Sampling frequency of the warped signal
h	Warping function
K	Number of samples of the warped discrete signal
\mathcal{M}	Time-frequency mask for modal filtering
N	Number of samples of the original discrete signal
\bar{t}	Continuous warped time

\mathcal{T}	Original time axis
$\bar{\mathcal{T}}$	Warped time axis
t_{min}	Minimal time of the original signal
t_{min}^h	Minimal time of the warped signal
t_{max}	Maximal time of the original signal
t_{max}^h	Maximal time of the warped signal
t_n	Time sample number n in the original time axis
\bar{t}_n	Time sample number n in the warped time axis
t_r	Time origin for warping

Mode dependent notations (mode m)

a_m	Modal amplitude
β_m	Imaginary part of the horizontal wavenumber
$f_{c,m}$	Cutoff frequency
k_m	Real part of the horizontal wavenumber
m	Mode number
M	Number of modes
ϕ_m	Modal phase
Ψ_m	Modal depth function
t_m	Modal travel time for an impulsive source at $t = 0$
τ_m	Modal travel time for a frequency modulated source with group delay $t_s(f)$
v_m	Group speed

I. INTRODUCTION

A. Context

The development of underwater acoustic signal processing was originally driven by military applications that require advanced sonar processing (Ainslie, 2010) to detect and localize quiet sources in an uncertain environment (Dosso and Wilmut, 2011). These days, acoustic signal processing also provides a major avenue for conducting oceanographic research. For example, active acoustics allows estimation of fish populations (Makris *et al.*, 2006; Stanton *et al.*, 2018) while passive acoustic monitoring (PAM) permits study of marine mammal seasonality and regional presence (Mellinger *et al.*, 2007). One factor common to all of these applications, civilian and military, is the need for robust algorithms that incorporate the complexity of the ocean environment, whose properties are only partially known and variable in time and space. Numerous attempts to adapt more advanced signal processing or beam-forming methods to underwater acoustic data [e.g., matched-field processing methods (Baggeroer *et al.*, 1993)] have met with limited success, because most methods require incorporating more knowledge about the oceanic propagation environment than is typically available. Furthermore, most advanced signal processing methods require the use of extended time-synchronized array hydrophones to perform spatial and temporal filtering. The deployment of such systems is awkward and expensive.

In this tutorial, we describe a relatively recent nonlinear signal processing method—termed *warping*—that is dedicated to the study of low-frequency ($f < 500$ Hz) transient sounds recorded in coastal environments (water depth

$D < 200$ m) after propagation over at least several kilometers (range $r > 1$ km). Numerous underwater sounds of interest fit into this category, including baleen whale vocalizations, airgun signals, or sounds from scientific tomography experiments. The objective of warping is to extract normally hidden features from the received signal that can then be used by other algorithms to localize the transient sound source, and/or to infer environmental information along the propagation track.

One probably knows the *warping* word from a certain famous science fiction TV show, in which warping is used to bend space, such that ships can travel faster than light. While this concept has been repudiated by the National Aeronautics and Space Administration (NASA, 2015), it nonetheless has concrete applications in signal processing, where space (or time) can be warped virtually, using a computer. What differentiates warping from other underwater signal processing methods is that it requires only a single hydrophone, and has proven robust to environmental uncertainty: it usually works even without detailed prior information about the environment. As will be seen shortly, the method works well in real ocean environments, even though the basic algorithm is derived from a simple idealized model of a shallow ocean.

The fact that this method requires only a single hydrophone has profound implications, particularly for bioacoustic studies. Passive acoustic recording systems are now routinely used to detect the seasonal and regional presence of whale species all over the world, including shallow continental shelf waters in the arctic and along the U.S. eastern seaboard. This is most often accomplished by deploying single-hydrophone recording packages over wide regions. The presence of baleen whale species in these data sets is identified by either manual or automated review of spectrograms, and then species-specific sounds are flagged (Leroy *et al.*, 2018; Moore *et al.*, 2006; Thode *et al.*, 2012). Subsequent localization of animal sounds from these data sets is often desirable, because source localization is a key step in establishing population density estimates and evaluating subtle responses to anthropogenic activities. However, this is rarely implemented, because most traditional source localization methods require the deployment of multiple hydrophones over wide spatial regions, and complex measurements of relative arrival times between sensors. In this tutorial, we demonstrate how localization information can be extracted from both baleen whale impulsive and frequency-modulated sounds from single-hydrophone recordings in shallow water. Although source localization is not the end goal of this tutorial, we will nonetheless illustrate how baleen whale sounds can be localized using warping methods. The conclusion of this demonstration is that hundreds of existing data sets may have exploitable baleen whale localization (and environmental) information embedded inside their recordings.

Bioacoustics is not the only application for warping. Traditional ocean acoustic experiments typically involve the use of array(s) of synchronized receivers. Hydrophone arrays are useful as they improve the signal-to-noise ratio

(SNR), as well as increase the spatial diversity of the sampled field. Arrays, and particularly vertical line arrays have been extensively used for water column tomography and/or geoacoustic inversion (Caiti *et al.*, 2006; Chapman, 2012). However, in shallow water, single hydrophone deployments also provide valuable data if the sources involved are sufficiently intense and broadband, because the frequency diversity of the signal can substitute for the spatial diversity normally sampled by a typical acoustic array (Hermand, 1999; Jesus *et al.*, 2000). The warping method presented in this tutorial further enables extraction of high resolution information that can be used as the core of single-hydrophone inversion schemes. To illustrate this, we will demonstrate how propagation information can be extracted from a tomographic source, and then used to localize it. More complex inverse problems, such as tomography or geoacoustic inversion, will be ignored to keep the focus of this tutorial manageable, but relevant key references will be provided.

While warping has demonstrated its utility and practicality, its learning curve can be steep. Warping is applicable to specific situations and does require expertise and some judgment in order to be used properly, as automated methods for bulk warping do not yet exist. Hence, the objective of this tutorial is to make warping understandable and available to any researcher that is interested in its application. To meet this objective, this tutorial includes extensive supplementary material¹ that provides both MATLAB code and data from several walk-through examples of biotic and abiotic sound sources. The aim is to provide the reader with the opportunity to easily try warping, to confirm they are using the technique properly, and to facilitate their application of the method on their own datasets.

B. Outline of the tutorial

This tutorial assumes that the reader has basic knowledge about time-frequency (TF) analysis. In particular, we assume that the reader knows how to generate and interpret a spectrogram, a basic TF representation that is widely used in ocean acoustics. We also mention that an additional non-technical overview of most of the concepts presented in this tutorial is provided in Bonnel (2018).

The tutorial is arranged so that most sections can be read separately. This introduces some redundancy, but overall should make the tutorial most useful to readers that are interested in specific topics. The remainder of the tutorial is organized as follows:

Section II covers the basic background required to understand the received signal, including the single receiver context (Sec. II A), modal propagation (Sec. II B), the Pekeris waveguide model (a simple model for coastal environments, Sec. II C), and time-frequency analysis (Sec. II D).

Section III details the warping theory. It starts with simple explanations (Sec. III A), introduces the warping general concepts (Sec. III B), and then details a specific warping adapted to our modal dispersion concept (Sec. III C).

Section IV presents the detailed warping algorithm—it is intended for readers with preliminary knowledge about warping, and can be skipped on a first reading. First, the numerical implementation of warping is presented (Sec. IV A). An example is then given, and remarks on the importance of time origin for warping are given (Sec. IV B).

Section V probably constitutes the heart of the tutorial for beginners. It presents practical uses of warping, serving as a template for using warping in specific scenarios. It starts with a quick explanation on warping parameters (Sec. V A). It continues with warping use for impulsive sources (Sec. V B), and emphasizes the importance of correctly choosing the time origin. It then presents warping use for sources that are not impulsive, but whose waveform is known (Sec. V C). Finally, the section ends by showing how warping can be used when the source is an unknown frequency modulation (Sec. V D), a typical situation for baleen whale calls.

Applications of warping are then covered in Sec VI. First, modal filtering and dispersion curve estimation are presented (Sec. VI A). Then, source localization is quickly reviewed (Sec. VI B).

Last but not least, Sec. VII presents experimental examples that cover both warping and source localization. The first example is an impulsive vocalization by a right whale (Sec. VII A). The second example is a non-impulsive

controlled tomography source, whose waveform is known (Sec. VII B). The last example is an unknown frequency modulation by a bowhead whale (Sec. VII C). The various steps of the analysis are shown, and the data used are provided in the supplementary material,¹ so that readers can process the data and confirm that their results match those shown in the tutorial.

The tutorial ends with concluding remarks in Sec. VIII and three Appendixes. A first Appendix delves into the difference between time and frequency warpings, which may be of interest to specialist readers. A second Appendix reviews the development of warping theory through the literature, to act as a guide for readers seeking further references. A third and final Appendix details the derivations that lead to numerical warping formulas.

Figure 1 shows the entire warping process as a flowchart, which provides a useful roadmap for the reader as the tutorial progresses.

II. UNDERSTANDING THE RECEIVED SIGNAL

A. Single receiver context

The context considered in this tutorial is the study of low-frequency transient signals ($f < 500$ Hz) that have propagated in shallow water ($D < 200$ m) over at least a few

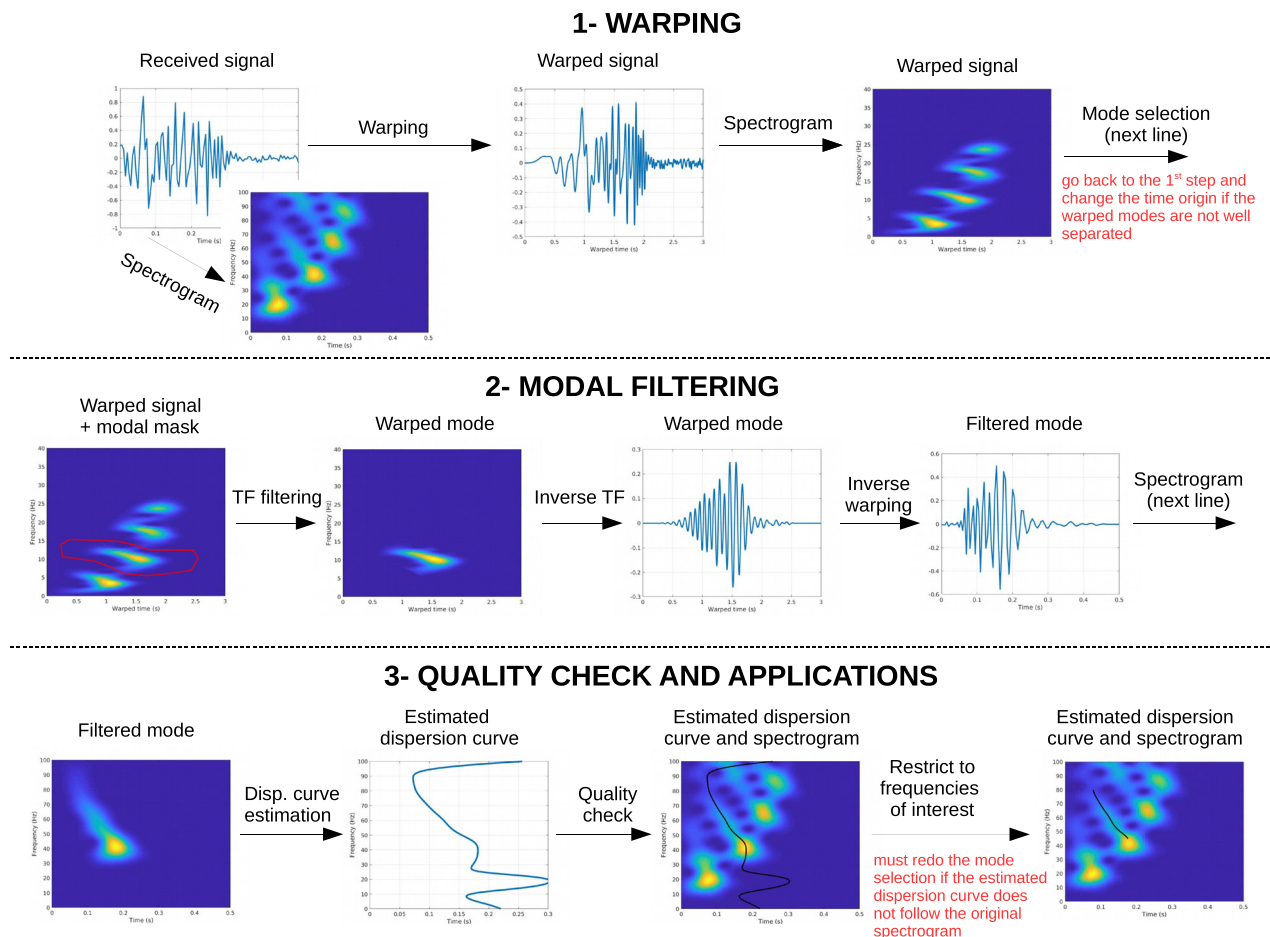


FIG. 1. (Color online) Flowchart of the modal filtering process using warping.

kilometers. This low-frequency/shallow water requirement ensures that acoustic propagation can be accurately modeled using *modal propagation*, which lies at the heart of the warping technique. While *modes* will be defined more precisely below, one can envision them as frequency-dependent arrivals generated from energy arriving along different routes between source and receiver. As a result, a signal received at a hydrophone will actually contain several distinct modal (arrivals) components. If these components can be isolated, one can use them to estimate the range and depth of the source (e.g., by comparing the relative arrival times and amplitudes of the arrivals) or infer details of the sound speed profile or ocean bottom composition (e.g., by measuring how the relative arrival times and amplitudes change with frequency). Extracting individual modes from a received signal is called *modal filtering*, and has been done historically by spatially filtering data on vertical line arrays (Buck *et al.*, 1998; Neilsen and Westwood, 2002; Tindle *et al.*, 1978).

The key point of this tutorial is that under most low-frequency shallow water environments, nonlinear signal processing (warping) and TF analysis can be used to conduct modal filtering on a single hydrophone. Once modes have been filtered, the fact that one has used a single receiver or an array becomes unimportant, and the source can be localized and/or the environment estimated using any previously developed modal based inversion scheme. The most popular method to do so is matched mode processing. It was originally proposed for array data (Wilson *et al.*, 1988; Yang, 1987), and was later extended to single receiver scenarios (Le Touzé *et al.*, 2008; Thode *et al.*, 2017).

B. Modal propagation

This section describes modal propagation in shallow water. Readers uninterested by further technical details can go directly to Sec. IIC. Readers who want to learn more about it are referred to Frisk (1994) and/or Jensen *et al.* (2011).

Acoustic propagation in shallow water is highly impacted by interactions with the sea surface and the seabed. This environment acts as an acoustic waveguide, which effectively operates on the signal as a linear time-invariant system. In other words, if one considers a source signal $s(f)$ emitted at depth z_s , the signal $y(f, z_s, z_r)$ received at depth z_r and range r appears to have been filtered by the environment,

$$y(f, z_s, z_r) = s(f)g(f, z_s, z_r, r), \quad (1)$$

where $g(f, z_s, z_r, r)$ is the environmental filtering effect. The quantity g has been assigned various names, including the transfer function or Green's function. Note that the inverse Fourier transform of g is usually called the impulse response. This term arises from the fact that if the source is a perfect impulse at time $t = 0$, then the received signal will simply become $y(t, z_s, z_r) = g(t, z_r, r)$, or the "response" of the medium to the impulse.

In the context considered here, shallow water (water depth $D < 200$ m) and low-frequencies ($f < 500$ Hz), the propagation is conveniently described by normal mode theory, provided that the ocean conditions (e.g., bathymetry, bottom composition) vary little with distance from the origin (the "range-independent" assumption). At ranges greater than a few acoustic wavelengths, the resulting acoustic field can then be interpreted as the sum of several modal components, with each mode propagating dispersively (i.e., the effective propagation speed varies with frequency). Formally stated,

$$g(f, z_s, z_r, r) = \sum_{m=1}^M a_m(f, z_s, z_r) e^{i\phi_m(f, r)}, \quad (2)$$

with M being the number of distinct propagating modes, with each mode having a unique amplitude $a_m(f, z_s, z_r)$ and a phase $\phi_m(f, r)$. This equation is derived from the acoustic wave equation using the classic separation of variables method in environments that are azimuthally symmetric and range-independent.

Before continuing, note that modal amplitude primarily depends on source/receiver depth, but not range, whereas the modal phase (and thus modal travel time) depends primarily on range, but not on source/receiver depth.

Of particular interest here is the modal phase

$$\phi_m(f, r) = rk_m(f), \quad (3)$$

where $k_m(f)$ represents the real part of the horizontal wavenumber of mode m . In other words, $k_m(f)$ is the spatial frequency of mode m . It solely depends on the environment (water depth, sound speed profile, seabed geoacoustic properties, etc.), but not on the experimental geometry. As a result, the modal travel time

$$t_m(f) = \frac{1}{2\pi} \frac{\partial \phi_m(f, r)}{\partial f} = \frac{r}{v_m(f)} \quad (4)$$

depends only on the range r and the environment. This environmental dependence is expressed through the modal group speed $v_m(f) = \partial f / \partial k_m(f)$.

On the other hand, the modal amplitude is given by

$$a_m(f, z_s, z_r) = Q \frac{\Psi_m(f, z_s) \Psi_m(f, z_r)}{\sqrt{k_m(f)} r} e^{-r\beta_m(f)}, \quad (5)$$

with $\Psi_m(f, z)$ being the modal depth function of mode m , $\beta_m(f)$ the imaginary part of the modal wavenumber, and Q a constant unimportant for our purposes here. The modal amplitude clearly depends on both the environment and on source/receiver depth through the modal depth function $\Psi_m(f, z)$. It also weakly depends on range through, but this is often ignored (Jensen *et al.*, 2011).

The modal travel time [Eq (4)] and amplitude [Eq (5)] have been given for the impulse response only. Let us now consider a source signal $s(f) = |s(f)|e^{i\phi_s(f)}$. Equation (1) shows that the received signal is a sum of modal components, which are all similarly impacted by the source amplitude and

phase. The received amplitudes become $|s(f)|a_m(f, z_s, z_r)$ and the received phase becomes $\phi_s(f) + \phi_m(f, r)$. As a result, the modal travel times for a general received signal become

$$\tau_m(f) = t_s(f) + \frac{r}{v_m(f)}, \quad (6)$$

where $t_s(f) = (1/2\pi)[\partial\phi_s(f)/\partial f]$ is the source time-frequency law. Note that if the source is impulsive, all the frequencies are emitted at a single time t_0 and $t_s(f) = t_0$. The quantity $t_s(f)$ is properly defined from a signal processing point of view as the source “group delay.” In the same manner, $\tau_m(f)$ is defined as the group delay of the received mode m . Because it marks a given mode’s location in the time-frequency (TF) domain (e.g., on a spectrogram, see Sec. II D), Eq. (6) is also called the TF dispersion curve.

Note that the derivations presented above assume that $a_m(f, z_s, z_r)$ is effectively an amplitude while $\phi_m(f, r)$ is effectively a phase. In other words, we assume that $a_m(f, z_s, z_r)$ —and thus $\Psi_m(f, z)$ —varies slowly with f , while $e^{j\phi_m(f, r)}$ oscillates more rapidly with respect to f . This is largely true for the context considered here (low-frequency propagation in shallow water). This decomposition may be questionable in other contexts, such as deep water (Emmetière *et al.*, 2018).

C. A simplified coastal environment

In this section, we present a simple model of a coastal environment. It will be used to illustrate modal propagation and to generate simulated signals to be used throughout the tutorial.

A straightforward way to acoustically model coastal environments is to consider an isovelocity fluid layer (the water), between a perfectly reflecting surface (the sea surface) and a semi-infinite isovelocity fluid basement (the seabed). This model is called the Pekeris waveguide, it carries the name of C. L. Pekeris, who first derived the associated equations (Pekeris, 1948). The model does not include any realistic range/depth dependence of the environment (e.g., water column stratification, seabed layers, etc.). Nonetheless, it produces realistic modal features. It can thus be used as an educational example. We will see later in the tutorial that it can also be used as the core of many localization algorithms.

A Pekeris waveguide is fully defined by the parameters of its water column and seabed. The following notations and nominal values will be used throughout the paper:

- water column: depth $D = 100$ m, sound speed $c_w = 1500$ m/s, density $\rho_w = 1000$ kg/m³;
- basement (sediment): sound speed $c_b = 1600$ m/s, density $\rho_b = 1500$ kg/m³.

Using these parameters, one can use simple numerical solvers (Jensen *et al.*, 2011) to find modal wavenumbers $k_m(f)$, group speeds $v_m(f)$, and depth functions $\Psi_m(f, z)$. These can be further combined to simulate a propagated signal in the frequency domain [using Eqs. (2), (3), and (5)] as well as in the time domain (going from the frequency domain to time domain with an inverse Fourier transform), or to directly predict modal travel time [using Eq. (6)].

Further details are not provided here. Rather, a MATLAB code to simulate propagation in a Pekeris waveguide is provided as supplementary material.¹

Figure 2 illustrates the propagation of a pulse in a Pekeris waveguide. Please note that some sediment attenuation (0.2 dB/λ) has been added in the seabed. This does not change the general propagation features, but slightly increases the modal separation, which makes the figure easier to understand.

The source signal considered here is a short broadband pulse. It lasts less than 0.1 s, and most of its frequency content is within 25 and 75 Hz. Because modal group speed depends on frequency, the duration of the pulse arrival increases with range. This can be understood by thinking about a marathon. All runners start at the same time, but the gap between runners increases as the race progresses, because their individual speeds are different. This can be seen on Fig. 2 after propagation over 5 km: the duration of the received pulse has nearly doubled. This effect, due to a frequency dependent speed, is called *dispersion*.

However, two different dispersions co-exist for modal propagation. For a given mode m_0 , the group speed $v_{m_0}(f)$ changes with frequency f , which is called intra-modal dispersion. As a result, the duration of a given mode increases with increasing range. On the other hand, for a given frequency f_0 , the group speed $v_m(f_0)$ changes with mode number m , which is called inter-modal dispersion. As a result, the gap between modes also increases with increasing range. We can further develop the racing analogy by considering several groups of runners (e.g., age classes), with runners within the same age class displaying similar speeds, but with different age classes displaying relatively larger differences in speed (with older age classes generally slower than younger age classes). If all the runners/groups start the race at the same time, one will not be able to distinguish the different age classes at the beginning. However, as the race continues, the different age classes will separate along the course. If the race lasts long enough, then the age classes may end up being completely separated in time, arriving past the finish line in distinct waves. Since runners within a given age class do not have exactly the same speed, timing gaps between runners in the same age class will also appear, causing a time spread in the finishing time within an age class. From a modal perspective, a running age class represents a given mode arrival, the time spread between runners within an age-class represents intra-modal dispersion, and the timing spread between age classes represents inter-modal dispersion. These concepts are further illustrated in Fig. 2. At 15 km range, one can see two modes that are nearly separated: mode 1 arrives between $t \simeq 0$ to $t \simeq 0.20$ s, while mode 2 arrives between $t \simeq 0.15$ to $t \simeq 0.40$ s. There is also a third mode, barely visible after $t \simeq 0.35$ s. At 30 km, the first two modes have become fully separated in time. This example also nicely illustrates how the duration (spread) of an individual mode arrival increases with range. This is particularly evident for mode 2, which lasts more than 0.4 s when $r = 30$ km.

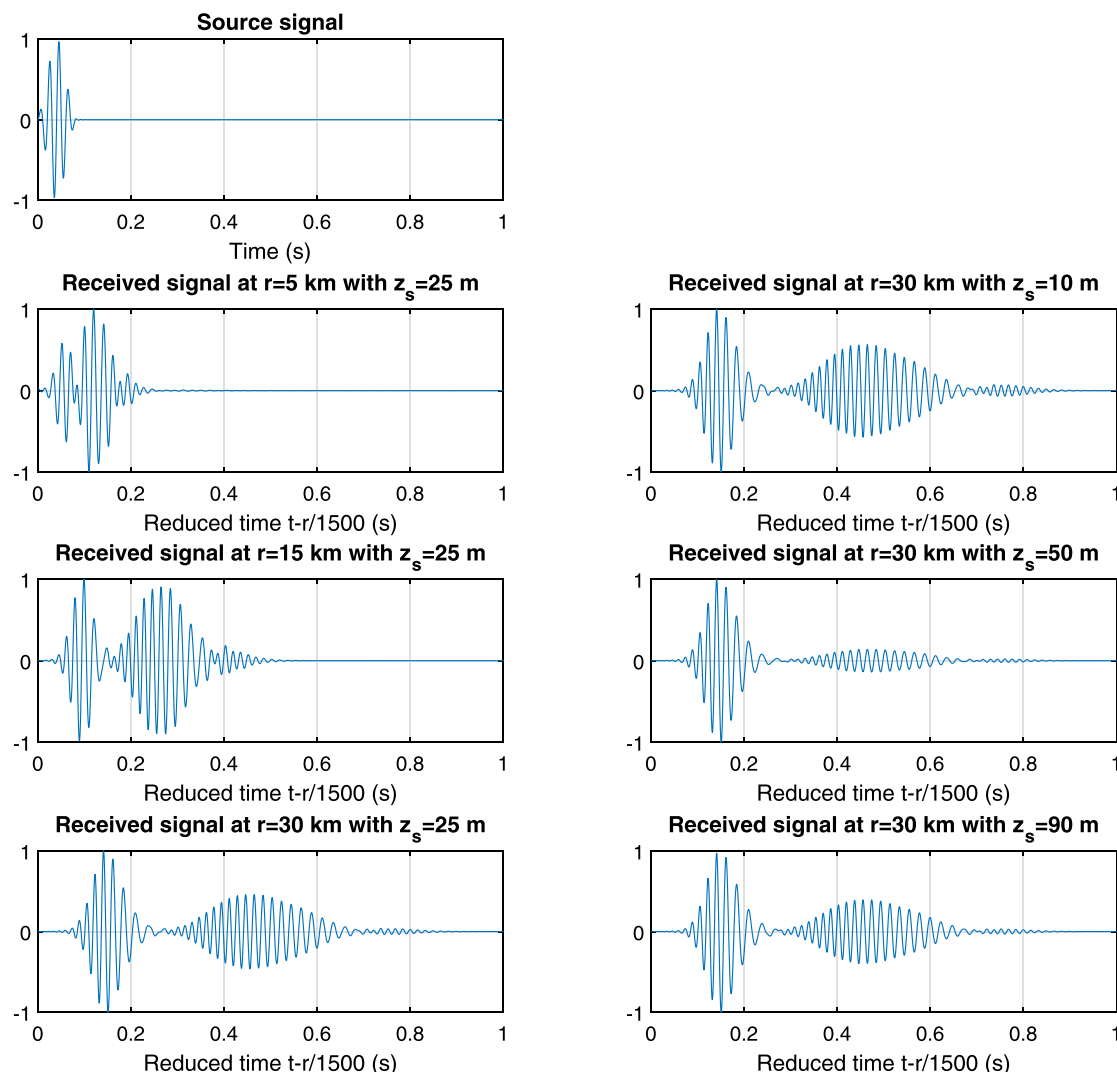


FIG. 2. (Color online) Simulated signal in a Pekeris waveguide with attenuation in the seabed. On each panel, the signal amplitude is arbitrarily normalized. The figure reproduces an example provided in [Jensen et al. \(2011\)](#) (pp. 636–640).

While mode duration/travel time changes with range, one can also see from the right column of Fig. 2 that the relative mode amplitudes depend on source/receiver depth. This is particularly evident for mode 2, as can be seen on the second column of Fig. 2. This phenomenon, mentioned in Sec. II B, is due to the oscillations of the modal depth functions (and thus the mode amplitudes) with depth. It will not be further explored in this tutorial.

D. Time-frequency representations

The example shown in Fig. 2 provides a first illustration of modal propagation. However, modal dispersion becomes easier to understand when the signal is plotted in the TF domain. Indeed, by using TF representation, one can directly visualize the modal dispersion. This is illustrated in Fig. 3, which presents idealized TF representations of underwater signals. The right part of the figure (with the warped signal) can be ignored for now.

The first line of Fig. 3 illustrates the propagation of an impulsive source. Such a source is a vertical line in the TF

domain, because all frequencies are emitted at the same time. After propagation, one can see several structures in the TF domain, because several modes are propagating. Each mode is a curve with high frequencies arriving before low frequencies. This is because in general, for shallow water environments, group velocity is higher for high frequencies, and thus travel time is smaller. Note that warping is based on the hypothesis that low order modes arrive first and that modes do not cross in the TF domain, a situation representative of most shallow-water modal propagation. Specific cases where the hypothesis is severely violated (and thus warping does not work at all) are discussed in the conclusion of the article. The remainder of the article assumes that this hypothesis (conventional shallow-water modal propagation) is verified, at least over the frequency band of interest. As a reminder, the formula for dispersion curves (the blue lines in Fig. 3) is given by Eq. (6).

The second line of Fig. 3 illustrates the propagation of a frequency modulated source. In the example considered here, the source signal is a linear downsweep: the frequency

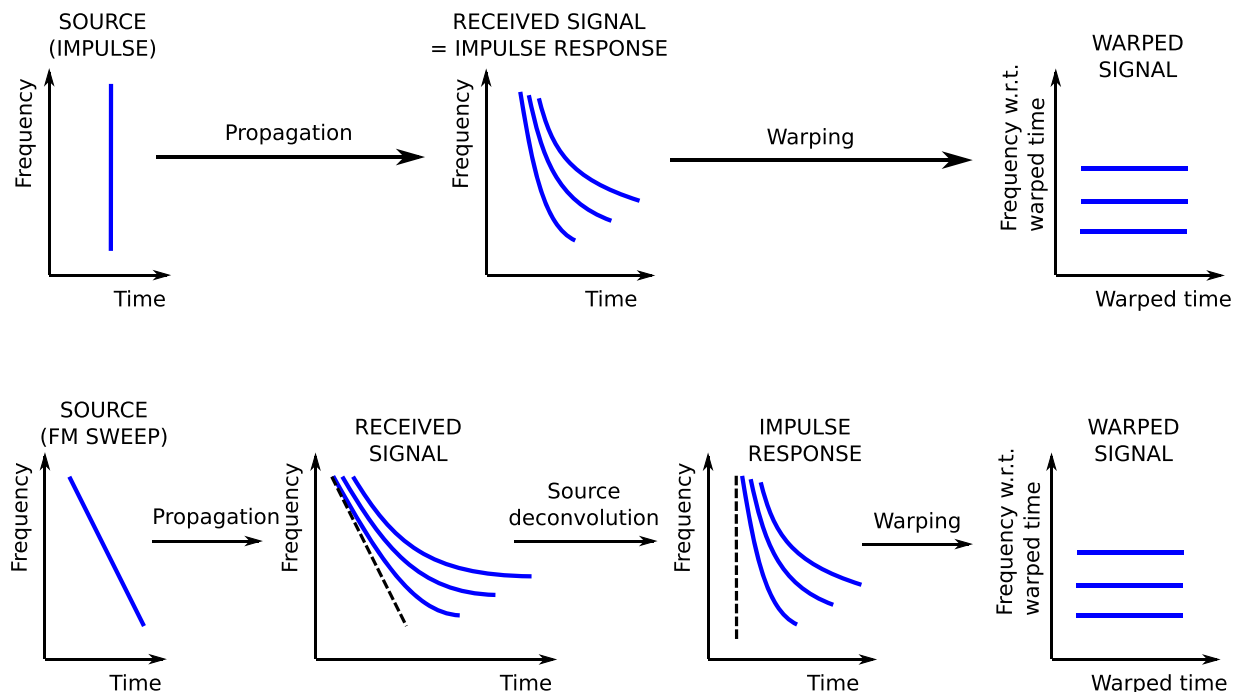


FIG. 3. (Color online) Time-frequency diagrams illustrating propagation and warping for an impulsive source (top line) and a frequency-modulated source (bottom line). Figure adapted from Bonnel *et al.* (2014).

gradually decreases with time. The interpretation of the received signal is the same as for the impulsive source. If one goes back to the race analogy, the impulsive source case is a race where all runners start together. On the other hand, the FM source context represents a race where runners have a staggered start. The delayed departure effectively delays their arrival time, which globally shifts the arrival pattern.

E. Spectrogram

Figure 3 is an ideal TF representation, which is impossible to obtain in reality. Any true TF representation is contaminated by TF uncertainty (Boashash, 2015), which makes it impossible to perfectly isolate a signal both in time and frequency. As a result, one cannot obtain infinitesimally thin lines—as in Fig. 3—to represent modes. Also, one will have to deal with interferences between modes, particularly at short ranges where mode time separations are small.

The spectrogram is one of the easiest TF representations to generate, and is the default choice for most practical applications. The spectrogram's main feature is that it minimizes the interferences between signal components (in our context, modes), and thus we will use the spectrogram in this tutorial. The trade-off is that a spectrogram provides a really poor TF resolution of these components (i.e., the mode dispersion curves). The dispersion curve estimation issue will be dealt with later (see Sec. VIA).

It is beyond the scope of this tutorial to detail how spectrograms are computed. Readers who want more background on it are referred to Boashash (2015). For practical matters, a MATLAB code to compute spectrograms is provided as supplementary material.^{1,2}

Examples of spectrograms simulated in a Pekeris waveguide are presented in Fig. 4. The figure shows the time signal in the upper panels, and the spectrograms in the middle panels. As explained before, modal time separations increase with range. At large ranges (i.e., 30 km), modes are well time-separated and thus barely interfere. At intermediate ranges (i.e., 15 km), one can still easily distinguish between modes, although there is some interference between them. However, at short ranges (i.e., 5 km), it becomes impossible to clearly distinguish between modes. Nonetheless, the strong, regular fluctuation in intensity vs frequency on the short-range spectrogram provides a clear indication that several modes exist in the signal, even if they are not distinctively separated in time. One may look at Crance *et al.* (2015) (their Fig. 3) for an experimental example of this phenomenon, where the considered signal is a fin whale vocalization recorded in the Arctic.

III. WARPING: THEORY

The main objective of warping is to facilitate modal separation of ocean waveguide signals, particularly for those produced at short- to mid-ranges. In a single receiver context, it is reasonable to resort to TF analysis to separate the modes. However, classical methods are not adapted to represent signals with non-linear time dependence, such as normal mode arrivals. As a result, warping can be used to “linearize” the modes, so that the transformed signal can then be conveniently processed using standard TF methods, such as the spectrogram.

A. Understanding warping

Warping is either a compression or stretching of a signal over time. In our context, the signal is recorded on a

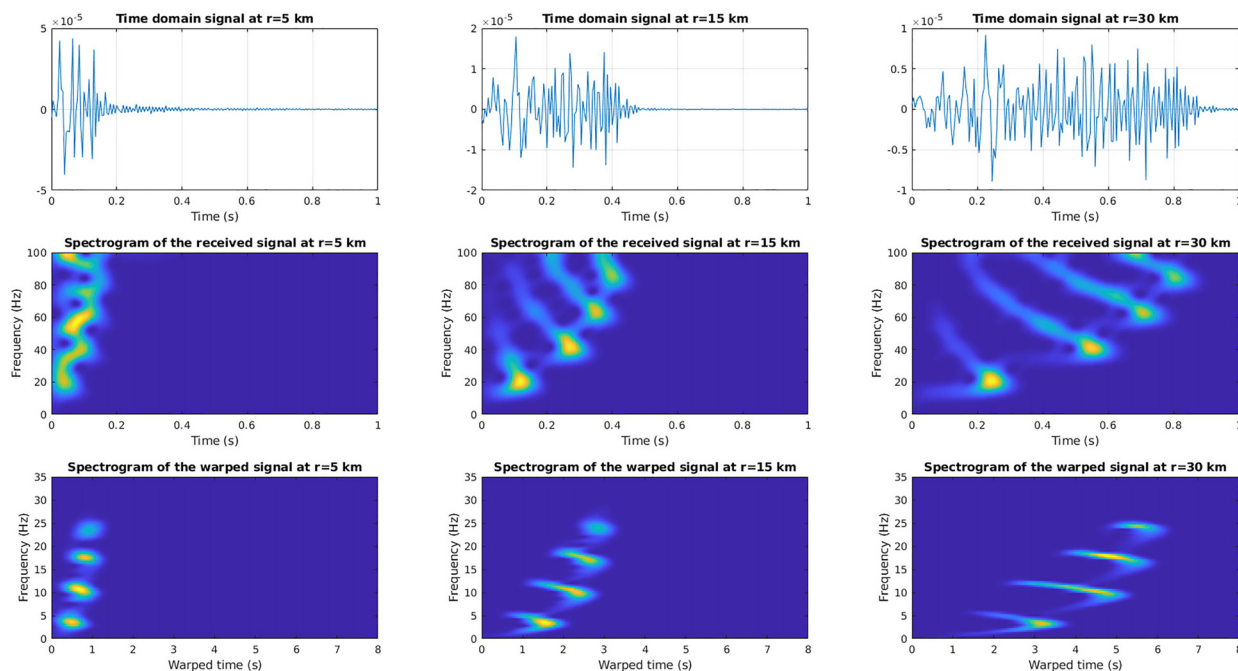


FIG. 4. (Color online) Simulated signal in the Pekeris waveguide at various ranges: 5 km (first column), 15 km (second column), and 30 km (third column).

single hydrophone as a function of time. As a result, the only way to compress/stretch the signal is to warp the time axis—or, equivalently, the frequency axis. If the signal were recorded on an array, one could also warp a spatial (or wave-number) axis. Mathematically, space and time warping are equivalent, but in this tutorial we only present time warping.

Before delving into the math, the warping effect on a continuous tone is illustrated in Fig. 5. The considered signal is color-coded, to better visualize the warping effect: the original signal contains an homogeneous ratio of colors. Figure 5(a) presents the original continuous tone signal. Figure 5(b) presents the signal after a linear compression. One can see that the signal duration is shorter, and the signal

is oscillating faster, i.e., frequency is higher. On the other hand, Fig. 5(c) presents the signal after a linear stretching. The signal duration is now longer, resulting in slower oscillations and thus a lower frequency. Although the lengths of the two warped signals [Figs. 5(b) and 5(c)] have been changed, the color ratio is still homogeneous. This is because the signals have been linearly (i.e., homogeneously) compressed/dilated.

Non-linear warping is illustrated in Fig. 5(d) (compression) and Fig. 5(e) (stretching). The result of the non-linear transformation is that the frequency of the warped signal varies with time. This is particularly evident in Fig. 5(e), where frequency increases as time evolves. This occurs

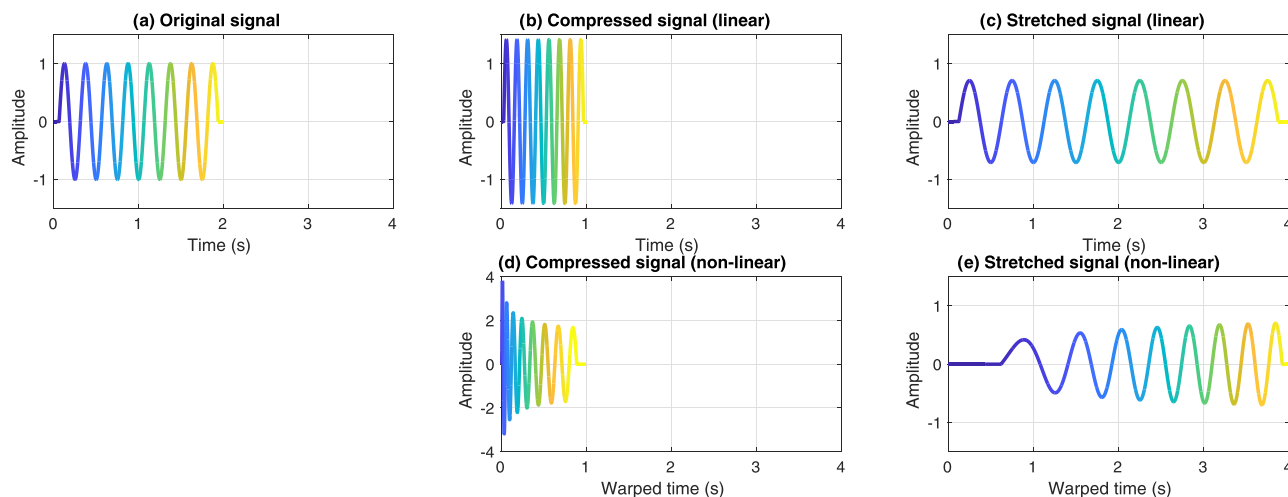


FIG. 5. (Color online) Illustration of warping. (a) Original signal, continuous tone. Linear warping, both (b) compression and (c) stretching. Non-linear warping, both (d) compression and (e) stretching. The scale of the time axis is the same for all the plots. The scale of the vertical axis is the same for all the plots, except for panel (d).

because the signal has been stretched non-uniformly over time. In the example presented here, the beginning of the signal has been stretched more than its end. This non-linear effect is further illustrated by the signal colors. Indeed, color ratio is homogeneous on the original signal [Fig. 5(a)] but becomes inhomogeneous after non-linear warping [Fig. 5(d) and 5(e)]. As an example, blue dominates the warped signal in Fig. 5(e).

One can also see from Fig. 5 that warping modifies the signal's amplitude [note that the scale of the vertical axis is the same for all panels but Fig. 5(d)]. As discussed below, the warping operation is designed so that it does not modify the overall energy of the signal. As a result, if a signal is shortened (compressed in time), its amplitude must increase in order to conserve energy. Consequently, the signal that has been linearly compressed [Fig. 5(b)] has a higher amplitude than the original signal [Fig. 5(a)]. For the same reason, the signal that has been linearly stretched [Fig. 5(c)] has a smaller amplitude than the original signal. For non-linear warping, the situation is more complex, and the amplitude must be adjusted over time, depending on the amount of compression/stretching that is occurring at that instant. One can see that the amplitude of the signal in Fig. 5(e) (non-linear stretching) is increasing with time. This is consistent with the previous observation, since the beginning of the signal has been stretched more than its end. In the same way, the amplitude of the signal in Fig. 5(d) (non-linear compression) is decreasing with time. This is because the beginning of the signal has been more compressed than its end.

B. Mathematical definition of warping

Mathematically, warping is a substitution. One replaces time t by something else, say $h(t)$, which is called a warping function. Considering an original signal $y(t)$, the warped signal $y_w(t)$ is obtained through

$$y_w(t) = \sqrt{|h'(t)|} y[h(t)], \quad (7)$$

where $h'(t)$ is the time derivative of the warping function. The factor $\sqrt{|h'(t)|}$ ensures energy conservation between y_w and y , it arises from integration by substitution while equating the energy of y with the energy of y_w . As a reminder, energy conservation has been discussed in Sec. III A, and can be visually observed in Fig. 5.

An important property of the warping function $h(t)$ is that it must be bijective (i.e., all points in a function are uniquely matched to all points in a second function). As a result, its (functional) inverse $h^{-1}(t)$ can be defined, and so warping can be reversed. Any signal warped with $h(t)$ can then be unwrapped using $h^{-1}(t)$ as the new warping function. In other words, there is a one-to-one correspondence between all points in y and y_w , so that it is possible to go from one to the other using either warping or inverse warping. In practice, warping (and inverse warping) are applied as a non-linear resampling, as detailed in Sec. IV.

The warping definition, Eq. (7), shows that the basic requirement for warping is to choose an appropriate warping function. To do so, one must remember the warping

objective (see the introduction of Sec. III), which is to “linearize modes,” or more generally, to linearize a signal. This is a shortcut for saying that we actually want to linearize the phase of the signal.

To understand this, let us consider a signal $y(t) = a(t)e^{j2\pi f_0 \Phi(t)}$, with $a(t)$ its amplitude, $\Phi(t)$ its (non-linear) phase, and f_0 a constant. To linearize the phase of $y(t)$ one must choose $h(t) = \Phi^{-1}(t)$ as the warping function. The warped signal becomes $y_w(t) = b(t)e^{j2\pi f_0 t}$, with $b(t) = \sqrt{|h'(t)|} a[h(t)]$ the amplitude of the warped signal, and $2\pi f_0 t$ its phase which is now linear. Going back to Fig. 5, one can imagine that the original signal with a non-linear phase is shown in Fig. 5(d) [or, equivalently, Fig. 5(e)], while the warped signal with a linear phase is shown in Fig. 5(a).

Before going further, note that our focus has been warping in time. However, one may also consider warping in frequency. In this case, one takes a signal in the frequency domain (i.e., after a Fourier transform), and then compresses/stretches the frequency axis (which is now the independent variable). This is not useful for the remainder of the tutorial, and thus will not be discussed further. However, an Appendix provides a comparison between time and frequency warping.

C. Dispersion based warping

In Sec. III B, we showed that an adequate choice of the warping function allows reaching the warping objective: linearizing the phase of the signal. However, to do so, one needs to know the phase of the signal $[\Phi(t)]$ in order to choose the correct warping function $[h(t) = \Phi^{-1}(t)]$. This, unfortunately, is a chicken and egg problem. One needs to know the expression of the phase to warp the signal. But if the phase is perfectly known, one usually does not need warping.

One way to circumvent the issue is to use an approximate model of $\Phi(t)$, knowing it is not perfect, but hoping it will be good enough for real-life applications. To do so in our underwater acoustics context, one needs to use knowledge about the underlying physics driving modal propagation.

However, our modal propagation context brings an extra complication. Several modes are propagating, and the phase of the modes is different from one mode to the next. As a result, one needs to either find a warping function that is adapted to every mode at once, or to define a different warping function for each mode.

As stated in the introduction of Sec. III, this tutorial focuses on the first option. We want to warp all the modes at once, so that the warped signal can be conveniently studied using a conventional spectrogram. The ideal results of warping are illustrated on the right side of Fig. 3. All the modes have been transformed into continuous tones and appear as horizontal lines in the TF domain.

1. Warping model: The isovelocity waveguide

The simplest model of shallow water propagation that captures the basic physics of dispersion is the so-called “ideal isovelocity” waveguide. It is a range-independent

waveguide with a perfectly reflecting surface, a constant water sound speed c_w , and a perfectly rigid seabed. In other words, the sea-surface is a perfect acoustic mirror, the water column is acoustically homogeneous, and the sound does not penetrate in the seabed.

The isovelocity waveguide is simpler than the Pekeris waveguide that was presented in Sec. II C, but less realistic. Its main interest here is that closed-form equations can be obtained for all the modal quantities. Of particular interest for us is the equation for modal phase, whose expression shows all modes can be warped using the same non-linear transform in time.

As a first step, let us write the received signal in the time domain as

$$y_{iso}(t) = \sum_{m=1}^{M_{iso}} a_{m_{iso}}(t) e^{j\phi_{m_{iso}}(t)}, \quad (8)$$

with M_{iso} the number of propagating modes in the ideal waveguide, $a_{m_{iso}}(t)$ the modal amplitude, and $\phi_{m_{iso}}(t)$ the modal phase.

For a source/receiver range r , the modal phase is given by

$$\phi_{m_{iso}}(t) = 2\pi f_{c,m} \sqrt{t^2 - t_r^2}, \quad t > t_r, \quad (9)$$

with $t_r = r/c_w$ (the earliest time at which any signal energy can reach the sensor), and $f_{c,m} = (2m - 1)c_w/4D$ (the cutoff frequency of mode m in the waveguide), and D the water depth.

Note that Eqs. (8) and (9) are given in the time domain, while the classical modal equations (2) and (3) are given in the frequency domain. Equation (8) is obtained from Eq. (2) using an inverse Fourier transform. The specific terms $a_{m_{iso}}(t)$ and $\phi_{m_{iso}}(t)$ can be analytically derived using a stationary phase approximation.

The interesting feature of the modal phase [Eq. (9)] is that it can be separated into a term ($2\pi f_{c,m}$) that depends on modal number m only, and a term $\xi(t) = \sqrt{t^2 - t_r^2}$ that depends on time t only. This specific separation of variables allows all modes to share the same warping transformation.

2. Warping function

To define the warping function let us assume propagation in an isovelocity waveguide and that the source is impulsive. In this case, the modal phase Eq. (9) can be used directly to define the warping function.

Following the derivation in Sec. III B, we will use the warping function $h(t) = \xi^{-1}(t)$, or

$$h(t) = \sqrt{t^2 + t_r^2}. \quad (10)$$

The function for inverse warping is obtained as $h^{-1}(t) = \xi(t)$, or

$$h^{-1}(t) = \sqrt{t^2 - t_r^2}. \quad (11)$$

Using $h(t)$ as the warping function, the warped signal becomes

$$y_{w_{iso}}(t) = \sum_{m=1}^{M_{iso}} b_{m_{iso}}(t) e^{j2\pi f_{c,m} t}, \quad (12)$$

with $b_{m_{iso}}(t) = \sqrt{|h'(t)|} a_{m_{iso}}[h(t)]$. Each warped mode is thus transformed into a continuous tone of frequency $f_{c,m}$, with amplitude $b_{m_{iso}}(t)$. What began as a series of frequency-modulated swept tones now appears as a set of parallel horizontal lines in the TF domain, as illustrated on the right side of Fig. 3.

Obviously, warped modes will be perfectly horizontal only if the signal under study has propagated within an actual ideal waveguide, which is never the case in reality. However, we have empirically observed that Eq. (10) is robust enough that when it is applied to real modes propagating in much more complex environments, the resulting transformed modes are still relatively tonal. This property is illustrated in the lower panels of Fig. 4 for propagation in a Pekeris waveguide. One can see that the warped modes are not the theoretically predicted pure tones (i.e., perfectly horizontal lines), but instead are tilted and slightly curved. Nonetheless, each warped mode is still separable using simple filtering methods. This robust behavior is characteristic of most shallow water experimental data, which will be illustrated through several examples below.

IV. WARPING: ALGORITHM

This section gives technical details about warping implementation. An important detail is that warping theoretically requires the knowledge of the source emission time, which is used as the time origin. Readers encountering warping for the first time can skip this section and go directly to Sec. V.

A. Numerical implementation of warping

This section summarizes all the formulas that enable a practical implementation of warping. A thorough derivation for these formulas is provided in Appendix C.

1. Warping

Let us consider a signal $y(t)$, with $t_{min} < t < t_{max}$. Its discrete version, sampled at frequency f_s , is denoted $y[n]$, with $n \in \llbracket 0; N - 1 \rrbracket$. The continuous warped signal, $y_w(t)$, is obtained through Eq. (7). We are interested here in its discrete version, denoted $y_w[k]$, with $k \in \llbracket 0; K - 1 \rrbracket$.

A convenient sampling frequency for the warped signal is $f_s^h = 2/\Delta t_N$, with

$$\Delta t_N = \frac{1}{f_s} \frac{t_{max}}{h^{-1}(t_{max})}. \quad (13)$$

The corresponding number of samples is

$$K = \text{ceil}\left(\left[h^{-1}(t_{max}) - h^{-1}(t_{min})\right] f_s^h\right), \quad (14)$$

with $\text{ceil}(x)$ the nearest integer greater than or equal to x . Finally, the k th sample of the warped signal is

$$y_w[k] = \sqrt{\frac{\bar{t}_k}{h(\bar{t}_k)}} y[h(\bar{t}_k)], \quad (15)$$

with $\bar{t}_k = k/f_s^h$, $h(\bar{t}_k) = (\bar{t}_k^2 + t_r^2)^{1/2}$, and the quantity $y[h(\bar{t}_k)]$ is obtained from the original discrete signal y through interpolation.

2. Inverse warping

Inverse warping can be seen as forward warping using $h^{-1}(t)$ as the warping function. However, the sampling frequency and the number of samples of the signal after inverse warping are already known: they are the same as for the original signal: f_s and N .

Let y_w denote the warped signal and y_u the unwarped signal, recovered from y_w using inverse warping. The n th sample of the unwarped signal is

$$y_u[n] = \sqrt{\frac{t_n}{h^{-1}(t_n)}} y_w[h^{-1}(t_n)], \quad (16)$$

with $t_n = t_{\min} + n/f_s$, $h^{-1}(t_n) = (t_n^2 - t_r^2)^{1/2}$, and the quantity $y_w[h^{-1}(t_n)]$ obtained via interpolation of the warped discrete signal y_w .

B. Example

As a simple example, let us consider a continuous tone and try warping. A 5 Hz sine wave, sampled at $f_s = 100$ Hz and lasting 2 s ($N = 201$) is plotted in Fig. 6. It is warped using the method presented above, using $t_r = r/c_w$, where $r = 10$ km and $c_w = 1500$ m/s. The warped signal, as well as the signal recovered after forward and inverse warping, are

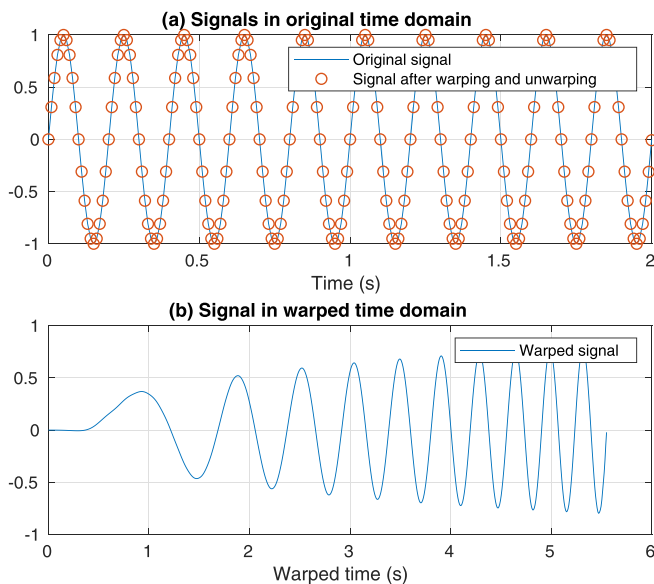


FIG. 6. (Color online) Example of warping and inverse warping on a tonal signal. Note that the time axes of the two subfigures are different.

also plotted in Fig. 6. One sees the perfect match between the original signal and the recovered signal, illustrating the reversibility of the warping operation.

Note that the time axis shown in Fig. 6(a) starts at $t = 0$ s. This is in violation of warping theory, which states that the signal of interest exists only for time $t > r/c$ [see Eq. (9)] which shows that the signal phase is not defined for $t \leq t_r$. However, in practice, one rarely knows the absolute time origin. A common practice is to set $t = 0$ as the initial sample of the signal of interest, as has been plotted here (and will be done for all the warping examples in this tutorial).

Nonetheless, the correct time origin needs to be used in the warping code. The algorithm provided as a supplementary material¹ uses a mathematically correct time axis, in that it assigns a time $t = r/c + 1/f_s$ to the first time sample. The accompanying plotting scripts, however, always plot the first sample as $t = 0$.

V. PRACTICAL CONSIDERATIONS WHEN WARPING

A. Warping parameters

The warping function, $h(t) = \sqrt{t^2 + t_r^2}$, requires the choice of a single parameter $t_r = r/c$. Since we will later be interested in source localization, one may think this involves circular reasoning, since range is something we eventually want to estimate, but the warping process needs an estimate of t_r , which in turn requires a range estimate. However, warping is always followed by inverse warping, which effectively removes any effect of the warping parameter $t_r = r/c$. As a result, any source localization result obtained with the extracted modes (which have been warped and unwarped) will be independent of t_r , and thus to the trial range chosen as a warping parameter.

In practice, warping results are only weakly sensitive to the choice of t_r , and it is not required to know the range nor the water sound speed to apply warping. All the signals presented in this tutorial, including the experimental ones, have been warped using $r = 10$ km and $c = 1500$ m/s (while true ranges are between 5 and 15 km). However, warping is much more sensitive to other factors, such as the choice of the time origin. This will be detailed below.

B. Impulsive signals and time origin identification

If the signal is impulsive, warping is straightforward, because the warping function has been defined with this assumption (see Sec. III C 2). An important step, though, is identifying the appropriate time origin.

As explained in Sec. IV B, one does not use the true time origin of the data. Rather, one identifies a time sample that corresponds to $t = r/c_w + 1/f_s$ and uses that as the start of the target signal. All the previous samples corresponding to $t \leq r/c_w$ are thus dropped before applying warping.

In an ideal waveguide without noise, no signal can exist before $t = r/c_w$. The acoustic energy only starts to arrive at $t = r/c_w$, and slowly decays for time $t > r/c_w$. In this case, identifying the appropriate sample to denote time origin of

warping is trivial: simply mark the time that the signal becomes non-zero.

However, real life is more complex, even if we still consider a noiseless environment. For example, the water column usually has a depth-dependent sound speed, so that c_w is not uniquely defined. Furthermore, acoustic energy also penetrates the seabed, which usually has a larger speed than the water column. As a result, even if the water column is isovelocity, acoustic energy will arrive at the receiver before $t = r/c_w$. This is illustrated in Fig. 7, which shows an impulsive signal simulated in a noisy Pekeris waveguide (see Sec. II C), with SNR = 20 dB. Time $t = r/c_w \simeq 6.67$ s and $t = r/c_b = 6.25$ s are identified with vertical lines. One can see acoustic energy before $t = r/c_w$, which corresponds to sound traveling through the seabed. These bottom arrivals are particularly visible for mode 1, with energy between 5 and 15 Hz at times between 6.3 and 6.75 s. The specific TF point where a dispersion curve bends is called the *Airy phase* (Jensen *et al.*, 2011, pp. 124–126). It is the last modal arrival, and thus corresponds to the group speed minimum. The acoustic energy at frequencies below the Airy phase mostly propagates within the seabed and is usually called a *ground wave*. It is thus highly attenuated and usually difficult to see in real data. However, in this case Airy phase and ground waves of mode 1 are clearly visible, they are respectively denoted by a black cross and a black ellipse in Fig. 7(b).

Regardless of these complications, when warping a signal one needs to make a guess and pick a time origin that is as close to $t = r/c_w$ as possible. We suggest using an iterative trial and error process. First, look at the original signal (both time domain and spectrogram) and try to assess the time of arrival of the highest frequencies, which tend to experience the least dispersion and thus approach t_r . Then, warp the signal using this time origin estimate. Last, verify the warping results by looking at the spectrogram of the

warped signal. If needed, change the time origin and iterate. Usually, a few iterations are enough to obtain a “adequate” warped signal Bonnel *et al.* (2017). By “adequate” we mean that the warped modes are relatively horizontal and cleanly separated on the warped spectrogram, and can thus be identified and filtered (see Sec. VI A).

The influence of time origin on warping is illustrated in Fig. 8. The signal considered here is the one shown in Fig. 7. Three different time origins are considered here: an accurate one which corresponds to $t = r/c_w$, an early one ($t = r/c_w - dt$), and a late one ($t = r/c_w + dt$), where $dt = 0.1$ s, which is about 10 samples. Such a large dt is probably too extreme for a realistic mistake, but it nonetheless provides intuitive insight. Even when using the correct time origin [Fig. 8(b)], one can see that the modes are not perfectly horizontal, due to the mismatch between the waveguide environment (a Pekeris waveguide), and the idealized analytical model used for warping (an ideal waveguide with a rigid bottom). However, the warped modes have become well-separated, when compared to the spectrogram in Fig. 7. When using a time origin that is too early [Fig. 8(d)], the modes are definitely not horizontal tones, but span a wider bandwidth across the warped spectrum. They thus overlap and interfere with each other, which will negatively impact further filtering. Finally, when the time origin is later than t_r [Fig. 8(f)], the modes become virtually horizontal. In this specific case, delaying the time origin helps compensate for the mismatch between the environment and the warping model, and is an ideal choice for separating modes 2 and higher. However, this improved separation comes at a price: mode 1 has vanished in the warped spectrogram. In shallow water propagation the first mode is generally the least impacted by dispersion, and as a result most of its energy arrival is close to t_r . It is thus typically removed from the signal if a late time origin is chosen.

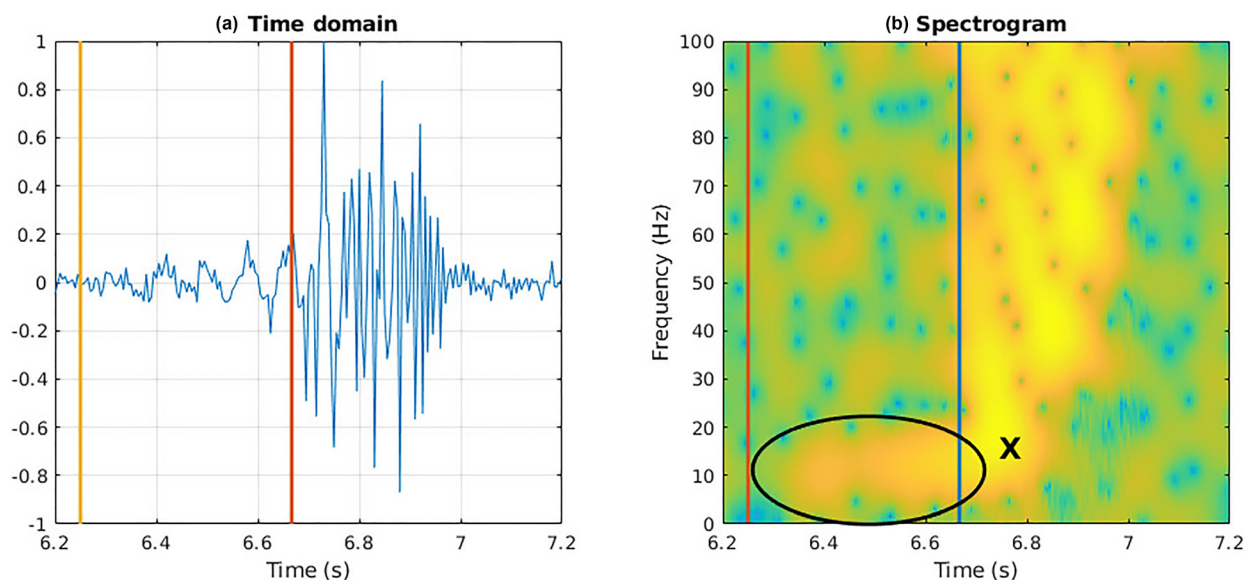


FIG. 7. (Color online) Impulsive signal simulated in a noisy Pekeris waveguide, with SNR = 20 dB. (a) Times series of the simulated signal and (b) corresponding spectrogram. On (b), the black cross shows the Airy phase of mode 1, and the black ellipse the corresponding ground wave.

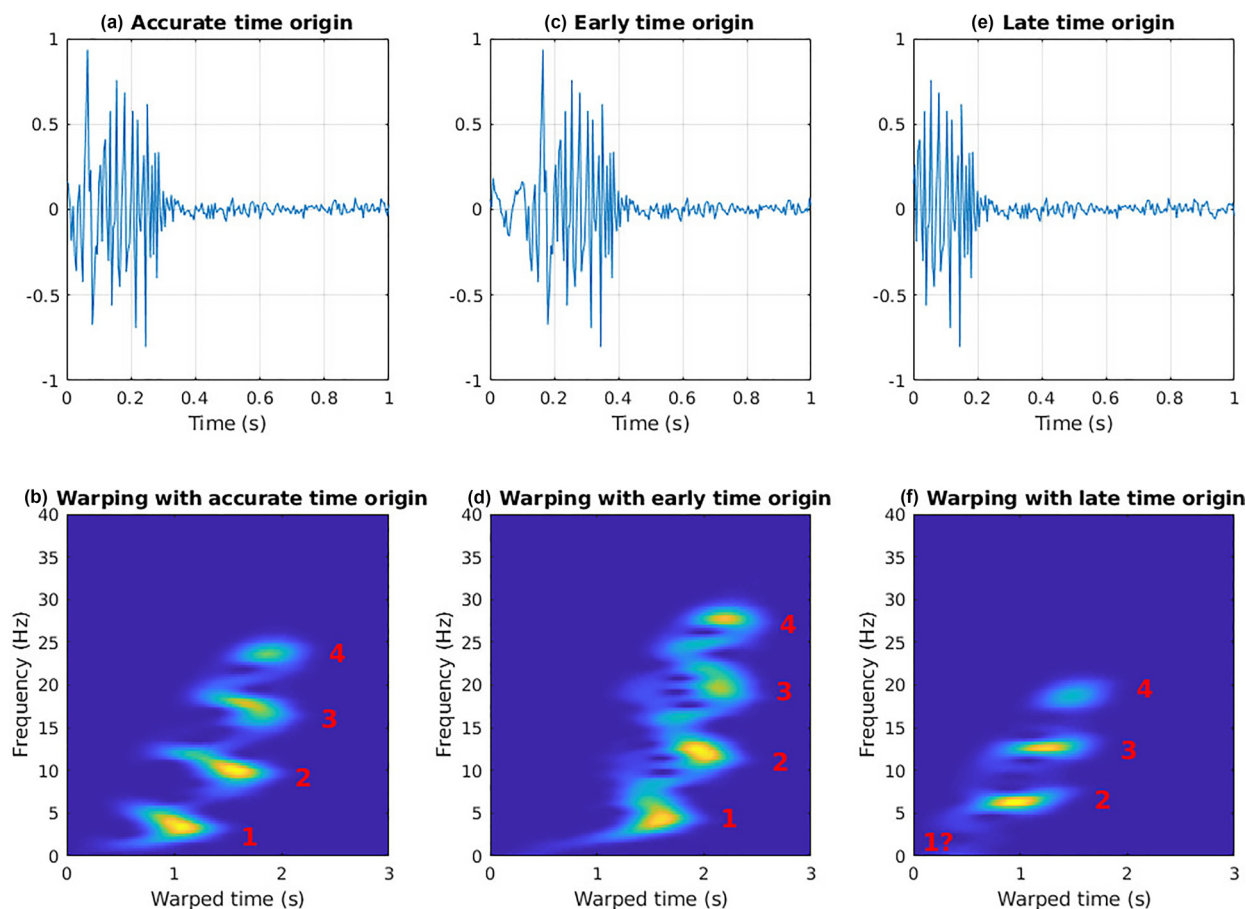


FIG. 8. (Color online) Impact of time origin on warping. (a) Time series with accurate time origin and (b) spectrogram of the corresponding warped signal. (c) Time series with an early time origin and (d) spectrogram of the corresponding warped signal. (e) Time series with a late time origin and (f) spectrogram of the corresponding warped signal.

In a real life scenario, we advise choosing the latest time origin that provides a decent modal separation without clipping mode 1. In certain cases, it may be useful to warp twice: the first using an early time origin for mode 1, and second using a later time origin for the other modes. When doing this though, one must be careful to time-align the recovered modes correctly.

C. Non-impulsive signals and source deconvolution

If the signal is not originally impulsive, it is necessary to compensate for the signal structure. In principle, if one had a very good knowledge of the source signal, one could perform source deconvolution. The basic idea is to divide Eq. (2) by $s(f)$ in order to completely cancel frequency-dependent variations arising from the source signal.

When $s(f)$ is well known, source deconvolution is straightforward. The only concern is to prevent division by zero (or small values), at frequencies where $s(f)$ is (near) null. The easiest and most classical deconvolution method is probably (Clayton and Wiggins, 1976)

$$y_{deconv}(f) = \frac{y(f)s^*(f)}{\max\{|s(f)|^2, \epsilon\}}, \quad (17)$$

with ϵ a small number, which is usually chosen as a small percentage of $\max\{|s(f)|^2\}$. This method is often called “water-level deconvolution”, with ϵ the water level parameter.

The importance of source deconvolution is illustrated below, using our noisy Pekeris waveguide from Fig. 7. However, here we incorporate a more complex source signal: a set of three perfect impulses with decreasing amplitude, and a separation of 75 ms between pulses. Although not fully realistic, this model captures key features of several controlled sources that are used for acoustical oceanography, such as light-bulbs (Heard *et al.*, 1997), combustive sound sources (CSS) (McNeese *et al.*, 2010), or explosives (Chapman, 1985). All these source systems produce an important impulsive signal that is often followed by secondary weaker impulsive signal(s), usually called bubble pulse(s).

The simulation results are illustrated in Fig. 9. The noisy impulse response and the corresponding warped signal are shown in the first column. Note that Fig. 9(b) is exactly the same as Fig. 8(b), in that we are using an accurate time origin, which allows us to visually assess the effect of source deconvolution.

The simulation that takes into account the source signal is shown in the second column of Fig. 9. One can see in the

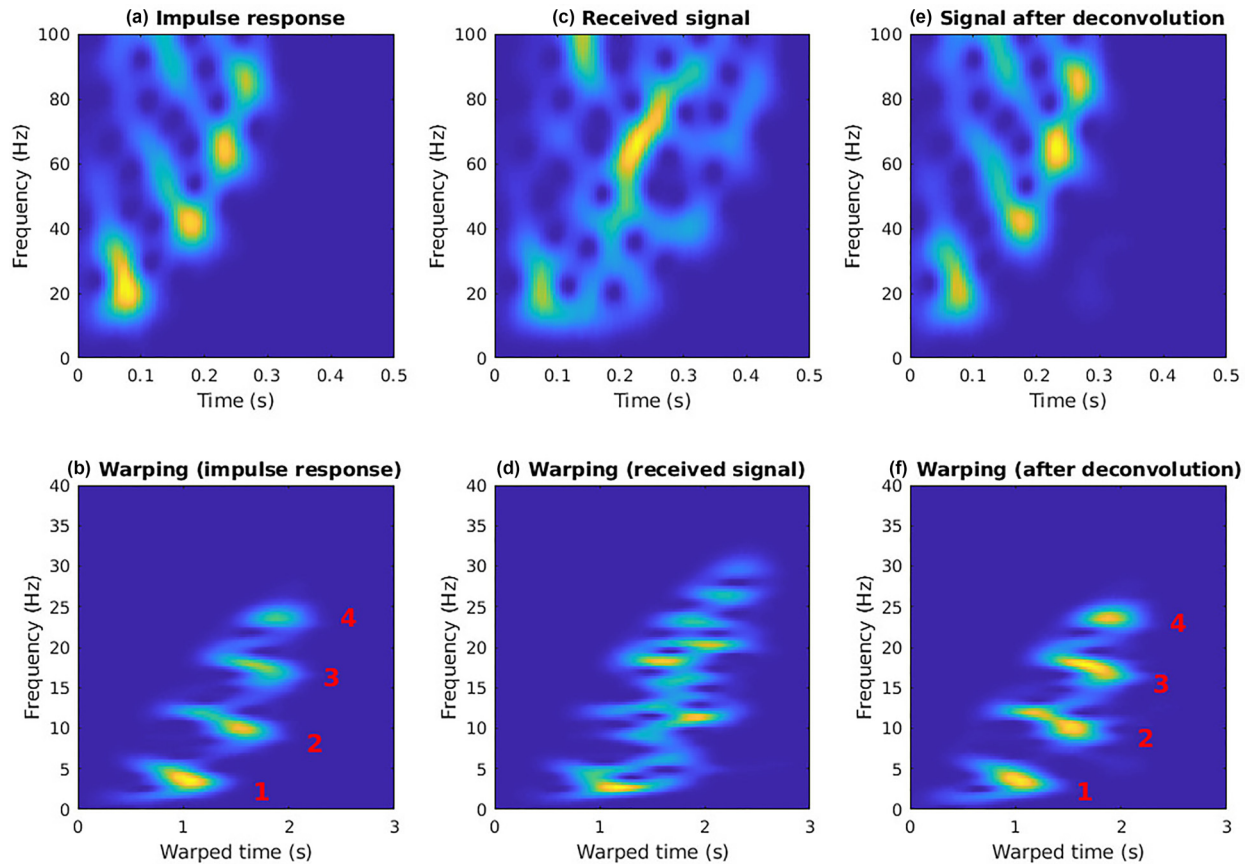


FIG. 9. (Color online) Importance of source deconvolution. (a) Spectrogram of the impulse response and (b) spectrogram of the corresponding warped signal. (c) Spectrogram of the received signal when the source is an impulse followed by bubble pulses, and (d) spectrogram of the corresponding warped signal. (e) Spectrogram of the signal after source deconvolution and (f) spectrogram of the corresponding warped signal.

original spectrogram that modal dispersion is barely visible. This is because modes from the main impulse are mixed with modes from the bubble pulses. This is further confirmed by the spectrogram of the warped signal, where modes cannot be identified.

As a last step, source deconvolution is applied, and results are illustrated in the third column of Fig. 9. One can now clearly see the modes, with spectrograms that are virtually equivalent to those obtained when working on the impulse response.

For *in situ* signals, source deconvolution requires a good measurement (or model) of the source signal. In a warping context, it has been successfully applied on light-bulb (Duan *et al.*, 2016) and CSS data (Bonnel *et al.*, 2019; Bonnel *et al.*, 2018), with $\epsilon \in [0.01, 0.1]$.

D. Frequency modulated signals and phase compensation

Source deconvolution is a powerful tool but it requires an accurate knowledge about the source signal structure, which is often not the case for bioacoustic signals. Here, we present an alternative deconvolution approach for frequency modulated (FM) source signals, where the precise nature of the FM signal is unknown. We find, practically speaking, that only a rough estimate of modulation phase is needed

and often can be estimated from the received signal. This rough estimate can subsequently be used to perform phase compensation, and then warping.

Even if one does not know the true frequency-dependent amplitude of the source signal, if one knows the source phase $\phi_s(f)$ or, equivalently, its time-frequency law $\tau_s(f) = (1/2\pi)[\partial\phi_s(f)/\partial f]$, then it is possible to compensate for the source phase in the received signal by computing (Bonnel *et al.*, 2014)

$$y_{pc}(f) = y(f)e^{-j\phi_s(f)}. \quad (18)$$

Such a process is not a true source deconvolution, as the amplitude of the received signal has not been corrected. However, the phase of $y_{pc}(f)$ has become the same as the phase of the channel impulse response. Warping can thus be applied on $y_{pc}(f)$.

In a real-life scenario where the source phase is unknown, the important point is to estimate $\phi_s(f)$ from the data. It is actually done by estimating the source TF law $\tau_s(f)$, which is later converted to $\phi_s(f)$ so that phase compensation can be performed.

To do so, a rough estimate of $\tau_s(f)$ can be obtained by manually tracing the arrival of mode 1 (or the lowest-order mode) on the received spectrogram. Since mode 1 is generally the least affected by dispersion, its TF shape on the

received signal usually provides a good initial estimate of the source TF law. In practice, $\tau_s(f)$ is usually approximated by a piecewise linear function, which is defined by manually selecting several TF points on the spectrogram, and then linearly interpolating between them.

A MATLAB code to perform this procedure is provided in the supplementary material.¹ Its use is illustrated here on a simple example. Once again, we re-use the noisy Pekeris waveguide model. The spectrogram of the environmental impulse response, as well as the spectrogram of the corresponding warped signal are presented in Figs. 10(a) and 10(b). These spectrograms are the same as those presented earlier, serving as a reference to visually assess the efficacy of phase compensation.

The source signal considered here is a non-linear FM sweep of constant amplitude. The source time-frequency law $\tau_s(f)$ is shown in Fig. 10(c) as a red curve. The spectrogram of the received signal, which combines both environmental dispersion and the source sweep, is shown in Fig. 10(c). The source $\tau_s(f)$ effectively tilts the whole TF pattern, as was explained in Sec. IID. The spectrogram of the corresponding warped signal is unusable [Fig. 10(d)].

As explained above, to perform phase compensation, we assume that the source TF law can be approximated by a

piecewise linear function. In this case, we show that a single linear piece is (nearly) enough to filter the modes. To perform phase compensation, we assume that the source TF law is a linear frequency modulation that goes from 100 to 0 Hz in 0.8 s, as illustrated by the black line in Fig. 10(c). The spectrogram of the signal after phase compensation is shown in Fig. 10(e), which one can see is qualitatively similar to the spectrogram of the impulse response [Fig. 10(a)]. The spectrogram of the corresponding warped signal is shown in Fig. 10(f), which demonstrates how the modes have been really well separated, although mode 3 is less tonal than in Fig. 10(b).

This result illustrates that warping is robust to uncertainty in the source signal, provided that the source signal has a relatively simple frequency-dependent phase. It can be used to separate modes when little information is available about the source, apart from the fact that it is frequency modulated. Warping is thus perfectly adapted to the study of signals like those of baleen whale vocalizations in shallow water. It has notably been applied to localize bowhead whale calls in the Arctic (Bonnell *et al.*, 2014; Warner *et al.*, 2016), which are frequency modulated signals with an unknown TF law (and thus an unknown phase).

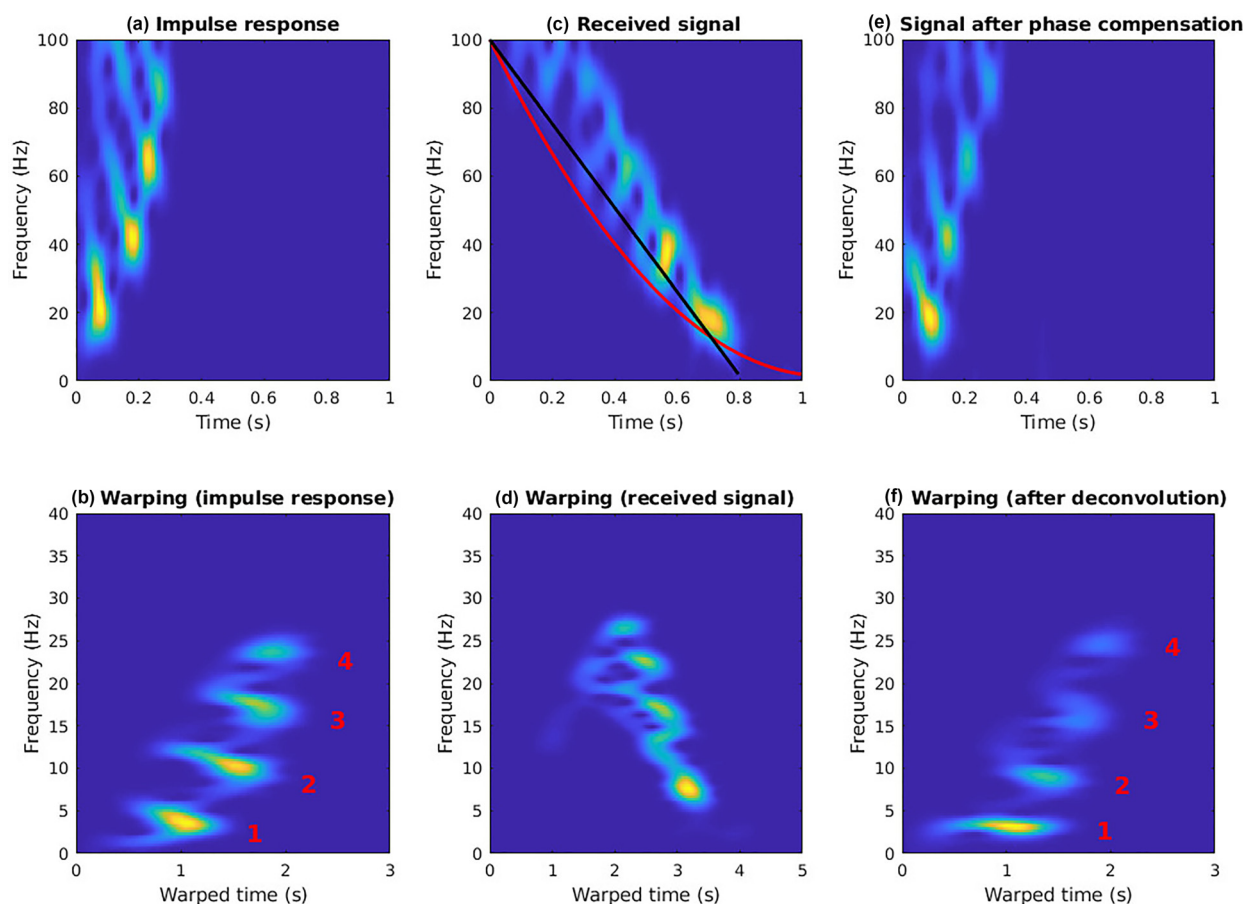


FIG. 10. (Color online) Importance of phase compensation. (a) Spectrogram of the impulse response and (b) spectrogram of the corresponding warped signal. (c) Spectrogram of the received signal when the source is a non-linear FM sweep, and (d) spectrogram of the corresponding warped signal. (e) Spectrogram of the signal after phase compensation and (f) spectrogram of the corresponding warped signal. On (c), the red curve is the true source instantaneous frequency, while the black line is the empirical estimate used to perform phase compensation.

VI. APPLICATIONS

Section V explained how to practically apply warping in various scenarios. In this section, we assume that warping has been correctly applied and we present some examples of post-processing applications. We show how warping can be used to filter modes and/or estimate a dispersion curve. Then we illustrate how estimated dispersion curves can be used for source ranging in a waveguide.

A. Modal filtering and dispersion curve estimation

We once again consider our favorite signal simulated in the noisy Pekeris waveguide, as in Sec. V. Choosing the time origin as best as possible, we thus reproduce the scenario presented in Fig. 8(a) (although the noise realization is not identical).

1. Warping

Because we want to illustrate the whole processing chain, we reproduce the warping steps presented earlier. The time series of the signal under study is presented in Fig. 11(a). The corresponding spectrogram is shown in Fig. 11(b). This spectrogram is not useful for modal filtering, but it will be used later to visually assess the quality of the filtering result. The received time series is then warped, and the result is shown in Fig. 11(c), along with its spectrogram in Fig. 11(d).

Before going further with post-processing, it is of paramount importance to obtain the best possible spectrogram

for the warped signal. This can be quantified by the separation of the warped modes: it is required to have them as horizontal—or at least as separated from each other—as possible. For real life scenarios, if the source is impulsive or if source deconvolution has been used, we advise trying several time origins (cf. Sec. VB), and choosing the iteration that gives the best spectrogram for the warped signal. A few iterations (~ 1 – 5) are usually adequate.

If the source has an unknown frequency modulation, then one has another degree of freedom to estimate, which is the source TF law. Once again, we advise an iterative trial-and-error process, iterating both on source TF law and time origin, until the warped modes are separated on a spectrogram. This process, while relatively cumbersome for warping beginners, can often be performed within a few minutes by experienced users. A click-and-play routine to do so is provided in the supplementary material.¹

2. Time-frequency filtering of warped modes

The next step is to filter a warped mode using TF filtering. TF filtering is easily done by defining a *mask* in the TF domain, which is 1 in an area of interest \mathcal{A} , and 0 elsewhere. Formally, we define $\mathcal{M}(t, f)$, with $\mathcal{M}(t, f) = 1$ if $(t, f) \in \mathcal{A}$, and $\mathcal{M}(t, f) = 0$ if $(t, f) \notin \mathcal{A}$. The TF representation under study is then multiplied by $\mathcal{M}(t, f)$, which effectively isolates the content of the signal that is in \mathcal{A} (Kozek and Hlawatsch, 1992). This is similar to basic bandpass filtering

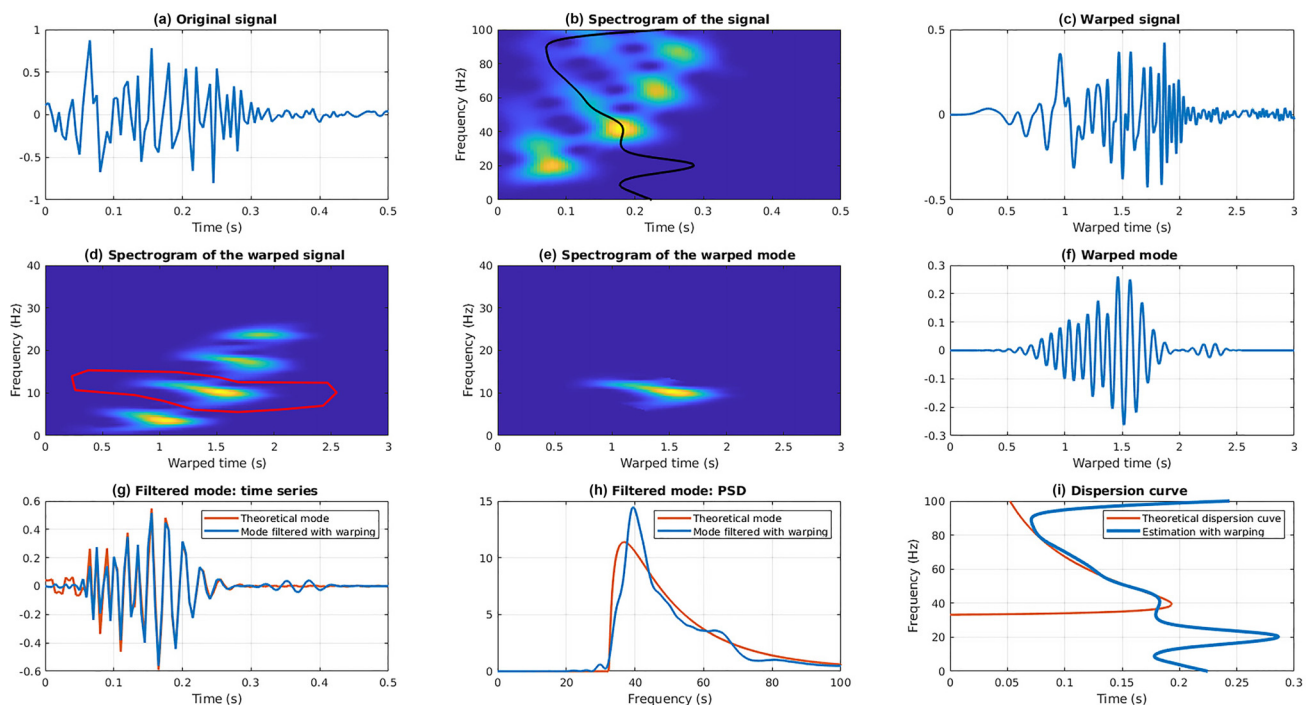


FIG. 11. (Color online) Modal filtering. (a) Time series of the received signal. (b) Spectrogram of the received signal. (c) Time series of the warped signal. (d) Spectrogram of the warped signal. The limit of the TF mask that is used to filter mode 2 is illustrated as a red polygon. (e) Spectrogram of warped mode 2. (f) Time series of warped mode 2. (g) Time series of mode 2 at the end of the processing chain (warping, filtering, inverse warping), and comparison with the theoretical time series. (h) Power spectral density (PSD) of mode 2 at the end of the processing chain, and comparison with the theoretical PSD. (i) Estimated dispersion curve for mode 2, and comparison with the theoretical dispersion curve. The estimated dispersion curve is also superimposed as a black curve on the original spectrogram [panel (b)].

performed in the frequency domain, except that a TF region is isolated instead of a simple frequency band.

For this application, TF filtering is performed on the spectrogram of the warped signal. The only trick is to define a region \mathcal{A} that isolates a warped mode. Although automatic processing can be used, we suggest initially defining the mask manually by looking at the spectrogram; a click-and-play routine to do so is provided in the supplementary material.¹ An example of a mask for filtering mode 2 is illustrated in Fig. 11(d) as a red contour: the area within the contour is the area \mathcal{A} . When the spectrogram of the warped signal is multiplied with this mask, the spectrogram of a single warped mode is obtained, as illustrated in Fig. 11(e).

The manual definition of the mask introduces another subjective operation in the modal filtering process. However, this one is relatively benign: if warping has been correctly performed, then the modes are well separated, and the exact selection of the mask does not matter. As an example, this is the case in Fig. 11(d). As long as the chosen mask is wide enough to encompass a mode, yet small enough not to overlap with another mode, then the TF filtering should be sufficient. In real life scenarios, one may still need to iterate this operation a few times. We do advise checking the quality of modal filtering by estimating the modal dispersion curve and comparing it with the original spectrogram, as illustrated in Fig. 11(b). The process to do so is explained below.

3. Isolating modes in the original time domain

After the masking process, we have a single warped mode in the TF domain. The following steps are to invert all the previous operations, except filtering. We first go from the TF domain to the time domain by computing an inverse short-time Fourier transform. The result, illustrated in Fig. 11(f), is the time series of a single warped mode. Such a warped mode can then be unwrapped using inverse warping, which leads to the time series of a single mode. This is the end result of modal filtering. The whole processing chain extracts the time series of a single mode from the original signal that combines all the modes.

The filtered mode 2 is presented in Fig. 11(f), which also shows the theoretical time series, obtained by direct simulation. There is an excellent match between the filtered and theoretical mode time series and their power spectral densities (PSD), as shown in Fig. 11(g). For practical applications, this match could be quantified using a normalized mean squared error, as is traditionally done in estimation theory.

The filtering result presented in Fig. 11 is not perfect: the amplitude of the filtered mode is slightly oscillating and the filtered mode peaks at a frequency different from the true one. This likely results from interference with other modes that have not been completely rejected by warping. Also, the filtered mode seems to have a cutoff frequency that is slightly too high. This is because the theoretical mode contains energy between the true cutoff and the Airy phase. This part of the mode is not modeled by warping, and thus is nearly impossible to filter. However, this part of the mode is

generally highly attenuated during propagation, and typically barely visible on real data (except when very powerful sound sources are used). This is unlikely to cause trouble to users that are interested in marine mammal localization, but is a known drawback for users that are interested in geoaoustic inversion.

4. Dispersion curve estimation

Once modes have been filtered, the game is won. As stated in the Introduction, filtered modes can be used as input data for many applications. However, one last processing step can be performed before turning to further applications: estimation of the modal dispersion curve (the TF positions of the modes). This is trivial to do once modes are filtered, as filtered modes are simple mono-component signals. The exercise is thus to estimate the instantaneous frequency (or group delay) of a mono-component signal. Many methods are available to do so (Boashash, 2015). We suggest computing the average time (first frequency moment) of the spectrogram of the filtered mode. A simple MATLAB code to do so is provided in the supplementary material.¹

Going back to the previous example, the estimated dispersion curve of mode 2 is presented in Fig. 11(i) and compared to the theoretical one. The estimation has been performed over the entire signal bandwidth, but the modal dispersion curve must be restricted to a frequency band of interest. Here, we have an excellent match over 45–80 Hz, which is the band where the mode is mostly energetic.

Restricting the dispersion curve to a frequency band of interest can be done automatically by setting a threshold on the filtered mode PSD, or manually by assessing the frequency band where there is a good match between estimated dispersion curve and spectrogram [see Fig. 11(b)]. Since warping requires several manual operations anyway, we advise doing this manually as well. In any case, even if an application does not require dispersion curve estimation, we still advise estimating the dispersion curves and comparing them with the original spectrogram. This is an easy quality check for assessing modal filtering performance for real life applications.

5. Further technical details

As a technical side note, let us look further at the errors between the estimated mode and the theoretical one. Visually, it is clear that these errors are quite small [Figs. 11(g), 11(h), 11(i)]. However, it is also clear that the error statistics are not trivial. In particular, the errors are correlated. As an illustration, one can look at the estimated dispersion curve on Fig. 11(i). The mode arrival time is consistently over-estimated between 60 and 70 Hz, which shows the high correlation of the error, at least in this frequency band. This must be taken into account for applications that require a proper statistical uncertainty characterization, such as Bayesian geoaoustic inversion.

Also, one can see that in the frequency band where the mode is energetic, the estimated dispersion curve [Fig. 11(i)] is better estimated than the mode amplitude [Fig. 11(g)],

which notably peaks at a wrong frequency. This probably explains why most successful applications of warping—and particularly environmental inversion which requires accurate environmental data—are based on estimated dispersion curves, rather than on raw filtered modes. In other words, warping is successful in recovering accurate relative timing between modes (phase), less so when extracting relative or absolute modal amplitudes.

Last but not least, modal filtering has been illustrated here for mode 2. To filter several modes, the process must be iterated. Usually, there is no need to warp the original signal several times, one can simply design a new mask corresponding to another mode [cf. Fig. 11(d)]. Still, on rare occasions, it may be necessary to change time origin and restart the modal filtering process from scratch (see Sec. VB).

B. Source localization

In this section, we illustrate how the modal dispersion curve, $\tau_m(f)$, can be used to solve an inverse problem. Because this is not the main focus of the tutorial but a mere illustration of warping capability, we keep this illustration as simple as possible. We thus focus on source localization, and more specifically the estimation of the source/receiver range. More advanced applications are covered by the citations given in Appendix B.

The concept of source localization using dispersion curves is relatively easy. On one hand, we have modal dispersion curves that have been estimated from the received signal; these will be called *data*. On the other hand, we have a propagation model that enables the simulation of dispersion curves; these will be called *replicas*. To localize the source, the idea is to iterate over many candidate source positions (as well as other environmental parameters if needed), and to quantify the degree of match between data and replicas. The position of the source is estimated by selecting the position of the simulated source that provides the best match between data and replicas.

Complex propagations models can (and should) be used to compute the replicas for optimal source localization. However, this is not within the scope of this tutorial, so a simple source localization will be illustrated using the Pekeris waveguide (see Sec. IIC) as an environmental model. We further assume that water depth D and water sound speed c_w are well-known, which is realistic for real life applications, although c_w usually depends on depth. The seabed sound speed c_b is unknown, and will be estimated as part of the inversion problem. The seabed density is also unknown, but it is arbitrary fixed at $\rho_b = 1600 \text{ kg/m}^3$ to keep the inversion simple. Last, but not least, source and receiver depths are considered unknown. However, a feature of acoustic waveguide propagation is that the source/receiver depths do not influence the modal phase term and thus do not impact the dispersion curves, so they can thus be ignored. As a result of this logic, two parameters will be estimated as part of the inversion process: source/receiver range r and seabed sound speed c_b . Estimated values are

noted with an underlying hat: \hat{r} and \hat{c}_b . Note that it is not expected for \hat{c}_b to be realistic. As in Collins and Kuperman (1991), the idea is to adapt the environmental model to help localization, not to perform a proper seabed geoacoustic inversion.

To formalize the localization method, we denote the data as $\tau_m^{data}(f)$ and the replicas as $\tau_m^{rep}(f, r, c_b)$. Note that replicas depend on r and c_b , as those are needed to compute the simulated dispersion curves. If the source is impulsive, the match between data and replicas can be directly quantified. However, the source and receiver are generally not synchronized. There is thus the need to include an unknown time shift parameter dt within the data/replicas comparison. Formally, using a simple least square fit,

$$[\hat{r}, \hat{c}_b, \hat{dt}] = \arg \min_{[r, c_b, dt]} \times \sum_m \sum_f (\tau_m^{data}(f) - \tau_m^{rep}(f, r, c_b) + dt)^2. \quad (19)$$

If the source is not impulsive but its waveform is known, then it can be localized using Eq. (19) after source deconvolution. If the source is an unknown frequency modulation, one cannot use Eq. (19) anymore. Indeed, the replicas now depend on the unknown source modulation $t_s(f)$, as shown in Eq. (6). One way to go around this issue is to match dispersion curve differences instead of dispersion curves,

$$[\hat{r}, \hat{c}_b] = \arg \min_{[r, c_b]} \sum_{m \neq n} \sum_f ([\tau_m^{data}(f) - \tau_n^{data}(f)] - [\tau_m^{rep}(f, r, c_b) - \tau_n^{rep}(f, r, c_b)])^2, \quad (20)$$

with $\tau_m^{rep}(f, r, c_b)$ replicas that are computed as if the source were impulsive, as in Eq (19).

Note that Eq. (20) can also be used to localize impulsive sources. Its benefit is that it removes the need to invert for dt . Its drawback is that it requires that the data contain at least two modes with a common frequency band, while Eq. (19) can be applied to a single mode, and/or to several modes that do not overlap in frequency.

The last step for source localization is to correctly perform the minimization involved in both Eqs. (19) and (20). This can be done using advanced optimization algorithms (Bonnans *et al.*, 2006). However, this is unnecessary here, as only 2 or 3 parameters are considered. The minimization can thus be done using a grid search. A simple MATLAB code to do so is given in the supplementary material.¹ Source localization (and warping) will be illustrated on experimental examples in Sec. VII.

VII. EXPERIMENTAL EXAMPLES

This section provides three examples of warping applications for experimental signals: an impulsive right whale vocalization, a controlled tomography source with known waveform, and a frequency modulated vocalization from a

bowhead whale whose TF law is not known *a priori*. They have been chosen to illustrate warping approaches that parallel the circumstances outlined in Sec. V.

The data and MATLAB code associated with these examples are provided as supplementary material.¹ Note that these examples are extracted from previous publications by the authors, and thus more detailed background information about the data and scenarios are available (Bonnell *et al.*, 2014; Bonnell *et al.*, 2018; Thode *et al.*, 2017).

A. Right whale gunshot

We first illustrate warping on a sound that is a pure impulse: the so-called “gunshot” sound (Crance and Berchok, 2016; Parks and Tyack, 2005) emitted by a North Pacific right whale. The data example used here was recorded in 2013 in the southeastern Bering Sea federally designated right whale critical habitat by the Alaska Fisheries Science Center, Marine Mammal Laboratory (Wright, 2017). This signal has been previously studied in Thode *et al.* (2017). It has notably been localized using single receiver matched mode processing (joint localization/environmental inversion): its estimated range from the hydrophone on which it was recorded is 8.7 km [Thode *et al.* (2017), last line of their Table I].

The spectrogram of the received signal is presented in Fig. 12(a). It shows modes that are quite well separated; indeed, modes could probably be filtered and isolated by a masking process applied on the original spectrogram. This is not always the case for baleen whale signals, but the call is used here to provide an easy means for a user to check that the warping process has been applied correctly. The spectrogram of the warped signal is presented in Fig. 12(b). It is interesting to see here that the warped mode 1 is widely spread in frequency, while the warped mode 4 is higher in frequency and clearly isolated from the other modes. This is a strong indication that the experimental environment differs

substantially from the isovelocity waveguide assumption used for warping. Nonetheless, modes can still be filtered, and the TF masks that are used are shown in Fig. 12(b). The corresponding estimated dispersion curves are shown in Fig. 12(a). They perfectly match the underlying spectrogram, illustrating the success of the modal filtering operation.

The estimated dispersion curves are then used as an input for localizing the gunshot. As the source is assumed to be impulsive, Eq. (19) is used for localization. The environmental model used to compute the replicas is a Pekeris waveguide with $D = 51$ m, $c_w = 1450$ m/s, $\rho_w = 1000$ kg/m³, and $\rho_b = 1600$ kg/m³. The other parameters required to compute the replicas (r , c_b and dt) are included in the inversion. The search space is as follows: $r \in [2, 16]$ km with 100 m steps, $c_b \in [1550, 2000]$ m/s with 10 m/s steps, and $dt \in [-7, -5]$ s with 0.01 s steps. Note that if time origin has been set properly for warping, then the expected value for dt is about r/c_w . One can start with wide search bounds and coarse steps for dt , and gradually narrow the bounds and decrease the steps.

Localization results are illustrated in Fig. 13. Figure 13(a) shows the match between the data and predicted replicas (i.e., replicas computed using the optimal localization/environmental parameters). The match is not perfect but looks good enough. It is unlikely that a better match can be obtained, as the experimental environment is complex, and the inversion is performed with a simplistic Pekeris waveguide. A one-dimensional slice of the least square fit is presented in Fig. 13(b). It shows a smooth global minimum, which suggests that the minimization was successful. The estimated range is $\hat{r} = 8.8$ km. It is consistent with results obtained by the detailed environmental inversion in Thode *et al.* (2017), illustrating that the simple procedure obtained here is enough to do as well (or as poorly) as more complex methods. Such a result is obtained because dispersion curves are highly sensitive to range, and relatively less sensitive to many environmental details, which makes dispersion curve inversions relatively robust.

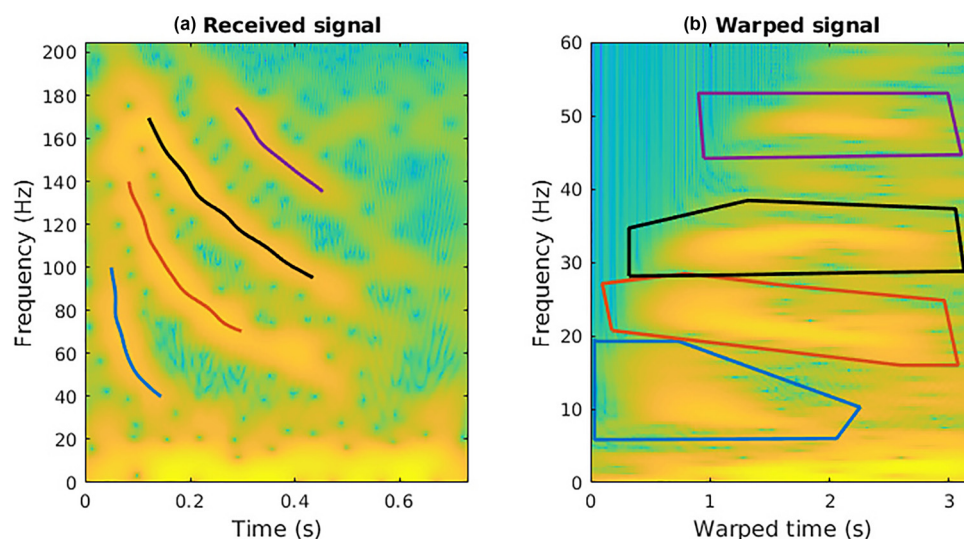


FIG. 12. (Color online) Gunshot signal analysis. (a) Spectrogram of the received signal and estimated dispersion curves. (b) Spectrogram of the warped signal and masks that are used to filter the modes.

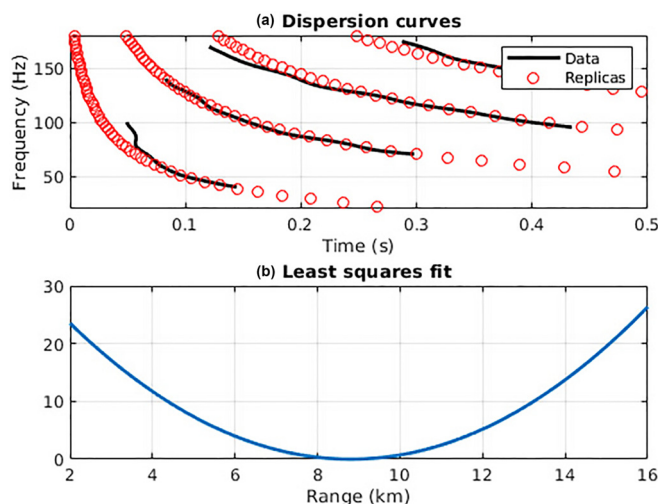


FIG. 13. (Color online) Gunshot localization. (a) Data and replicas. (b) One-dimensional slice of the least square fit, with $c_b = \hat{c}_b$ and $dt = \hat{dt}$.

Before going further, note that localization results should always be checked by examining the data/replica match and the error surface. If the data/replica match is very poor or the least squares fit is not smooth, or if one of the estimated parameters is found at a boundary of the search grid, then the inversion probably performed poorly, and further investigation is required.

B. Combustive sound source

For our second example, we consider the sound from a Combustive Sound Source (CSS) (McNeese *et al.*, 2010) recorded during the Seabed Characterization Experiment

that took place on the New England mud patch in March/April 2017 (Knobles and Wilson, 2017). The specific signal presented here has previously been studied in Bonnel *et al.* (2018). Its range, known through GPS measurement, is 4.8 km.

The CSS source is known to be a strong impulse, preceded by a weak precursor and followed by several bubble pulses. It was monitored at close range during the experiment, and thus a source deconvolution can be performed before warping. The time series of the received signal is shown in Fig. 14(a), and the corresponding signal is shown in Fig. 14(b). The TF modal dispersion is highly contaminated by the source waveform. However, the spectrogram of the signal obtained after source deconvolution, shown in Fig. 14(c), is much cleaner. Warping can then be applied and modes can subsequently be filtered. The spectrogram of the warped signal is shown in Fig. 14(d), and the estimated dispersion curves are superimposed in Fig. 14(c). It is very interesting to see that many modes have been estimated, and that a gap exists between the low-order modes and the high-order modes. While the first nine modes are modes 1 to 9 with little doubt, it is impossible to identify the last five high-order modes without further effort. The curious reader may read Bonnel *et al.* (2018) to learn that the estimated high-order modes are modes 14 to 18. This example is a perfect illustration of warping's capability to filter modes, including highly dispersive high-order modes.

The estimated dispersion curves are then used as an input for localizing the CSS signal. However, inversion is restricted to the first four modes and to the frequency band 100–300 Hz. Because mode 1 is not estimated in this frequency band, it is effectively excluded from the localization.

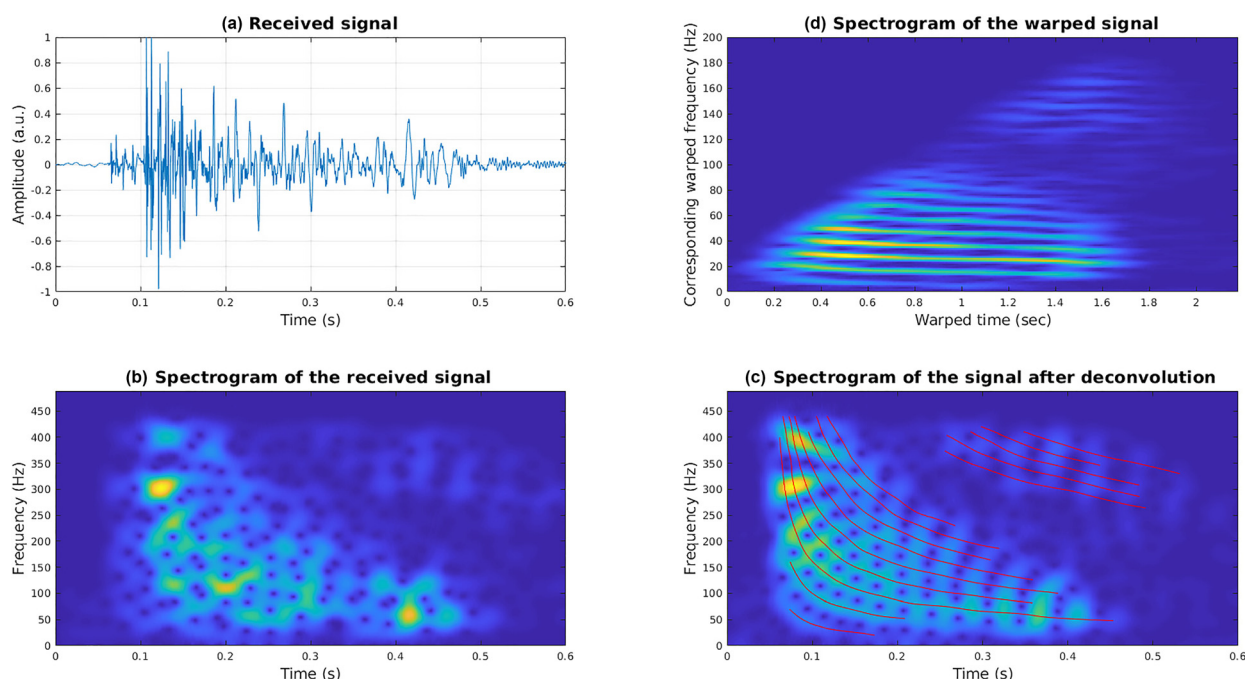


FIG. 14. (Color online) CSS signal analysis. (a) Received signal. (b) Spectrogram of the received signal. (c) Spectrogram of the signal after source deconvolution, and estimated dispersion curves. (d) Spectrogram of the signal after source deconvolution and warping. Figure from Bonnel *et al.* (2018).

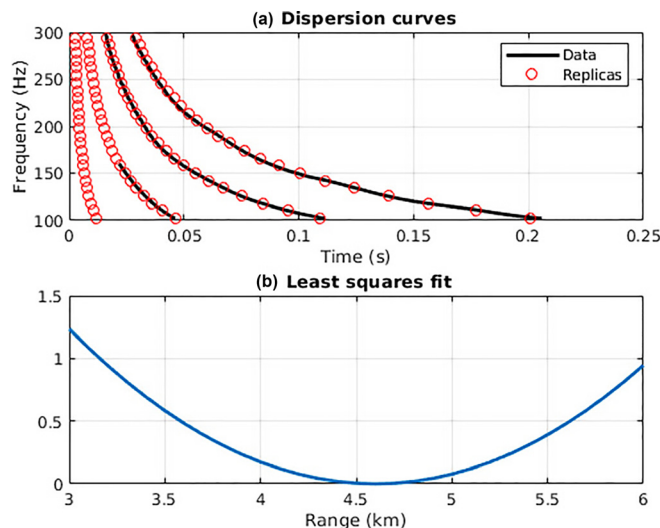


FIG. 15. (Color online) CSS localization. (a) Data and replicas. (b) One-dimensional slice of the least square fit, with $c_b = \hat{c}_b$ and $dt = \hat{dt}$. Note that the data presented on panel (a) is actually a subset of the estimated dispersion shown on Fig. 14(d); data is restricted to modes 1 to 4 and to the frequency band 100–300 Hz.

As source deconvolution has been performed, Eq. (19) is used for localization. The environmental model used to compute the replicas is a Pekeris waveguide with $D = 69.5$ m, $c_w = 1464.5$ m/s, $\rho_w = 1000$ kg/m³, and $\rho_b = 1600$ kg/m³. Other parameters are included in the inversion. The search space is as follows: $r \in [3, 6]$ km with 100 m steps, $c_b \in [1470, 2000]$ m/s with 10 m/s steps, and $dt \in [-4, -2]$ s with 0.01 s steps.

Localization results are illustrated in Fig. 15. Figure 15(a) shows the match between the data and predicted replicas, and Fig. 15(b) shows a slice of the least square fit. Both are good, suggesting that the localization scheme is successful. The estimated range, $\hat{r} = 4.6$ km is fully consistent with the GPS ground truth.

Interestingly, the data/replica match is nearly perfect [Fig. 15(a)]. This is because in the considered frequency band and for the considered modes, the experimental environment can be well-approximated with a Pekeris waveguide. Actually, the experimental environment is more complicated than that (Bonnell *et al.*, 2019), and such a match would have been impossible over a broader frequency band, and/or for a greater number of modes. Localizing the source using all the modes would have required the use of an environmental model much more complex than the Pekeris one.

C. Bowhead whale upsweep

The third and final example in this tutorial is a bowhead whale upsweep recorded off Deadhorse (Alaska) by Scripps Institution of Oceanography in August 2010 (Thode *et al.*, 2012). The specific signal presented here has previously been studied in Bonnell *et al.* (2014). Its estimated range, measured with a distributed array of vector sensors, is 14.1 ± 1.8 km (Bonnell *et al.*, 2014, call 7 in their Table I).

The signals emitted by bowhead whales are not stereotyped, and thus the exact TF structure of the original call is unknown. It is thus impossible to do source deconvolution. As an alternative, phase compensation (Sec. VD) must be used before warping.

The spectrogram of the received signal is presented in Fig. 16(a). The source TF law is estimated with three linear pieces. As explained in Sec. VD, an easy way to do so is to roughly follow the TF contour of mode 1. The chosen contour is plotted as a black curve in Fig. 16(a). Phase compensation is then performed, and the resulting spectrogram is plotted in Fig. 16(b). The result is definitely not perfect, but good enough to perform warping. The warped spectrogram is presented in Fig. 16(c). The TF masks that are used to filter the modes are also shown as red polygons. The corresponding estimated

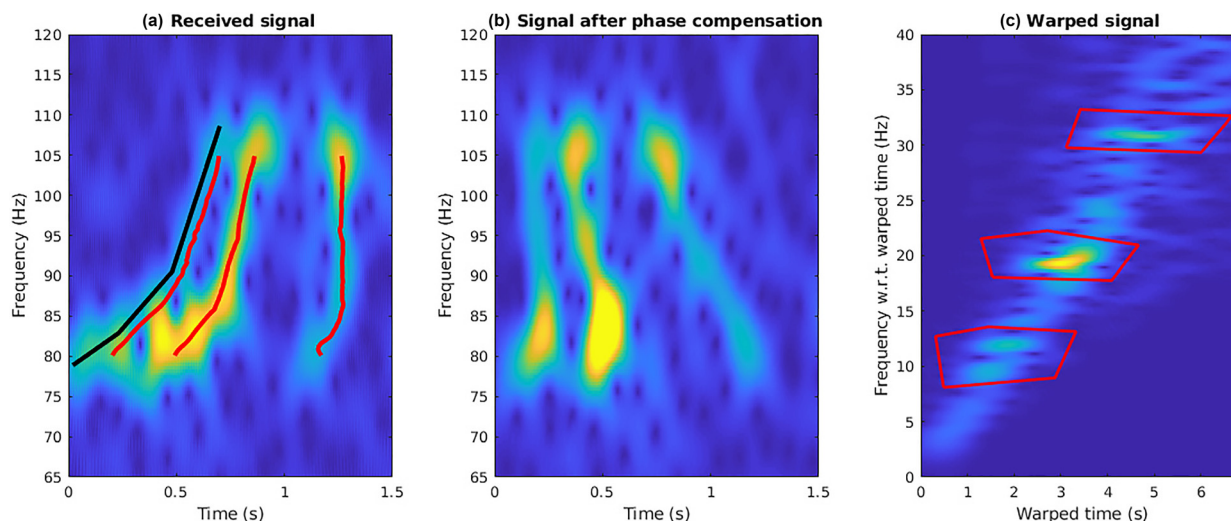


FIG. 16. (Color online) Bowhead whale upsweep analysis. (a) Received signal with estimated dispersion curves (red) and estimated source TF law (black). (b) Spectrogram of the signal after phase compensation. (c) Spectrogram of the signal after phase compensation and warping. Figure adapted from Bonnell *et al.* (2014).

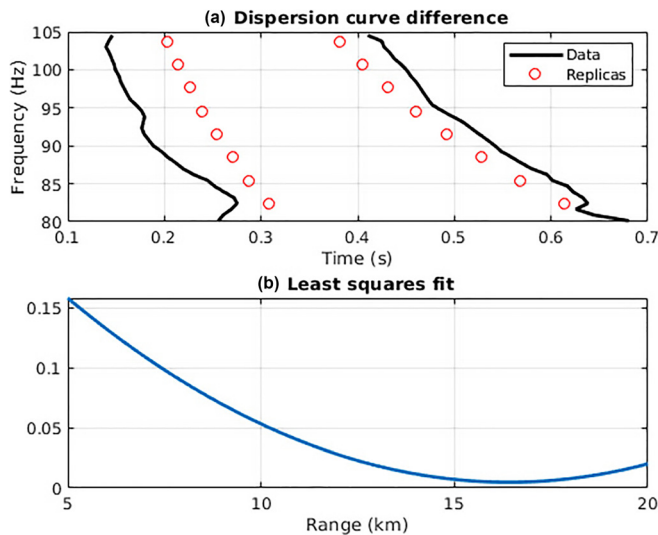


FIG. 17. (Color online) Bowhead whale upsweep localization. (a) Data and replicas. (b) One-dimensional slice of the least square fit, with $c_b = \hat{c}_b$.

dispersion curves are shown in Fig. 16(a). They perfectly match the underlying spectrogram, illustrating the success of the modal filtering operation. Note that such a result was obtained after several manual iterations to tune the estimated source TF law [Fig. 16(a)], the time origin of the signal after phase compensation [Fig. 16(b)] and the TF masks [Fig. 16(c)].

The next step is source localization. Because the exact TF law of the source is unknown, Eq. (20) is used. The raw dispersion curves of modes 1, 2, and 3 are combined to form dispersion curve differences between modes 2–1 and modes 3–2. The corresponding data are shown on Fig. 17(a). The environmental model used to compute the replicas is a Pekeris waveguide with $D = 55$ m, $c_w = 1442$ m/s, $\rho_w = 1000$ kg/m³, and $\rho_b = 1600$ kg/m³. The search space is as follows: $r \in [5, 20]$ km with 100 m steps, and $c_b \in [1550, 2000]$ m/s with 10 m/s steps.

Localization results are illustrated in Fig. 17. Figure 17(a) shows the match between the data and predicted replicas, and Fig. 17(b) shows a slice of the least square fit. The estimated range, $\hat{r} = 16.4$ km, is fully consistent with the independent estimates obtained with a distributed array of vector sensors.

Interestingly, the data/replica match is far from perfect [Fig. 17(a)]. One dispersion curve difference is underestimated, while the other is overestimated. However, the overall behavior of the least square fit is still smooth, which suggests that the two errors are compensating to provide a decent localization result. The misfit between data and predicted replica is due to the experimental environment, which has a range-dependent bathymetry. It is thus largely different from the Pekeris waveguide used for inversion. Nonetheless, dispersion curves are so sensitive to range that the localization result is correct.

VIII. CONCLUSION

This tutorial has reviewed both theoretical and practical uses of warping to study low-frequency transient sounds in

shallow water. Warping can be used to filter modes from a single receiver, an operation that had traditionally been performed with dense vertical line arrays. This reduction in hardware requirements has opened the door for applying sophisticated acoustical techniques to single-hydrophone data sets, including marine mammal localization or geoacoustic inversion. The tutorial covered in detail modal filtering using warping, and presented a simple source localization algorithm as a direct application of warping.

An important point of warping is that it requires a model of the sound propagation. In this tutorial, we have presented a warping operator based on an ideal isovelocity waveguide. Such a model has been widely used in the literature, and is known to enable modal filtering in various real-life scenarios. Its main advantages are that it is simple but robust, and allows warping all the modes with a single transform. While this simple transform works well in most shallow-water environments, we have noted that isovelocity warping may not work in environments dominated by strong reflection (e.g., a sharp thermocline), or situations where modal arrivals actually cross in the time-frequency domain.

However, warping operators can also be based on more complex models. This has been explored in the literature. Proposed models include the beam-displacement-ray-mode theory (Niu *et al.*, 2014a,b), an approximated Pekeris waveguide (Le Touzé *et al.*, 2009), or a waveguide invariant approximation (Bonnell *et al.*, 2013b; Qi *et al.*, 2015). Although these new warping operators are interesting in theory, they have not been used extensively for concrete applications. As far as we know, no study has compared their respective performance in terms of modal filtering. The potential gain brought by these more complex operators remains an open question. Nonetheless, the waveguide invariant approach is particularly promising for studying signals in strongly refracting environments, i.e., with a negative waveguide invariant. This is yet to be demonstrated on experimental data.

Interestingly, these new warping operators still warp all the modes at once. This is a feature which makes them promising candidates for advanced modal filtering. On the other hand, these operators are unable to handle specific dispersion features, such as modes crossing in the TF domain. Such a behavior may be due to a strong Airy phase, and/or to a given mode being ducted in either the water column or the seabed. A specific warping operator, based on a numerical approximation of the dispersion curves, has been proposed to handle these situations (Bonnell *et al.*, 2017). It improves filtering performance for mode 1, but not for the other modes. Filtering modes that cross in the TF domain remains an open question. This will need to be solved for shallow water environments when the sound speed profile has a gradient strong enough to fully trap low order modes. Such a situation is classical in the Arctic during summer but also happens at lower latitudes under specific oceanographic conditions. Spectrograms of experimental data with this specific features can be found in (Michalopoulou and Pole, 2016) or (Roth *et al.*, 2012).

Last but not least, this tutorial is based on the underlying assumption that the signal under study is a transient signal (such as an impulse or a frequency-modulated sweep) recorded on a single hydrophone. If the signal is not transient, such as broadband noise by a ship, then the method cannot be applied. In this case, a solution may be to work with the autocorrelation of the signal. Indeed, warping the autocorrelation of the signal has been proposed in Qi *et al.* (2015) and Zhou *et al.* (2014). The method seems very promising but, as far as we know, it has never been applied on broadband noise signals. Here again, there is the need for an experimental study. If the method is successful, it will open a new research avenue. Potential applications include ship localization, and/or tomography using sources of opportunity.

Overall, the single receiver capacity offered by warping opens the door to new experimental designs with single-hydrophone units that can be spread over an area of interest, forming a wide, non-synchronized, distributed array. In this configuration, warping may be used on each single-hydrophone unit independently, removing the need of coherent processing (and thus time synchronization) along the array. Interestingly, single hydrophone units are also easy to integrate into cheap robotic platforms. Because warping removes the need for time synchronization between sensors, robots may be used as a swarm to form a large array. However, no bulk automated algorithm yet exists for warping, and thus manual trial and error is a required procedure at present in order to obtain the best modal extractions. The automation issue will need to be solved before warping can be automatically applied on robotic platforms. Nonetheless, warping allows revisiting existing long-term datasets, and may open possibilities for localization/environmental information in situations that were previously not envisioned when the data were initially collected.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (Task Force Ocean, project N00014-19-1-2627) and by the North Pacific Research Board (project 1810). Original warping developments were supported by the French Delegation Generale de l'Armement.

APPENDIX A: TIME WARPING AND FREQUENCY WARPING

This appendix quickly reviews the difference between time and frequency warping. The starting point of dispersion-based warping is the need for a model of the modal phase. Because modal propagation is conveniently defined in the frequency domain through Eq. (2), we usually have a frequency domain model for the phase $\phi_m(f)$. As a result, it is natural to consider frequency warping, with warping function $\phi_m^{-1}(f)$. However, the definition of modal travel time through group speed is a stationary phase approximation. The same approximation can also be used to obtain a time domain expression for the modal phase, $\phi_m(t)$.

As a result, time warping also becomes a viable alternative, with warping function $\phi_m^{-1}(t)$.

Note that warping that operates in the time domain is usually dubbed “time warping,” although its results are conveniently displayed in the frequency domain. On the other hand, warping that operates in the frequency domain is usually dubbed “frequency warping,” although its results are conveniently displayed in the time domain. In this tutorial, we consider *time warping*. By warping the time axis, we modify the frequency content of modes, and thus obtain a result that is conveniently described in the frequency domain.

Nonetheless, the use of time warping versus frequency warping in a given context can still be questioned. As an example, Fig. 18 illustrates the time and frequency warping principles for three different signals. Time and/or frequency warping can be chosen, depending on the signal's TF distribution (i.e., dispersion curve). If, at any given instant, only one frequency exists, then the signal can be studied with time-warping (e.g., signal 1 in Fig. 18), and the TF shape of the warped signal becomes that of a continuous tone. On the other hand, if any given frequency corresponds with just a single time, then the signal can be studied with frequency-warping (e.g., signal 3 in Fig. 18), and the TF shape of the warped signal becomes an impulse. Last but not least, if the dispersion curve is bijective, then the signal can be studied with either time or frequency warping (e.g., signal 2 in Fig. 18).

In an underwater acoustics context, if one ignores 3D propagation effects, then the time of arrival of each mode at a given frequency is uniquely defined (as for signals 2 and 3 in Fig. 18). As a result, frequency warping can always be defined. If a closed-form expression for the modal phase $\phi_m(f)$ is not available, then one can resort to a numerical approximation, as obtained with a modal propagation code. This has been illustrated in Bonnel *et al.* (2017).

On the other hand, time warping can only be applied for simple propagation models where the frequency of a mode is uniquely defined for a given time (as for signals 1 and 2 in Fig. 18). As a counter example, this condition is violated in non-ideal waveguides because of the Airy phase behavior: at a given instant, a mode can be excited at two distinct frequencies (as signal 3 in Fig. 18).

Simple propagation models exist for which both time warping and frequency warping can be defined. The isoveLOCITY model considered in this tutorial is such a model. Both time warping and frequency warping are presented in Bonnel *et al.* (2010). Interestingly, time warping and frequency warping defined using this model have very different properties. As shown in this paper, time warping allows warping all the modes at once. On the other hand, frequency warping must be performed mode by mode (Bonnel *et al.*, 2010). This particular behavior does not seem particularly intuitive. It illustrates the importance of studying both time and frequency warping, because despite their analogous behavior, they may have fundamentally different properties.

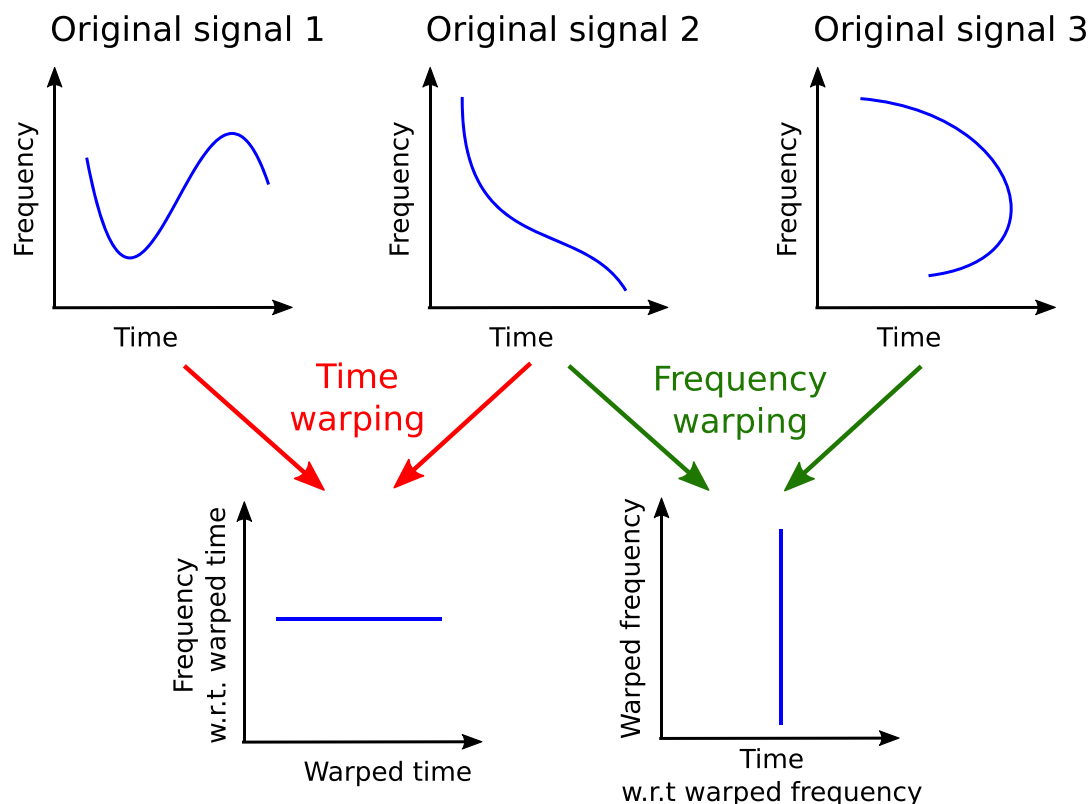


FIG. 18. (Color online) Comparison between time-warping and frequency-warping. Signal 1 must be warped using time-warping. Signal 3 must be warped using frequency-warping. Signal 2 can be warped using either time-warping or frequency-warping. The figure is from [Bonnel et al. \(2017\)](#).

APPENDIX B: BRIEF LITERATURE OVERVIEW

1. Motivations for warping

The central objective of separating modes with a single receiver with TF analysis is not new. A key notion here is the modal TF *dispersion*. This simply refers to the frequency-dependent time of arrival of the modes, which sometimes can be visualized on a simple spectrogram. Modal TF dispersion was first studied in the 1960s to understand the propagation of explosive sounds in shallow water ([Barakos, 1962](#); [Ewing et al., 1959](#)). An interesting phenomenon—which is explained in Sec. II B—is that the modal separation naturally increases with range. As a result, at large ranges, modal arrivals become naturally separated in time, and a simple spectrogram is enough to visualize their TF dispersion. This theory was used for pioneering studies in the 1960s ([Barakos, 1962](#); [Ewing et al., 1959](#)). In the 1980s, fundamental studies by [Zhou et al.](#) showed how the modal dispersion, as seen at long range in the time domain, could be used to infer seabed properties ([Zhou, 1985](#); [Zhou et al., 1987](#)). Since then, with the advance of numerical signal processing and of TF analysis, the techniques have advanced further. Modal TF dispersion has notably been used for geoaoustic inversion ([Potty et al., 2000](#); [Potty et al., 2004](#); [Rajan and Becker, 2010](#)) and for ranging marine mammals ([Abadi et al., 2014](#); [Munger et al., 2011](#); [Wiggins et al., 2004](#)). However, these studies have been performed on sources at relatively distant ranges, so that modes

were clearly time-separated on a conventional spectrogram. At shorter ranges, the modal arrivals blend together in time, requiring further signal processing development to extract modal information.

In a sense, the topic at hand is similar to the single channel source separation problem in signal processing. A single hydrophone (i.e., channel) is available, and the received modes can be seen as different sound sources that need to be separated. For speech processing, single channel source separation is usually solved by training statistical models on existing data [e.g., [Grais et al. \(2014\)](#), [Jang and Lee \(2003\)](#), and [Ozerov et al. \(2007\)](#)]. Unfortunately, existing data with reliable labels are very sparse in ocean acoustics, notably because environmental impact and noise degradation both impede reproducibility. As a result, alternative methods must be found.

Another option is to use more complex TF methods [e.g., [Taroudakis and Tzagkarakis \(2004\)](#)], and/or to post-processed traditional spectrograms with further signal processing methods [e.g., [Michalopoulou and Pole \(2016\)](#)]. Of particular interest here, theoretical advances in TF analysis for signals with dispersive group delays ([Papandreou-Suppappola et al., 2001](#)) have led to the development of TF methods specifically adapted for modal propagation ([Bonnel et al., 2013b](#); [Chen et al., 2003](#); [Hong et al., 2005](#); [Le Touzé et al., 2009](#)). Unfortunately, these methods have found very limited applications, one interesting exception being a geoaoustic inversion study by [Potty et al. \(2008\)](#). We believe

this is because the advanced TF literature is cryptic for non-specialists, and in turn enigmatic for the end-users who may need it.

Another option to analyze modal propagation at short range is to use classical TF analysis, but to transform the signal beforehand. This is particularly useful when using invertible transforms, so that one can filter the modes in the transformed space, and then go back to the original space. This idea is similar to classical bandpass filtering in the frequency domain, except that we are now considering transforms that are more complex than the Fourier transform.

Warping, the topic of this tutorial, is an example of advanced domain transform (Baraniuk and Jones, 1995). When combined with a physical model of the propagation, warping can be used to estimate modal dispersion (Bonnel et al., 2010) but also to filter modes, and thus to recover separated modal time series (Bonnel et al., 2017).³

2. Warping in ocean acoustics

The theoretical use of warping to filter modes (Le Touzé et al., 2009) and localize an impulsive sound source with a single receiver (Le Touzé et al., 2008) was first introduced by Le Touzé et al. The first warping applications for marine mammal localization (Bonnel et al., 2008) and geoacoustic inversion (Bonnel and Chapman, 2011) have been proposed by Bonnel et al. Since then, warping has been adopted in the underwater acoustic community, and has been used for various applications.

Warping has notably been used for environmental estimation studies, mostly seabed geoacoustic inversion (Bonnel et al., 2013a; Bonnel et al., 2019; Bonnel et al., 2012; Dong et al., 2017; Duan et al., 2016; Feng-Hua et al., 2014; Petrov, 2014; Zeng et al., 2013), but also water column tomography (Ballard et al., 2014), as well as joint estimation of water column and seabed properties (Warner et al., 2015). On a more basic research point of view, it is interesting to note that warping has also been used to estimate modal depth functions (Bonnel et al., 2011; Thode and Bonnel, 2015), as well as to filter modes from noise interferometry data (Sergeev et al., 2017; Tan et al., 2018).

Moreover, warping has been used for source localization in shallow water (Le Touzé et al., 2008; Qi et al., 2015; Zhou et al., 2014), and the performance of such source localization methods have been theoretically derived (Le Gall et al., 2017). The warping based localization methods have found application in bioacoustics. In particular, warping has been used to range baleen whales with a single sensor (Bonnel et al., 2014; Crance et al., 2015) and/or to estimate their calling depth (Thode et al., 2017). In the context where several non-synchronized hydrophones were available, warping has been used to estimate the latitude/longitude position of marine mammals (Bonnel et al., 2008; Warner et al., 2016).

A final application of warping is its utility as an alternative to source deconvolution to cancel the bubble pulse of a nearly impulsive source (Niu et al., 2013). It has also been

applied on vertical line array data to estimate the array tilt (Thode and Bonnel, 2015; Lu et al., 2017). Although this tutorial focuses on underwater acoustics applications, it is interesting to note that warping has been applied in many other fields. It has notably been used to study ultrasonic Lamb waves (De Marchi et al., 2009; De Marchi et al., 2010; Xu et al., 2012; Xu et al., 2014) with a procedure that is very similar to the one presented here: propagation models are used to predict dispersion curves, which are later used to define warping operators. Warping is also heavily used for speech processing (Aikawa, 1991; Benzeghiba et al., 2007; Zhan and Westphal, 1997).

APPENDIX C: NUMERICAL IMPLEMENTATION OF WARPING

1. Time and frequency domains

Before performing warping on sampled data using a computer, the first practical step is correctly identifying the time domain and the frequency domain of the original signal $y(t)$ and of the warped signal $y_w(t)$. This subsection largely reproduces a discussion that can be found in French in (Le Touzé, 2007).

Let us consider a signal $y(t)$. We further assume that the time domain of y is $\mathcal{D}_t = [t_{\min}, t_{\max}]$, and that its frequency domain is $\mathcal{D}_f = [f_{\min}, f_{\max}]$. Also, recall that warping is implemented as if the environment were an ideal waveguide, so that the signal is described by Eqs. (8) and (9). As a result, $t_{\min} > t_r$. For convenience, we now time-shift the time axis so that $t_{\min} = [t_r]^+$. In other words, we have $\mathcal{D}_t = [t_r, t_{\max}]$.

As a next step, we will define the domains of the warped signal, $y_w(t)$. First, let us denote \mathcal{T} the original time axis, and $\bar{\mathcal{T}}$ the warped time axis. Because of the warping definition, Eq. (7), any time \bar{t} in $\bar{\mathcal{T}}$ corresponds to $h(\bar{t})$ in \mathcal{T} . Conversely, any time t in \mathcal{T} corresponds to $h^{-1}(t)$ in $\bar{\mathcal{T}}$. Noting that $h^{-1}(t)$ is a monotonically increasing function, one finds that $s_w(t)$ has a finite time domain

$$\begin{aligned} \mathcal{D}_t^h &= [t_{\min}^h, t_{\max}^h] = [h^{-1}(t_{\min}); h^{-1}(t_{\max})] \\ &= [0; \sqrt{t_{\max}^2 - t_r^2}]. \end{aligned} \quad (C1)$$

The next step is to assess the frequency domain of the warped signal: $\mathcal{D}_f^h = [f_{\min}^h, f_{\max}^h]$.

To do so, one uses Eq. (12) which states that modes are warped onto their cutoff frequencies $f_{c,m}$. As a result, f_{\min}^h and f_{\max}^h respectively correspond to the cutoff frequencies of the minimal mode number (m_{\min}) and maximal mode number (m_{\max}) of the original signal: $f_{\min}^h = f_{c,m_{\min}}$ and $f_{\max}^h = f_{c,m_{\max}}$.

Although they fully define the frequency domain of the warped signal, these formulas are not convenient from a signal processing point of view. They notably contain environmental features (e.g., D , through $f_{c,m}$) which are irrelevant here, as the warping function does not use them. In the following, we seek formulas for f_{\min}^h and f_{\max}^h that are based on signal/warping features (e.g., t_{\max}, f_{\max}) rather than on environmental features ($f_{c,m}$).

To do so, the first step is to obtain the modal dispersion curve from Eq. (9) by computing the mode instantaneous frequency

$$\nu_{m_{iso}}(t) = \frac{1}{2\pi} \frac{\partial \phi_{m_{iso}}(t)}{\partial f} = \frac{(2m-1)c_w}{4D} \frac{t}{\sqrt{t^2 - t_r^2}}. \quad (C2)$$

Then, Eqs. (C2) and (11) are combined to define a time-frequency dependent (non-integer) mode number

$$m(t, f) = \frac{2D}{c_w} \frac{h^{-1}(t)}{t} f + \frac{1}{2}. \quad (C3)$$

Because $m(t, f)$ is increasing with t and f , one finds that

$$m_{min} = \frac{2D}{c_w} \frac{h^{-1}(t_{min})}{t_{min}} f_{min} + \frac{1}{2} = \frac{1}{2} \quad (C4)$$

and

$$m_{max} = \frac{2D}{c_w} \frac{h^{-1}(t_{max})}{t_{max}} f_{max} + \frac{1}{2}. \quad (C5)$$

Recalling that $f_{max}^h = f_{c, m_{max}}$, one can define f_{max}^h using only the maximal time and maximal frequency of the original signal

$$f_{max}^h = \frac{(2m_{max} - 1)c_w}{4D} = \frac{h^{-1}(t_{max})}{t_{max}} f_{max}. \quad (C6)$$

Similarly, one can show that $f_{min}^h = 0$. This leads to the final definition of the frequency domain of the warped signal

$$\mathcal{D}_f^h = \left[0, \frac{h^{-1}(t_{max})}{t_{max}} f_{max} \right]. \quad (C7)$$

Equations (C1) and (C7) fully define the time and frequency domain of the warped signal, using time and frequency features from the original signal. Because both time domain and frequency domain are finite, we can now proceed with questions related to sampling.

2. Discrete warping

Let us consider $y[n]$, the discrete version of the original signal $y(t)$. We assume it is sampled at frequency f_s and has N samples. We thus have $y[n] = y(t_n)$ with $t_n = t_{min} + n/f_s$ and $n \in [[0; N-1]]$.

As explained in Appendix C 1, warping moves the sample at location t_n in \mathcal{T} to a new location $\bar{t}_n = h^{-1}(t_n)$ in $\bar{\mathcal{T}}$. Because h^{-1} is non-linear, sampling in the warped domain is irregular. The warped signal must thus be interpolated on a regular time grid. This raises the question of the minimal sampling frequency for the warped signal.

a. Sampling frequency

If the original signal has been correctly sampled, its sampling frequency is $f_s > 2f_{max}$. As a result, Eq. (C6)

becomes $f_{max}^h < [h^{-1}(t_{max})/t_{max}](f_s/2)$, and thus the sampling frequency of the warped signal

$$f_s^h > \frac{h^{-1}(t_{max})}{t_{max}} f_s. \quad (C8)$$

Interestingly, one can obtain this result using some intuitive thought. The time derivative $[h^{-1}]'(t) = t/\sqrt{t^2 - t_r^2}$ is decreasing and positive. As a result, the spacing of the irregular time grid in $\bar{\mathcal{T}}$ decreases as \bar{t} increases. In other words, $\Delta \bar{t}_n = \bar{t}_n - \bar{t}_{n-1}$ decreases as n increases. The smallest spacing is thus $\Delta \bar{t}_N = h^{-1}(t_N) - h^{-1}(t_{N-1})$, and it is tempting to naively use it as a regular sampling step for the warped signal.

In fact,

$$\Delta \bar{t}_N = \Delta t \frac{h^{-1}(t_N) - h^{-1}(t_{N-1})}{\Delta t} \simeq \Delta t [h^{-1}]'(t_{max}), \quad (C9)$$

with $\Delta t = 1/f_s$ the regular sampling period in \mathcal{T} . Because $[h^{-1}]'(t_{max}) = t_{max}/h^{-1}(t_{max})$, then

$$\Delta \bar{t}_N = \frac{1}{f_s} \frac{t_{max}}{h^{-1}(t_{max})}. \quad (C10)$$

This happens to be the minimal bound of Eq. (C8), and thus $\Delta \bar{t}_N$ can be used as a regular sampling step in the warped domain. In practice, if computation time is not an issue, one may use $f_s^h = 2/\Delta \bar{t}_N$ for a better visual representation of the warped signal.

b. Number of samples

The number of samples K of the warped signal is obtained from its time domain \mathcal{D}_t^h and sampling frequency f_s^h :

$$K = \text{ceil} \left([h^{-1}(t_{max}) - h^{-1}(t_{min})] f_s^h \right), \quad (C11)$$

with $\text{ceil}(x)$ the nearest integer greater than or equal to x .

As stated in Appendix C 1, one usually chooses $t_{min} = [t_r]^+$. As a result $h^{-1}(t_{min}) = 0$, and thus $K = \text{ceil}(h^{-1}(t_{max}) f_s^h)$.

c. Interpolation

The discrete version of the warped signal $y_w[k]$ is thus sampled at frequency f_s^h with K samples, with $\bar{t}_k = k/f_s^h$, $k \in [[0; K-1]]$.

Remembering the discussion in Appendix C 1, any time \bar{t} in $\bar{\mathcal{T}}$ corresponds to $h(\bar{t})$ in \mathcal{T} . The value of the warped signal $y_w[m]$ is thus given by the value of the original signal at time $h(k/f_s^h)$. Because it is unlikely that a sample exists for the original signal at $t = h(k/f_s^h)$, the value at this specific time is obtained by interpolation. Although linear interpolation has been used in the past and gives satisfactory results, it sometimes creates high frequency artifacts in the warped signal (Arisdakessian, 2014). Therefore, we advise using the Whittaker-Shannon interpolation (Abdul, 1977) to

obtain the value of $y(t)$ at any time t from its sampled version,

$$y(t) = \sum_{n=0}^{N-1} y[n] \text{sinc}(tf_s - n). \quad (\text{C12})$$

d. Energy conservation

As a last step, one must compute the multiplicative factor $|h'(t)|^{1/2}$ to conserve energy. It comes directly from Eq. (10) and compensates for the time axis stretching/compression. Because $h'(t) = t/h(t)$, this factor is given as $\sqrt{|h'[k/f_s^h]|} = \sqrt{\bar{t}_k/h(\bar{t}_k)}$ for the k th sample of the warped signal.

As a summary, the k th sample of the warped signal is

$$y_w[m] = \sqrt{\frac{\bar{t}_k}{h(\bar{t}_k)}} s[h(\bar{t}_k)], \quad (\text{C13})$$

with $\bar{t}_k = k/f_s^h$, $h(\bar{t}_k) = (\bar{t}_k^2 + t_r^2)^{1/2}$, and the quantity $y[h(\bar{t}_k)]$ is obtained from the original discrete signal y through interpolation.

3. Inverse warping

Inverse warping can be seen as forward warping using $h^{-1}(t)$ as the warping function. However, for practical applications, one deals with an original sampled signal (N samples, sampling frequency f_s), warps it (K samples, sampling frequency f_s^h), and then unwarps it. As a result, inverse warping does not require estimation of numbers of samples and sampling frequency. These are known *a priori*: they are identical to the properties of the original signal (N samples, sampling frequency f_s).

The only steps that need consideration when unwarping a signal are proper interpolation and energy conservation. Let y_w denote the warped signal and y_u the unwarped signal, recovered from y_w using inverse warping. The n th sample of the unwarped signal simply becomes [via Eq. (C13)]

$$y_u[n] = \sqrt{\frac{t_n}{h^{-1}(t_n)}} y_w[h^{-1}(t_n)], \quad (\text{C14})$$

with $t_n = t_{\min} + n/f_s$, $h^{-1}(t_n) = (t_n^2 - t_r^2)^{1/2}$, and the quantity $y_w[h^{-1}(t_n)]$ obtained via interpolation of the warped discrete signal y_w .

¹See supplementary material at <https://doi.org/10.1121/10.0000937> for the MATLAB toolbox associated with this tutorial.

²The code provided here has actually been extracted from a free online toolbox (Auger *et al.*, 1996), and slightly modified so that it can run as a standalone MATLAB routine

³Note that warping is actually a preliminary step to compute some of the adapted TF representations that have been cited earlier (Bonnell *et al.*, 2013b; Le Touzé *et al.*, 2009; Papandreou-Suppappola *et al.*, 2001).

Abadi, S., Thode, A., Blackwell, S., and Dowling, D. (2014). "Ranging bowhead whale calls in a shallow-water dispersive waveguide," *J. Acoust. Soc. Am.* **136**(1), 130–144.

- Abdul, J. (1977). "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE* **65**(11), 1565–1596.
- Aikawa, K. (1991). "Speech recognition using time-warping neural networks," in *Neural Networks for Signal Processing Proceedings of the 1991 IEEE Workshop*, IEEE, pp. 337–346.
- Ainslie, M. A. (2010). *Principles of Sonar Performance Modelling* (Springer, Berlin).
- Arisdakessian, C. (2014). "Bowhead whale localization in the Beaufort Sea," Technical Report, ENSTA Bretagne & Scripps Institution of Oceanography.
- Auger, F., Lemoine, O., Goncalves, P., and Flandrin, P. (1996). "The time-frequency toolbox" tfb.nongnu.org/index.html (Last viewed March 18, 2020).
- Baggeroer, A., Kuperman, W., and Mikhalevsky, P. (1993). "An overview of matched field methods in ocean acoustics," *IEEE J. Ocean. Eng.* **18**(4), 401–424.
- Ballard, M. S., Frisk, G. V., and Becker, K. M. (2014). "Estimates of the temporal and spatial variability of ocean sound speed on the new jersey shelf," *J. Acoust. Soc. Am.* **135**(6), 3316–3326.
- Barakos, P. (1962). "Experimental determination of compressional velocity for the bottom layer by the dispersion method," *J. Acoust. Soc. Am.* **34**(12), 1919–1926.
- Baraniuk, R., and Jones, D. (1995). "Unitary equivalence: A new twist on signal processing," *IEEE Trans. Sign. Process.* **43**(10), 2269–2282.
- Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvét, D., Fissore, L., Laface, P., Mertins, A., Ris, C., Rose, R., Tyagi, V., and Welkensk, C. (2007). "Automatic speech recognition and speech variability: A review," *Speech Commun.* **49**(10–11), 763–786.
- Boashash, B. (2015). *Time-frequency Signal Analysis and Processing: A Comprehensive Reference* (Academic, New York).
- Bonnans, J.-F., Gilbert, J. C., Lemaréchal, C., and Sagastizábal, C. A. (2006). *Numerical Optimization: Theoretical and Practical Aspects* (Springer Science & Business Media, New York).
- Bonnell, J. (2018). "Warping sound in the ocean: A far-out theoretical concept has practical applications," *Oceanus Magazine*, www.whoi.edu/oceanus/feature/warping-sound-in-the-ocean (Last viewed March 18, 2020).
- Bonnell, J., Caporale, S., and Thode, A. (2017). "Waveguide mode amplitude estimation using warping and phase compensation," *J. Acoust. Soc. Am.* **141**(3), 2243–2255.
- Bonnell, J., and Chapman, N. (2011). "Geoacoustic inversion in a dispersive waveguide using warping operators," *J. Acoust. Soc. Am.* **130**, EL101–EL107.
- Bonnell, J., Dosso, S., and Chapman, R. (2013a). "Bayesian geoacoustic inversion of single hydrophone light bulb data using warping dispersion analysis," *J. Acoust. Soc. Am.* **134**, 120–130.
- Bonnell, J., Dosso, S., Eleftherakis, D., and Chapman, R. (2019). "Trans-dimensional inversion of modal dispersion data on the New England mud patch," *IEEE J. Ocean. Eng.* **45**, 116–130.
- Bonnell, J., Gervaise, C., Nicolas, B., and Mars, J. (2012). "Single-receiver geoacoustic inversion using modal reversal," *J. Acoust. Soc. Am.* **131**(1), 119–128.
- Bonnell, J., Gervaise, C., Roux, P., Nicolas, B., and Mars, J. (2011). "Modal depth function estimation using time-frequency analysis," *J. Acoust. Soc. Am.* **130**, 61–71.
- Bonnell, J., Le Touzé, G., Nicolas, B., and Mars, J. (2013b). "Physics-based time-frequency representations for underwater acoustics: Power class utilization with waveguide-invariant approximation," *IEEE Sign. Process. Mag.* **30**(6), 120–129.
- Bonnell, J., Le Touzé, G., Nicolas, B., Mars, J., and Gervaise, C. (2008). "Automatic and passive whale localization in shallow water using gunshots," in *IEEE/MTS, Oceans'08, Quebec, Canada*, IEEE, pp. 1–6.
- Bonnell, J., Lin, Y.-T., Eleftherakis, D., Goff, J., Dosso, S., Chapman, R., Miller, J., and Potty, G. (2018). "Geoacoustic inversion on the New England mud patch using warping and dispersion curves of high-order modes," *J. Acoust. Soc. Am.* **143**(5), EL405–EL411.
- Bonnell, J., Nicolas, B., Mars, J., and Walker, S. (2010). "Estimation of modal group velocities with a single receiver for geoacoustic inversion in shallow water," *J. Acoust. Soc. Am.* **128**, 719–727.
- Bonnell, J., Thode, A., Blackwell, S., Kim, K., and Macrander, A. (2014). "Range estimation of bowhead whale (*Balaena mysticetus*) calls in the arctic using a single hydrophone," *J. Acoust. Soc. Am.* **136**(1), 145–155.

- Buck, J. R., Preisig, J. C., and Wage, K. E. (1998). "A unified framework for mode filtering and the maximum a posteriori mode filter," *J. Acoust. Soc. Am.* **103**(4), 1813–1824.
- Caiti, A., Chapman, N. R., Hermand, J.-P., and Jesus, S. M. (2006). *Acoustic Sensing Techniques for the Shallow Water Environment: Inversion Methods and Experiments* (Springer Science & Business Media, New York).
- Chapman, N. R. (1985). "Measurement of the waveform parameters of shallow explosive charges," *J. Acoust. Soc. Am.* **78**(2), 672–681.
- Chapman, N. R. (2012). "Geoacoustic perspectives: What have we learned about modelling sound propagation in the ocean bottom," in *Proceedings of Acoustics 2012*, Fremantle, Australia.
- Chen, C., Miller, J., Boudreaux-Bartels, G., Potty, G., and Lazauski, C. (2003). "Time-frequency representations for wideband acoustic signals in shallow water," in *Proceedings of OCEANS 2003*, Vol. 5, pp. SP2903–SP2907.
- Clayton, R. W., and Wiggins, R. A. (1976). "Source shape estimation and deconvolution of teleseismic bodywaves," *Geophys. J. Int.* **47**(1), 151–177.
- Collins, M. D., and Kuperman, W. (1991). "Focalization: Environmental focusing and source localization," *J. Acoust. Soc. Am.* **90**(3), 1410–1422.
- Crance, J., and Berchok, C. (2016). "Stereotyped, repetitive gunshot call patterns produced by the north Pacific right whale, *Eubalaena japonica*," *J. Acoust. Soc. Am.* **140**(4), 3295–3296.
- Crance, J. L., Berchok, C. L., Bonnel, J., and Thode, A. M. (2015). "Northeasternmost record of a north Pacific fin whale (*Balaenoptera physalus*) in the Alaskan Chukchi Sea," *Polar Biol.* **38**, 1767–1773.
- De Marchi, L., Marzani, A., Caporale, S., and Speciale, N. (2009). "Ultrasonic guided-waves characterization with warped frequency transforms," *IEEE Trans. Ultrasonics Ferroelectr. Freq. Control* **56**(10), 2232–2240.
- De Marchi, L., Ruzzene, M., Xu, B., Baravelli, E., and Speciale, N. (2010). "Warped basis pursuit for damage detection using lamb waves," *IEEE Trans. Ultrasonics Ferroelectr. Freq. Control* **57**(12), 2734–2741.
- Dong, H., Badiey, M., and Chapman, N. R. (2017). "Geoacoustic inversion of argon data under influence of internal waves," *IEEE J. Oceanic Eng.* **42**(3), 632–638.
- Dosso, S. E., and Wilmot, M. J. (2011). "Bayesian multiple-source localization in an uncertain ocean environment," *J. Acoust. Soc. Am.* **129**(6), 3577–3589.
- Duan, R., Chapman, N. R., Yang, K., and Ma, Y. (2016). "Sequential inversion of modal data for sound attenuation in sediment at the new jersey shelf," *J. Acoust. Soc. Am.* **139**(1), 70–84.
- Emmetière, R., Bonnel, J., Géhant, M., Cristol, X., and Chonavel, T. (2018). "Understanding deep-water striation patterns and predicting the waveguide invariant as a distribution depending on range and depth," *J. Acoust. Soc. Am.* **143**(6), 3444–3454.
- Ewing, M., Mueller, S., Landisman, M., and Sato, Y. (1959). "Transient analysis of earthquake and explosion arrivals," *Pure Appl. Geophys.* **44**(1), 83–118.
- Feng-Hua, L., Bo, Z., and Yong-Gang, G. (2014). "A method of measuring the *in situ* seafloor sound speed using two receivers with warping transformation," *Chin. Phys. Lett.* **31**(2), 024301:1–4.
- Frisk, G. V. (1994). *Ocean and Seabed Acoustics: A Theory of Wave Propagation* (Pearson Education, Upper Saddle River, NJ).
- Grais, E. M., Sen, M. U., and Erdogan, H. (2014). "Deep neural networks for single channel source separation," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 3734–3738.
- Heard, G. J., McDonald, M., Chapman, N., and Jaschke, L. (1997). "Underwater light bulb implosions: A useful acoustic source," in *Oceans '97. MTS/IEEE Conference Proceedings*, IEEE, Vol. 2, pp. 755–762.
- Hermand, J. (1999). "Broad-band geoacoustic inversion in shallow water from waveguide impulse response measurements on a single hydrophone: Theory and experimental results," *IEEE J. Oceanic Eng.* **24**(1), 41–66.
- Hong, J., Sun, K., and Kim, Y. (2005). "Dispersion-based short-time Fourier transform applied to dispersive wave analysis," *J. Acoust. Soc. Am.* **117**, 2949–2960.
- Jang, G.-J., and Lee, T.-W. (2003). "A maximum likelihood approach to single-channel source separation," *J. Mach. Learn. Res.* **4**(Dec), 1365–1392.
- Jensen, F., Kuperman, W., Porter, M., and Schmidt, H. (2011). *Computational Ocean Acoustics*, 2nd ed. (American Institute of Physics, New York).
- Jesus, S., Porter, M., Stéphan, Y., Démoulin, X., Rodríguez, O., and Coelho, E. (2000). "Single hydrophone source localization," *IEEE J. Oceanic Eng.* **25**(3), 337–346.
- Knobles, D., and Wilson, P. (2017). "Sbcex: Seabed characterization experiment" www.arlut.utexas.edu/sbcex/index.html (Last viewed March 18, 2020).
- Kozek, W., and Hlawatsch, F. (1992). "A comparative study of linear and nonlinear time-frequency filters," in *Proceedings of the IEEE-SP International Symposium*, pp. 163–166.
- Le Gall, Y., Socheleau, F.-X., and Bonnel, J. (2017). "Performance analysis of single-receiver matched-mode localization," *IEEE J. Oceanic Eng.* **44**(1), 193–206.
- Le Touzé, G. (2007). "Localisation de source par petits fonds en UBF (1–100 Hz) à l'aide d'outils temps-fréquence," Ph.D. thesis, Institut National Polytechnique de Grenoble, Grenoble.
- Le Touzé, G., Nicolas, B., Mars, J., and Lacoume, J. (2009). "Matched representations and filters for guided waves," *IEEE Trans. Sign. Process.* **57**(5), 1783–1795.
- Le Touzé, G., Torras, J., Nicolas, B., and Mars, J. (2008). "Source localization on a single hydrophone," in *OCEANS 2008*, IEEE, pp. 1–6.
- Leroy, E. C., Samaran, F., Stafford, K. M., Bonnel, J., and Royer, J.-Y. (2018). "Broad-scale study of the seasonal and geographic occurrence of blue and fin whales in the southern Indian Ocean," *Endang. Spec. Res.* **37**, 289–300.
- Lu, L., Ren, Q., Hu, T., and Ma, L. (2017). "Array shape estimation method based on extraction of single normal mode," in *Acoustics 2017*, Perth, Australia.
- Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S., and Nero, R. W. (2006). "Fish population and behavior revealed by instantaneous continental shelf-scale imaging," *Science* **311**(5761), 660–663.
- McNeese, A. R., Sagers, J. D., Wilson, P. S., and Knobles, D. P. (2010). "An investigation of the combustive sound source," in *Proceedings of Meetings on Acoustics*, ASA, Vol. 9, p. 005002.
- Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (2007). "An overview of fixed passive acoustic observation methods for cetaceans," *Oceanography* **20**(4), 36–45.
- Michalopoulou, Z.-H., and Pole, A. (2016). "Sediment sound speed inversion with time-frequency analysis and modal arrival time probability density functions," *J. Acoust. Soc. Am.* **140**(1), EL131–EL136.
- Moore, S. E., Stafford, K. M., Mellinger, D. K., and Hildebrand, J. A. (2006). "Listening for large whales in the offshore waters of Alaska," *BioScience* **56**(1), 49–55.
- Munger, L., Wiggins, S., and Hildebrand, J. (2011). "North Pacific right whale up-call source levels and propagation distance on the southeastern Bering Sea shelf," *J. Acoust. Soc. Am.* **129**, 4047–4054.
- NASA (2015). "Is warp drive real?," www.nasa.gov/centers/glenn/technology/warp/warp.html (Last viewed March 18, 2020).
- Neilsen, T., and Westwood, E. (2002). "Extraction of acoustic normal mode depth functions using vertical line array data," *J. Acoust. Soc. Am.* **111**(2), 748–756.
- Niu, H., Zhang, R., and Li, Z. (2014a). "A modified warping operator based on BDRM theory in homogeneous shallow water," *Sci. China Phys. Mech. Astron.* **57**(3), 424–432.
- Niu, H., Zhang, R., and Li, Z. (2014b). "Theoretical analysis of warping operators for non-ideal shallow water waveguides," *J. Acoust. Soc. Am.* **136**(1), 53–65.
- Niu, H.-Q., Zhang, R.-H., Li, Z.-L., Guo, Y.-G., and He, L. (2013). "Bubble pulse cancelation in the time-frequency domain using warping operators," *Chin. Phys. Lett.* **30**(8), 084301.
- Ozerov, A., Philippe, P., Bimbot, F., and Gribonval, R. (2007). "Adaptation of Bayesian models for single-channel source separation and its application to voice/music separation in popular songs," *IEEE Trans. Audio Speech Lang. Process.* **15**(5), 1564–1578.
- Papandreou-Suppappola, A., Murray, R., Iem, B., and Boudreaux-Bartels, G. (2001). "Group delay shift covariant quadratic time-frequency representations," *IEEE Trans. Sign. Process.* **49**(11), 2549–2564.
- Parks, S. E., and Tyack, P. L. (2005). "Sound production by North Atlantic right whales (*Eubalaena glacialis*) in surface active groups," *J. Acoust. Soc. Am.* **117**(5), 3297–3306.

- Pekeris, C. (1948). "Theory of propagation of explosive sound in shallow water," *Mem. Geo. Soc. Am.* **27**, 1–116.
- Petrov, P. (2014). "A method for single-hydrophone geoacoustic inversion based on the modal group velocities estimation: Application to a waveguide with inhomogeneous bottom relief," in *Days on Diffraction (DD)*, 2014, IEEE, pp. 186–191.
- Potty, G., Miller, J., Dahl, P., and Lazauski, C. (2004). "Geoacoustic inversion results from the ASIAEX East China Sea experiment," *Oceanic Engineering, IEEE Journal of* **29**(4), 1000–1010.
- Potty, G., Miller, J., Lynch, J., and Smith, K. (2000). "Tomographic inversion for sediment parameters in shallow water," *J. Acoust. Soc. Am.* **108**, 973–986.
- Potty, G., Miller, J., Wilson, P., Lynch, J., and Newhall, A. (2008). "Geoacoustic inversion using combusive sound source signals," *J. Acoust. Soc. Am.* **124**, EL146–EL150.
- Qi, Y. B., Zhou, S. H., Zhang, R. H., and Ren, Y. (2015). "A waveguide-invariant-based warping operator and its application to passive source range estimation," *J. Comput. Acoust.* **23**(01), 1550003.
- Rajan, S., and Becker, K. (2010). "Inversion for range-dependent sediment compressional-wave-speed profiles from modal dispersion data," *IEEE J. Oceanic Eng.* **35**(1), 43–58.
- Roth, E. H., Hildebrand, J. A., Wiggins, S. M., and Ross, D. (2012). "Underwater ambient noise on the Chukchi Sea continental slope from 2006–2009," *J. Acoust. Soc. Am.* **131**(1), 104–110.
- Sergeev, S., Shurup, A., Godin, O., Vedenev, A., Goncharov, V., Mukhanov, P. Y., Zabolot, N., and Brown, M. (2017). "Separation of acoustic modes in the Florida straits using noise interferometry," *Acoust. Phys.* **63**(1), 76–85.
- Stanton, T. K., Lee, W.-J., and Baik, K. (2018). "Echo statistics associated with discrete scatterers: A tutorial on physics-based methods," *J. Acoust. Soc. Am.* **144**(6), 3124–3171.
- Tan, T., Godin, O. A., Lefebvre, A., Beaute, W., Katsnelson, B. G., and Yarina, M. (2018). "Characterizing the seabed by using noise interferometry and time warping," in *Proceedings of Meetings on Acoustics 176ASA, ASA*, Vol. 35, p. 070001.
- Taroudakis, M., and Tzagkarakis, G. (2004). "On the use of the reassigned wavelet transform for mode identification," *J. Comput. Acoust.* **12**(02), 175–196.
- Thode, A., Bonnel, J., Thieury, M., Fagan, A., Verlinden, C., Wright, D., Berchok, C., and Crance, J. (2017). "Using nonlinear time warping to estimate North Pacific right whale calling depths in the Bering Sea," *J. Acoust. Soc. Am.* **141**(5), 3059–3069.
- Thode, A., Kim, K., Blackwell, S., Greene, C., Jr., Nations, C., McDonald, T., and Macrander, A. (2012). "Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys," *J. Acoust. Soc. Am.* **131**, 3726–3747.
- Thode, A. M., and Bonnel, J. (2015). "Experimental demonstration of time warping to invert for array tilt and mode shape on a vertical array in a shallow arctic environment," in *OCEANS 2015, Genova (Italy)*, IEEE, pp. 1–7.
- Tindle, C., Guthrie, K., Bold, G., Johns, M., Jones, D., Dixon, K., and Birdsall, T. (1978). "Measurements of the frequency dependence of normal modes," *J. Acoust. Soc. Am.* **64**(4), 1178–1185.
- Warner, G. A., Dosso, S. E., Dettmer, J., and Hannay, D. E. (2015). "Bayesian environmental inversion of airgun modal dispersion using a single hydrophone in the Chukchi Sea," *J. Acoust. Soc. Am.* **137**(4), 2361–2362.
- Warner, G. A., Dosso, S. E., Hannay, D. E., and Dettmer, J. (2016). "Bowhead whale localization using asynchronous hydrophones in the Chukchi Sea," *J. Acoust. Soc. Am.* **140**(1), 20–34.
- Wiggins, S., McDonald, M., Munger, L., Moore, S., and Hildebrand, J. (2004). "Waveguide propagation allows range estimates for North Pacific right whales in the Bering Sea," *Can. Acoust.* **32**, 146–154.
- Wilson, G., Koch, R., and Vidmar, P. (1988). "Matched mode localization," *J. Acoust. Soc. Am.* **84**, 310–320.
- Wright, D. (2017). "Passive acoustic monitoring of the critically endangered eastern north Pacific right whale (*Eubalaena japonica*). Final Report for the Marine Mammal Commission," technical report.
- Xu, K., Ta, D., Hu, B., Laugier, P., and Wang, W. (2014). "Wideband dispersion reversal of Lamb waves," *IEEE Trans. Ultrasonics Ferroelectr. Freq. Control* **61**(6), 997–1005.
- Xu, K., Ta, D., Moilanen, P., and Wang, W. (2012). "Mode separation of Lamb waves based on dispersion compensation method," *J. Acoust. Soc. Am.* **131**(4), 2714–2722.
- Yang, T. (1987). "A method of range and depth estimation by modal decomposition," *J. Acoust. Soc. Am.* **82**(5), 1736–1745.
- Zeng, J., Chapman, N. R., and Bonnel, J. (2013). "Inversion of seabed attenuation using time-warping of close range data," *J. Acoust. Soc. Am.* **134**(5), EL394–EL399.
- Zhan, P., and Westphal, M. (1997). "Speaker normalization based on frequency warping," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, Vol. 2, pp. 1039–1042.
- Zhou, J.-x. (1985). "Normal mode measurements and remote sensing of sea-bottom sound velocity and attenuation in shallow water," *J. Acoust. Soc. Am.* **78**(3), 1003–1009.
- Zhou, J.-x., Zhang, X.-z., Rogers, P. H., and Jarzynski, J. (1987). "Geoacoustic parameters in a stratified sea bottom from shallow-water acoustic propagation," *J. Acoust. Soc. Am.* **82**(6), 2068–2074.
- Zhou, S., Qi, Y., and Ren, Y. (2014). "Frequency invariability of acoustic field and passive source range estimation in shallow water," *Sci. China Phys. Mech. Astron.* **57**(2), 225–232.