3-2006

Doppler-only Multistatic Radar

Dustin G. Mixon

DOPPLER-ONLY MULTISTATIC RADAR

THESIS

Dustin G. Mixon

Second Lieutenant, USAF

AFIT/GAM/ENC/06-01

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

WRIGHT-PATTERSON AIR FORCE BASE, OHIO

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the United States Government.

AFIT/GAM/ENC/06-01

DOPPLER-ONLY MULTISTATIC RADAR

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science

Dustin G. Mixon, B.S.
Second Lieutenant, USAF

March 2006

AFIT/GAM/ENC/06-01

DOPPLER-ONLY MULTISTATIC RADAR

Dustin G. Mixon, B.S.

Second Lieutenant, USAF

Approved:

Matthew C. Fickus Thesis Advisor	Date
Mark A. Abramson Committee Member	Date
Laura R. C. Suzuki Committee Member	Date

AFIT/GAM/ENC/06-01

Abstract

In order to estimate the position and velocity of a target, most multistatic radar systems require multiple independent target measurements, such as angle-of-arrival, time-of-arrival, and Doppler information. Though inexpensive and reliable, Doppler-only systems have not been widely implemented due to the inherent nonlinear problem of determining a target's position and velocity from their measurements. We solve this problem. In particular, we first establish the lack of observability in the Doppler-only bistatic system, thereby demonstrating the need for multiple transmitters and/or receivers. Next, for a multistatic system with a sufficient number of transmitter-receiver pairs, we invoke classical optimization techniques, such as gradient-descent and Newton's method, to quickly and reliably find a numerical solution to the system of nonlinear Doppler equations. Finally, we indicate a best design for the transmitter-receiver constellation to be employed in the aforementioned optimization.

Acknowledgements

The student who claims independence is either a liar or a fool. I am truly thankful to all who have supported me throughout my thesis research endeavors. In particular, I would like to thank my extremely helpful advisor, Dr Matthew Fickus, who was always happy to provide clear direction. Further gratitude goes to the rest of my thesis committee, Lt Col Mark Abramson and Maj Laura Suzuki, whose multiple dimensions of support permitted sound research on my topic. I would also like to thank Will and Paul, who not only introduced me to this problem, but also aided my appreciation for its application.

It can be easy to take for granted those who are most significant. I am thankful for my deep and abiding friendships, which have tremendously animated my time here at AFIT. My parents have also given me much to be thankful for, including a solid, healthy upbringing. They have never failed to support me in any of my pursuits; I am extremely grateful for this and their ever-abundant love. Finally, I would like to thank He who matters most, as it is through His mercy that I have even the privilege to breathe.

Dustin G. Mixon

Table of Contents

Page

List of Figures

DOPPLER-ONLY MULTISTATIC RADAR

I. Introduction

Radar is the process of estimating the attributes of an object by measuring its effect on the passage of electromagnetic signals. *Multistatic* radar employs multiple transmitters and/or receivers, while combining data at a central location [39, p. 4]. Most modern multistatic radar systems use multiple receiver sites, each collecting several independent target measurements, such as the time-of-arrival, angle-of-arrival and Doppler shift of the reflected signals. These measurements are then combined to estimate the *target state*; that is, the target's parameters of interest, such as position and velocity.

Of the many multistatic systems operating today, few, if any, rely solely on Doppler shift information, depending, at least in part, upon time and/or angle measurements. The purpose of this thesis is to determine the extent to which a target state can be estimated from Doppler-only measurements in a multistatic system. Though we show that use of a single transmitter and receiver is insufficient, we further establish that a Doppler-only multistatic system is feasible, provided enough transmitters and receivers are available. To be precise, we show that determining the target state from Doppler shifts is equivalent to solving a nonlinear optimization problem, namely the minimization of a certain objective function. We then solve this problem numerically, using an original analytic derivation of the objective function's gradient and Hessian.

1.1 History of Multistatic Radar

Multistatic radars were first theoretically addressed in 1917. The August edition of *The Electrical Experimenter* included editor Hugo Gernsbach's interview of Nikola Tesla on "subjecting [submerged] enemy submarines" [30]. Tesla said:

> ...consider that a concentrated ray from a searchlight is thrown on a balloon at night. When the spot of light strikes the balloon, the latter at once becomes visible from many different angles. The same effect would be created with the electric ray if properly applied. When the ray struck the rough hull of a submarine it would be reflected, but not in a concentrated beam — it would spread out; which is just what we want. Suppose several vessels are steaming along in company; it thus becomes evident that several of them will intercept the reflected ray and accordingly be warned of the presence of the submarine... [34]

Despite Tesla's anticipation for such technology, five years passed with no successful implementation of his idea. Then, in September 1922, U.S. Navy civilian engineer Dr Albert Hoyt Taylor and his assistant Leo Clifford Young, both of the United States Naval Aircraft Radio Laboratory at Anacostia, DC, conducted tests at the Naval Air Station, where "audible maxima and minima caused by reflections from steel buildings were observed." After this success, they put the receiver in a car and drove it across the Potomac River. This time, reflections occurred from several objects, including a modest-sized wooden steamer passing along the river. Despite Taylor and Young's success, their work received little support, and they instead turned their focus towards ionic sounding experiments, a precursor of modern radar experiments [39, pp. 17-18].

By the early 1930s, the employment of radar had become worldwide. Interestingly, these original radars were primitive forms of multistatic radar, as conventional radar would require the then-undeveloped technology of switching an antenna from transmit to receive mode. However, the invention of the synchronizer in the mid-1930s quickly shifted the world's focus to just such monostatic radar, as dealing with multistatic radar proved cumbersome. Consequently, of the many important

developments in radar made during the Second World War, few were made in the field of multistatic radar.

The Cold War induced an era of nuclear fear in the United States. Once the Soviet Union launched Sputnik I in 1957, detecting and tracking such satellites became a major national security issue [26]. To help remedy the situation, a wide range of space-monitoring systems were deployed, and many are still in use. Many of these sensors were designed to warn of strategic missile attacks [32]. Multistatic radar technology proved to be quite useful in this effort. For example, the Multistatic Measurements System, completed in 1980, used time-of-arrival and Doppler information from multiple transmitters and receivers to locate and track airborne objects such as ballistic missiles [39, p. 55].

Another significant application of multistatic radar was the continuous-wave radar "fence" known as the Navy Space Surveillance System (NavSpaSur) [32]. This system was developed in 1958 to detect and track satellites which transmitted signals, as well as those that were "quiet." It consists of three transmitters and six receivers from California to Georgia, as well as a computational center at the NavSpaSur headquarters in Dahlgren, Virginia. The system uses angle measurements from two sites to estimate a target's location/velocity [39, p. 39], and is capable of doing so for basketball-sized objects up to 7,500 miles in altitude [26].

Today, NavSpaSur remains an essential element of the U.S. Space Command Detection and Tracking System, as it is used for several purposes. It helps maintain and update the database of orbiting objects/debris [26] and provides continuous surveillance of space objects orbiting over the contiguous United States [32]. Further, NavSpaSur is "the only space surveillance system which provides satellite vulnerability data to fleet units" [32].

In 1999, Lockheed Martin created a multistatic system, called Silent Sentry 2, that uses time-of-arrival and Doppler measurements of everyday broadcast signals to detect and track targets. Also, there are multiple multistatic systems, described

throughout the Internet, that are used by radio-astronomers to track meteors entering the atmosphere.

Having given a broad sense of the history of multistatic radar, we now review the current literature, emphasizing the work related to Doppler-only systems.

1.2 Literature Review

Little has been written on Doppler-only multistatic radar. However, there is a great deal of literature on other forms of radar, some of which is pertinent to our problem. The types of radar are defined by how the transmitter, receiver and target are related. For example, the typical *active* radar is characterized by a collocated transmitter and receiver. A radar is *passive* when the target is either the transmitter or the receiver. Finally, *semi-active* radar has a distinct transmitter, receiver and target. Note that multistatic radar is often considered semi-active, since the receivers tend to have little or no control of transmission. Though the literature sometimes refers to this form of radar as "passive," we shall strictly use "passive" to denote the transmitting or receiving target scenario.

1.2.1 Active Radar. Monostatic radar has been exhaustively researched. However, the significant structural differences between monostatic and bistatic systems make using monostatic results challenging. Nevertheless, there exist useful results for our problem. For example, in 1980, Levanon [24] discussed derivatives of Doppler-only information, assuming a constant target velocity. Here, he provided explicit expressions for range, velocity, and the angle between the object's position and velocity in terms of the Doppler shift and its derivatives. Two years later, in response to this paper, Webster [36] explained how three Doppler measurements, under the same assumptions, suffice to discern these target state parameters. Chapter IV of this thesis generalizes this use of Doppler derivatives.

1.2.2 Passive Radar. There is a significant amount published in the literature about passive radar. So much, in fact, that our review is broken into three areas: observability, target tracking and stochastic analysis.

1.2.2.1 Observability. Observability refers, in this context, to the existence of a unique tracking solution. Initial discussions [14, 22] on target observability in passive radar date back to the early 1970s, and focused on general nonlinear systems. In 1981, Shensa [31] used geometric considerations to derive sufficient observability conditions in Doppler-only tracking. Since then, it seems that addressing observability with passive radar was strictly limited to angle measurements, that is, until Becker [4] discussed observability from angle and *frequency* measurements. In doing so, he provided simpler and more general necessary and sufficient observability conditions.

1.2.2.2 Target Tracking. One problem, given by Weinstein and Levanon in 1980, was to track a transmitter moving in a ballistic trajectory. Understanding ballistic motion, they were able to determine a weighted least-squares estimate of the track parameters [38]. Seven years later, Statman and Rodemich [33] introduced a method for real-time estimation; that is, a method which did not involve computationally expensive iterative least-squares estimation. This method was used in the simulation of an area weapons effects system that the Jet Propulsion Laboratory developed for the U.S. Army.

In 1982, Weinstein considered the case of a passive radar in which an array of collinear receivers yields Doppler measurements for tracking [37]. Knowing the time differences of arrival permitted estimation of position from the Doppler information, along with subsequent velocity estimation.

The 1990s proved to be a developing decade for Doppler-only passive radar. Research shifted from special cases to more applicable results. In particular, Chan and Jardine gave, analogous to [24] and [36], a method for estimating target state

purely from *sonar* Doppler and rate of Doppler change information, see [6]. Radar and sonar applications are very similar, differing only in signal type and medium of passage. They used a grid-search to determine an initial estimate and a Kalman filter to reduce the search for later estimates while tracking. In [8], Chan and Towers introduced a more efficient way to search by using intermediate equations to effectively decrease the search dimension. Chan and Towers [9] applied this idea to Doppler-only passive radar with multiple receivers, dubbed "sequential localization," so that initial estimation could occur earlier in real time.

Levesque and Bondaryk [25] later provided some insight on sequential localization, as they applied it to state estimation of submarine targets. In particular, they found that the use of more than four receivers is helpful only in certain array configurations. Armstrong and Holeman [3] considered estimation of a slightly more generalized target state, where target acceleration is assumed to be some nonzero constant. With the help of time-differences-of-arrival information, Ho and Xu [17] established an algebraic, noniterative solution to the target state. Finally, Becker [5] generalized previously studied two-dimensional tracking with angle and Doppler measurements by considering the problem in three dimensions. In particular, he applied the problem to Airborne Warning and Control System (AWACS) scenarios.

1.2.2.3 Stochastic Analysis. When considering a radar signal, there is inevitably noise in the reception. The noise is often modelled as random, which is then addressed in research through stochastic analysis. For example, in 1979, Schultheiss and Weinstein [29] exploited the structure of Gaussian noise to estimate Doppler shift in the received signal. Abel [1] used stochastic analysis to determine an optimal array configuration of receivers for target state estimation. Jauffret and Bar-Shalom [20] estimated the target state from bearing and Doppler information from a low signal-to-noise ratio environment, where false detections are common. Lastly, Chan and Rudnicki [7] expanded on this by using instrumental variables, developed from history measurements, to yield a recursive, unbiased target estimate.

1.2.3 Semi-Active Radar. There is much discussed in the literature about using bistatic radar cross sections for imaging [12, 13, 23, 40, 41]. There is also a small amount of research on the incorporation of clutter in bistatic radar [16, 21, 42]. This thesis will not delve further into such research, as we are only concerned with estimating the target state.

Much of the contemporary research in bistatic radar has focused on the ambiguity function. Tsao *et al.* analyze this function [35], as it seems to arise when observing a low-velocity target in white Gaussian noise. Ringer and Frazer [28] used the ambiguity function to help assess the feasibility of multistatic radar; their conclusion was that such radar is, in fact, viable. Griffiths *et al.* [15] recently gave some measurements of the ambiguity functions and commented on their form regarding bistatic radar systems for localization.

Ironically, the research that has shown the most promise in Doppler-only multistatic radar used angle information along with Doppler measurements. In 1995, Howland [18] demonstrated a working system that would detect, locate and track airborne targets. After detecting the Doppler shift of a specific target over time, known as the *trace*, he used a Kalman filter to help complete the trace. He then used a modified Levenberg-Marquardt algorithm to determine a target state estimate that would minimize the difference between expected and measured data. Four years later, he published a more elaborate paper [19] that introduced the use of a genetic algorithm to initialize the minimization close enough to the true optimal target state estimate.

Other research includes Dommermuth's model of bistatic Doppler shift as a probabilistic quantity [11]. Here, he uses this stochastic assumption to achieve a probability distribution for the target state. Two stochastic filters for bistatic surveillance are presented in [16]. One uses a Kalman filter and Doppler information, while the other uses a particle filter to both track and classify targets. We will not pursue any of this further, as we shall attempt to discern the mathematical structure em-

bedded within multistatic and bistatic radar, and thereby retrieve a deterministic solution.

1.3 Thesis Outline

In Chapter II, we introduce the basic principles of the simplest case of multistatic radar, known as bistatic radar. Here, we explain what the Doppler effect actually is and what it tells about the target. This chapter motivates the remainder of the thesis. Chapter III explains the calculus we will need to proceed. We then assess the observability of the bistatic system with only Doppler information in Chapter IV. Next, Chapter V provides a minimization algorithm to determine the target state given multistatic Doppler measurements. Finally, Chapter VI considers both the asymptotics of the objective function used in our minimization algorithm and the optimality of multistatic configurations of transmitters and receivers.

II. Basic Principles of Doppler-Only Bistatic Radar

In this chapter, we present a mathematical treatment of the physical phenomenon known as the Doppler effect. The results presented herein are commonly known — we only include them for ease of understanding and completeness.

The simplest semi-active radar system is one with a single stationary transmitter and a distinct stationary receiver; this is called a *bistatic* system. The transmitter sends an omnidirectional signal with an assumed constant speed of light c. Once the signal collides with an object, it is emitted a second time from the collision location and both the original and reflected signals reach the receiver.

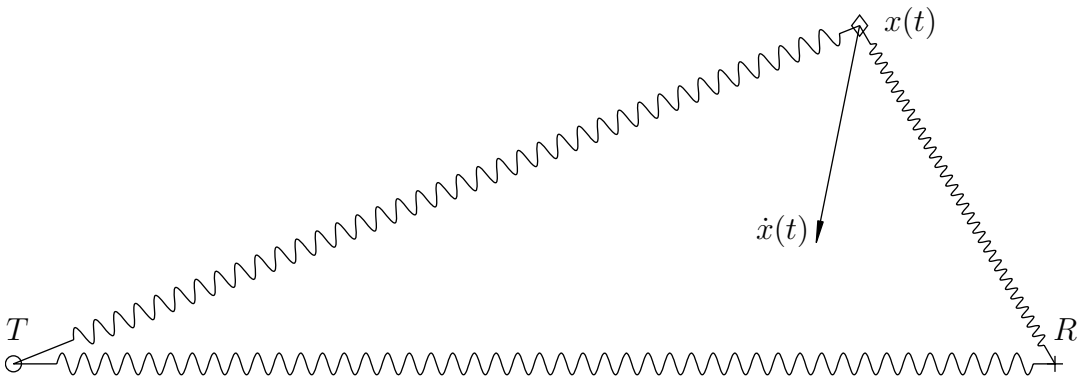

Figure 1 The Bistatic System

To be precise, consider the arrangement depicted in Figure 1, with a fixed transmitter and receiver located at positions $T, R \in \mathbb{R}^N$, respectively, and a moving object whose location at time t is given by the function $x : \mathbb{R} \to \mathbb{R}^N$. Let s_T denote the transmitted signal, s_R the directly received signal, and s_x the reflected signal. Though the received signal will be some superposition of s_R with s_x, we assume that s_R and s_x have been correctly identified. In practice, one usually relies upon well-established signal processing-based techniques to accomplish this decomposition.

As the direct signal travels from T to R at the speed of light, we have

$$s_R(t) = s_T(t - |R - T|/c). \tag{1}$$

Similarly, the reflected signal is

$$s_x(t) = s_T(t - \varrho(t)/c), \tag{2}$$

where $\varrho(t)$ is the length of the path travelled by the reflected signal which was received at time t. As the receiver and transmitter are stationary, s_R is simply a time-delayed version of s_T. However, even when T and R are fixed, the delay in the reflected signal, namely $\varrho(t)/c$, will not be constant in general. That is, s_x is a time-delayed version of s_T, where the delay itself is a function of time. In the case where s_T is sinusoidal, the nonconstant time-delay causes a phenomenon known as the Doppler effect.

2.1 Derivation of the Doppler Effect

For the sake of mathematical simplicity, let us assume that our transmitted signal is complex with unit amplitude

$$s_T(t) = \mathrm{e}^{2\pi \mathrm{i} \alpha(t)},$$

where $2\pi\alpha(t)$ is the angle of the signal in the complex plane at time t. It follows that $\alpha(t + h) - \alpha(t)$ is the number of cycles that the signal has traversed over an interval of time $[t, t + h]$, and therefore the ratio

$$\frac{\alpha(t + h) - \alpha(t)}{h}$$

is the mean frequency over the same interval. Since the instantaneous frequency is defined to be the limit of average frequencies as the length of time over which the

average was taken tends to zero, we have

$$\mathrm{freq}(s_T(t)) \equiv \lim_{h\to 0} \frac{\alpha(t+h)-\alpha(t)}{h} = \dot{\alpha}(t). \tag{3}$$

In this case, the directly received (1) and reflected signals (2) are also sinusoidal:

$$s_R(t) = s_T(t-|R-T|/c) = \mathrm{e}^{2\pi \mathrm{i}\alpha(t-|R-T|/c)},$$
$$s_x(t) = s_T(t-\varrho(t)/c) = \mathrm{e}^{2\pi \mathrm{i}\alpha(t-\varrho(t)/c)}.$$

It follows from (3) that the frequency of the direct signal is

$$\mathrm{freq}(s_R(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\alpha(t-|R-T|/c) = \dot{\alpha}(t-|R-T|/c), \tag{4}$$

while the frequency of the reflected signal is

$$\mathrm{freq}(s_x(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\alpha(t-\varrho(t)/c) = \dot{\alpha}(t-\varrho(t)/c)(1-\dot{\varrho}(t)/c). \tag{5}$$

We now note that in practice, a receiver in a bistatic system may only have knowledge of $s_R(t)$ and $s_x(t)$, with little to no knowledge of the original transmission. That is, one may know the frequencies (4) and (5) without having the original frequency $\dot{\alpha}(t)$. Fortunately, this difficulty can be overcome when the frequency of the original signal changes slowly. In particular, if the original frequency is constant, namely $\dot{\alpha}(t) = \beta$ for all t, then (4) and (5) give

$$\mathrm{freq}(s_x(t)) = \beta(1-\varrho'(t)/c) = \beta - \frac{\beta}{c}\dot{\varrho}(t) = \mathrm{freq}(s_R(t)) - \frac{1}{\lambda}\dot{\varrho}(t) \tag{6}$$

where $\lambda = c/\beta$ is the wavelength of the original signal. We now note that, even in the case where the frequency of the original transmission is constant, the frequency of the reflected signal s_x is not exactly that of the direct signal s_R. In fact, the gap between these two received frequencies is proportional to the rate of change of the

reflected distance $\varrho(t)$ with respect to time. This change in frequency, caused by the nonconstant time-delay in the expression for s_x, is known as a Doppler shift, and is explicitly derived from (6):

$$\text{freq}(s_x(t)) - \text{freq}(s_R(t)) = -\frac{1}{\lambda}\dot{\varrho}(t). \tag{7}$$

Thus, if one, through appropriate time-frequency analysis, explicitly computes this shift, one may, in fact, determine something about the target, namely $\dot{\varrho}(t)$. Unfortunately, the exact relationship between $\dot{\varrho}(t)$ and the position of the object $x(t)$ is quite complicated. Nevertheless, when the object is moving significantly slower than the speed of light, the quantity $\dot{\varrho}(t)$ is extremely close to another quantity which has a relatively simple dependence upon $x(t)$, which we now discuss.

2.2 Start-Stop Approximation

For the bistatic system depicted in Figure 1, recall that the quantity $\varrho(t)$ represents the length of the path travelled by the reflected signal which was received at time t. From the point of view of the receiver operator, this quantity is difficult to determine, as the target's location is unknown. However, when the speed of the target is significantly less than the speed of light, the distance $\varrho(t)$ is almost exactly the same as the distance from transmitter to moving body to receiver at the time of reception, namely the bistatic distance

$$\rho(t) = |x(t) - T| + |x(t) - R|. \tag{8}$$

As $\rho(t) \cong \varrho(t)$ when $|\dot{x}(t)| \ll c$, one intuitively expects $\dot{\rho}(t) \cong \dot{\varrho}(t)$, and so the Doppler shift given in (7) is approximately proportional to the rate of change of the bistatic distance. To discuss these matters more formally, consider the following

proposition, recalling the standard dot product on $\mathbb{R}^N$,

$$x \cdot y \equiv \sum_{n=1}^{N} x_n y_n,$$

which leads to the Euclidean norm,

$$|x| \equiv \sqrt{x \cdot x} = \left[\sum_{n=1}^{N} x_n^2 \right]^{1/2}.$$

Proposition 2.2.1. *If $x : \mathbb{R} \to \mathbb{R}^N$ is continuously differentiable and $\sup |\dot{x}(t)| < c$, then there exists a unique $t_0 \geq 0$ such that $|x(t_0)| = ct_0$.*

Proof: To show that t_0 exists, note that if $|x(0)| = 0$, we are done. Suppose $|x(0)| > 0$, let $a \equiv \sup |\dot{x}(t)|$, and consider the continuous function defined by $y(s) \equiv |x(s)| - cs$ for all $s \geq 0$. Thus, $y(0) = |x(0)| > 0$. Furthermore, taking t such that $t > |x(0)|/(c-a)$, we have

$$|x(t)| \leq |x(0)| + \left| \int_0^t \dot{x}(t)\, \mathrm{d}t \right| \leq |x(0)| + at < ct,$$

and so $y(t) < 0$. The Intermediate Value Theorem then gives $t_0 \in (0, t)$ such that $y(t_0) = 0$, that is, $|x(t_0)| = ct_0$. To show that t_0 is unique, suppose $t_0 < t_1$ both satisfy the condition, namely $ct_0 = |x(t_0)|$ and $ct_1 = |x(t_1)|$. The Reverse Triangle Inequality then gives

$$c(t_1 - t_0) = |x(t_1)| - |x(t_0)| \leq |x(t_1) - x(t_0)| = \left| \int_{t_0}^{t_1} \dot{x}(t)\, \mathrm{d}t \right| \leq a(t_1 - t_0). \qquad (9)$$

Since $t_1 > t_0$, (9) implies $c \leq a$, a contradiction. ■

Returning to the bistatic system of Figure 1, note that for any time $t \in \mathbb{R}$, Proposition 2.2.1 guarantees that there exists a unique time $\tau < t$ such that

$$|x(\tau) - R| = c(t - \tau). \qquad (10)$$

This time τ corresponds to the moment when the signal which was received at time t had reflected off of the moving object. Thus, the distance travelled by the signal received at time t was actually the bistatic distance (8) at time $\tau < t$, that is, $\varrho(t) = \rho(\tau)$. This implies that

$$\dot{\varrho}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\rho(\tau) = \dot{\rho}(\tau)\frac{\mathrm{d}\tau}{\mathrm{d}t}. \tag{11}$$

Further, since

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}\Big[|x(\tau) - R|\Big] &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left[\sum_{n=1}^{N}(x_n(\tau) - R_n)^2\right]^{1/2} \\ &= \left[\sum_{n=1}^{N}(x_n(\tau) - R_n)^2\right]^{-1/2}\sum_{n=1}^{N}(x_n(\tau) - R_n)\dot{x}_n(\tau) \\ &= \frac{x(\tau) - R}{|x(\tau) - R|}\cdot\dot{x}(\tau),\end{aligned} \tag{12}$$

we have, by differentiating (10), that

$$\left(\frac{x(\tau) - R}{|x(\tau) - R|}\cdot\dot{x}(\tau)\right)\frac{\mathrm{d}\tau}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}|x(\tau) - R| = c\left(1 - \frac{\mathrm{d}\tau}{\mathrm{d}t}\right).$$

From here, rearranging gives

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = c\left[\frac{x(\tau) - R}{|x(\tau) - R|}\cdot\dot{x}(\tau) + c\right]^{-1},$$

so that, considering (11), we have by Triangle and Reverse Triangle Inequalities that

$$\left|\frac{\dot{\varrho}(t)}{\dot{\rho}(\tau)} - 1\right| = \left|\frac{\mathrm{d}\tau}{\mathrm{d}t} - 1\right| \leq \frac{|\dot{x}(\tau)|}{c - |\dot{x}(\tau)|}.$$

That is, $\dot{\rho}(t) \cong \dot{\varrho}(t)$ when $|\dot{x}(t)| \ll c$, and so the Doppler shift given in (7) is approximately proportional to the rate of change of the bistatic distance, as desired. Let us now determine this rate of change.

Considering (12), we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[|x(t)-T|\Big] = \frac{x(t)-T}{|x(t)-T|}\cdot\dot{x}(t), \quad \frac{\mathrm{d}}{\mathrm{d}t}\Big[|x(t)-R|\Big] = \frac{x(t)-R}{|x(t)-R|}\cdot\dot{x}(t),$$

so that the Doppler shift is approximately proportional to the rate of change of the bistatic distance

$$\dot{\rho}(t) = \left(\frac{x(t)-T}{|x(t)-T|} + \frac{x(t)-R}{|x(t)-R|}\right)\cdot\dot{x}(t). \tag{13}$$

Note that (13) is symmetric between the transmitter and receiver locations, meaning that interchanging their locations will not change the Doppler measurement. Also notice that (13) is an underdetermined vector differential equation. It follows that if enough bistatic measurements are made using multiple distinct transmitters and receivers, a solution for $x(t)$ could possibly be resolved if only an initial condition were established.

Figure 2 illustrates the two vectors whose dot product yields the rate of change of the bistatic distance. The ellipse depicted about foci T and R is a curve of constant

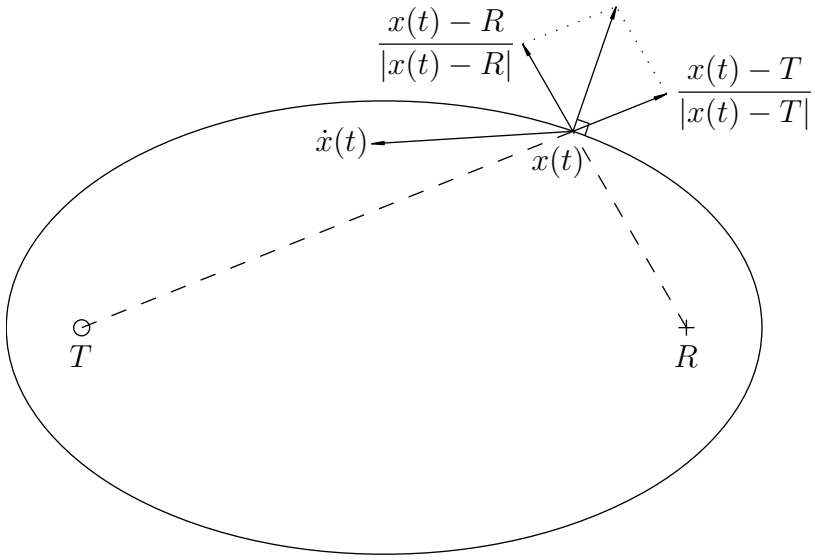

Figure 2 Geometry of the Bistatic System

bistatic distance. In $\mathbb{R}^3$, the contour takes the form of a prolate spheroid with the

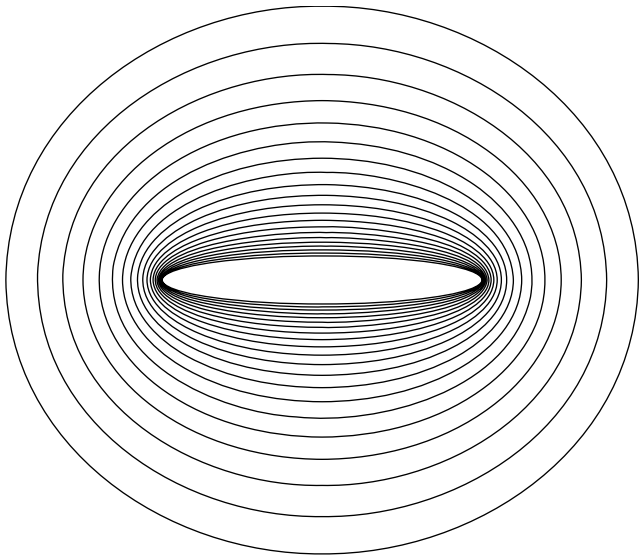

Figure 3 Curves of Constant Bistatic Distance

same foci. Note that the sum of the bearing vectors is orthogonal to this contour. This is given by the fact that the gradient of a functional is orthogonal to its level curves. In particular, the prolate spheroid is the level curve of ρ and the bearing sum is its gradient.

Thus, the quantity $\dot{\rho}(t)$ represents a scaled version of the "amount" of $\dot{x}(t)$ that is crossing the level curve $|x-T|+|x-R| = \rho(t)$ at time t. In particular, if the target is moving towards the inside of the ellipsoid, then the dot product $\dot{\rho}(t)$ is positive, and s_x, in turn, has a higher frequency than s_R. Meanwhile, if the target is moving towards ellipsoids of greater size, then the bistatic distance $\rho(t)$ is increasing. In this case, $\dot{\rho}(t) > 0$, and so the Doppler shift is negative. Finally, an object travelling along the surface of such an ellipsoid would produce no Doppler shift at all.

It is important to understand that these confocal ellipsoids are not parallel by any means. Consider the list of contours illustrated in Figure 3. We notice two asymptotic tendencies of the ellipsoids. Ellipsoids that tend closer to the foci approach the line segment connecting the foci, while the larger ellipsoids approach

the sphere centered between the foci of radius $\rho/2$. Let us now consider an example of the bistatic distance derivative.

2.3 An Example

Consider the bistatic system illustrated in Figure 1, in which the transmitter is located at T, the receiver at R, and target at $x(t)$. If the reflected signal is received at time $t = 0$, then by Proposition 2.2.1, there exists a unique time $\tau < 0$ at which the target reflected this signal. Furthermore, this time satisfies

$$-c\tau = |x(\tau) - R|,$$

where c is the speed of light. Suppose the target is travelling with constant velocity. Then

$$x(t) = x + th,$$

for all t, for some vectors $x, h \in \mathbb{R}^N$, where N is the spatial dimension. Solving for the time τ of reflection, we have

$$c^2\tau^2 = |x + \tau h - R|^2 = |x - R|^2 - 2(x - R) \cdot h\tau + |h|^2\tau^2,$$

or equivalently,

$$(c^2 - |h|^2)\tau^2 + 2(x - R) \cdot h\tau - |x - R|^2 = 0.$$

Solving for $t < 0$ via the Quadratic Formula yields

$$\tau = -\frac{1}{c^2 - |h|^2}\left\{(x - R) \cdot h + \left[[(x - R) \cdot h]^2 - (c^2 - |h|^2)|x - R|^2\right]^{1/2}\right\}. \qquad (14)$$

The speed of light being large, we see that $\tau \cong 0$, as desired. Applying the previous section's notation, the distance actually travelled by the signal received at time zero

is expressed by

$$\varrho(0) = |x(\tau) - T| + |x(\tau) - R|,$$

for τ given in (14) above. The actual distance between the transmitter and the receiver at time zero is given by

$$\rho(0) = |x - T| + |x - R|.$$

For $\tau \cong 0$, these two distances are extremely close, and so there is no great harm in using $\rho(0)$ instead of $\varrho(0)$.

Let us now determine $\dot{\rho}(t)$ for all t. To do this, we will first consider the term

$$\frac{x(t) - T}{|x(t) - T|} \cdot \dot{x}(t).$$

By completing a square, we see that the square of the denominator is

$$\begin{aligned} |x(t) - T|^2 &= |x - T + th|^2 \\ &= |x - T|^2 + 2(x - T) \cdot ht + |h|^2 t^2 \\ &= |h|^2 \left[\left(\frac{|x - T|}{|h|} \right)^2 - \left(\frac{(x - T) \cdot h}{|h|^2} \right)^2 + \left(t + \frac{(x - T) \cdot h}{|h|^2} \right)^2 \right] \\ &= |h|^2 B_T^2 \Big[1 + [(t - A_T)/B_T]^2 \Big], \end{aligned} \tag{15}$$

where

$$A_T = -\frac{(x - T) \cdot h}{|h|^2}, \quad B_T = \left[\left(\frac{|x - T|}{|h|} \right)^2 - \left(\frac{(x - T) \cdot h}{|h|^2} \right)^2 \right]^{1/2}.$$

Note that the radicand of B_T is positive by the Cauchy-Schwartz Inequality. We also have numerator

$$(x(t) - T) \cdot \dot{x}(t) = (x + th - T) \cdot h$$

$$= |h|^2 t + (x - T) \cdot h$$
$$= |h|^2 B_T[(t - A_T)/B_T]. \tag{16}$$

Combining (15) and (16) gives the fraction:

$$\frac{x(t) - T}{|x(t) - T|} \cdot \dot{x}(t) = |h| \left[\frac{(t - A_T)/B_T}{\sqrt{1 + [(t - A_T)/B_T]^2}} \right].$$

Simple replacement of T with R yields a similar expression:

$$\frac{x(t) - R}{|x(t) - R|} \cdot \dot{x}(t) = |h| \left[\frac{(t - A_R)/B_R}{\sqrt{1 + [(t - A_R)/B_R]^2}} \right],$$

and so we have the following rate of change of bistatic distance:

$$\dot{\rho}(t) = |h| \left[\frac{(t - A_T)/B_T}{\sqrt{1 + [(t - A_T)/B_T]^2}} + \frac{(t - A_R)/B_R}{\sqrt{1 + [(t - A_R)/B_R]^2}} \right].$$

We see that this rate is merely the sum of two shifted and dilated functions of the form

$$\frac{t}{\sqrt{1 + t^2}}.$$

Thus, even in this case where the target's travel is linear, that is, $\ddot{x} = 0$, there are still five unknown quantities that determine the Doppler shift. Therefore, any successful application of Doppler-only radar will require several measurements over time and/or multiple transmitter-receiver pairs; that is, multistatic radar.

2.4 What Remains

In Chapter IV, we determine the theoretical limitations of using multiple time steps in a bistatic system. Mathematically, having multiple time measurements is similar to having knowledge of the time derivatives of the Doppler signal $\dot{\rho}(t)$. We will see that even in this case, while assuming constant velocity, the most one can

ever expect to determine about the target's position is the target's distances from the transmitter and receiver. Discovering this fact, we see that a Doppler-only requirement forces us to use a multistatic system. But, before we can attempt to consider either system, we must first develop a better understanding of the underlying calculus.

III. Distance and Bearing Derivatives

As discussed in the previous chapter, Doppler information is proportional to the rate at which the bistatic distance changes. Such rates of change are, by definition, derivatives. In this chapter, we develop the mathematical groundwork upon which the rest of our work will be based. Though the techniques used in this chapter are standard tools of multivariable Taylor series and tensor analysis, we could not find the results given below in the literature. As our motivating problem is the only one of which we know that requires these derivatives, we believe these results could be original.

3.1 Fréchet Derivatives

The Fréchet derivative of a differentiable function $f : \mathbb{R}^{N_1} \to \mathbb{R}^{N_2}$ at x is the linear operator $\mathrm{D}f(x) : \mathbb{R}^{N_1} \to \mathbb{R}^{N_2}$ such that

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - \mathrm{D}f(x)h|}{|h|} = 0.$$

Meanwhile, for any $M > 1$, the Mth Fréchet derivative of f at x is recursively defined to be the symmetric M-multilinear operator $\mathrm{D}^M f(x) : \mathbb{R}^{M \times N_1} \to \mathbb{R}^{N_2}$ that satisfies

$$\lim_{h \to 0} \frac{1}{|h|^M} \left| f(x+h) - \sum_{m=0}^{M} \frac{1}{m!} \mathrm{D}^m f(x) h^m \right| = 0.$$

Moreover, when f is real-valued and $(M+1)$-times Fréchet differentiable at x, then the $(M+1)$th derivative of f at x in the directions $(h_1, \ldots, h_{M+1})$ is the dot product of final direction h_{M+1} with the gradient of the Mth derivative in the directions $(h_1, \ldots, h_M)$:

$$\mathrm{D}^{M+1} f(x)\{h_m\}_{m=1}^{M+1} = [\nabla_x (\mathrm{D}^M f(x)\{h_m\}_{m=1}^{M})] \cdot h_{M+1}. \tag{17}$$

Thus, the "entries" of $\mathrm{D}^M f(x)$, given by $\mathrm{D}^M f(x)\{e_{k_m}\}_{m=1}^M$, are the Mth mixed partial derivatives of f. The background material for these generalized derivatives, including a generalized Product Rule and Chain Rule, is given in the appendix.

3.2 The Euclidean Norm Functional

The first derivative of the paraboloid $|\cdot|^2 : \mathbb{R}^N \to \mathbb{R}$ is

$$\mathrm{D}|x|^2 h = \nabla |x|^2 \cdot h = 2x \cdot h, \tag{18}$$

where $\nabla |x|^2 = 2x$ is the gradient of $|x|^2$. Meanwhile, the second derivative of $|x|^2$ is

$$\mathrm{D}^2|x|^2(h_1, h_2) = 2h_1 \cdot h_2, \tag{19}$$

since the Hessian of $|x|^2$ is $2I$. All higher order derivatives of $|x|^2$ are zero:

$$\mathrm{D}^M|x|^2\{h_m\}_{m=1}^M = 0, \ M > 2. \tag{20}$$

Using the Product Rule, one may then use these derivatives of $|x|^2$ to compute the derivatives of $|x|$. In particular, using the Product Rule, (18) becomes

$$2x \cdot h = \mathrm{D}|x|^2 h = 2|x|\mathrm{D}|x|h,$$

and so the first derivative of the Euclidean norm functional is

$$\mathrm{D}|x|h = \frac{x}{|x|} \cdot h = b \cdot h \tag{21}$$

where $b = x/|x|$. Similarly, under two applications of the Product Rule, (19) becomes

$$2h_1 \cdot h_2 = \mathrm{D}^2|x|^2(h_1, h_2) = 2|x|\mathrm{D}^2|x|(h_1, h_2) + 2\mathrm{D}|x|h_1\mathrm{D}|x|h_2. \tag{22}$$

Using the expression of the first derivative (21), we may further simplify (22),

$$2h_1 \cdot h_2 = 2|x|\mathrm{D}^2|x|(h_1, h_2) + 2\,(b \cdot h_1)\,(b \cdot h_2)\,,$$

and so the second derivative of the norm is

$$\mathrm{D}^2|x|(h_1, h_2) = \frac{1}{|x|}[h_1 \cdot h_2 - (b \cdot h_1)(b \cdot h_2)]. \tag{23}$$

Continuing in this manner, we may use the Product Rule as well as the first and second derivatives of $|x|$ above to find the third derivative. Specifically, we have

$$\begin{aligned}
0 &= \frac{1}{2}\mathrm{D}^3|x|^2(h_1, h_2, h_3) \\
&= |x|\mathrm{D}^3|x|(h_1, h_2, h_3) + \mathrm{D}|x|h_3\mathrm{D}^2|x|(h_1, h_2) \\
&\quad + \mathrm{D}|x|h_2\mathrm{D}^2|x|(h_1, h_3) + \mathrm{D}|x|h_1\mathrm{D}^2|x|(h_2, h_3) \\
&= |x|\mathrm{D}^3|x|(h_1, h_2, h_3) + \frac{b \cdot h_3}{|x|}[h_1 \cdot h_2 - (b \cdot h_1)(b \cdot h_2)] \\
&\quad + \frac{b \cdot h_2}{|x|}[h_1 \cdot h_3 - (b \cdot h_1)(b \cdot h_3)] + \frac{b \cdot h_1}{|x|}[h_2 \cdot h_3 - (b \cdot h_2)(b \cdot h_3)],
\end{aligned}$$

and so the third derivative of the norm is

$$\begin{aligned}
\mathrm{D}^3|x|(h_1, h_2, h_3) = \frac{1}{|x|^2}[&3(b \cdot h_1)(b \cdot h_2)(b \cdot h_3) \\
&- (b \cdot h_1)(h_2 \cdot h_3) - (b \cdot h_2)(h_1 \cdot h_3) - (b \cdot h_3)(h_1 \cdot h_2)].
\end{aligned} \tag{24}$$

Although it is cumbersome to determine the Mth derivative of the Euclidean norm functional in this way, we are able to observe a pattern in (21), (23) and (24): each expression contains dot products of the bearing b with the directions h_m, as well as dot products of the directions with each other.

This pattern inspired the following result, in which we find a closed form expression for a general Mth-order derivative of $|x|$. As far as we know, this result is

original. However, due to the long history and wealth of literature on the topic, as well as the relative simplicity of the techniques used in our proof, we suspect the result may already be known, perhaps as folklore. The result itself makes use of the double factorial function, recursively defined as

$$n!! \equiv \begin{cases} n(n-2)!! & \text{for } n \geq 1, \\ 1 & \text{for } n = 0, -1. \end{cases}$$

Thus, for n positive and even, $n!!$ is the product of all positive even numbers less than or equal to n, while for n positive and odd, $n!!$ is the product of all positive odd numbers less than or equal to n.

We also introduce a new symbol to remove the need for tensor notation in our expression for the M-multilinear function $\mathrm{D}^M|x|$. In particular, let $\{h_m\}_{m=1}^M \subseteq \mathbb{R}^M$ and take $A \subseteq \{1, \ldots, M\}$ such that the cardinality of A, denoted $|A|$, is even. Let $P(A)$ be the set of all partitions of A into pairs. The distinct product of $\{h_k\}_{k \in A}$ is

$$\mathrm{dp}\{h_k\}_{k \in A} \equiv \sum_{\mathcal{P} \in P(A)} \prod_{\{h_i, h_j\} \in \mathcal{P}} (h_i \cdot h_j).$$

Thus, the distinct products of a collection of two and four vectors are

$$\begin{aligned} \mathrm{dp}(h_1, h_2) &= h_1 \cdot h_2, \\ \mathrm{dp}(h_1, h_2, h_3, h_4) &= (h_1 \cdot h_2)(h_3 \cdot h_4) + (h_1 \cdot h_3)(h_2 \cdot h_4) + (h_1 \cdot h_4)(h_2 \cdot h_3), \end{aligned}$$

respectively, while the distinct product of a collection of six vectors is

$$\begin{aligned} \mathrm{dp}(h_1, h_2, h_3, h_4, h_5, h_6) = {} & (h_1 \cdot h_2)(h_3 \cdot h_4)(h_5 \cdot h_6) + (h_1 \cdot h_2)(h_3 \cdot h_5)(h_4 \cdot h_6) \\ & + (h_1 \cdot h_2)(h_3 \cdot h_6)(h_4 \cdot h_5) + (h_1 \cdot h_3)(h_2 \cdot h_4)(h_5 \cdot h_6) \\ & + (h_1 \cdot h_3)(h_2 \cdot h_5)(h_4 \cdot h_6) + (h_1 \cdot h_3)(h_2 \cdot h_6)(h_4 \cdot h_5) \\ & + (h_1 \cdot h_4)(h_2 \cdot h_3)(h_5 \cdot h_6) + (h_1 \cdot h_4)(h_2 \cdot h_5)(h_3 \cdot h_6) \end{aligned}$$

$$
\begin{aligned}
&+ (h_1 \cdot h_4)(h_2 \cdot h_6)(h_3 \cdot h_5) + (h_1 \cdot h_5)(h_2 \cdot h_3)(h_4 \cdot h_6) \\
&+ (h_1 \cdot h_5)(h_2 \cdot h_4)(h_3 \cdot h_6) + (h_1 \cdot h_5)(h_2 \cdot h_6)(h_3 \cdot h_4) \\
&+ (h_1 \cdot h_6)(h_2 \cdot h_3)(h_4 \cdot h_5) + (h_1 \cdot h_6)(h_2 \cdot h_4)(h_3 \cdot h_5) \\
&+ (h_1 \cdot h_6)(h_2 \cdot h_5)(h_3 \cdot h_4).
\end{aligned}
$$

By convention, the distinct product of the empty set is one: $\mathrm{dp}\emptyset \equiv 1$.

Theorem 3.2.1. *For any $M \geq 1$, the Mth derivative of the Euclidean norm is*

$$
\mathrm{D}^M |x| \{h_m\}_{m=1}^M = \frac{1}{|x|^{M-1}} \sum_{\ell=0}^{\lfloor M/2 \rfloor} (-1)^{M+\ell+1} (2M-2\ell-3)!! \sum_{\substack{A \subseteq \{1,\ldots,M\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k \in A} \prod_{j \notin A} \left(h_j \cdot \frac{x}{|x|} \right).
$$

Proof: We first verify the result when $M = 1$. As stated above, we have by convention that $(-1)!! \equiv 1$ and $\mathrm{dp}\emptyset = 1$, and so our conjectured expression for $\mathrm{D}|x|h$ is

$$
\frac{1}{|x|^0} (-1)^2 (-1)!! \, \mathrm{dp}\emptyset \left(h \cdot \frac{x}{|x|} \right) = h \cdot \frac{x}{|x|},
$$

which is consistent with our previous derivation of the first derivative (21). Similarly, for $M = 2$, our conjectured expression for $\mathrm{D}^2|x|(h_1, h_2)$ is

$$
\begin{aligned}
\frac{1}{|x|^1} \left[(-1)^3 1!! \, \mathrm{dp}\emptyset \left(h_1 \cdot \frac{x}{|x|} \right) \left(h_2 \cdot \frac{x}{|x|} \right) + (-1)^4 (-1)!! \, \mathrm{dp}(h_1, h_2) \right] \\
= \frac{1}{|x|} \left[h_1 \cdot h_2 - \left(h_1 \cdot \frac{x}{|x|} \right) \left(h_2 \cdot \frac{x}{|x|} \right) \right]
\end{aligned}
$$

which is consistent with our earlier computation of the second derivative (23). As our result holds for $M = 1, 2$, we proceed by induction, supposing the claim holds for some $M \geq 2$. Letting

$$
c_{M,\ell} \equiv (-1)^{M+\ell+1} (2M - 2\ell - 3)!!, \tag{25}
$$

we have

$$\mathrm{D}^M|x|\{h_m\}_{m=1}^M = \frac{1}{|x|^{M-1}} \sum_{\ell=0}^{\lfloor M/2 \rfloor} c_{M,\ell} \sum_{\substack{A \subseteq \{1,\dots,M\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k\in A} \prod_{j \notin A} \left(h_j \cdot \frac{x}{|x|} \right). \tag{26}$$

To begin evaluating the $M+1$ derivative, recall that by the symmetry of mixed partial derivatives, the $(M+1)$-multilinear function D^{M+1} is symmetric, that is

$$\mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1} = \mathrm{D}^{M+1}|x|\{h_{\sigma(m)}\}_{m=1}^{M+1}$$

for any permutation σ of the indices $\{1,\dots,M+1\}$. In particular, we may make each h_m the final direction, in order to exploit the recursive means of computing higher-order derivatives, as given in (17):

$$\begin{aligned}
\mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1} &= \frac{1}{M+1} \sum_{n=1}^{M+1} \mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1} \\
&= \frac{1}{M+1} \sum_{n=1}^{M+1} \mathrm{D}^{M+1}|x|(h_1,\dots,h_{n-1},h_{n+1},h_{M+1},h_n) \\
&= \frac{1}{M+1} \sum_{n=1}^{M+1} [\nabla(\mathrm{D}^M|x|(h_1,\dots,h_{n-1},h_{n+1},h_{M+1}))] \cdot h_n
\end{aligned} \tag{27}$$

The summands of (27) are given by the inductive hypothesis (26):

$$\begin{aligned}
&\mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1} \\
&\quad = \frac{1}{M+1} \sum_{n=1}^{M+1} \nabla \left[\frac{1}{|x|^{M-1}} \sum_{\ell=0}^{\lfloor M/2 \rfloor} c_{M,\ell} \sum_{\substack{A \subseteq \{1,\dots,M\} \\ n \notin A,\, |A|=2\ell}} \mathrm{dp}\{h_k\}_{k\in A} \prod_{j \notin A \cup \{n\}} \left(h_j \cdot \frac{x}{|x|} \right) \right] \cdot h_n.
\end{aligned} \tag{28}$$

Interchanging summation and distributing the gradient in (28) then gives

$$\mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1}$$
$$= \frac{1}{M+1}\sum_{\ell=0}^{\lfloor M/2\rfloor} c_{M,\ell}\sum_{n=1}^{M+1}\sum_{\substack{A\subseteq\{1,\dots,M\}\\ n\notin A,\,|A|=2\ell}} \mathrm{dp}\{h_k\}_{k\in A}\nabla\left[\frac{1}{|x|^{M-1}}\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)\right]\cdot h_n. \quad (29)$$

We now directly evaluate the gradient in (29), using the Product Rule:

$$\nabla\left[\frac{1}{|x|^{M-1}}\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)\right]\cdot h_n \quad (30)$$
$$= (\nabla|x|^{1-M})\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)\cdot h_n + \frac{1}{|x|^{M-1}}\nabla\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)\cdot h_n$$
$$= (1-M)|x|^{-M}\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)(\nabla|x|\cdot h_n) \quad (31)$$
$$+ \frac{1}{|x|^{M-1}}\sum_{j\notin A\cup\{n\}}\prod_{i\notin A\cup\{n\}\cup\{j\}}\left(h_i\cdot\frac{x}{|x|}\right)\left[\nabla\left(h_j\cdot\frac{x}{|x|}\right)\right]\cdot h_n. \quad (32)$$

The gradient in (31) is actually a special case of our overall computation — it represents a first order derivative of the norm function, as derived above in (21):

$$\nabla|x|\cdot h_n = \mathrm{D}|x|h_n = h_n\cdot\frac{x}{|x|}. \quad (33)$$

Likewise, the gradient in (32) is related to the second derivative of the norm (23):

$$\left[\nabla\left(h_j\cdot\frac{x}{|x|}\right)\right]\cdot h_n = [\nabla(\mathrm{D}|x|h_j)\cdot h_n$$
$$= \mathrm{D}^2|x|(h_j,h_n)$$
$$= \frac{1}{|x|}\left[h_j\cdot h_n - \left(h_j\cdot\frac{x}{|x|}\right)\left(h_n\cdot\frac{x}{|x|}\right)\right]. \quad (34)$$

Substituting (33) for (31) and (34) for (32), (30) becomes

$$\nabla\left[\frac{1}{|x|^{M-1}}\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)\right]\cdot h_n = \frac{1-M}{|x|^M}\prod_{j\notin A}\left(h_j\cdot\frac{x}{|x|}\right) \quad (35)$$

$$+\frac{1}{|x|^M}\sum_{j\notin A\cup\{n\}}\prod_{i\notin A\cup\{n\}\cup\{j\}}\left(h_i\cdot\frac{x}{|x|}\right)(h_j\cdot h_n)$$
$$-\frac{1}{|x|^M}\sum_{j\notin A\cup\{n\}}\prod_{i\notin A}\left(h_i\cdot\frac{x}{|x|}\right). \tag{36}$$

We now note that (36) is a sum of multiple copies of the same vector. In particular, the number of indices j which are not in $A\cup\{n\}$ is the total number of indices, namely $M+1$, minus the number of indices in $A\cup n$, namely $2\ell-1$. Thus, (36) becomes

$$-\frac{1}{|x|^M}\sum_{j\notin A\cup\{n\}}\prod_{i\notin A}\left(h_i\cdot\frac{x}{|x|}\right)=-\frac{M-2\ell+2}{|x|^M}\prod_{i\notin A}\left(h_i\cdot\frac{x}{|x|}\right). \tag{37}$$

Collecting common terms, we obtain a simplified version of (35):

$$\nabla\left[\frac{1}{|x|^{M-1}}\prod_{j\notin A\cup\{n\}}\left(h_j\cdot\frac{x}{|x|}\right)\right]\cdot h_n$$
$$=\frac{1}{|x|^M}\left[\sum_{j\notin A\cup\{n\}}\prod_{i\notin A\cup\{n\}\cup\{j\}}\left(h_i\cdot\frac{x}{|x|}\right)(h_j\cdot h_n)-(2M-2\ell-1)\prod_{j\notin A}\left(h_j\cdot\frac{x}{|x|}\right)\right]. \tag{38}$$

Using (38), we may return to our expression for the $(M+1)$th derivative (29):

$$(M+1)|x|^M\mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1}$$
$$=\sum_{\ell=0}^{\lfloor M/2\rfloor}c_{M,\ell}\sum_{n=1}^{M+1}\sum_{\substack{A\subseteq\{1,\dots,M\}\\ n\notin A,\,|A|=2\ell}}\mathrm{dp}\{h_k\}_{k\in A}\sum_{j\notin A\cup\{n\}}\prod_{i\notin A\cup\{n\}\cup\{j\}}\left(h_i\cdot\frac{x}{|x|}\right)(h_j\cdot h_n) \tag{39}$$
$$-\sum_{\ell=0}^{\lfloor M/2\rfloor}c_{M,\ell}\sum_{n=1}^{M+1}\sum_{\substack{A\subseteq\{1,\dots,M\}\\ n\notin A,\,|A|=2\ell}}\mathrm{dp}\{h_k\}_{k\in A}(2M-2\ell-1)\prod_{j\notin A}\left(h_j\cdot\frac{x}{|x|}\right). \tag{40}$$

To continue, we simplify (39) by noting

$$\sum_{n=1}^{M+1} \sum_{\substack{A\subseteq\{1,\ldots,M+1\}-\{n\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k\in A} \sum_{j\notin A\cup\{n\}} (h_j \cdot h_n) \prod_{i\notin A\cup\{n\}\cup\{j\}} \left(h_i \cdot \frac{x}{|x|}\right)$$

$$= 2(\ell+1) \sum_{\substack{B\subseteq\{1,\ldots,M+1\} \\ |B|=2(\ell+1)}} \mathrm{dp}\{h_k\}_{k\in B} \prod_{j\notin B} \left(h_j \cdot \frac{x}{|x|}\right), \quad (41)$$

since, for a given $B \subseteq \{1,\ldots,M+1\}$, $|B| = 2\ell+1$, we may pull out any of the $\ell+1$ products $h_j \cdot h_n$, and what remains is a term from the distinct product of $A \subseteq \{1,\ldots,M+1\} - \{n\}$, $|A| = 2\ell$. The extra "2" factor comes from the fact that we count both $h_j \cdot h_n$ and $h_n \cdot h_j$, thus over-counting the terms of a distinct product.

Similarly, we may simplify (40) by noting

$$\sum_{n=1}^{M+1} \sum_{\substack{A\subseteq\{1,\ldots,M+1\}-\{n\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k\in A} \prod_{j\notin A} \left(h_j \cdot \frac{x}{|x|}\right)$$

$$= (M-2\ell+1) \sum_{\substack{B\subseteq\{1,\ldots,M+1\} \\ |B|=2\ell}} \mathrm{dp}\{h_k\}_{k\in B} \prod_{j\notin B} \left(h_j \cdot \frac{x}{|x|}\right), \quad (42)$$

since, for a given $B \subseteq \{1,\ldots,M+1\}$, $|B| = 2\ell$, one may write B as $A \subseteq \{1,\ldots,M+1\} - \{n\}$, $|A| = 2\ell$. Specifically, each such B can be written as such an A exactly $M-2\ell+1$ times — one for each $n \in \{1,\ldots,M+1\} - A$.

Substituting (41) and (42) into (39) and (40), respectively, we have

$$\begin{aligned}
&\mathrm{D}^{M+1}|x|\{h_m\}_{m=1}^{M+1} \\
&= \frac{2(\ell+1)}{(M+1)|x|^M} \sum_{\ell=0}^{\lfloor M/2\rfloor} c_{M,\ell} \sum_{\substack{B\subseteq\{1,\ldots,M+1\} \\ |B|=2(\ell+1)}} \mathrm{dp}\{h_k\}_{k\in B} \prod_{j\notin B} \left(h_j \cdot \frac{x}{|x|}\right) \\
&\quad - \frac{(2M-2\ell-1)(M-2\ell+1)}{(M+1)|x|^M} \sum_{\ell=0}^{\lfloor M/2\rfloor} c_{M,\ell} \sum_{\substack{B\subseteq\{1,\ldots,M+1\} \\ |B|=2\ell}} \mathrm{dp}\{h_k\}_{k\in B} \prod_{j\notin B} \left(h_j \cdot \frac{x}{|x|}\right).
\end{aligned}$$

Considering (26), this equivalently means that the coefficients of the $(M+1)$th derivative can be recursively obtained in terms of the coefficients of the Mth derivative:

$$c_{M+1,\ell} = \frac{1}{M+1}[2\ell c_{M,\ell-1} - (2M-2\ell-1)(M-2\ell+1)c_{M,\ell}] \tag{43}$$

for all ℓ. Recalling the definition of the Mth coefficients (25), we now verify that the $(M+1)$th coefficient is of the same form:

$$\begin{aligned}
c_{M+1,\ell} &= \frac{1}{M+1}[2\ell c_{M,\ell-1} - (2M-2\ell-1)(M-2\ell+1)c_{M,\ell}] \\
&= \frac{1}{M+1}[2\ell(-1)^{M+\ell}(2M-2\ell-1)!! \\
&\qquad\qquad - (2M-2\ell-1)(M-2\ell+1)(-1)^{M+\ell+1}(2M-2\ell-3)!!] \\
&= \frac{(-1)^{M+\ell}}{M+1}[2\ell(2M-2\ell-1)!! + (M-2\ell+1)(2M-2\ell-1)!!] \\
&= \frac{(-1)^{M+\ell}}{M+1}(M+1)(2M-2\ell-1)!! \\
&= (-1)^{(M+1)+\ell+1}(2(M+1)-2\ell-3)!!,
\end{aligned}$$

proving the result. ■

Corollary 3.2.2. *For any $M \geq 1$,*

$$\begin{aligned}
&\mathrm{D}^M|x|h^M \\
&= \frac{|h|^M}{|x|^{M-1}} \sum_{\ell=0}^{\lfloor M/2 \rfloor} (-1)^{M+\ell+1} \frac{M!}{2^M(2M-2\ell-1)} \binom{2M-2\ell}{M-\ell,\ M-2\ell,\ \ell} \left(\frac{h\cdot x}{|h||x|}\right)^{M-2\ell}.
\end{aligned}$$

Proof: The previous result gives

$$\mathrm{D}^M|x|\{h_m\}_{m=1}^M = \frac{1}{|x|^{M-1}} \sum_{\ell=0}^{\lfloor M/2 \rfloor} (-1)^{M+\ell+1}(2M-2\ell-3)!! \sum_{\substack{A\subseteq\{1,\ldots,M\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k\in A} \prod_{j\notin A} \left(h_j \cdot \frac{x}{|x|}\right).$$

Taking $h_m = h$ for all $m = 1, \ldots, M$ gives

$$\prod_{j \notin A} \left(h_j \cdot \frac{x}{|x|} \right) = \left(h \cdot \frac{x}{|x|} \right)^{M-2\ell}$$

since the number of $j \in \{1, \ldots, M\}$ such that $j \notin A$ is $M - 2\ell$. Next, note that for any $\ell = 1, \ldots, \lfloor M/2 \rfloor$ and any $A \subseteq \{1, \ldots, M\}$ such that $|A| = 2\ell$, the total number of distinct dot products of $\{h_k\}_{k \in A}$ is

$$\binom{M}{2\ell} \binom{2\ell}{2, \ldots, 2} \frac{1}{\ell!},$$

as there are $\binom{M}{2\ell}$ ways to pick the terms to be formed into dot products, $\binom{2\ell}{2,\ldots,2}$ ways to break them into pairs, remembering to divide by $\ell!$, so as to count each product of ℓ distinct dot products exactly once. Moreover, since $h_k = h$ for all $k \in A$, every term of $\mathrm{dp}\{h_k\}_{k \in A}$ is a product of ℓ copies of $h \cdot h$, namely $|h|^{2\ell}$. Thus,

$$\begin{aligned} &\mathrm{D}^M |x| \{h_m\}_{m=1}^M \\ &= \frac{1}{|x|^{M-1}} \sum_{\ell=0}^{\lfloor M/2 \rfloor} (-1)^{M+\ell+1} (2M - 2\ell - 3)!! \binom{M}{2\ell} \binom{2\ell}{2, \ldots, 2} \frac{1}{\ell!} |h|^{2\ell} \left(h \cdot \frac{x}{|x|} \right)^{M-2\ell}. \end{aligned}$$

Notice that, assuming $M > 2$,

$$\begin{aligned} &(2M - 2\ell - 3)!! \binom{M}{2\ell} \binom{2\ell}{2, \ldots, 2} \frac{1}{\ell!} \\ &= \frac{(2M - 2\ell - 3)!}{2^{M-\ell-2}(M - \ell - 2)!} \frac{M!}{(M - 2\ell)!(2\ell)!} \frac{(2\ell)!}{2^\ell \ell!} \\ &= \frac{M!}{2^{M-2}} \frac{(2M - 2\ell - 2)(2M - 2\ell - 1)(2M - 2\ell)}{(2M - 2\ell - 2)(2M - 2\ell - 1)(2M - 2\ell)} \\ &= \frac{M!}{2^M (2M - 2\ell - 1)} \frac{(2M - 2\ell)!}{(M - \ell)!(M - 2\ell)!\ell!} \\ &= \frac{M!}{2^M (2M - 2\ell - 1)} \binom{2M - 2\ell}{M - \ell,\ M - 2\ell,\ \ell}, \end{aligned}$$

and so we have our desired result for $M \geq 3$. Evaluation at $M = 1, 2$ also holds. ■

The results above allow us to take Mth derivatives of the Euclidean norm functional. This will become useful as we establish their connection to time derivatives of the Doppler information in the following chapter. They may also be used in a Taylor approximation of the Euclidean norm functional. The following figures illustrate what appears to be Taylor series convergence in two dimensions. That is, we consider

$$\sum_{m=0}^{M} \frac{\mathrm{D}^m|x|h^m}{m!}, \tag{44}$$

where $\mathrm{D}^m|x|h^m$ is given by Corollary 3.2.2. In particular, the graphs on the left show this Taylor approximation (44) for different numbers of terms M, whereas the graphs on the right show the error when the Euclidean norm functional is subtracted from this approximation. In these graphs, the darker shades represent lower values, while the lighter shades represent higher values. The center of each of the graphs is an arbitrary vector $x \neq 0$, and the circle drawn in the righthand graphs is given by the set of points $\{x+h : h \in \mathbb{R}^2,\ |h| = |x|\}$. Furthermore, the origin is the leftmost point of each of these circles. In the righthand graphs, gray represents points close to zero; that is, points in which the truncated Taylor series is close to the Euclidean norm that it is approximating. As one expects, it appears that these Taylor polynomials do converge to the norm function on a neighborhood of x as M grows large. Let us now consider derivatives of the bearing function, as they will arise in Chapter VI.

3.3 Derivatives of the Bearing Function

For any positive integer N, the bearing function $b : \mathbb{R}^N \to \mathbb{R}^N$ is defined by

$$b(x) = \frac{x}{|x|}.$$

The following theorem follows naturally from Theorem 3.2.1.

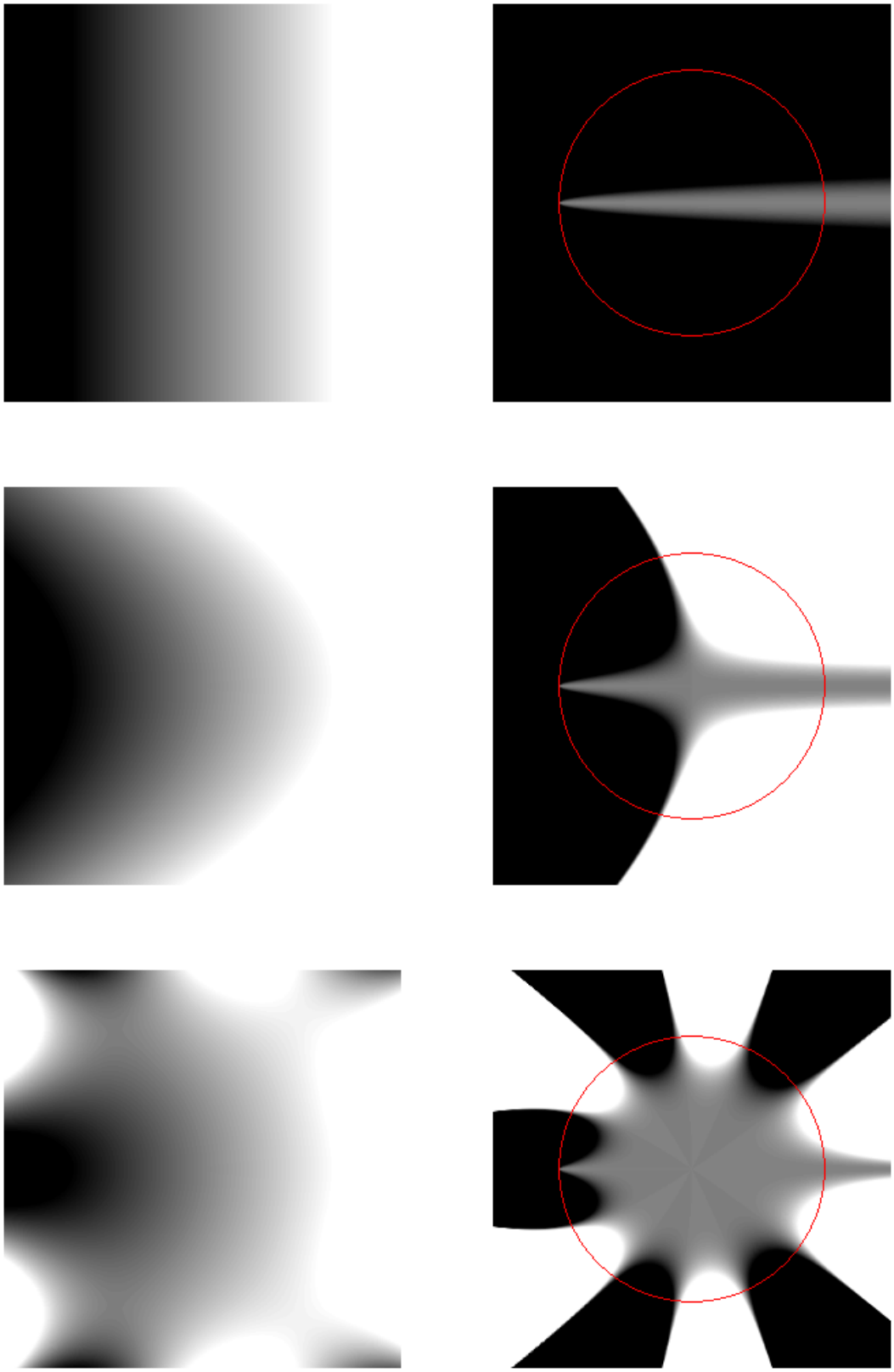

Figure 4 Taylor Series Approximation (1, 2, 6 terms)

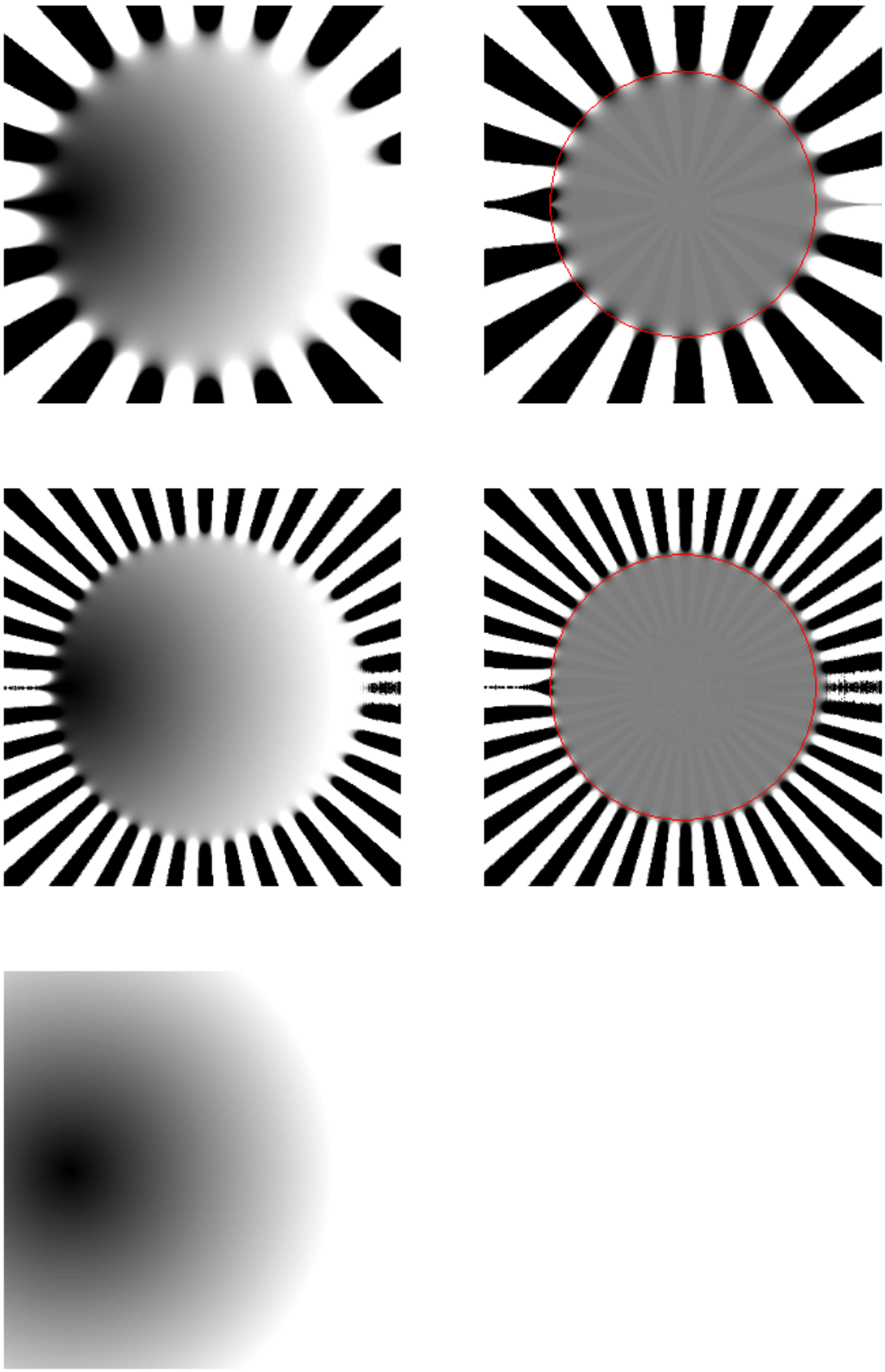

Figure 5 Taylor Series Approximation (20, 40 terms), Euclidean Norm Functional

Theorem 3.3.1. *For any $M \geq 1$,*

$$\begin{aligned}\mathrm{D}^M \frac{x}{|x|}\{h_m\}_{m=1}^M = \frac{1}{x^M} \sum_{\ell=0}^{\lceil M/2 \rceil} (-1)^{M+\ell}(2M-2\ell-1)!! \\ \times \Bigg\{ \sum_{k=1}^{M} \Bigg[\sum_{\substack{A \subseteq \{1,\dots,M\}-\{k\} \\ |A|=2(\ell-1)}} \mathrm{dp}\{h_{k'}\}_{k' \in A} \prod_{j \notin A \cup \{k\}} \left(h_j \cdot \frac{x}{|x|} \right) \Bigg] h_k \\ + \sum_{\substack{A \subseteq \{1,\dots,M\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k \in A} \Bigg[\prod_{j \notin A} \left(h_j \cdot \frac{x}{|x|} \right) \Bigg] \frac{x}{|x|} \Bigg\}.\end{aligned}$$

Proof: The result is immediate when $M = 1$. For $M \geq 2$, fix any $h_0 \in \mathbb{R}^N$. We then consider derivatives of the real-valued function $\dfrac{x}{|x|} \cdot h_0$, recalling $\dfrac{x}{|x|} \cdot h_0 = \mathrm{D}|x|h_0$:

$$\mathrm{D}^M \left(\frac{x}{|x|} \cdot h_0 \right) \{h_m\}_{m=1}^M = \mathrm{D}^M \left(\mathrm{D}|x|h_0 \right) \{h_m\}_{m=1}^M = \mathrm{D}^{M+1}|x|\{h_m\}_{m=0}^M.$$

By Theorem 3.2.1, we therefore have

$$\begin{aligned}&\mathrm{D}^M \left(\frac{x}{|x|} \cdot h_0 \right) \{h_m\}_{m=1}^M \\ &= \frac{1}{|x|^M} \sum_{\ell=0}^{\lceil M/2 \rceil} (-1)^{M+\ell}(2M-2\ell-1)!! \sum_{\substack{A \subseteq \{0,\dots,M\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k \in A} \prod_{j \notin A} \left(h_j \cdot \frac{x}{|x|} \right) \\ &= \frac{1}{|x|^M} \sum_{\ell=0}^{\lceil M/2 \rceil} (-1)^{M+\ell}(2M-2\ell-1)!! \\ &\quad \times \Bigg\{ \sum_{k=1}^{M} \sum_{\substack{B \subseteq \{1,\dots,M\}-\{k\} \\ |B|=2(\ell-1)}} \mathrm{dp}\{h_{k'}\}_{k' \in B} (h_k \cdot h_0) \prod_{j \notin B \cup \{0\} \cup \{k\}} \left(h_j \cdot \frac{x}{|x|} \right) \\ &\qquad + \sum_{\substack{A \subseteq \{1,\dots,M\} \\ |A|=2\ell}} \mathrm{dp}\{h_k\}_{k \in A} \Bigg[\prod_{j \notin A-\{0\}} \left(h_j \cdot \frac{x}{|x|} \right) \Bigg] \left(h_0 \cdot \frac{x}{|x|} \right) \Bigg\}.\end{aligned}$$

Since our choice for h_0 is arbitrary, our theorem naturally follows in this case. ■

As an immediate corollary, note the first derivative to the bearing function is

$$\mathrm{D}\frac{x}{|x|}(h) = \frac{1}{|x|}\left[h - \frac{h\cdot x}{|x|^2}x\right].$$

Finally, after proceeding as we did in the proof of Corollary 3.2.2, one arrives at the following corollary.

Corollary 3.3.2. *For any $M \geq 2$,*

$$\mathrm{D}^M \frac{x}{|x|}h^M = \frac{|h|^M}{|x|^M}\sum_{\ell=0}^{\lceil M/2\rceil}(-1)^{M+\ell}\frac{M!}{2^M}\binom{2M-2\ell}{M-\ell,\ M-2\ell,\ \ell} \times \left(\frac{h\cdot x}{|h||x|}\right)^{M-2\ell}\left[\frac{2\ell}{M-2\ell+1}\left(\frac{h\cdot x}{|h||x|}\right)\frac{h}{|h|} + \frac{x}{|x|}\right].$$

Now that we have calculated these derivatives, we can apply them to the remainder of this thesis. The first of these applications is to establish the observability of the Doppler-only bistatic system.

IV. Observability of the Doppler-Only Bistatic System

Bistatic radar uses two antennas at different locations: one for transmission and the other for reception. In this chapter, we will mathematically derive the limitations of Doppler-only bistatic radar. In particular, we show that even for a target moving at a constant velocity, the Doppler information provided by a single transmitter and receiver is wholly inadequate in determining the target's position at any time. This is not to say that Doppler derivative information is useless, as we do discover an original result that does narrow down the target's location to some extent.

As we are restricted to a single transmitter and a single receiver, our hope in being able to determine the target state lies with making multiple measurements over time. However, rather than assume we have measured $\dot{\rho}(t_k)$ for M distinct times, we instead make the mathematically simpler assumption that we have knowledge of $M-1$ derivatives of $\dot{\rho}$ at a single time t. For example, instead of having $\dot{\rho}(t_0)$ and $\dot{\rho}(t_1)$, we assume we have $\dot{\rho}(t)$ and $\ddot{\rho}(t)$ for some t. Therefore, in practical applications where only the Doppler is given explicitly, the techniques of this chapter would first require the estimation of the derivatives of the Doppler shift. This would most likely be accomplished by measuring $\dot{\rho}(t)$ at several distinct times, and then using finite differences of these measurements to approximate $\rho^{(M)}(t)$.

In this chapter, we first establish the relationship between derivatives of the Euclidean norm functional and time derivatives of the bistatic distance. Applying this understanding to the Doppler information will then establish the extent to which one can estimate the target state.

4.1 Derivative Relationship

Suppose $x : \mathbb{R} \to \mathbb{R}^N$ is a function of t. By the Chain Rule, as given in the appendix, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}|x(t)| = \mathrm{D}|x|\dot{x}(t)\Big|_{x=x(t)} = \frac{x(t)}{|x(t)|} \cdot \dot{x}(t),$$

and further,

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}t^2}|x(t)| &= \frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{D}|x|\dot{x}(t)]\Big|_{x=x(t)} \\
&= \mathrm{D}|x|\ddot{x}(t)\Big|_{x=x(t)} + \mathrm{D}^2|x|(\dot{x}(t), \dot{x}(t))\Big|_{x=x(t)} \\
&= \frac{x(t)}{|x(t)|} \cdot \ddot{x}(t) + \frac{1}{|x(t)|}\left[|\dot{x}(t)|^2 - \frac{(x(t) \cdot \dot{x}(t))^2}{|x(t)|^2}\right].
\end{aligned}$$

Assuming $\ddot{x}(t) = 0$ for all t, we inductively see that

$$\frac{\mathrm{d}^M}{\mathrm{d}t^M}|x(t)| = \mathrm{D}^M|x|\dot{x}(t)^M\Big|_{x=x(t)},$$

as all other terms will contain the multilinear operators D^m, $m < M$, evaluated at second or higher degree derivative of x, all of which are identically zero. In particular, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}|x(t)| &= \frac{x(t)}{|x(t)|} \cdot \dot{x}(t), \\
\frac{\mathrm{d}^2}{\mathrm{d}t^2}|x(t)| &= \frac{1}{|x(t)|}\left[|\dot{x}(t)|^2 - \frac{[x(t) \cdot \dot{x}(t)]^2}{|x(t)|^2}\right].
\end{aligned}$$

Interestingly, one can iteratively arrive at the same conclusion without using Fréchet derivatives. To do this, let $x, y : \mathbb{R} \to \mathbb{R}^N$ be n-times differentiable. Then, suppressing dependence on t, repeated application of the Product Rule yields a Binomial

Theorem version of the Product Rule:

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}(x \cdot y) = \sum_{k=0}^{n} \binom{n}{k} x^{(k)} \cdot y^{(n-k)}.$$

It follows that if $g : \mathbb{R} \to \mathbb{R}$ is defined by $g(t) \equiv |x(t)|$ for all t, then

$$\sum_{k=0}^{n} \binom{n}{k} g^{(k)} g^{(n-k)} = \sum_{k=0}^{n} \binom{n}{k} x^{(k)} \cdot x^{(n-k)}.$$

Suppose our parameterized path x has constant velocity, that is, $\ddot{x} = 0$. Then, considering the case $n = 1$ for our above equation gives

$$g\dot{g} + \dot{g}g = x \cdot \dot{x} + \dot{x} \cdot x,$$

and so

$$\dot{g} = \frac{x \cdot \dot{x}}{g} = \frac{x \cdot \dot{x}}{|x|}.$$

For $n = 2$, we have

$$g\ddot{g} + 2\dot{g}\dot{g} + \ddot{g}g = 2\dot{x} \cdot \dot{x},$$

and so

$$\ddot{g} = \frac{|\dot{x}|^2 - (\dot{g})^2}{g} = \frac{1}{|x|}\left[|\dot{x}|^2 - \frac{(x \cdot \dot{x})^2}{|x|^2}\right],$$

as expected. This same technique may be used to iteratively find $g^{(m)}$ for all m.

4.2 Bistatic Observability

To reiterate, assuming $\ddot{x} = 0$, we have

$$\frac{\mathrm{d}^M}{\mathrm{d}t^M}|x(t)| = \mathrm{D}^M |x| \dot{x}(t)^M \Big|_{x=x(t)}.$$

If we then apply Corollary 3.2.2 to the case where the $x(t)$ from the corollary is taken to be $x(t) - T$, and then $x(t) - R$, we obtain the following expression for the Mth derivative of the bistatic distance:

Theorem 4.2.1. *We have*

$$\frac{(-1)^{M+1}2^M}{M!}\rho^{(M)}(t) = s(t)\sum_{\ell=0}^{\lfloor M/2\rfloor}\frac{(-1)^\ell}{2M-2\ell-1}\binom{2M-2\ell}{M-\ell,\ M-2\ell,\ \ell} \times [p_T(t)^{M-1}\cos^{M-2\ell}\theta_T(t) + p_R(t)^{M-1}\cos^{M-2\ell}\theta_R(t)],$$

where $s(t) = |\dot{x}(t)|$, $p_y(t) = \dfrac{|\dot{x}(t)|}{|x(t)-y|}$ *and* $\cos\theta_y(t) = \dfrac{\dot{x}(t)\cdot(x(t)-y)}{|\dot{x}(t)||x(t)-y|}$, *for* $y = T, R$.

Recall that we are assuming the derivatives $\rho^{(M)}(t)$ are given data for all M, and so we only have unknowns $s(t)$, $p_T(t)$, $p_R(t)$, $\cos\theta_T(t)$, and $\cos\theta_R(t)$. Suppressing dependence on t, the first five equations are

$$\begin{aligned}
\dot{\rho} &= s(\cos\theta_T + \cos\theta_R)\\
-2\ddot{\rho} &= s[p_T\cos^2\theta_T + p_R\cos^2\theta_R - 2(p_T + p_R)]\\
\frac{1}{3}\rho^{(3)} &= s[(p_T^2\cos^3\theta_T + p_R^2\cos^3\theta_R) - (p_T^2\cos\theta_T + p_R^2\cos\theta_R)]\\
-\frac{1}{3}\rho^{(4)} &= s[5(p_T^3\cos^4\theta_T + p_R^3\cos^4\theta_R) - 6(p_T^3\cos^2\theta_T + p_R^3\cos^2\theta_R) + (p_T^3 + p_R^3)]\\
\frac{1}{15}\rho^{(5)} &= s[7(p_T^4\cos^5\theta_T + p_R^4\cos^5\theta_R) - 10(p_T^4\cos^3\theta_T + p_R^4\cos^3\theta_R)\\
&\qquad + 3(p_T^4\cos\theta_T + p_R^4\cos\theta_R)].
\end{aligned}$$

Clearly, it is difficult to solve this system of equations. Gaussian elimination, algebraic geometry and numerical techniques have all been applied in search of the solution, with no successful conclusion.

But for now, suppose we acquired enough derivatives to resolve these unknowns. Dividing s by p_T and p_R would yield $|x - T|$ and $|x - R|$, respectively. Having the distances of x from both the transmitter and receiver, the location $x(t)$ would be determined up to a circle. Meanwhile, as $\cos\theta_T$ and $\cos\theta_R$ give the an-

gles between $\dot{x}(t)$ and $x(t) - T$ and $x(t) - R$, respectively, knowledge of these five constants would also completely determine the components of $\dot{x}(t)$ which lie in the plane spanned by $x(t) - T$ and $x(t) - R$. Since we would also have $|\dot{x}(t)|$, this would determine $\dot{x}(t)$ up to two distinct points.

However, we note that the symmetry in our polynomials is such that we can never distinguish p_T from p_R, and so we could only hope to resolve location up to two symmetric circles. We therefore conclude, even with infinitely differentiable Doppler information coupled with a known constant-velocity target, that the Doppler-only bistatic system can only ever at most determine the target's position up to a circle, and from that, determine the target's velocity up to two distinct vectors. Thus, even with the harsh assumption that $\ddot{x}(t) = 0$ for all t, the Doppler-only bistatic system is insufficient for target observability.

Therefore, if one wishes to have a Doppler-only radar system in which a target's position can be completely determined, one is forced to employ multiple transmitters and/or receivers, that is, a nontrivial bistatic system.

V. Doppler-Only Multistatic Target State Estimation

Let us now consider the case where we have measurements from multiple transmitter-receiver pairs at a fixed time. Given enough pairs, we determine the target's position and velocity — a previously unsolved problem. To be precise, we characterize the target state as the global minimizer of a certain objective function. We then determine the gradient and Hessian of this objective function, and, further, provide a simple minimization algorithm to find this global minimizer, thereby yielding a target state estimate. These results constitute a significant improvement over previous results; both those given in the previous chapter, as well as those found elsewhere in the literature. Here, not only do we completely determine the position and velocity of a target, but we also do so without assuming the velocity of the target is constant.

In this chapter, let N be the spatial dimension, typically either 2 or 3, and P be the number of transmitter-receiver pairs. We will use a dot $(\cdot)$ and $|\cdot|$ to indicate the dot product and norm for $\mathbb{R}^N$, respectively; and $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$, similarly, for $\mathbb{R}^P$. Let $\{T_p\}_{p=1}^P \subseteq \mathbb{R}^N$ be the locations of the transmitters, and $\{R_p\}_{p=1}^P \subseteq \mathbb{R}^N$ the locations of the receivers. In the theory below, we do not assume the transmitter and receiver locations are necessarily distinct — that is, one may have $T_i = T_j$ for $i \neq j$ and similarly for receivers. For example, a real-world system consisting of a single receiver and four transmitters would have T_1, T_2, T_3 and T_4 all being distinct, while $R_1 = R_2 = R_3 = R_4$.

5.1 The Objective Function

We define the *bearing matrix* to be a function of position $B : \mathbb{R}^N \to \mathbb{R}^{P\times N}$ whose pth row is given by

$$[B_{p,\cdot}(x)]^T = \frac{x - T_p}{|x - T_p|} + \frac{x - R_p}{|x - R_p|}. \tag{45}$$

Recall that if Doppler measurements are without error, we have

$$\left[\frac{x(t)-T_p}{|x(t)-T_p|}+\frac{x(t)-R_p}{|x(t)-R_p|}\right]\cdot\dot{x}(t)=\dot{\rho}_p(t). \tag{46}$$

Letting $d:\mathbb{R}\to\mathbb{R}^P$ be defined by

$$d(t)=\begin{bmatrix}\rho_1(t)\\ \rho_2(t)\\ \vdots\\ \rho_P(t)\end{bmatrix},$$

we may express the system of equations given by (46) for $p=1,2,\ldots,P$ as a single matrix-vector equation:

$$B(x(t))\dot{x}(t)=d(t). \tag{47}$$

If P is large enough, then the system of differential equations in (47) is not underdetermined. If only we had an initial condition, namely, if we actually knew $x(t)$ for some t, we could track the target using numerical ordinary differential equation solvers. This approach is a subject of ongoing research, and is not discussed further here. Indeed, we instead show that given enough transmitter-receiver pairs, the information contained in (47) is enough to determine the position of the target, even when the Doppler measurements are only given for a single time. In particular, we fix t and attempt to solve the equation

$$B(x)v=d \tag{48}$$

for the target's position $x\in\mathbb{R}^N$ and velocity $v\in\mathbb{R}^N$. We note that (48) is equivalent to solving

$$\|B(x)v-d\|=0,$$

and pose it as a minimization problem; that is, we wish to find the target state (x^*, v^*) that satisfies

$$\|B(x^*)v^* - d\| = \min_{x,v\in\mathbb{R}^N} \|B(x)v - d\|. \tag{49}$$

Note that a solution to (49) will exist even when a solution to (48) does not, which could happen when $P > N$ and d has been corrupted, either by measurement error or quantization.

Since our expression for d is linear in v, solving (49) is equivalent to minimizing the function $\Upsilon : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\Upsilon(x) \equiv \min_{v\in\mathbb{R}^N} \|B(x)v - d\|^2. \tag{50}$$

For a fixed $x \in \mathbb{R}^N$, the minimizer of the linear least-squares problem $\|B(x)v - d\|^2$ is given by

$$v^*(x) = [B(x)]^\dagger d, \tag{51}$$

where $[B(x)]^\dagger = [(B(x))^T B(x)]^{-1}(B(x))^T$ is the pseudoinverse of $B(x)$. Note that we assume $B(x)$ is of full rank so that this pseudoinverse exists. Furthermore, with the projection onto the orthogonal complement of the range of matrix $B(x)$ given by

$$P_{B(x)^\perp} \equiv I - B(x)[B(x)]^\dagger,$$

it follows, by direct substitution of (51) into (50), that

$$\Upsilon(x) = \|P_{B(x)^\perp} d\|^2. \tag{52}$$

Figure 6 illustrates a two-dimensional example of our objective function, where the lighter regions indicate values of the objective function $\Upsilon(x)$ which are close to zero, and the darker regions correspond to larger values of $\Upsilon(x)$. In this example, a "○"

denotes a transmitter location, "+" denotes the receiver location and "◇" denotes the target location.

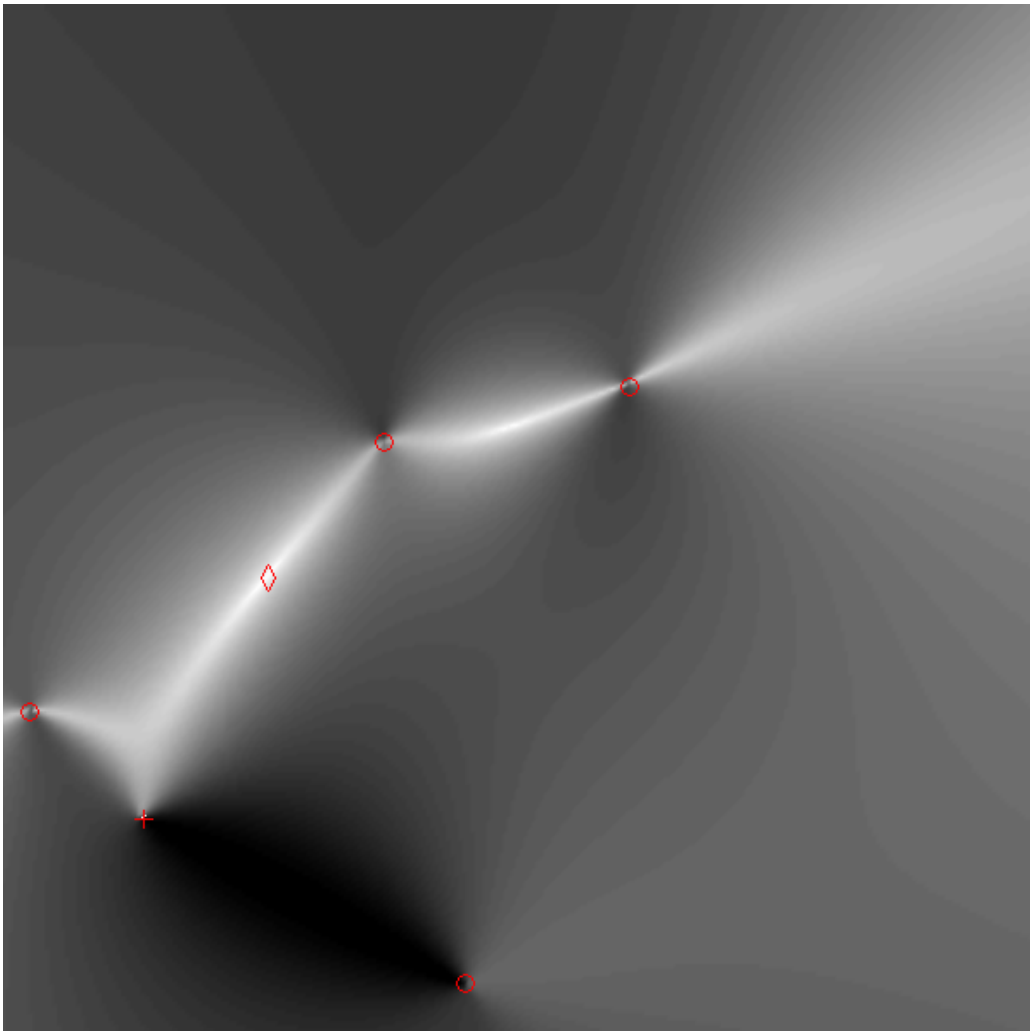

Figure 6 Example of Objective Function

In a real-life application, one would know the positions of the transmitters and receiver, as well as the Doppler information, all of which will yield the objective function. However, the true location of the target would not be known. In other words, one would not see the diamond in Figure 6, and instead would need to determine its location by finding the lightest point of the graph. The purpose of this chapter is to show how to find the diamond; that is, how to minimize this objective function, as its global minimum corresponds to the target state in question.

To minimize (52), first note that

$$0 \leq \Upsilon(x) = \|P_{B(x)^\perp} d\|^2 \leq \|P_{B(x)^\perp}\|^2 \|d\|^2 \leq \|d\|^2,$$

for all $x \in \mathbb{R}^N$. Furthermore, we know that the value of the objective function Υ evaluated at the target's true location is, in fact, zero. It follows that the target's location gives a global minimum to the objective function. Unfortunately, as (52) is defined in terms of a nonlinear matrix-valued function $B(x)$, it appears unlikely that any closed form expression for the minimizers can be found. We therefore focus on a numerical approach to find the minimizers.

5.2 The Gradient of the Objective Function

To reiterate, the location of the target is the minimizer of (52). In this section, we explicitly compute the gradient of (52), the natural first step in solving any such optimization problem. As the existing methods for estimating the target state in a Doppler-only multistatic system are completely different than our own, this computation is original.

Throughout the proof of the next result and the rest of this thesis, we denote partial derivatives parenthetically with the following superscript convention. For a function g of several real variables, let

$$g^{(n_1,n_2,\ldots,n_k)}(x) \equiv \frac{\partial^k g(x)}{\partial x_{n_k} \cdots \partial x_{n_2} \partial x_{n_1}}.$$

We note this notation is nonstandard, as $g^{(n)}$ usually denotes an nth order derivative with respect to the same variable.

Theorem 5.2.1. *The gradient of the objective function* (52) *is given by*

$$\nabla \Upsilon(x) = -2 \left\{ \sum_{p=1}^{P} [P_{B(x)^\perp} d]_p \left[\frac{P_{(x-T_p)^\perp}}{|x - T_p|} + \frac{P_{(x-R_p)^\perp}}{|x - R_p|} \right] \right\} [B(x)]^\dagger d.$$

Proof: We first notice the following alternative expression for the objective function (52), based upon the fact that $P_{B(x)^\perp} = P^T_{B(x)^\perp} = P^2_{B(x)^\perp}$. Suppressing dependence on x:

$$\Upsilon = \|P_{B^\perp} d\|^2 = \langle P_{B^\perp} d, P_{B^\perp} d\rangle = \langle P^T_{B^\perp} P_{B^\perp} d, d\rangle = \langle P_{B^\perp} d, d\rangle = d^T P_{B^\perp} d.$$

Since d does not depend upon x_n, we have

$$\frac{\partial \Upsilon(x)}{\partial x_n} \equiv \Upsilon^{(n)}(x) = d^T P^{(n)}_{B(x)^\perp} d = \langle P^{(n)}_{B(x)^\perp} d, d\rangle. \tag{53}$$

Thus, finding $\Upsilon^{(n)}(x)$ reduces to finding $P^{(n)}_{B(x)^\perp}$. Note that for a matrix-valued function $A(x)$ such that $A(x)$ has a matrix inverse $[A(x)]^{-1}$ for all x, we have $A(x)[A(x)]^{-1} = I$, and so the Product Rule yields

$$([A(x)]^{-1})^{(n)} = -[A(x)]^{-1} A^{(n)}(x) [A(x)]^{-1}.$$

Suppressing dependence on x, the Product Rule further gives

$$\begin{aligned}
P^{(n)}_{B^\perp} &= [I - BB^\dagger]^{(n)} \\
&= [I - B(B^T B)^{-1} B^T]^{(n)} \\
&= -B^{(n)}(B^T B)^{-1} B^T - B[(B^T B)^{-1}]^{(n)} B^T - B(B^T B)^{-1}(B^T)^{(n)} \\
&= -B^{(n)} B^\dagger + B(B^T B)^{-1}(B^T B)^{(n)}(B^T B)^{-1} B^T - (B^\dagger)^T (B^{(n)})^T \\
&= -B^{(n)} B^\dagger + (B^\dagger)^T [(B^{(n)})^T B + B^T B^{(n)}] B^\dagger - (B^\dagger)^T (B^{(n)})^T \\
&= -B^{(n)} B^\dagger + (B^\dagger)^T (B^{(n)})^T B B^\dagger + (B^\dagger)^T B^T B^{(n)} B^\dagger - (B^\dagger)^T (B^{(n)})^T \\
&= -B^{(n)} B^\dagger + (B^\dagger)^T (B^{(n)})^T P_B + P^T_B B^{(n)} B^\dagger - (B^\dagger)^T (B^{(n)})^T \\
&= P_B B^{(n)} B^\dagger - B^{(n)} B^\dagger + (P_B B^{(n)} B^\dagger)^T - (B^{(n)} B^\dagger)^T \\
&= -P_{B^\perp} B^{(n)} B^\dagger - (P_{B^\perp} B^{(n)} B^\dagger)^T.
\end{aligned} \tag{54}$$

Substituting (54) into the expression for $\Upsilon^{(n)}(x)$ above (53) gives

$$\begin{aligned}\frac{\partial \Upsilon(x)}{\partial x_n} &= \langle P^{(n)}_{B^\perp} d, d\rangle \\ &= -\langle P_{B^\perp} B^{(n)} B^\dagger d, d\rangle - \langle (P_{B^\perp} B^{(n)} B^\dagger)^T d, d\rangle \\ &= -\langle P_{B^\perp} B^{(n)} B^\dagger d, d\rangle - \langle d, P_{B^\perp} B^{(n)} B^\dagger d\rangle \\ &= -2\langle P_{B^\perp} B^{(n)} B^\dagger d, d\rangle \\ &= -2\langle B^{(n)} B^\dagger d, P_{B^\perp} d\rangle. \end{aligned} \tag{55}$$

What remains is to determine partial derivatives of $B(x)$. Letting $h = e_n$, where e_n is the nth column of the $N \times N$ identity matrix, Theorem 3.3.1 gives

$$\frac{\partial}{\partial x_n}\ \frac{x}{|x|} = \frac{1}{|x|}\left[e_n - \frac{e_n \cdot x}{|x|^2} x\right] = \frac{P_{x^\perp} e_n}{|x|}. \tag{56}$$

Since the pth row of $B(x)$ is given by (45), we see from (56) that

$$B^{(n)}_{p,\cdot}(x) = e_n^T \left[\frac{P_{(x-T_p)^\perp}}{|x - T_p|} + \frac{P_{(x-R_p)^\perp}}{|x - R_p|}\right]. \tag{57}$$

Thus, substituting (57) into (55) gives the following explicit inner product evaluation:

$$\begin{aligned}\frac{\partial \Upsilon(x)}{\partial x_n} &= -2\langle B^{(n)}(x)[B(x)]^\dagger d, P_{B(x)^\perp} d\rangle \\ &= -2\sum_{p=1}^{P} [B^{(n)}(x)[B(x)]^\dagger d]_p [P_{B(x)^\perp} d]_p \\ &= -2e_n^T \left\{\sum_{p=1}^{P} [P_{B(x)^\perp} d]_p \left[\frac{P_{(x-T_p)^\perp}}{|x - T_p|} + \frac{P_{(x-R_p)^\perp}}{|x - R_p|}\right]\right\} [B(x)]^\dagger d. \end{aligned}$$

As $\Upsilon^{(n)}(x)$ is the nth component of $\nabla \Upsilon(x)$, the proof is complete. ■

While the gradient is all that is needed in many optimization algorithms, better performance can usually be obtained by making use of curvature information; that is, with second derivatives.

5.3 The Hessian of the Objective Function

As in the previous section, methods and results below are unique to this thesis, not appearing in the preexisting literature.

Theorem 5.3.1. *The Hessian of the objective function* (52) *is given by*

$$[\nabla^2 \Upsilon(x)]_{n,m} = -2d^T [P_{B(x)^\perp} B^{(n)}(x)[B(x)]^\dagger]^{(m)} d,$$

where the middle factor is given by

$$\begin{aligned}
&[P_{B(x)^\perp} B^{(n)}(x)[B(x)]^\dagger]^{(m)} \\
&= -[P_{B(x)^\perp} B^{(m)}(x)[B(x)]^\dagger + (P_{B(x)^\perp} B^{(m)}(x)[B(x)]^\dagger)^T] B^{(n)}(x)[B(x)]^\dagger \\
&\quad + P_{B(x)^\perp} B^{(n,m)}(x)[B(x)]^\dagger \\
&\quad + P_{B(x)^\perp} B^{(n)}[B(x)]^\dagger [(P_{B(x)^\perp} B^{(m)}(x)[B(x)]^\dagger)^T - B^{(m)}(x)[B(x)]^\dagger], \qquad (58)
\end{aligned}$$

while $B^{(n)}(x)$ *is given by its pth row in* (57)*:*

$$B^{(n)}_{p,\cdot}(x) = e_n^T \left[\frac{P_{(x-T_p)^\perp}}{|x - T_p|} + \frac{P_{(x-R_p)^\perp}}{|x - R_p|} \right],$$

and $B^{(n,m)}(x)$ *is similarly given by*

$$\begin{aligned}
B^{(n,m)}_{p,\cdot}(x) = &- \left[\frac{P_{(x-T_p)^\perp}}{|x - T_p|} (e_m e_n^T + e_n e_m^T) + \left(e_n^T \frac{P_{(x-T_p)^\perp}}{|x - T_p|} e_m \right) I \right] \frac{x - T_p}{|x - T_p|} \\
&- \left[\frac{P_{(x-R_p)^\perp}}{|x - R_p|} (e_m e_n^T + e_n e_m^T) + \left(e_n^T \frac{P_{(x-R_p)^\perp}}{|x - R_p|} e_m \right) I \right] \frac{x - R_p}{|x - R_p|}. \qquad (59)
\end{aligned}$$

Proof: The entries of the Hessian are given by partial derivatives of the entries of the gradient. That is, considering (55), we see that entries of the Hessian are of the form

$$\Upsilon^{(n,m)}(x) \equiv \frac{\partial^2 \Upsilon(x)}{\partial x_m \partial x_n} = \frac{\partial}{\partial x_m} \Upsilon^{(n)}(x) = \frac{\partial}{\partial x_m} [-2d^T P_{B(x)^\perp} B^{(n)}(x)[B(x)]^\dagger].$$

Further, since d does not vary with x, we have by the Product Rule that

$$\Upsilon^{(n,m)}(x) = -2d^T[P_{B(x)^\perp}B^{(n)}(x)[B(x)]^\dagger]^{(m)}d.$$

To express the middle term, we have by the Product Rule that,

$$[P_{B^\perp}B^{(n)}B^\dagger]^{(m)} = P_{B^\perp}^{(m)}B^{(n)}B^\dagger + P_{B^\perp}B^{(n,m)}B^\dagger + P_{B^\perp}B^{(n)}(B^\dagger)^{(m)}, \tag{60}$$

suppressing dependence on x. We also notice that with the Product Rule, $BB^\dagger = P_B$ gives

$$B^{(m)}B^\dagger + B(B^\dagger)^{(m)} = P_B^{(m)}.$$

Subtracting $B^{(m)}B^\dagger$ and multiplying the resulting equation on the left by $B^\dagger$ gives

$$(B^\dagger)^{(m)} = B^\dagger[P_B^{(m)} - B^{(m)}B^\dagger]. \tag{61}$$

Substituting (61) into (60) yields

$$[P_{B^\perp}B^{(n)}B^\dagger]^{(m)} = P_{B^\perp}^{(m)}B^{(n)}B^\dagger + P_{B^\perp}B^{(n,m)}B^\dagger + P_{B^\perp}B^{(n)}B^\dagger[P_B^{(m)} - B^{(m)}B^\dagger]. \tag{62}$$

Now, we substitute (54) into (62) to get

$$\begin{aligned}[P_{B^\perp}B^{(n)}B^\dagger]^{(m)} = &-[P_{B^\perp}B^{(m)}B^\dagger + (P_{B^\perp}B^{(m)}B^\dagger)^T]B^{(n)}B^\dagger + P_{B^\perp}B^{(n,m)}B^\dagger \\ &+ P_{B^\perp}B^{(n)}B^\dagger[P_{B^\perp}B^{(m)}B^\dagger + (P_{B^\perp}B^{(m)}B^\dagger)^T - B^{(m)}B^\dagger].\end{aligned} \tag{63}$$

Here, we notice that

$$B^\dagger P_{B^\perp} = (B^TB)^{-1}B^T[I - B(B^TB)^{-1}B^T] = 0,$$

and so (63) reduces to (58).

At this point, the expression for $B^{(n)}$ given in the result is simply a restatement of (57) from the proof of the previous result. We therefore have only to determine $B^{(n,m)}$. Considering Theorem 3.3.1, we have

$$\begin{aligned}
&\frac{\partial^2}{\partial x_m \partial x_n}\frac{x}{|x|} \\
&= \mathrm{D}^2 \frac{x}{|x|}(e_n, e_m) \\
&= \frac{1}{|x|^2}\left\{3\left(e_n \cdot \frac{x}{|x|}\right)\left(e_m \cdot \frac{x}{|x|}\right)\frac{x}{|x|} \right. \\
&\qquad\qquad \left. - \left[\left(e_n \cdot \frac{x}{|x|}\right)e_m + \left(e_m \cdot \frac{x}{|x|}\right)e_n + (e_n \cdot e_m)\frac{x}{|x|}\right]\right\} \\
&= \frac{1}{|x|^3}\left\{\frac{3}{|x|^2}x e_n^T x e_m^T x - \left[e_m e_n^T x + e_n e_m^T x + x e_n^T e_m\right]\right\} \\
&= -\frac{1}{|x|^3}\left\{\left[e_m e_n^T x - \frac{x e_m^T x e_n^T x}{|x|^2}\right] + \left[e_n e_m^T x - \frac{x e_n^T x e_m^T x}{|x|^2}\right] + \left[x e_n^T e_m - \frac{x e_n^T x e_m^T x}{|x|^2}\right]\right\} \\
&= -\frac{1}{|x|^3}\left\{\left[e_m - x\frac{e_m^T x}{|x|^2}\right]e_n^T + \left[e_n - x\frac{e_n^T x}{|x|^2}\right]e_m^T + e_n^T\left[e_m - x\frac{e_m^T x}{|x|^2}\right]I\right\}x \\
&= -\frac{1}{|x|^3}\left[P_{x^\perp}(e_m e_n^T + e_n e_m^T) + (e_n^T P_{x^\perp} e_m)I\right]x. \qquad (64)
\end{aligned}$$

As the rows of the bearing matrix $B(x)$ are given by (45), our expression for (59) follows from (64). ■

Assuming Υ has sufficient smoothness, the Hessian must be symmetric by the symmetry of mixed partial derivatives. However, the expression above does not make this fact obvious. To convince the reader that the expression for the Hessian given in Theorem 5.3.1 is correct, we now verify that it is indeed symmetric; that is, $[\nabla^2\Upsilon(x)]_{n,m} = [\nabla^2\Upsilon(x)]_{m,n}$.

Let $A \in \mathbb{R}^{P\times P}$ be a matrix. Then, we have

$$d^T A d = \langle d, Ad\rangle = \langle A^T d, d\rangle = \langle d, A^T d\rangle = d^T A^T d. \qquad (65)$$

Further, rearranging (58) gives

$$[P_{B^\perp} B^{(n)} B^\dagger]^{(m)} = P_{B^\perp} B^{(n,m)} B^\dagger \tag{66}$$

$$- P_{B^\perp} B^{(m)} B^\dagger B^{(n)} B^\dagger - P_{B^\perp} B^{(n)} B^\dagger B^{(m)} B^\dagger \tag{67}$$

$$- (P_{B^\perp} B^{(m)} B^\dagger)^T B^{(n)} B^\dagger + P_{B^\perp} B^{(n)} B^\dagger (P_{B^\perp} B^{(m)} B^\dagger)^T. \tag{68}$$

We see that the expressions in (66) and (67) are symmetric with respect to m and n. Finally, (68) satisfies

$$\begin{aligned} &- (P_{B^\perp} B^{(m)} B^\dagger)^T (P_{B^\perp} B^{(n)} B^\dagger) + (P_{B^\perp} B^{(n)} B^\dagger)(P_{B^\perp} B^{(m)} B^\dagger)^T \\ &\qquad = [-(P_{B^\perp} B^{(n)} B^\dagger)^T (P_{B^\perp} B^{(m)} B^\dagger) + (P_{B^\perp} B^{(m)} B^\dagger)(P_{B^\perp} B^{(n)} B^\dagger)^T]^T, \end{aligned}$$

that is, taking the transpose of (68) is equivalent to interchanging m and n. By (65), this implies that the Hessian is symmetric, as desired.

5.4 A Minimization Algorithm

With expressions for the gradient and Hessian of the objective function established in Theorems 5.2.1 and 5.3.1, respectively, we now use these results to determine the target state in a Doppler-only multistatic system, provided we have enough transmitter-receiver pairs. Specifically, recall that the target's position is the minimizer of objective function $\Upsilon(x)$. Using standard, well-known techniques of optimization, we now numerically determine this minimizer. Though the expressions for the gradient and Hessian of $\Upsilon(x)$ given in the previous sections are new, our application of them in this section makes judicious use of ideas that appear in well-known optimization textbooks, such as [27].

Recall that the gradient is the direction of greatest instantaneous increase, and the negative gradient that of greatest decrease. *Gradient descent* is a greedy optimization algorithm, taking steps in the direction of greatest instantaneous decrease,

that is, negative of the gradient direction. A more efficient optimization algorithm is *Newton's method*. Here, the idea is to approximate the objective function by its second order multivariable Taylor polynomial

$$\Upsilon(x+h) \cong \Upsilon(x) + [\nabla\Upsilon(x)]^T h + \frac{1}{2} h^T [\nabla^2\Upsilon(x)] h,$$

and move in the direction of the vertex of this paraboloid, given by the vector $-[\nabla^2\Upsilon(x)]^{-1}\nabla\Upsilon(x)$, assuming that the Hessian is, in fact, positive-definite.

However, the Hessian of objective function $\Upsilon(x)$ is not always positive-definite. Therefore, in defining the iteration of our algorithm

$$x_{k+1} = x_k + \alpha_k p_k, \tag{69}$$

where p_k is the search direction and α_k is the step length, we need to be careful in establishing p_k, as we cannot always employ Newton's method. We therefore arrive at the following piecewise-defined search direction:

$$p_k = \begin{cases} -\nabla\Upsilon(x_k) & \text{when } \nabla^2\Upsilon(x_k) \text{ is not positive-definite} \\ -[\nabla^2\Upsilon(x_k)]^{-1}\nabla\Upsilon(x_k) & \text{when } \nabla^2\Upsilon(x_k) \text{ is positive-definite} \end{cases}.$$

We need to further establish how the step length α_k is determined. To ensure some sense of convergence [27, pp. 41–42], we use a backtracking line search, given by the following algorithm:

$\alpha = 1;$
`while` $\Upsilon(x_k + \alpha p_k) > \Upsilon(x_k) + \epsilon\alpha\nabla\Upsilon(x_k)^T p_k$
    $\alpha = \alpha/2;$
`end`
$\alpha_k = \alpha;$

where $0 < \epsilon < 1$.

In practice, this algorithm is quite effective. Figures 7 and 8 show two simulations that employ this algorithm to find the target's location, given only Doppler information. In these graphs, lines that connect iteration points are of different shades, depending on how the search direction was chosen according to our piecewise-defined rule. Specifically, lighter lines indicate gradient-based search directions, whereas darker lines indicate use of the Newton direction.

By contrast, Figure 9 depicts a simulation in which our algorithm converges to a local minimum, that is, a location other than the target's. This illustrates that our algorithm still requires some fine-tuning. More specifically, we recommend a reasonable grid-search to determine starting points for multiple iterations. Since the functions are relatively inexpensive to evaluate, performing multiple runs from different starting points can help avoid local minima.

Recall Figure 6, which first illustrated an example of our objective function $\Upsilon(x)$. Figure 10 depicts, in white, the set of all points which, when used as a starting point for our algorithm, induce an iteration that converges to the true location of the target from Figure 6. It appears that if a starting point is chosen close enough to the true location of the target, indicated by a diamond, then it would be reasonable to expect that our algorithm will converge to this location. Next, Figure 11 expresses the number of steps necessary to converge to any local minimum of the graph in Figure 6. In particular, darker points represent starting points whose iteration took longer to converge. Notice that points close to the true target location converge quickly, as expected. Interestingly, this graph appears to have fractal behavior.

In summary, our suggestion for acquiring the optimal target state estimate starts with applying a grid-search. This grid-search will yield the first point in our iteration. To find the following point, one first determines a search direction according to the piecewise-defined rule, which depends on whether the Hessian is positive-definite. Next, the step length is determined via backtracking, and so we arrive at the next iteration point. Continuing the iteration in this manner seems to

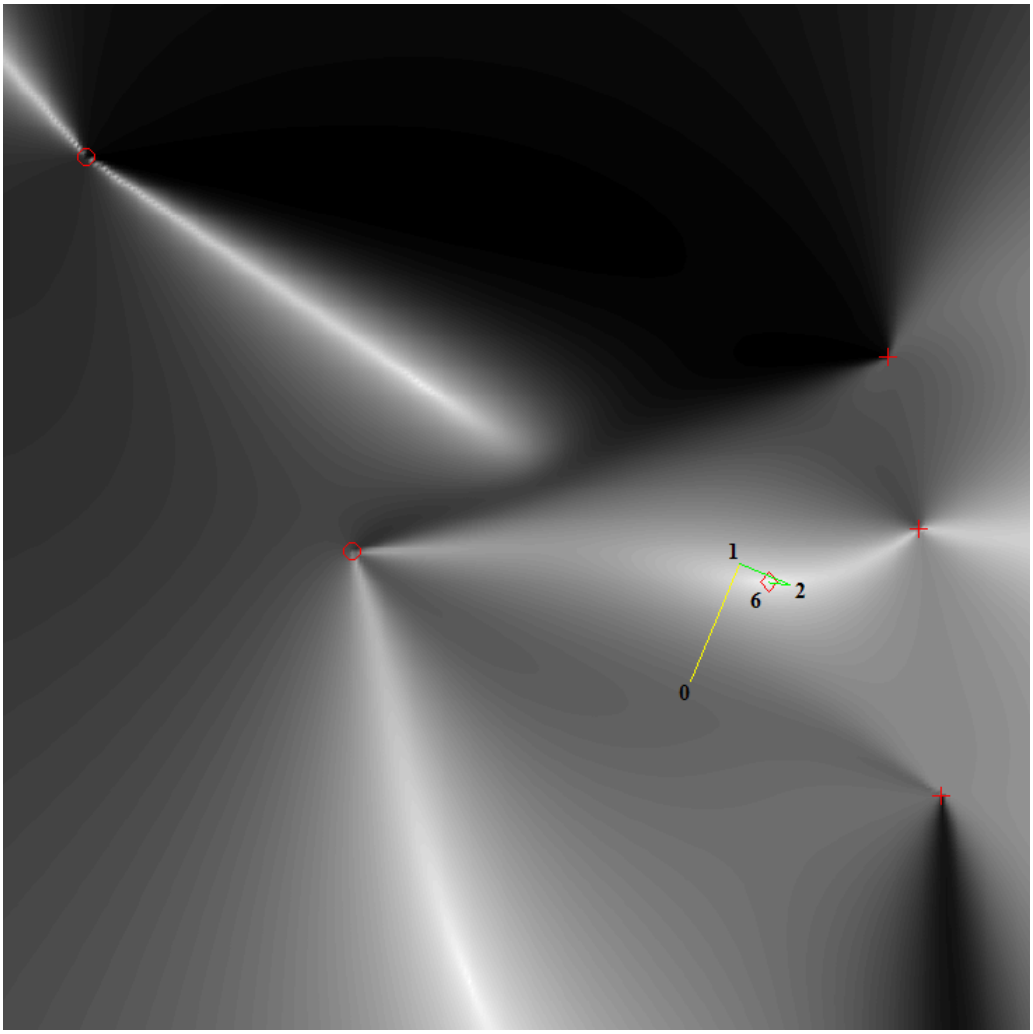

Figure 7 Example of Convergence to Global Minimum

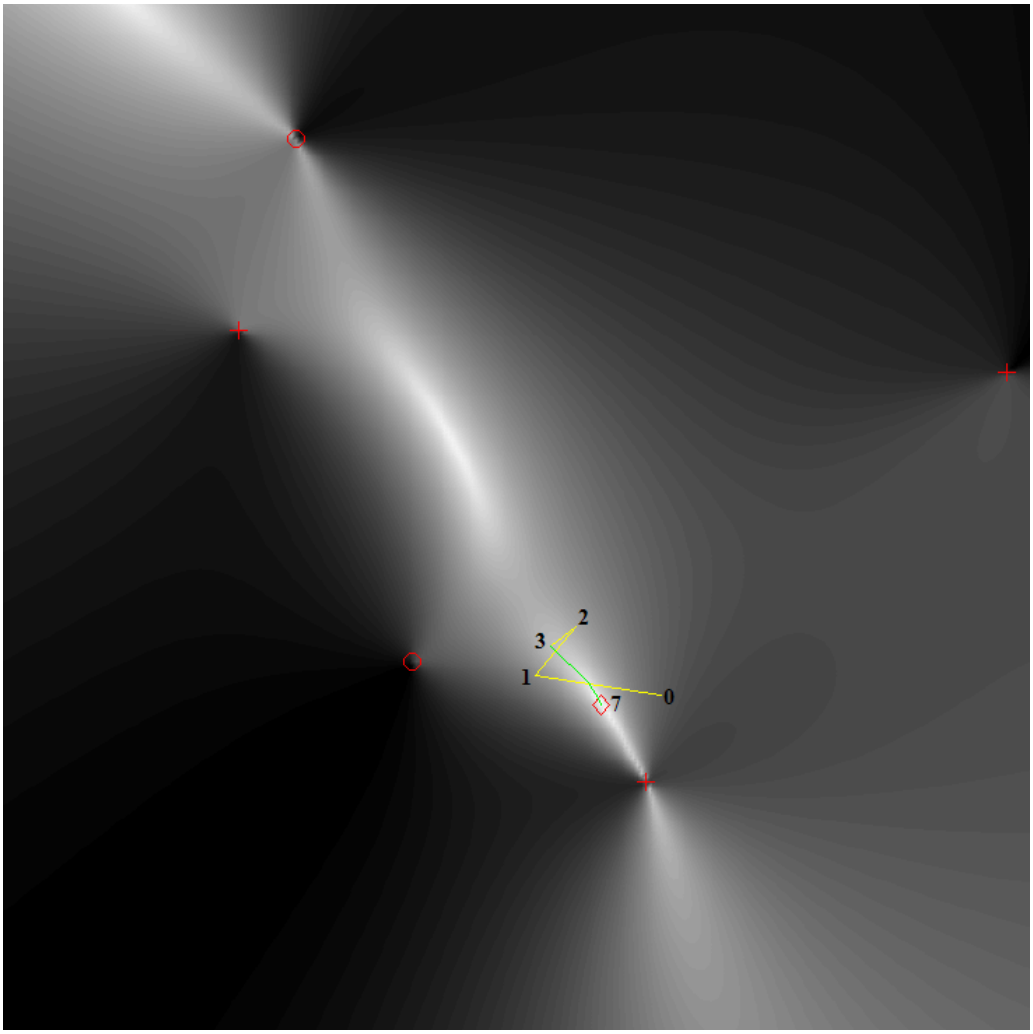

Figure 8 Example of Convergence to Global Minimum

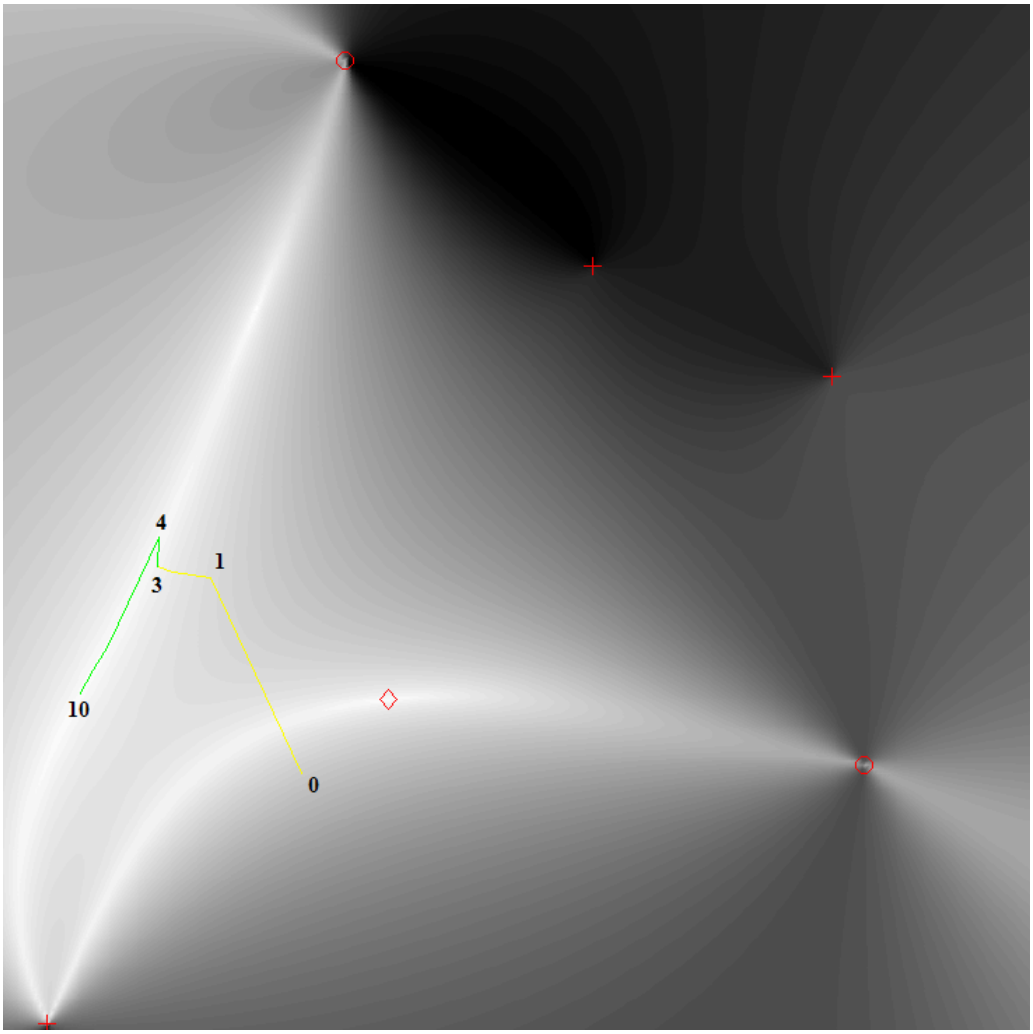

Figure 9 Example of Convergence to Local Minimum

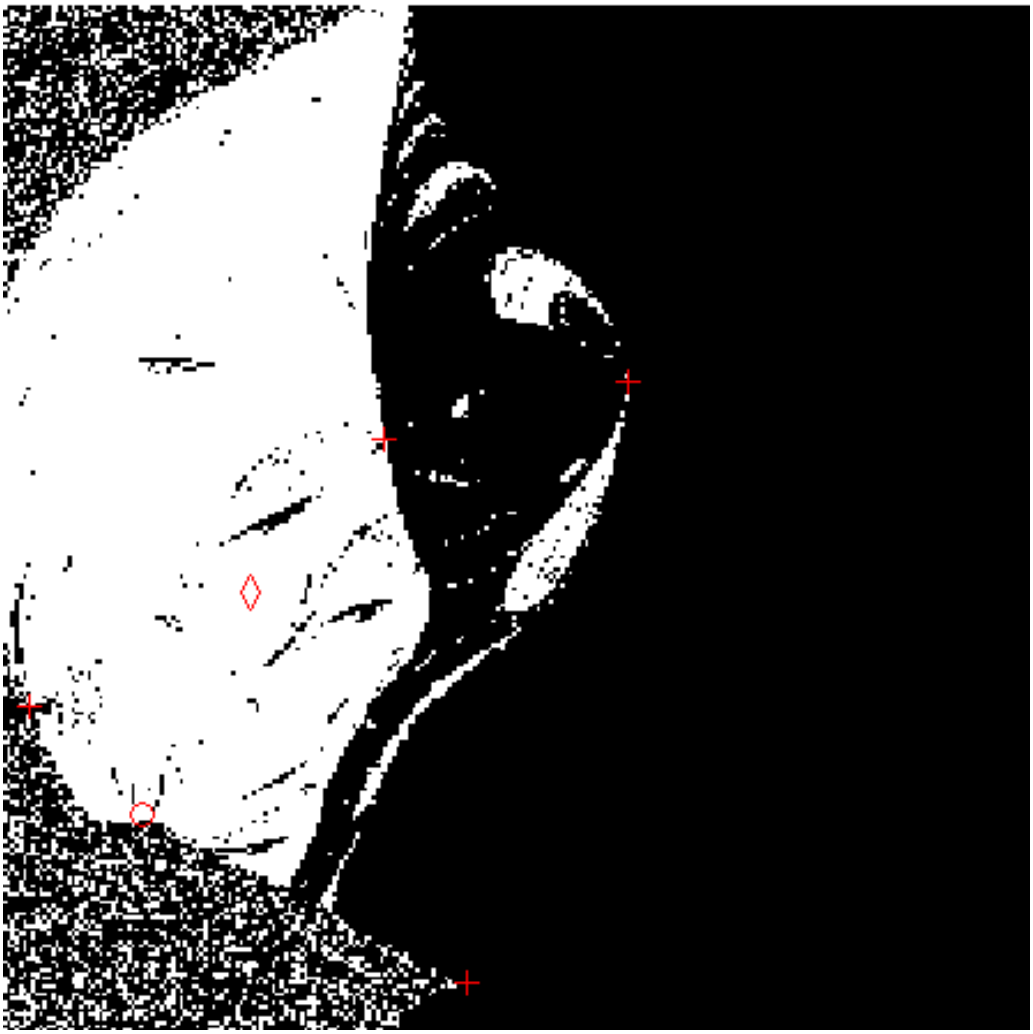

Figure 10 Starting Points for Convergence to Global Minimum (White)

frequently yield convergence to a local, or even global, minimizer provided the iterations do not diverge to infinity. Further, the chances of converging to the optimal target state estimate naturally increase with a finer initial grid-search. In the following section, we will mathematically assess the convergence of algorithms similar to our own.

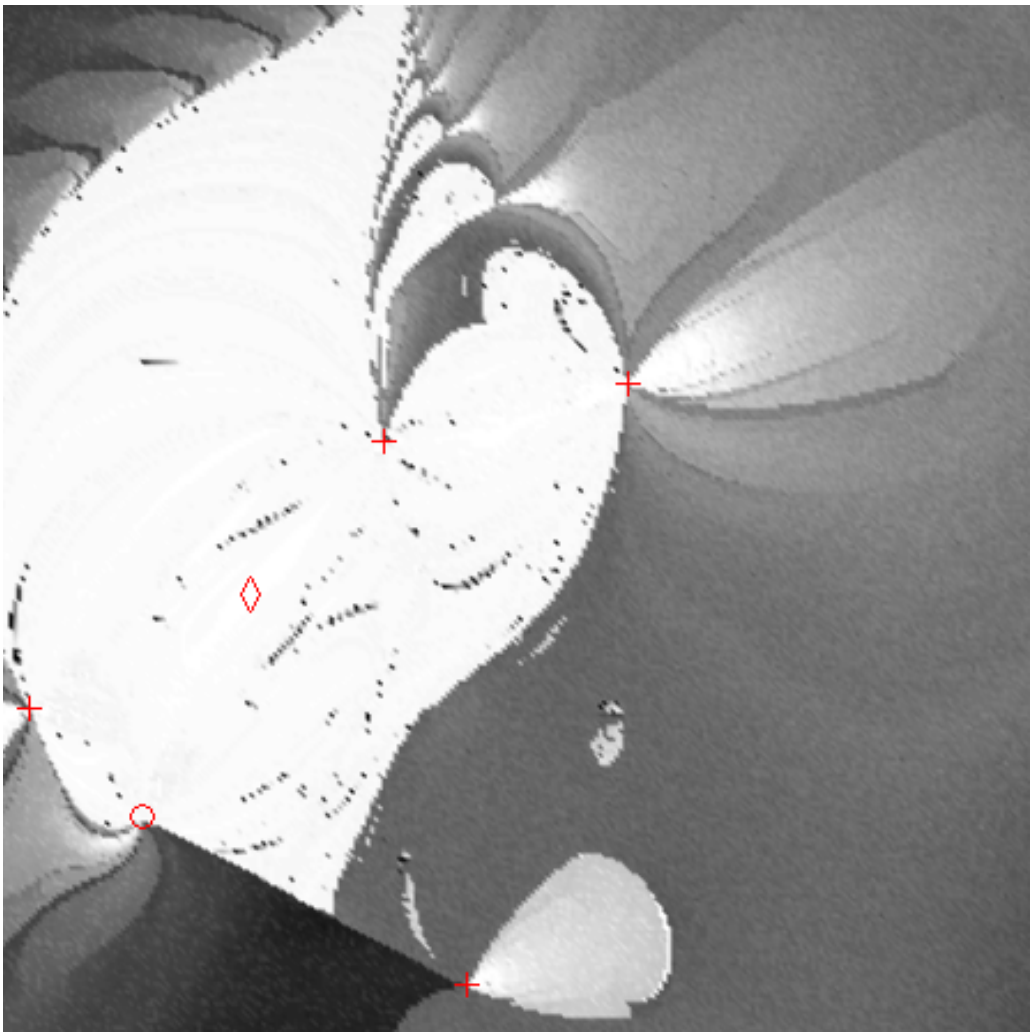

Figure 11 Number of Steps to Converge

5.5 Convergence Issues

An iteration $\{x_k\}_{k=0}^{\infty}$, given by (69), is defined by a search direction p_k and a step length α_k. To assure convergence, Wolfe Conditions are often used to determine the step length [27, p. 39]. These are sufficient decrease and curvature conditions, given by

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k [\nabla f(x_k)]^T p_k, \tag{70}$$

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 [\nabla f(x_k)]^T p_k. \tag{71}$$

With this information in mind, we come to the following theorem [27, p. 43].

Theorem 5.5.1. *Consider any iteration of the form* (69)*, where p_k is a descent direction and α_k satisfies the Wolfe Conditions,* (70) *and* (71)*. Suppose that f is bounded below in $\mathbb{R}^n$ and that f is continuously differentiable in an open set $\mathcal{N}$ containing the level set $\mathcal{L} \equiv \{x : f(x) \leq f(x_0)\}$, where x_0 is the starting point of the iteration. Assume also that the gradient ∇f is Lipschitz continuous on $\mathcal{N}$; that is, there exists a constant $L > 0$ such that*

$$\|\nabla f(x) - \nabla f(\tilde{x})\| \leq L\|x - \tilde{x}\|,$$

for all $x, \tilde{x} \in \mathcal{N}$. Then,

$$\sum_{k=0}^{\infty} \left(\nabla f(x_k) \cdot \frac{x_{k+1} - x_k}{|x_{k+1} - x_k|} \right)^2 < \infty.$$

Recall the induced operator norm $\|\cdot\|_2$ is defined by

$$\|A\|_2 \equiv \sup_{\|v\|=1} \|Av\|,$$

for all matrices A. We also recall the Frobenius norm $\|\cdot\|_F$, defined by

$$\|A\|_F \equiv \left[\sum_{i,j} a_{ij}^2 \right]^{1/2},$$

for all matrices A. We therefore come to the following theorem, which applies Theorem 5.5.1 to our objective function Υ, given in (52).

Theorem 5.5.2. *Let C be a constellation such that $B(x)$ has full column rank for all $x \in \mathbb{R}^N - C$. Then, every bounded iteration $\{x_k\}_{k=0}^{\infty}$ satisfying Wolfe Conditions*

(70) *and* (71)*, and further satisfying*

$$\inf_{\substack{k\in\mathbb{N}\\ y\in C}} |x_k - y| > 0$$

and

$$\frac{\nabla\Upsilon(x_k)}{|\nabla\Upsilon(x_k)|}\cdot\frac{x_{k+1}-x_k}{|x_{k+1}-x_k|} \leq -\delta$$

for all k*, for some* $\delta > 0$*, also satisfies* $\lim_{k\to\infty} |\nabla\Upsilon(x_k)| = 0$.

Proof: We first show that the Hessian of our objective function is bounded on a set which contains our iteration. We then show that this implies that the gradient of our objective function is Lipschitz continuous. Finally, we add other functions of Lipschitz continuous gradients, defined to be zero at each of the points of our iteration, to our objective function to illustrate our desired result.

Pick an iteration $\{x_k\}_{k=0}^{\infty}$, and let

$$\varepsilon \equiv \frac{1}{3}\min\left\{\inf_{\substack{k\in\mathbb{N}\\ y\in C}} |x_k - y|, \frac{1}{2}\min_{y_1,y_2\in C}|y_1 - y_2|\right\}$$

and

$$R \equiv 3\max\left\{\sup_{k\in\mathbb{N}} |x_k - \mu|, \quad \varepsilon + \max_{y\in C}|y-\mu|\right\},$$

where μ is the center of the constellation C, defined by

$$\mu \equiv \frac{1}{2P}\sum_{p=1}^{P}(T_p + R_p).$$

We first show that the Hessian of the objective function is bounded on

$$\overline{\mathcal{N}}(\varepsilon, R) \equiv \left[\mathbb{R}^N - \bigcup_{y\in C} B(y,\varepsilon)\right] \cap \overline{B}(\mu, R).$$

Recall that $\|A\|_2 \leq \|A\|_F$, see [10, p. 28]. Considering Theorem 5.3.1, the Hessian expression yields

$$|[\nabla^2 \Upsilon(x)]_{n,m}|^2 \leq 2\|d\|^4 \|[P_{B(x)^\perp} B^{(n)}(x)[B(x)]^\dagger]^{(m)}\|_2^2. \tag{72}$$

We have by (59) that

$$|B_{p,\cdot}^{(n,m)}(x)| \leq \frac{3}{|x - T_p|} + \frac{3}{|x - R_p|} \leq \frac{6}{\varepsilon},$$

and so

$$\|B^{(n,m)}(x)\|_2 \leq \|B^{(n,m)}(x)\|_F = \left(\sum_{p=1}^{P} |B_{p,\cdot}^{(n,m)}(x)|^2 \right)^{1/2} \leq \frac{6\sqrt{P}}{\varepsilon}.$$

Also note, from (57), that

$$|B_{p,\cdot}^{(n)}(x)| \leq \frac{1}{|x - x_{p_T}|} + \frac{1}{|x - x_{p_R}|} = \frac{2}{\varepsilon},$$

and so, similarly,

$$\|B^{(n)}(x)\|_2 \leq \frac{2\sqrt{P}}{\varepsilon}.$$

We wish to also determine an upper bound for $\|[B(x)]^\dagger\|_2$. Since $B(x)$ has full column rank for all $x \in \mathbb{R}^N - C$, we know $[B(\cdot)]^\dagger$ is continuous. Further, $\overline{\mathcal{N}}(\varepsilon, R) \subset \mathbb{R}^N - C$ is compact. It follows that $\|[B(\cdot)]^\dagger\|_2$ achieves its maximum, which we shall denote, q, in $\overline{\mathcal{N}}(\varepsilon, R)$. It follows, in summary, that

$$\begin{aligned} \|B^{(n,m)}(x)\|_2 &\leq \frac{6\sqrt{P}}{\varepsilon}, \\ \|B^{(n)}(x)\|_2 &\leq \frac{2\sqrt{P}}{\varepsilon}, \\ \|[B(x)]^\dagger\|_2 &\leq q, \end{aligned}$$

and so by (58) and the Triangle Inequality, we have

$$\|[P_{B(x)^\perp} B^{(n)}(x)[B(x)]^\dagger]^{(m)}\|_2 \leq \frac{16q^2P}{\varepsilon^2} + \frac{6q\sqrt{P}}{\varepsilon}.$$

Finally, (72) gives

$$\begin{aligned}
\|\nabla^2\Upsilon(x)\|_2 &\leq \|\nabla^2\Upsilon(x)\|_F \\
&= \left[\sum_{n=1}^{N}\sum_{m=1}^{N} |[\nabla^2\Upsilon(x)]_{n,m}|^2\right]^{1/2} \\
&\leq \left[\sum_{n=1}^{N}\sum_{m=1}^{N} 2\|d\|^4\left(\frac{16q^2P}{\varepsilon^2} + \frac{6q\sqrt{P}}{\varepsilon}\right)^2\right]^{1/2} \\
&= N\sqrt{2}\|d\|^2\left(\frac{16q^2P}{\varepsilon^2} + \frac{6q\sqrt{P}}{\varepsilon}\right), \qquad (73)
\end{aligned}$$

as desired.

Since the Hessian of our objective function is bounded on $\overline{\mathcal{N}}(\varepsilon, R)$, we claim that our gradient is consequently Lipschitz continuous on this same region. Pick $a, b \in \overline{\mathcal{N}}(\varepsilon, R)$ and minimum-length path $r : [0,1] \to \mathbb{R}^N$ such that $r(0) = a$ and $r(1) = b$. Then, if $a \neq b$, we have, by the Fundamental Theorem of Line Integrals, that

$$\begin{aligned}
\frac{|\nabla\Upsilon(b) - \nabla\Upsilon(a)|}{|b-a|} &= \frac{1}{|b-a|}\,|\nabla\Upsilon(r(1)) - \nabla\Upsilon(r(0))| \\
&= \frac{1}{|b-a|}\left|\int_0^1 \nabla^2\Upsilon(r(t))\dot{r}(t)\,\mathrm{d}t\right| \\
&\leq \frac{1}{|b-a|}\int_0^1 \|\nabla^2\Upsilon(r(t))\|_2|\dot{r}(t)|\,\mathrm{d}t \\
&\leq \frac{K}{|b-a|}\int_0^1 |\dot{r}(t)|\,\mathrm{d}t,
\end{aligned}$$

where K denotes the bound (73). Notice that $\int_0^1 |\dot{r}(t)|\,\mathrm{d}t$ is the length of the path $r[0,1]$. Further, this shortest path is usually a straight line from a to b. However,

if the straight line from a to b does not lie entirely in $\overline{\mathcal{N}}(\varepsilon, R)$, due to the fact that it passes through an ε-ball about a point in the constellation, then the path of minimum length is no longer than that which curves along the geodesic of the ε-ball's surface. Moreover, the ratio

$$\frac{1}{|b-a|}\int_0^1 |\dot{r}(t)|\,\mathrm{d}t$$

is maximized when a and b are located at opposite ends of a common ε-ball. It follows that choosing $L \equiv K\pi/2$ implies Lipschitz continuity of the gradient on $\overline{\mathcal{N}}(\varepsilon, R)$, as desired.

We now consider twice-differentiable functions $s_1, s_2 : \mathbb{R} \to \mathbb{R}$ with the following properties:

$$\begin{aligned} s_1(t) &> \|d\|^2,\ t \in [0, \varepsilon^2] \qquad & s_2(t) &= 0,\ t \in [0, R^2/4] \\ s_1(t) &= 0,\ t \in [4\varepsilon^2, \infty) \qquad & s_2(t) &> \|d\|^2,\ t \in [R^2, \infty). \end{aligned}$$

Then, we have

$$\nabla^2 s_i(|x|^2) = 2\dot{s}_i(|x|^2)I + 4\ddot{s}_i(|x|^2)xx^T$$

for $i = 1, 2$. We see that these Hessians are bounded on $\mathbb{R}^N - B(0, \varepsilon)$ and $\overline{B}(0, R)$, respectively. Let us define another function $\widehat{\Upsilon} : \mathbb{R}^N \to \mathbb{R}$ by

$$\widehat{\Upsilon}(x) \equiv \Upsilon(x) + s_2(|x-\mu|^2) + \sum_{y \in C} s_1(|x-y|^2). \tag{74}$$

Since each of the terms in (74) have a bounded Hessian, we know $\widehat{\Upsilon}$ also has a bounded Hessian. Furthermore, the iteration $\{x_k\}_{k=0}^{\infty}$ also satisfies Wolfe Conditions (70) and (71) on $\widehat{\Upsilon}$, since $\{x_k\}_{k=1}^{\infty} \subset \mathrm{int}[\overline{\mathcal{N}}(2\varepsilon, R/2)]$, and $\widehat{\Upsilon}(\tilde{x}) = \Upsilon(\tilde{x})$ for all $\tilde{x} \in \mathrm{int}[\overline{\mathcal{N}}(2\varepsilon, R/2)]$. We see that $\widehat{\Upsilon}$ is continuously differentiable in the open set $\mathcal{N} \equiv \mathrm{int}[\overline{\mathcal{N}}(\varepsilon, R)]$, which contains all level sets of the form $\{x : \widehat{\Upsilon}(x) \leq \widehat{\Upsilon}(x_0)\}$. Further, since $\nabla\widehat{\Upsilon}$ is Lipschitz continuous, by bounded Hessian on $\mathcal{N}$, Theorem

5.5.1 gives

$$\sum_{k=0}^{\infty} \left(\nabla \Upsilon(x_k) \cdot \frac{x_{k+1} - x_k}{|x_{k+1} - x_k|} \right)^2 < \infty.$$

It follows that

$$\left(\frac{\nabla \Upsilon(x_k)}{|\nabla \Upsilon(x_k)|} \cdot \frac{x_{k+1} - x_k}{|x_{k+1} - x_k|} \right)^2 |\nabla \Upsilon(x_k)|^2 \to 0,$$

and since our theorem statement gives

$$\left(\frac{\nabla \Upsilon(x_k)}{|\nabla \Upsilon(x_k)|} \cdot \frac{x_{k+1} - x_k}{|x_{k+1} - x_k|} \right)^2 \geq \delta^2 > 0,$$

we have $|\nabla \Upsilon(x_k)|^2 \to 0$, as desired. ■

Note the condition on the constellation that was used in the statement of the theorem: that $B(x)$ has full column rank for all $x \in \mathbb{R}^N - C$. It is important to note that not all constellations have this characteristic. For example, no constellation of one or two pairs satisfies this condition, since the bearing matrix at a point between a transmitter-receiver pair will have rank of at most one. Furthermore, consider a constellation in which each transmitter and receiver is used exactly once in a pair, and in which the open line segments between each of these pairs all intersect at a point $\hat{x}$. Then, we see that $B(\hat{x}) = 0$, and so $B(\hat{x})$ consequently has rank zero. For each of these examples, Theorem 5.5.2 does not apply.

However, Theorem 5.5.2 does say something meaningful. For non-pathological constellations containing enough transmitter-receiver pairs, any iteration which satisfies the Wolfe Conditions (70) and (71), which neither diverges to infinity nor is attracted to constellation points, and which tends to move in the general direction opposite the gradient, does, in fact, converge. As the algorithm given in the previous section does not take the curvature condition (71) into account, the previous result does not absolutely guarantee its convergence. However, experimentation indicates that convergence is the rule, rather than the exception.

VI. Further Applications of Bearing Derivatives

The results of this chapter are new, and represent the beginnings of future work on the study of the objective function $\Upsilon(x)$, introduced in the previous chapter. To be precise, we study both the asymptotic behavior of $\Upsilon(x)$ for a fixed constellation, as well as begin the important task of determining the optimal placement of transmitters and/or receivers in a Doppler-only multistatic system.

6.1 Asymptotic Analysis

In order to effectively optimize the objective function of the previous chapter, it is important to have an idea of where the local minima lie. More specifically, it is important that they are not too far away from the constellation of transmitters and receivers, or else the objective function becomes unstable. These things considered, we want a symbolic approximation to the objective function at infinity, that is, an asymptotic approximation.

6.1.1 Asymptotic Analysis of the Multistatic Bearing Operator. In order to get an asymptotic approximation for Υ, we first need an asymptotic approximation for B. Considering the form of the rows of B, this means that we need an approximation for $x/|x|$. We already have a Taylor series for $x/|x|$, and amazingly this same computation yields the asymptotics we desire, that is, the Taylor series at infinity. In the rest of the chapter, we shall have $\mathbf{1}_P \in \mathbb{R}^P$ denote the column vector of ones.

Theorem 6.1.1. *For any $\mu \in \mathbb{R}^N$,*

$$B(x) = \frac{2}{|x-\mu|}\left\{\mathbf{1}_P(x-\mu)^T + [\mathbf{1}_P\mu^T - \overline{X}]P_{(x-\mu)^\perp}\right\} + O\left(\frac{1}{|x|^2}\right)$$

as $|x| \to \infty$. *Further, when* μ *is taken to be the center of the constellation, defined to be*

$$\mu \equiv \frac{1}{2P} \sum_{p=1}^{P} (T_p + R_p),$$

the ranges of $\mathbf{1}_P (x-\mu)^T$ *and* $[\mathbf{1}_P \mu^T - \overline{X}] P_{(x-\mu)^\perp}$ *are orthogonal subspaces of* $\mathbb{R}^P$.

Proof: Taylor's Theorem gives that

$$\left| \frac{x+h}{|x+h|} - \sum_{m=0}^{M} \frac{1}{m!} \mathrm{D}^m \frac{x}{|x|} h^m \right| \leq \frac{1}{(M+1)!} \max_{|k| \leq |h|} |\mathrm{D}^{M+1} f(x+k) h^{M+1}|.$$

Considering Corollary 3.2.2 and letting $M = 1$, we therefore have

$$\begin{aligned}
&\left| \frac{x+h}{|x+h|} - \left[\frac{x}{|x|} + \frac{1}{|x|} P_{x^\perp} h \right] \right| \\
&\leq \frac{1}{2} \max_{|k| \leq |h|} \frac{|h|^2}{|x+k|^2} \left| 3 \left(\frac{h \cdot (x+h)}{|h||x+k|} \right)^2 \frac{x+k}{|x+k|} - 2 \left(\frac{h \cdot (x+h)}{|h||x+k|} \right) \frac{h}{|h|} - \frac{x+k}{|x+k|} \right| \\
&\leq \frac{3|h|^2}{(|x| - |h|)^2},
\end{aligned}$$

for $|h| < |x|$. Considering the symmetry between x and h, we also have

$$\left| \frac{x+h}{|x+h|} - \left[\frac{h}{|h|} + \frac{1}{|h|} P_{h^\perp} x \right] \right| \leq \frac{3|x|^2}{(|h| - |x|)^2}, \quad |h| > |x|.$$

That is, for large $|h|$, we have

$$\frac{x+h}{|x+h|} = \frac{h}{|h|} + \frac{1}{|h|} P_{h^\perp} x + O\left(\frac{1}{|h|^2} \right).$$

Pick $\mu \in \mathbb{R}^N$. Then, for large $|x|$, it follows that

$$\frac{x - T_p}{|x - T_p|} = \frac{(\mu - T_p) + (x - \mu)}{|(\mu - T_p) + (x - \mu)|} = \frac{x-\mu}{|x-\mu|} + \frac{1}{|x-\mu|} P_{(x-\mu)^\perp} (\mu - T_p) + O\left(\frac{1}{|x|^2} \right),$$

and similarly

$$\frac{x-R_p}{|x-R_p|} = \frac{x-\mu}{|x-\mu|} + \frac{1}{|x-\mu|} P_{(x-\mu)^\perp}(\mu - R_p) + O\left(\frac{1}{|x|^2}\right),$$

so that

$$[B_{p,\cdot}(x)]^T = 2\frac{x-\mu}{|x-\mu|} + \frac{1}{|x-\mu|} P_{(x-\mu)^\perp}(2\mu - T_p - R_p) + O\left(\frac{1}{|x|^2}\right)$$

and so

$$[B(x)]^T = \frac{2}{|x-\mu|}\left\{(x-\mu)\mathbf{1}_P + P_{(x-\mu)^\perp}[\mu\mathbf{1}_P^T - \overline{X}^T]\right\} + O\left(\frac{1}{|x|^2}\right).$$

Our asymptotic expression for $B(x)$ naturally follows. Now, pick any $u, v \in \mathbb{R}^N$ and set

$$\mu = \frac{1}{2P}\sum_{p=1}^{P}(T_p + R_p).$$

Then, we have

$$\begin{aligned}
&\langle \mathbf{1}_P(x-\mu)^T u, [\mathbf{1}_P\mu^T - \overline{X}]P_{(x-\mu)^\perp} v\rangle \\
&= \langle (x-\mu)^T u, \mathbf{1}_P^T[\mathbf{1}_P\mu^T - \overline{X}]P_{(x-\mu)^\perp} v\rangle \\
&= \left\langle (x-\mu)^T u, \left[P\mu^T - \left(\sum_{p=1}^{P}\frac{T_p+R_p}{2}\right)^T\right] P_{(x-\mu)^\perp} v\right\rangle \\
&= \langle (x-\mu)^T u, 0\rangle \\
&= 0,
\end{aligned}$$

thereby yielding orthogonality. ■

6.1.2 Asymptotic Analysis of the Objective Function. To better understand our objective function, we will study its asymptotic behavior. We have the following results.

Corollary 6.1.2. *We have* $\lim_{|x|\to\infty} \nabla\Upsilon(x) = 0$.

Proof: Clearly $B(x) \to \frac{2}{|x-\mu|}\mathbf{1}_P(x-\mu)^T$, and so $B^\dagger d \to \frac{\mathbf{1}_P^T d}{2|x-\mu|}(x-\mu)$. Thus, since

$$\nabla\Upsilon(x) = -2\left\{\sum_{p=1}^{P}[P_{B(x)^\perp}d]_p\left[\frac{P_{(x-T_p)^\perp}}{|x-T_p|}+\frac{P_{(x-R_p)^\perp}}{|x-R_p|}\right]\right\}[B(x)]^\dagger d,$$

we have $\nabla\Upsilon(x) \to 0$. ■

Corollary 6.1.3. *We have*

$$\lim_{|x|\to\infty}\left\langle\frac{x}{|x|},\frac{\nabla\Upsilon(x)}{|\nabla\Upsilon(x)|}\right\rangle = 0.$$

Proof: Since $B^\dagger d \to \frac{\mathbf{1}_P^T d}{2|x-\mu|}(x-\mu)$ and $P_{(x-x_0)^\perp} \to P_{x^\perp}$ for all $x_0 \in \mathbb{R}^N$, we see from our gradient expression that

$$\left\langle\frac{x}{|x|},\frac{\nabla\Upsilon(x)}{|\nabla\Upsilon(x)|}\right\rangle \to \left\langle\frac{x}{|x|},P_{x^\perp}\left(\frac{-B^\dagger(x)d}{|P_{x^\perp}B^\dagger(x)d|}\right)\right\rangle = 0,$$

as desired. ■

Conjecture 6.1.4. *There exists a unique constellation center* $\mu \in \mathbb{R}^N$ *such that*

$$\lim_{|x|\to\infty}\left\langle x-\mu,\frac{\nabla\Upsilon(x)}{|\nabla\Upsilon(x)|}\right\rangle = 0.$$

Further,

$$\mu = \frac{1}{2P}\sum_{p=1}^{P}(T_p+R_p).$$

This conjecture claims that the level curves of our objective function approach asymptotes which all intersect at a unique constellation center, expressible in terms of the constellation.

6.2 Optimizing the Doppler-Only Multistatic Constellation

6.2.1 Constellation Objective Function. Consider a field of transmitters. How does one determine how to best situate a corresponding field of receivers to best track a given target? One could consider the impact of transmitter-receiver distances on the amplitude of Doppler-shifted signals. However, we will simply consider the effectiveness of the algorithms described in the previous chapter.

In particular, we want to measure how well our algorithms perform, given a specific constellation. In our noise-free situation, we have only to be concerned with the form of the objective function, which is the only portion of our algorithms which depends explicitly on the constellation. Consider the condition number

$$\kappa(B_C(x)) \equiv \|B_C(x)\|_2 \|[B_C(x)]^\dagger\|_2,$$

where $B_C(x)$ is the bearing matrix at point $x \in \mathbb{R}^N$, given constellation C. This number provides a bound on how inaccurate $v^* = [B_C(x)]^\dagger d$ will be after our numerical solution. But, this simply measures the "badness" of the constellation at a specific point $x \in \mathbb{R}^N$. To include all points in our objective function, we consider a weight distribution $F : \mathbb{R}^N \to \mathbb{R}$ and subsequent objective function

$$\int_{\mathbb{R}^N} \kappa(B_C(x)) dF(x).$$

We see that this function shall effectively measure the "badness" of a constellation in our context, as long as our distribution appropriately weights $\mathbb{R}^N$. However, it will be difficult to determine the gradient of this function, as the condition number, having no closed form in terms of the entries of B, is probably impossible to differentiate symbolically. Rather than attempting numerical derivatives, we shall consider

$$K(B_C(x)) \equiv \|B_C(x)\|_F \|[B_C(x)]^\dagger\|_F.$$

Notice

$$\kappa(B_C(x)) \leq K(B_C(x)) \leq N\kappa(B_C(x)),$$

and so κ and K are of similar size. We further note that finding the gradient of

$$\Theta(C) \equiv \int_{\mathbb{R}^N} [K(B_C(x))]^2 dF(x)$$

will be much more manageable.

6.2.2 Gradient Descent. In order to establish minima of our objective function $\Theta(C)$, we will employ the gradient descent method. In this setting, the iteration is of vectors which list the locations of the receivers in a constellation. Therefore, visually, the iteration will resemble a multibody problem, where the receivers move in different directions to decrease the value of the objective function. But, before we can employ this algorithm, we must first find the gradient. Consider the following theorem.

Theorem 6.2.1. *The gradient of our expression is given by*

$$\nabla_{R_p}\Theta = 2\int_{\mathbb{R}^N} \frac{P_{(x-R_p)^\perp} Q(x) e_p dF(x)}{|x - R_p|},$$

where

$$Q \equiv \left\{[\mathrm{Tr}(B^TB)](B^TB)^{-2} + [\mathrm{Tr}(B^TB)^{-1}]I\right\} B^T.$$

Proof: We first note, suppressing dependence on C, that

$$\begin{aligned} B'(x) &\equiv \frac{\partial}{\partial R_p(n)} B(x) \\ &= e_p \frac{\partial}{\partial R_p(n)} \left(\frac{x - T_p}{|x - T_p|} + \frac{x - R_p}{|x - R_p|} \right)^T \\ &= \frac{-1}{|x - R_p|} e_p (P_{(x-R_p)^\perp} e_n)^T. \end{aligned}$$

Further, suppressing dependence on x,

$$\begin{aligned}\frac{\partial}{\partial R_p(n)}\|B\|_F^2 &= \mathrm{Tr}\left[\frac{\partial}{\partial R_p(n)}(B^TB)\right] \\ &= \mathrm{Tr}[(B')^TB + B^TB'] \\ &= \mathrm{Tr}(B^TB').\end{aligned}$$

Considering our expression for B', we have

$$\begin{aligned}\frac{\partial}{\partial R_p(n)}\|B(x)\|_F^2 &= \mathrm{Tr}[(B(x))^TB'(x)] \\ &= \frac{-2}{|x-R_p|}\mathrm{Tr}[(B(x))^Te_pe_n^TP_{(x-R_p)^\perp}] \\ &= \frac{-2}{|x-R_p|}e_n^TP_{(x-R_p)^\perp}(B(x))^Te_p,\end{aligned}$$

and so

$$\nabla_{R_p}\|B(x)\|_F^2 = \frac{-2}{|x-R_p|}P_{(x-R_p)^\perp}(B(x))^Te_p.$$

Also, suppressing dependence on x,

$$B^\dagger(B^\dagger)^T = (B^TB)^{-1}.$$

Thus,

$$\begin{aligned}\frac{\partial}{\partial R_p(n)}\|B^\dagger\|_F^2 &= \frac{\partial}{\partial R_p(n)}\mathrm{Tr}[B^\dagger(B^\dagger)^T] \\ &= \frac{\partial}{\partial R_p(n)}\mathrm{Tr}[(B^TB)^{-1}] \\ &= \mathrm{Tr}\left[\frac{\partial}{\partial R_p(n)}(B^TB)^{-1}\right] \\ &= \mathrm{Tr}\left[-(B^TB)^{-1}\left(\frac{\partial}{\partial R_p(n)}(B^TB)\right)(B^TB)^{-1}\right] \\ &= \mathrm{Tr}[-(B^TB)^{-1}(B^T)'(B^\dagger)^T - B^\dagger B'(B^TB)^{-1}] \\ &= -2\mathrm{Tr}[B^\dagger B'(B^TB)^{-1}]\end{aligned}$$

$$= \quad -2\mathrm{Tr}[B'(B^TB)^{-2}B^T],$$

and considering our expression for B', we have

$$\begin{aligned}
\frac{\partial}{\partial R_p(n)}\|[B(x)]^\dagger\|_F^2 &= -2\mathrm{Tr}\left[\left(\frac{\partial}{\partial R_p(n)}B(x)\right)[(B(x))^TB(x)]^{-2}(B(x))^T\right] \\
&= \frac{2}{|x-R_p|}\mathrm{Tr}[e_pe_n^TP_{(x-R_p)^\perp}[(B(x))^TB(x)]^{-2}(B(x))^T] \\
&= \frac{2}{|x-R_p|}e_n^TP_{(x-R_p)^\perp}[(B(x))^TB(x)]^{-2}(B(x))^Te_p.
\end{aligned}$$

It follows that

$$\nabla_{R_p}\|[B(x)]^\dagger\|_F^2 = \frac{2}{|x-R_p|}P_{(x-R_p)^\perp}[(B(x))^TB(x)]^{-2}[B(x)]^Te_p.$$

We further see that

$$\nabla_{R_p}\|B(x)\|_F^2\|[B(x)]^\dagger\|_F^2 = \frac{2P_{(x-R_p)^\perp}Q(x)e_p}{|x-R_p|},$$

where Q is as defined in the proposition statement. From here, our theorem naturally follows. ■

6.2.3 Using Gradient Descent. In practice, this algorithm seems to be extremely consistent. The following convergence examples were created with a Gaussian weight function of mean T.

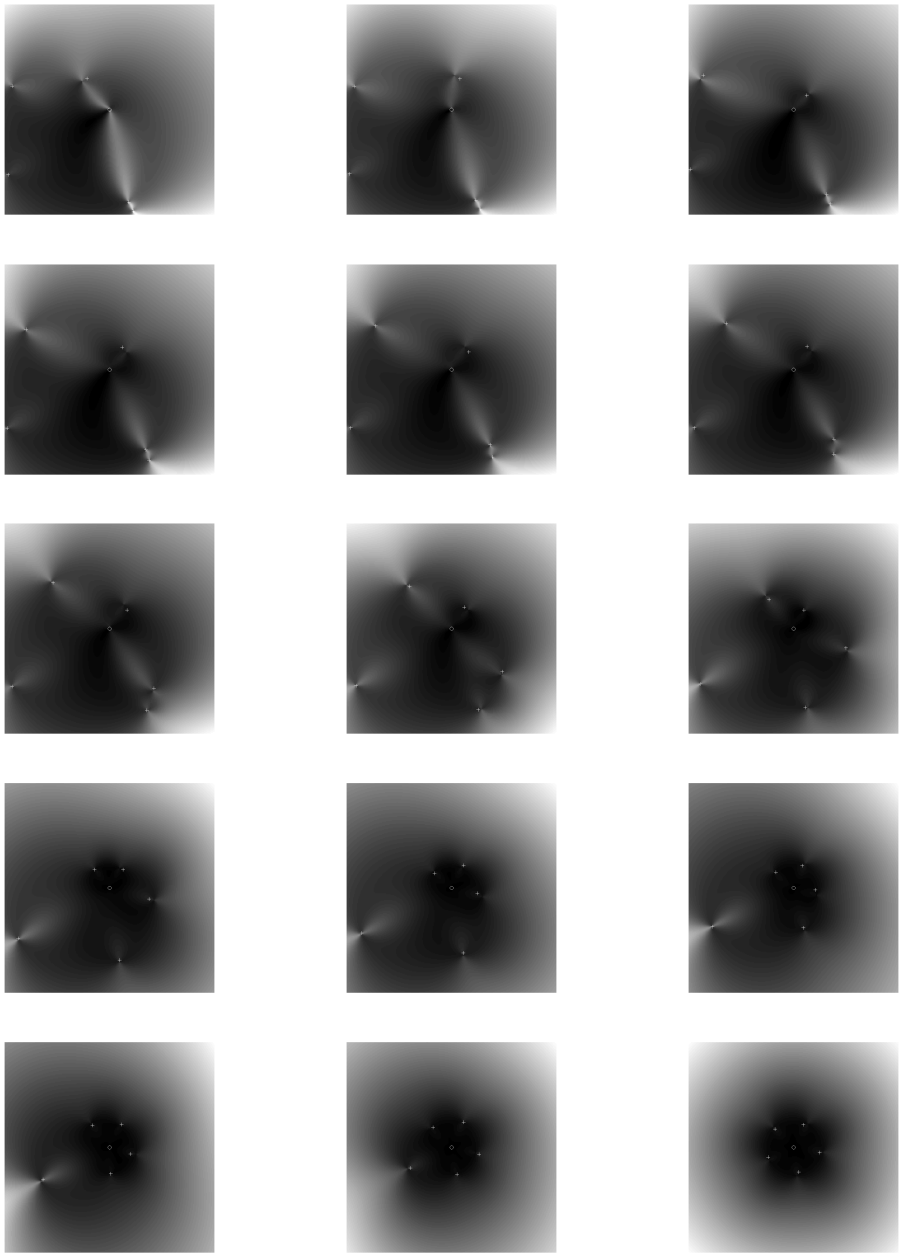

Figure 12 Example of Constellation Convergence with 5 Receivers

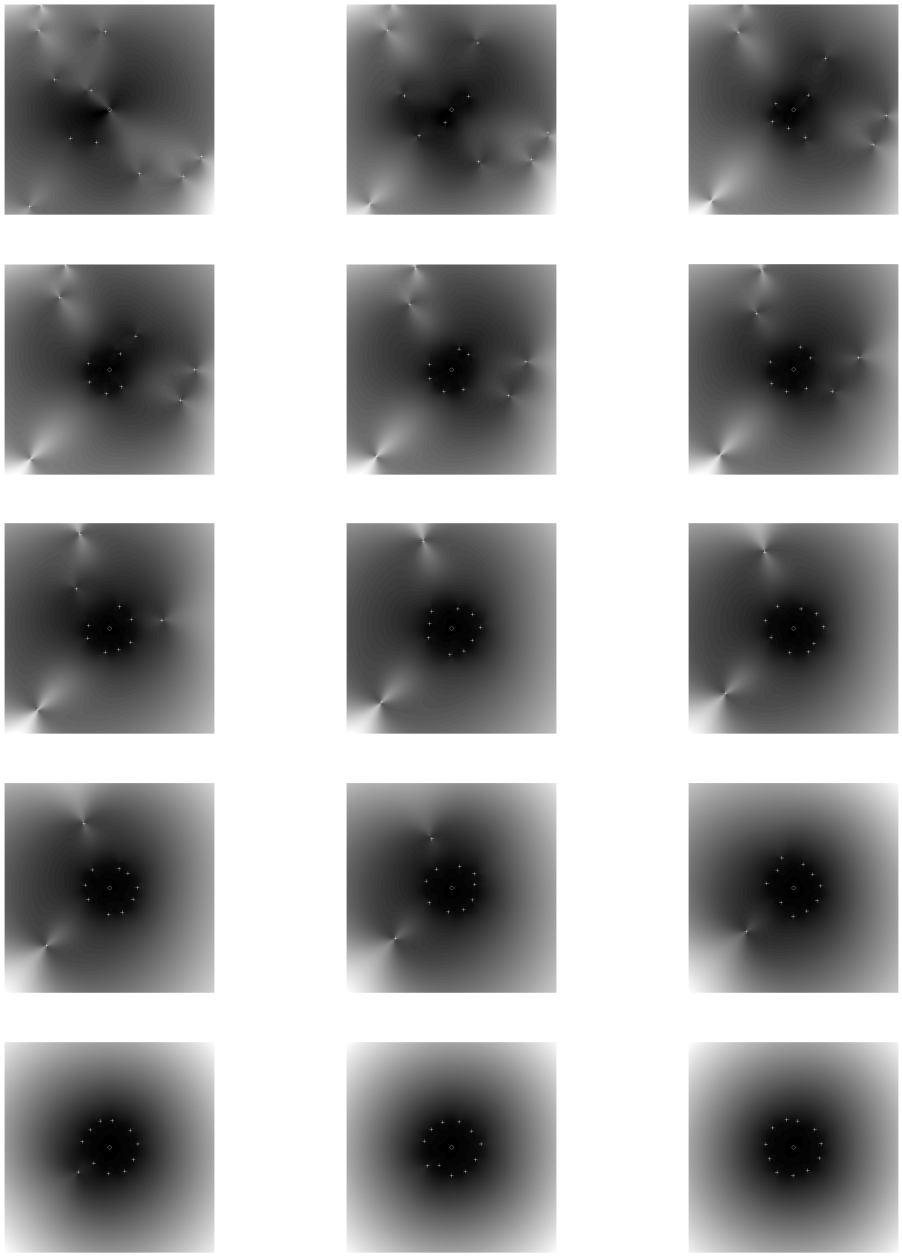

Figure 13 Example of Constellation Convergence with 11 Receivers

6.2.4 Apparent Discoveries. As the reader can gather from these convergence examples, the minimum of our objective function seems unique. Also, considering the convergence tendency of a constellation appears to be independent of its size, we arrive at the following conjecture.

Conjecture 6.2.2. *Suppose $N = 2$, and let Θ have Gaussian weight function of mean T. Then, the one-transmitter constellation which minimizes Θ has each of its receivers uniformly distributed along a circle centered at T.*

One could make a similar conjecture for the case where $N = 3$, but we have not studied this case enough to settle the ambiguity of an analogous uniform distribution on the sphere.

VII. Conclusion

This thesis gives the first deterministic solution to the problem of finding a target's position and velocity from Doppler shift information alone. We have shown that any solution to this problem requires the use of multiple transmitters and receivers. Moreover, when multiple fixed transmitters and receivers are available, we have shown that determining the target state from Doppler shifts is equivalent to solving a nonlinear optimization problem. In particular, a target's position is the minimizer of a specific objective function. Using standard techniques from multivariable calculus and linear algebra, we have discovered a previously unknown formula for every derivative of the Euclidean norm, and used this calculation to determine the gradient and Hessian of this objective function. We then employed our analytic, closed-form expressions for the gradient and Hessian to numerically find the minimizer of our objective function. In so doing, we have found a quick and reliable method for determining a target's position from Doppler-only multistatic measurements. The only noteworthy limitation on the accuracy of our algorithm is the precision with which the Doppler shifts are measured. These results pave the way for a real-world implementation of a Doppler-only multistatic system in which targets can be precisely located and tracked.

Further research in this area could begin with determining the effect of noisy Doppler measurements on our algorithm. In particular, one could study how sensitive the objective function's global minimum is to perturbations in the Doppler information. Another natural next step is to generalize the objective function given in Chapter V so as to include multiple Doppler measurements over time, as well as over multiple transmitter-receiver pairs. Such consideration could potentially lessen the required number of transmitters and/or receivers. Another research topic is the extension of our notion of constellation optimality, introduced in Chapter VI, to more practical cases, that is, constellations with more than one transmitter. Further

generalizations should also consider the rate of decay of a signal's strength, so as to maximize the area over which a Doppler-only multistatic system is useful. Finally, this thesis has left a list of open questions for further research: whether a system of Doppler equations can have a closed-form solution, whether the constellation center bears the given asymptotic property, and whether an optimal single-transmitter constellation in two dimensions is circular.

Appendix A. Multivariable Derivatives

A.1 Derivatives of Functionals

Let $f : \mathbb{R}^N \to \mathbb{R}$ be a differentiable functional. Then, we consider $\mathrm{D}f : \mathbb{R}^N \to L(\mathbb{R}^N, \mathbb{R})$ such that for each $x \in \mathbb{R}^N$, we have linear functional $\mathrm{D}f(x) : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\mathrm{D}f(x)u \equiv \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} = \nabla f(x) \cdot u,$$

for each $u \in \mathbb{R}^N$. Further, we iteratively define $\mathrm{D}^{M+1} f \equiv \mathrm{D}(\mathrm{D}^M f)$. We see that

$$\begin{aligned}
\mathrm{D}^2 f(x)(h_1, h_2) &\equiv [\mathrm{D}^2 f(x) h_2] h_1 \\
&= \left[\mathrm{D}^2 f(x) \sum_{n=1}^{N} h_2(n) e_n \right] h_1 \\
&= \sum_{n=1}^{N} h_2(n) [\mathrm{D}^2 f(x) e_n] h_1 \\
&= \sum_{n=1}^{N} h_2(n) \left[\frac{\partial}{\partial x_n} \mathrm{D}f(x) \right] h_1 \\
&= \sum_{n=1}^{N} h_2(n) \frac{\partial}{\partial x_n} [\mathrm{D}f(x) h_1] \\
&= \nabla [\mathrm{D}f(x) h_1] \cdot h_2.
\end{aligned}$$

One can inductively see that

$$\mathrm{D}^{M+1} f(x) \{h_m\}_{m=1}^{M+1} = \nabla [\mathrm{D}^M f(x) \{h_m\}_{m=1}^{M}] \cdot h_{M+1}.$$

We therefore have some understanding of how to take an Mth derivative of some differentiable functional on $\mathbb{R}^N$. To do so, one must "build up" to this derivative by first taking lower derivatives. Another way to look at the Mth derivative in

directions $\{h_m\}_{m=1}^M$ is by the following evaluation rule:

$$\begin{aligned} \mathrm{D}^M f(x)\{h_m\}_{m=1}^M &= \mathrm{D}^M f(x) \left\{ \sum_{k_m=1}^M h_m(k_m) e_{k_m} \right\}_{m=1}^M \\ &= \sum_{k_1,\ldots,k_M} \left[\prod_{m=1}^M h_m(k_m) \right] \mathrm{D}^M f(x)\{e_{k_m}\}_{m=1}^M. \end{aligned}$$

Here, derivatives can be taken in $\{h_m\}_{m=1}^M$ once derivatives are calculated in the standard directions $\{e_m\}_{m=1}^M$. Let us now consider the product rule for this form of differentiation.

A.2 The Product Rule

Consider functions $f, g : \mathbb{R}^N \to \mathbb{R}$. Then, we have first derivative

$$\begin{aligned} \mathrm{D}(fg)(x)h &= \nabla(fg)(x) \cdot h \\ &= [f(x)\nabla g(x) + g(x)\nabla f(x)] \cdot h \\ &= f(x)(\nabla g(x) \cdot h) + g(x)(\nabla f(x) \cdot h) \\ &= f(x)\mathrm{D}g(x)h + g(x)\mathrm{D}f(x)h, \end{aligned}$$

second derivative

$$\begin{aligned} \mathrm{D}^2(fg)(x)(h_1,h_2) &= \nabla(f(x)\mathrm{D}g(x)h_1 + g(x)\mathrm{D}f(x)h_1) \cdot h_2 \\ &= [\mathrm{D}g(x)h_1\nabla f(x) + f(x)\nabla\mathrm{D}g(x)h_1 \\ &\qquad + \mathrm{D}f(x)h_1\nabla g(x) + g(x)\nabla\mathrm{D}f(x)h_1] \cdot h_2 \\ &= \mathrm{D}g(x)h_1\mathrm{D}f(x)h_2 + f(x)\mathrm{D}^2g(x)(h_1,h_2) \\ &\qquad + \mathrm{D}f(x)h_1\mathrm{D}g(x)h_2 + g(x)\mathrm{D}^2f(x)(h_1,h_2), \end{aligned}$$

and third derivative

$$\begin{aligned}
\mathrm{D}^3(fg)(x)(h_1,h_2,h_3) &= \nabla \mathrm{D}^2(fg)(x)(h_1,h_2)\cdot h_3 \\
&= f(x)\mathrm{D}^3 g(x)(h_1,h_2,h_3) + \mathrm{D}f(x)(h_3)\mathrm{D}^2 g(x)(h_1,h_2) \\
&\quad + \mathrm{D}f(x)(h_2)\mathrm{D}^2 g(x)(h_1,h_3) + \mathrm{D}f(x)(h_1)\mathrm{D}^2 g(x)(h_2,h_3) \\
&\quad + \mathrm{D}^2 f(x)(h_1,h_2)\mathrm{D}g(x)(h_3) + \mathrm{D}^2 f(x)(h_1,h_3)\mathrm{D}g(x)(h_2) \\
&\quad + \mathrm{D}^2 f(x)(h_2,h_3)(h_1) + \mathrm{D}^3 f(x)(h_1,h_2,h_3)g(x).
\end{aligned}$$

One can inductively see that

$$\mathrm{D}^M(fg)(x)\{h_m\}_{m=1}^M = \sum_{k=0}^{M} \sum_{\substack{A\subseteq\{1,\ldots,M\} \\ |A|=k}} \mathrm{D}^k f(x)\{h_i\}_{i\in A}\mathrm{D}^{M-k} g(x)\{h_j\}_{j\notin A}.$$

Notice that if $h_m = h$ for all m, then our product rule would follow the Binomial Theorem, as expected.

A.3 The Chain Rule

The chain rule naturally extends to our derivatives, as expected. Consider the following proposition.

Proposition A.3.1. *Suppose $x : \mathbb{R} \to \mathbb{R}^N$ is differentiable at t, and $f : \mathbb{R}^N \to \mathbb{R}$ is differentiable at $x(t)$. Then, we have*

$$\frac{d}{dt} f(x(t)) = \mathrm{D}f(x(t))\dot{x}(t).$$

Proof: Let Δt be an increment in t, Δx an increment in $x(t)$ and Δf in $f(x(t))$. Then, we see that

$$|\Delta x| = |\dot{x}(t)|\Delta t + o(\Delta t) = \left[|\dot{x}(t)| + \frac{o(\Delta t)}{\Delta t}\right]\Delta t$$

$$\Delta f \;=\; \nabla f(x(t)) \cdot \Delta x + o(|\Delta x|) = \left[\nabla f(x(t)) \cdot \frac{\Delta x}{|\Delta x|} + \frac{o(|\Delta x|)}{|\Delta x|}\right] |\Delta x|,$$

and so we have

$$\frac{\Delta f}{\Delta t} = \left[\nabla f(x(t)) \cdot \frac{\Delta x}{|\Delta x|} + \frac{o(|\Delta x|)}{|\Delta x|}\right] \left[|\dot{x}(t)| + \frac{o(\Delta t)}{\Delta t}\right].$$

It follows that

$$\begin{aligned}
\frac{d}{dt} f(x(t)) &= \lim_{\Delta t \to 0} \frac{\Delta f}{\Delta t} \\
&= \lim_{\Delta t \to 0} \nabla f(x(t)) \cdot \frac{|\dot{x}(t)|}{|\Delta x|} \Delta x \\
&= \nabla f(x(t)) \cdot \dot{x}(t) \\
&= \mathrm{D} f(x(t)) \dot{x}(t),
\end{aligned}$$

as desired. ■

Further, an extended chain rule also holds.

Proposition A.3.2. *We have*

$$\begin{aligned}
&\frac{d}{dt} \mathrm{D}^n f(x(t))(y_1(t), \ldots, y_n(t)) \\
&\quad = \mathrm{D}^{n+1} f(x(t))(y_1(t), \ldots, y_n(t), \dot{x}(t)) \\
&\qquad + \sum_{j=1}^{n} \mathrm{D}^n f(x(t))(y_1(t), \ldots, \dot{y}_j(t), \ldots, y_n(t)).
\end{aligned}$$

Proof: Let $\mathbf{N} \equiv \{1, \ldots, N\}$. Then, using the notation for mixed partial derivatives from Chapter V, since

$$\mathrm{D}^n f(x(t))(y_1(t), \ldots, y_n(t)) = \sum_{a \in \mathbf{N}^n} f^{(a)}(x(t)) \prod_{i=1}^{n} y_i(t)(a_i),$$

we have

$$
\begin{aligned}
&\frac{d}{dt}\mathrm{D}^n f(x(t))(y_1(t),\ldots,y_n(t)) \\
&= \frac{d}{dt}\sum_{a\in\mathbf{N}^n} f^{(a)}(x(t))\prod_{i=1}^{n} y_i(t)(a_i) \\
&= \sum_{a\in\mathbf{N}^n}\left[\left(\frac{d}{dt}f^{(a)}(x(t))\right)\prod_{i=1}^{n} y_i(t)(a_i) + \sum_{j=1}^{n} f^{(a)}(x(t))\dot{y}_j(t)(a_j)\prod_{i\neq j} y_i(t)(a_i)\right] \\
&= \sum_{a\in\mathbf{N}^n}\left(\nabla f^{(a)}(x(t))\cdot\dot{x}(t)\right)\prod_{i=1}^{n} y_i(t)(a_i) + \sum_{a\in\mathbf{N}^n}\sum_{j=1}^{n} f^{(a)}(x(t))\dot{y}_j(t)(a_j)\prod_{i\neq j} y_i(t)(a_i) \\
&= \nabla\left[\sum_{a\in\mathbf{N}^n} f^{(a)}(x(t))\prod_{i=1}^{n} y_i(t)(a_i)\right]\cdot\dot{x}(t) + \sum_{j=1}^{n}\mathrm{D}^n f(x(t))(y_1(t),\ldots,\dot{y}_j(t),\ldots,y_n(t)) \\
&= \nabla\mathrm{D}^n f(x(t))(y_1(t),\ldots,y_n(t))\cdot\dot{x}(t) + \sum_{j=1}^{n}\mathrm{D}^n f(x(t))(y_1(t),\ldots,\dot{y}_j(t),\ldots,y_n(t)) \\
&= \mathrm{D}^{n+1} f(x(t))(y_1(t),\ldots,y_n(t),\dot{x}(t)) + \sum_{j=1}^{n}\mathrm{D}^n f(x(t))(y_1(t),\ldots,\dot{y}_j(t),\ldots,y_n(t)),
\end{aligned}
$$

as desired. ■

Bibliography

[1] Abel, J. "Optimal sensor placement for passive source localization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 5, pp. 2927–2930, Apr. 1990.

[2] Arfken, G. *Mathematical Methods for Physicists*. 3rd ed. Orlando FL: Academic, 1985.

[3] Armstrong, B. and R. Holeman. "Target tracking with a network of Doppler radars," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 33–48, Jan. 1998.

[4] Becker, K. "A general approach to TMA observability from angle and frequency measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, pp. 487–494, Jan. 1996.

[5] Becker, K. "Three-dimensional target motion analysis using angle and frequency measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, pp. 284–301, Jan. 2005.

[6] Chan, Y. and F. Jardine. "Target localization and tracking from Doppler-shifted measurements," *IEEE J. Oceanic Eng.*, vol. 15, pp. 251–257, July 1990.

[7] Chan, Y. and S. Rudnicki. "Bearings-only and Doppler-bearing tracking using instrumental variables," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 1076–1083, Oct. 1992.

[8] Chan, Y. and J. Towers. "Passive localization from Doppler shifted frequency measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 2, pp. 1465–1468, Apr. 1991.

[9] Chan, Y. and J. Towers. "Sequential localization of a radiating source by Doppler-shifted frequency measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 1084–1090, Oct. 1992.

[10] Datta, B. *Numerical Linear Algebra and Applications*. New York: Brooks/Cole, 1995.

[11] Dommermuth, F. "Probabilistic modeling of bistatic Doppler shift," *IEE Proc. Radar, Sonar Navig.*, vol. 148, pp. 348–352, Dec. 2001.

[12] Germond, A. and J. Saillard. "Nine polarimetric bistatic target equations," *IEE Electronic Letters Online*, vol. 33, pp. 1494–1495, Aug. 1997.

[13] Glaser, J. "Some results in the bistatic radar cross section (RCS) of complex objects," *Proceedings of the IEEE*, vol. 77, pp. 639–648, May 1989.

[14] Griffith, E. and K. Kumar. "On the observability of nonlinear systems," *J. Math. Anal. Appl.*, vol. 35, pp. 135–147, 1971.

[15] Griffiths, H., C. Backer, H. Ghaleb, R. Ramakrishnan and E. Willman. "Measurement and analysis of ambiguity functions of off-air signals for passive coherent location," *IEE Electronic Letters Online*, vol. 39, pp. 1005–1007, June 2003.

[16] Herman, S. and P. Moulin. "A particle filtering approach to FM-band passive radar tracking and automatic target recognition," *Proceedings of the IEEE Aerospace Conference*, vol. 4, pp. 1789–1808, 2000.

[17] Ho, K. and W. Xu. "An accurate algebraic solution for moving source location using TDOA and FDOA measurements," *IEEE Trans. Signal Processing*, vol. 52, pp. 2453–2463, Sept. 2004.

[18] Howland, P. "Passive tracking of airborne targets using only Doppler and DOA information," *IEE Colloquium on Algorithms for Target Tracking*, pp. 37–39, May 1995.

[19] Howland, P. "Target tracking using television-based bistatic radar," *IEE Proc. Radar, Sonar Navig.*, vol. 146, pp. 166–174, June 1999.

[20] Jauffret, C. and Y. Bar-Shalom. "Track formation with bearing and frequency measurements in clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, pp. 999–1010, Nov. 1990.

[21] Klemm, R. "Comparison between monostatic and bistatic antenna configurations for STAP," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, pp. 596–608, Apr. 2000.

[22] Kou, S., D. Elliott and T. Tarn. "Observability of nonlinear systems," *Information and Control*, vol. 22, pp. 89–99, 1973.

[23] Lanterman, A. and D. Munson. "Deconvolution techniques for passive radar imaging," *Proc. SPIE*, vol. 4727, pp. 166–177, Aug. 2002.

[24] Levanon, N. "Some results from utilizing Doppler derivatives," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-16, pp. 727–729, Sept. 1980.

[25] Levesque, I. and J. Bondaryk. *Performance Issues Concerning Doppler-Only Localization of Submarine Targets*. Saclant Undersea Research Centre le Spezia (Italy), July 2000 (A179983).

[26] "NAVSPASUR." Unpublished military article online. n. pag. `http://www.nrl.navy.mil`. 17 July 2005.

[27] Nocedal, J. and S. Wright. *Numerical Optimization*. New York: Springer, 1999.

[28] Ringer, M. and G. Frazer. "Waveform analysis of transmissions of opportunity for passive radar," *International Symposium on Signal Processing and its Applications*, vol. 5, pp. 511–514, Aug. 1999.

[29] Schultheiss, P. and E. Weinstein. "Estimation of differential Doppler shifts," *J. Acoust. Soc. Am.*, vol. 66, pp. 1412–1419, Nov. 1979.

[30] Secor, H. "Tesla's views on electricity and the war," *Electrical Experimenter*, vol. 5, 1917.

[31] Shensa, M. "On the uniqueness of Doppler tracking," *J. Acoust. Soc. Am.*, vol. 70, pp. 1062–1064, Oct. 1981.

[32] "Space Surveillance." Unpublished article online. n. pag. `http://www.globalsecurity.org`.17 July 2005.

[33] Statman, J. and E. Rodemich. "Parameter estimation based on Doppler frequency shifts," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 31–39, Jan. 1987.

[34] Swords, S. *Technical History of the Beginnings of RADAR*, IEE History of Technology Series, vol. 6. London, 1986.

[35] Tsao, T., M. Slamani, P. Varshney, D. Weiner, H. Schwarzlander and S. Borek. "Ambiguity function for a bistatic radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, pp. 1041–1051, July 1997.

[36] Webster, R. "An exact trajectory solution from Doppler shift measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-18, pp. 249–252, Mar. 1982.

[37] Weinstein, E. "Optimal source localization and tracking from passive array measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 69–76, Feb. 1982.

[38] Weinstein, E. and N. Levanon. "Passive array tracking of a continuous wave-transmitting projectile," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-16, pp. 721–726, Sept. 1980.

[39] Willis, N. *Bistatic Radar*. Technology Service: Silver Spring, MD, 1995.

[40] Wu, Y. and D. Munson. "Multistatic passive radar imaging using the smoothed pseudo Wigner-Ville distribution," *International Conference on Image Processing*, vol. 3, pp. 604–607, 2001.

[41] Wu, Y. and D. Munson. "Multistatic synthetic aperture imaging of aircraft using reflected television signals," *Proc. SPIE*, vol. 4382, pp. 1–12, Aug. 2001.

[42] Zhang, Y. and B. Himed. "Effects of geometry on clutter characteristics of bistatic radars," *IEEE Radar Conferences*, pp. 417–424, 2003.