SYSTEM IDENTIFICATION FOR PRECISION CONTROL OF A WINGSAILED GPS-GUIDED CATAMARAN

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Gabriel Hugh Elkaim

December 2001

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Dr. Bradford W. Parkinson, Principal Advisor

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Dr. Per Enge

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Dr. J. David Powell

Approved for the University Committee on Graduate Studies

Abstract

This thesis details the Atlantis project, whose aim is the design, development, and experimental testing of an autonomous wind-propelled marine craft. Functionally, such a vehicle is the marine equivalent of an unmanned aerial vehicle (UAV), and would serve similar purposes. The Atlantis project has been able to demonstrate an advance in control precision of a wind-propelled marine vehicle from typical commercial autopilot accuracy of 100 meters to an accuracy of better than one meter with a prototype based on a modified Prindle-19 light catamaran. The project involves substantial innovations in three areas: wind-propulsion system, overall system architecture, and sensors.

The wind-propulsion system is a rigid wing-sail mounted vertically on bearings, mass balanced to allow free rotation in azimuth about a stub-mast. Aerodynamic torque about the stub-mast is trimmed using a flying tail mounted on booms aft of the wing. This arrangement allows the wing-sail to automatically attain the optimum angle to the wind, and weathervane into gusts without inducing large heeling moments.

The sensor system uses differential Global Positioning System (DGPS) augmented by a low-cost attitude system based on accelerometer- and magnetometer-triads for position and velocity measurements. Accurate attitude determination is required to create a synthetic position sensor that is located at the center-of-gravity (c.g.) of the boat, rather than at the Global Positioning System (GPS) antenna location. A high-performance estimator/controller was implemented and tested on the full-scale prototype. The identified controllers were able to perform remarkably well, in the presence of wind and waves, tracking the desired line to within 0.3 meters (~1 foot).

To my brother, Avi, who would have enjoyed this.

Aron "Avi" Moshe Elkaim
1965 — 1993

Acknowledgments

I started this project with an unrealistic view of the effort it would require. I was even less aware of what it would require from my colleagues, friends, and family. Without their contributions, at various times and in countless ways, the Atlantis Project would never have occurred.

My advisor, Professor Bradford W. Parkinson, gave me the freedom to pursue my dream, the guidance to keep it on track, and the inspiration (and sometimes prodding) to drive it to completion. My readers, Professors Per Enge and J. David Powell, provided context and a voice of realism along the way and then read the entire thesis, improving it significantly. My thanks, too, to the two professors on my defense committee, Professor Juan Alonso and my committee chairman, Professor Chris Gerdes.

My family — my parents, Meyer and Helen, my sister, Estreilla, and my brother-in-law, Larry, my nieces, Ariella and Anica, and the family dog, Sasha — all provided tremendous unconditional love and support. No matter the time or weather, I could always count on my father to be available, helping to assemble the Atlantis or as support for yet another experimental trial. Special thanks to my wife, Silvia Grossmann, who designed the beautiful (and award winning) logo for this project, as well as providing love, laughs, a vision of the future, and a home for the past six years.

All of the students of the WAAS, LAAS, and Deere projects deserve mention, but especially I'd like to thank Chad, Konstantin, Lee, Demoz, Jon Fay, Andy Rekow, Eric, Sherman, Guttorm, Sasha, and Sharon. I'd like to thank Professors Juan Alonso and Ilan

Kroo for their help with the airfoil design, and Professor Ed Carryer for his smart product design class that taught me pretty much all I know about electronics.

My proof-readers, Pauline, my brother-in-law, Larry, and my father, Meyer, helped make this thesis coherent and cohesive. The HEPL staff, Jan, Jerri, Dee, Pete, Jeff, Jennifer, Mike, and Robin, allowed me to spend my time focused on the research itself. Brad Parkinson's administrative assistants over the years, Denise, Vivian, and Noriko, kept me abreast of his schedule and travels. Thanks to the Aero/Astro department staff, Sally, Yasmeen, Meg, and Ralph, who made sure that I remained a student in good standing. Art Lange at Trimble Navigation provided some of the hardware used on the project.

I am grateful to Stanford University for assembling so many people I now count as friends, and for the opportunity to tackle something of this size and scope. No matter where I go in life, I will always look back on this time at Stanford with great fondness.

Finally, I'd like to thank the Office of Technology Licensing (OTL) at Stanford, and specifically Mary Watanabe, for arranging the birdseed funding that enabled this entire project. I hope that you find the money well-spent.

1 Introduction

This thesis details the four year course of the Atlantis Project, which began in March of 1997 with the concept of an unmanned, autonomous, Global Positioning System (GPS)-guided, Wingsail-propelled sailboat. This is a "systems" thesis, achieving substantial innovations in the areas of wind-propulsion, overall system architecture, sensors, system identification, and control.

Functionally, the Atlantis is the marine equivalent of an unmanned aerial vehicle and would serve similar purposes such as remote monitoring, unmanned ferrying, environmental monitoring, surveillance, and minefield mapping. The Atlantis project has been able to demonstrate an advance in control precision of a wind-pro-

pelled marine vehicle from typical commercial autopilot accuracy of 100 meters to an accuracy of better than one meter. This quantitative improvement enables new applications, including unmanned station-keeping for navigation or communication purposes, autonomous “dock-to-dock” capabilities, emergency “return unmanned” functions, and many others still to be developed. The prototype is based on a modified Prindle-19 light catamaran, pictured below in Figure 1-1.

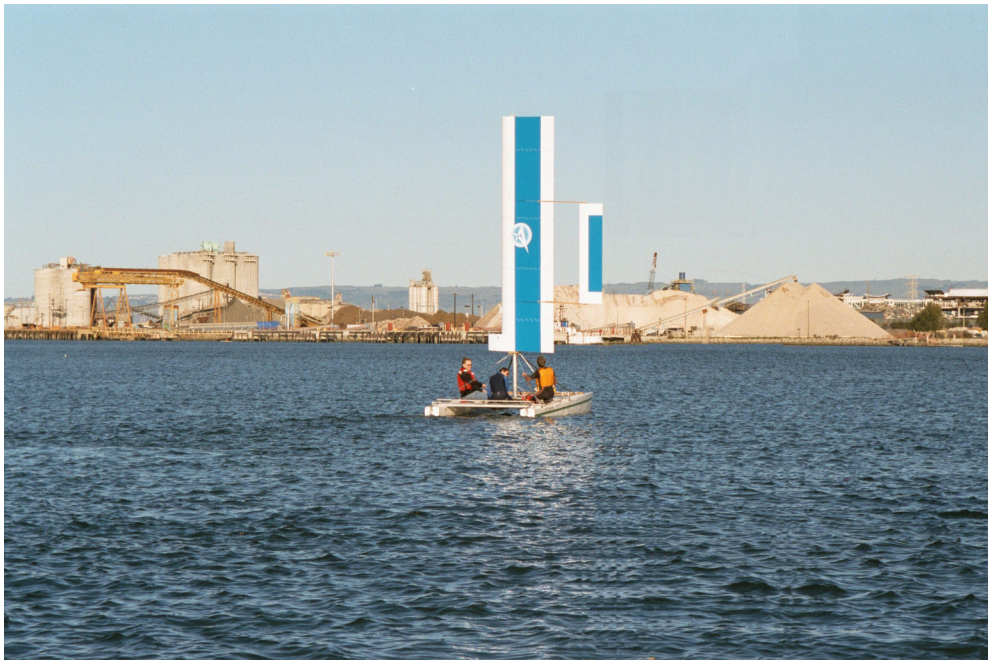

Figure 1-1 The prototype Atlantis GPS-autonomous sailing vessel on her maiden voyage in Redwood City harbor, January 2001. Note the vertical wingsail with tail, and the members of her crew acting as live ballast in case of a sudden wind gust.

The wind-propulsion system is a rigid wingsail mounted vertically on bearings to allow free rotation in azimuth about a stub-mast. Aerodynamic torque about the stub-mast is trimmed using a flying tail mounted on booms joined to the wing. This arrangement allows the wing-sail to automatically attain the optimum angle to the wind and weathervane into gusts without inducing large heeling moments. Modern airfoil design allows for

an increased lift-drag (L/D) ratio over a conventional sail, providing increased thrust while reducing the overturning moment.

The system architecture is based on distributed sensing and actuation with a high-speed digital serial bus connecting the various modules together. Sensors are sampled at 100 Hz, and a central main computer performs the estimation and control tasks at 5 Hz. This bandwidth has been demonstrated to be sufficient for the precise control of the catamaran while well within the capability of off-the-shelf hardware. The distributed architecture is both more robust and less expensive than systems that employ a high-speed, and often analog, star-configuration topology with centralized sensor interpretation and actuation.

The sensor system uses differential GPS (DGPS) augmented by a low-cost attitude system based on accelerometer- and magnetometer-triads for position and velocity measurements. Accurate attitude determination is required to create a synthetic position sensor that is located at the center-of-gravity (CG) of the boat rather than at the GPS antenna location. This is sometimes called the “lever arm” correction.

Experimental trials recorded sensor and actuator data intended to excite all system modes. A simple kinematic system model, based only on geometry, was used as a baseline for comparison. A system model was assembled using Observer/Kalman System Identification (OKID) techniques. Utilizing the kinematic model as a baseline reference, a novel method of modeling the system as velocity invariant was developed such that the gain is input scaled rather than gain-scheduled. This methodology, in turn, was used to generate velocity invariant experimental data for system identification purposes.

Experimental tests were performed, requiring the catamaran to sail on a precise track through the water, in the presence of currents, wind, and waves. Using the identified system models, a high-performance estimator/controller was implemented and tested on

the full-scale prototype. These controllers were indeed quite successful, tracking the line to within 0.3 meters.

SECTION 1.1 MOTIVATIONS

In general, unmanned vehicle systems are useful in their ability to remove humans from dangerous environments, relieve them of tedious tasks, or simply go to locations otherwise inaccessible or inhospitable. Examination of the tasks currently associated with unmanned aerial vehicles (UAV) demonstrates this explicitly. With the use of wind propulsion, an autonomous sailboat is the marine equivalent of a UAV with an infinite loiter time. Many tasks that cannot currently be performed with UAVs due to either payload or duration constrains can be accomplished with the Atlantis. With electric power generated via a wind generator or solar cells, and propulsion via the wind, the Atlantis does not burn any fuel to stay on station or move from one place to another. Thus the autonomous sailboat can be used as an observation platform for weather- or environmental-monitoring, or as a surveillance platform for port observation.

With the ability to station keep (or remain over one specific site), the Atlantis could replace deep ocean buoys which normally drift up to distances of a mile on their long mooring lines. Also, without the need to anchor the Atlantis as a buoy, damage to the ocean floor with cable anchors would be avoided. Because of the autonomous navigation and propulsion capabilities, Atlantis-type buoys could be repositioned rapidly and inexpensively.

A secondary benefit of the work on wind propulsion is increased efficiency. Though initially conceived to solve the problem associated with large sail forces and actuators, the wingsail can increase the sailing efficiency of almost any sailboat. Indeed, sail-

assisted propulsion in general can greatly reduce the fuel required to move from one place to another. Experimental trials conducted on the sugar freighter *Lace* in 1983 demonstrated fuel reduction of 34% over normal (diesel) propulsion [24]. This was accomplished using a conventional cat-claw sail that was not particularly efficient (see Figure 1-10). Further research studies have shown that fishing vessels under true operational conditions saved over 50% of their fuel costs, but were difficult to use [21]. The Bergeson study, published in 1985, theorizes that greater fuel savings could have been realized had the sails been easier to use [21]. This is due to the already high workload found on most commercial fishing vessels [22].

With a self-trimming wingsail, most of the difficulty of using sails disappears. Additionally, the efficiency should be much greater than conventional sails. In a recent study, a Dutch company proposed to build a wingsailed oil tanker with five over the top tacking wings (they are currently seeking funding). This design was projected to reduce fuel consumption under full load by more than 50% (Figure 1-17) [109].

SECTION 1.2 WINGSAIL HISTORY

The concept of using a wing upon a sailboat has been around almost as long as aircraft themselves. The history of wingsailed craft is relevant to this body of work as it demonstrates how the problem of wind propulsion was solved in the past. Many previous designers have come to the false conclusion that adequate lift coefficient could only be achieved with an asymmetric (cambered) wing. This led them to designs that tack by flipping up over the top of the mast, often leaving the wing inclined at 45 degrees (the flip does not rotate the wing a full 180 degrees). The disadvantage of this arrangement is that the weight associated with the flipping mechanism is usually large enough to negate any increased propulsive efficiency by virtue of increase hull drag. Also, with the inclined

wing designs only the vertical projection of the wing acts to propel the boat. This again results in a loss of propulsive efficiency. The history of both symmetrical and asymmetric designs is presented here, as well as advances in the areas of sail-assisted propulsion. Though this section is rather detailed, it shows the advancement of the wingsail from the early 1920's to almost present day. During this span of time, almost every conceivable configuration has been built and tested. By studying the failures and successes of such a long history, advantageous choices can be made in selecting the exact type of wingsail for this specific application.

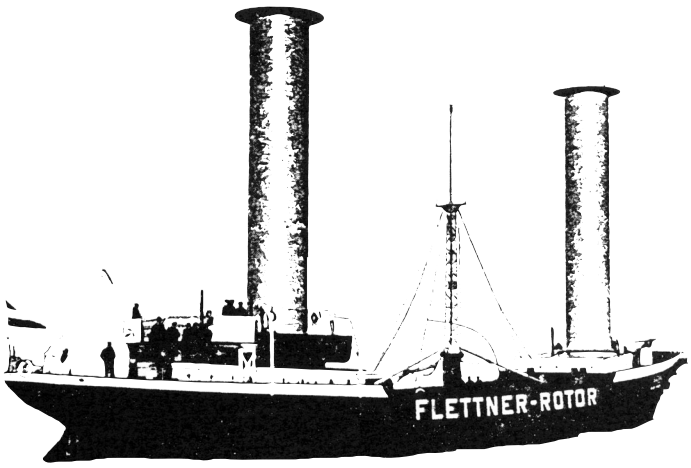

Figure 1-2 The *Baden Baden*, outfitted with two Magnus rotors in 1926 by Anton Flettner. The rotors are spun using a motor, and act as lifting sections due to the Coanda effect. She sailed across the Atlantic in 1927, from Germany to New York City.

The earliest proposal for using a wing on a sailboat was in 1922 by Anton Flettner [49]. After achieving renown with his servo-tab actuated rudder, Flettner proposed replacing the sails on a sailing ship with a self-trimming wing, very similar in design to the wingsail on the Atlantis. During the course of investigating the performance of the proposed wing, Flettner became enamoured with the Magnus rotor (a cylindrical wing based on powered lift via forced rotation [23]) and abandoned the wingsail idea. In 1926, Flettner outfitted the *Baden Baden*, a medium sized sailing vessel (Figure 1-2), with two mag-

nus rotors and sailed across the Atlantic ocean to New York. With the advent of the Great Depression, Flettner did not pursue the idea further. Later wartime pursuits led Flettner to great fame in the area of helicopter research.

Following Flettner, the next independent effort was undertaken by Utne in Norway with the construction of the *Flaunder* [1]. The *Flaunder* was a one-man sailing kayak with a self-trimming wingsail of conventional layout (Figure 1-3). Utne remarked on the sailing qualities of the wingsail, most of which remain evident as characteristics of the Atlantis:

> I had a lot of fun sailing *Flaunder.* She behaved almost like a motor boat and could go backwards magnificently. She could also sail about 20 degrees from the wind with adequate steering speed. I also made a rather important discovery: sudden gusts of wind which would heel neighboring sailing boats almost to the gunwale were in most cases no increase in wind speed but only changes of wind direction. In such cases, the sail of the *Flaunder* turned to a new position with hardly any change in the boat's heeling [1].

Unfortunately, with the outbreak of World War II and the Nazi occupation of Norway, the *Flaunder* was destroyed as a potential war machine. No further progress was noted.

In the United States, Carl in 1951 patented the concept of a wing to replace the conventional sail [31]. Initially he designed a rather complicated mechanical system to raise (or step) the wing, but discovered that it was easier to leave the wing stepped but free to rotate. The Carl design used an aluminum wing that was trimmed via a series of conventional ropes, winches, and pulleys. It took no advantage of aerodynamic surfaces for trim, instead literally replaced the sail with an aluminum wing of similar proportions (Figure 1-4).

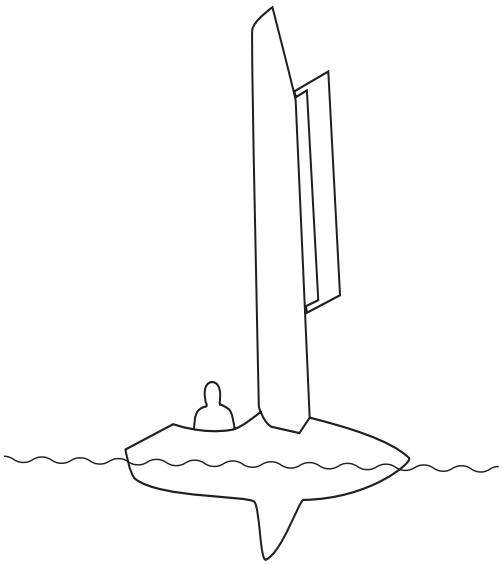

Figure 1-3 The *Flaunder*, built by Utne in Norway in 1941. Utne modified a sailing kayak to use a self-trimming rotating wing sail. The section is an RAF section, and characteristics were not particularly suited towards a wing sail. The *Flaunder* sailed very well, demonstrating the ability to have the wing sail respond independently from the hull in gusts. She was destroyed by the Nazis as a possible machine of war.

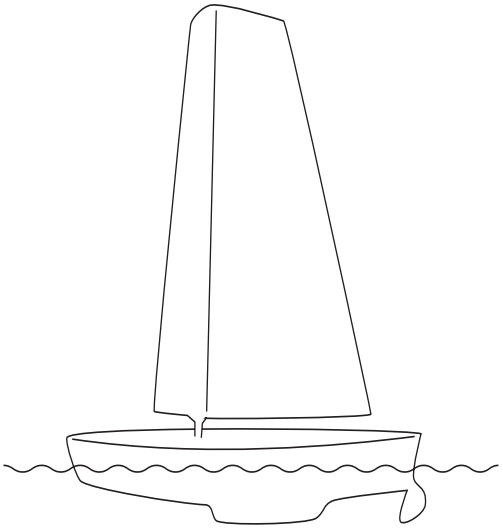

Figure 1-4 The aluminum wing sail patented by Carl in 1951. Rather than use the aerodynamics of the sail for trim, the Carl wing replaces a cloth sail of similar size and shape. It uses a conventional set of winches and pulleys to trim the wing. In spite of this, Carl remarked on how high into the wind he was able to make headway.

In England, an aerospace engineer named John Walker designed and built a series of wingsailed trimarans culminating with the *PlaneSail* in 1968 (see [81], [89], [111],

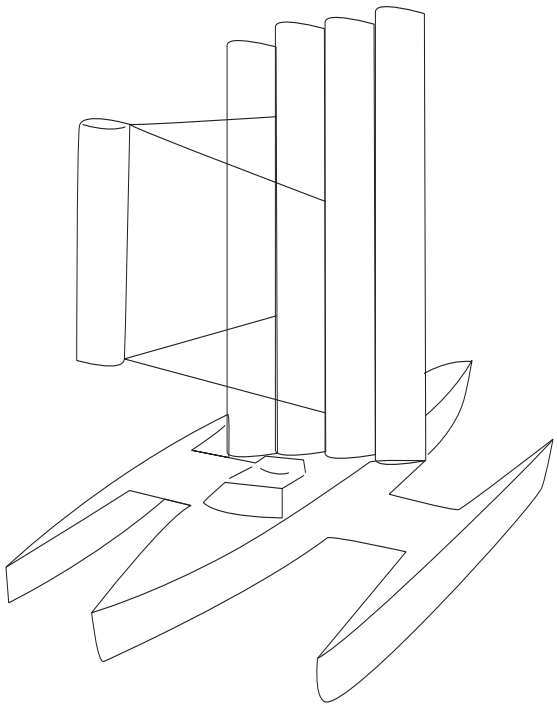

Figure 1-5 The trimaran *PlaneSail*, designed and built by aeronautical engineer John Walker in 1968. She uses a venetian blind arrangement of four vertical wings with a trailing tail. This is a self-trimming set-up very similar to the *Flaunder*. The wings are NACA 0012 section, which perform poorly in this regime. The Planesail did, however, change the way wings on sailboats were perceived.

[139], and [140]). This trimaran used a series of four NACA 0012 sections arranged as a quadra-plane with a trailing tail for self-trimming (see Figure 1-5). Walker continued to refine his designs, with another trimaran the *Flyer*, in 1972, using a single multi-element wing with both leading and trailing edge devices (Figure 1-6). In 1990, with much fanfare, he sailed the *Blue Nova* (Figure 1-7) across the Atlantic. During the crossing, the *Blue Nova* passed through hurricane *Candice*, proving that the wingsail could be quite robust and did not need to be reefed during adverse weather. This was yet another self-trimming trimaran with a simple biplane wing and a trailing tail. [The *Blue Nova* is currently being

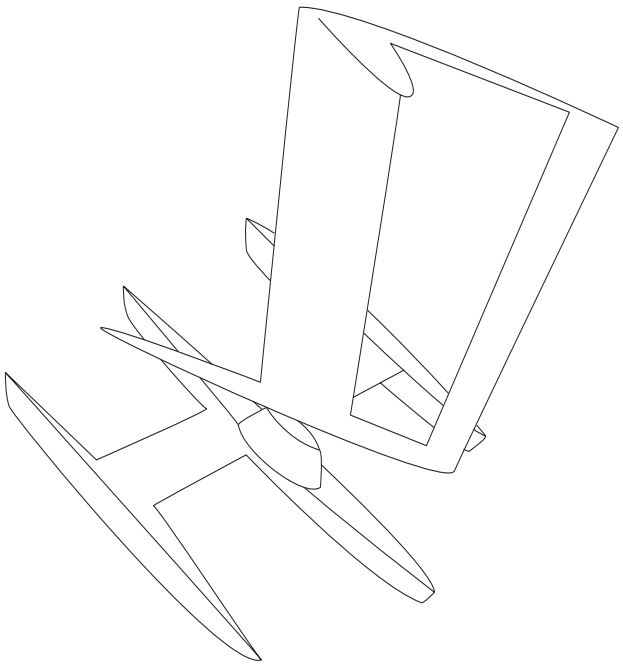

Figure 1-6 The trimaran *Flyer*, built by John Walker in 1972. A continuation of the work that began with the *Planesail*, the *Flyer* uses a single wing and tail. The wing section was a proprietary design, using both leading edge slats and trailing edge flaps. The forward boom included a counterweight that brought the mass center of the wing in line with the lift force.

outfitted to circumnavigate the globe with a crew of disabled sailors.] Based on the success of the *Blue Nova* crossing, Walker went into business marketing the *Zephyr*, a 42' self-trimming trimaran with a multi-element single sail and trailing tail (Figure 1-8). In 1996, the first production *Zephyr*, named the *Ocean Tern*, was sold. It was also the last one to be built. Walker Wingsail Ltd. declared bankruptcy in 1998 and disappeared from the public eye.

Walker published many papers championing the cause of the wingsail for various commercial vessels, see [139], [140], and [141]. Unfortunately, his papers did not detail any of the technical aspects of the wingsail or substantiate any claims of performance. Thus, thirty years of experience and development of this concept has been lost.

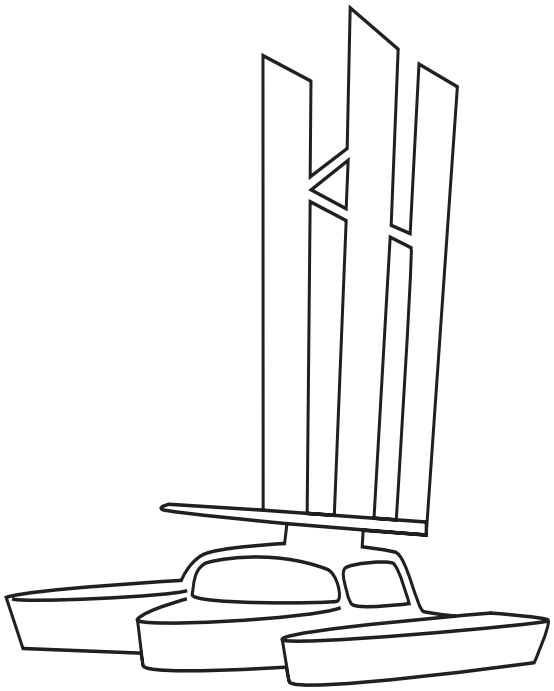

Figure 1-7 The trimaran *Blue Nova* built by John Walker in 1990. The wing sail was a biplane arrangement with a trailing tail, similar to *Planesail*. The *Blue Nova* was sailed across the Atlantic and during the crossing passed through the hurricane *Candice*, proving that the wing sail could be quite robust and did not need to be reefed during adverse weather. The *Blue Nova* has resurfaced as the sailing vessel for the first group of handicapped sailors to attempt a circumnavigation.

Perhaps the seminal event of the wingsail world occurred in 1972, when the first wingsailed catamaran, *Ms. Nylex*, won the "little America's Cup" for Class "C" catamarans. Up until this point in time, wingsails had remained an oddity in the sailing world. The success of *Ms. Nylex* demonstrated to the sailing world unequivocally that a wing really was more efficient than a conventional sail. *Ms. Nylex* replaced the sail with a simple rigid wing, trimming it in a conventional manner with ropes and winches (Figure 1-9).

In 1979, Bergeson [22] published a report for the Merchant Marine stating that several proposed wind assisted propulsion schemes were viable and would result in a reduction of fuel costs anywhere from 34% to greater than 50%. Following this report, a small sugar freighter, the *Lace*, was outfitted with a simple cat-claw sail (a triangular cloth sail, pictured in Figure 1-10). The actual operating log shows a reduction of 36% in fuel

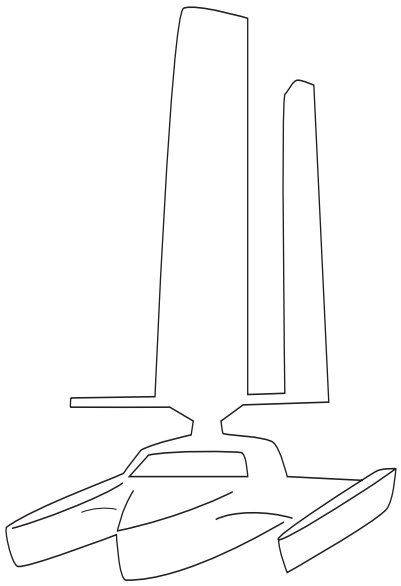

Figure 1-8 The trimaran *Zephyr*, built by Walker Wingsail Ltd. in 1996. This was the production version of the *Blue Nova*, and was intended to revolutionize the yacht market. The first and last one produced was the *Ocean Tern*, which today is harbored in San Diego, California. Walker Wingsail Ltd. went into bankruptcy in 1998 and the jigs and tooling for the *Zephyr* were auctioned off by creditors.

use, with an accompanying increase in transit speed. Furthermore, when the main engine of the *Lace* failed 3 days outside of port, she was able to come in under sail alone at a reduced speed, obviating the need for emergency tow insurance.

In 1981 the Japanese ship, *Shin Aitoko Maru,* a small chemical carrier (Figure 1-11), was launched with two steel cam-shell doors that opened up to form two very inefficient low-aspect ratio square sails [143]. Originally, an entire fleet of these ships was to be built, but by the mid-1980's the oil crisis was over and the price of diesel had dropped to the point that sail assisted propulsion fell out of favor. Indeed, most of the research on sails and wings for sail assisted propulsion died out by 1986.

In 1983, Fekete and his group at McGill University in Canada, analyzed and built a conventional self-trimming wingsail on a canoe [48] (Figure 1-12). They actually went to the effort of measuring the performance of the wing, which was based on a NACA 0012

Figure 1-9 *Ms. Nylex*, winged Class "C" catamaran, built in 1972. She raced and won the "Little America's Cup" in 1972. The wing has a simple NACA 0015 section, but with a large 60% trailing edge flap. The wing was trimmed conventionally, with the usual assortment of winches, pulleys, and lines. The ability to trim the top and bottom of the wing separately, effectively managing the twist of the wing, was considered crucial to the victory.

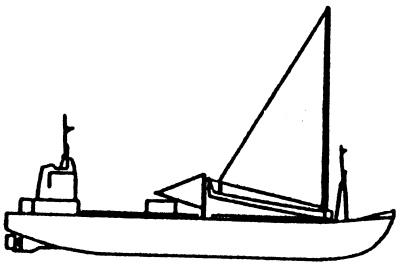

Figure 1-10 The *Lace*, a small sugar freighter, was outfitted with a simple cat-claw sail in 1980. This simple triangular cloth sail was controlled via hydraulic winches controlling wire rope attached to the boom. The *Lace's* log showed actual fuel savings of over 34% on routes that were not particularly advantageous to wind propulsion.

airfoil section. They experienced great difficulty in measuring the baseline performance of the wing due to the unsteadiness of the wind but eventually generated some nominal performance figures. After correcting for differences in hull length and sail area, they showed that the performance of the wing sailed canoe was superior to the performance of a con-

Figure 1-11 The chemical carrier, *Shin Aitoko Maru*, was built in Japan and launched in 1981. The Japanese fleet had several wind assisted freighters during the early 1980's. The *Shin Aitoko Maru* used two metal clamshell doors that opened up and produced a sail with an effective aspect ratio of approximately 0.7. This led to very poor performance on all points of sail except those directly downwind.

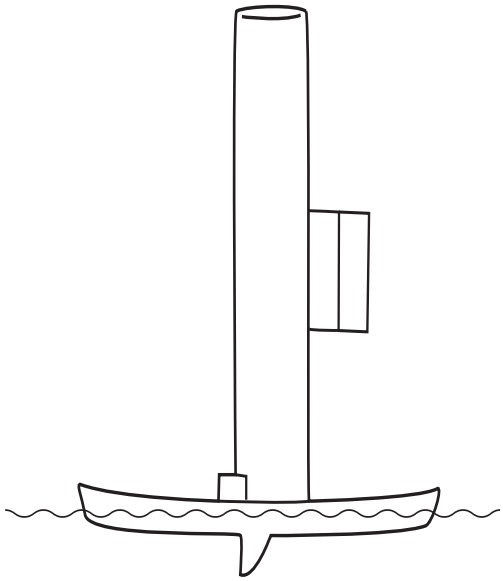

Figure 1-12 The Fekete group at McGill University analyzed and build this wing sailed canoe in 1983. They used a NACA 0012 section which is very poor for the Reynolds number range in question. They painstakingly measured the performance of the wing and concluded that it outperformed the sail on a Laser on all points of sail except straight downwind. They eventually added a wing extension to increase both the aspect ratio and the total wing sail area.

ventionally sailed Laser. Their effort was the first formal analysis of the wing sail as a method of wind propulsion. They failed, however, to take into account the Reynold's number effects of the operational airfoil section, thus experiencing performance well below what could be achieved utilizing a better airfoil design.

In 1985, the Cousteau society began to explore the possibilities of wind assisted propulsion to reduce the fuel costs (and environmental impact) of their ocean exploration vessels. Though they considered using wingsails, they quickly discarded the idea in favor of the "more efficient" Magnus rotor. The claimed efficiency of the Magnus rotor stemmed from a coefficient of lift (C_L) of 12 or more with end plates, which came directly from the powered nature of the lift generated. Once the correction for the power injected by spinning the cylinder was factored in, then the Magnus rotor was no longer very efficient compared to a wing (though it was still much better than a sail). In the end, after much research based on the aspiration of boundary layer as opposed to spinning the cylinder [33], the Cousteau society built the *Calypso II* with twin turbosails—elongated cardioid shaped cylinders with a moving flap and the boundary layer sucked off via a vacuum pump (see Figure 1-13).

Figure 1-13 The Cousteau Society turbosailed vessel *Calypso II* uses a modern version of the Magnus rotor. Instead of spinning the cylinders to provide lift, turbosails suck the boundary layer off with vacuum pumps through a porous surface. In addition, a movable flap is used to effectively "tack" the turbosails.

Perhaps the most famous winged boat in the popular press was the *Stars & Stripes*, a 60' catamaran with a fixed wing the semi-span of a Boeing 747. This wing was designed

by John Roncz and built by Burt Rutan. John Roncz and Burt Rutan have collaborated extensively on wing design and construction of such famous projects as the non-stop around the world aircraft, *Voyager*. The boat was built to answer a poorly framed challenge from the New Zealanders for the America's Cup in 1988. The *Stars & Stripes*, with Dennis Connor as the skipper, easily won the races against a conventionally sailed monohull of gargantuan proportions (in excess of 100' at the waterline). Anecdotal evidence was that neither Connor nor his crew was comfortable sailing the *Stars & Stripes*, and all were in fact terrified of capsizing. In addition, since the wing was not self-trimming, constant adjustments to the various flaps and leading edge slots were required. Due to the unsteadiness of the wind, it is unlikely that they were able to hold the maximum coefficient of lift without stall for any length of time.

Figure 1-14 The wingsailed catamaran *Stars and Stripes*. This catamaran was built to answer a poorly framed challenge for the America's Cup race in 1988. The wing was designed by Jon Roncz and built by Burt Rutan. The *Stars and Stripes* won the 1988 America's Cup easily, but both captain and crew were never entirely comfortable sailing her. The wing was fixed as opposed to self-trimming, and was very mechanically complicated.

Ross, et. al, in [118] analyzed, designed, and built a high performance wing for a Class “C” catamaran. The analysis for this project was very formal, with the correct Reynold's number, and using the best available (at the time) computation fluid dynamics (CFD) codes. His group, however, decided upon a complex, three element airfoil section, with a moveable leading edge slot, main section, and a moveable double slotted flap. Although this allowed a coefficient of lift (C_L) of over 3.0, it also required an asymmetric section. Thus, the Ross design required an inclined wing which was tacked over the top by turning the entire wing parallel to the horizon and switching the top and bottom ends on each tack. Because the mechanisms required to accomplish this are very complex and correspondingly heavy, any advantage gained in propulsion was lost in hull drag due to the increased weight.

In 1996, Cornell University students embarked on the RAFT project [34]. Forty-

Figure 1-15 Cornell RAFT project was an undergraduate effort to build a high performance wing for a catamaran. The wing uses carbon fiber ribs with a mylar covering to achieve light weight. The wing is asymmetrical, with a leading edge slot and a trailing edge flap. The wing is tacked "over-the-top" by pivoting the aft section at the vertical center behind the leading edge slat section.

eight students and three professors were to build an over the top tacking wing for a light catamaran (Figure 1-15). The initial design used a two-element section with a leading edge slot and trailing main section. The idea was to pivot the main section about the horizontal axis but leave the leading slat fixed. Though no technical papers have been published, the web site indicated that they eventually completed construction and preliminary testing. The web site was unclear about the final form of the wing. Inquiries to the professors and staff involved have remained unanswered.

In 1999, Boatek in England began manufacturing a self-trimming, asymmetric, over-the-top tacking wingsail that could be hoisted up a conventional mast [26], and

flipped over for tacking or jibing (Figure 1-16). The key to his design was an underwater fin keel that moved in conjunction with the wingsail to counteract heeling moments.

Figure 1-16 The Boatek wing fitted to a day-sailing catamaran. The wing is hoisted up a conventional mast and is asymmetric. The asymmetric design is intended to have a higher lift coefficient than a symmetrical wing. It is tacked “over-the-top” and has a corresponding underwater fin that counteracts the healing moment induced by the wing. Several designs have emerged from this designer, including something akin to a wind surfer, and several larger catamarans.

Several other wingsail proposals have surfaced within the past few years, including a solar-electric wingsail ferry in Australia that would use solar panel covered wings for wind assisted electric propulsion, an oil tanker with six large vertical wings and a 50' deep swing keel, with each wing split into four vertical sections each rotating about the mast in an over the top fashion (Figure 1-17), and an 80' foil-borne trimaran with a 120' self-trimming wing with a cloth jib on the front [104] (Figure 1-18). Clearly, the wingsail is an idea that refuses to die.

SECTION 1.3 RESEARCH FOR THIS THESIS

The research for this thesis is broken down into four distinct areas, as follows:

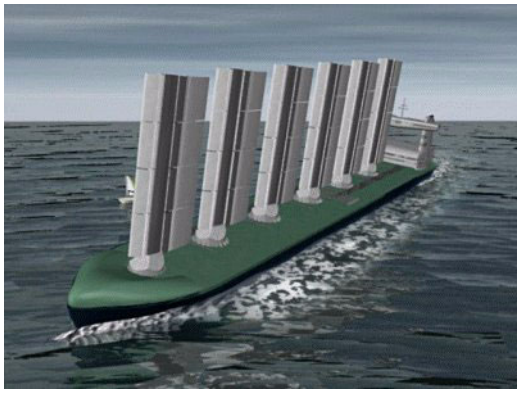

Figure 1-17 The Dutch Pelmatic wingsailed oil tanker proposal. This oil tanker is to use five independent wingsails that are tacked by splitting each wing into four vertical sections that independently rotate about the mast in an over the top fashion. The design includes two 50' deep swing keels to make sailing efficient. The wingsails are expected to reduce fuel consumption by over 50%.

Figure 1-18 An artist rendition of the *Volantis*, an 80 feet long foil bourne trimaran with a rigid self-trimming wing sail that includes a cloth jib. It is intended to set a new Pacific crossing speed record, and also to be used to tow microphone arrays in an autonomous mode. This is a commercial project with preliminary funding. They have not yet built a prototype.

a) *System Architecture*. Utilizes a network of sensors and actuators to achieve a robust control architecture. Much of this work is in the area of low cost sensors and actuators, particularly with significant attention paid to calibration issues.

b) *Attitude System.* This includes a new formulation for the solution of Wahba's problem [144] of determining attitude from vector matching as well as implementation and testing of this algorithm using low cost sensors. Novel methods for calibrating scale-factor and bias errors on a low-cost three-axis magnetometer are developed and experimentally demonstrated.

c) *Wind Propulsion System.* An optimized wing airfoil section, configuration analysis, structural analysis, and implementation of the wingsail propulsion are developed.

d) *System Identification*. The exploration of system identification methods and robust control structures that include velocity invariance are explored in order to demonstrate experimental results of autonomous control.

SECTION 1.4 PRIOR ART

Two prior attempts at autonomous sailboats have been recorded in the literature. The first is the **S**tation **K**eeping **A**utonomous **M**obile **P**latform (SKAMP) (Figure 1-19) developed in 1968 by the RCA Astro-Electronics division [130], [121]. The SKAMP utilized a curving ring-shaped wing with NACA 0030 airfoil sections and a semi-submerged elastomeric hull. The projected top speed of the SKAMP was 4 knots and the position accuracy was estimated to be 0.2 nautical miles (roughly 300 meters). The main position-

ing system used was the Navy's navigation satellite system (known as TRANSIT), yielding a position fix every 90 minutes. The SKAMP included 360 degree roll stability and was intended for covert surveillance operations. Once the project was publicized, however, the Department of Defense pulled the funding and the project died. Since the paucity of published data included no actual sailing data it remains unclear if the SKAMP ever sailed. Pictures of the SKAMP at Long Beach Harbor showed the boat on dry land, but never sailing.

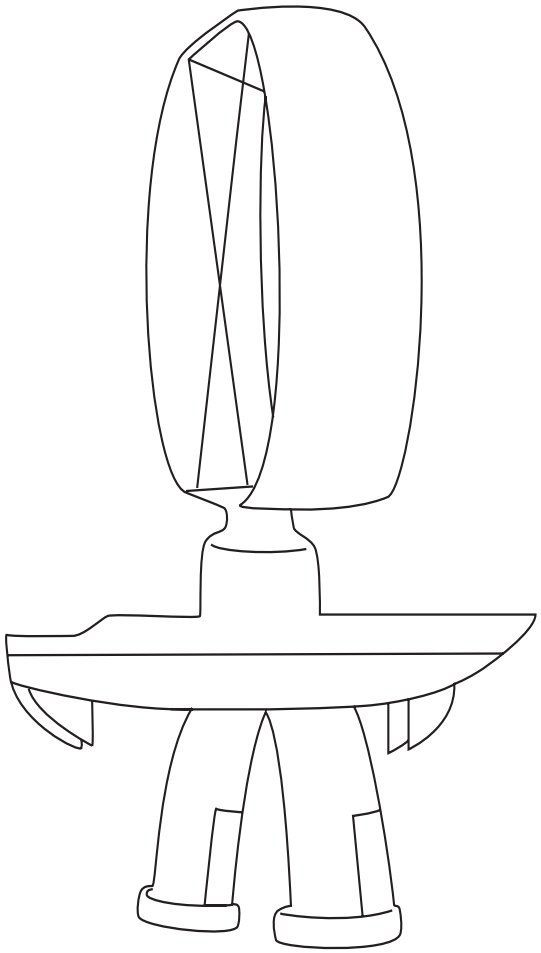

Figure 1-19 The RCA Astro-electronics division SKAMP project, developed in 1968. The SKAMP was a mobile surveillance platform that was wind propelled. It had several innovations such as a semi-submerged elastomeric hull, a curved ring wing with NACA 0030 sections, and weighted keels that gave it 360 degree roll stability.

The only other published autonomous sailboat attempt is the *Relationship* project (Figure 1-20) out of the Technical University of Furtwangen, Germany [137]. This project started in 1995, with the intention of performing a zero-handed circumnavigation (of Earth) using a conventional sail. The entire university participated in the project, with 168 students, 12 faculty, 10 outside consultants, and a budget of just under two million dollars. They completed the construction of their balsa cored composite hulls in 1996. Initially, the intention was to sail autonomously; however, after much trouble with a shakedown cruise to Lisbon, the project changed to a remote control mode through a satellite link. Regulatory difficulties with respect to sailing a boat that cannot, by definition, be aware of its surroundings led them to flag the *Relationship* as a "rudder-disabled" vessel. This, however, has not satisfied the maritime authorities, and it is currently moored in the Azores, not having moved in over 18 months.

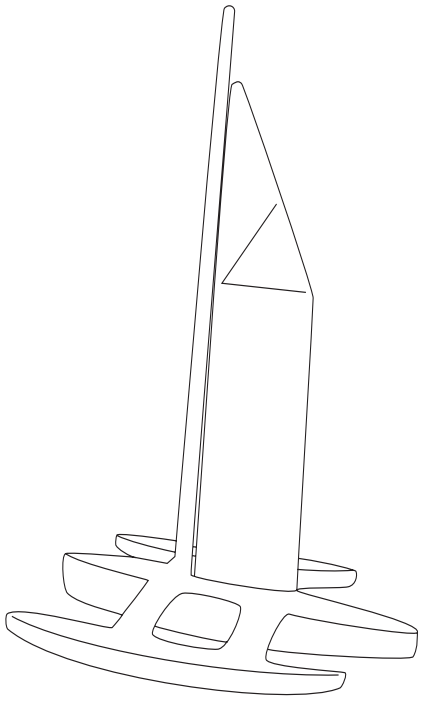

Figure 1-20 The *Relationship* from the University of Furtwangen, Germany. This was an attempt at a zero-handed circumnavigation. A conventionally sailed trimaran, the sail and rudder controls were first intended to be autonomous. After several setbacks, the *Relationship* was controlled remotely from Furtwangen, and currently remains stationary in the Azores.

SECTION 1.5 THESIS CONTRIBUTIONS

This thesis represents the sum total of all of the different contributions in the areas of structures, sensors, actuators, aerodynamics, guidance, and control that were required to bring the project to fruition. As a systems thesis, it represents significant contributions in all three disciplines within the Aeronautics and Astronautics department: Structures, Fluids, and Guidance Navigation Control (GNC). The main contributions are:

- Conception, design, construction, and experimental demonstration of an autonomous sailboat capable of precision control to better than 0.3 meters.
- Development of the methodology to identify robust plant models and controllers that are invariant under velocity changes.
- Description of an optimization scheme for symmetric wingsail section based on requirements unique to sailing vehicles.
- Development and experimental demonstration of a novel quaternion based attitude estimation algorithm from vector observations.*
- Development and experimental demonstration of a novel method for calibrating any three-axis sensor that requires no external reference.†

SECTION 1.6 THESIS STRUCTURE

This thesis is laid out as descriptions of the various steps required to build Atlantis. For the sake of brevity and readability, some detail is omitted. The general scope of the project should, however, remain clear.

*. Contribution in conjunction with Dr. Demoz Gebre-Egziagbher

†. Contribution in conjunction with Dr. Demoz Gebre-Egziagbher

Chapter 1 contains the introduction, history of wingsailed vessels, prior art, and general overview of this thesis.

Chapter 2 presents the system layout, details of the communications, and a general introduction to the various sensors and actuators that make Atlantis an autonomous vehicle.

Chapter 3 describes the sensors and actuators and the physics behind the measurements. This chapter also includes the calibration of the sensors and the actuators, except for the magnetometers, which are detailed later.

Chapter 4 covers the attitude system, including error analysis, and magnetometer calibration. Performance for simulated data and flight test data from an instrumented aircraft (Queenair) testbed are shown.

Chapter 5 details the propulsion system. Included is configuration analysis, airfoil requirements and design, flap/chord optimization, and a wing structural analysis.

Chapter 6 describes the system identification methodology and control design. It describes the observer Kalman filter identification (OKID) methodology and its application to this thesis, as well as methods for incorporating velocity invariance into the identification and control design.

Chapter 7 details the results for both the trolling motor experimental runs as well as the data for the sailing experiments performed.

Chapter 8 states the conclusions of this work along with recommendations for further areas of research.

2 System Description

In order to validate the concepts presented in this thesis, a prototype system has been built for experimental confirmation. The prototype is based on a heavily modified Prindle-19 day-sailing catamaran. The Prindle is better suited to automatic control due to the centerboards and symmetric hulls used in its design.

The catamaran is 7.2 meters long, 3 meters wide, and was originally equipped with a sloop rig sail with 17 square meters of sail area. The sloop rig has been replaced by a rigid self-trimming wingsail with a total wing area of 7.65 square meters. Directional control is effected by rudders at the end of each hull and two

retractable centerboards approximately ½ meter behind the main cross beam.

Several sensors and actuators are installed on the hulls and the entire sailing system (mast, boom, main and jib sails) has been replaced with a vertical self-trimming wing (wingsail) suspended by spherical roller bearings on a two meter aluminum stub-mast (Figure 2-1). The wingsail design and analysis are found in Chapter 5.

This chapter covers the sensors, actuators, and other systems required to achieve autonomous control. Note that Chapter 3 details the calibration and linearization of the sensors and actuators, while this chapter includes only an overview and description.

The prototype system, called the Atlantis, is based on a high-speed network of many sensors, actuators, and microcontrollers. What follows is a brief description of the system operation in order to place the more detailed descriptions of sensors and actuators in context. A photograph of the Atlantis appears in Figure 2-1, showing the wing and tail, attitude system, Guidance-Navigation-Control (GNC) computer, and rudder actuator.

The attitude system (detailed in Chapter 4)—a three-axis magnetometer, two-axis accelerometer, and Siemens 515 microcontroller—is mounted alongside the Global Positioning System (GPS) antenna. The electronics that make up the attitude system are contained inside a waterproof Pelican case on a wooden crossbeam at the forward stay location. The Pelican case ensures that the electronics function in the heavy spray environment at the front of the catamaran. The wooden crossbeam is added for increased structural rigidity of the hulls, counteracting stresses induced by the wing.

The GNC computer, a Pentium class laptop, is placed inside another waterproof case, along with the Trimble Ag122 GPS receiver. The GNC computer is equipped with a parallel port device (or “dongle”) that allows communication over the Controller Area

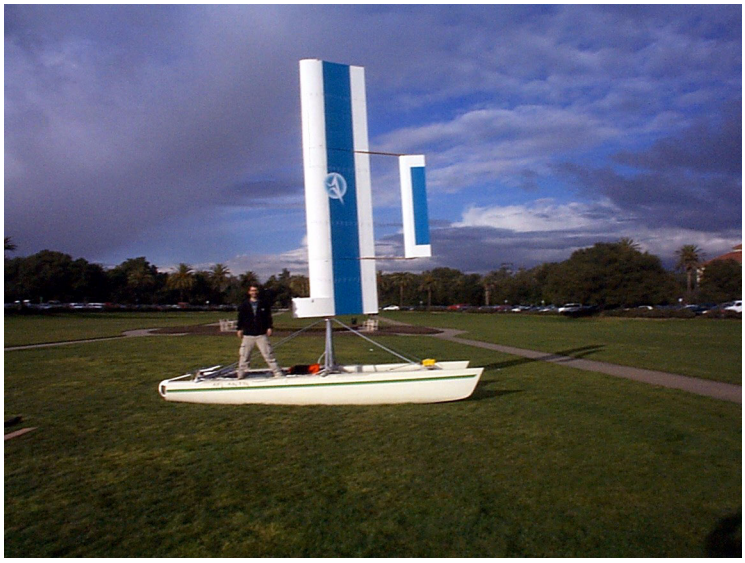

Figure 2-1 The Atlantis, assembled on the Stanford Oval. The hulls and crossbeams are originally from a Prindle-19 day-sailing catamaran. The author is included in the picture to give a sense of scale. The wing is free to pivot on bearings and is trimmed by aerodynamic loads on the fully flying tail.

Network (CAN) bus. A DC/DC power converter insures that the laptop draws current from the boat power bus rather than its own internal batteries.

Inside the starboard hull, underneath the rear inspection cover, are two Siemens 505 microcontrollers, one for the hullspeed and rudder angle sensors, the other for the rudder actuator. There is a Standard Marine Electronics through-hull speed sensor that pierces the bottom of the starboard hull. The rudder angle sensor is a magnetic flux sensor (called a “LoHet”) between two magnets on the upper rudder hinge. The rudder actuator is a fractional horsepower DC motor, with a lead screw assembly, constrained to one plane of rotation. The actuator is commanded by the microcontroller through an Infineon H-bridge mosfet drive using pulse width modulation (PWM) to control the direct battery current.

Inside the stub-mast, a Mercotac slip ring allows 360 degree rotation without twisting the four wires that connect all of the components of the Atlantis. The slip ring is

located just above the bearings at the top of the lower wing section. The wing itself is built in three sections that are assembled on site (Figure 2-2).

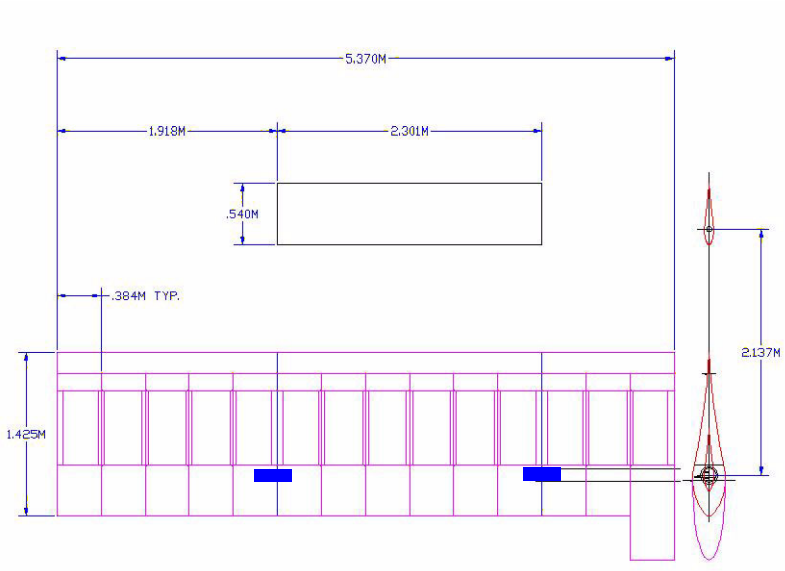

Figure 2-2 An engineering layout of the wingsail. This drawing shows the dimensions of the wing, the three sections, and the rib layout pattern. On the right side it shows the overhead view, including the electronics pod at the front of the wingsail. The aluminum tongues that joint the sections together are represented by the blue solid rectangles, and are joined to the reinforced spar caps with 5/16” stainless steel bolts.

The lower section contains an electronics pod with the batteries, ballast, battery-charging electronics, and, as previously noted, the slip rings. Each wing section contains a microcontroller and DC motor to control the trailing edge flap of that section. The lower section also contains the anemometer microcontroller, a Standard Marine Electronics anemometer transducer head attached to the top of the electronics pod lid performs double duty as both wing speed sensor and an angle of attack sensor for the wing. The lower section has an additional Honeywell three-axis magnetometer to sense the angle of the wing relative to the boat centerline in addition to the anemometer controller.

The middle section of the wing contains the microcontroller and drive electronics for the tail in addition to the microcontroller and drive motor for the trailing edge flap. The

upper section contains only the microcontroller and DC motor necessary to control the upper trailing edge flap. The top two sections of the wing contain flotation balloons inside the structure of the wing to prevent the Atlantis from “turtling” in the event of a capsize (turtling or turning turtle is defined as coming to rest on the water with the wing pointed straight down). Additionally, the three wing sections are wired with coaxial cable to allow the GPS antenna to be mounted at the top of the mast.

Each section of the wing is covered in a polyester skin that is intended to be water-proof. Furthermore, each section was sealed as best as possible to provide independent flotation in the event of a capsize. Depending on the speed of rotation at impact, the wing might not have the structural strength to survive the impact without damage. Fortunately, this design feature was never tested under operational circumstances.

SECTION 2.1 ATLANTIS OPERATION

Assembly: The three sections of the Atlantis’ wing are bolted together with aluminum tongues at the spar caps. The current method of assembling the wing onto the boat is to first join the top two sections and the tail, tightening the four 5/16” stainless steel bolts. The entire lower section of the wing remains permanently attached to the stub-mast and hulls. If a crane is available, the top two sections are hoisted above the lower section, and the wing is fitted together and bolted into place. Lacking a crane, the entire boat (hulls, stub-mast, spider and lower section) is tipped over onto the port side, and the wing joined horizontally. Then the entire boat and wing are returned to its normal, upright position. This usually involves removing the boat from the trailer, assembling it, and lifting it back onto the trailer after this has been done.

Operation: Once the Atlantis has been safely placed in the water, the GNC computer box is attached to the CAN network and the RF (GPS antenna) cable. An external keyboard is used to send commands to the GNC computer. At this point, the main power breaker is turned on, allowing 12V power to reach all of the components. With the GNC computer removed from its watertight case, a system check is performed to verify that all systems are functional and that the GPS receiver is generating position and velocity fixes.

The GNC computer is then placed into its watertight box. Up to this point, the wing is pointing into the wind and the wingsail is generating no thrust. In order to begin sailing, the tail is commanded to an angle of attack (the trailing edge of the tail is pointed in the direction of desired travel). As the boat accelerates, and the wing will pivot into the relative wind, pointing more towards the prow of the catamaran. Steering is accomplished through the rudders which are controlled via keyboard commands to the GNC computer.

If the path desired includes a tack or a jibe, the tail is centered during the maneuver when the centerline of the boat is aligned with the wind. The position of the tail is then reversed from its previous position to complete the tack or jibe. Full details of the mechanics of wingsailing are given in Section 7.1, “tacking, jibing, and self-trimming,” on page 272.

Once the desired starting point can be reached through a straight-line path (no tacks or jibes required), then the automatic control is engaged. A course is then assigned to the catamaran; sail from the coordinates “A” to the coordinates “B.” The Atlantis will at this point sail itself towards the first waypoint. Upon reaching the waypoint, the GNC computer will command a line-following controller to track a straight line towards the next waypoint. The control on this straight line segment minimizes cross-track error, compensating for wind, current, and sensor biases.

SECTION 2.2 OVERALL SYSTEM ARCHITECTURE

One of the innovations of the Atlantis project is the method of connecting all of the sensors and actuators with a high-speed serial network. This is true for all sensors and actuators except the GPS receiver, which communicates with the GNC computer directly through a serial bus. The high-speed serial network used for this project is the Controller Area Network (CAN) bus, originally designed by Bosch for car electronics. While modern cars use networked sensors, very few (if any) of the control functions rely on the network for critical message passing. Neither do they leverage the presence of the network to calibrate or linearize sensors at the point of the sensor.

Topologically, the CAN network appears to be a shared memory space, i.e., each message object has a unique identity, and is available to all nodes on the network. Several advantages over a conventional "star" topology can be gained by using a networked system. In the traditional system, a central computer communicates to all the sensors and actuators via an Analog-to-Digital/Digital-to-Analog (A/D-D/A) card. Unless the system is very small, long wires with shielding must be used to prevent noise corruption. Isolating sensors with weak signals and reducing cross-talk between sensors and actuators becomes increasingly difficult as the power levels increase in the actuators. Trouble-shooting and debugging grows increasingly difficult due to the large number of wires (sometimes hundreds) necessary to connect the system.

By contrast, the Atlantis uses a four-wire bus to connect all sensors and actuators. The wires are power (+12V) and ground, and CAN_hi and CAN_low, the differential signaling lines used by the CAN network for communications. Since a CAN message is received by all nodes on the network, communication integrity can be verified with minimal effort. With communications established, faults can be isolated to smaller modules quickly, and inexpensive modules can be swapped out for replacements. Also, the ability

to synthesize CAN messages means that any module can be made to appear in a hardware-in-the-loop simulation for testing and trouble-shooting.

Another advantage of the network bus architecture is that any sensor or actuator may be physically moved without requiring any reprogramming or rewiring. It is simply connected to the existing network, and works without further configuration. Lastly, sensors and actuators are linearized, calibrated, and converted to engineering units at the microcontroller next to the sensor or actuator (detailed in Chapter 3). This reduces the workload on the central processor, and in the case of a distributed control system, the architecture negates the need for a central processor completely.

SECTION 2.3 CAN BUS DETAILS

The CAN protocol is an International Standards Organization standard (ISO 11898) for serial data communication. As previously stated, the protocol was developed by Bosch for automotive applications. Today, CAN has gained widespread acceptance and is used in industrial automation as well as automotive and mobile machines. The CAN standard includes a physical layer and a data-link layer which defines a few different message types, arbitration rules for bus access and methods for fault detection and fault confinement. The CAN protocol as defined by ISO 11898 is summarized as follows:

- The physical layer uses differential transmission on a twisted pair wire.
- A non-destructive bit-wise arbitration is used to control access to the bus.
- The messages are small (at most eight data bytes) and are protected by a checksum.

- There is no explicit address in the messages; instead, each message carries a numeric value which controls its priority on the bus and may also serve as an identification of the contents of the message.
- An elaborate error handling scheme is provided that results in retransmitted messages when they are not properly received.
- There are effective means for isolating faults and removing faulty nodes from the bus.

The CAN protocol is a good basis for designing distributed control systems. The CAN arbitration method ensures that each CAN node only deals with messages relevant for that node. A distributed control system can be described as a system where the processor capacity is distributed among all nodes in a system, as opposed to centralized (traditional) control.

The CAN bus is a broadcast type of bus which means that all nodes can “hear” all transmissions. There is no way to send a message to just a specific node. The CAN hardware, however, provides local filtering so that each node may react to only the pertinent messages. The bus uses Non-Return To Zero (NRZ) filtering with bit-stuffing. The modules are connected to the bus in a wired-AND fashion: if just one node is driving the bus to a logical 0, then the entire whole bus will be in that state regardless of the number of nodes transmitting a logical 1.

The message arbitration multiple access protocol is a carrier sense multiple access with non-destructive arbitration (CSMA/NDA) and is an *Aloha* derivative in some ways similar to Ethernet. The message arbitration (the process in which two or more CAN controllers agree on which one is to use the bus) is of great importance in order to fully utilize the available bandwidth for data transmission. Any CAN controller may start a transmission when it has detected bus idle state. This may result in two or more controllers starting

a message (almost) simultaneously. The conflict is easily resolved. The transmitting nodes monitor the bus while they themselves are sending. If a node detects a dominant level when it is sending a recessive level itself, it will immediately quit the arbitration process and become a receiver instead (Note that this is physically implemented by having the dominant level be ground, and the recessive level as + voltage; the dominant transmitter sinks all other nodes trying to send a high level). The arbitration is performed over the whole Arbitration Field and when that field has been sent, exactly one transmitter is left on the bus. This node continues the transmission as if nothing has happened. The other potential transmitters will try to retransmit their messages when the bus next becomes available. No time is lost in the arbitration process (see Figure 2-3).

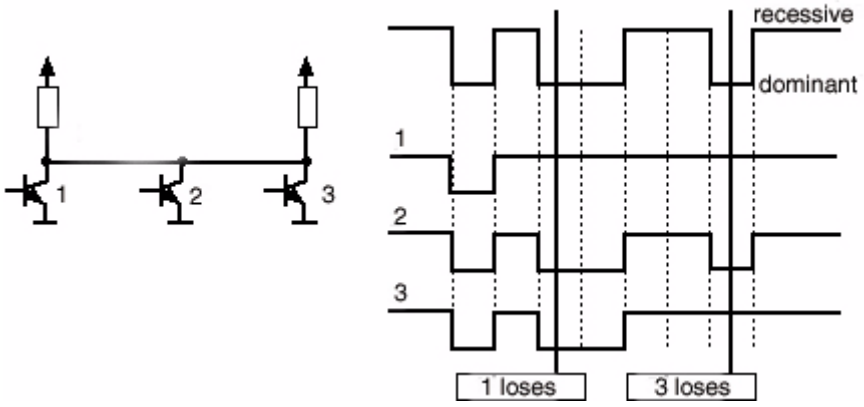

Figure 2-3 Illustration of the CAN non-destructive arbitration. All three nodes begin to transmit on the line. The first bit is dominant (low), and is identical for all three nodes. The second bit recessive (high) for all three nodes. The third bit is recessive for Node 1, which loses arbitration and reverts to receive mode. Nodes 2 and 3 continue to match bits up until bit 7, where Node 3 transmits recessive and Node 2 transmits dominant. Node 3 loses arbitration, reverts to a receiver, and Node 2 takes control of the bus for transmission.

An important condition for this bit-wise arbitration to succeed is that no two nodes may transmit the same Arbitration Field. There is one exception to this rule: if the message contains no data, then any node may transmit that message. Since the bus is wired-AND and a Dominant bit is logically 0, it follows that the message with the numerically lowest Arbitration Field will win the arbitration.

It is worth noting once again that there is no explicit address in the CAN messages. Each CAN controller will pick up all traffic on the bus, and using a combination of hardware filters and software, will determine whether the message is “pertinent”. In fact, there is no notion of message addresses in CAN. Instead, the content of each message is identified by an identifier which is present somewhere in the message. CAN messages are said to be “contents-addressed.”

A conventional message address would read, “message for node X.” A contents-addressed message reads, “data labeled X.” The difference between these two concepts is small but significant. The contents of the arbitration field is, per the ISO standard, used to determine the message priority on the bus. All CAN controllers will also use the arbitration field as a key in the hardware filtration process. Reemphasizing, in most contents-addressed systems the arbitration field defines the contents of the message.

Error handling is built into the CAN protocol and is of great importance for the performance of a CAN system. The error handling aims at detecting errors in messages appearing on the CAN bus so that the transmitter can retransmit an erroneous message. Every CAN controller along a bus will try to detect errors within a message. If an error is found, the discovering node will transmit an error flag, thus destroying the bus traffic. The other nodes will detect the error caused by the error flag (if they have not already detected the original error) and take appropriate action, i.e., discard the current message. Each node maintains two error counters: the transmit error counter and the receive error counter. There are several rules governing how these counters are incremented and/or decremented.

In order to prevent a single transmitter from inadvertently jamming the CAN bus, a transmitter detecting a fault increments its transmit error counter faster than the listening nodes will increment their receive error counter. This is because there is a good chance

that it is the transmitter itself that is at fault. When any error counter rises over a certain value, the node will first become “error passive.” That is, it will not actively destroy the bus traffic when it detects an error. It then becomes “bus off,” which means that the node does not participate in the bus traffic at all.

Using the error counters, a CAN node not only detects faults but also performs error confinement. The CAN protocol defines no less than five different ways of detecting errors. Two of these work at the bit level, and the other three at the message level.

Bit Monitoring: Each *transmitter* on the CAN bus monitors (i.e., reads back) the transmitted signal level. If the bit level read differs from the one transmitted, a bit error is signaled. (No bit error is raised during the arbitration process.)

Bit Stuffing: When five consecutive bits of the same level have been transmitted by a node, it will add a sixth bit of the opposite level to the outgoing bit stream. The receivers will remove this extra bit. This is done to avoid excessive DC components on the bus, but it also gives the receivers an extra opportunity to detect errors: if more than five consecutive bits of the same level occurs on the bus, a stuff error is signaled.

Frame Check: Some parts of the CAN message have a fixed format, i.e., the standard defines exactly what levels must occur and when. These parts which will be explained below are the CRC Delimiter, ACK Delimiter, End of Frame, and also the Intermission, though there are more specific error checking rules for this last message type. If a CAN controller detects an invalid value in one of these fixed fields, a form (or format) error is signaled.

ACKnowledgement Check: All nodes on the bus that correctly receive a message (regardless of its “pertinence”) are expected to send a dominant level in the so-called acknowledgement slot in the message. The transmitter will send a recessive level here. If

the transmitter cannot detect a dominant level in the ACK slot, an Acknowledgement Error is signaled. (Note that an ACK error will not increase the error counters. When the bus is first powered, an active node cannot put itself into a "bus-off" state while transmitting and waiting for another node on the network to respond).

Cyclic Redundancy Check: Each message features a 15-bit Cyclic Redundancy Checksum (CRC), and any node that detects a different CRC in the message than which it has calculated itself will signal a CRC error.

The CAN bus typically runs at a bit-rate of between 25 kilobits/second and 1 megabit/second. Due to the overhead of the identifier and CRC blocks, the data rate is slower than the bit-rate by 36%. Another feature of the priority based message arbitration is that the time of arrival for the highest priority message is deterministic. At a bit-rate of 1 Mbit/s, this message will arrive at all other nodes a maximum of 134 microseconds from transmission (this assumes that the longest possible message has just won the arbitration when the highest priority message starts to transmit).

On the Atlantis, the CAN bus is running at 500 kbits/s. There are a total of 17 microcontrollers on the Atlantis, each transmitting several messages at a rate of 100 Hz. This results in a roughly 28% load on the CAN bus. This number is somewhat misleading as all microcontrollers burst transmit all of their data simultaneously resulting in a 100% bus load for 28% of the time and no traffic on the CAN bus until the next synchronization message. This results in a data rate of 13.6 kilobytes/second.

The CAN bus is a built-in hardware peripheral of the Siemens microcontrollers. The microcontroller boards include CAN transceivers that convert the signals from the microcontroller pins to the differential signaling used on the physical media of the CAN bus. There exist myriad CAN enabled microcontrollers that are inexpensive, standardized, and easily programmable in "C." The CAN network is usually set up as a special location

in memory. Writing to the memory results in a message being issued on the network and reading the memory checks for the arrival of a new message. In addition, most microcontrollers are set up with a hardware filter that allows only messages whose arbitration field matches a specified pattern to generate an interrupt to the microcontroller. This reduces the overhead on the microcontroller and allows a timely response to messages of interest.

SECTION 2.4 SENSORS AND ACTUATORS

Each node on the system is a sensor or an actuator, and is responsible for either measuring some physical parameter of the system or causing the system to respond to a signal. What follows is a brief description of each sensor and actuator, with a brief explanation of the physical process of the sensor as required. Figure 2-4 shows a block diagram of the Atlantis system with the CAN bus and various sensors and actuators.

SECTION 2.4.1 ANEMOMETER AND WEATHERVANE

There are two sensors for determining the local wind vector. This is essential, since without knowledge of wind direction, we cannot determine if a desired waypoint is achievable directly or not. The magnitude of the wind velocity vector is determined by the anemometer and its relative direction comes from the weathervane.

The anemometer and weathervane are a commercial transducer unit manufactured by Standard Marine Electronics. An aluminum mount was machined to allow the transducer to be mounted at various placed such as on the wooden forward crossbeam, on the electronics pod cover, or at the top of the wing underneath the GPS antenna. The transducer unit is pictured in Figure 2-5 below.

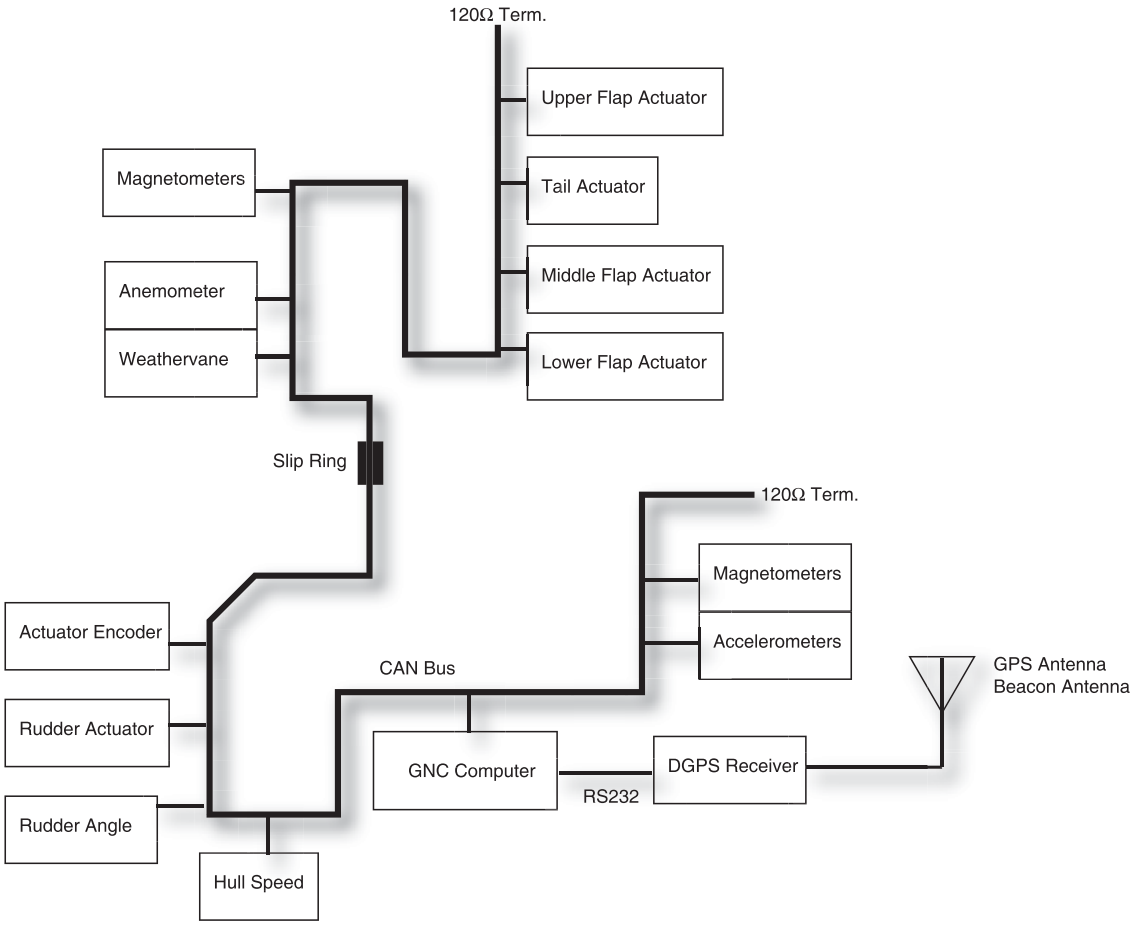

Figure 2-4 System block diagram of the Atlantis. The GPS receiver is DGPS receiver that receives both GPS signals and the USCG Beacon differential broadcasts. It communicates to the main GNC computer through a serial port. All other communication between sensors and actuators is through the CAN bus. The slip ring is used to keep the wires from the CAN bus from winding up as the wing rotates.

The weathervane operates simply through a drag-rudder method. The (comparatively) large surface generates a stabilizing force any time the weathervane is not aligned into the relative wind, and returns it quickly to point directly into the relative wind, since the moment of inertia is very small. The entire weathervane assembly is mass-balanced about the pivot point to ensure that the weathervane measures wind direction uncorrupted by deviations in pitch or roll of the platform. Further analysis of the weathervane is found in Section 3.2 on page 64.

Figure 2-5 Standard Marine Electronics anemometer and weathervane transducer. The weathervane outputs an analog quadrature signal with a 5 volt sine and cosine components. It is capable of recording absolute position of the weathervane from 0 to 360 degrees without wind-up. The anemometer produces a square wave proportional to the wind speed. The greater the frequency of the square wave, the greater the wind speed.

The anemometer functions through differential drag on the cups. The drag of the open cup is greater than the drag on the closed cup, thus the differential in drag, coupled with a wind velocity, results in rotation. The speed of this rotation is not, in fact, proportional to wind velocity, but rather proportional to the dynamic pressure of the flow field. The wingsail responds to the same effective (indicated) airspeed. If required, the indicated airspeed can be corrected utilizing absolute pressure and temperature. Details of the anemometer, along with calibration and validation data are presented in Section 3.1 on page 57.

SECTION 2.4.2 HULLSPEED

In order to generate data for the water current, the velocity of the Atlantis through the water must be known (as GPS will provide velocity referenced to Earth). Thus, a hullspeed transducer is required. The hullspeed transducer used on this project is a commercially available transducer, fabricated by Standard Communications Electronics Corporation (Figure 2-6). It utilizes a four-armed paddle wheel made out of a magnetic ceramic material. The paddle wheel housing is such that the paddle remains semi submerged in the flow passing the hull. A Hall-effect sensor detects the passing of the oppo-

site paddle arm and generates an electrical signal. The paddle is assumed to be incapable of having any motion with respect to the passing water (i.e., the drag of the water is so great that the paddle tip moves as another integral part of the water stream), and thus the rate of pulses is proportional to the inverse of the water velocity.

Figure 2-6 Standard Communications Electronics hullspeed sensor, pictured before installation on starboard hull. This through hull sensor remains semi-submerged in the flow of the water around the hull. It has a magnetic four-armed paddle wheel and a Hall-effect sensor to produce a square wave whose frequency is proportional to the velocity of the water flowing past the sensor (or the velocity of the sensor through the water).

Note that in the Atlantis, the hullspeed sensor was placed along the centerline of the starboard hull, approximately 1 meter behind the centerboard, and 0.4 meters in front of the rudder leading edge. Details of the hullspeed sensor and calibration are given in Section 3.3 on page 67.

SECTION 2.4.3 RUDDER ANGLE

An integral part of the control system is the knowledge of rudder position. That is, in order to generate an accurate control model, a sensor that can resolve the position of the rudder accurately is required. The sensor utilized is a LoHet, which senses magnetic flux. The LoHet is mounted vertically on the pivot axis of the starboard rudder, and a set of powerful rare-earth magnets is mounted on either-side of the rudder, with the flux lines of the magnetic field perpendicular to the LoHet. The LoHet was made waterproof by “pot-

ting" it in epoxy. Two high strength ceramic magnets are mounted on aluminum tangs attached to the rudder itself. This installation is pictured in Figure 2-7. This method of resolving rudder angle is used for several reasons, but mainly because the sensor is water-proof and the measurement is unaffected by submersion in salt water. Full details of the LoHet, along with calibration methodologies are given in Section 3.4 on page 72.

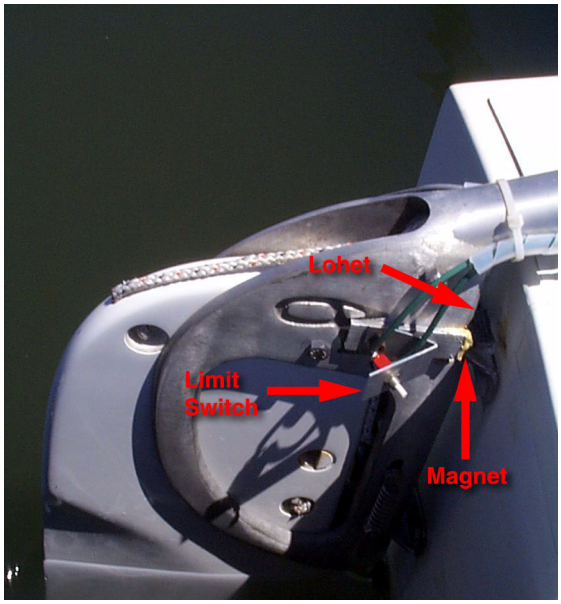

Figure 2-7 Close up of the starboard rudder. The LoHet angle sensor is mounted on the hinge line and "potted" in epoxy for waterproofing. Two aluminum tangs were manufactured to hold the ceramic magnets that create the magnetic field for the LoHet sensor. Limit switches were installed to prevent the rudder actuator from forcing the rudder to an extreme angle which would have damaged the magnets.

Another possible sensor for rudder angle is to keep track of the number of turns of the rudder actuator. The rudder actuator is equipped with an optical encoder that records 500 pulses per revolution. There are two main difficulties with using the encoder counts to keep track of rudder angle. The first is that the encoder keeps track of relative position since the microcontroller has been reset. This results in an ambiguity about the initial off-set from zero. This can be resolved with an additional sensor or switch to reset the counter at some "home" position. The second disadvantage is that the number of counts has a

rather complex mapping to the actual rudder angle. While this can be mapped, there are no inherent advantages over the Lohet that was used.

SECTION 2.4.4 ATTITUDE SYSTEM

Another integral requirement for the control of Atlantis is an attitude system. The primary sensor required is for azimuth, or heading, but in order to extract the true azimuth full attitude is needed. This comes from the fact that a magnetometer or compass will deviate from the true heading measurement if pitched or rolled. The heading measurement is made from the arctangent of the North and East components of the magnetic field. Earth's magnetic field around San Francisco, California, has a dip of approximately 70 degrees. Thus, pitching or rolling the magnetometer causes the North and East components of the measurement to be skewed by the portion that they are sensing the "down" component of the magnetic field. Additionally, the ideal location for the position measurement (GPS antenna) would be at the vessel center of gravity, however, the center of gravity (CG) in general has a very poor view of the sky making the placement of the GPS antenna at the CG location impractical.

However, utilizing both position and attitude, a "synthetic" sensor can be created by measuring the position at a convenient location and correcting for the lever arm between actual position of the antenna and its desired position on the vehicle using the full attitude solution. The actual attitude system is based on a novel solution of vector matching, or Wahba's problem, based on a gyro-free quaternion solution. This is the subject of Chapter 4, and all the details—including algorithm development and simulation—are presented there. Figure 2-8 gives a view of the overall attitude box which contains the attitude system. The yellow box is a Pelican case, capable of being submerged to a depth of 10 meters without flooding.

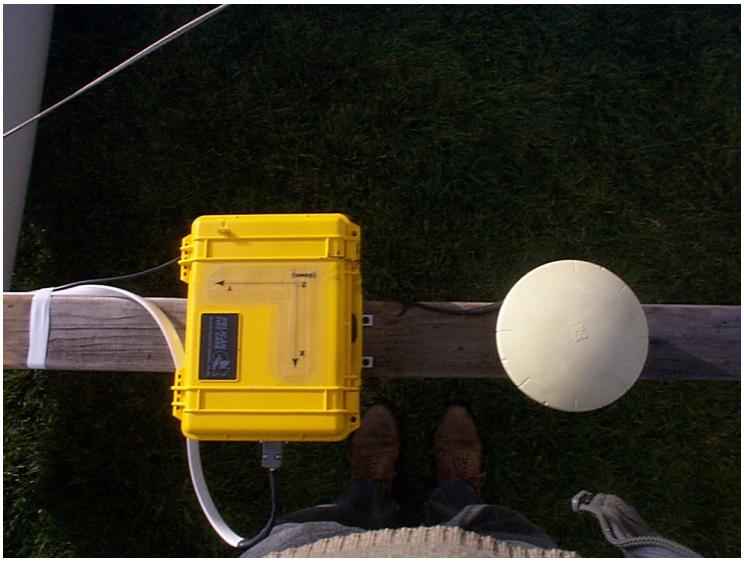

Figure 2-8 The forward wooden crossbeam as viewed from above. The attitude system is contained within the yellow Pelican case waterproof housing. Inside the Pelican case, a two-axis micromachined accelerometer and a three-axis hybrid magnetometer are connected to a Siemens 515 microcontroller. The GPS antenna is located to the right of the picture, and houses both the GPS and the Coast Guard beacon antennae.

SECTION 2.4.4.1 ACCELEROMETERS

There are two sensor sections of the attitude system, an orthogonal magnetometer triad, and an orthogonal two-axis accelerometer set. The accelerometers are a set of micro-machined, capacitively sensed accelerometers made by Analog Devices. The Analog Devices ADXL202 is a low-cost, low-power (250 microamperes per axis), complete two-axis accelerometer with a digital output on a single monolithic chip. The outputs are Duty Cycle Modulated (DCM) signals whose duty cycles (ratio of pulse width to period) are proportional to the acceleration in each of the two sensitive axes and are measured directly with the microcontroller counter. No analog-to-digital converter or glue logic is required. The DCM has a resolution of approximately 14 bits. The ADXL202 is ideally suited for a wide variety of applications that need to measure gravity (tilt), vibration, shock and motion. Full details of the accelerometer can be found on the Analog Devices data sheets, given in [2]. The bandwidth of the ADXL202 may be set from 0.01 Hz to 6 kHz, as

required, and the typical noise floor is 500 µg divided by the square root of Hz. Figure 2-9 shows the microcontroller, accelerometers and magnetometers together in the installed package.

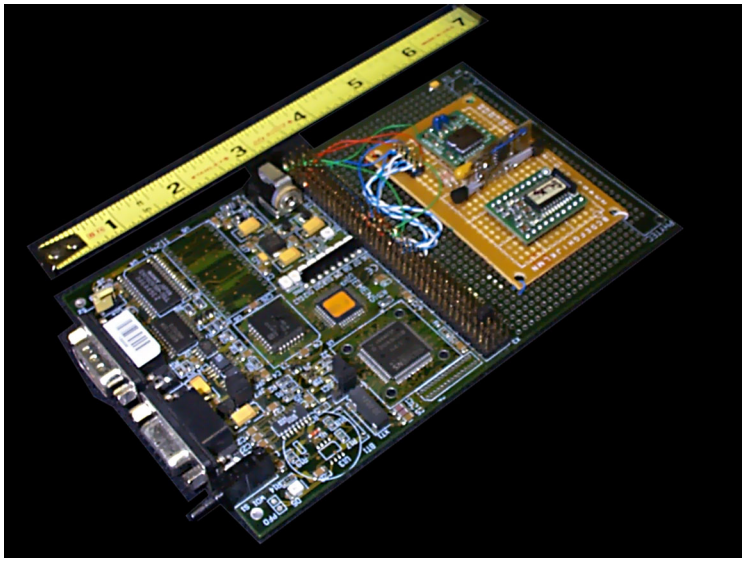

Figure 2-9 The attitude system microcontroller, accelerometers and magnetometers. The attitude system contains an Analog Devices ADXL202 two-axis micromachined accelerometer and a Honeywell HMR 2003 three-axis strap-down magnetometer. The attitude is resolved using an iterated solution to Wahba's problem in the quaternion domain.

SECTION 2.4.4.2 MAGNETOMETERS

The magnetometers are a low-cost, three axis magnetometer triad developed by the Honeywell Corporation. The HMR2003 is a hybrid perma-alloy magneto-resistive sensor, with built in degaussing straps. This allows the polarity of the sensor to be reversed periodically. Each polarity reversal results in an opposite measurement. By differencing successive measurements after polarity reversal the sensitivity is doubled while at the same time the bias in the measurement is significantly reduced. This is due to the fact that the common mode cancels after polarity reversals, whereas the true magnetic field measurement is doubled. Each axis on the HMR2003 is capable of reading +/- 2 gauss (where the Earth's magnetic field is approximately 0.5 gauss), and gives an analog output. In order

for the HMR2003 to operate well, a strong set and reset pulse of approximately 4 amps must periodically be sent into the sensor. These pulses are sent at a rate of 200 Hz in order to generate a differenced pair of subsequent measurements at 100 Hz. This requires some power mosfet components and a large capacitor to act as a local charge reservoir. The typical accuracy of the magnetometers is approximately 3 degrees (1-σ) before any averaging or smoothing. Section 4.7 on page 122 details the calibration of the magnetometers in order to use them effectively in the attitude solution.

SECTION 2.4.5 GPS RECEIVER

The most important sensor on the Atlantis is the GPS receiver. The GPS receiver provides position and velocity information at 5 Hz. and is the basis for the entire control system function. The GPS receiver is a Trimble Ag122 receiver with a built-in coast-guard beacon receiver. Details of the GPS receiver testing and accuracy modelling are found in Section 3.5 on page 76.

The Global Position System (GPS) is the most recent, and most accurate, in a long lineage of radio-navigation aides (see [106], [80], and [50]). The GPS constellation consists of a nominal 24 satellites in six orbital planes (Figure 2-10), each in a semi-sidereal orbit (crossing the same point on Earth's surface every 12 sidereal hours). Each satellite is equipped with a very accurate atomic clock synchronized with a master ground station clock. All 24 satellites broadcast a synchronized spread spectrum signal that contains information about the satellite constellation and the identity of the broadcasting satellite. Utilizing the difference in arrival times of the signals, a user receiving the signals can solve for his/her own position and a time offset of their own local clock from the GPS system time. This is referred to as "stand-alone GPS."

The accuracy of unaided GPS ranges from 5 to 15 meters based on the geometry of the satellites, the state of the iono- and tropo-spheres, and other noise sources (the signal

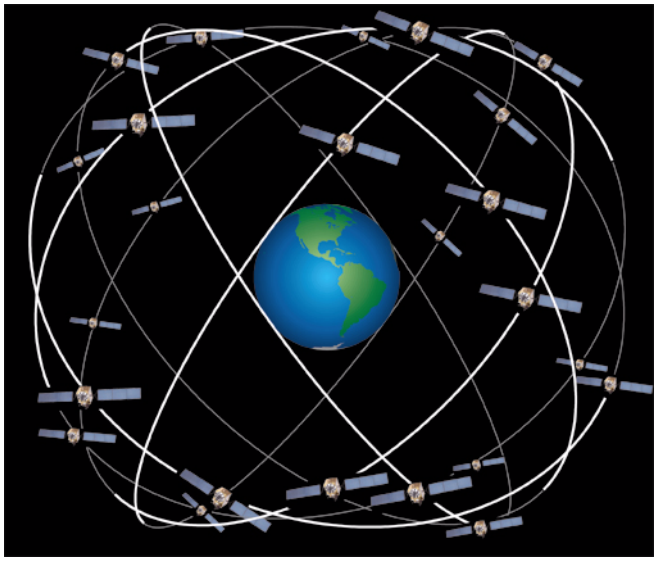

Figure 2-10 A snapshot of the GPS constellation which consists of 24 satellites in 6 orbital planes. The satellites are in semi-sidereal orbits, crossing the same position on Earth every 12 hours. The satellites each carry a high precision atomic clock and are synchronized to each other. This allows a user on Earth to solve for their position using four ranging signals from the GPS constellation.

was used to be degraded intentionally by the Department of Defense, but this has ceased since May 2000). While this is more than adequate for most navigation requirements, precision control cannot be achieved with this coarse navigation solution. Likewise, high integrity requirements such as aircraft landing systems cannot be met by the unaided GPS solution. Based on the requirement to improve both the accuracy and the integrity of the navigation solution, several differential correction methods have been developed.

The basic premise of differential GPS (DGPS) is that the errors seen by a user and a reference station are essentially the same. A reference station has its location previously surveyed, and encoded into its electronics. Thus, when the reference station solves for its own position, it is immediately aware of the difference between its known (surveyed) position and the calculated position. The reference station then generates a set of “corrections” to the signal that will bring its calculated position and its surveyed position together. As long as the assumption that these errors are common holds true, all local GPS users can use these same “corrections” to bring their increase the accuracy of their own GPS posi-

tion. This remains valid as long as the reference station and the user are neither spatially nor temporally decorrelated. Spatial decorrelation refers to corrections that become less and less useful as the user and the reference station become more distant from each other, and temporal decorrelation refers to corrections that become less useful due to latency of the broadcast corrections. (This is, of course, a simplistic explanation of DGPS. Corrections are made to the signal of each satellite visible by the reference station, and the user applies these corrections as appropriate.)

Several differential systems are either in use today or will be available in the near future. Perhaps the most ambitious of these is the Wide Area Augmentation System (WAAS) which is operational now and slated by the FAA to be certified for aircraft use in 2003. This system uses several reference stations at very large distances (~1500 kilometers) to generate global corrections and achieve 1-5 meter accuracies [46]. In addition to WAAS, the United States Coast Guard and International Light House Association have placed code-differential beacons to cover virtually every populated coastline and navigable waterway in the United States and Europe [136]. This network provides a 1-3 meter accuracy within 250 nautical miles of the broadcasting station and is operational now.

SECTION 2.4.6 WINGSAIL ANGLE

In order to determine the angle of the wing, referenced to the centerline of the boat, another Honeywell HMR2003 three-axis magnetometer is placed inside the electronics pod of the wingsail. Since the roll and pitch of the wing and boat are constrained by the stub-mast to be the same, then both the attitude system magnetometers and the wingsail magnetometers will have their horizontal components corrupted in the same way by the dip of the magnetic field. The only difference in the arctangents of the *X*- and *Y*-components of the magnetic field from the two magnetometers can be from the azimuth difference of the wing from the boat. That is—while the arctangent of the *X*- and *Y*-components

of either the attitude magnetometers or the wingsail magnetometers will be neither the true heading of the hulls or the true heading of wingsail—the difference in arctangents will be the angle between the hulls and the wingsail. The scale factor errors will be dependent on the roll and pitch angles, but can be mapped through normal calibration techniques. An Infineon SAB505CA microcontroller is used to sample the magnetometer and report the relative angles to the main GNC computer.

SECTION 2.4.7 RUDDER ACTUATOR

The rudders are driven electronically by a brushed DC-motor and lead screw

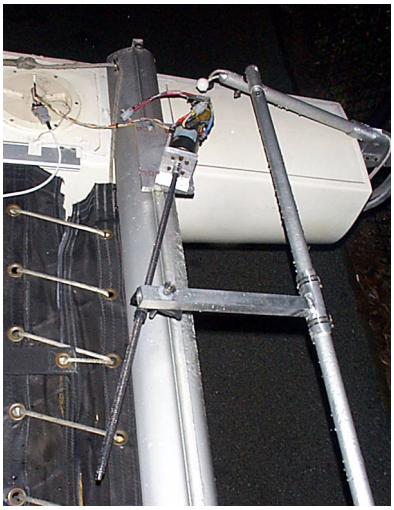

Figure 2-11 A Pitman 24V motor is used as the rudder actuator, as seen from above. The actuator is a fractional horsepower DC motor with a built-in encoder and a lead screw assembly fitted between the motor and the rudders. The motor is controlled via pulse width modulation (PWM) as commanded by an Infineon microcontroller acting through an Infineon 5-amp H-bridge that regulates the motor speed. This allows the rudder slew rate to be varied up to +/- 25 degrees/ second.

assembly. The motor is a 24 volt Pitman direct drive motor de-rated to 12 volts, with a stainless steel lead screw attached to the end. The entire motor lead screw assembly is constrained in the pitch and roll degrees of freedom (with respect to the catamaran deck) but is free to pivot in the azimuth direction. The lead screw macpherson nut is made out of

delrin, and is attached to the existing linkage that connects the tiller the rudders (Figure 2-11).

The motor drive assembly is controlled via pulse width modulation (PWM) from an Infineon SAB505-CA microcontroller driving a Siemens TLE5205S 5-amp H-bridge drive. Essentially, this allows the entire bus voltage (12 volts) and full amperage (6 amps when stalled) to be sent through the motor on demand. With this setup, rudder angle slew rates of ~25 degrees/second are achieved. Full details of the rudder actuator linearization mapping are found in Section 3.6 on page 82.

SECTION 2.4.8 TAIL ACTUATOR

The tail actuator consists of an identical motor to the rudder actuator, being driven

Figure 2-12 A view of the installed flap actuator in the lower wing section before covering. The flap and tail actuators are identical. The actuators are Pitman 24V DC motors, de-rated to 12V. A stainless steel lead screw assembly is used to move a delrin macpherson nut. Stainless steel pushrods are attached to the macpherson nut, which are connected to control horns (not shown) on the flaps and tail.

in an identical fashion through the same kind of H-bridge (Figure 2-12). This was done to retain modular hardware as much as possible through Atlantis. The lead screw assembly is machined out of an aluminum channel, with a 3/16” threaded rod serving as the lead screw. The macpherson nut is machined from delrin, and has a stainless steel pushrod that exits the wing and is attached to a control horn on the tail. There are three limit switches

that provide an on-off feedback to the microcontroller. The two end limit switches prevent the motor from jamming the lead screw or breaking the tail. The center or home switch is triggered every time the actuator has the tail centered. The results in a reset to the microcontroller counter that keeps track of encoder counts which re-zeros the tail position.

SECTION 2.4.9 FLAP ACTUATORS

There are three trailing edge flap actuators distributed along the wing. They are

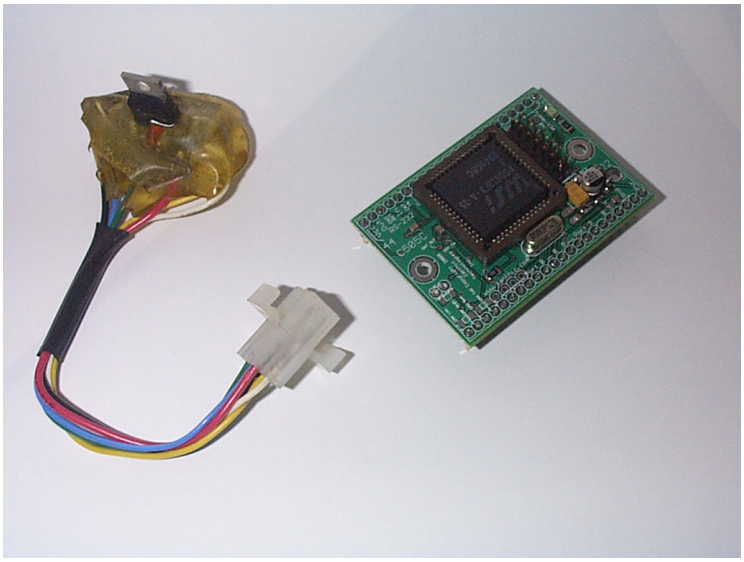

Figure 2-13 A close up view of the Infineon SAB505-CA microcontroller and Siemens 5-amp H-bridge used to control the rudder, flap, and tail actuators. The microcontroller commands the H-bridge through the Capture and Compare hardware functions. This allows very fine resolution of the PWM control of the motors. The H-bridge is a modular unit, and has been "potted" in epoxy to make it waterproof.

identical to the tail actuator in every way except for the length of the pushrod. Each is controlled by its own Infineon SAB505-CA microcontroller and H-bridge chip and is individually addressable, thereby providing a splitting of the flap motion which is necessary to prevent a capsize. In the event of excessive roll angle (indicating an imminent capsize) the flap on the upper section would be reversed to pull the wing back upright while the lower sections maintain the wing camber for the increased lift coefficient. Additionally, this

allows the loading of the wing to be “tuned” to maximize the usable lift. These actuators are made in a modular fashion in order to reduce the parts count, as well as facilitating the quick replacement a faulty actuator.

3 Sensor Calibration and Actuator Linearization

In this chapter the non-linearities of both the sensors and actuators are considered. One of the basic tenets of feedback control systems is that a great amount of aggregate knowledge is available on how to control a linear time invariant (LTI) systems (a system whose dynamics can be described by a linear system of equations). Although techniques exist to deal with non-linear control systems, most of the time the non-linear systems are dealt with by linearization about some nominal control point. In order to realize a linear system from a non-linear one, a perturbation analysis of the system is performed about an equilibrium point. This is usually sufficient for linearization of the dynamics of the system plant. However, the sensors and actuators must be calibrated and linear-

ized in order to provide accurate data for the linearized system plant and actuators.

Non-linearities in sensors and actuators present design challenges. However, leveraging the presence of distributed microcontrollers, each sensor or actuator is calibrated and linearized at the point of data collection. The basic method of sensor linearization is to first digitally low-pass an oversampled data stream, and then match curves based on a lookup table to compensate for biases and scale factors. On certain sensors, there is a requirement to correct for non-linearity and even coupling between two sensors. Note that the calibration is performed off-line and stored in the non-volatile RAM of the microcontrollers.

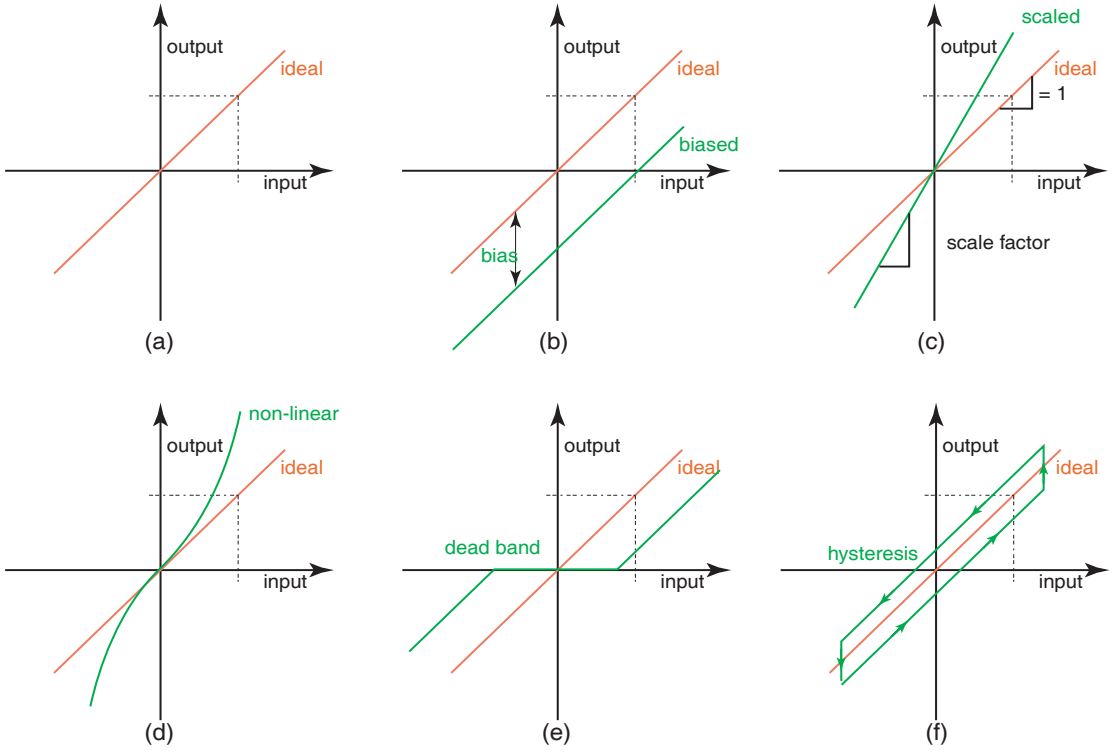

Figure 3-1 Sensor responses presented with common errors and non-linearities. The first frame (a) is an ideal sensor, where the output is exactly proportional to the input. The second frame (b) shows the true output offset from the ideal by a bias error. The third frame (c) shows the output slope greater than one as by a scale-factor error. In the lower figures, a generic non-linear response is shown (d), along with a dead-band (e) and hysteresis (f). Any real sensor will always be corrupted by errors such as these.

Figure 3-1 shows various sensor corruptions from an ideal, linear sensor response. The top left figure, “a” is an ideal sensor. In frame “b”, the presence of a bias is examined. The top right figure, “c,” shows a scale-factor error. On the lower line of images, the left most image, “d,” shows a simplified non-linear response. The middle image, “e,” shows a typical dead-band response, and the last image, “f,” demonstrates the presence of hysteresis. Sensors frequently have dynamics such as lags and transient responses as well. All of these errors need to be addressed for each and every sensor and actuator; the actuators have a reverse kind of response: rather than provide an electrical signal proportional to a sensation, they provide force in proportion to an electrical signal.

Again, in order to facilitate the discussion that follows, a block diagram of the system is reproduced from Chapter 2, below in Figure 3-2.

SECTION 3.1 ANEMOMETER

The anemometer is a Standard Communications Electronics three cup anemomenter. It senses wind speed through differential drag on the open faces versus closed faces of the cups. See Figure 2-5 for a photograph of the transducer head. A schematic of the transducer head is pictured below in Figure 3-3.

The relevant parameters are the wind speed, V_0, the radius of the cups, r, and the drag on the cups facing and opposing the wind. Define the drag of the cup facing the wind as C_{Dopen} and the drag of the cup pointing away from the wind as $C_{Dclosed}$. The angular rate of the anemometer is defined as Ω The velocity of the cup swinging towards the wind

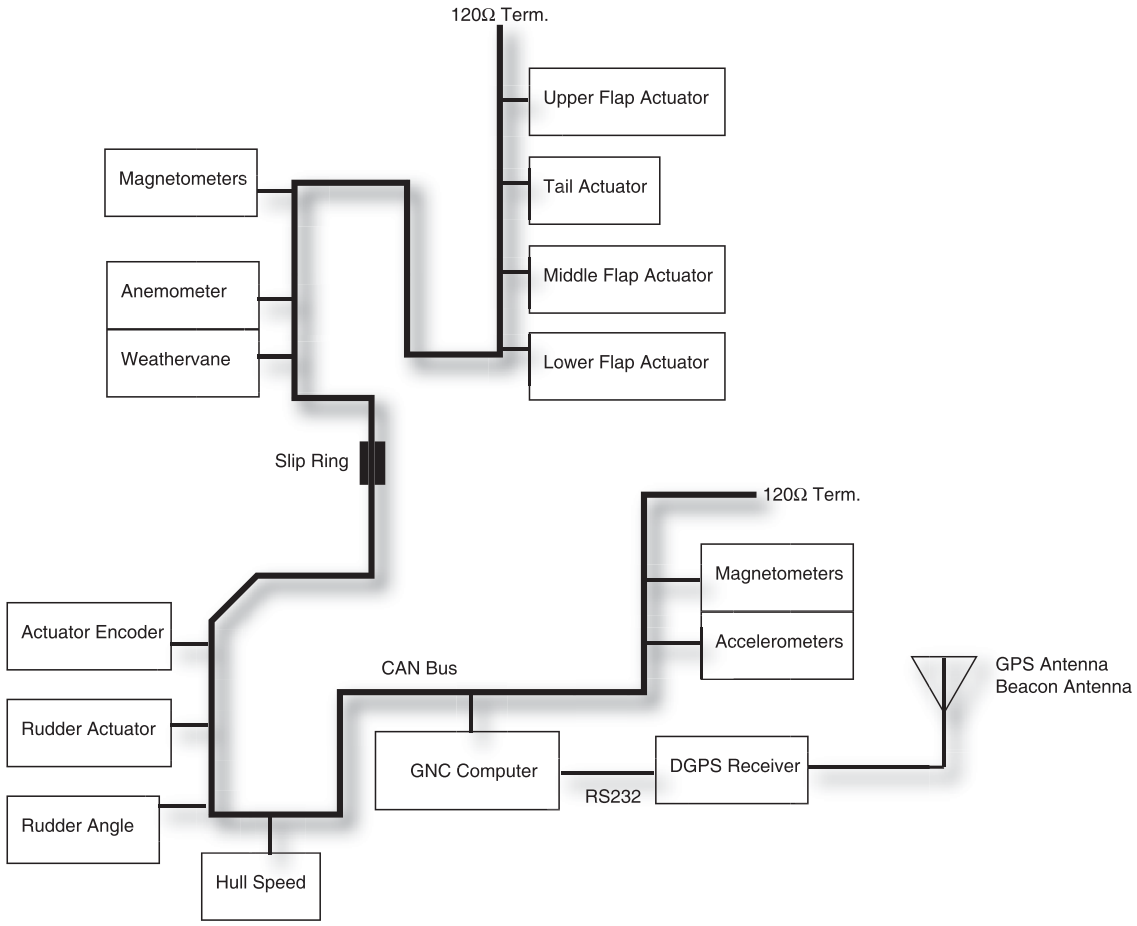

Figure 3-2 System block diagram of the Atlantis. The GPS receiver is DGPS receiver that receives both GPS signals and the USCG Beacon differential broadcasts. It communicates to the main GNC computer through a serial port. All other communication between sensors and actuators is through the CAN bus. The slip ring is used to keep the wires from the CAN bus from winding up as the wing rotates.

is $r\Omega+V_0$. Whereas the velocity of the cup retreating from the wind is $r\Omega-V_0$. Thus, balancing the drag forces on both cups:

$$\frac{1}{2}\rho(r\Omega + V_0)^2 SC_{Dclosed} = \frac{1}{2}\rho(r\Omega - V_0)^2 SC_{Dopen} \qquad \text{(EQ 3.1)}$$

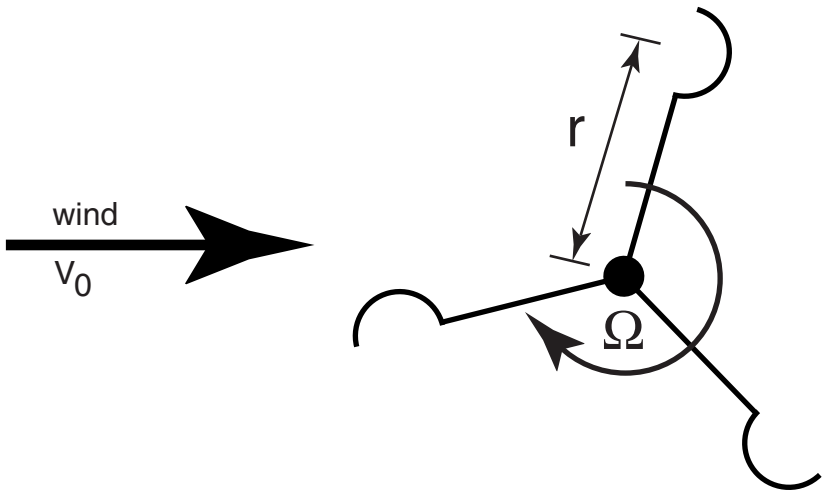

Figure 3-3 Schematic of the anemometer transducer. The open face of the cup has greater drag than the closed face of the cup. Thus, as the wind speed increase, the rotation rate increases. This occurs because the forces balance due to the increased relative wind on the closed face of the cup versus the decreased relative wind on the open face. That is, the velocity of the open cup face is the wind velocity minus the product of the angular velocity of the anemometer and its radius. The same is true of the closed face, but the cup velocity is added to the wind speed.

where ρ is the density of the air (1.025 kg/m^3 at sea-level), S is the area of the cups in square-meters. Defining:

$$\lambda \equiv \frac{C_{Dopen}}{C_{Dclosed}} \qquad \text{(EQ 3.2)}$$

and substituting into Equation 3.1, leads to:

$$V_0{}^2 + 2V_0 r\Omega\frac{(1+\lambda)}{(1-\lambda)} + r^2\Omega^2 = 0 \qquad \text{(EQ 3.3)}$$

where:

$$\lambda \cong 3.5 \qquad \text{(EQ 3.4)}$$

Equation 3.3 can be factored in the usual manner and the negative root is discarded. This yields a monotonic, linear relationship between V_0 and Ω, within the range of speeds of interest. A full investigation of the cup anemometer can be found in [71].

The electrical interface to the anemometer is such that a square wave is produced twice each revolution of the anemometer. The frequency of this square wave is proportional to the ambient wind velocity. In order to measure the frequency of the square wave, a timer is triggered off of the rising edge of this pulse. While this works well for steady and increasing winds, it suffers from a problem of observability when the anemometer stops. Once the anemometer has stopped, no rising edge triggers the interrupt on the microcontroller, and thus the anemometer continues to read the last value indefinitely. In order to compensate for this shortcoming, software within the microcontroller generates an interrupt periodically while monitoring the output of the anemometer, thus smoothly ramping the measured velocity down with an exponential decay until the next read value. While this is not perfect, it leads to very smooth, repeatable data on wind velocity, which is a primary parameter for control of the wing.

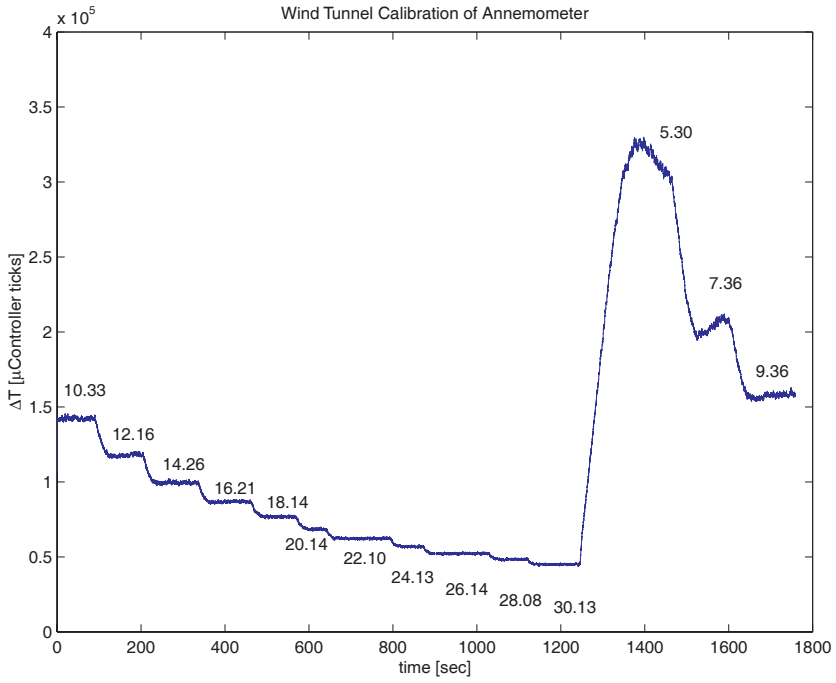

Figure 3-4 Anemometer response data taken during a wind tunnel calibration run. The anemometer was placed in a small wind tunnel and the velocity was recorded. The numbers are the velocities presented by the wind tunnel control computer. Using this data, a linear fit was used to calibrate the anemometer.

SECTION 3.1.1 METHODOLOGY

In order to calibrate the anemometer a map between the pulse frequency, or time difference of arrival of subsequent pulses, and the true wind velocity must be created. In order to generate this map the anemometer was placed in a small wind tunnel in the Durand facility at Stanford University. The wind tunnel was set for a given wind speed and the speed allowed to stabilize. Data was taken throughout and later reduced. The actual data from the run is presented in Figure 3-4. Note that the transient response of the wind tunnel is readily visible in the data.

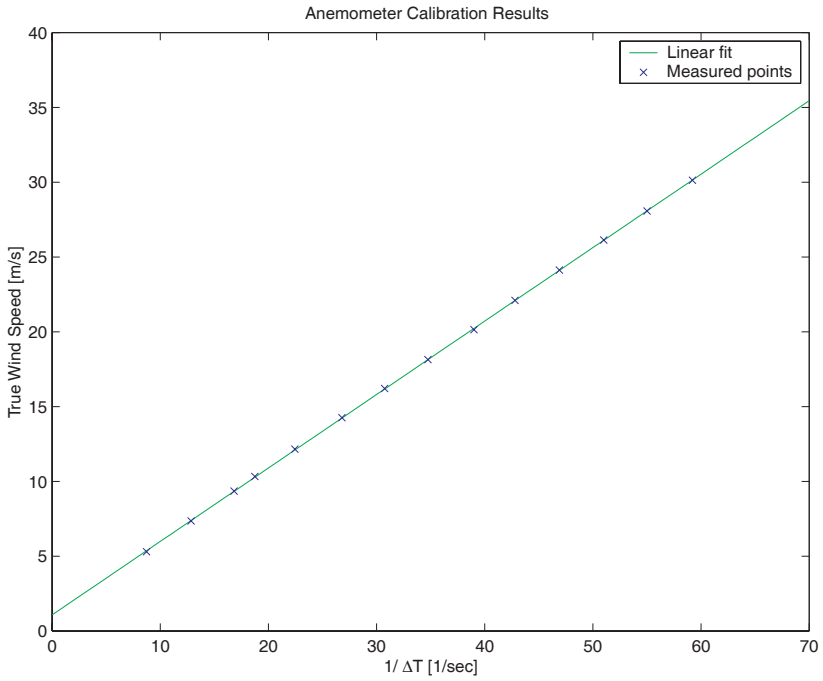

Figure 3-5 Anemometer data from wind tunnel test and linear fit of the same. The linear fit shows an excellent match to the actual points collected from the wind tunnel data. From this linear fit it is evident that the minimum wind speed that can be measured by the anemometer is slightly greater than one meter/second. Also of note is that the x-axis is the inverse of the time between pulses, thus as the wind speed slows, the anemometer reading will be fixed at the previous value until the next pulse arrives.

SECTION 3.1.2 CALIBRATION

The data taken in the wind tunnel has been smoothed, filtered and plotted out against the set speed of the wind tunnel. Based on the data, a linear fit is found to be adequate for calibrating the anemometer velocity response. The plot of the data, along with the linear fit through the points, is shown in Figure 3-5. Thus, the final linear fit reduces to a simple linear offset and scale factor, which then translates the interrupt time difference, Δt, into a wind velocity. The actual equation is:

$$V_{wind} = 0.491032\left(\frac{1}{\Delta t}\right) + 1.079468 \qquad \text{(EQ 3.5)}$$

where V_{wind} is in meters/second and Δt is in seconds.

SECTION 3.1.3 EXPERIMENTAL VALIDATION

Figure 3-6 Acura Integra with anemometer and GPS antenna installed for experimental validation of the linear calibration of the anemometer. Velocity was recorded from GPS, and from the anemometer as well. Flow effects due to the shape of the car, as well as any ambient wind, will distort the measured airspeed. Tests were performed on calm nights to minimize the ambient wind.

In order to validate the calibration of the anemometer, the anemometer was affixed to a vehicle (a 1988 Acura Integra, see Figure 3-6) and data was taken while the car accelerated and decelerated to various speeds. The first thing to note is that the car was able to decelerate much faster than the anemometer due to the built up angular momentum of the three-cup anemometer. The speed recorded for the vehicle is the GPS velocity which

would also be the velocity recorded by the anemometer, except for wind velocity. Attempts to validate the anemometer occurred only on days that had very little ambient wind speed. Also, it is accepted that the airflow around the car is distorted from the circulation induced by the body of the car. These effects are demonstrated to be minor.

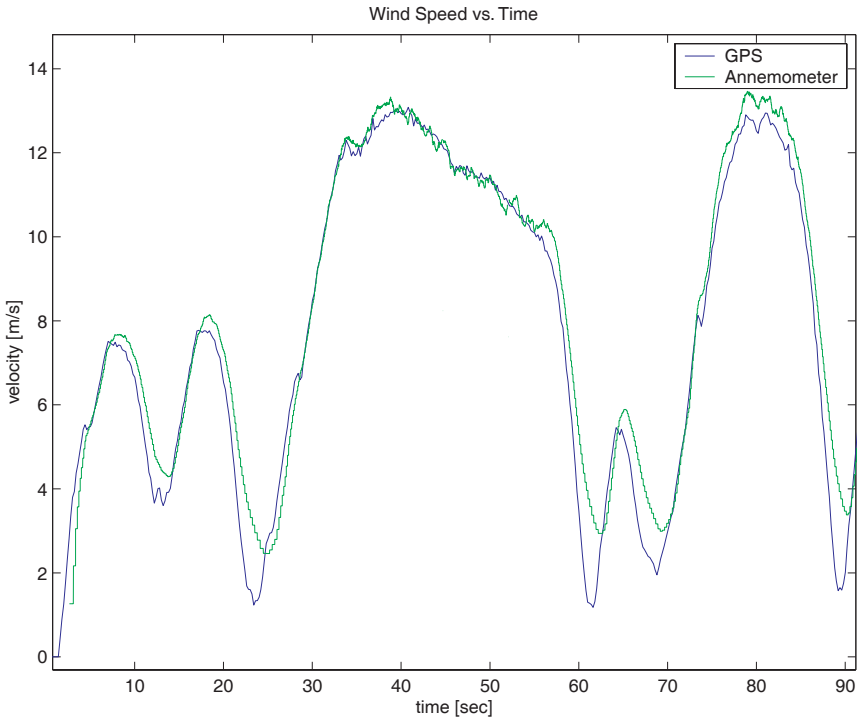

Figure 3-7 Anemometer and GPS velocity data from the experimental validation tests run using the Acura Integra. While the overall agreement is excellent, several notable features are evident. Firstly, the overshoot and inability of the anemometer to decrease in speed quickly. These effects are due to the angular momentum stored in the anemometer. Also, the ragged features of the high speed data is due to the buffeting of the airflow around the car at that speed.

The results of one of several typical validation runs are presented in Figure 3-7. As can be seen, the results are in excellent agreement, including the noted lags in deceleration each time the car decelerated quickly. The excellent agreement is crucial because using GPS velocity and the anemometer and weathervane (explained below) allows the recovery of the true wind speed. This will become essential in providing a measurement of sailing efficiency as well as a required variable to normalize wingsail performance data.

SECTION 3.2 WEATHERVANE

The second part of the anemometer is the weathervane which is used to indicate wind direction. Note that the weathervane is later used as an angle-of-attack indicator for the wingsail. By mounting it on the pod cover, the true wind direction is recovered using the wing-to-centerline angle plus the angle-of-attack. The weathervane itself is a mass-balanced fin that will point directly into the relative wind, with little or no overshoot on the transient response. Internally, it uses two magnetic sensors to deliver a quadrature signal (sine and cosine of the angle). Both of these signals are sampled and an arctangent is used to generate the final angle.

SECTION 3.2.1 METHODOLOGY

In order to calibrate the weathervane, pictured in Figure 2-5, a bob weight is attached first to the nose of the weathervane, and then to the tail, as the weathervane is swung through a series of angles from +90 degrees to -90 degrees. The setup is pictured in Figure 3-8. Using this setup, data was gathered for 19 different angles (every 10 degrees) with the weathervane pointed vertically upwards and again downwards. The data gathered from these trials is presented below, in Figure 3-9.

SECTION 3.2.2 CALIBRATION

Based on the plot of sine and cosine portions of the quadrature signal shown in Figure 3-9, it is seen that a linear fit is best. Higher order fits appear to give no appreciable advantage. The plot for the sine is shown in Figure 3-10, and for the cosine in Figure 3-11. Given these linear fits, the net equation for the weathervane is a simple two equation linear fit based on the received value of the A/D readings from the microcontrollers.

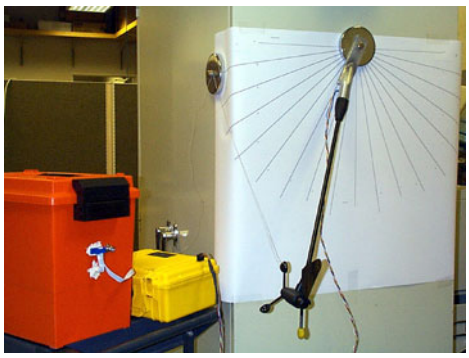

Figure 3-8 Experimental calibration setup for the weathervane. The weathervane was weighted with a nose weight, and the anemometer arm aligned to a precise angle from vertical. Both the sine and cosine components were recorded. Next, the weight was transferred to the tail of the weathervane, and the experiment repeated. Thus, a full 360 degree range was recorded with a sample taken every 10 degrees.

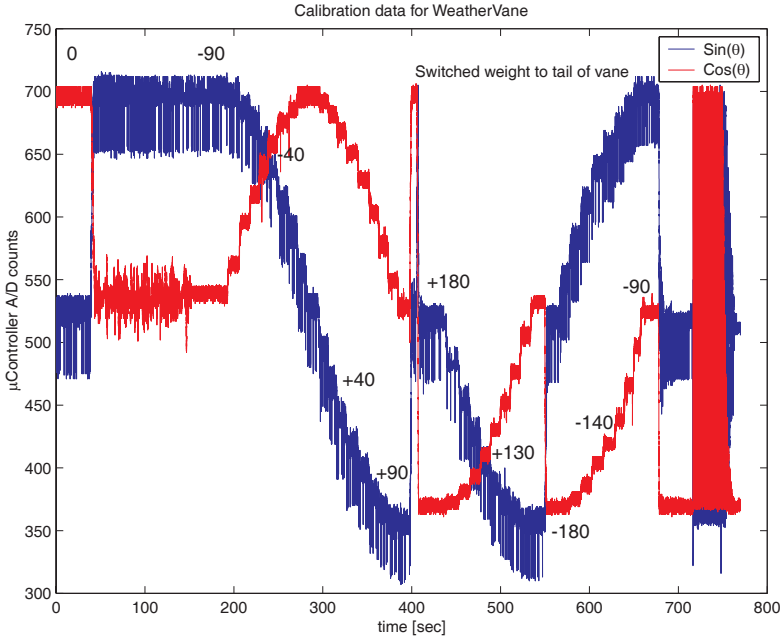

Figure 3-9 Results of the weathervane calibration showing the raw microcontroller A/D counts versus time. Both the sine and cosine quadrature signals are presented. The data contains occasional spikes, though it appears much worse in the graph above due to the limit on pixel width of the lines. Note that at the end of the data collection the weathervane is spun around the axis several times at high speed to generate the realistic limits of the voltage range.

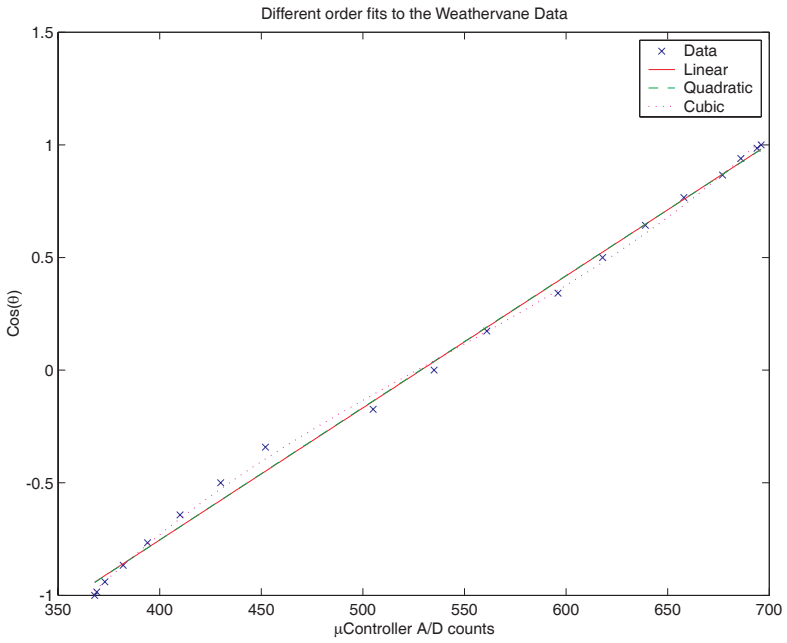

Figure 3-10 A Linear fit of the sine data of the weathervane from the data presented in the previous graph. Note that the real data is scattered, but does not have any obvious pattern. Thus, in spite of using higher order fits to the data, the best fit appears to be the linear one.

SECTION 3.2.3 RESULTS

The microcontroller that samples the weathervane data uses a 10-bit A/D for reading both quadrature signals. The results are oversampled at 100 Hz, and then digitally low-pass filtered down to 5 Hz. using an 8th order Bessel filter. The final calibration of the weathervane is implemented in three steps. The three steps are detailed below in Equation 3.6, Equation 3.7, and Equation 3.8.

$$\sin\psi = -0.005839(A/D) + 3.06016 \quad \text{(EQ 3.6)}$$

$$\cos\psi = 0.0058617(A/D) - 3.098351 \quad \text{(EQ 3.7)}$$

$$\psi = \text{atan}(\cos\psi, \sin\psi) + 0.007882 \quad \text{(EQ 3.8)}$$

Ψ is the angle of the wind with respect to the axis of the anemometer mount, and the result in radians. This uses the four quadrant arctangent function, based on the actual sign of the cosine and sine functions. This allows the output of the weathervane to be com-

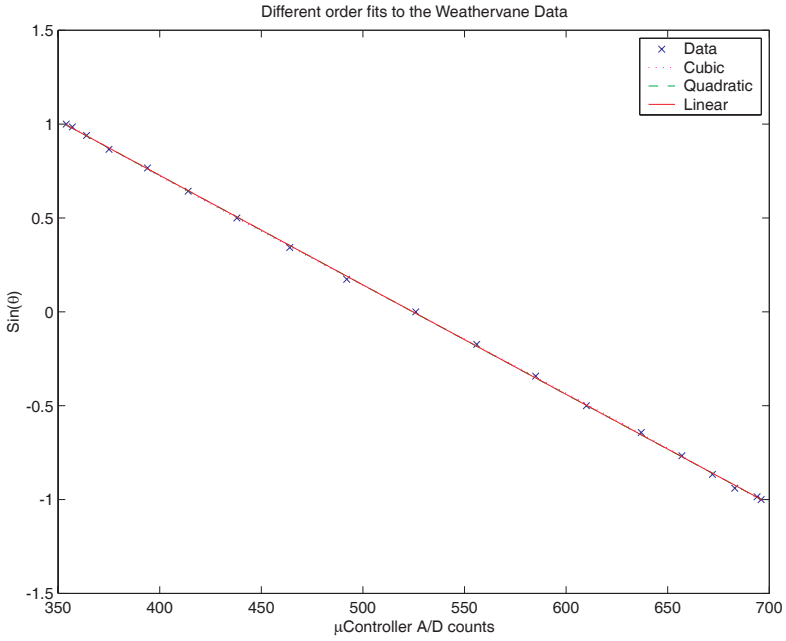

Figure 3-11 A Linear fit of the cosine data of the weathervane from the data presented in the raw weathervane data plot. Note that the data is quite linear, and that higher order fits do not increase the accuracy of the calibration. As for the sine data, the linear fit is best.

pletely calibrated and used to sense either wind direction or wing angle of attack, depending upon whether the anemometer is mounted fixed relative to the wing (angle of attack) or fixed relative to the hulls (true wind direction).

SECTION 3.3 HULLSPEED

The hullspeed sensor is a Standard Communications Electronics through-hole transducer. It is mounted at the center of the starboard hull, about a meter behind the centerboard and forward of the starboard rudder, such that the paddle wheel sticks straight down into the water at the center of the starboard hull. Functionally, the hullspeed transducer works almost identically to the anemometer. It depends on drag in the portion of the paddle-wheel submerged in the flow in a similar manner to the differential cup drag of the anemometer. That is, the hullspeed sensor contains a cruciform paddle-wheel that is half

outside the hull in the water and half contained within the hullspeed sensor housing. The drag on the exposed part of the paddle-wheel causes the paddle-wheel to rotate. The paddle-wheel itself is made from a ceramic magnetic material and its motion is picked up by a Hall effect sensor on the inside of the hullspeed transducer. Simplistically, the water flow is assumed to be solid and the paddle-wheel rolls through it. The factory specification is that there are 22000 pulses per nautical mile.

Ideally, the hullspeed sensor would be calibrated the same way as the anemometer, in a flume (the water equivalent of a wind tunnel). Unfortunately, no flume capable of housing the entire catamaran was available (the sensor itself cannot be calibrated separately due to the interference effects of the hulls). Therefore another method had to be found to calibrate the hullspeed sensor.

Figure 3-12 The Atlantis propelled by a trolling motor, as setup for hullspeed calibration. A series of measurements were made using the GPS velocity to calibrate the hullspeed sensor. It was found that the rudder angle had a very large impact on the hullspeed reading (the aforementioned coupling), and thus the calibrated hullspeed equations use both the rudder angle and raw hullspeed sensor measurements as inputs.

SECTION 3.3.1 METHODOLOGY

Using test data from the catamaran propelled by a trolling motor, the hullspeed can be arrived at by averaging speed over an out and back round trip course, assuming that the current remains constant over that time. The assumption is that the velocity detected by GPS would be greater than the hullspeed when travelling with the current, but less than the hullspeed when travelling against the current. Thus the average of the round-trip would yield a point to calibrate the hullspeed sensor. In order to calibrate the hullspeed sensor, velocity derived from the GPS (over an out and back course) and the raw hullspeed output was recorded for several voltage settings on the trolling motor. This was done to generate a number of velocity points.

SECTION 3.3.2 CALIBRATION

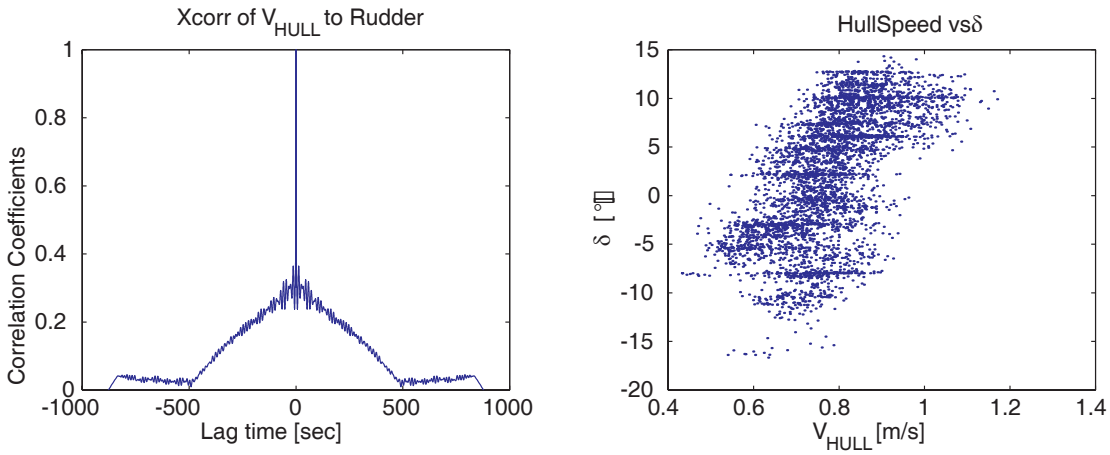

Figure 3-13 Cross-correlation of hullspeed versus rudder angle, and a plot of hullspeed versus rudder angle. The cross-correlation shows that rudder angle has a strong influence on the hullspeed measurement, and this can be confirmed by the slanted nature of the cloud of points in the hullspeed versus rudder angle plot. The calibration of the hullspeed sensor utilizes both the raw hullspeed sensor and the rudder angle as inputs to determine the true hullspeed.

The results of these calibration runs indicated that several large errors in the hullspeed sensor remained. Glauert's lifting line theory (simplified aerodynamics, [59]) suggests that the major contribution to errors in the hullspeed sensor would be an angle of

attack on the centerboards, or consequentially, any angle on the rudders. In order to verify this, the cross-correlation of the hullspeed reading and the rudder angle was computed and is plotted on the left in Figure 3-13. Additionally, the raw hullspeed versus rudder angle is plotted on the right. If there were no influence of the rudder angle on the hullspeed reading, there would be a symmetric cloud of points with a random distribution. Instead, a slant of the cloud is readily visible. As a first approximation, a linear fit (via least squares) is used to remove the influence of the rudder angle on hullspeed. While this removed the influence of the rudder on the hullspeed measurement, the measurement still proved to be quite noisy, and relatively unusable for current prediction. Therefore, another method was devised to predict current that did not rely solely on the hullspeed sensor.

SECTION 3.3.3 RESULTS

The final formula for calibrated hullspeed is given by:

$$V_{hull}\big|_{raw} = 0.0926\left(\frac{1}{\Delta t}\right) \quad \text{(EQ 3.9)}$$

and the final, true calibrated hullspeed is given by:

$$V_{hull}\big|_{true} = 1.486(V_{hull}\big|_{raw}) - 0.0145(\delta_{rud}) - 0.00045(\delta_{rud})^2 \quad \text{(EQ 3.10)}$$

where V_{hull} is in meters/second, and δ_{rud} is defined as the rudder angle in radians. This calibration is used to correct the raw hullspeed sensor reading and a typical pass produces the plot in Figure 3-14. Note that the noise level on the hullspeed sensor make detection of a current rather difficult. However, if the data is presented as a histogram with the best Gaussian fit through the distribution, the presence of the current is quite clear as the difference of the measured means of both distributions (Figure 3-15).

Thus the hullspeed becomes a rather secondary sensor, useful for confirming the presence of a current, but not useful for quantifying said current. Instead, crosstrack current is estimated using the measured azimuth versus the ground track, and using the mag-

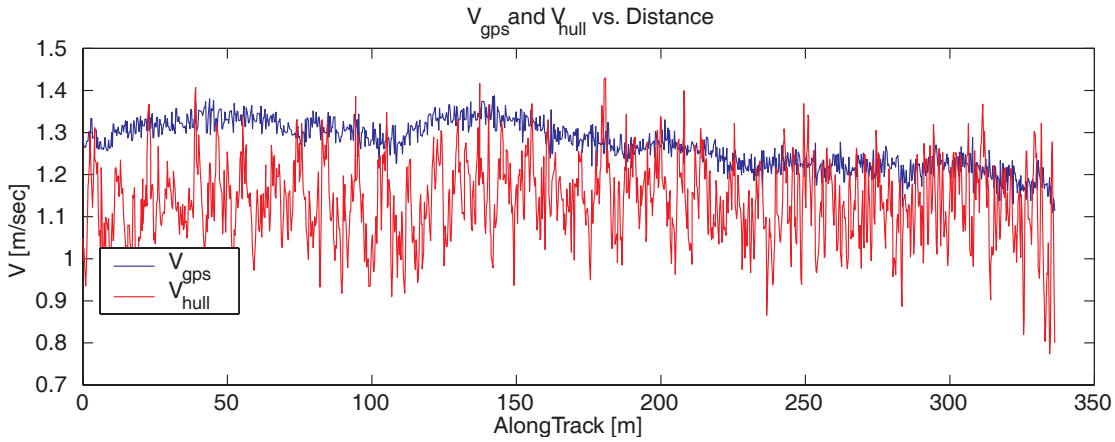

Figure 3-14 Calibrated hullspeed performance compared to GPS velocity. Note that the hullspeed sensor contains much more noise than the GPS velocity measurement. The vector difference between the hullspeed and the GPS velocity is the current. Unfortunately, extracting the value of the current is difficult with this level of noise on the hullspeed sensor.

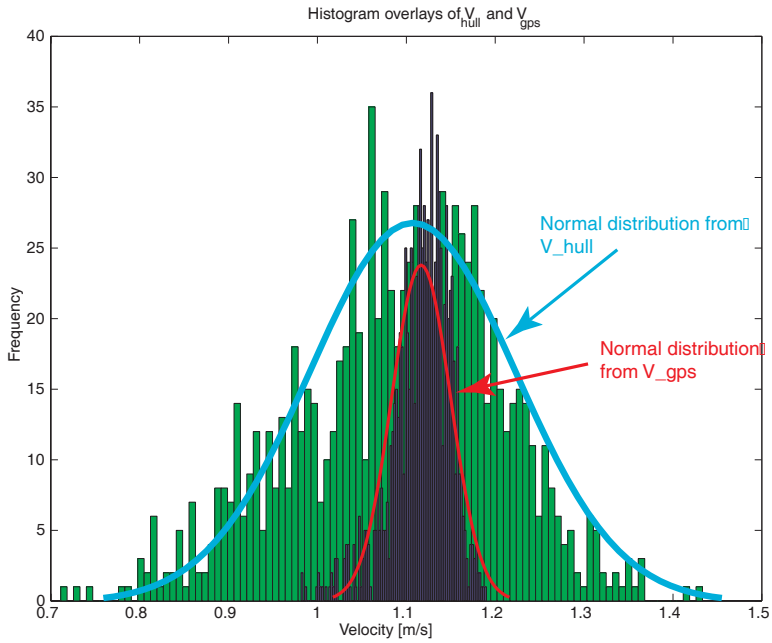

Figure 3-15 Histograms of hullspeed and GPS velocity of a trolling motor pass. Though the hullspeed distribution is much wider due to the increased noise levels on that sensor. However, the mean of the fitted distribution (cyan line) is definitely lower than that of the GPS velocity. This difference is the water current, and causes the bias shift in the previous plot.

nitude of the GPS velocity as truth. This yields only information of current perpendicular to the present path. However, because the control system regulates cross-track error the

perpendicular current is the only one of interest. Currents along track either slow progress along the desired line, or increase the velocity travelled along the line. In either case this does not induce errors in the line tracking performance.

SECTION 3.4 RUDDER ANGLE

The rudder angle is a crucial measurement in terms of controlling the Atlantis. Based on the simple kinematic model presented in Section 6.2 on page 251, it is one of the three required states. This data in turn is necessary to determine the identified model. Essentially the rudder angle measures the integrated actuator output. While it is possible to operate the Atlantis without the rudder angle sensor (through the use of an estimator) it would operate at reduced capacity, in the sense that Sailboat Technical Errors (STE) would be larger. The sensor used to measure the rudder angle is a LoHet, an analog magnetic field sensor that produces a voltage in proportion to the magnetic flux across the plane of the sensor. Schematically, the LoHet is shown in Figure 3-16.

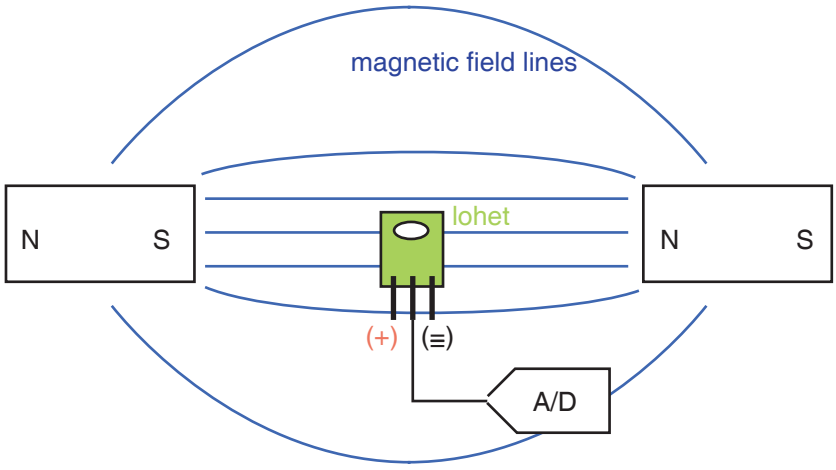

Figure 3-16 LoHet schematic diagram. The LoHet sensor produces an output proportional to the magnetic field flux across the sensing element. Thus, by placing the sensing element within the gap between two magnets, a signal proportional to sine or cosine of the angle can be extracted. The advantage of the LoHet is that the magnetic field is unaffected by salt-water submersion, and the output is quite linear.

SECTION 3.4.1 METHODOLOGY

Calibrating the rudder angle LoHet is simply a matter of using a set of architectural triangles to set the rudder at various angles with respect to the flat starboard hull end and record the output data. After taking several points of data, the raw calibration data is plotted, and a quadratic fit proves to be better than a linear fit for extrapolating the true angle of the rudders. The raw data is pictured in Figure 3-17, and the various fits to the raw data

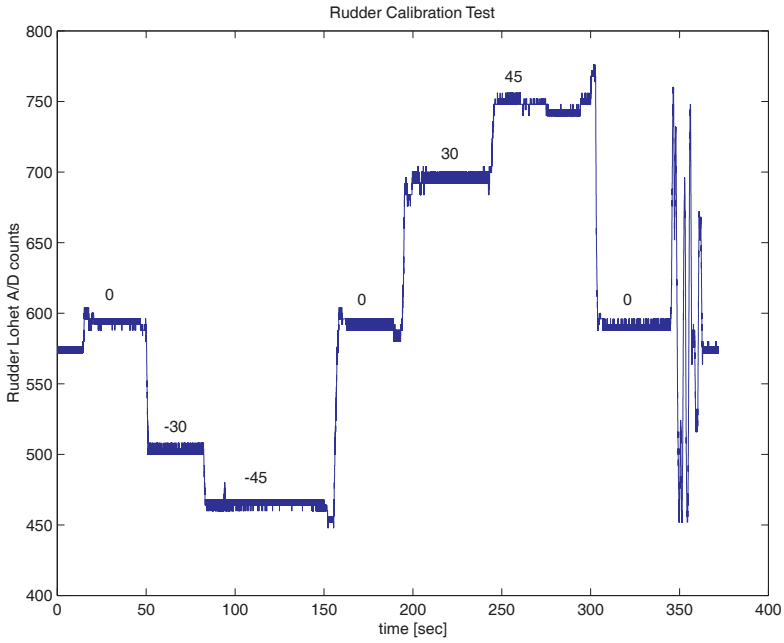

Figure 3-17 Raw rudder calibration data recorded using architectural triangles to set the data. Close inspection of the data reveals that the LoHet counts oscillate within +/- one quanta of the true value, resulting in the thick lines. The end of the data run includes some fast oscillations to check the tracking of the LoHet sensor.

are pictured in Figure 3-18.

It can be seen that a quadratic fit is much better, in this case, than the standard linear fit. The quadratic fit is used with the caveat that this might be overfitting the data as the number of distinct angles used is small. Because there are only five distinct angles, a fifth

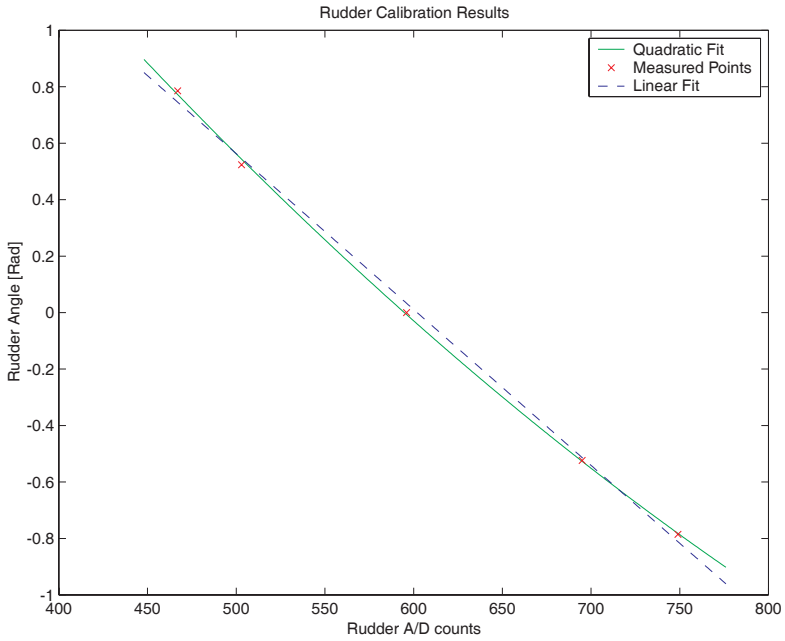

Figure 3-18 Rudder calibration linear and quadratic fits to the data. The quadratic fit produces a better match to the data without too many degrees of freedom in the matching equation. Higher order polynomials better match the data points, but tend to be erratic in between the points and when extrapolating beyond the measured data. The quadratic fit is a good compromise between complexity and accuracy.

order fit would exactly match the five points, but would be highly erratic outside of the points themselves. Thus, the second order fit is considered to be a favorable compromise.

SECTION 3.4.2 CALIBRATION

The final formula for the calibrated rudder angle is:

$$\delta_{rud} = 3.4518\times10^{-6}(A/D)^2 - 0.0097(A/D) + 4.55423 \qquad \text{(EQ 3.11)}$$

where δ_{rud} is in radians, and the A/D count is a positive integer number from 1 to 1023 (10-bit A/D). However, this process only calibrates the geometric angle of the rudders; the effective rudder angle remains unknown. The effective rudder angle is the azimuth rate in still water which should correspond to the geometric rudder angle. In order to determine

the effective rudder angle and bias, a series of constant radius turns was performed with the catamaran on the water (Figure 3-19).

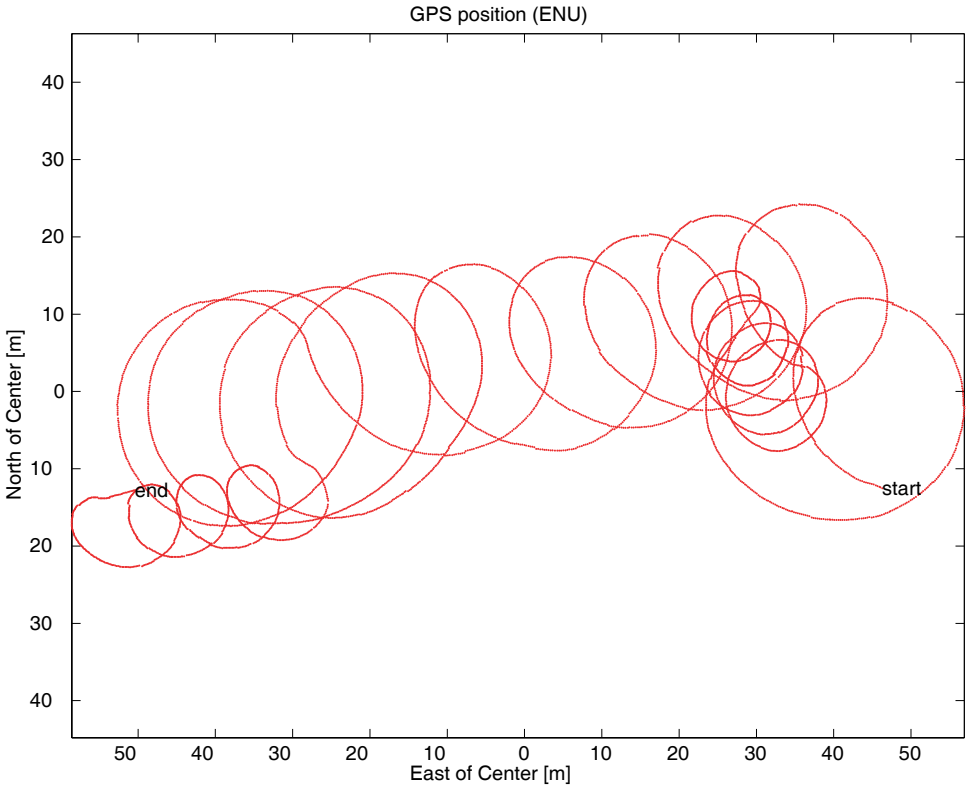

Figure 3-19 Typical Atlantis trajectory as viewed from above for testing effective rudder angle. By making turns of constant radius, and removing the drift due to the current, the bias and scale factors that relate the geometric rudder angle to the effective rudder angle can be computed. Note that turns were performed in both clockwise and counterclockwise rotations, and that the current direction varies. This technique is referred to as calibrating "set and drift."

A simple least squares estimator is set up to remove the effects of current, which is seen to be dragging the path into a spiral instead of a perfect circle. Once the current is removed from the data, the effective radius and azimuth rate can be computed and plotted. This data, shows that the geometric angle is an excellent measure. It is, however, corrupted by a very small bias in the system. Thus,

$$\delta_{rud}\big|_{effective} = \delta_{rud}\big|_{geom} + 0.012153 \qquad \text{(EQ 3.12)}$$

where both the effective rudder angle and the geometric rudder angle are measured in radians. Note that the small bias is easily accounted for by hull misalignment or rudder toe-in (as there is a rudder on each hull).

SECTION 3.5 GPS RECEIVER

The GPS receiver is the Atlantis' primary sensor, giving position, velocity, and bearing (as opposed to heading) information. It is a Trimble Ag122 differential GPS receiver with an integrated H-field antenna that receives the United States Coast Guard differential corrections. The International Light House association maintains a series of low-frequency (300 KHz) beacons around the coastal waterways of the entire world. In the United States, these beacons are installed and maintained by the United States Coast Guard. The entire coast of the continental United States and all of its navigable waterways are within range of one or more of these beacons.

In the case of the GPS receiver no calibration is possible due to the fact that the differential corrections are being provided by the United States Coast Guard. That is, an independent measurement of either velocity or bearing that is more accurate than GPS is very difficult and expensive to use for this purpose. It is possible, however, to quantify the noise statistics on the position measurements so that the controller can be fine tuned based on the observed properties.

SECTION 3.5.1 METHODOLOGY

Two tests were performed in order to quantify the noise statistics of the GPS receiver. The first was a typical static survey test. The GPS receiver was set to use the local Coast Guard corrections and data was recorded over a 24 hour period. This test was

performed on the roof of the Hansen Physics Laboratory (HEPL), at Stanford University, in order to remove possible effects of multipath due to reflectors near the antenna. By locating the test above the height of most buildings the errors due to multipath were expected to be minimal.

The second test was performed to quantify the results of the GPS receiver in a dynamic case while the antenna was under motion. In order to do this, a spin rig was constructed allowing the GPS antenna to rotate through a 360 degree arc in azimuth without winding up wires. This spin rig has a diameter of approximately 1.4 meters. The GPS receiver and antenna were mounted on the spin rig and spun around repeatedly. The noise statistics are based on a best fit circle to the data and a histogram of the radial error.

SECTION 3.5.2 EXPERIMENTAL SETUP

To generate error statistics for the GPS receiver, the static test was performed on the roof of the HEPL building for a period of 24 hours. The Trimble Ag122 GPS receiver was set to record data once every second and the data was streamed to a file. Figure 3-20, shows the static setup of the antenna and receiver on the roof.

In order to test the GPS receiver while under motion a spin rig was built. The motion in question is to rotate the antenna around at the end of an arm (i.e, a merry-go-round) while recording data. Thus, the GPS antenna was placed on the end of an arm, and rotated around in azimuth to gather data. The entire setup was balanced on a lazy susan bearing and counterweighted by a battery so as to be mass balanced about the pivot axis of the bearing. The GPS receiver was set to take data at 5 Hz, using differential corrections supplied by the Coast Guard beacon at Pigeon Point (located approximately 40 nautical miles away). The entire setup was battery powered, thus no slip rings were required. The entire spin rig was spun by hand, with an attempt to keep a fairly constant velocity of about one revolution every 5-7 seconds (Figure 3-21). Further refinement of this experi-

Figure 3-20 GPS setup for static test. The GPS receiver was located on the roof of a building on Stanford University campus. The roof was chosen to be above the surrounding buildings in order to minimize the possibility of multipath errors. Data was recorded once per second for 24 hours. This was used to determine the static position statistics for this GPS receiver.

ment is possible through the use of a motor to provide a precise rotation rate, but it was deemed unnecessary for this work.

SECTION 3.5.3 RESULTS

The data gathered from the static test indicated that for a 24 hour period the errors look very white. Figure 3-22 shows the plot of the points obtained from the 24 hour period, as well as the histogram of East position errors. The measured error standard deviations (1-σ) are 0.36 meters in East position, 0.34 meters in North position, and 0.42 meters in Up position. Because the firmware in the GPS receiver turns off the velocity measurement below a certain threshold (known as position-velocity filtering), the raw

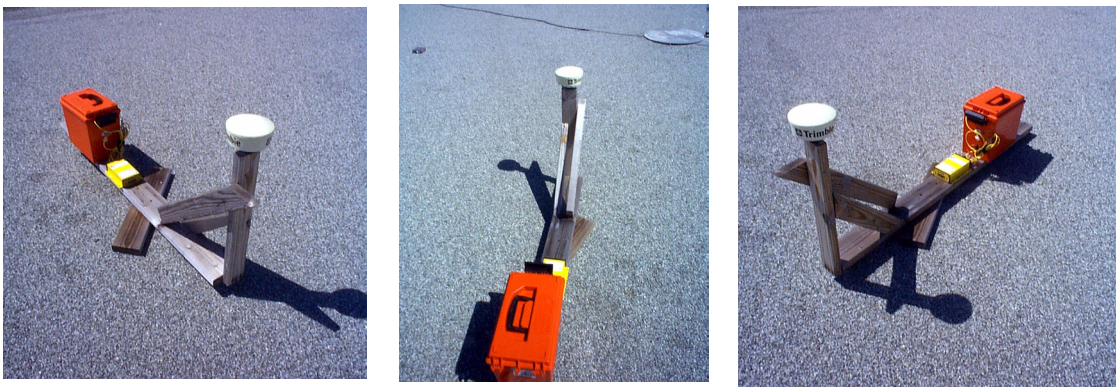

Figure 3-21 The GPS dynamic spin rig. This spin rig was used to test the GPS position performance under dynamic conditions. The spin rig is a merry-go-round, balanced on a lazy susan bearing. The yellow box is the receiver, and the white object on the post the antenna. The entire system is battery powered, obviating the need for slip rings. The rig was spun by hand, attempting to keep a steady angular rate of one revolution every 5-7 seconds.

velocity errors are not available for a static test. However, by subtracting the successive position error measurements and dividing by time (numerical differentiation), it is possible to arrive at a mean drift rate for the static case, which was found to be approximately 0.0015 meters/second. With this drift rate, it can be safely predicted that the dynamics of the catamaran will be significantly faster such that the controller will respond to the set point motion smoothly and slowly. Note that this analysis assumes that the mean of the positions for 24 hours represents the true position as no alternative means for positioning the GPS antenna is available.

The results for the dynamic case are more difficult to interpret. The position is spread in a circular pattern. The mean of the best circular path is taken to be the actual measurement of the true diameter of the spin rig. Once the spin rig diameter circle is removed from the data (transforming to polar coordinates and then back again after the mean radius has been removed), the usual statistics are compiled. However, the measurement has been squared and square-rooted, causing a distortion in the statistics. Certainly, at this point, the distribution is Chi-squared rather than the usual Gaussian. In order to present the data in familiar terms, the standard deviations are multiplied by the square root

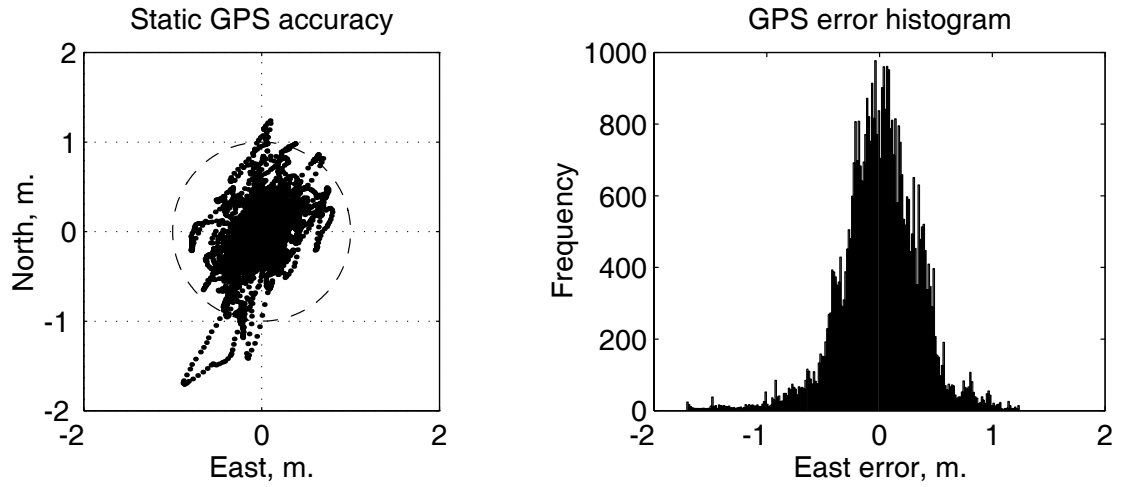

Figure 3-22 GPS errors from static test. The left plot shows the raw North and East data. The dotted circle is a circle of one meter radius. Thus, the standard deviation of the errors are 0.36 meters East, 0.34 meters North, and 0.42 meters Up. The histogram of the East errors on the right shows that the errors look Gaussian. The assumption of Gaussian error models appears justified based on this data.

of two in order to correct for the distortion. Due to the fact that velocity about the circle was controlled rather crudely, no direct measurements of the velocity statistics could be made. A histogram of the velocity shows that on average, the speed was approximately 0.6 meters/second (Figure 3-23).

The position information is taken to be the variation in radial position, as the angular position is a result of the spinning of the boom. In Figure 3-24, the circular path of the antenna is shown very clearly. Transforming this to polar coordinates, the histogram of the radial position can be plotted, as shown in Figure 3-25.

By transforming the spin rig position data into polar coordinates, the mean of the radius can be seen to be 0.7009 meters or very close to half of the measured diameter of 1.4 meters. The next step is to remove the radius and transform the measurements back into the East-North-Up coordinate frame, as displayed in Figure 3-26.

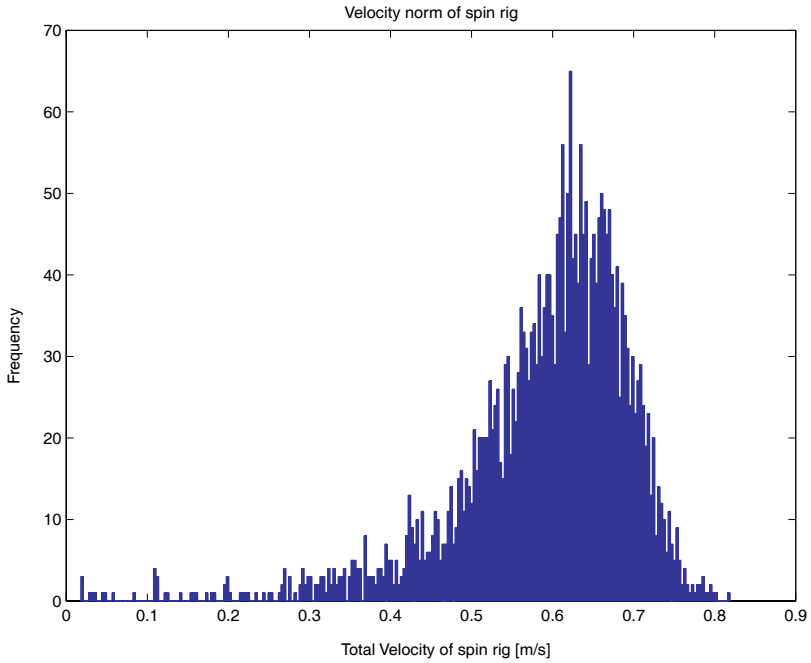

Figure 3-23 Histogram of GPS spin rig velocity. This is the magnitude of the velocity vector presented as a histogram. The spin rig was spun by hand, and an effort was made to keep the velocity constant. The resulting velocity (and later radius) corresponds to an angular rate consistent with the desired rate of one revolution every 5-7 seconds.

The last step is the calculation of the standard deviations of the remaining East and North positions. As observed from the histograms of the residuals, the distributions are no longer Gaussian, but rather Chi-squared due to the squaring and square-rooting that take place in the Cartesian to polar to Cartesian transformations (Figure 3-27).

Thus, the statistics for the East and North residual data (after radius removal) are 0.109 meters and 0.115 meters, respectively. However, if the results are multiplied by the square root of two, as suggested by the Chi-squared-like distribution, the pseudo-East standard deviation is 0.33 meters and the pseudo-North standard deviation is 0.34 meters corresponding very closely to the static test.

Lastly, it is interesting to note that the velocity averaged 0.67 meters/second, with a radius of 0.7 meters. This corresponds to a rotational speed of approximately 8 revolu-

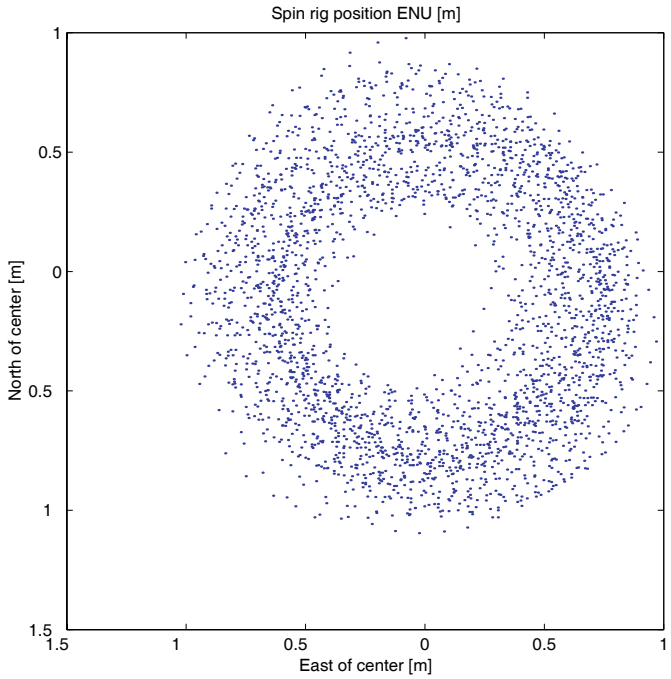

Figure 3-24 GPS spin rig position data in ENU coordinates. The GPS data was sampled five times per second, and the data plotted as seen from above. The presence of the circle centered within the data is clearly evident, as is the radial error spread. This will form the basis of the statistical analysis of the dynamic GPS data.

tions per minute, or a rotation every 7.5 seconds, which is quite close to the original estimate. Because the GPS data is filtered, one would expect to see an amplitude roll-off and a phase distortion as the frequency of rotation increased. The rotational frequency of the spin rig is much less than the frequency at which the amplitude distortion would be visible. Further experimentation could give an interesting view of the dynamics of the internal GPS filter.

SECTION 3.6 RUDDER ACTUATOR

The most difficult, and crucial component to linearize is the rudder actuator. As the primary actuator for the control system, the majority of the system input is injected into

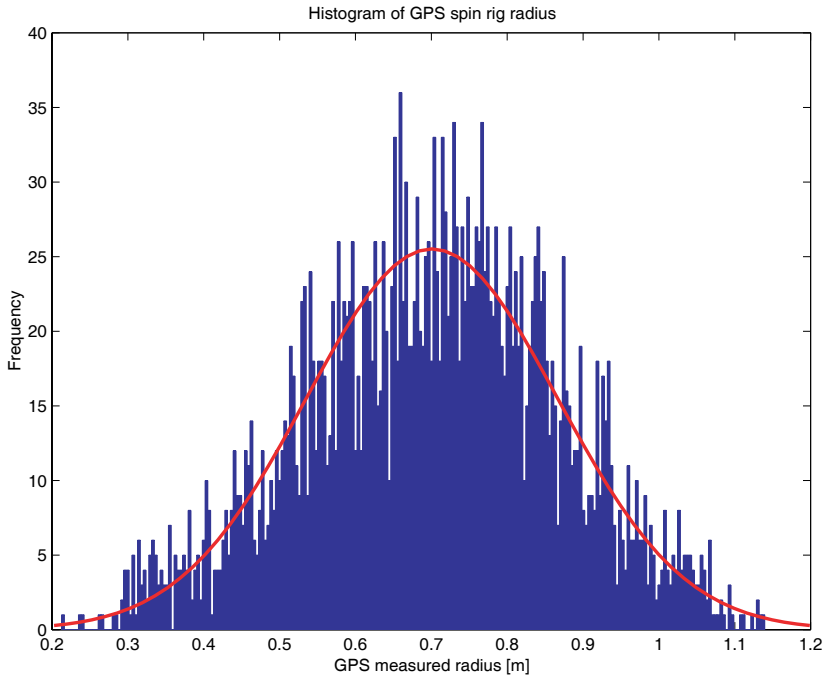

Figure 3-25 Histogram of GPS spin rig radial position data. The histogram of the data shows clearly the centered circle is approximately 0.7 meters in radius. This corresponds very well to the measured diameter of the circle as 1.4 meters. Note that the errors in radius look, and can be assumed, to be Gaussian.

the plant through this actuator. Thus any non-linearity left in the actuator will make system identification problematic at best. In addition, the actuator, by virtue of its construction, possesses some non-linearities that prove challenging to remove.

Basically, the non-linearities of the rudder actuator are too difficult to remove through either mechanical or simple calibration techniques. Instead, the solution consists of mapping of the non-linearities and using this to “inverse-map” the actuator, thereby removing the visibility of the non-linearity from either the system identification algorithms or the control system itself.

As long as the class of non-linearities can be cast into the form of repeatable, memory-less, mappable errors, the inverse mapping technique works flawlessly. In the case of the rudder actuator, this includes stiction (static friction) and backlash (hysteresis) in the

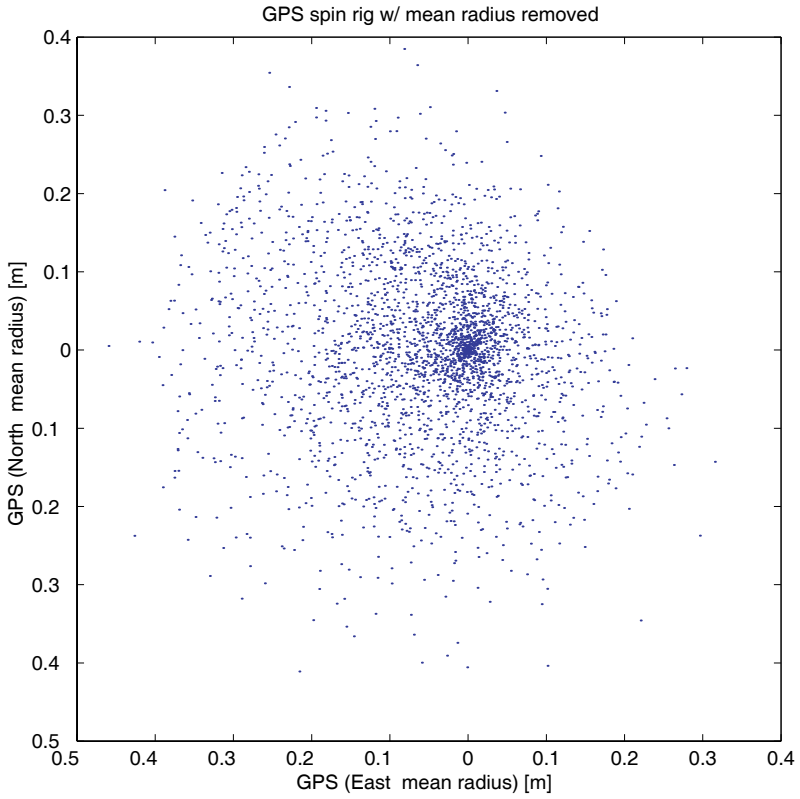

Figure 3-26 GPS spin rig position data with the mean radius removed. The previous GPS spin rig "donut" data has been transformed into polar coordinates, then the mean of the radius was removed from the data. Following this, the data was re-transformed back into cartesian coordinates. This allows the normal statistics to be compiled from the residual data.

lead screw assembly. Geometry changes in the lead screw to rudder bar linkage account for additional non-linearities.

SECTION 3.6.1 METHODOLOGY

In order to map the non-linearities of the rudder actuator, in essence creating a map from computer input value to rudder angle slew rate, a series of water trials was performed with the rudder slewed back and forth at different rates. The pulse width modulation (PWM) rate for the motor, controlled by a 16 bit timer in the microcontroller, was set to overflow every 26667 clock cycles (yielding a nominal PWM frequency of 100 Hz).

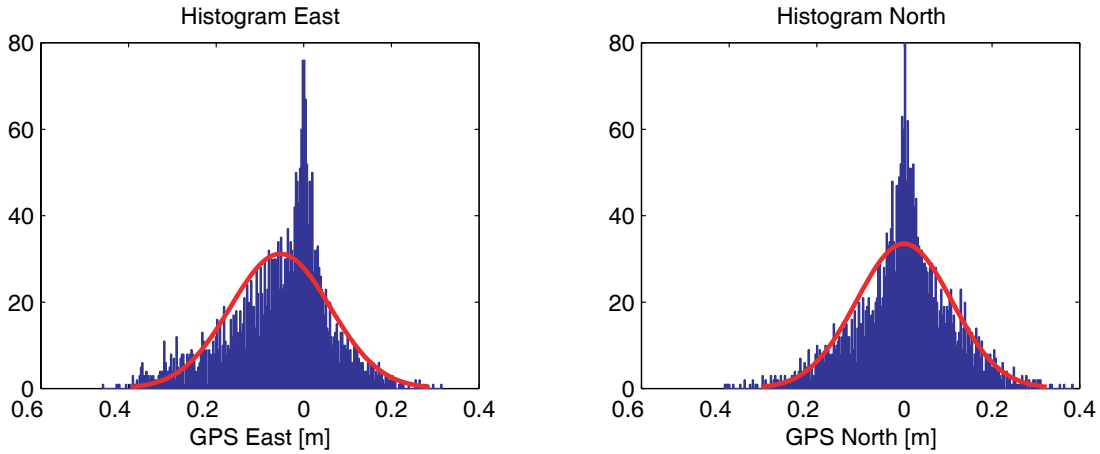

Figure 3-27 Histograms of GPS spin rig East and North data with mean radius removed. The red line is the best fit gaussian distribution to the curves. Note the excessive spikes in the data, typical of the Chi-squared distribution. This is due to the squaring and square-rooting that occurs in the transformation between polar and cartesian coordinate. In order to present the relevant statistics in a comparable form, they must be multiplied by the square root of two. The pseudo-standard deviation is 0.33 meters East, and 0.34 meters North.

Therefore there are 26666 possible values to use for PWM. Since mapping all of them would be far too time intensive, a subset of 50 values, corresponding to 2% increments, was used to map the actuator performance.

Figure 3-28 shows a typical path followed during one of several trials to map the actuator. Note that the PWM values are used both for port and starboard rudder slew directions. The odd path in position is due to the essentially slow oscillation of the rudders, induced by the controller to map the rudder actuator response. As the path is essentially the distance double integration of the rudder angle, no pattern can be seen even though the rudder slew rate is essentially sinusoidal.

SECTION 3.6.2 CALIBRATION

Using the recorded data from the rudder actuator mapping trials, the PWM value and the rudder angle are recorded and time-tagged according to GPS time. All the information needed to map the rudder actuator are present in the logged files: the PWM value,

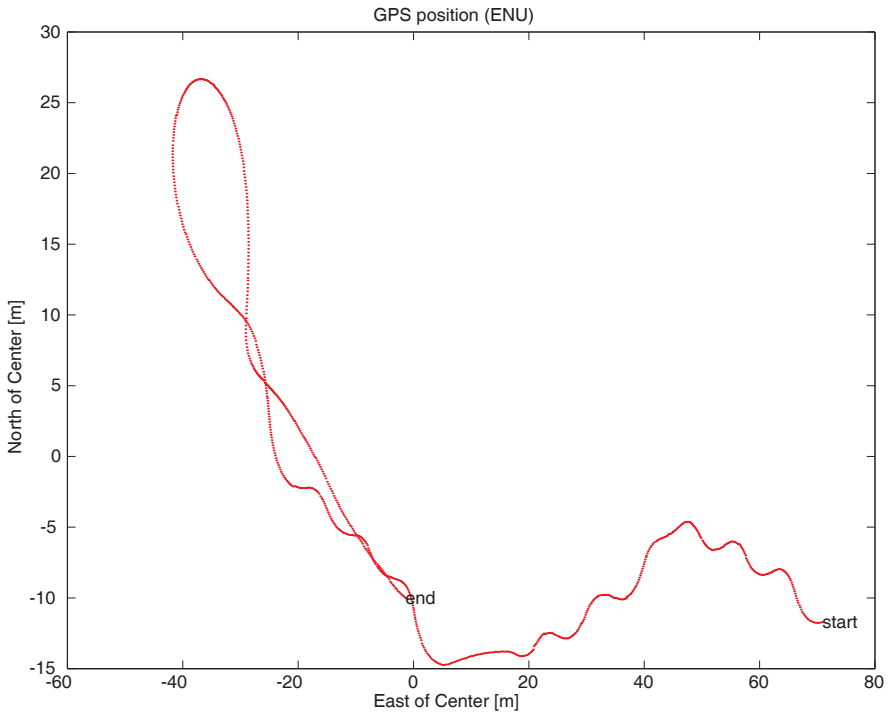

Figure 3-28 Typical catamaran path during rudder actuator mapping trials. The rudders were slewed to port and starboard at specific rates to map out the rudder actuator non-linearities. The odd path is due to the slow oscillation of the rudders due to the input commands. This is one of several passes used for the actuator mapping.

the rudder angle, and time. A method to estimate the rudder angle slew rate from the plots of the rudder angle versus time is required. Due to the nature of these experiments there was a tremendous amount of data to be reduced. Thus, fitting the slope of the rudder angle versus time line by hand was not an option. Below, the PWM and rudder angle data from the pass pictured in Figure 3-28 is displayed in Figure 3-29.

Software written in the MATLAB environment was used to process the numerous control passes. A least squares estimate of each slew rate section was used to extract the slew rate and match it up with the appropriate PWM value. Note that the PWM value is assigned a positive (+) value for slew rates to starboard, and a negative (-) value for slew rates to port. This is based on the North-East-Down (NED) coordinate system used for attitude computations (see Chapter 4). An enlarged section of the plot above is shown in

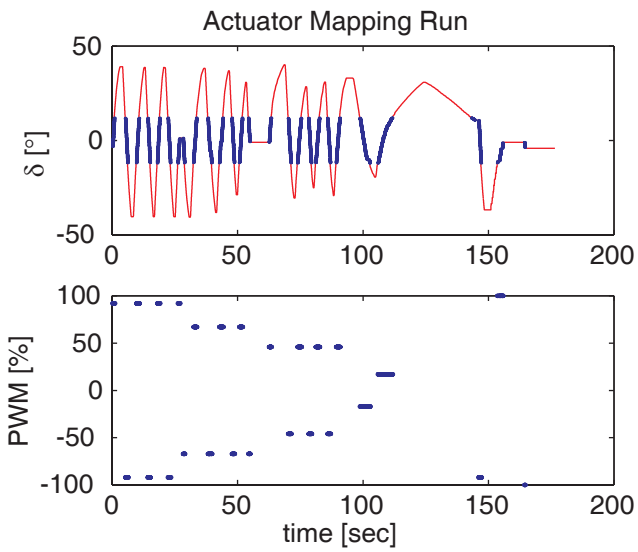

Figure 3-29 PWM and δ for the rudder actuator trial. The slew rate for the rudders in plotted in the top plot, with the blue sections of the plots used in the automatic slope calculation. Due to the very large amount of data, the process of reducing the rudder angle to rudder slew rate was automated in MATLAB. The lower plot shows the corresponding PWM command with starboard direction defined as positive.

Figure 3-30, along with the recorded values for the PWM normalized and expressed as a percent.

Based on evaluating the slew rate performance of the rudder actuator, a map is generated from all of the data points. That map is shown in Figure 3-31, which plots the PWM value expressed as a signed percentage versus the rudder angle slew rate in degrees/second. The actuator is capable of deflecting the rudder at over 25 degrees/second, allowing the rudder to traverse from stop to stop in a little over 3 seconds. One of the first things to note is that there is a rather large dead band in the middle of the plot. This occurs because at very low PWM settings, not enough torque generated by the motor armature to overcome the friction in the rest of the actuator system. Thus no motion occurs. Furthermore, both the positive and negative branches of the PWM map are non-linear, appearing to be of cubic or higher order. Also, though they appear very similar, the positive and negative branches are not symmetrical and each requires a separate polynomial to map.

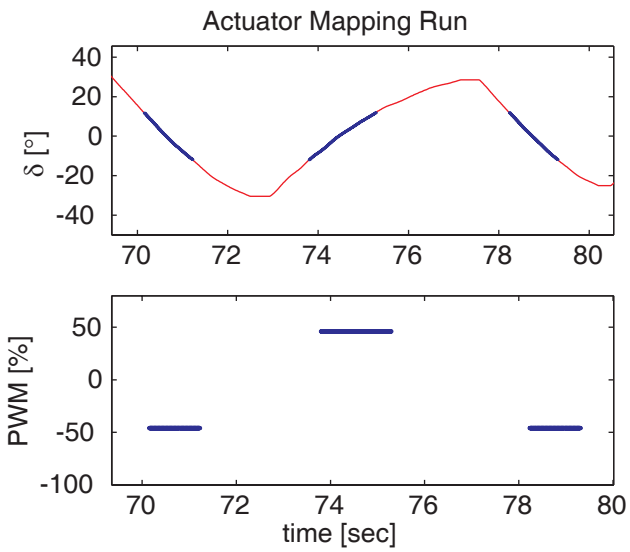

Figure 3-30 Enlargement of the PWM vs. δ mapping for rudder actuator mapping. The blue dots are used to estimate the slope of the line (the rudder angle slew rate) and the lower plot is the PWM signal reduced to a percentage. These points are assembled into a large meta table that is used to map out the rudder actuator non-linearities.

SECTION 3.6.3 RESULTS

Based on the data in Figure 3-31, the mean lines in a least squares sense for both sides (negative and positive) of the PWM response are plotted. The mapping equations are:

$$\dot{\delta} = 0.41692\times10^{-4}(PWM)^3 + 0.01130(PWM)^2 + 1.0627(PWM) + 8.188747 \quad \text{(EQ 3.13)}$$

for PWM values of less than -10% and,

$$\dot{\delta} = 0.39232\times10^{-4}(PWM)^3 - 0.010307(PWM)^2 + 0.9832364(PWM) - 0.9198531 \quad \text{(EQ 3.14)}$$

for values of greater than +10%. Note that this maps the PWM to rudder angle slew rate in radians per second.

An inverse mapping has also been developed. This inverse mapping is required because the control output will be rudder rate in radians/second, and this value will be sent along the network to the rudder actuator microcontroller. The rudder actuator microcon-

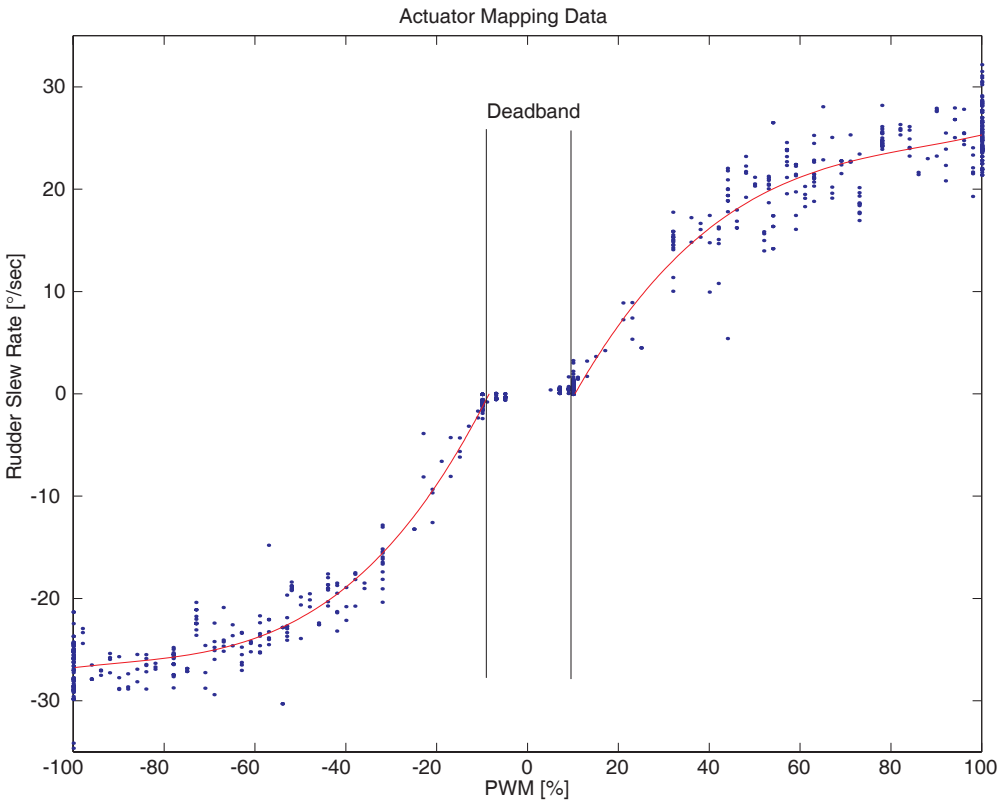

Figure 3-31 Full experimental rudder actuator non-linearity mapping. Several interesting features are present. The blue dots represent experimental data, and the red lines the polynomial curves fit to that data. Note the deadband between +/- 10% PWM. Also note that there are two separate cubic fits, one for positive rudder angle slew rates, and another for negative.

troller, however, must map this value back into a PWM value from its available 26666 possibilities. Equation 3.13 and Equation 3.14 are used to report back to the control software on the GNC computer the actual rudder slew rate achieved. The inverse mapping is based on a table lookup of precomputed values due to the difficulty of iterating for roots of a fourth order polynomial equation with an 8-bit fixed point microcontroller. The final mapping for the rudder actuator is depicted graphically in Figure 3-32.

Thus the rather gross non-linearities of the rudder actuator are mapped back into a linear block through a transformation that simply checks if bounds are outside the dead-

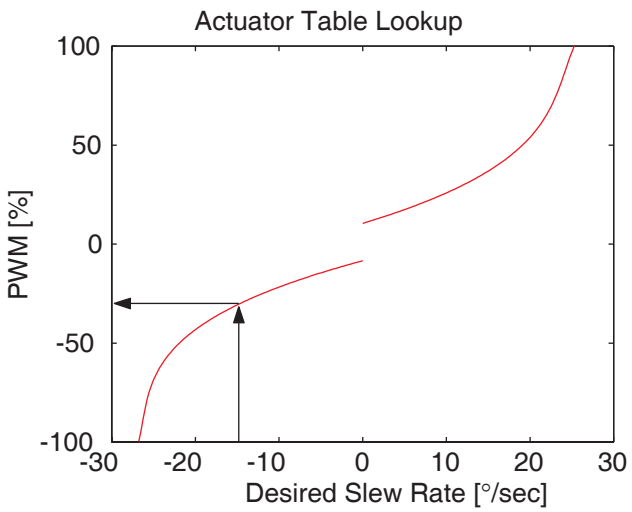

Figure 3-32 Rudder actuator table lookup function. This function is the inverse mapping of the rudder actuator non-linearities. This is done so that the output of the control system can be sent as a real value to the actuator subsystem. The actuator subsystem receives the desired value of the rudder angle rate (in radians/second) and converts this to a PWM value based on the above table. The example shown is for a -15 degree/second rudder angle rate which corresponds to a PWM value of -28%.

band, and then applies one of two cubic functions to return the correct value for the PWM value based on the sign of the desired slew rate.

Note that this type of inverse mapping is a technique is not necessarily the most robust technique to linearize the actuator. As the motor ages, temperatures increase, and the original calibration drifts, the actuator will have to be periodically recalibrated lest the inversion technique no longer work. In practice, this kind of calibration is rarely used outside of research prototype systems. Instead, a feedback loop is wrapped around the motor and encoder making it a servo-motor with a defined time lag. The performance penalty induced by the motor lag is accepted as the price to pay for an actuator that will continue to operate linearly over its lifetime.

4 Attitude System

Attitude determination is a requirement for most navigation and control problems. Classically, this problem has been solved by an Attitude Heading Reference System (AHRS) using gyroscopes that are updated by gravity sensors (pitch and roll) and magnetic field sensors (yaw) with low-pass filters to attenuate errors incurred during turns. Rate gyroscopes are prone to bias- and random-walk errors resulting from the integration of wide-band noise. Successful AHRS requires very expensive sensors that have exceptional bias stability. The sensor cost for this kind of attitude determination has limited such AHRS to very expensive applications. In cost-sensitive applications, filtering techniques are applied to bound the attitude error growth when using lessor sensors. With the recent

proliferation of low cost inertial sensors along with position and attitude sensors like the Global Positioning System (GPS), it has become viable to construct inexpensive attitude determination systems which have attitude errors that are bounded in time.

Recently, a considerable amount of effort has been directed at developing low cost systems for attitude determination. For example, [83] discusses an inexpensive attitude determination system for aviation applications. This attitude determination system does not employ any inertial sensors. Instead, it relies on a kinematic model of the vehicle along with GPS position and velocity measurements derived from a single GPS antenna to generate what is termed "pseudo-attitude." A different approach is taken by [63], where a triad of inexpensive automotive-grade rate gyroscopes are fused with an ultra-short baseline GPS attitude determination system. In this system, the high bandwidth attitude information derived from integrating the output of the rate gyroscope was blended with the low bandwidth GPS attitude solution using a complementary filter. However, because the rate gyroscopes used possess poor bias stability, the GPS attitude system was used to continuously update the estimate of the gyro biases. In order to further reduce the cost of this attitude system, an offshoot was created with a single GPS baseline for yaw, with roll and pitch obtained by a non-linear acceleration vector matching [64]. This work is further complemented by [65], where the use of magnetometers is offered as an alternative to the single GPS baseline for yaw. This can be seen as a precursor to the work contained in this thesis, with the primary differences being a more formal mathematical basis and an integration of the magnetometer triad during accelerated motion.

Vector matching in multiple coordinate frames is often referred to as Wahba's problem. First published in 1965 [144], Wahba proposed an attitude solution by matching two non-zero, non-colinear vectors that are known in one coordinate frame, and measured in another. Several solutions to this method of attitude determination have been proposed and implemented [10], [11], [73], usually on satellites with star-tracker sensors. In this

chapter, a novel method of solving Wahba's problem is demonstrated, in which the measurement equation can be cast into a standard form. By using the quaternion representation of attitude, the algorithm does not require the solution of transcendental equations (as would be the case with Euler angles) and is suitable for implementation in a small, inexpensive microcomputer. Specifically, by measuring the body-fixed magnetic field and the body-fixed gravitational acceleration, the attitude of a moving object is tracked both in simulation, and experimentally with post-processed data.

SECTION 4.1 MOTIVATIONS

The need for a complete three-axis attitude reference unit on a marine surface vessel is not immediately obvious. However, there are a few pressing reasons apparent upon closer examination. Since heading is sensed magnetically, the roll and pitch can alter the measured heading by as much as 10-15 degrees under normal operating conditions. This is due to the downward component of the magnetic field (dip) corrupting the horizontal components of the measurement. This precludes accurate heading information which is a requirement for high precision control. Further discussion can be found in Chapter 6.

The other motivation comes from the unique lever arm problems associated with GPS control of a marine vessel. Ideally, the GPS antenna would be placed on the highest point of the craft, yielding an unobstructed view of the sky. This would give the best view of the satellites and minimize multipath effects. However, the point of the vessel that needs to be controlled, and hence its position and velocity sensed, is the center of mass of the vessel. Only on a simple point mass do the center of mass and the highest point coincide. Thus, the position of the center of mass will always be the sensed position of the antenna corrected for some lever arm position. In the case of a sail-propelled vehicle, the ideal location for the GPS antenna is at the top of the mast.

A synthetic sensor is created by fusing the GPS position data with the known lever arm and attitude data to place a virtual antenna at the center of mass of the vehicle, which can be below structures that would normally obstruct the GPS signal and can even be below the water. Without this synthetic sensor, roll and pitch of the vessel in waves can cause the antenna to swing with the mast. This creates phantom inputs into the control algorithm that would lead the boat to turn into the heel, further exacerbating the problem. Furthermore, this can make a non-minimum phase situation occur where control inputs first cause a direct excursion from the desired control path. Non-minimum phase system possess the disadvantage that optimal control methods are always band-limited due to initial response in the wrong direction.

SECTION 4.2 REFERENCE FRAME REVIEW

The attitude of a vehicle is the difference between the navigation frame, usually fixed to Earth at the location of the vehicle in a traditionally North-East-Down coordinate system, and a coordinate frame fixed to the body of the vehicle (Figure 4-1). Traditionally, the body frame has the x-axis pointed out the nose of the vehicle, the y-axis out the right side, and the z-axis down.

Specifically, attitude describes the relative orientation of two (usually orthogonal) unit vector triads. The attitude simply describes what rotations one must perform to move from one description to the other. The reference frame and body frame are used to define a reference attitude to call zero on the various parameterizations. For instance, traditionally, the North-East-Down (NED) reference frame attached to Earth is considered as the inertial frame. While this works perfectly well for an aircraft flying a few hundred miles, this frame quickly loses its utility as a reference if the object of interest were a satellite. For a satellite, one might desire to know the attitude relative to a coordinate frame that tracks the

satellite and has its x-axis pointing along the orbit, its z-axis pointing towards the center of Earth, and the y-axis pointing out of the orbit plane.

Alternately, a coordinate frame that is attached to the center of Earth and points towards a distant star might be useful to measure effects on a satellite that is in a heliocentric orbit. This coordinate frame would be difficult to use for an interplanetary probe, which would probably be better served by a sun-fixed frame, pointing out towards galactic center.

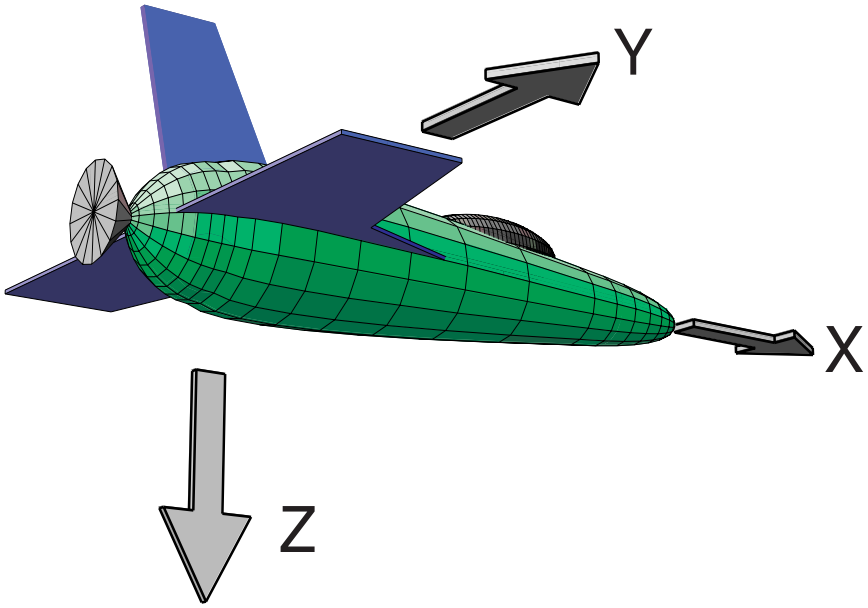

Figure 4-1 The body frame rigidly attached to the vehicle. This is the standard aircraft coordinate system definition. X-axis out the nose, Y-axis out the right wing, and Z-axis down. The body frame remains fixed with respect to the vehicle. Other frames are used as appropriate, for instance, an inertial frame attached to Earth, with its X-axis pointing North, Y-axis pointing East, and Z-axis pointing down. Transformations between these coordinate frames are used to keep values correct in their respective frames.

A reference frame is simply that: an orientation to call zero, something relative to which attitude may be measured. In the case of the attitude problem discussed here, it is convenient to use the North-East-Down frame and refer to this as the navigation frame. Reemphasizing, a body frame is used that is rigidly attached to the vehicle, with the x-axis

pointing out the nose, the y-axis out the right wing or starboard side, and the z-axis down (as shown in Figure 4-1).

SECTION 4.3 ATTITUDE PARAMETERIZATION

Having clarified that attitude is simply the relative orientation of one frame to another, there are several methods used for describing the relative orientation between frames. The three main parameterizations for attitude are: *Euler angles*, *Direction Cosine Matrix*, and *quaternions*. Several textbooks such as [86], [80], and [60], provide an excellent background that is reviewed here only for completeness.

Euler angles are the most intuitive parameterization. Three angles, yaw, pitch, and roll, are used to describe the attitude. Starting with the vehicle aligned to the navigation frame, the vehicle is first yawed about the z-axis, as in Figure 4-2 (top frame). Then the vehicle is pitched about the new body y-axis (Figure 4-2 middle frame), and finally the vehicle is rolled about the resulting x-axis (Figure 4-2 bottom frame). This is referred to as a [3-2-1] Euler angle set due to the order of rotation axes, and is made unique by limiting yaw to +/- 180 degrees, pitch to +/- 90 degrees, and roll to +/- 180 degrees.

The parameterization of attitude is used in equations to transform vector quantities from one coordinate frame to the other (the so-called transformation matrices), and in the propagation or integration of attitude based on some known or measured rotation rate. The generalized equations for a rigid spinning body, known as Euler's equations, are independent of parameterization of attitude, and are expressed as:

$$\frac{d}{dt}\omega = [I]^{-1}[M - [\omega\mathrm{x}]I\omega] \qquad \text{(EQ 4.1)}$$

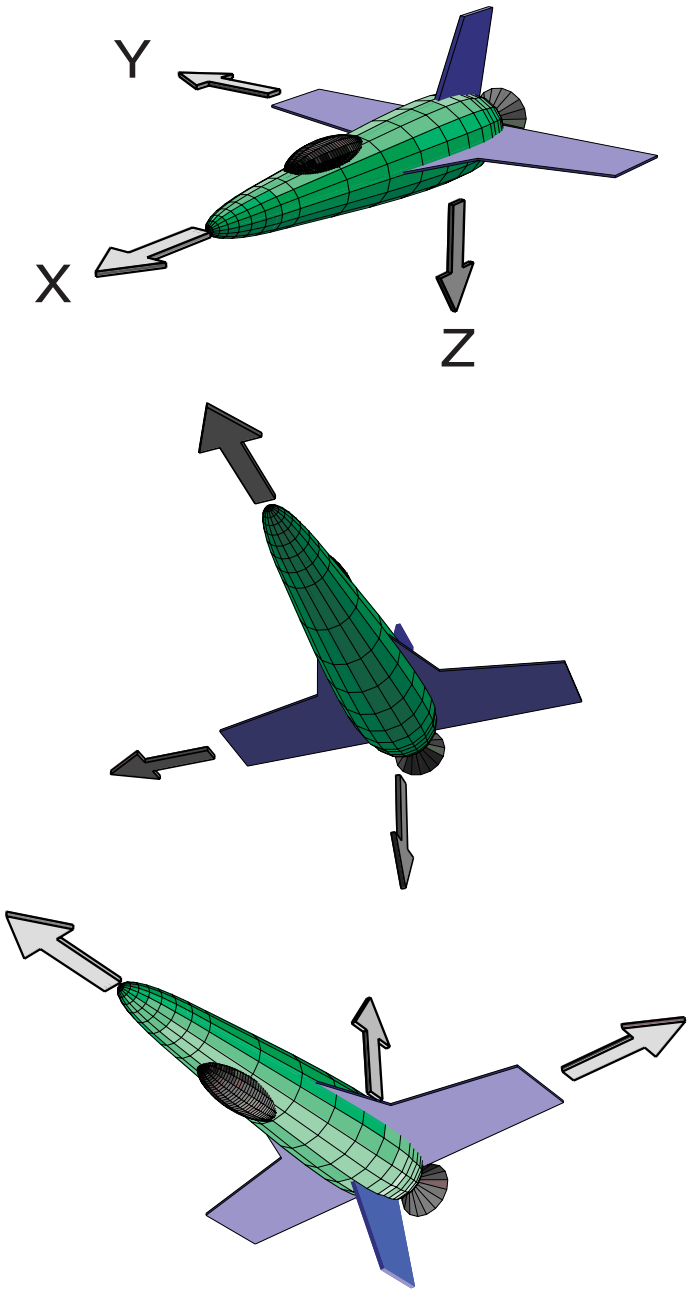

Figure 4-2 The Euler angle parameterization for attitude is usually defined as a [3-2-1] rotation. That is, first yaw (Ψ) about the navigation frame z-axis (top drawing), then pitch (Θ) about the intermediate frame y-axis, and finally roll (Φ) about the resulting body x-axis, to the final orientation.

where ω is the angular rate of the body, I is the moment of inertia matrix, and M is the applied external torque. The quantity $[\omega x]$ is the cross-product matrix, defined as

$$[\omega x] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \quad \text{(EQ 4.2)}$$

The transformation matrix from navigation to body coordinates, expressed as a function of Euler angles is:

$$\overset{n \to b}{C} = \begin{bmatrix} \cos\psi\cos\theta & \sin\psi\cos\theta & -\sin\theta \\ -\sin\psi\cos\theta + \cos\psi\sin\theta\sin\phi & \cos\psi\cos\phi + \sin\psi\sin\theta\sin\phi & \cos\theta\sin\phi \\ \sin\psi\sin\theta + \cos\psi\sin\theta\cos\phi & -\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi & \cos\theta\cos\phi \end{bmatrix} \quad \text{(EQ 4.3)}$$

where the angles are yaw (ψ), pitch (θ), and roll (ϕ). The propagation equation for the time rate of change of Euler angles with the body rates as inputs is:

$$\frac{d}{dt}\begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\phi & \sin\phi\cos\theta \\ 0 & -\sin\phi & \cos\phi\cos\theta \end{bmatrix}^{-1} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad \text{(EQ 4.4)}$$

where

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \quad \text{(EQ 4.5)}$$

and p, q, and r are the body fixed rotation rates as measured by on-board gyroscopes.

There are several advantages to Euler angles. They are immediately intuitive and an efficient method of parameterizing attitude. Furthermore, the angles are typically a required output of an attitude system. However, in order to transform a vector from one frame to another, a Direction Cosine Matrix (DCM) must be formed from the Euler angles and that DCM is full of transcendental functions as demonstrated in Equation 4.3. Also, when integrating the Euler angles to propagate attitude using a known or measured rota-

tion rate, the resulting equations are highly non-linear and again full of transcendental functions (Equation 4.4). Lastly, there is an ambiguity when the pitch angle is at either +/- 90 degrees. In a mathematical sense, there is a singularity, and yaw and roll cannot be defined, though the sum of yaw and roll and their difference remain well behaved. This is known as "Gimbal Lock" and in the days of mechanical gyroscopes was a concern for attitude system designers.

Another method for parameterizing attitude is through the DCM directly. The DCM is the transformation matrix between the navigation and body frames and thus is almost always in a form needed for further computation. The DCM is an orthonormal matrix, meaning that its transpose and its inverse are the same, and each of its columns is the unit normal basis vector of the navigation frame expressed in body coordinates.

The time propagation equation for the DCM can be written as:

$$\frac{d}{dt}\overset{n\to b}{C} = -[\omega\times]\overset{n\to b}{C} = \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix}\overset{n\to b}{C} \qquad \text{(EQ 4.6)}$$

where again,

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \qquad \text{(EQ 4.7)}$$

and *p, q*, and *r* are the body fixed rotation rates.

The clear advantage of the DCM is that the transformation matrix is already formed and the time propagation equation is a straightforward function with no transcendental functions or other obvious problems. The disadvantages are threefold. First, the DCM has no intuitive value whatsoever and discerning the rotation via the DCM is difficult at best. Second, the time propagation equation only contains three equations but the

DCM contains nine unique elements. The remaining six equations come from the constraint equations governing an orthonormal DCM. That is, the DCM must have its inverse equal to its transpose and each of its rows and columns must have unit length. These equations are non-linear and require much additional computation to resolve. In practice, this means that due to numerical instability during integration, the DCM must be constantly re-normalized and re-orthogonalized. This adds computational overhead, and often gives poor results if the integration is allowed without the re-normalization. Finally, there is a certain inelegance in carrying around nine parameters to describe only three angles.

The last parameterization discussed in this thesis is quaternions. Quaternions arise from the observation that any rotation between two coordinate frames can be recast as a single rotation about some unit vector. Thus quaternions represent the axis and angle of rotation between the navigation and body frames. The history and practical use of quaternions is discussed at length in Section 4.4 on page 102. The advantages of quaternions are efficiency (four parameters versus nine in the DCM), and the fact that both transformation and time propagation equations remain linear and computationally easy (requiring only simple addition and multiplication functions).

The transformation of a vector quantity from the navigation- to the body-frame can be accomplished with two quaternion multiplications:

$$r_n = q \otimes \begin{bmatrix} 0 \\ r_b \end{bmatrix} \otimes q^* \qquad \text{(EQ 4.8)}$$

where q is the base quaternion, and q^* is the quaternion's complement. The vector, r, is some arbitrary vector expressed in the navigation and body frames as denoted by the subscript n or b, respectively.

Likewise, the time propagation equations are equally simple:

$$\frac{d}{dt}q = -\frac{1}{2}\Omega q \qquad \text{(EQ 4.9)}$$

where Ω is defined as:

$$\Omega = \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\omega^T \\ \omega & [\omega\times] \end{bmatrix} \qquad \text{(EQ 4.10)}$$

where ω is again,

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \qquad \text{(EQ 4.11)}$$

and the rotation rates are body-fixed.

Thus the time propagation matrix, Ω, is simply the assembly of the ω, the cross-product matrix of ω, and the transpose of ω.

Again, the lack of intuition with quaternions is a disadvantage. In and of itself, an axis and angle representation is intuitive, but neither the axis nor the angle are immediately obvious from an initial inspection of the quaternion. Quaternions are, however, efficient, have no singularities, and are easily implemented on a digital computer without the need for transcendental functions. The attitude system assembled for the Atlantis was implemented using the quaternion attitude estimation algorithm developed in this chapter.

Essentially, the choice of parameterization is driven by the requirements of the system. If most of the attitude computation is integration or propagation, quaternions have an advantage. However, if the integration times are short, or the output is used by a human, then the quaternions would have to be converted to Euler angles to be useful, and thus

there is an advantage to remaining with the Euler angle parameterization in this case. Lastly, “gimbal lock” is not a problem with modern electronics, and can be dealt with by the use of some simple switching logic in the software.

SECTION 4.4 QUATERNION REVIEW

Quaternions are actually hyper-complex numbers of rank four and were invented by Sir William Rowan Hamilton in 1848. He was working on extending the simple complex number to have a vector imaginary part, as opposed to a mere scalar. The quaternion, along with the operations of quaternion addition and multiplication, forms a non-commutative division ring [86]. Though the application of this system of mathematics eluded his lifetime, the properties of the quaternion math are wonderfully adept at rotations, transformations, and keeping track of attitude. Specifically, quaternion math permeates the fields of satellite attitude control and computer graphics.

The quaternion is an “axis and angle” representation of the rotation between two coordinate frames. That is, the vector part of the quaternion can be seen to indicate an axis about which the object is rotated, and the scalar part the magnitude of that rotation. Mathematically, this can be seen from the properties of the DCM or transformation matrix:

$$tr\left[\overset{n\to b}{C}\right] = 1 + 2\cos\Phi \qquad \text{(EQ 4.12)}$$

where *tr* is the trace operator, and Φ is the total angle between the two frames. Also,

$$\frac{\overset{n\to b}{C} - \left(\overset{n\to b}{C}\right)^T}{2} = \sin\Phi \begin{bmatrix} 0 & a_z & -a_y \\ -a_z & 0 & a_x \\ a_y & -a_x & 0 \end{bmatrix} = -[a\times]\sin\Phi \qquad \text{(EQ 4.13)}$$

where a is the unit vector about which the frame rotates the total angle, Φ, to become the second frame. Intuitively, this can be visualized with the following thought experiment:

visualize the human right hand with the index finger pointing straight out, the thumb up, and the second finger bent at the top knuckle so as to point to the left. This defines the unit triad. Notice that almost any two orientations can be accomplished by bending the wrist and rotating the entire arm at the elbow. Realize that this is a single rotation about some strange axis. Thus, any rotation can be converted to a single rotation about some (arbitrary) axis. With this in mind, the definition of the quaternion as a representation or parameterization of attitude is shown:

$$q = \begin{bmatrix} \cos\frac{\Phi}{2} \\ a_x \sin\frac{\Phi}{2} \\ a_y \sin\frac{\Phi}{2} \\ a_z \sin\frac{\Phi}{2} \end{bmatrix} = \begin{bmatrix} q_0 \\ \vec{q} \end{bmatrix} \quad \text{(EQ 4.14)}$$

Because the rotation can be positive about the vector, a, or negative about the opposite vector, $-a$, the quaternion is unchanged by sign. Thus, in terms of attitude:

$$q = -q \quad \text{(EQ 4.15)}$$

Likewise, in order for the quaternion to represent attitude, it must be a unit length quaternion, or:

$$q^T q = q q^T = 1 \quad \text{(EQ 4.16)}$$

which can be verified by substituting the base formulation of the quaternion, Equation 4.14, into this equation. The quaternion complement, or inverse, represents a reverse rotation from the one specified. As can be seen from substituting $-\Phi$ for Φ in Equation 4.14,

$$q^{-1} = q^* = \begin{bmatrix} q_0 \\ -\vec{q} \end{bmatrix} = \begin{bmatrix} -q_0 \\ \vec{q} \end{bmatrix} \quad \text{(EQ 4.17)}$$

The last two basic quaternion math requirements are for multiplication and transformation. The quaternion multiplication represents successive rotation in *reverse* order of multiplication, thus $q \otimes s$ represents first a rotation through the quaternion, s, and then further rotation through the quaternion, q. The vector equation for quaternion multiplication is:

$$r = q \otimes s = \begin{bmatrix} q_0 s_0 - (\vec{q} \cdot \vec{s}) \\ q_0 \vec{s} + s_0 \vec{q} + \vec{q} \times \vec{s} \end{bmatrix} \quad \text{(EQ 4.18)}$$

which can also be written as:

$$q \otimes s = \left(q_0 I + \begin{bmatrix} 0 & -\vec{q}^T \\ \vec{q} & [\vec{q}\ \times] \end{bmatrix} \right) s \equiv [Q_+(q)]s \quad \text{(EQ 4.19)}$$

or alternately as:

$$q \otimes s = \left(s_0 I + \begin{bmatrix} 0 & -\vec{s}^T \\ \vec{s} & [-\vec{s}\ \times] \end{bmatrix} \right) s \equiv [Q_-(s)]q \quad \text{(EQ 4.20)}$$

where I is the [4 x 4] identity matrix. As previously stated, the quaternion multiplication and inverse properties can also be used to transform a vector quantity into a new coordinate frame:

$$r_n = q \otimes \begin{bmatrix} 0 \\ r_b \end{bmatrix} \otimes q^* \quad \text{(EQ 4.21)}$$

The transformation matrix can then be written as:

$$\overset{n \to b}{C} = \begin{bmatrix} 2q_0^2 + 2q_1^2 - 1 & 2q_1 q_2 - 2q_0 q_3 & 2q_1 q_3 + 2q_0 q_2 \\ 2q_1 q_2 + 2q_0 q_3 & 2q_0^2 + 2q_2^2 - 1 & 2q_0 q_1 - 2q_2 q_3 \\ 2q_1 q_3 - 2q_0 q_2 & 2q_0 q_1 + 2q_2 q_3 & 2q_0^2 + 2q_3^2 - 1 \end{bmatrix} \quad \text{(EQ 4.22)}$$

or, in matrix form:

$$\overset{n \to b}{C} = (2q_0^2 - 1)I - 2q_0[\vec{q}\ \times] + 2\vec{q}\vec{q}^T \quad \text{(EQ 4.23)}$$

These equations provide all of the necessary tools to derive the attitude algorithm, and will be used in later sections to demonstrate the quaternion estimation of Wahba's problem. For completeness, the transformation from quaternions to Euler angles is also included:

$$\Psi = \operatorname{atan}\frac{2q_1q_2 + 2q_0q_3}{2q_0^2 + 2q_1^2 - 1} \quad \text{(EQ 4.24)}$$

$$\Theta = \operatorname{asin}(2q_0q_2 - 2q_1q_3) \quad \text{(EQ 4.25)}$$

$$\Phi = \operatorname{atan}\frac{2q_2q_3 + 2q_0q_1}{2q_0^2 + 2q_3^2 - 1} \quad \text{(EQ 4.26)}$$

these are often required when the output of the attitude system is used directly by a human operator.

SECTION 4.5 VECTOR MATCHING

In 1965, Grace Wahba proposed that given a minimum of two vector measurements in the body frame, and a known value for the vectors in the reference frame, then the rotation between the body frame and the reference frame can be uniquely determined as long as the vectors are non-colinear and of non-zero length [144]. Another way of stating this is that a (non-orthogonal) triad can be formed by each of the two vectors and the cross-product between the two. With measurements of the components of the two vectors, the rotation necessary to force the vectors in line with this triad can be computed. The necessity of a non-zero value for the cross-product forces the two vectors to be non-colinear, and of non-zero length. If only one vector is present, the rotation can only be resolved to two of the necessary three angles.

Wahba's problem, as vector matching has become known, is used on many satellites for attitude information. The sensors are usually high quality star trackers that give

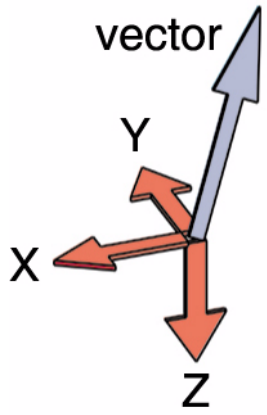

Figure 4-3 A vector is shown in the inertial coordinate frame. This will be the basis for an example of the mechanics of Wahba's problem and how it is solved. The blue vector can be any vector quantity that is known in the inertial frame, such as magnetic field or the acceleration of gravity.

measurements to the line-of-sight of known stars. Often, these are complemented with high-quality gyroscopes in order to smooth the attitude solution.

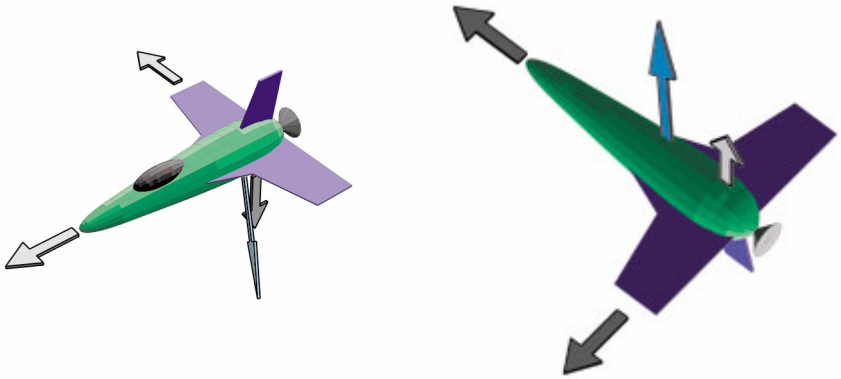

Figure 4-4 The same vector as previously defined in the inertial frame, as measured in the body coordinates (on the left side). In order to line up the vector to the inertial "known" state, the body must be rotated to line the vector up. This rotation is exactly that which transforms body quantities to inertial quantities. Of course, with only one vector, there remains an ambiguity of rotation about the vector itself. Thus, two or more non-colinear vectors are required.

Visually, the solution to Wahba's problem is simple. Examine the triad and vector pictured in Figure 4-3. The blue vector is some quantity that is known in the inertial coor-

dinate frame. Measuring the same components of that vector in the body frame, rotate the body until the vector lines up with what it is known to be in the inertial coordinates. At this point, the rotation is exactly that which transforms a quantity from the body to the inertial frame. This is pictured in Figure 4-4.

Several different algorithms have been used to solve Wahba's problem. Most notable are the papers by Bar-Itzhack [9], [10], [11], [12], and [13], Creamer [35], Lefferts [87], and Markley [95], [96]. These authors present algorithms that use singular value decompositions and other advanced matrix techniques to solve directly for the rotation matrix. The method used in this thesis is different from all of them and will be developed in detail in the following section.

SECTION 4.6 GYROSCOPE-FREE ATTITUDE ESTIMATION

In order to minimize the cost of the attitude system for the Atlantis project, it was decided not to include gyroscopes in the attitude solution. Two things make this reasonable. First, the bandwidth requirement for attitude was fairly low, at 10 Hz; and second, the rotation rates expected from the Atlantis during her trial runs were expected to be very small.

The vector matching algorithm, as previously explained, requires two non-colinear, non-zero length vectors in order to function. For convenience, the vectors used were Earth's magnetic field and total gravitational acceleration. Earth's magnetic field has been extensively mapped and the gravitational field can be assumed to point straight down.

In order to measure the body-fixed components of the magnetic field, a three-axis strap-down magnetometer was used. Body fixed acceleration was measured by a triad of micromachined accelerometers. The accelerometers, however, do not measure gravity

directly. Instead, they measure the specific force acting upon them. In this case, this means that they measure the vector difference of any acceleration in the inertial frame and gravity. Thus, the specific force in the navigation frame, i.e., the acceleration of the vehicle minus gravity must be known.

$$\vec{f} = \vec{a} - \vec{g} \qquad \text{(EQ 4.27)}$$

In an aircraft, as later examples will demonstrate, the inertial acceleration is estimated using the difference in subsequent GPS velocity measurements divided by the time elapsed between measurements (numerical differentiation of velocity to yield acceleration). This provides a good estimate of inertial acceleration, but the accuracy degrades as velocity decreases. This is due to the fact that the noise on the GPS velocity measurement is constant. As the velocity decreases, the portion of the measurement that is noise increases (decreasing signal to noise ratio). At very low velocities, it is unclear if the object is in motion or the output is noise from the GPS signal. This is further exacerbated by the differencing of subsequent measurements to generate a pseudo-derivative. As the noise is white, subtraction cannot be assumed to cancel the noise. In the specific case of the Atlantis, the accelerations were so small that the noise in the GPS velocity measurement was far larger than the signal being estimated. Thus, in the actual Atlantis hardware, the GPS acceleration was not used. The Atlantis was assumed to be accelerating at such a small fraction of one g that this omission caused only minor errors.

SECTION 4.6.1 MATHEMATICAL FORMULATION

The development of the novel attitude algorithm is straight forward and concludes with the attitude problem recast into a standard form to which most of the conventional filtering tools can be applied.

First, assume that the vector measurements are known in the navigation frame (denoted by the superscript *n*), and measured in the body frame (denoted by the superscript *b*). Thus, the magnetic field in the navigation frame is denoted as M^n and the magnetic field measured in the body frame is denoted as M^b.

The problem is one of attitude estimation in the quaternion domain. Thus there are three quaternions required: the true quaternion, q_{true}, the estimated attitude quaternion, $\hat{q}$, and the error quaternion relating the previous two, q_e. The relationship between the three is:

$$q_{true} = q_e \otimes \hat{q} \quad \text{(EQ 4.28)}$$

Thus the estimated attitude is again rotated slightly more by the error quaternion to become the true quaternion. In a conventional sense, q_e might be referred to as a correction quaternion rather than an error quaternion, but in order to remain consistent with prior research in [35], [43], and [126], it is defined here as the error quaternion. Note that because the error quaternion, q_e, is small, it can be approximated as:

$$q_e \cong \begin{bmatrix} 1 \\ \vec{q}_e \end{bmatrix} \quad \text{(EQ 4.29)}$$

where the scalar part of the quaternion is set to one. Note also that this is consistent with both the definition found in Equation 4.14 and the assumption that the angle of rotation, Φ, is small.

The first part of the mathematical formulation will focus on the magnetic field with the understanding that the same exact formulation applies to the total acceleration. The use of magnetic field is entirely arbitrary and was done because a convenient sensor was available in this problem. The formulation will remain correct for any vector quantity that can be known in the navigation frame and measured in the body frame.

The attitude formulation begins with a simple transformation of a vector from the body frame to the navigation frame:

$$\vec{M}^n = {}^{n\to b}C\,\vec{M}^b \qquad \text{(EQ 4.30)}$$

or, in the quaternion domain:

$$\begin{bmatrix} 0 \\ \vec{M}^n \end{bmatrix} = q_{true} \otimes \begin{bmatrix} 0 \\ \vec{M}^b \end{bmatrix} \otimes q^*_{true} \qquad \text{(EQ 4.31)}$$

substituting in the definition of q_{true} from Equation 4.28,

$$\begin{bmatrix} 0 \\ \vec{M}^n \end{bmatrix} = (q_e \otimes \hat{q}) \otimes \begin{bmatrix} 0 \\ \vec{M}^b \end{bmatrix} \otimes (q_e \otimes \hat{q})^* \qquad \text{(EQ 4.32)}$$

which can be expanded to:

$$\begin{bmatrix} 0 \\ \vec{M}^n \end{bmatrix} = q_e \otimes \left(\hat{q} \otimes \begin{bmatrix} 0 \\ \vec{M}^b \end{bmatrix} \otimes \hat{q}^* \right) \otimes q_e^* \qquad \text{(EQ 4.33)}$$

where the center part, the transformation of the body-fixed magnetic field into the navigation frame via the estimate of the true attitude, is referred to as $\hat{M}^n$, i.e.,

$$\begin{bmatrix} 0 \\ \hat{M}^n \end{bmatrix} \equiv \left(\hat{q} \otimes \begin{bmatrix} 0 \\ \vec{M}^b \end{bmatrix} \otimes \hat{q}^* \right) \qquad \text{(EQ 4.34)}$$

$\hat{M}^n$ is the estimate of the magnetic field expressed in the navigation frame. Making this substitution for $\hat{M}^n$:

$$\begin{bmatrix} 0 \\ \vec{M}^n \end{bmatrix} = q_e \otimes \begin{bmatrix} 0 \\ \hat{M}^n \end{bmatrix} \otimes q_e^* \qquad \text{(EQ 4.35)}$$

Note that using the transformation matrix or DCM notation, Equation 4.35 can be rewritten as:

$$\vec{M}^n = \overset{b \to n}{C}(q_e)\hat{M}^n \quad \text{(EQ 4.36)}$$

Substituting Equation 4.23 for the DCM, transposing the equation to change from body to navigation, using the definition of q_e provided in Equation 4.29, and discarding the higher order terms, consistent with the assumption of a small vector portion of the error quaternion yields,

$$\vec{M}^n = (I + 2[\vec{q}_e \times])\hat{M}^n \quad \text{(EQ 4.37)}$$

which can be rearranged as:

$$\underbrace{\vec{M}^n - \hat{M}^n}_{\delta\vec{M}^n} = 2[\vec{q}_e \times]\hat{M}^n \quad \text{(EQ 4.38)}$$

with the left hand side of Equation 4.38 denoted as $\delta\vec{M}^n$. Using the property of cross products that $A \times B = -B \times A$, and applying this to the right hand side of the equation:

$$\delta\vec{M}^n = -2[\hat{M}^n \times]\vec{q}_e \quad \text{(EQ 4.39)}$$

Note that Equation 4.39 has the form of the standard measurement equation of:

$$z = Hx \quad \text{(EQ 4.40)}$$

where, including the terms from the total acceleration:

$$z = \begin{bmatrix} \delta\vec{M}^n \\ \delta\vec{A}^n \end{bmatrix} \quad \text{(EQ 4.41)}$$

and

$$H = \begin{bmatrix} -2[\hat{M}^n \times] \\ -2[\hat{A}^n \times] \end{bmatrix} \quad \text{(EQ 4.42)}$$

and x is, of course, the vector portion of the error quaternion, $\vec{\hat{q}}_e$.

Thus, the very non-linear equation for transforming a measured quantity from the body frame to a known quantity in the navigation frame has been recast in a standard form. The measurement, z, and the observation matrix, H, both change with every time step. This does not present a problem as standard time-varying linear solution schemes are well known.

SECTION 4.6.2 ITERATED LEAST SQUARES

One method of estimating the attitude quaternion from the development above is to use an iterated least squares. In this methodology, as soon as a measurement is recorded, the solution to the measurement equation is recomputed and the solution is "walked" into convergence. The steps for the iterated least squares solution are as follows:

(1) Measure the body-fixed magnetic field and specific force: $\vec{M}^b$ and $\vec{A}^b$.

(2) Assume $\hat{q} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T$ and $q_e = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T$.

(3) Form the vector $\hat{M}^n$ by quaternion multiplication: $\hat{M}^n = \hat{q} \otimes \vec{M}^b \otimes \hat{q}^*$.

(4) Repeat Step (3) for $\hat{A}^n$.

(5) Form the vector $\delta\vec{M}^b$ where $\delta\vec{M}^b = \vec{M}^b - \hat{M}^n$.

(6) Repeat Step (5) for $\delta\vec{A}^b$.

(7) Form the measurement vector, z, by stacking the results of Steps (5) and (6).

(8) Form the observation matrix, H, as in Equation 4.42.

(9) Form the pseudo-inverse of H, $H^\dagger = [H^T H]^{-1} H^T$

(10) Compute the new vector part of the error quaternion, $\vec{\hat{q}}_e = \alpha[H^T H]^{-1} H^T z$

(11) Update the quaternion estimate: $\hat{q}(+) = q_e \otimes \hat{q}(-)$

(12) Return to Step (3) and repeat until converged.

The tuning parameter, α, in Step (10) is used to smooth the solution by using only part of the entire correction computed for the vector part of the error quaternion. The pseudo-inverse of H is invertible only when $\vec{M} \times \vec{A} \neq 0$, which mathematically restates the requirement that the two vectors be non-colinear and non-zero. To validate the algorithm, a Monte-Carlo simulation was performed where a random starting attitude was given to the algorithm. The body-fixed measurements were corrupted with appropriate levels of sensor noise, and the algorithm was allowed 100 iterations to converge with a tuning parameter, α, of 1/10. Figure 4-5 shows a time history for the attitude quaternion components during a single run in these series of simulations. As can be seen, the convergence to the correct attitude is rapid and smooth.

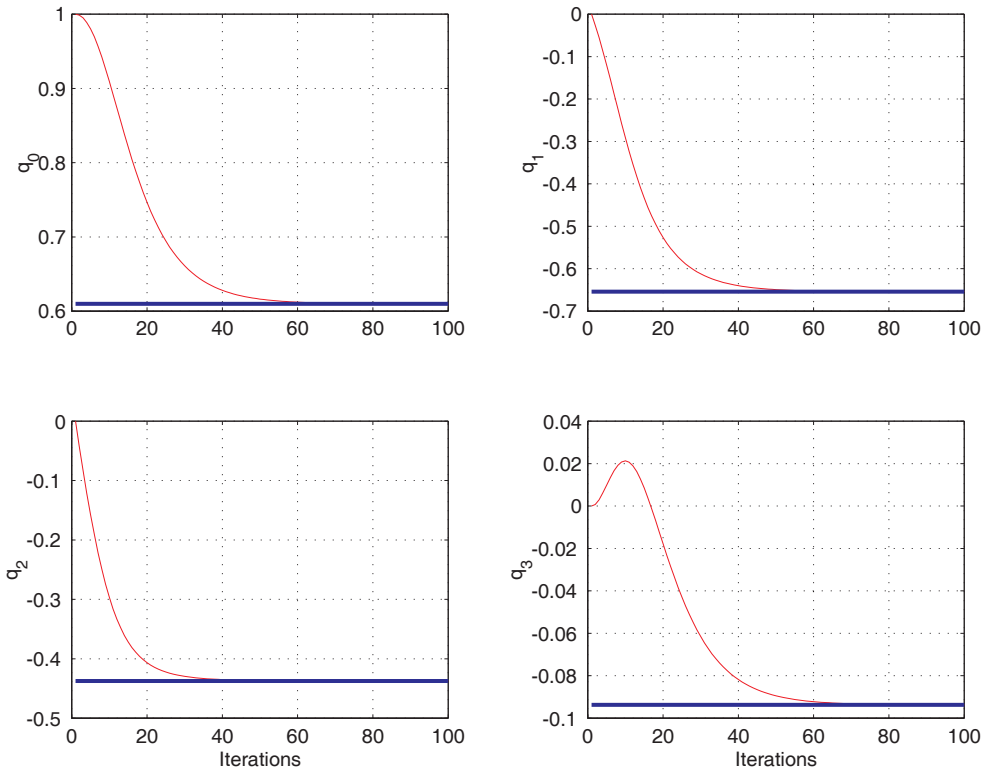

Figure 4-5 Quaternion convergence from Iterated Least Squares Monte-Carlo simulation. The body is set to a random orientation as indicated by the blue constant lines. The measured body fixed components of the magnetic field and gravity are calculated. The Iterated Least Squares algorithm is run for 100 times with a tuning parameter, α, of 1/10. Note that the solution is converged by 60 iterations.

In order to visualize the convergence of the algorithm, since the quaternions themselves do not lend an intuitive picture, a unit normal triad is attached to the simulated vehicle, and the trace of the tips of this triad plotted on the surface of a sphere. That is, the tip of a unit vector pointing out the nose of the aircraft is computed at every step in the iteration, and plotted on the surface of the sphere, along with the unit vector pointing out the right wing, and the one pointed down. That sphere, along with the traces, appears in Figure 4-6. The smoothness of the convergence is readily apparent in this view.

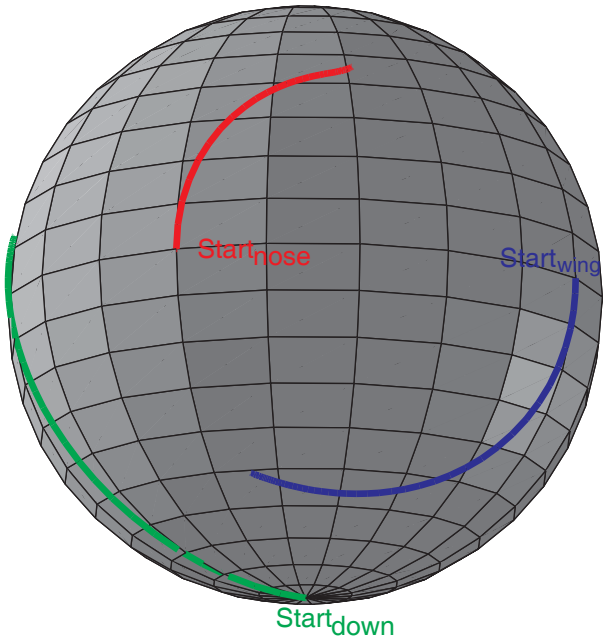

Figure 4-6 Iterated Least Squares solution convergence mapped onto a sphere. The orthogonal triad of unit vectors that point out the nose of the aircraft (in red), the right wing of the aircraft (in blue) and the bottom of the aircraft (in green) are rotated by the Iterated Least Squares algorithm at each step until convergence. Note that this algorithm is globally convergent.

The initial points, designated in red as "start nose," in blue as "start wing," and in green as "start down" correspond to the nose pointing north, and the aircraft level, as indicated by Step (2) of the iterated least squares algorithm, where $\hat{q} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T$. The final

location of the three unit vectors is the true attitude. In order to see this more clearly, the sphere portrayed in Figure 4-6 is unwrapped into a Mercator projection in Figure 4-7, below.

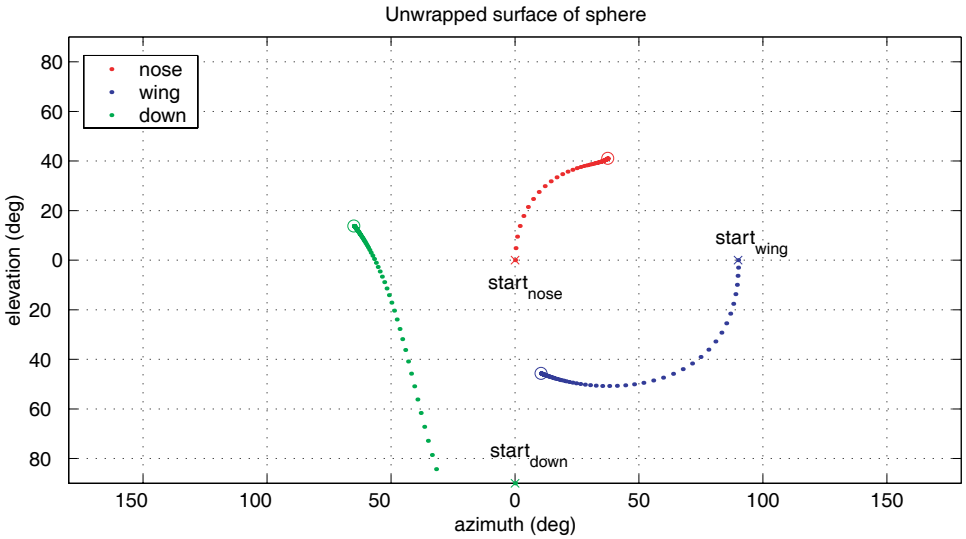

Figure 4-7 Iterated Least Squares solution sphere unwrapped via Mercator projection. Again, the unit normal triad of a unit vector out the nose of the aircraft (in red), out the right wing of the aircraft (in blue) and down through the bottom of the aircraft (in green) are plotted from the initial start of the Iterated Least Squares algorithm to convergence. Note that the initial large step in the down vector is an artifact of the unwrapping of the sphere into a rectangle.

The figures presented are of a typical run during the Monte-Carlo simulation. The iterated least squares quaternion estimation algorithm solution to Wahba's problem was numerically demonstrated to be globally convergent, in this static case, always converging to the correct attitude regardless of the initial guess on attitude. This was demonstrated by extensive simulation in which the true attitude was varied at 2.5 degree intervals through the entire space of possible attitudes. Furthermore, the tuning parameter is found to be quite robust, ensuring convergence to ε within n iterations as long as α is within the range of [1/n to 2].

SECTION 4.6.3 KALMAN FILTER

The iterated least squares solution works very well, but it is a snapshot solution based only on the current measurements of the vectors in the body frame. There is, however, more information available in the form of past attitude information. Typically, a Kalman filter is used to smoothly join the measurements with a dynamic model of the vehicle attitude. A Kalman filter is simply the linear-optimal estimator that minimizes the state error covariance for all time by properly blending prior information and current measurements. Certain assumptions about the broadband (white and Gaussian) nature of the model error (process noise) and sensor error (measurement noise) statistics are required for the optimality of the Kalman filter to hold.

In order to implement a Kalman filter, equations accounting for the dynamics must be included in the formulation. If angular rate measurements are available from gyroscopes, a dynamic model for the quaternion attitude estimate based on the kinematics of the attitude problem will be included [35]. If a dynamic model for the rate of change of the attitude quaternion is not included, then a lag will be introduced into the attitude solution. The lag, however, will be inconsequential if the dynamics are assumed to have a low frequency content. In this case, the quaternion errors may be modeled as an exponentially correlated or Gauss-Markov process. The dynamics of the state variable of interest, $\hat{q}_e$, can be written as:

$$\frac{d}{dt}\hat{q}_e = F\hat{q}_e + Gw \qquad \text{(EQ 4.43)}$$

where the state transition matrix, F, is:

$$F = -\frac{1}{\tau}I_{3\times 3} \qquad \text{(EQ 4.44)}$$

and the noise input matrix, G, is:

$$G = -\tau F = I_{3\times 3} \qquad \text{(EQ 4.45)}$$

and w is the white noise input and τ is the time constant of the exponential correlation.

If the problem is then formulated as a time-varying Kalman filter, using the state variable of interest and the measurements and measurement matrix as previously defined, the measurement update equation is stated as:

$$\hat{\vec{q}}_e^{(+)} = \hat{\vec{q}}_e^{(-)} + L(z - H\hat{\vec{q}}_e^{(-)}) \quad \text{(EQ 4.46)}$$

where the (-) and (+) superscripts denote before and after the measurement update, respectively. The measurement vector, z, is defined in Equation 4.41 and the measurement matrix is defined in Equation 4.42. The time-varying Kalman gain, L, is:

$$L = P^{(-)}H^T(H^TP^{(-)}H + R_v)^{-1} \quad \text{(EQ 4.47)}$$

where P is the state error covariance matrix, and R_v is the measurement noise matrix. The time covariance propagation is carried out via:

$$P^{(+)} = \Phi P^{(-)}\Phi^T + C_d \quad \text{(EQ 4.48)}$$

where Φ is the discrete equivalent of the state transition matrix, F, and C_d is the discrete process noise matrix. Finally, the measurement update of the covariance is carried out using the standard Kalman filter covariance update equation as found in [57]:

$$P^{(+)} = (I - LH)P^{(-)} \quad \text{(EQ 4.49)}$$

Note that if gyroscopes are present, then Equation 4.43 through Equation 4.46 change substantially. A very complete treatment of gyroscope integration within the framework of quaternion attitude determination can be found in [56].

SECTION 4.6.4 SIMULATION RESULTS

To confirm the performance of the time-varying Kalman filter formulation, a simulation was performed. This simulation includes both attitude and gross motion dynamics, and is such that each Euler angle and velocity can be ascribed a sinusoidal function with

individual means, amplitudes, and frequencies. This allows the motion differentiation to yield the exact acceleration, which can then be corrupted by noise to simulate actual sensors.

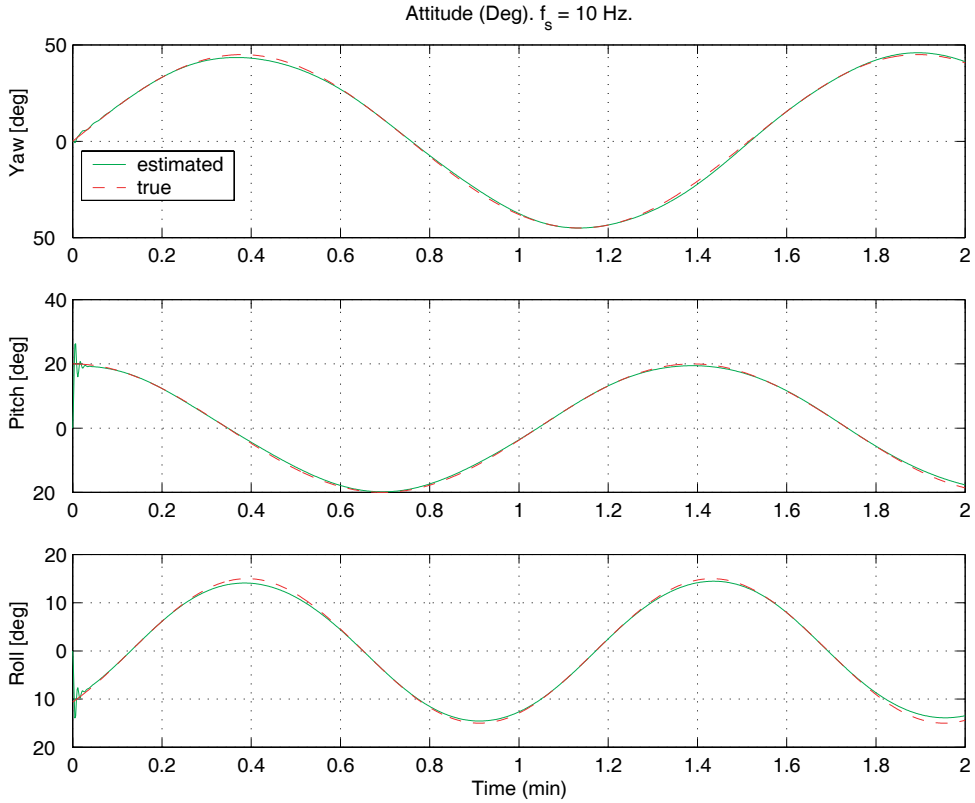

Figure 4-8 Simulation of the time-varying Kalman filter gyroscope-free attitude solution, in Euler angles. This is a model free Kalman filter solution that assumes the dynamics of the state transition matrix can be adequately modelled as a Gauss-Markov (exponentially correlated) process. Note the initial transients at the beginning as the filter converges on steady state gains.

Note that in this case, the accelerometers on board are not measuring gravity, but rather the vehicle specific force ($\hat{f} = \hat{a} - \hat{g}$). This is due to the fact that accelerations and gravitation cannot be distinguished as postulated by Einstein's equivalence principle. Thus, the local specific force is in fact, $\hat{a} - \hat{g}$, where both are expressed in the navigation frame. The local inertial acceleration, a, is obtained via differencing GPS velocity mea-

surements (in a numerical differentiation scheme). This has been included in the simulation.

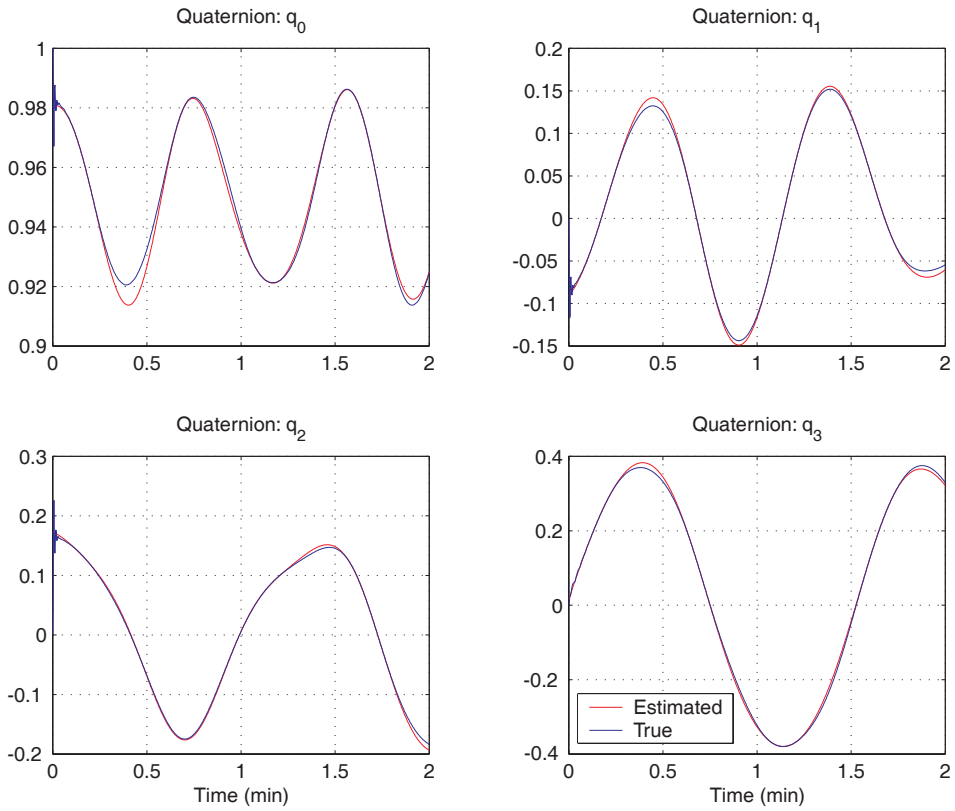

Figure 4-9 Simulation of the time-varying Kalman filter attitude solution, in quaternions. This is the same simulation as presented in the previous figure, but shown parameterized in quaternions. Again, note the initial transient response before the filter locks on to the steady state gains, and the small lag in the response at 0.4 minutes.

The simulation was run for approximately two minutes of simulated time, with the vehicle attitude oscillating back and forth. Figure 4-8 shows the three Euler angles, pitch, yaw, and roll, in degrees, with time along the x-axis. The initial transients show the filter locking on to the correct solution, and close inspection of the time around 0.4 minutes shows that the solution lags the true solution slightly at the inflection point of the roll and pitch motion.

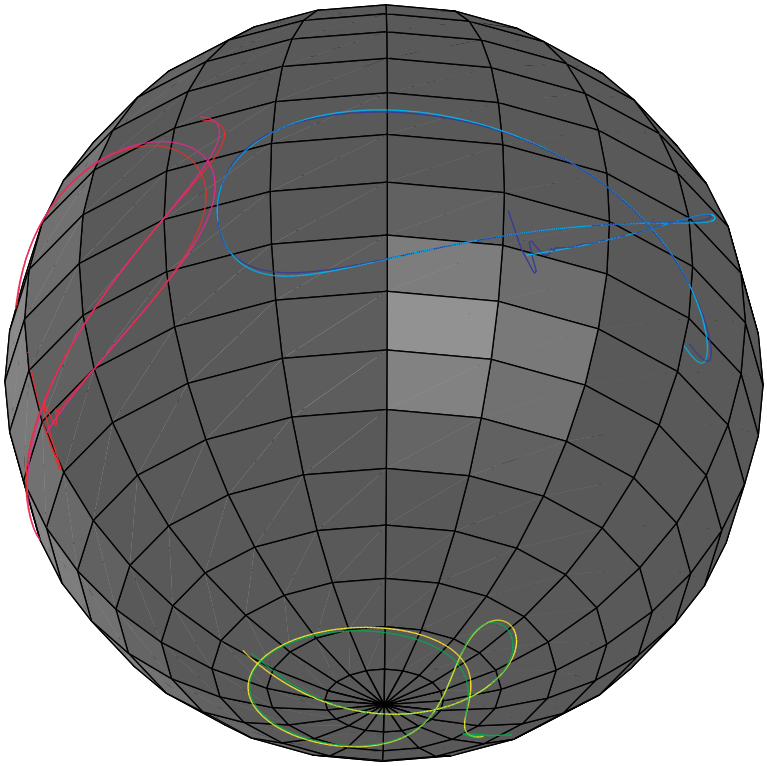

Figure 4-10 Simulation of the time-varying Kalman filter attitude solution, unit normal triad on sphere. In this figure, dynamic tracking (as opposed to a static solution) is shown. The true out the nose unit normal is shown in magenta, and the filtered version shown in red. The true out the right wing vector is shown in cyan, with the filtered version in blue, and the corresponding down vectors are in yellow and green.

Figure 4-9 shows the same simulation but in this case displaying the solution in the quaternion domain. The scalar part of the attitude quaternion estimate, $\hat{q}$, is placed in the upper left corner, with the three components of the vector portion shown. Since the quaternion has a scalar portion very close to one, the remaining vector components can be interpreted as the sine of roll, pitch, and yaw. Again, note the initial transient, and the lag at 0.4 minutes.

Figure 4-10 shows the trace of the estimated unit normal triad of the body-fixed axes on the surface of a sphere, along with the true attitude unit normal trace, for the same simulation. Here the transients of the initial convergence can be seen very clearly. This is

evidence that the estimator poles are fast, but lightly damped. Figure 4-11 shows the same information as Figure 4-10, but has unwrapped the surface of the sphere via a mercator projection. This projection distorts the areas near the top and bottom of the sphere, such that the information on the down axis is relatively poor, but the transients and lags show up very clearly on the nose and wing traces.

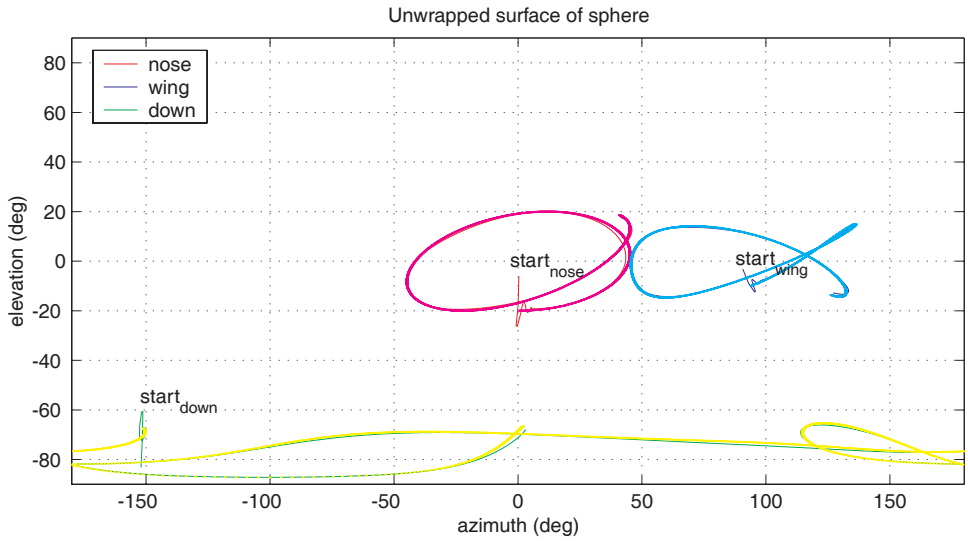

Figure 4-11 Simulation of the time-varying Kalman filter attitude solution, unit normal triad on sphere unwrapped using a mercator projection. This is the same data as the previous figure unwrapped. The large distortion on the "down" trace is due to the projection used. The initial transients and lag in the estimated attitude show up very well in the "nose" and "wing" traces.

The time-varying Kalman filter performs very well, even in the presence of gross sensor noise, and high frequency dynamics. The solution locks on to the true attitude quickly and demonstrates very little lag, in spite of the fact that the dynamic model is essentially a low-pass filter.

There is a trade-off in the tuning of the Gauss-Markov parameter, τ. A larger τ corresponds to longer correlation times and slows down the tracking response. This, however, also smooths the solution and effectively averages out noise. Decreasing τ improves tracking performance at the expense of making the solution more sensitive to noise. The initial

transients in the simulated response could be reduced greatly by initially using the static iterated least squares solution to begin the filter from a good initial guess of attitude.

SECTION 4.7 MAGNETOMETER CALIBRATION

In this application, one of the two known vectors for the vector matching algorithm is Earth's magnetic field. Driving this selection are the recent developments in materials technology that make low-cost three-axis strap-down magnetometers a reality. These sensors are, however, noisy and prone to both scale-factor and bias errors. Furthermore, the local magnetic field is corrupted by nearby objects. In this light, an introduction to the errors common to magnetometers and traditional methods for calibrating them will be presented, along with a novel two-step calibration algorithm.

SECTION 4.7.1 INTRODUCTION

Magnetometers are instruments used for measuring the strength and direction of magnetic fields. They are used extensively in aircraft navigation, marine navigation, and oilfield borehole applications. The magnetometers are used to determine the vehicle heading, where heading is defined as the angle formed between the longitudinal axis of the vehicle projected onto the horizontal plane and magnetic north. In most applications, normally a pair of magnetometers are mounted perpendicular to each other or a triad of magnetometers is mounted orthogonally. In this configuration, the magnetometers are used to measure the strength of Earth's magnetic field vector in body coordinates from which heading is computed. When the vehicle is straight and level (pitch and roll angles equal zero), magnetic heading is determined using:

$$\psi = \text{atan}\left(\frac{B_x^b}{B_y^b}\right) \qquad \text{(EQ 4.50)}$$

where B_x^b and B_y^b represent measurements of Earth's magnetic field vector resolved in a coordinate system attached to the vehicle's body. The superscript "b" denotes a vector measurement made in the body coordinate frame. Errors in making the measurements B_x^b and B_y^b directly corrupt the accuracy of the heading obtained from Equation 4.50. The process of calibrating magnetometers involves identifying and removing the errors in these measurements. This section explores a novel method to calibrate these magnetometers.

SECTION 4.7.2 MAGNETOMETER MEASUREMENT ERROR MODELS

Magnetometers come in many different forms and operate on different principles. A new generation of low-cost magnetometers used in this research is Anisotropic Magneto-resistive (AMR) sensors. These units have a sensing element that is made from a nickel-iron alloy (or Permalloy) whose electrical resistance changes in the presence of magnetic fields. The Permalloy material is normally deposited on thin silicon wafers which can be bulk manufactured in a form suitable for commercial integrated circuit packages. The sensing element of the magnetometer used in this research has dimensions on the order of 10 mm per side [32]. The output of these devices is an analog voltage proportional to the strength of the magnetic field.

The mathematical model for the output of such a strap-down magnetometer triad is:

$$\hat{B}^b = C_m C_{sf} C_{si} (B^b + \delta B^b) \qquad \text{(EQ 4.51)}$$

where B^b represents the *actual* or true magnetic field vector while $\hat{B}^b$ represents the *measured* magnetic field vector. The term δB^b represents the hard iron biases. The three leading matrices of the right-hand side of Equation 4.51 account for misalignment, scale factor errors, and soft iron, respectively.

The magnetic field that is used in heading determination is Earth's magnetic field. In most practical applications there will be other unwanted magnetic fields corrupting the measurements of the magnetometer triad. These unwanted fields are normally generated by ferromagnetic materials with permanent magnetic fields (or "hard irons") that are part of the aircraft structure or equipment installed near the magnetometer. These unwanted magnetic fields are superimposed on the output of the magnetometers' measurement of Earth's magnetic field. The effect of this superposition is to bias the magnetometer output.

If the unwanted magnetic fields are time invariant, they are referred to as "hard iron" errors and can be represented by a vector quantity, δB^b. If the strength and direction of these unwanted magnetic fields is known, their effect can be removed to un-bias the magnetometer readings. It should be noted that the unwanted magnetic fields can also be caused by items external to the vehicle, but since the vehicle is usually moving, the effect of such fields will be temporary. Furthermore, external fields are unlikely to be present in the air or on the water when using this application for marine or flight vehicles. Thus, errors due to external fields can usually be safely neglected (Note that some areas, such as iron ore deposits can cause local deviations in the magnetic field of up to 10 degrees). Items inside the vehicle can generate unwanted magnetic fields that are time varying. For example, such an item would be a current carrying wire. If the current through the wire is time varying, the resulting magnetometer bias will also be time varying and difficult to calibrate. Fortunately, such errors can be managed by taking care during installation of the magnetometers.

There are materials that generate magnetic fields in response to externally applied fields. The field generated by these materials, called "soft irons," can vary over a wide range depending on both the magnitude and direction of the applied external magnetic field. If such materials are present, they will generate a magnetic field that will be super-imposed on the magnetometer output as they generate their own magnetic field in

response to Earth's magnetic field. Since the orientation of Earth's magnetic field vector relative to the soft iron materials fixed inside the vehicle changes with vehicle attitude, this gives rise to a varying bias on the magnetometer output.

In a simple one-dimensional case, the magnitude of the soft iron response is proportional to the external magnetic field. The constant of proportionality is a property of the soft iron material and it is referred to as the material's magnetic susceptibility. In this work it will be assumed that this simple linear relationship is sufficient. However, in some cases there can be appreciable hysteresis. In most soft iron materials the hysteresis is small enough that the linear model is sufficient. In a hysteresis-free, three-dimensional case, a 3×3 matrix is required as opposed to a constant of proportionality. This is the C_{si} matrix in Equation 4.51. In most vehicle installations, magnetometers are installed in parts of the vehicle where errors due to soft iron will be small, and thus C_{si} is assumed to be the identity matrix.

Ideally, the three magnetometers that make up the triad are identical. In reality, however, this may not be the case; each magnetometer will have different sensitivities. That is, when all three magnetometers are subjected to an identical magnetic field, the observed output from each will not be the same due to scale factor errors. Calibrating the scale factor error involves determining the multiplicative factor that must be applied to each magnetometer such that the outputs will be the same when subjected to identical magnetic fields. In the three-dimensional case, the multiplicative constant is the 3×3 diagonal matrix, C_{sf}.

The last remaining error source in the Equation 4.51 is the misalignment error matrix, C_m. In an ideal installation, the magnetometer triad will be mounted in perfect alignment with the body axis of the vehicle. If reasonable alignment is not achieved, however, errors in the magnetometers' outputs will be present due to cross coupling of the

magnetic field axes. If care is taken during the installation process, the misalignment between the triad and the vehicle body axes can be minimized (Note that in a real world case, it would be far easier to estimate the misalignment error rather than try to achieve perfect mechanical mounting). Rotating the body full circle while on a level platform can extract the pitch and roll misalignments, but not the yaw. For the analysis that follows it is assumed that the magnetometer misalignment has been removed and hence the term C_m is the identity matrix. In most vehicle applications, the two largest error sources addressed in magnetometer calibration algorithms are hard iron biases and scale factor errors.

SECTION 4.7.3 MAGNETOMETER SWINGING

Two-magnetometer heading determination systems have been used extensively in navigation applications. Traditionally the sensors used in these systems were flux-gate or flux-valve magnetometers. The method of calibration in the heading domain has been known for over a century [27] and has been used to calibrate modern solid-state magnetometers. This method of calibration is based on perturbation of the basic heading equation (Equation 4.50) and a substitution of the error equations (Equation 4.51). This results in the following heading error equation:

$$\delta\psi = A + B\sin\psi + C\cos\psi + D\sin 2\psi + E\cos 2\psi \qquad \text{(EQ 4.52)}$$

This is, in effect, a truncated Fourier series where the Fourier coefficients are functions of the hard and soft iron errors. Estimation of the Fourier coefficients is accomplished by a procedure called “swinging.” The procedure involves leveling and rotating the vehicle containing the magnetometer through a series of *N known* headings as shown schematically in Figure 4-12. At each known k^{th} heading, the heading error, $\delta\psi_k$, is computed. These values will be used to form the system of equations (Equation 4.53). The

Moore-Penrose pseudo-inverse, or batch least squares solution of Equation 4.53 yields estimates for the coefficients *A* through *E*.

$$\begin{bmatrix} \delta\psi_1 \\ \delta\psi_2 \\ \dots \\ \delta\psi_N \end{bmatrix} = \begin{bmatrix} 1 & \sin\psi_1 & \cos\psi_1 & \sin 2\psi_1 & \cos 2\psi_1 \\ 1 & \sin\psi_2 & \cos\psi_2 & \sin 2\psi_2 & \cos 2\psi_2 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \sin\psi_N & \cos\psi_N & \sin 2\psi_N & \cos 2\psi_N \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \\ E \end{bmatrix} \quad \text{(EQ 4.53)}$$

There are three major drawbacks with this calibration method which necessitated the development of the new calibration algorithm. A shortcoming of the method which becomes apparent when examining Equation 4.50 is that heading is a required input. Since heading errors due to hard and soft iron errors are not constant but heading dependent, the heading input into the algorithm will be corrupted by a non-constant bias. Thus, the first drawback is that another independent measurement of heading is required when calibrating magnetometers using this method.

When performing this calibration for a magnetometer triad installed in an aircraft, the standard practice is to use a compass rose painted on the tarmac as shown in Figure 4-12, as the secondary independent heading measurement. The second drawback is that some of the coefficients, *A* through *E,* are functions of the Earth's local magnetic field strength. That is, if the vehicle with the magnetometers is expected to travel over a large geographic area, there will be large variations in the Earth's local magnetic field vector and multiple calibrations must be performed (in practice the errors are simply tolerated). Each calibration will yield a location dependent coefficient set *A* through *E* that can be scheduled as needed. Lastly, this method can only calibrate heading in the two dimensional case. If the magnetometers are not going to be used in a heading determination system or if the system consists of three orthogonal magnetometers, then this calibration algorithm is not applicable.

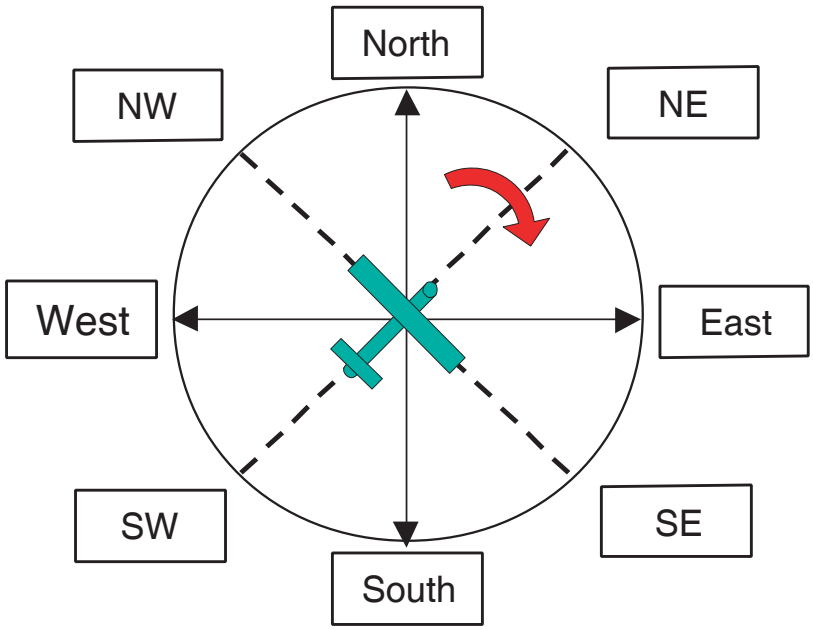

Figure 4-12 Graphical depiction of magnetometer calibration via "swinging." The aircraft is set upon a known heading reference (usually a painted compass rose at an airport) and "swung" through different known headings. Based on the readings of the magnetometers at each heading, a set of Fourier coefficients can be computed to solve for the biases in heading.

SECTION 4.7.4 ALGORITHM DEVELOPMENT

Unlike the swinging algorithm, a new calibration method has been developed that requires no external reference, nor is it location dependent. The algorithm will be detailed first in the two dimensional case, and then extended to three dimensions. The calibration method is based on the observation that the locus of error-free measurements from two perpendicularly mounted magnetometers is a circle. This can be shown as follows:

$$(B_x^b)^2 + (B_y^b)^2 = B_H^2\cos^2\psi + B_H^2\sin^2\psi = B_H^2 \qquad \text{(EQ 4.54)}$$

where B_H is the magnitude of the horizontal component of Earth's magnetic field vector. Equation 4.54 is the equation of a circle with its origin at the center and radius equal to the magnitude of the horizontal component of Earth's magnetic field. The magnitude of the radius varies with latitude, longitude and altitude because Earth's magnetic field vector

varies with geographic location. This variation of Earth's magnetic field vector is well known and well modeled. The magnitude of the Earth's local magnetic field vector is determined using the 1999 International Geomagnetic Reference Field model [17].

The effect of the various magnetometer errors described in Equation 4.51 is to alter the shape of the circle described in Equation 4.54. Hard iron biases, for example, shift the origin of the circle. This can be shown mathematically if we consider the case where the x and y components of Earth's magnetic field vector are offset by hard iron biases, δB_{x_0} and δB_{y_0}, respectively. In this case, the measured field strengths, $\hat{B}_x^b$ and $\hat{B}_y^b$, are given by:

$$\begin{aligned} \hat{B}_x^b &= B_x^b + \delta B_{x_0} \\ \hat{B}_y^b &= B_y^b + \delta B_{y_0} \end{aligned} \qquad \text{(EQ 4.55)}$$

By substituting the terms of Equation 4.55 into Equation 4.54, it follows that:

$$(\hat{B}_x^b - \delta B_{x_0})^2 + (\hat{B}_y^b - \delta B_{y_0})^2 = B_H^2 \qquad \text{(EQ 4.56)}$$

which is still the equation for a circle of radius B_H, but is now centered on the point $(\delta B_{x_0}, \delta B_{y_0})$. Scale factor errors cause the body x- and y-magnetometer measurements to be different when both are subjected to a magnetic field of the same strength. This can be expressed mathematically as follows:

$$\begin{aligned} \hat{B}_x^b &= (1 + sf_x) B_H \cos\psi \\ \hat{B}_y^b &= -(1 + sf_y) B_H \sin\psi \end{aligned} \qquad \text{(EQ 4.57)}$$

which can be rearranged as:

$$\left(\frac{\hat{B}_x^b}{1 + sf_x}\right)^2 + \left(\frac{\hat{B}_y^b}{1 + sf_y}\right)^2 = B_H^2 \qquad \text{(EQ 4.58)}$$

which is simply the equation of an ellipse, centered on the origin, with semi-major and semi-minor axes determined by the parameters sf_x and sf_y. When the output of the magne-

tometer is corrupted both by scale factor errors and by hard iron biases, the resulting locus is an ellipse with its center at the point $(\delta B_{x_0}, \delta B_{y_0})$, and the full equation is:

$$\left(\frac{\hat{B}_x^b - \delta B_{x_0}}{1 + sf_x}\right)^2 + \left(\frac{\hat{B}_y^b - \delta B_{y_0}}{1 + sf_y}\right)^2 = B_H^2 \qquad \text{(EQ 4.59)}$$

In this analysis, the misalignment error is considered to be the identity matrix. However, if misalignment were present, it would have the effect of rotating the semi-major and semi-minor axes with respect to the cartesian x- and y-coordinate system.

The calibration algorithm developed is a parameter estimation problem. When an aircraft with a pair of magnetometers is rotated through 360 degrees about its yaw axis, the locus of body fixed measurements will be an ellipse. The algorithm attempts to fit the best ellipse (in the least squares sense) to the measured data. If misalignment and soft iron errors are minimized during installation, then the parameter estimation unknowns are the hard iron errors and scale factor errors. In terms of the mathematics of the estimation process, the hard iron errors correspond to the center of the ellipse and the scale factor errors correspond to the size of the major and minor axes of the ellipse. Furthermore, extension of this algorithm to the case of three magnetometers is straight forward: instead of the parameters of an ellipse, the parameters of an ellipsoid are estimated.

There are two ways to carry out the parameter estimation. The first method is to use a classical approach to estimating parameters in a non-linear system of equations such as a non-linear least squares solution (such as the Extended Kalman Filter (EKF) described in [57] and [133]). Specifically, this estimator is based on linearizing the locus described in Equation 4.59. The governing non-linear equations are linearized such that the perturbations of the parameter of interest are estimated. The estimated perturbations are added to non-perturbed variables and the estimation is repeated until convergence is

achieved. It is an iterative procedure and requires a good initial guess of the scale factors and hard iron biases to begin the algorithm.

The EKF approach was simulated extensively in [44] and found to be sensitive to primarily three effects: a good initial guess of scale factor and biases was required for convergence, large sampling noise caused the EKF to diverge, and a large portion of the ellipsoid surface was required for convergence.

The second method for parameter estimation uses a non-linear, two-step estimator. The non-linear two-step estimator, which solves the parameter identification problem by breaking the estimation process into two steps, is an adaptation of the estimator introduced in [62]. In the first step, a set of variables (called first step states) are defined as a combination of the various parameters to be estimated. The estimation problem is linear in these first step state variables and therefore retains the desirable properties of linear systems. Following estimation of the first step states, the second step states, which are the scale factors and hard iron biases, are extracted from the first state steps through algebraic manipulation. The art in this process consists of properly identifying the first step states.

Derivation of the equations for the non-linear, two-step estimator begins by expanding the locus equation Equation 4.59. By multiplying up from the denominator and rearranging, the equation can be rewritten in matrix form as:

$$\left[(\hat{B}_x^b)^2\right] = \left[2\hat{B}_x^b \; -(\hat{B}_y^b)^2 \; 2\hat{B}_y^b \; 1\right] \begin{bmatrix} \delta B_{x_0} \\ k_2 \\ k_2 \delta B_{y_0} \\ k_1 - (\delta B_{x_0})^2 - k_2(\delta B_{y_0})^2 \end{bmatrix} \quad \text{(EQ 4.60)}$$

where:

$$k_1 = (1 + sf_x)^2 B_H^2 \quad \text{(EQ 4.61)}$$

$$k_2 = \frac{(1 + sf_x)^2}{(1 + sf_y)^2} \quad \text{(EQ 4.62)}$$

Equation 4.60 represents a single point. If multiple measurements are taken and the left hand side of the equation is stacked with all of the measurements, as is the first matrix on the right hand side of the equation, then the parameters of the vector can be obtained in a least squares sense using the pseudo-inverse. Once the parameter vector is known, then algebraic manipulation will extract the relevant scale factors and biases using Equation 4.61 and Equation 4.62. In order to gauge the effectiveness of this algorithm, a simulation study was carried out.

The simulation is shown in Figure 4-13. The original centered circle is shown in red. Scale factors then distort the circle into an ellipse (blue), and biases finally move the center of the ellipse (green). The green dots represent noisy sampled data of the final ellipse. Using this sampled data, the algorithm is run to determine the scale factors and biases. The only additional information that is available to the algorithm (other than the green points themselves) is the known radius of the red circle. Using the estimated scale factors and biases, the green points are transformed to remove the bias and scale factor effects. The result of that transformation is the red dots scattered along the original circle. Note that the entire ellipse was not required (the amount of the circle required depends on the level of noise, more noise required more information i.e. more of the circle). Again, the only additional information required was the radius of the original circle.

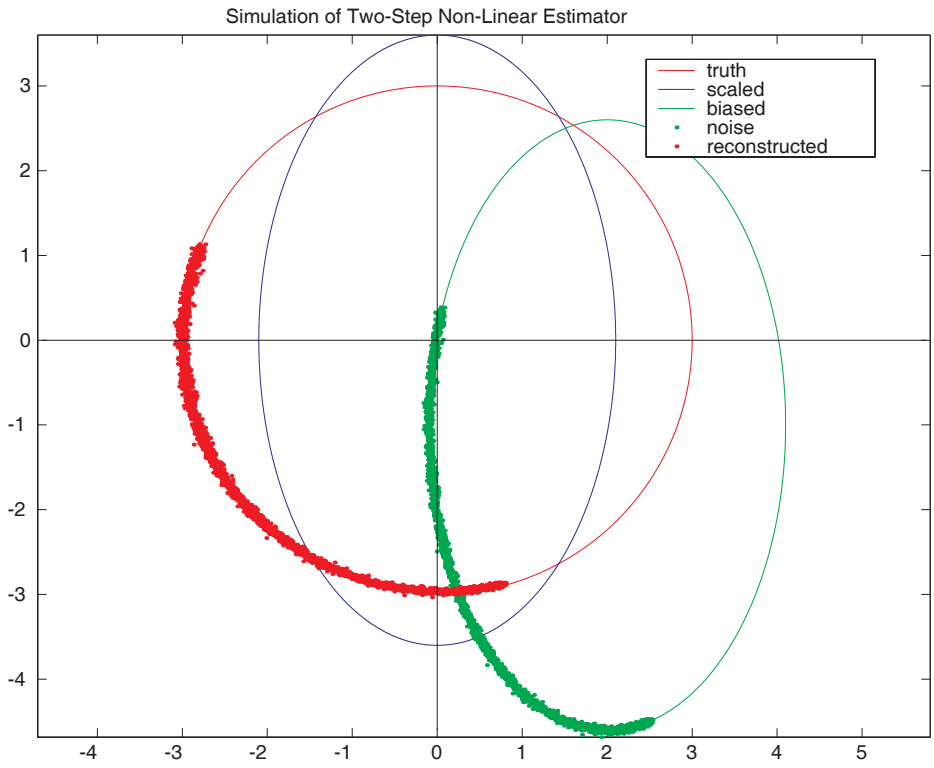

Figure 4-13 Simulation of the two-step calibration algorithm for a 2-D case. The red solid line represents the true response of the magnetometer. It is then distorted by scale factors, stretching the red circle into the blue ellipse. Then it is corrupted by biases, shifting the blue ellipse to the off centered green ellipse. Noisy data from that ellipse (green dots) are run through the calibration algorithm, resulting in a calibration that remaps those points into the red dots, on the original response circle.

The three dimensional case is a simple extension of the two dimensional case. The governing ellipse equation becomes an equation for an ellipsoid:

$$\left(\frac{\hat{B}_x^b - \delta B_{x_0}}{1 + sf_x}\right)^2 + \left(\frac{\hat{B}_y^b - \delta B_{y_0}}{1 + sf_y}\right)^2 + \left(\frac{\hat{B}_z^b - \delta B_{z_0}}{1 + sf_z}\right)^2 = \|B\|^2 = R^2 \qquad \text{(EQ 4.63)}$$

and this can be rearranged into the following matrix form:

$$\left[-(\hat{B}_x^b)^2\right] = \left[-2\hat{B}_x^b\ (\hat{B}_y^b)^2\ -2\hat{B}_y^b\ (\hat{B}_z^b)^2\ -2\hat{B}_z^b\ 1\right]\begin{bmatrix}\delta B_{x_0}\\ k_2\\ k_2\delta B_{y_0}\\ k_3\\ k_3\delta B_{z_0}\\ k_4\end{bmatrix} \quad \text{(EQ 4.64)}$$

where:

$$k_1 = (1 + sf_x)^2 R^2 \quad \text{(EQ 4.65)}$$

$$k_2 = \frac{(1 + sf_x)^2}{(1 + sf_y)^2} \quad \text{(EQ 4.66)}$$

$$k_3 = \frac{(1 + sf_x)^2}{(1 + sf_z)^2} \quad \text{(EQ 4.67)}$$

$$k_4 = -k_1 + (\delta B_{x_0})^2 + k_2(\delta B_{y_0})^2 + k_3(\delta B_{z_0})^2 \quad \text{(EQ 4.68)}$$

In this case, samples of the points along the surface of the ellipsoid are stacked, and the pseudo-inverse is used to extract the right most column of Equation 4.64. Then algebraic manipulations using Equation 4.65, Equation 4.66, Equation 4.67, and Equation 4.68 are performed to yield the biases and scale factors which most perfectly transform the sampled data into a sphere of radius, *R*, centered on the origin.

SECTION 4.7.5 EXPERIMENTAL SETUP

Figure 4-14 shows the experimental setup used to validate the calibration algorithm on the three-axis strap-down magnetometers. The yellow box on the end of the wooden boom contains a Honeywell HMR-2003 three-axis strap-down magnetometer hybrid. It is placed on the wooden boom to remove it from the magnetic fields generated by the power supply for the inertial navigation system (INS) that is located behind the

Figure 4-14 Experimental setup to validate magnetometer calibration algorithm. The yellow box at the end of the wooden boom contains the Honeywell HMR-2003 three-axis magnetometers and a microcontroller. It was placed out at the end of the wooden boom to minimize the effect of the magnetic fields of the INS power supply located beneath the orange box. Data was taken by pitching, rolling, and yawing the entire setup repeatedly.

screen, below the orange case. The INS below the orange box is a Honeywell YG-1851 navigation grade INS with rate gyroscopes whose drift rates are less than 0.01 degree/ hour. The INS provides a truth measurement. Because the drift rates on the INS gyroscopes are so low, it is able to detect Earth's rotation rate, and lock onto its own attitude. Once this process is complete, the INS will continue to track attitude with a very high degree of accuracy (Note that this kind of accuracy is available only on very, very expensive sensors).

In order to generate points of the surface of the ellipsoid, the entire setup pictured above was rolled, pitched, and yawed in order to measure the various components of Earth's magnetic field. Note that the easiest motion to achieve was yaw and thus most of the point traces are loops around the surface of the ellipsoid. This is due to the fact that Earth's magnetic field has a dip of approximately 70 degrees in the San Francisco area (i.e., the magnetic field vector points into the ground at approximately 70 degrees). Thus a

simple yaw motion with the magnetometers perfectly level will trace out a circle on the top of the ellipsoid. As the magnetometer is pitched up and then yawed again, a new circle will be traced which will translate down the surface of the ellipsoid toward the positive y-axis. Rolls will move the circle toward the negative x-axis.

SECTION 4.7.6 EXPERIMENTAL RESULTS

The setup pictured in Figure 4-14 was tilted up and down, side to side, and yawed

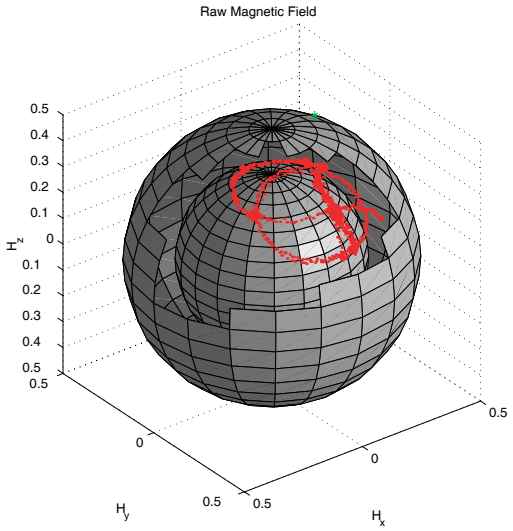

Figure 4-15 Plot of the measured components of Earth's magnetic field before calibration, along with a sphere of radius equal to the magnitude of Earth's magnetic field. This is the raw data recorded from the magnetometers on the wooden boom in the previous figure. In order to see the actual trace, portions of the sphere had to be removed, as the scale factors are less than one. In addition, close inspection will reveal that the trace lies to the right and behind the smaller sphere, indicating the presence of bias errors as well.

about several times in each direction over the course of an hour. The body-fixed magnetic field was recorded along with true attitude from the INS. The three components of the uncorrected magnetic field are plotted and a reference sphere of radius $\|\vec{B}\|$ is plotted in Figure 4-15. Because the scale factors are less than one, the trace can not be seen. Thus, portions of the outer reference sphere are removed in order to see the actual trace and a

smaller reference sphere inserted into the plot. Note that in addition to the scale factors, careful inspection shows the trace to lie above and off to the right side of the smaller, inner reference sphere. This indicates biases on the z- and y-axes.

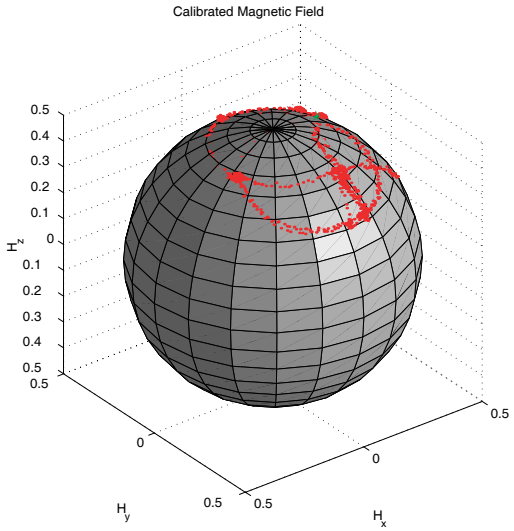

Figure 4-16 Plot of the measured components of Earth's magnetic field after calibration, along with a sphere of radius equal to the magnitude of Earth's magnetic field. Note the excellent agreement between the data and the surface of the sphere. This indicated that the algorithms has matched the biases and scale factors very well indeed.

Figure 4-16 shows the same data plotted after calibration. Note that the trace is now very well aligned with the surface of the reference sphere and that it does not penetrate the surface. This shows that the algorithm performs phenomenally well using even crude, noisy data. In order to better view the agreement of the calibrated data to the surface of the sphere, Figure 4-17 shows the magnitude of the measured magnetic field before (in blue) and after calibration (in red) through time. What can be seen here is the very poor agreement with a constant magnitude magnetic field before calibration and an excellent agreement afterwards. Indeed, the post calibration residuals are on the order of 0.005 gauss, or 1% of Earth's total magnetic field. This is well within the magnetometer specifications. Histograms of the residuals show them to be wide-band noise that could be low-pass filtered for a smoother solution. The residuals are slightly larger towards the end

of the experiment. This is due to dynamics; at this point the wooden boom was highly loaded and most likely flexing.

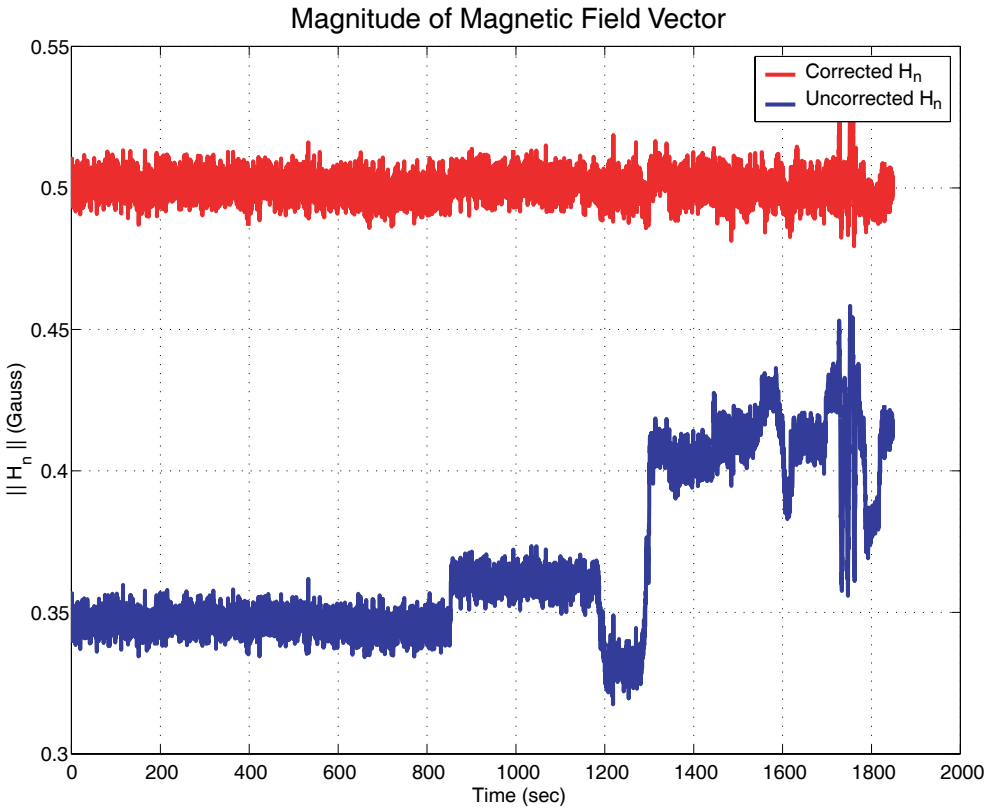

Figure 4-17 Magnitude of body-fixed magnetic field measurements before and after calibration. The initial period was used to generate static statistics on the sensors, and starting at approximately 800 seconds, the large motion was begun. Note that the pre-calibration data (shown in blue) was not even remotely constant, but that the post calibration data (shown in red) gives an excellent measure of the strength of the magnetic field.

The final check of the algorithm was performed in the heading domain. The heading error residual was computed by taking the difference between the heading generated by the magnetometer triad (post calibration) and the heading information from the INS. The errors were found to have a standard deviation of approximately 3 degrees and were again distributed as wide-band noise which could be filtered if this heading information were to be used for vehicle guidance.

Note that while this new calibration algorithm was developed and tested on a strap-down three-axis magnetometer, it is completely general. It could be used to calibrate any three-axis strap-down sensor as long as the magnitude of the total measurement remained constant and was known.

SECTION 4.8 EXPERIMENTAL VALIDATION

With the magnetometer calibrated and attitude estimation algorithms ready, a flight test was set up to validate the algorithms using real sensors and actual flight parameters. The flight test vehicle is a Beechcraft/Raytheon QueenAir, twin engine propeller driven aircraft. This aircraft has been heavily instrumented and has been used extensively in WAAS and LAAS experimentation at Stanford University. It is pictured below, in Figure 4-18.

SECTION 4.8.1 FLIGHT TEST SETUP

A block diagram of the experimental setup used to validate this algorithm is shown in Figure 4-19. Experimental validation of the two vector matching algorithm was performed on a Beechcraft/Raytheon QueenAir test aircraft (Figure 4-18). This aircraft is equipped with several high quality sensors. For attitude reference a Honeywell YG1851 IRU navigation-grade (1 nm/hr. drift) inertial navigation sensor (INS) unit was used to record position and attitude at 50 Hz. The data from the INS was recorded on a separate computer and later aligned with the GPS-time-tagged data. A second, less expensive inertial measurement unit (CrossBow DMU-FOG) was used to send data to the main computer at a rate of 125 Hz. These measurements are also aligned to GPS time internally.

Figure 4-18 Beechcraft/Raytheon QueenAir used for experimental testing of attitude algorithm. This aircraft is equipped with a short baseline GPS attitude system, a Honeywell HMC-2300 digital three-axis magnetometer, Crossbow DME-FOG, and a navigation grade INS. The aircraft was instrumented using the gyroscope-free quaternion attitude estimator and flown through various steep turns. The attitude was found to be in excellent agreement with the INS.

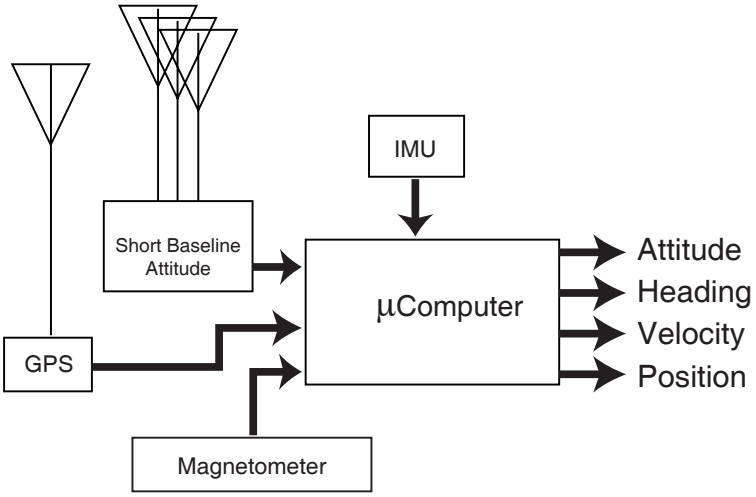

Figure 4-19 Simple block diagram of hardware used for experimental validation of attitude system. The aircraft, a Beechcraft/Raytheon QueenAir is equipped with many high quality sensors. A short baseline attitude system developed at Stanford University, along with a Honeywell navigation grade INS provide attitude. In addition, a Crossbow fiber-optic attitude heading reference system and a Honeywell HMC-2300 digital three-axis magnetometer are all synchronously sampled and aligned with GPS time.

The raw output of the Crossbow DMU accelerometers was used as the body-fixed accelerometer measurements, $\tilde{g}^b$. The local level acceleration, $\tilde{a}$, was computed by differencing velocities derived from GPS augmented by WAAS (again, in a numerical differentiation scheme). This measurement was used to correct the aircraft apparent gravity (accelerometer specific force) measurements during turns. A low cost magnetometer triad (Honeywell HMR 2300) was used to measure Earth's magnetic field vector in the body frame, $\vec{\tilde{B}}^b$. The output of the magnetometer was recorded serially on the main computer and aligned with GPS-time at 1 Hz.

The known portion of the vector measurements, Earth's gravitational field and magnetic field vector in the navigation frame (i.e., $\tilde{g}^n$ and $\vec{\tilde{B}}^n$), are well known and readily modeled. Earth's magnetic field was modeled using Schmidt-normalized coefficients from the 1995 International Geomagnetic Reference Field [17] as a function of GPS-position. The gravitational field of Earth was assumed to remain down in the local level coordinate frame (navigation frame), and to have a magnitude of 9.81 m/s^2. The accelerometers of the DMU are MEMS accelerometers similar to those used in the final system.

The test flight lasted approximately 45 minutes from takeoff to landing and included many steep turns. Pitch and roll doublets were also performed in order to deviate from straight and level flight. Following the flight, the data was aligned to GPS-time in post-processing, and the magnetometer and accelerometer data were aligned to each other in order to simulate running the system in real time. Data from the Honeywell INS was interpolated to match the sample points of the magnetometer and accelerometers. Finally, the difference of GPS velocity measurements (numerical differentiation) were used to obtain the measurement of $\tilde{a}^n$. This measurement could have been improved by using a simple Kalman smoother, but for this experiment the added complexity was not deemed justified.

The accelerometer bias and scale factor calibration was already performed internally by the Crossbow DMU, thus no further processing of the accelerometer data was required. The magnetometer required calibration for misalignment errors, hard and soft iron errors (bias), and scale factors errors. The method for calibrating the magnetometer was discussed previously in Section 4.7 on page 122.

The Kalman filter implementation, previously described, was used to estimate the attitude using the post-processed data. An initial guess of the quaternion estimate, $\hat{q}$, and the initial error quaternion, q_e, was used and the Kalman filter run sequentially through the post-processed data. It is important to note that the Kalman filter is providing an estimate of q_e, not $\hat{q}$. The output of interest is estimated attitude, $\hat{q}$, and is computed from q_e continuously.

SECTION 4.8.2 FLIGHT TEST RESULTS

Figure 4-20 compares the attitude versus time for the flight. The green line is the attitude computed by the quaternion based solution to Wahba's problem. Note that this estimate is using the Gauss-Markov based Kalman filter with no gyroscopes. Truth is presented as the red dashed line, and is the attitude recorded by the Honeywell INS. In general, agreement with the INS is good, considering the cost of the overall system. On this scale, the estimated and true attitude appear to track each other with virtually no lag whatsoever. Close inspection of the lower of the three panels will reveal rolls in excess of 50 degrees. Several quick roll reversals, as well as pitch doublets were included in the flight to exercise the attitude estimation system. Pitch excursions in excess of +/- 10 degrees are noted in the middle panel. While the agreement between the estimated and true attitude is indeed good, several areas of poor tracking (most notably in pitch) can be seen.

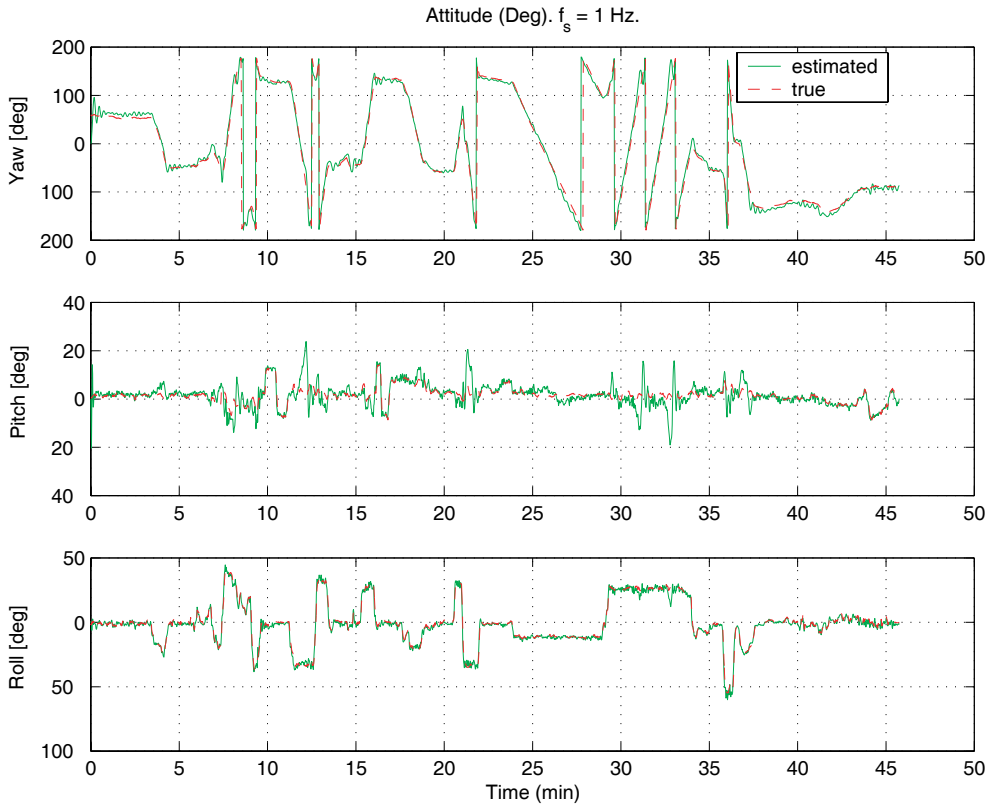

Figure 4-20 Performance of attitude estimation algorithm on actual flight data, presented in Euler angles. The algorithm does an excellent job of matching the INS output (red). A few notable exceptions can be seen, especially in pitch (for example, at 12 minutes into the flight). Note that at this time the yaw is at -90 degrees and decreasing, i.e., the aircraft is turning from west to south, and the local apparent down lines up closely with the magnetic field. This violates the initial assumption of two non-colinear vectors.

For the entire flight, the means and standard deviations for the vector matching Kalman filter (truth provided by INS) are summarized in Table 4-1. There are three main

	μ [deg]	σ [deg]
yaw	0.8313	10.507
pitch	-0.0959	3.6945
roll	-0.7570	2.0624

Table 4-1: Statistics from the gyroscope-free quaternion attitude estimation algorithm on flight test data given for Euler angles (degrees). Much of the large standard deviations is due to the excursions explained in the previous figure. Even with this deviation from the basic premise of the algorithm, the performance is excellent.

causes for the relatively large errors. The first is poor placement of the magnetometer inside the fuselage of the aircraft. The logical place to place the magnetometers are at the wing tip or on top of the rudder as far away from the electronics as possible. The second source of errors is due to a difference in time base between the recorded measurements of the INS and the attitude system. Without synchronous measurements, residual errors from interpolation remain present. The third reason is vector alignment as explained below.

Several times in the flight, the pitch measurement appears to diverge from the INS but then recovers. Close inspection of Figure 4-20 will show that the yaw oscillates from the true attitude at the same time. These divergences appear periodically, at 8, 13, 17, and 22 minutes into the flight, and during the time from 29-33 minutes and 36-37 minutes. The reason that the algorithm performs poorly during these transients is that during a coordinated turn the maneuvering aircraft periodically aligns the body z-axis with the magnetic field of Earth. This condition violates the basic requirement that the two vectors be non-colinear. During this transient moment, while the aircraft nose swings through due east or due west, the magnetic field vector is aligned with the aircraft specific force (apparent gravity), $\hat{a} - \hat{g}$. During this transient, there is insufficient information to determine aircraft attitude. The Kalman filter, however, continues to perform albeit at a degraded performance level even with the lack of gyroscopes. Rate gyroscopes would allow the attitude determination algorithm to bridge these outages without the large transients, but with increased cost and complexity.

It is important to clarify that the large transients are not a failure of the algorithm itself, but rather an operation outside the assumptions of the basic Wahba's problem. The roll angle is such that the local apparent gravity is approximately equal to the magnetic dip and the yaw angle lines up the aircraft apparent gravity vector with Earth's magnetic field. This only happens when the nose of the aircraft is turning through a yaw angle of +/- 90 degrees. Close inspection of the results verifies that this is indeed the case.

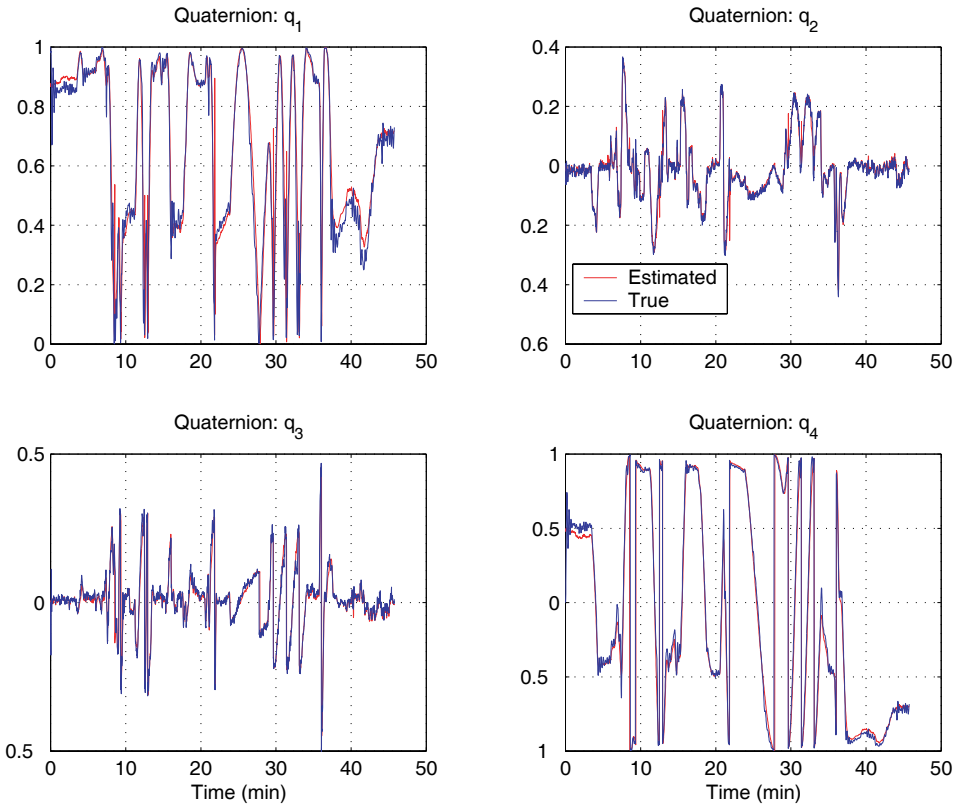

Figure 4-21 Performance of gyroscope-free quaternion attitude estimation algorithm on flight test data, presented as quaternions. Note that in the quaternion domain, the excursions in pitch are much harder to locate. Again, the performance of the algorithm matches the INS quite well, even during some rather "sporty" maneuvering.

Without these transients, where the basic assumptions of the algorithm are not violated, the tracking performance is much better. In order to find a best case scenario for the algorithm, the longest continuos section of maneuvering flight in which this vector aligning did not occur was used to gather statistics. As this turned out to be a four minute section with large turns and moderate pitch excursion, it is felt that the statistics are valid. Thus the section of the flight from 22 to 26 minutes of the flight, which includes a minute long steep banked turn, is used to compute best case scenario statistics for the attitude

algorithm. The tracking performance is much better and is summarized in Table 4-2 below:

	μ [deg]	σ [deg]
yaw	-0.8225	3.286
pitch	0.9517	1.3801
roll	-0.9959	1.2596

Table 4-2: Statistical performance of the gyroscope-free quaternion attitude estimation algorithm over best-case scenario on real data, presented as Euler angles (degrees). These statistics are compiled over a section of flight that included maneuvering, but did not swing the local apparent down in line with the magnetic field. The residual errors are very small, but are likely caused by a temporal misalignment between the magnetometer data and the INS data. Actual performance on the sailboat showed errors that were much smaller.

This should be considered as the best-case performance of the gyroscope-free quaternion-based attitude determination algorithm. Finally, the aircraft environment is very poor for the performance of the magnetometer, given the large electromagnetic transients that cannot be calibrated out of the measurements. Also, the interpolation and alignment of the INS data introduces some errors that are difficult to quantify.

SECTION 4.9 STATIC BOAT TEST

In order to differentiate between the aircraft tests and the expected performance that would be encountered on the sailboat, a series of static tests were performed. In order to perform the experiments, the boat was set upon its trailer at the center of a larger parking structure. Every attempt was made to ensure that the boat was level and oriented to point the nose of the boat towards true north. This was done with a magnetic compass and a large bubble level.

With the boat leveled and pointed north, the boat attitude system was turned on and allowed to collect data for approximately six hours. Unfortunately, no truth system

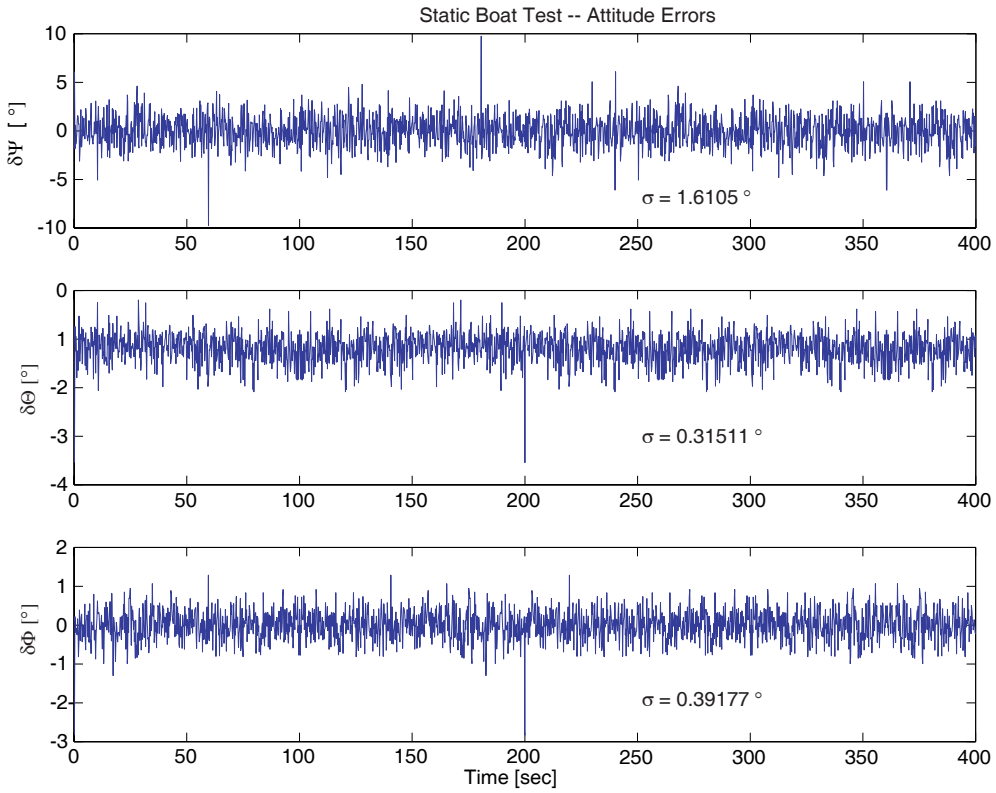

Figure 4-22 Static boat attitude errors. These errors are based on placing the boat trailer in a parking garage and leveling the boat as best as possible. A few minutes of typical data are presented. Because there is no independent source of truth, only the standard deviations of the attitude errors are reported. This shows that the standard deviations are approximately 1.5 degrees in yaw and less than 0.5 degrees in both pitch and roll.

was available, so that the bias in the attitude system could not be directly measured. Note that the biases were small and most likely attributed to poor alignment of the boat with north and level than errors in the sensors, but this cannot be separated. Based on this data, the standard deviations of the attitude system are reported in Table 4-3 below.

	σ [deg]
yaw	1.6105
pitch	0.3151
roll	0.3918

Table 4-3: Standard deviations of the static attitude data taken from the sailboat upon its trailer in the parking garage. Note that due to the lack of an independent measurement of truth, the mean of the error is not reported. Rather, only the standard deviation is presented for yaw, pitch, and roll.

SECTION 4.10 CONCLUSIONS

An attitude determination system was built for the Atlantis that is based on two vector measurements of non-zero, non-colinear vectors. The algorithm is based on a quaternion formulation of Wahba's problem, whereby the error quaternion (q_e) becomes the observed state and can be cast into a standard linear measurement equation. Using Earth's magnetic field and gravity as the two measured quantities, a low-cost attitude determination system was developed into a real-time attitude system. An iterated least squares solution to the attitude determination problem was simulated on static cases and shown to be globally convergent. A time-varying Kalman filter implementation of the same formulation is tested on simulated dynamic data, as well as on experimental data from a maneuvering aircraft.

In addition, a non-linear two-step estimation algorithm for calibrating solid-state strap-down three-axis magnetometers was developed. This calibration algorithm is general, and allows estimation of any three-axis vector measurement errors. If the sensor is perfectly aligned with the body axes (or the misalignment matrix has been separately estimated), then the procedure for estimating the errors does not require any external reference, only the magnitude of the vector that is being measured. In the specific case presented in this thesis, the estimated errors are used to calibrate the magnetometer for the attitude system. This calibration technique is both robust and accurate; the robustness was demonstrated experimentally by converging even in the presence of considerable sensor noise. Further, the algorithm required only a small portion of the ellipsoid locus for the estimation process to work.

The best case performance of the system in the aircraft was roughly 3 degrees in yaw and 1 ½ degrees for pitch and roll. Rate gyroscopes would improve the response, allowing the true dynamics to be tracked through the outages, but with the penalty of more

expensive hardware. Static tests on the catamaran showed much better performance, with errors of 1 ½ degrees in yaw, and less than ½ degree in pitch and roll, with no tendency to drift whatsoever. This low cost system was able to reliably provide a full attitude solution at a 10 Hz rate (limited by the GPS velocity output rate) at a fraction of the cost of gyroscope-based systems.

5 Propulsion System

The most visibly unique aspect of the Atlantis project is the wingsail propulsion system, as shown in Figure 5-1. This chapter details the design evolution that results in the unique configuration of the wingsail. The design considerations and goals are: equivalent performance to the original sail system, low actuation force, and the ability to precisely control the resulting system.

A sloop rig sail can achieve a maximum lift coefficient of 0.8 if the jib and sail are perfectly trimmed. Realistically, an operating maximum lift coefficient is 0.6. The design goal of the Atlantis wing is to achieve a maximum lift coefficient of 1.8. Since this allows the wing to generate three times the force of an equivalently

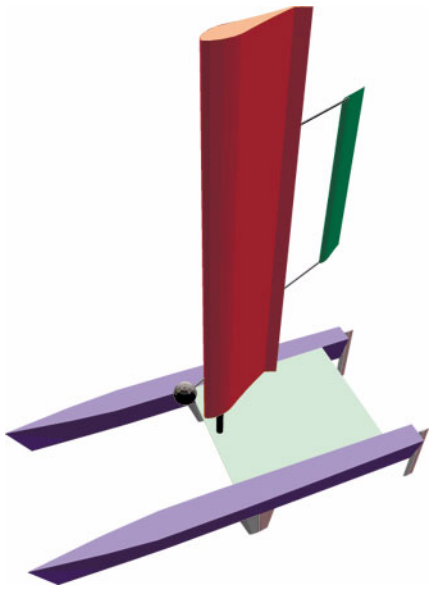

Figure 5-1 The engineering model of the Atlantis. The wing sail is 5.37 meters tall and has a chord of 1.45 meters. The self-trimming tail is used to balance the aerodynamic moments. The model includes a spherical mass attached to the leading edge of the wing to bring the mass center of the wing/tail combination in line with the stub-mast. In the actual prototype, the ball mass was replaced with an electronics pod attached to the forward end of the lower wing section.

sized sail, the wing area is reduced to one third of the area of the original sails. Because the drag characteristics of the wing are much improved, the performance of the wingsailed catamaran should be superior to the original configuration. At worst, the wing will yield equivalent performance.

Fundamental to the goal of autonomous operation is the requirement that the actuation of the sail be simple. In the case of a conventional sail, this would be extremely expensive in terms of actuator cost and power required as the forces required are quite large. Additionally, the complex nature of the aerodynamics of a sail makes any sort of precise control of the sail difficult to accomplish. In order to achieve precision control of the catamaran, the disturbances generated by the propulsion system must be minimized. Fundamentally, this requirement forces the design away from a conventional sail.

SECTION 5.1 DESIGN SPACE EVOLUTION

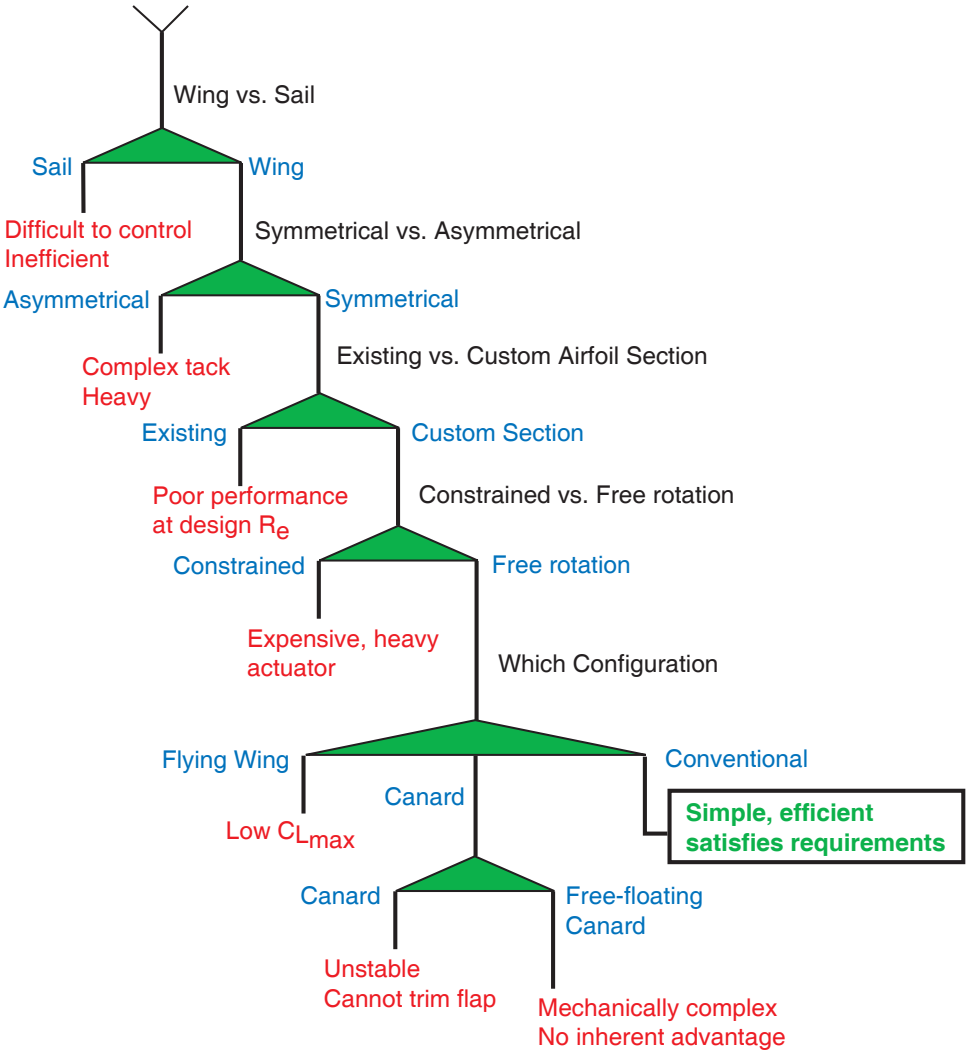

Figure 5-2 The design evolution of the propulsion system. The design choices are on the right of the figure in black. The choices are designated by the green triangles, with the winning choice to the right in blue. The red text explains the problem with the losing choice. The series of choices lead the design to a self-trimming wingsail with a conventional tail, using a custom designed airfoil section for the appropriate Reynolds number.

Figure 5-2 shows the design evolution of the wingsail for the Atlantis. The design choices are on the right of the figure in black. The choices are designated by the green tri-

angles, with the losing choice in black to the left, and the winning choice to the right in blue. The red text explains the problem with the losing choice.

The steps of the evolution are each detailed in later sections. The design evolution begins with a choice between a conventional cloth sail or a rigid wingsail. Then the choice is between a symmetrical or asymmetrical section. Following the symmetry choice, one must choose between an existing section and a custom designed airfoil section. With the section design complete, the next issue is to trim the wing aerodynamically or mechanically. Lastly, four possible configurations for the wing and trimming surface are considered.

The series of choices lead the design to a self-trimming wingsail with a conventional tail, using a custom designed airfoil section for the appropriate Reynolds number. The remainder of this chapter considers each choice in detail.

SECTION 5.2 WINGSAIL DESCRIPTION

The chosen wingsail is 5.37 meters tall and has a chord of 1.45 meters. It is constructed in three sections: the lower section which includes the forward electronics/ballast pod, the middle section to which the tail is attached by twin booms, and the upper section. The wingsail is built entirely of marine grade plywood covered in polyester fabric and is suspended by a spherical roller bearing at the top of the stub-mast. It is stabilized by a needle roller bearing around the stub-mast at the bottom of the wing. This allows the wing to rotate freely through 360 degrees without significant resistance. An engineering diagram of the wing is shown in Figure 5-3.

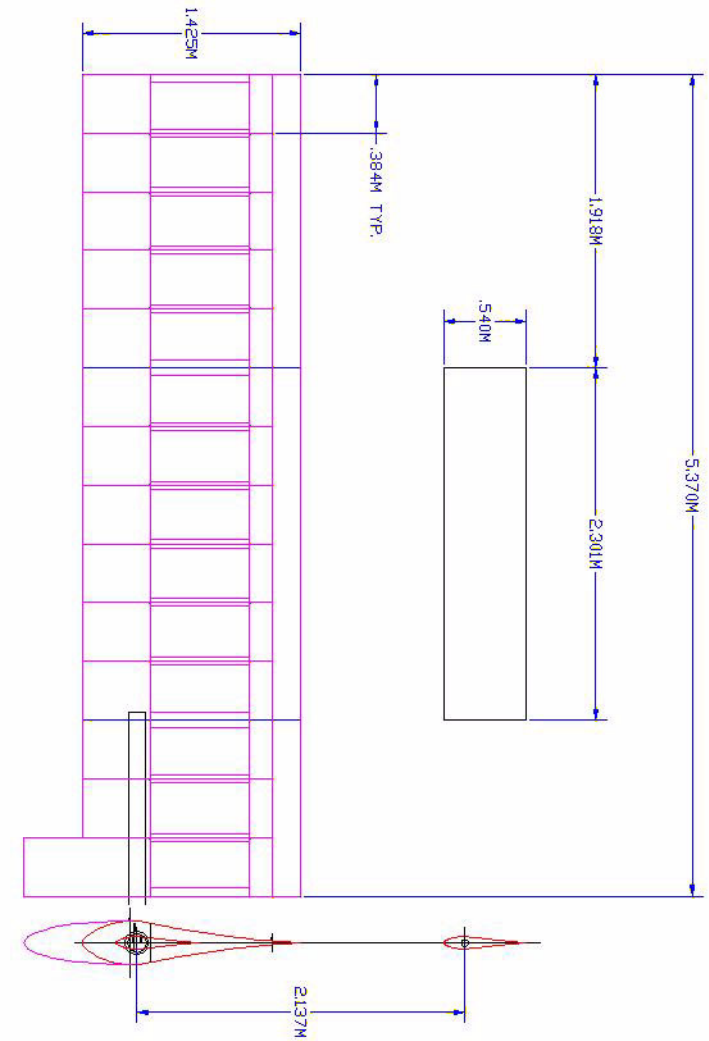

Figure 5-3 An engineering layout of the wingsail. This drawing shows the dimensions of the wing, the three sections, and the rib layout pattern. On the bottom it shows the overhead view, including the electronics pod at the front of the wingsail. Details of the wing structure and construction can be found at the end of the chapter.

SECTION 5.3 WING VERSUS SAIL

There are three main reasons to use a wing instead of a sail: efficiency, less actuation force required, and self-trimming.

The first and most obvious is that a rigid wing is far more efficient than a cloth sail. Though some attention needs to be given to Reynolds number effects, the coefficient of lift, C_L, has a maximum of 1.8 for the Atlantis wingsail versus typically 0.8 for a perfectly trimmed sloop rig (jib and mainsail). Also, the Lift/Drag (L/D) ratio of the Atlantis wingsail section is in the 30 - 100 range, whereas the L/D of the conventional sail is in the 3 - 5 range. Further, a cloth sail suffers from aeroelastic collapse when pointed high into the wind (the sail is said to be luffing). This causes a great deal of drag when sailing close-hauled and effectively limits how high the boat can point into the wind. The rigid wing, by contrast, suffers no aeroelastic problems; it can point straight into the wind with very little drag, no flapping, no whipping about, and no noise, while effectively reefing the wing. In fact, the feathered wing-tail combination has much less drag than the bare mast. This is demonstrated in Figure 5-4, which shows two sections (cylinder and airfoil) that have the same net drag (including both viscous and pressure forces). Because the two sections have the same drag, the ability to reef a sail (or reduce the area of the sail) is moot when using a rigid wing because the wing has far less aerodynamic load on it than the bare mast itself.

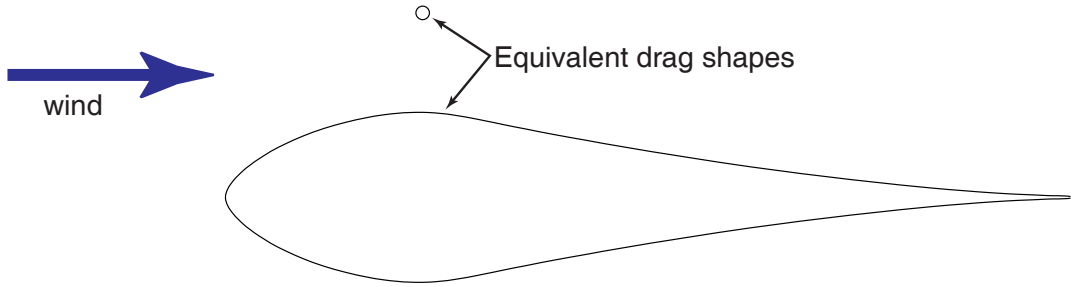

Figure 5-4 Demonstration of the equivalent drag sections at Reynolds number of 229,000. The small solid cylinder and the airfoil section have the same total drag (including both skin friction and pressure drag terms). A rigid wing need not be capable of reefing (or reducing its total area) in order to protect the boat. As demonstrated above, the wing (if allowed to pivot freely) will have much less force on it than the bare mast itself.

The second main reason to use a wingsail for propulsion is less force is required to actuate the wing itself. A cloth sail is fixed to the mast, and trimmed from the boom. Since the center of pressure of the sail is aft of the leading edge, the trim force must overcome a

portion of the lift of the sail. Inspection of a conventional sailboat shows a large block-and-tackle with eight or more loops of line attached to the boom is required to trim the main sail. With a winch, an additional 8:1 mechanical advantage is required to hold the boom in. To control this effectively in an automatic manner, a very large and fast-acting actuator is required. These types of actuators quickly become very expensive and a typical one would cost more than the entire budget for the project. By contrast, the wing can be designed to pivot near the center of pressure of the wing itself. The wingsail is turned to an angle of attack either directly or through an auxiliary trimming surface. In either case, this is accomplished with a small DC motor and can be actuated quickly and inexpensively. The cost effectiveness of this design is the main reason it was used for this project.

The third main advantage of the wingsail over the conventional sail is the ability to make the wingsail self-trimming. The benefit of this is that the wing will absorb gusts without transmitting the force of the gusts through to the guidance system. By decoupling the propulsion system from the guidance system through passive stability (self-trimming), the control system design is greatly simplified. Through proper arrangement of the flying surfaces, the wingsail will readjust automatically to a change in either wind speed or wind direction, with no intervention from pilot or control system.

The self-trimming capability makes the wingsail ideal for an autonomous sailboat because it eliminates the requirement for a very large and fast acting actuator to constantly retrim the sails. The only time that direct intervention into the trim control of the wing is required is when the wing crosses the longitudinal centerline of the boat. During this maneuver, the flap and tail are reversed from their previous positions. Note that in a conventional sense this corresponds to tacking (when the wind is from the front of the boat) and jibing (when the wind is from the stern of the boat). The maneuvers using the wingsail are very gentle and controlled because the bearings allow the sail to rotate 360 degrees

about the mast without interference and the wing can point straight into the wind with no ill effects.

Conventional sails have one serious advantage: due to their sharp leading edge, they tend to be insensitive to Reynolds number variation. This alone may explain why they have persisted on modern designs even after the preponderance of evidence has demonstrated that wings are vastly superior. The other advantage that cloth sails may have over rigid wings is weight: for sails below a certain size, a rigid wing will almost certainly be heavier. This is due to the square-cubed law with respect to the strength of structures. Above a mast height of approximately 20 meters, the structure of the mast could just as easily be incorporated into the spar of a wing. In [93], a race between two similar catamarans with a rigid wing and a conventional sail was analyzed. The winged catamaran had superior aerodynamic thrust on all points of sail, but the difference of ~150 pounds required a wind speed of greater than 8 knots for the superior aerodynamics to result in superior boat speed. The greater weight led to greater drag on the hulls due to the extra displacement of water. In the race, all legs that were raced at wind speeds greater than 8 knots were won by the winged catamaran, but all legs below 8 knots were won by the conventionally sailed catamaran.

SECTION 5.4 REYNOLDS NUMBER EFFECTS

As mentioned previously, the Reynolds number effects of the wingsail section design must be accounted for in order to maximize the efficiency of the wing. Ignoring these Reynolds number effects has been the largest failing of wingsails to date, resulting in sections with poor performance in the field, and, in turn, delaying the transition to rigid wings on sailboats.

The Reynolds number, R_e, is defined as:

$$R_e = \frac{\rho V L}{\mu} \quad \text{(EQ 5.1)}$$

where ρ is the density of the medium, V is the velocity of the flow, L is the characteristic length, and μ is the viscosity of the medium. The Reynolds number represents the ratio of kinematic or inertial forces to the viscous forces in the fluid (that is, the ratio of force required to push the fluid out of the way versus the force required to slip through the "goo-eyness"). Typically, insect flight has R_e on the order of 100s to 1000s, bird flight and models in the 100,000s, small aircraft in the millions range, and large aircraft in the tens of millions range. The Reynolds number characterizes the flow's ability to negotiate the curves of a section without separation. Illustrative of this is Figure 5-5 which is reproduced from [123] and demonstrates the different drag characteristics of a 2-D cylinder as a function of Reynolds number.

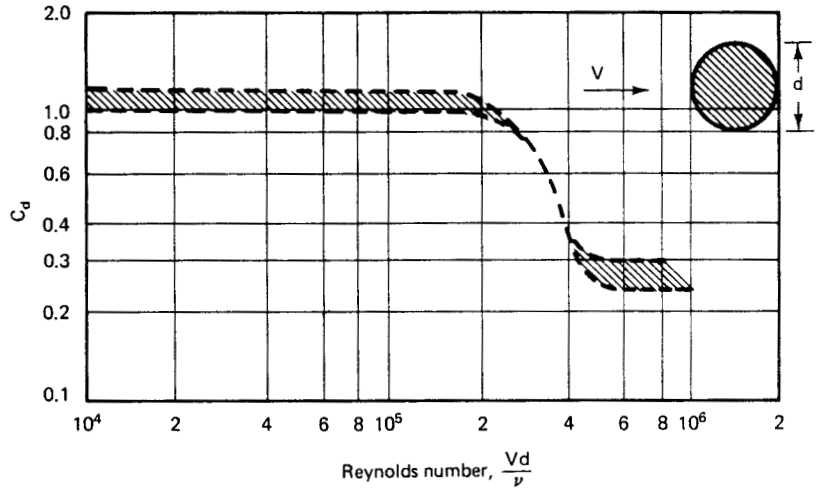

Figure 5-5 Effect of Reynolds number on the drag coefficient of a 2-D cylinder. At low Reynolds numbers, the drag remains constant. This is consistent with separation occurring just aft of the maximum diameter of the cylinder. As the Reynolds number increases, the now turbulent flow is able to negotiate the curvature better. This causes the drag coefficient to drop until the point that the flow remains attached approximately 1/3 of the way down the back side of the cylinder, at which point the drag coefficient once again becomes constant with increasing Reynolds number.

In the case of airfoil sections, several effects come into play at low Reynolds numbers that make design of high lift sections difficult. Most of these are discussed at length in [122]. Flow about an airfoil at low Reynolds numbers is almost entirely laminar. Thus the flow can withstand neither sharp radii nor severe adverse pressure gradients without separation (and consequential very large drag rise). In the case of airfoil sections, the flow separates, but then reattaches causing a laminar separation bubble whose flow eddy results in a very large increase in the base drag of the section. Furthermore, as soon at the section is put at an increased angle of attack, the laminar separation bubble bursts, causing large scale flow separation and effectively limiting the maximum lift coefficient, C_L, attainable.

In [25], the designers demonstrate a knowledge of the difficulties in designing good sections at these Reynolds numbers, but fail to capitalize on this knowledge and find an appropriate design. They correctly identify the proper Reynolds number range for sail operation as 200,000 to 1.2 million, and complain that "good low Reynolds number aerodynamic data applicable to sails is not readily available." However, low Reynolds number computer codes had not yet reached maturity at the time they were investigating superior sails. Thus, while correctly identifying the problem, they did not find the ideal solution.

The reason that low Reynolds number airfoil sections do not exist for this range has to do with the unique requirements of sailing vehicles. Typically, this Reynolds number range corresponds to small model airplanes, usually gliders. The differences are subtle and will be exposed in detail forthwith.

First, both the model glider and the sailboat require a high lift/drag (L/D) ratio. In a glider, this corresponds to glide distance, or a minimum sink condition. In a sailboat, this corresponds to the ability to point upwind. Second, both a model glider and a sailboat require a high maximum C_L. In the case of the glider, this corresponds to slow flight while

circling tightly in thermals; in a sailboat, the configuration is maximum speed while sailing across or down wind.

At this point, the requirements are essentially the same and there should be a large body of work on appropriate sections that can be used for the sailboat wing. Given the constraints of designing a low Reynolds number airfoil section for the wingsail, there are a few details to consider. The wingsail section must be symmetrical so that the Atlantis may sail on both port and starboard tacks.

SECTION 5.5 SYMMETRY

An airfoil section can be made either symmetrical or asymmetrical. An asymmetrical section can always achieve a higher maximum lift coefficient and a higher lift/drag ratio than a symmetric section. Symmetric sections have the advantage of identical lift characteristics with both positive and negative angles of attack. Symmetry arguments become important in sailing vessels because a sailboat is required to sail equally well on both port and starboard tacks and thus the section must be symmetrical. The model glider is rarely required to fly inverted, and certainly not for long periods of time. Thus model glider sections are always asymmetrical in order to maximize the L/D.

Certain sailboats, including the designs in [118], [26], and [34], attempt to capture the maximum L/D by using an asymmetrical section, but then tack and jibe "over-the-top." This means that the wing is pinned midway up its span, then flipped to a horizontal position, and finally the bottom and top ends are then switched as the tack or jibe is completed. This is demonstrated in Figure 5-6. Needless to say, this results in an extremely heavy structure at the pin joint as well as an exposed support or mast which greatly

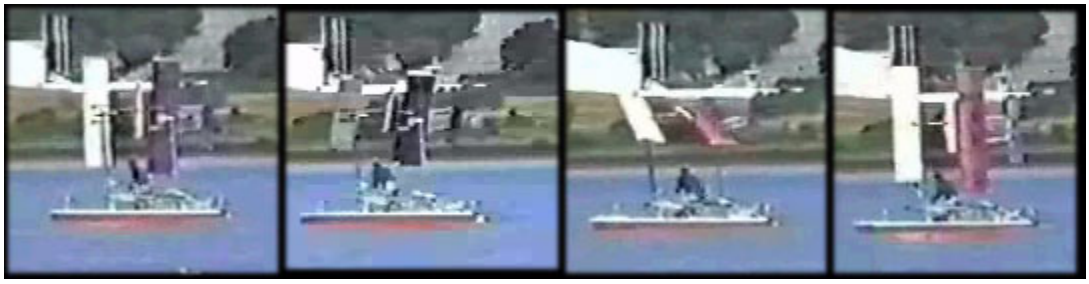

Figure 5-6 Frames from video showing the Boatek wingsail performing an over-the-top tack. From left to right, the boat is first turned into the wind. As the angle to the wind approaches zero, the wing (in this case, wings) are pivoted horizontally midway up the wing. As the boat turns through the wind, the wings are lowered to the opposite position from the first frame and secured. The top and bottom ends of the wings having been swapped, the tack is now complete.

increases the overall drag on the superstructure of the boat. It also makes wing control during this maneuver difficult in strong winds.

Using modern airfoil design techniques and a simple plain flap, one can achieve very close to the maximum C_L of an asymmetrical section. Thus, the increased weight, complexity, drag, and loss of the ability to self-trim in an asymmetrical design seem hardly worth the effort. Indeed, the ease of handling a symmetric section which does not pivot horizontally about the mast allows an increase in wing area, thus making up for the lower maximum lift coefficient. Although some continue to advocate over-the-top designs, they seem to stem more from novelty than an true understanding of aerodynamic trade-offs.

SECTION 5.6 AIRFOIL SECTION DESIGN

The first step in designing the best performing airfoil section is determining the appropriate Reynolds number, then achieving the best lift with the most benign characteristics. Second, it is desirable for the section to achieve a maximum lift coefficient of 1.8 at a Reynolds number range of 200,000 to 250,000. This can be aided by a simple plain flap

of constant flap/chord ratio. The pitching moment coefficient must be small with the flap in trail so as to be easily balanced by the tail. Third, the greater the lift/drag ratio, the better will be the upwind performance of the Atlantis.

In order to match the total force on the wing at a wind speed of 5 knots with a theoretical lift coefficient of 1.8, Equation 5.1 is solved for a resulting Reynolds number of 229,000. Figure 5-7 shows the wind velocity required to achieve this Reynolds number as a function of angle from the true wind. This is based on the assumption that the sailboat can sail at one third of the true wind speed. The figure shows that the range of minimum wind speeds is from 3.8 to 6 knots.

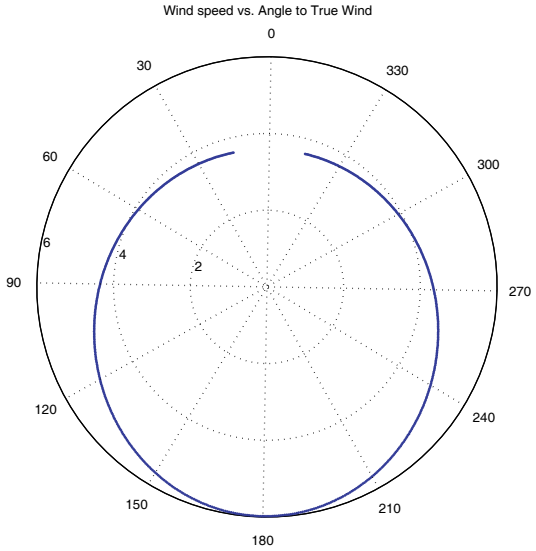

Figure 5-7 Polar plot of the true windspeed versus the angle to the true wind. This plot shows the velocity of the wind to reach a minimum speed necessary to have a Reynolds number of at least 229000. The sailboat is assumed to be able to sail at one third the speed of the true wind. This results in a range of 3.8 knots at 15 degrees to the true wind to 6 knots running directly before the wind.

The wing has one third the area of the sails, but generates three times the lift at its design point. This was chosen to enable a comparison of performance between the wing and sail. Note that the final design gets better as the Reynolds number increases. The diffi-

cult thing to achieve is low Reynolds number performance. The same airfoil section can achieve a higher coefficient of lift at greater Reynolds number.

In order to achieve the desired goals of maximum lift coefficient of 1.8, lift/drag ratio of better than 20, and optimization for a Reynolds number of 229,000, a rather unusual design emerges. First, in order to achieve the high lift coefficients at low Reynolds numbers, a very thick section is required, where the entire lift is generated on the forward section, typical of the Liebeck “rooftop” sections. The boundary layer requires a trip-strip that will force the transition from laminar to turbulent, placed symmetrically on the top and bottom surfaces. Typically, these trip-strips are a thick material with a zig-zag leading edge that is affixed to the surface at the desired location. The zig-zag causes a small-scale vortex to form which pulls in the higher energy flow outside of the boundary layer, and though viscous drag increases, separation (and thus form drag) is delayed.

In addition to the short, flat pressure distribution on the section, the entire aft portion of the section is given to pressure recovery of the flow preventing flow separation from the section surface. Thus the back three quarters of the section do not contribute at all to the lift, but merely ensure that the airflow can recover to free stream conditions gracefully.

The tail section is also designed to have a flat, rooftop pressure distribution, using trip-strips to force a boundary layer transition. The design goals for the tail are simply to achieve a maximum lift coefficient of 0.5 at a Reynolds number of 44,000. The design methodologies and tools for the wing and tail sections are virtually identical, as explained later.

SECTION 5.6.1 ANALYSIS TOOLS

In order to design the wing and tail sections, modern computational fluid dynamics (CDF) computer codes are used to predict performance and refine the design of the sections. The two main codes utilized for this are Ilan Kroo's PANDA and Mark Drela's XFOIL.

PANDA, which is an acronym from **P**rogram for **AN**alysis and **D**esign of **A**irfoils, was developed by Professor Ilan Kroo in the 1980's at Stanford University[84]. The program computes and graphically displays the pressure distribution (in coefficient form) on airfoil sections in subsonic flow. For a particular airfoil with coordinates stored in a standard text file, the program calculates the inviscid pressure distribution over the airfoil at a specified angle of attack and Mach number; lift and pitching moment about the 1/4-chord point are also computed. The analysis is done with remarkable speed (less than a second) so that the effects of changes in angle of attack or airfoil geometry can be studied easily. The program also computes the boundary layer properties based on this inviscid pressure distribution. The location of transition, laminar or turbulent separation, and total drag are computed based on integral boundary layer methods. It is possible to specify a position for "transition grit" or "trip-strip" on the upper and lower surfaces to force transition or model surface roughness.

A major feature of the PANDA program is its provision for rapidly changing the airfoil geometry. This is done by positioning the cursor over the part of the airfoil to be changed and clicking the mouse button. A smoothly-faired bump (with specified but editable height and width) is added to the section at this point and the new pressure distribution is quickly redrawn (the normalized pressure is referred to as the coefficient of pressure, *Cp*). In this way the airfoil can be rapidly reshaped to produce a desirable *Cp* distribution.

XFOIL 1.0 is a computational fluid dynamics (CFD) code that was written by Mark Drela in 1986 at the Massachusetts Institute of Technology (MIT), see [39], [40], and [41]. The main goal was to combine the speed and accuracy of high-order panel methods with the fully-coupled viscous/inviscid interaction method used in the more sophisticated codes developed by Drela and Giles. A fully interactive interface was employed from the beginning to make it much easier to use than the traditional batch-type CFD codes. Several inverse modes and a geometry manipulator were also incorporated early in XFOIL's development, making it a fairly general airfoil development system.

XFOIL is a much more full-fledged code than PANDA, able to operate well into the low Reynolds number regimes with excellent predictive capabilities. It also includes the ability to use either free or forced boundary layer transitions and to predict lift and drag polars to just beyond the maximum lift coefficient.

SECTION 5.6.2 WING SECTION

Section development starts with a National Advisory Committee for Aeronautics (NACA, the predecessor of the present NASA) 00xx section to probe the design space. The NACA section is then modified using PANDA until reasonable performance was achieved. At this point, the section coordinates are transferred to XFOIL which is used to iterate on the pressure distribution and boundary layer trip-strip location until the desired results are achieved.

The first attempt used a NACA 0015 airfoil section. Originally developed in the 1930's, the NACA 0015 is a symmetrical section with a thickness to chord ratio of 15%, and designed as a turbulent section. While this airfoil section is known to have poor performance at low Reynolds numbers, as the *de facto* standard for symmetrical sections, it functions as a benchmark against which to compare all other attempts. Further, land yacht designers are using NACA 00xx sections almost exclusively in their successful designs.

Part of their rationale behind this choice is the observation that the main drag source is not parasitic drag but rather induced drag. Since induced drag is largely a function of the aspect ratio of the wing and the load carried by the wing, the effect of airfoil section is minimal. This gross analysis, however, fails to take into consideration the loading variation of the wing and the problems of stall and separation. While the wing might be flying at a coefficient of lift below stall, sections of it might be above due to variations in wind speed with height (wind gradient) or effective twist due to the same effect. These problems can only be addressed with high maximum lift coefficient and the NACA 0015 simply cannot provide it. Figure 5-8 shows the poor performance of the NACA 0015 at low Reynolds numbers, where the flow is largely laminar. Note the laminar separation bubble ("LS") on the top surface, the turbulent separation ("TS") indicating a stall, and the rather low lift/drag ratio at this condition.

The laminar separation bubble, indicated by the "LS" in Figure 5-8, can be seen more clearly if we zoom into the section. In Figure 5-9, we take a closer look at the laminar separation bubble and can clearly see the effect of the boundary layer growth and subsequent contraction as the flow reattaches following the laminar separation. While the presence of the laminar separation bubble is invisible from a macroscopic view, it nonetheless affects the entire flow of the section. The maximum lift coefficient attainable is directly related to the presence or absence of the laminar separation bubble and the manner in which the section stalls is also driven by its presence. When a laminar separation bubble is present, the stall is likely to occur at the point of the bubble rather than at the trailing edge. This results in a sudden loss of lift and increase in drag, rather than a gradual loss of lift and increase in drag.

The final design, after many iterations, results in a rather unusual shape. First, the final wing section is enormously thick, with a thickness to chord ratio of over 21%. The distribution of that thickness is predominately toward the nose of the section. This is con-

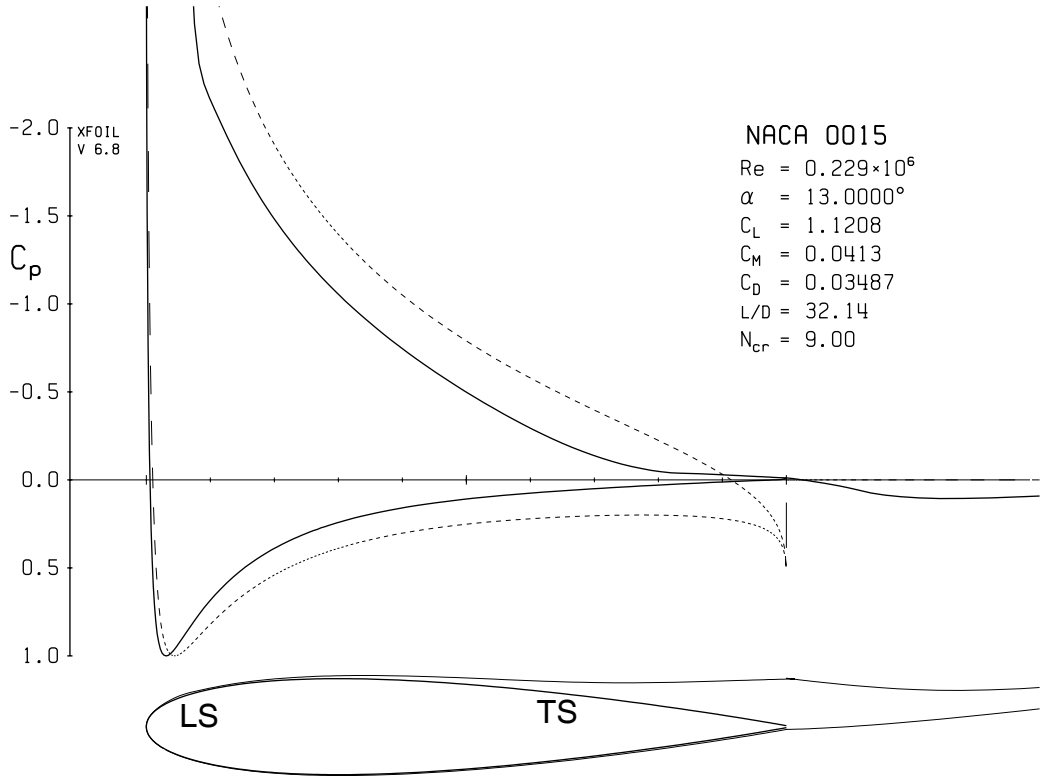

Figure 5-8 XFOIL results for NACA 0015 airfoil at Reynolds number of 229,000 and C_L of 1.12. This is a turbulent boundary layer section. Note the very sharp pressure spike corresponding to the leading edge of the airfoil (with a $C_p < -4.0$ for a lift coefficient of 1.12) that will most likely cause separation. Indeed the laminar separation bubble is marked by "*LS*" and the trailing edge turbulent separation indicative of stall is marked "*TS*."

Figure 5-9 Close up view of the laminar separation bubble on NACA 0015 airfoil at Reynolds number of 229,000. At this scale, the enlargement and then, farther along, contraction of the boundary layer is clearly visible. Inside the enlarged section an eddy vortex is stationary and consumes energy in its rotation. This results in increased drag. As the angle of attack increases, the vortex tightens and eventually bursts, resulting in turbulent boundary separation and stall.

sistent with the requirement that most of the lift is generated at the front part of the section, in front of the boundary layer trip-strip, while the entire aft section is there only for pressure recovery.

Close inspection of the section will show that the post boundary trip curvature is in fact concave, making construction using a normal cloth covering somewhat of a challenge. As the cloth covering shrinks, it will tend to pull off of the curved rear section of the airfoil since a straight line connecting the point of maximum cross section and the end just before the flap hinge is shorter than the actual surface. Looking at the pressure distribution in Figure 5-10, one can immediately see the design challenges that were presented and how they were solved. Note the absence of either laminar separation bubbles or turbulent separation at the end of the section. This is at a C_L of 1.04, with no flaps deployed.

Again, it is important to point out the salient features of the pressure distribution shown in Figure 5-10. Observe the flat top of the pressure distribution, corresponding to a uniform suction on the upper front surface. The pressure begins its recovery just after the trip strip located at the 22% chord point and very smoothly recovers back to free stream pressure without separation. Note that the flow is actually accelerating on the lower surface below the stagnation point. This causes the upward slope of the lower line in the pressure distribution, indicating some suction existing at the maximum chord point of the final wing section. Also, just after the trip-strip lies a very smooth pressure recovery all the way to the rear point of the airfoil section. Reemphasizing, there are no laminar separation bubbles and no turbulent separation. This airfoil section is not close to stall but will stall gently from the rear progressing forward, resulting in a very gradual loss of lift and increase in drag. This is important due to the varying nature of the wind (which is evidenced by data presented in Chapter 7). When the wind is highly variable, a conventional section like the 0015 will often abruptly stall and lose lift.

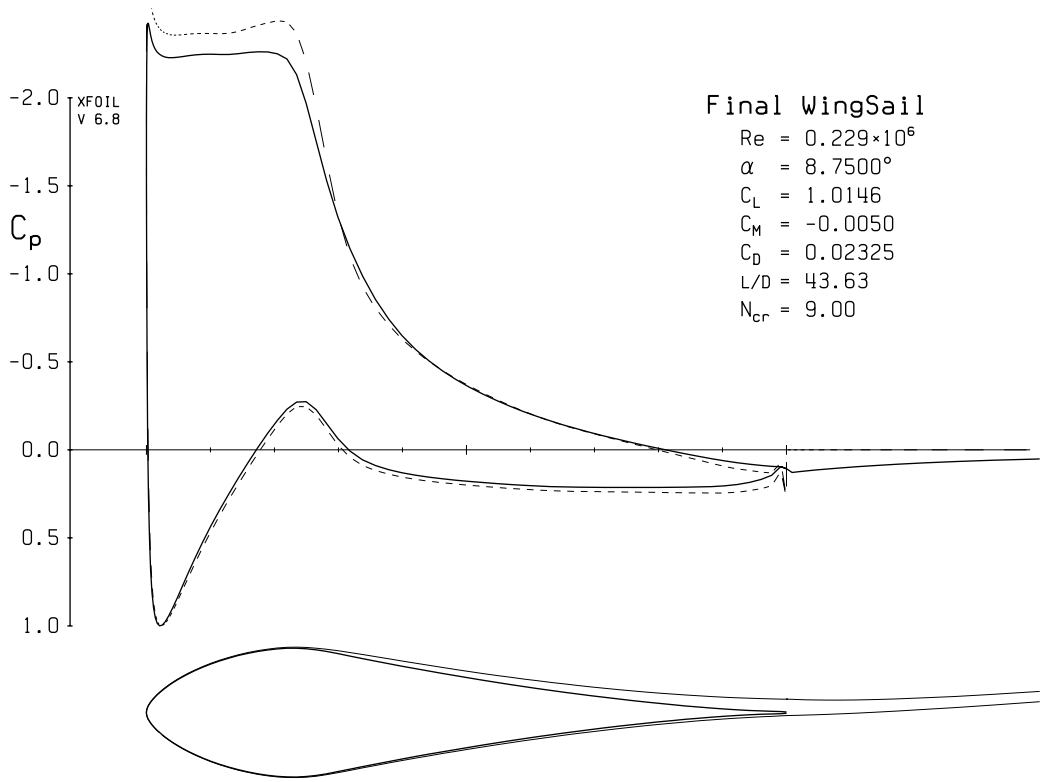

Figure 5-10 Final wing sail airfoil section and pressure distribution, Reynolds number of 229,000 and a coefficient of lift of 1.0. The pressure distribution is shown in the standard manner, with $-C_p$ along the y-axis, and the normalized chord along the x-axis. This section demonstrates a "rooftop" pressure distribution that rises immediately to a value of -2.5 and stays there for the 25% of the airfoil section. There, the boundary layer is tripped to force a transition to a turbulent section, and a long slow pressure recovery is used to prevent separation.

SECTION 5.6.3 TAIL SECTION

The design methodology for the tail section is virtually identical to that of the main wing section. The differences here are that the tail section will not be flapped, and because of its narrower chord, the design Reynolds number is much lower, around 44,000. Again, the same methods are employed, first using PANDA and then converging on the final design with XFOIL. A trip-strip is needed and has been placed at the 20% chord-wise location due to the inability to withstand adverse pressure gradients. Figure 5-11 demonstrates the flat forward rooftop pressure distribution along with the gradual recovery to

free stream pressures. Further investigation demonstrates that the tail has an expected lift coefficient of 0.75 before turbulent separation begins at the rear of the section. This once again allows a gradual and smooth change in the lift and drag characteristics of the section without compromising the maximum lift that can be generated. Note that there are no laminar separation bubbles, and at the lift coefficient of 0.5, there is no trace of turbulent separation at the rear of the section. Thus, all of the design requirements of the tail section are met and the relatively large thickness to chord ratio allows a robust structure to be built using conventional materials such as foam, plywood, and polyester fabric.

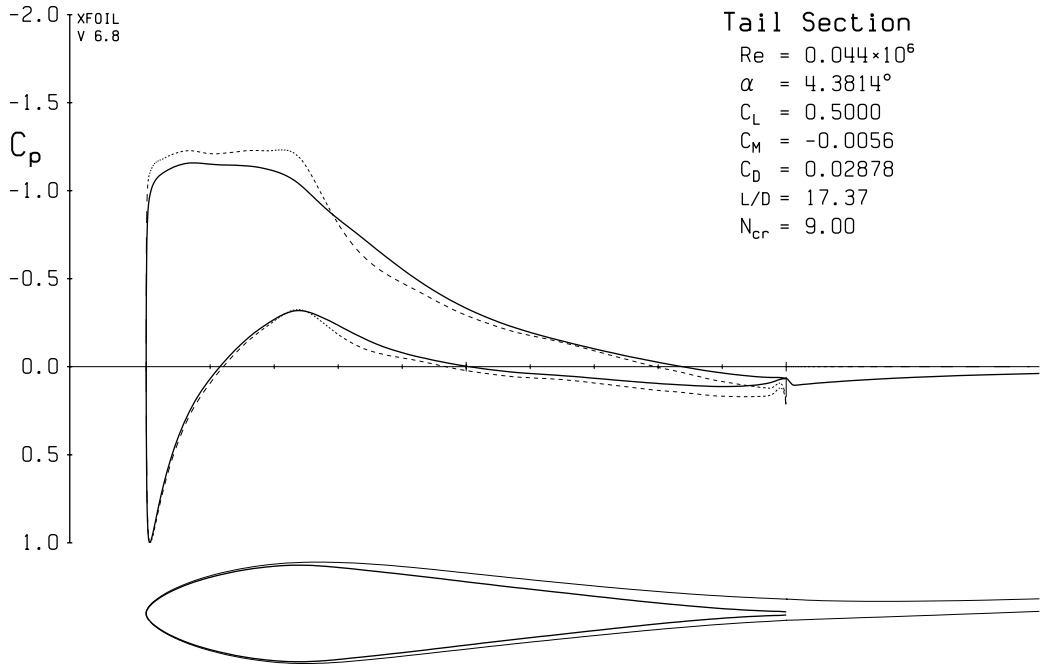

Figure 5-11 Pressure distribution of the final tail section at a Reynolds number of 44,000 and a lift coefficient of 0.5. This section is very similar to the final wing section. It shows the same “rooftop” flat forward section on the pressure distribution. The lift coefficient is 0.5, with no trace of laminar separation bubbles nor any turbulent separation. Further analysis using XFOIL indicates that this section can reach a C_L of 0.75 before stall.

SECTION 5.6.4 FLAP/CHORD RATIO

In order to increase the coefficient of lift of the main wing section and obviate the need for “over-the-top” tacking and jibing, a simple plain flap is used to increase the chamber of the wing. Figure 5-13 shows the pressure distribution with the flap deployed at 45 degrees. Note that the flow separates off the back of the flap causing an increase in drag. Unfortunately, at these low Reynolds numbers, the flow cannot negotiate the curvature of the flap hinge regardless of where it is placed on the airfoil section. This means that the low pressure side of the flap will separate as soon as it is deflected more than a degree or so. With this constraint, the issue becomes one of trading the separated flow and subsequent drag for increased effective chamber of the section and increased lift. Thus, the low Reynolds number flow pushes the design toward a very small flap/chord ratio and large deflection. In other words, a small trailing edge tab deflected a great deal will turn the flow enough to give effective chamber, while giving the flow only the smallest area from which to separate.

In order to find the optimum flap/chord ratio, a grid point search is performed using XFOIL in order to find the minimum drag at a coefficient of lift of 1.8. The flap/chord ratio is varied from 1% to 40% in 1% increments and the optimum is found to be at 13%. Figure 5-12 shows the results of these computations. Both the maximum attainable lift coefficient as well as the lift to drag ratio at that lift coefficient are presented. It can be seen that both reach there maximum close to a 13% flap to chord ratio.

The final shape for the main wing section is presented in Figure 5-13, with the flap deflected 45 degrees. The aggregate plots of the lift to drag coefficients for the final section with the flap deployed can be found in Figure 5-14. The plot shows that there exists an “efficient boundary” where the lift/drag ratio is maximized for a given lift distribution. This will then become the basis of control: once the desired lift coefficient is determined, the correct flap setting can be chosen to minimize the drag.

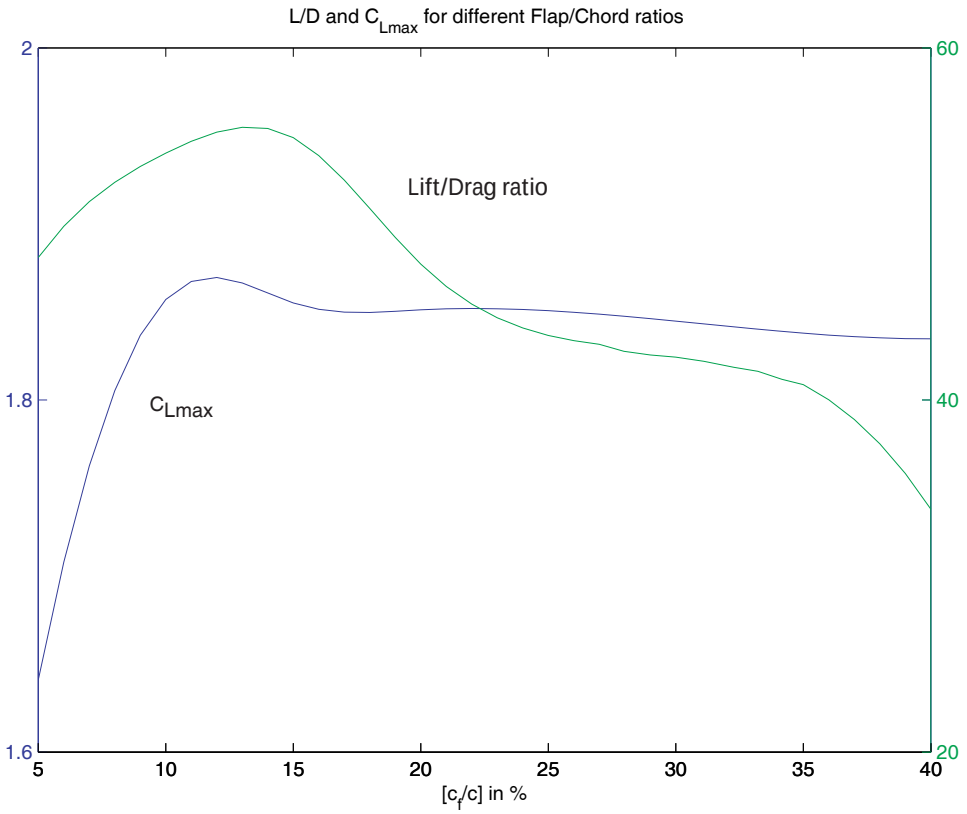

Figure 5-12 Results of the grid point search for optimum flap performance. Maximum attainable lift coefficient is plotted in blue and the lift to drag ratio is plotted in green. Note that while they both have a maximum in between 10% and 15% flap to chord ratios, the lift/drag maximum is much sharper in the area of 13%.

Note that above a C_L of 1.8, the drag continues to increase without any further increase in lift. This is expected from the increase in separation of the flow, and as predicted is gradual. Looking at the data in a different way, it is useful to visualize the lift/drag ratio as a function of either lift, drag, or angle of attack. Note that the angles of attack involved are uniformly small, implying that the control over the tail must be precise or the tail, and subsequently the wing sail, will be stalled for the duration of the sail test. The lift/drag ratios are plotted against the above mentioned parameters in Figure 5-15. It can be seen to have a peak at an angle of attack (α) of 2-3 degrees, with the flap set at 25 degrees. This, then, is the most efficient configuration at which to sail upwind. All other points of

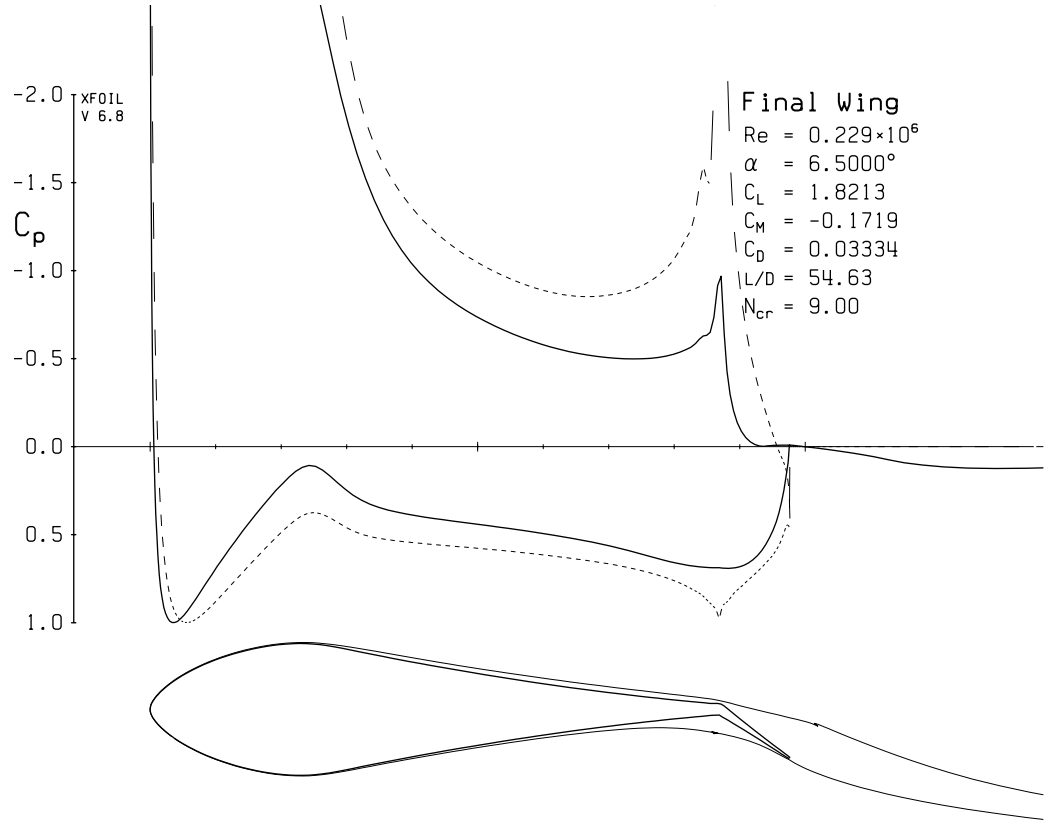

Figure 5-13 Pressure distribution of main wing sail section with flap deployed, Reynolds number of 229,000 and a coefficient of lift of 1.8. In order to preserve the lift/drag ratio of the section with the flap deployed, while attaining a high C_{Lmax}, a small trailing edge flap is used. At this Reynolds number, any flap deflection will result in separation. Thus, a narrow chord flap is deflected a large amount to generate a high effective camber. At the same time, this design minimizes the area of separation, and hence drag.

sail require that the wing sail provide the maximum force and then be modulated downwards as the threat of capsize increases.

SECTION 5.7 WING/TAIL CONFIGURATION ANALYSIS

Two possibilities exist for actuating the wing and controlling its angle of attack. The first possibility is to control the angle of attack mechanically using an actuator that rotates the wing about the mast. This has the advantage of quick actuation, and correct

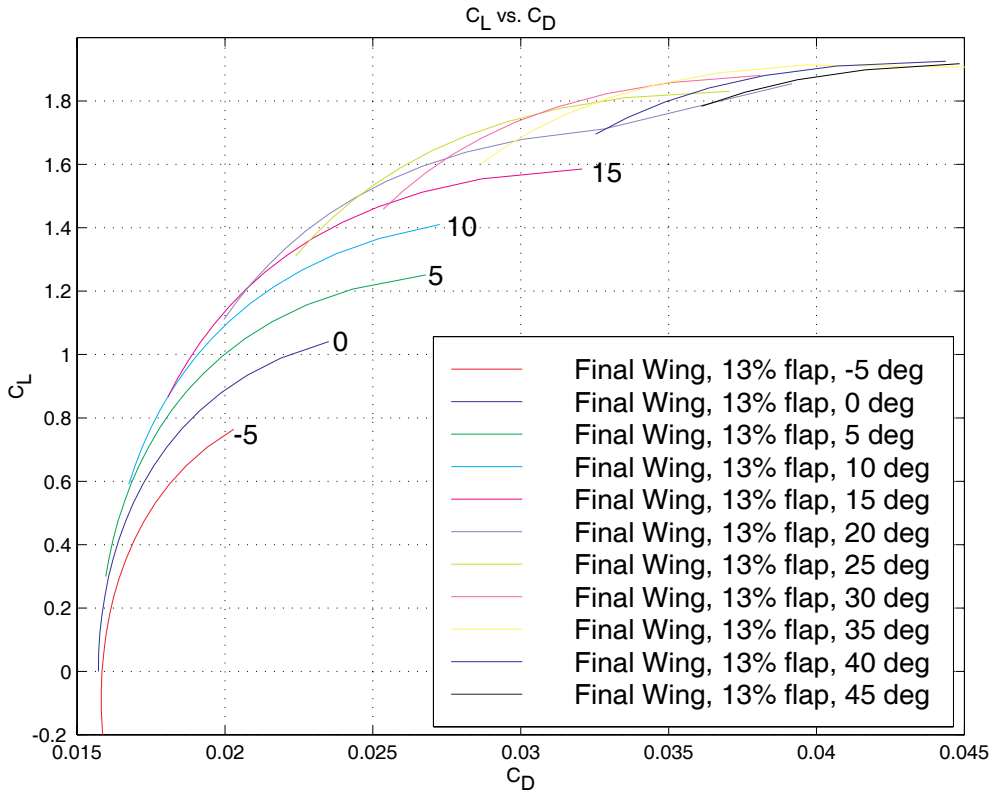

Figure 5-14 Coefficients of lift vs. drag for the final wing section with flap deployed at a Reynolds number of 229,000. For a given C_L, there is a unique flap angle that yields the minimum drag for that lift coefficient forming an efficient boundary This will later be used to control the wing; once the desired lift is set, the flap is tuned for minimum drag based on that lift coefficient.

placement of the rotation axis can keep the forces low. However, the variability of the wind will require high frequency actuation to keep the wing correctly trimmed. Furthermore, the entire range of angles of attack between zero lift and stall is less than 10 degrees. This translates into the actuator requirement to track the wind very closely indeed.

The other possibility for angle of attack control is to use an auxiliary surface to trim the wing aerodynamically. The auxiliary surface can take the form of a tail behind the wing (conventional), a tail in front of the wing (canard), or attached to the trailing edge of the wing (flying wing). The actuator requirement in this case is to move the trimming surface only. By designing the auxiliary surface in such a way as to have the wing/surface

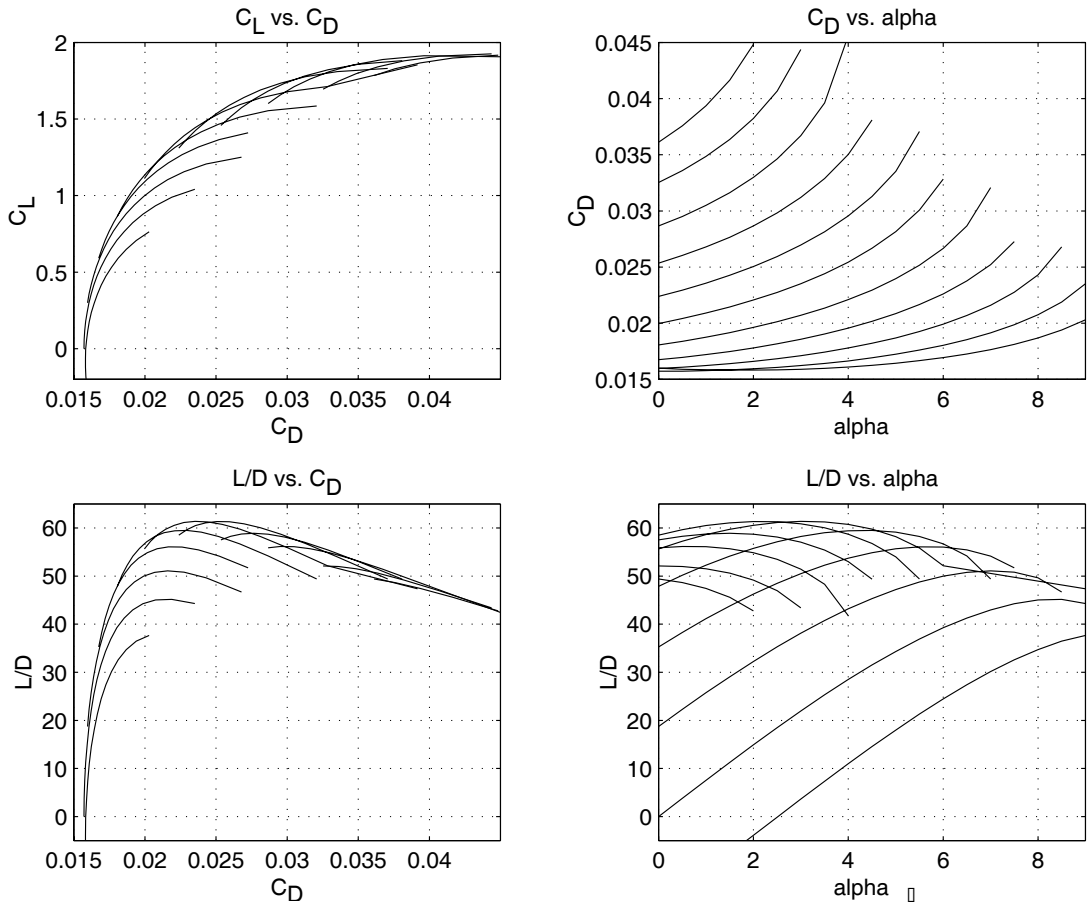

Figure 5-15 Aerodynamic polars for the main wing sail with flap deployed at a Reynolds number of 229,000. The lift/drag ratio peaks at around 62 with a corresponding angle of attack of 2 degrees. This corresponds to a lift coefficient of approximately 1.3 and a flap deflection of 20 degrees. Note that at this flap deflection, the lift/drag ratio remains high over the range of 0 to 4 degrees angle of attack.

assembly be passively stable with respect to angle of attack, the entire system will track the relative wind automatically. This is a great advantage over active control in terms of actuation effort, simplicity of design, and overall performance.

With the main wing section and tail section designed, the various arrangements of wing and tail needed for the Atlantis to sail can now be considered. The first requirement for wing and tail consideration is stability with respect to change in wind direction or velocity. Based on the design of the wing section and the conclusions reached in

Section 5.6.4 on page 172, the configuration must be able to hold a C_L of 1.8 with a flap deflection of 45 degrees. This is important because the deflected flap will cause an increase in the pitching moment of the wing section about the mast. Formally, these requirements can be written as:

$$C_m = 0 \quad \text{(EQ 5.2)}$$

The pitching moment of the entire wing/tail system about the mast should be zero. That is, the system is in trim, and:

$$\frac{d}{d\alpha}(C_m) < 0 \quad \text{(EQ 5.3)}$$

The change in pitching moment with a change in angle of attack should be negative. The wing/tail system should be stable with respect to angle of attack.

The other considerations are mechanical complexity, control power, and a minimum swept radius of the farthest point away from the mast. The minimum swept radius constraint is due to the fact that in order to remove the coupling between angle of attack of the wing sail and heeling angles, the wing/tail assembly must be mass balanced about the mast so that pitch and roll angles do not induce changes in angle of attack. A tail heavy wing/tail assembly would result in an increase in angle of attack with roll angle, thus inducing instability in close hauled conditions.

Fekete et al. in [48] perform a simplified analysis of the conventional and canard configurations. In this section, that analysis is refined by using higher fidelity models for downwash and induced drag as well as correcting for aspect ratio effects. Furthermore, two other configurations, the “flying wing” and the “free-floating canard,” are analyzed using the same tools.

SECTION 5.7.1 METHODOLOGY

In order to keep the problem tractable, simplified lifting line theory, as discussed in Glauert [59], is used for the entire analysis. This provides a simple yet accurate and consistent method of accounting for flap deflection and three dimensional flow effects in the form of an aspect ratio correction. Lifting line theory assumes that the airfoil is flat and imposes a bound vortex to induce circulation and create the corresponding (inviscid) lift. The drag considerations are those caused by induced drag and those associated with changing the momentum of the flow by altering its direction. In the specific case of this configuration analysis, first consider the main wing with flap deployed.

Define:

$$\Theta_f \equiv \text{acos}\left(2\left[\frac{c_f}{c}\right] - 1\right) \tag{EQ 5.4}$$

where $\left[\frac{c_f}{c}\right]$ is the flap/chord ratio, optimized to be 0.13, thus yielding a value for Θ_f of 2.404 radians. The quantity Θ_f is used in:

$$\alpha_\delta \equiv 1 - \frac{\Theta_f - \sin\Theta_f}{\pi} \tag{EQ 5.5}$$

which, yields a value for α_δ of 0.449 and in

$$\mu_1 \equiv \frac{1}{2}\left(1 - \left[\frac{c_f}{c}\right]\right)\frac{\sin\Theta_f}{\pi - \Theta_f + \sin\Theta_f} \tag{EQ 5.6}$$

which yields μ_1 equal to 0.2075. These values are used in the fundamental equations for calculating lift and pitching moment in the following way:

$$C_L = C_{L_\alpha}\alpha + C_{L_\delta}\delta \tag{EQ 5.7}$$

where α is the airfoil angle of attack and δ is the flap deflection. However, from lifting line theory:

$$C_{L_\alpha} = \frac{2\pi AR}{AR + 2} \quad \text{(EQ 5.8)}$$

$$C_{L_\delta} = \alpha_\delta C_{L_\alpha} \quad \text{(EQ 5.9)}$$

where AR is the aspect ratio (defined as the square of the span divided by the area, or the ratio of span to chord for a rectangular wing), and α_δ is defined in Equation 5.5. Note that the pitching moment is defined as:

$$C_{M_\delta} = \mu_\delta C_{L_\alpha} = -\mu_1 \alpha_\delta \left[\frac{c_f}{c}\right] C_{L_\alpha} \quad \text{(EQ 5.10)}$$

where μ_1 is defined in Equation 5.6 and $\left[\frac{c_f}{c}\right]$ is the flap/chord ratio. In summary:

$$\left[\frac{c_f}{c}\right] = 0.13 \quad \text{(EQ 5.11)}$$

$$\Theta_f = 2.404 \quad \text{(EQ 5.12)}$$

$$\alpha_\delta = 0.449 \quad \text{(EQ 5.13)}$$

$$\mu_1 = 0.2075 \quad \text{(EQ 5.14)}$$

$$\mu_\delta = -0.0121 \quad \text{(EQ 5.15)}$$

When combined with Equation 5.7 through Equation 5.10 give the basic relationships required to analyze the various configurations.

SECTION 5.7.2 CONVENTIONAL LAYOUT

The conventional layout is what would occur if a normal airplane were sliced in half down the longitudinal axis, turned on its side, and affixed to the mast through the quarter chord point of the wing. The arrangement, pictured in Figure 5-16 has the wing forward, followed by a tail some distance back. This has the immediate disadvantage of being tail heavy. This requires ballast forward to place the center of mass at the quarter chord point of the main wing. In terms of a wing, ballast is useless weight. Because the

weight must be attached to the wing, it raises the center of gravity of the boat. This makes the design more prone to capsizing. Additionally, the swept radius of the tail is quite far back. This means that in close quarters (such as berthing), the tail may swing out beyond the catamaran hulls and strike an adjacent ship.

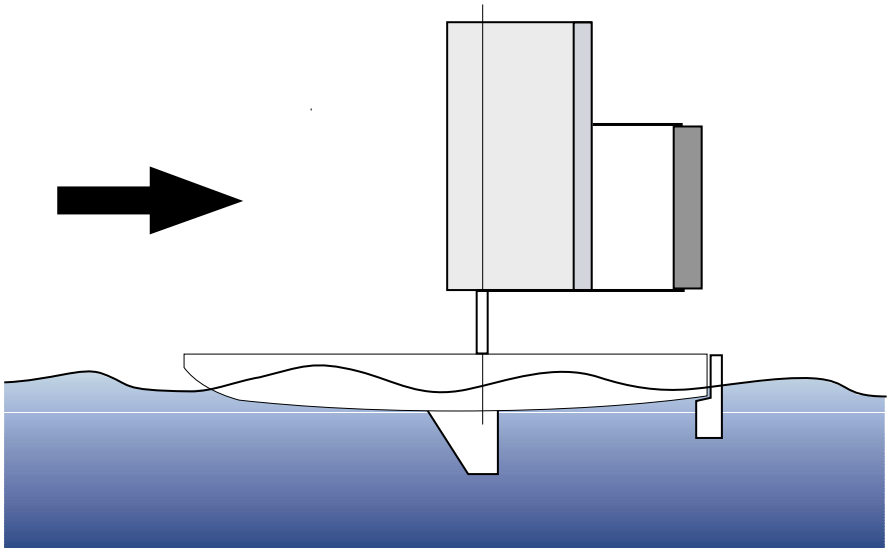

Figure 5-16 Conventional configuration for the wing sail and tail arrangement. This is the equivalent of a conventional airplane sliced in half down the length of the airplane, turned sideways and mounted on the stub-mast. This configuration has the inherent disadvantage that the wing design is tail heavy. This requires ballast to bring the mass center of the wing/tail assembly in line with the stub-mast. Additionally, this configuration has the farthest point of the wing/tail far away from the stub-mast. Thus, it sweeps out a large radius, making it impossible to use external stays on the stub-mast above the wing.

A top down view of Figure 5-16 is presented in Figure 5-17 and shows all of the force and moment vectors acting on the wing sail and tail. Note that in order to balance or trim the wing sail, the moment about the pivot point must be zero. To guarantee passive gust stability, the derivative of the moment equation with respect to the angle of attack must be negative. This implies that a perturbation of the angle of attack in a positive sense will cause a negative (or nose down) pitching moment which will reduce the angle of attack, and likewise a negative angle of attack perturbation will cause an increase in the pitching moment (nose up) and will increase the angle of attack. Also note that the refer-

ence to up and down is simply a convention to relate the wing terminology back into the intuitive reference of flight. In fact, there is no up or down, rather port or starboard, and that the important feature which is missing from these equations which would be present if this were in fact an aircraft are the gravity terms. They do not, however, come into play in this stability analysis.

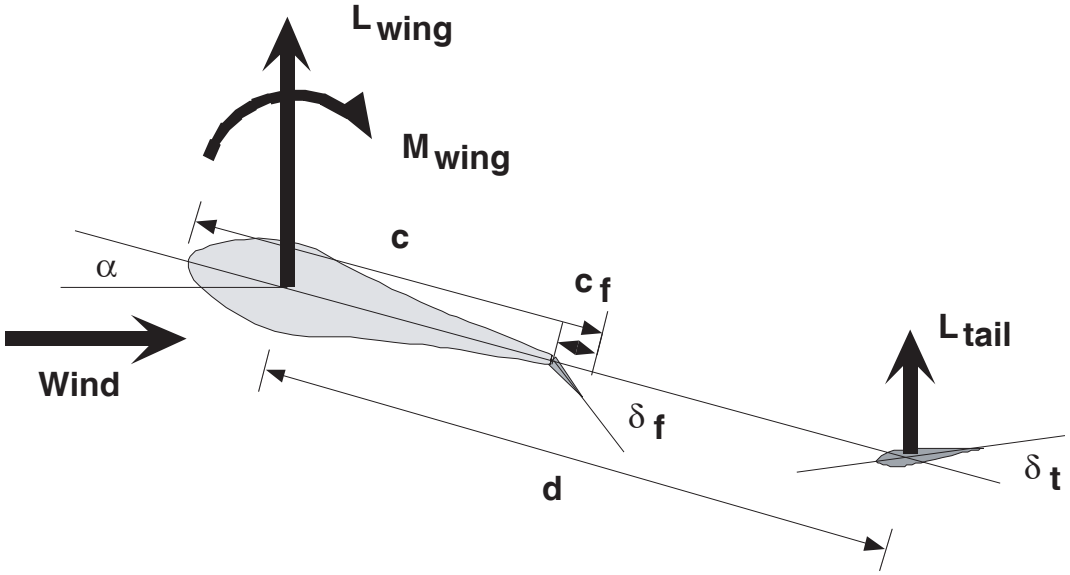

Figure 5-17 Force vectors on the conventional configuration. The forces and moments on the conventional configuration are displayed along with the relevant angles and distances. The wing is "flying" at an angle of attack, α, which in turn generates lift and pitching moment associated mostly with the trailing edge flap. This pitching moment must be resisted by the lift force on the tail.

With Figure 5-17 as a reference, expand the various terms in the equations:

$$L_W = \frac{1}{2}\rho V^2 S_w (C_{L_\alpha})_W (\alpha + \alpha_\delta \delta_f) \qquad \text{(EQ 5.16)}$$

where L_w is the lift of the wing, ρ is the density of the air, V is the free stream wind velocity, S_w is the wing area, α is the angle of attack of the wing, and δ_f is the flap deflection. Likewise:

$$M_W = \frac{1}{2}\rho V^2 S_w c_w (C_{L_\alpha})_W \mu_\delta \delta_f \qquad \text{(EQ 5.17)}$$

where M_w is the pitching moment about the quarter chord point (defined as positive [+] nose up), c_w is the wing chord length, and μ_δ is defined in Equation 5.10. Lastly:

$$L_T = \frac{1}{2}\rho V^2 S_T (C_{L_\alpha})_T \left(\alpha + \delta_T - \frac{\varepsilon}{2}\right) \quad \text{(EQ 5.18)}$$

where L_T is the lift of the tail, S_T is the tail area, δ_T is the tail deflection relative to the main wing chord line, and ε is the downwash angle of the flow off of the main wing.

Using Equation 5.16 through Equation 5.18, it is possible to define the moment balance about the pivot point (located at the quarter chord, or aerodynamic center):

$$\sum M_\Theta = M_w - d \times L_T \quad \text{(EQ 5.19)}$$

which can be expanded to:

$$M_\Theta = \frac{1}{2}\rho V^2 S_w c_w (C_{L_\alpha})_W \mu_\delta \delta_f - \frac{1}{2}\rho V^2 S_T d (C_{L_\alpha})_T \left(\alpha + \delta_T - \frac{\varepsilon}{2}\right) \quad \text{(EQ 5.20)}$$

Rearranging and expanding:

$$C_{M_\Theta} = \frac{M_\Theta}{\frac{1}{2}\rho V^2 S_w c_w} = (C_{L_\alpha})_W \mu_\delta \delta_f - \left(\frac{d}{c_w}\right)\left(\frac{S_T}{S_w}\right)(C_{L_\alpha})_T \left(\left[\frac{AR_w + 1}{AR_w + 2}\right]\alpha + \delta_T - \frac{\alpha_\delta \delta_f}{AR_w + 2}\right) \quad \text{(EQ 5.21)}$$

Let the following definitions be made:

$$\bar{d} \equiv \frac{d}{c_w} \quad \text{(EQ 5.22)}$$

$$\bar{S} \equiv \frac{S_T}{S_w} \quad \text{(EQ 5.23)}$$

Substituting the definitions above and rearranging Equation 5.21:

$$C_{M_\Theta} = \left[(C_{L_\alpha})_W \mu_\delta + \bar{d}\bar{S}(C_{L_\alpha})_T \frac{\alpha_\delta}{AR_w + 2}\right]\delta_f - \bar{d}\bar{S}(C_{L_\alpha})_T \frac{AR_w + 1}{AR_w + 2}(\alpha + \delta_T) \quad \text{(EQ 5.24)}$$

Where the downwash angle is:

$$\frac{\varepsilon}{2} = \frac{\alpha + \alpha_\delta \delta_f}{AR_w + 2} \quad \text{(EQ 5.25)}$$

the wing lift curve slope is:

$$(C_{L_\alpha})_W = 2\pi \frac{AR_w}{AR_w + 2} \quad \text{(EQ 5.26)}$$

and the tail lift curve slope is:

$$(C_{L_\alpha})_T = 2\pi \frac{AR_T}{AR_T + 2} \quad \text{(EQ 5.27)}$$

The stability criteria is that the change in pitching moment coefficient with respect to a change in angle of attack be negative. Thus it follows from Equation 5.24:

$$\frac{\partial C_{M_\Theta}}{\partial \alpha} = -\overline{dS}(C_{L_\alpha})_T \frac{AR_w + 1}{AR_w + 2} \quad \text{(EQ 5.28)}$$

Thus, the stability criteria is met for all $d > 0$ (i.e., the tail must be behind the main wing, and it will be stable). Using the results from XFOIL and the results of Equation 5.11 through Equation 5.15, the trim condition is evaluated with a maximum lift coefficient of 1.8 for the main wing with the flap deflected 45 degrees, and a maximum lift coefficient of 0.75 for the tail. This results in a minimum trim condition, such that the tail is at its maximum lift coefficient at the same instant that the wing reaches its maximum lift coefficient. The minimum trim condition is realized for a "normalized tail volume" of:

$$\overline{dS} > 0.00528\pi^2 \quad \text{(EQ 5.29)}$$

Thus in the case of the conventional wing/tail arrangement, both the stability criteria and the trim condition can be met, as long as the normalized tail volume satisfies Equation 5.29. The swept radius of the conventional configuration is the distance from the quarter chord to the end of the tail, which turns out to be:

$$R_s = c_w\left[\bar{d} + \frac{3}{4}\bar{c}\right] \quad \text{(EQ 5.30)}$$

where $\bar{c}$ is the ratio of the tail chord to the wing chord. Note that while $\bar{c}$ is fairly small, the distance from the mast to the quarter chord of the tail can be quite large, and the swept

radius is too large to place inside the guy wires of the original existing mast. Thus, the conventional arrangement requires a free standing mast whose supports are entirely below the pivot point of the main wing sail.

SECTION 5.7.3 CANARD

An alternate configuration is the “canard,” where the tail is placed in front of the wing as pictured in Figure 5-18. The immediate advantage professed for this arrangement is that both the main wing and the tail are lifting in the same direction and therefore must be more efficient. In an aircraft, this turns out to be untrue. A canard aircraft has trim and stall problems that must be dealt with and can usually be designed for either passive stability, or efficiency (i.e., reduced induced drag), but never both.

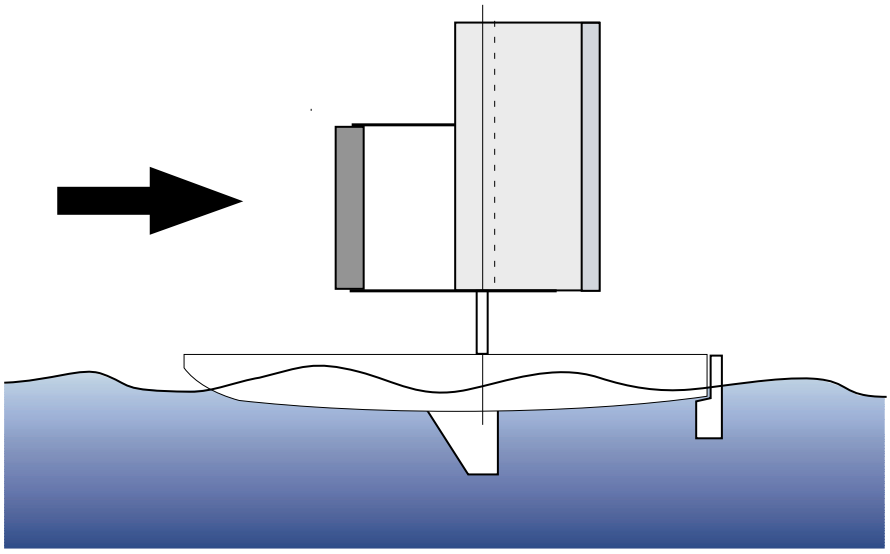

Figure 5-18 Canard configuration of the main wing sail and tail. The canard configuration has the trim surface (or tail) in front of the main wing. The advantage of this is that is can be made to have its mass center coincide with the stub-mass. Also, there exists the possibility that the swept radius can be made such that it is possible to have the entire canard/wing assembly fit inside guy wires that stabilize the mast.

The overwhelming advantage a canard has for the sailboat propulsion system is that it is more easily balanced about its neutral point, making the entire setup lighter. Also,

depending on the distances that occur for trim and stability, it is possible that the radius swept by the canard arrangement can be made small. If the swept radius is small enough to fit within the existing guy wires of the original mast, then the canard can be fit around the existing mast like a sleeve. This would negate the need for a free-standing stub-mast making the structure of the mast much easier to design. Figure 5-19 shows the vectors and key distances on the canard configuration. Once again, it is required that the moment balance be zero (trim) and that the change in moment be negative (stability).

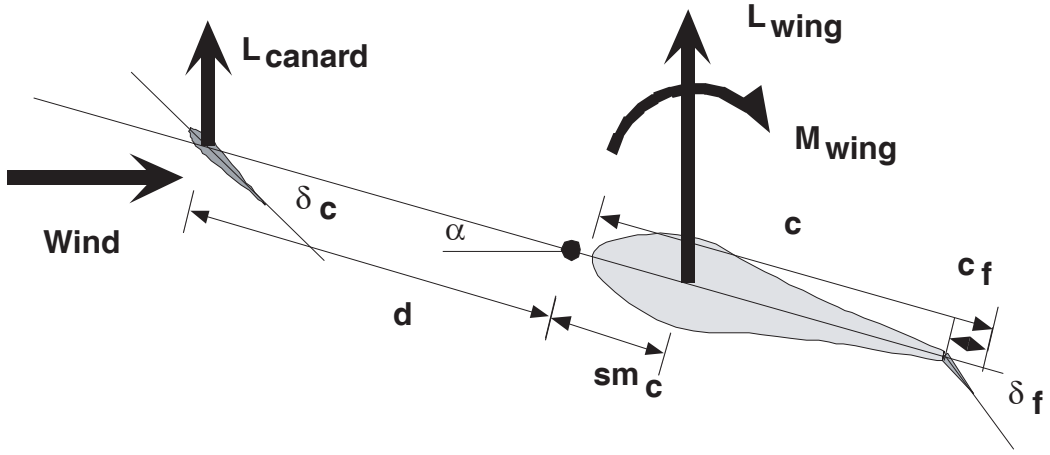

Figure 5-19 Force vectors on the canard configuration. The canard configuration has the purported advantage of efficiency, due to the fact that both the main wing and canard lift in the same direction. It truth, canard configurations can be made to be efficient (low induced drag) or passively stable, but never both. As a wingsail, the canard offers the advantage of already having its mass center near the pivot point. This minimized the ballast required and results in a lighter overall rig.

With the Figure 5-19 as a reference, expand the force and moment terms:

$$L_W = \frac{1}{2}\rho V^2 S_w (C_{L_\alpha})_W (\alpha + \alpha_\delta \delta_f) \qquad \text{(EQ 5.31)}$$

The lift on the wing, L_w, is identical to the conventional arrangement, as in Equation 5.16. The moment on the wing, M_w, is also identical to Equation 5.17

$$M_W = \frac{1}{2}\rho V^2 S_w c_w (C_{L_\alpha})_W \mu_\delta \delta_f \qquad \text{(EQ 5.32)}$$

The lift on the canard, L_c, is however, slightly different from the lift on the tail:

$$L_c = \frac{1}{2}\rho V^2 S_c (C_{L_\alpha})_c (\alpha + \delta_c) \quad \text{(EQ 5.33)}$$

where S_c is the area of the canard, and δ_c is the canard angle relative to the canard/wing chord line. With these basic definitions and the distances as defined in Figure 5-19, the moment balance for stability and trim can be formulated as follows:

$$\sum M_\Theta = M_w - sm \times L_w - d \times L_c \quad \text{(EQ 5.34)}$$

which can be expanded to:

$$M_\Theta = \frac{1}{2}\rho V^2 S_w c_w (C_{L_\alpha})_W \mu_\delta \delta_f - \frac{1}{2}\rho V^2 S_c d (C_{L_\alpha})_c (\alpha + \delta_c) - \frac{1}{2}\rho V^2 S_w sm (C_{L_\alpha})_W (\alpha + \alpha_\delta \delta_f) \quad \text{(EQ 5.35)}$$

and can be rewritten in a coefficient form and non-dimensionalized as:

$$C_{M_\Theta} = \frac{M_\Theta}{\frac{1}{2}\rho V^2 S_w c_w} = (C_{L_\alpha})_W \mu_\delta \delta_f - \left(\frac{d}{c_w}\right)\left(\frac{S_T}{S_w}\right)(C_{L_\alpha})_c (\alpha + \delta_c) - \left(\frac{sm}{c_w}\right)(C_{L_\alpha})_W (\alpha + \alpha_\delta \delta_f) \quad \text{(EQ 5.36)}$$

Rearranging and grouping the first and third terms of the right hand side of the equations:

$$C_{M_\Theta} = (C_{L_\alpha})_W [\mu_\delta + \overline{sm}\alpha_\delta]\delta_f - \{\overline{sm}(C_{L_\alpha})_W + \overline{dS}(C_{L_\alpha})_c\}\alpha - \overline{dS}(C_{L_\alpha})_c \delta_c \quad \text{(EQ 5.37)}$$

where the definitions are same in Equation 5.22 and Equation 5.23, except for:

$$\overline{sm} \equiv \frac{sm}{c_w} \quad \text{(EQ 5.38)}$$

Differentiating Equation 5.37 with respect to α, the stability criteria is:

$$\frac{\partial C_{M_\Theta}}{\partial \alpha} = -\{\overline{sm}(C_{L_\alpha})_W + \overline{dS}(C_{L_\alpha})_c\} < 0 \quad \text{(EQ 5.39)}$$

Note here that both $\overline{sm}$ and $\overline{d}$ are defined as positive rear of the pivot point, so that in fact the $\overline{d}$ in Equation 5.39 is negative, whereas $\overline{sm}$ is positive and $\overline{S}$ is much less than one. At this point, no mention has been made of the correction for the downwash induced on the main wing or the upwash on the canard. The correction is not as straight forward as the downwash for the conventional setup in Equation 5.25. Instead, the downwash is cal-

culated from biplane theory developed in the 1920s. Using Munk's stagger theorem, the downwash of the canard on the wing has roughly the same effect as the wing would have on the canard if the flow were reversed. Thus, the downwash correction shows up in the lift curve slope of the main wing.

$$(C_{L_\alpha})_c = 2\pi \frac{AR_c}{AR_c + 2} \quad \text{(EQ 5.40)}$$

and

$$(C_{L_\alpha})_w = 2\pi \frac{AR_w}{AR_w + 2}\left(1 - \frac{2(C_{L_\alpha})_c \bar{S}}{\pi AR_w}\right) \quad \text{(EQ 5.41)}$$

Thus, in the case of the canard, the trim condition and stability are related. In other words, using the results from the XFOIL code and evaluating for a given coefficient of lift and angle of flap deflection, the canard can be made to trim and be passively stable. The equations are solved simultaneously to yield the parameters (setting δ_f and δ_c determines α) and the required static margin ($\overline{sm}$) and canard or tail volume ($\overline{dS}$).

The problem occurs with the robustness of this solution. The solution to the equations exists for only one flap deflection, δ_f. For instance, if the trim and stability criteria are solved for a flap deflection of 45 degrees and a coefficient of 1.8, then when the flap is deflected less than 45 degrees, the canard configuration is not stable, and will attempt to swap ends. Furthermore, if the trim and stability criteria are solved for zero flap deflection, then at a high lift coefficient the canard will not be able to hold the main wing at a sufficient angle of attack. The canard itself will not be able to generate sufficient lift to balance the nose-down pitching moment induced by the large lift on the main wing.

Unfortunately, the canard configuration must choose to be stable, or to have a high maximum lift coefficient, but cannot do both. Thus the canard configuration is not acceptable for this project. If the additional complexity of a flap on the canard is included, then it

is possible to adjust the flap on the canard in such a way as to compensate for the additional moment generated by the deflected flap on the main wing.

This is exactly the same problem that canard aircraft have. Very few of them have flaps on the main wing for decreased landing speeds. In fact, those few that do have flaps on their main wings look to some rather exotic solutions, such as having a variable sweep canard that changes the longitudinal center of the canard as the flaps are deployed. For an excellent treatment of the subtleties of canard designs, see [85] and [98].

SECTION 5.7.4 FLYING WING

If the desire is to minimize the swept radius of the wing, then certainly the flying wing would represent the optimal approach. Flying wings, however, almost always rely on washout of the tips to provide stability. That is, the tips of the rearward swept wings are twisted nose down and act somewhat like a conventional tail. To make a symmetrical flying wing sail requires both trim and stability without any wing twist whatsoever. This represents a difficult design challenge.

One of the main differences of the flying wing from other configurations is that the wing is tapered, rather than rectangular. Due to this fact, the lift and moment equations will be formulated for an infinitesimal of the wing shown in Figure 5-20. The tip to chord ratio is defined as τ, and the sweep of the quarter chord line as Θ_s. With these definitions:

$$\bar{y} \equiv \frac{y}{b} \qquad \text{(EQ 5.42)}$$

where y is the distance from the base of the wing to the tip, b is the semi-span, and $\bar{y}$ is the normalized coordinate that is zero at the base and one at the tip. From Equation 5.42, it follows from the differentiation that:

$$dy = b d\bar{y} \qquad \text{(EQ 5.43)}$$

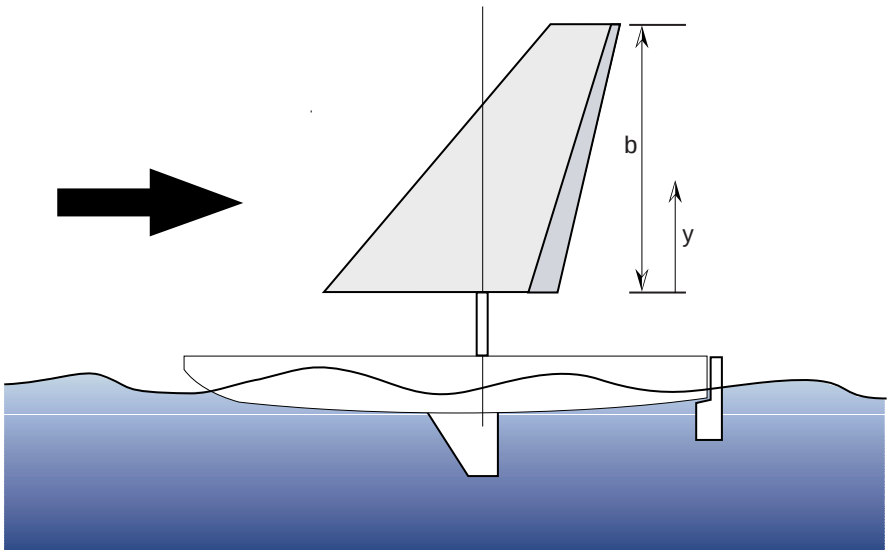

Figure 5-20 Flying wing configuration for wing sail propulsion. The flying wing has many obvious advantages. The flying wing can be made mass balanced with little or no additional ballast. Additionally, it can have the minimum swept radius of any design. The difficult is in achieving both trim and stability with no twist. Flying wings on aircraft rely on wing twist to provide both stability and trim.

The chord as a function of the normalized height is:

$$c(\bar{y}) = \frac{2\bar{c}}{\tau + 1}((\tau - 1)\bar{y} + 1) \qquad \text{(EQ 5.44)}$$

where $\bar{c}$ is the mean chord and τ is the taper ratio of the wing. The location of the quarter chord point of the wing with respect to the pivot, x_{qc}, is:

$$x_{qc}(\bar{y}) = AR\bar{c}\tan\Theta_s\bar{y} + x_0 \qquad \text{(EQ 5.45)}$$

With these basic geometric relationships defined, it is now possible to analyze the moment balance on the flying wing configuration to determine if it is indeed viable. With Figure 5-21 as a reference, the equations for lift and moment of the thin slices of the wing are:

$$L_w = \int_0^1 l_w d\bar{y} = \int_0^1 \frac{1}{2}\rho V^2 c(\bar{y}) C_{L_\alpha}(\alpha + \alpha_\delta \delta_f) b d\bar{y} \qquad \text{(EQ 5.46)}$$

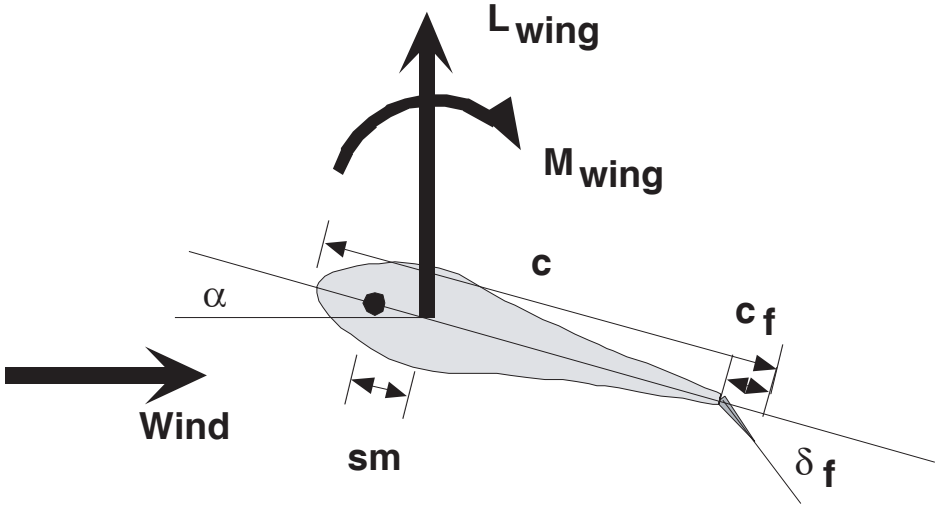

Figure 5-21 Lift and moment vectors for the flying wing configuration. This is for an infinitesimal slice through the horizontal plane. These total values will be the result of integration along the span of the wing. If the flying wing can be made to balance (trim) and be passively stable, then it will have the smallest swept radius of any of the configurations.

and

$$M_w = \int_0^1 m_w d\bar{y} = \int_0^1 \frac{1}{2}\rho V^2 c(\bar{y}) c(\bar{y}) C_{L_\alpha}(\mu_\delta \delta_f) b\, d\bar{y} \qquad \text{(EQ 5.47)}$$

In order for the trim criteria to be satisfied, it is required that:

$$\int_0^1 m_\Theta d\bar{y} = \int_0^1 (m_w - x_{qc}(\bar{y}) l_w) d\bar{y} = 0 \qquad \text{(EQ 5.48)}$$

Substituting Equation 5.43 through Equation 5.47 into the above equation, integrating and evaluating the limits, the moment balance is shown to be:

$$C_{m_\Theta} = \frac{M_\Theta}{\frac{1}{2}\rho V^2 S_w \bar{c}} = C_{L_\alpha}\left\{\left[\frac{4}{3}\frac{\tau^2+\tau+1}{(\tau+1)^2}\right]\mu_\delta \delta_f - \frac{1}{3}\left[AR\frac{2\tau+1}{\tau+1}\tan\Theta_s + 3\bar{x}_0\right](\alpha + \alpha_\delta \delta_f)\right\} \qquad \text{(EQ 5.49)}$$

Differentiating the moment equation with respect to the angle of attack, α, the stability criteria is:

$$\frac{\partial C_{M_\Theta}}{\partial \alpha} = -\left\{ AR\frac{2\tau + 1}{3(\tau + 1)}\tan\Theta_s + \bar{x}_0 \right\} < 0 \qquad \text{(EQ 5.50)}$$

Thus, as long as the sweep angle is positive, and the offset from the mean aerodynamic quarter chord point is less than the first part of the equation above, stability is ensured. The problem occurs when trimming the wing. Due to stability considerations, the second term in Equation 5.49 is negative, thus the first term must be positive in order to trim the wing. Note that the taper ratio, τ, can vary only from 0 to 1, and thus the term $\left[\frac{4}{3}\frac{\tau^2 + \tau + 1}{(\tau + 1)^2}\right]$ will always be positive. In order for the flying wing to have a net zero moment on the wing, the quantity ($\mu_\delta \delta_f$) must be positive, but from Equation 5.15, μ_δ is negative. That is, δ_f must be negative.

Unfortunately, this means that the only way for the flying wing to be both passively stable and have zero net moment is for the flap angle to be negative (or reflexed). Since the high lift coefficient depends on a positive displacement of the flap, the maximum lift coefficient that can be achieved with the flying wing configuration is approximately 0.8 based on XFOIL results. This performance is actually worse than the cloth sail and cannot be accepted.

The stability bound can be found to be a relationship of x_0 and τ for any given Θ_s. As an example of such a relationship, Figure 5-22 is provided. The figure shows the stability bound for a Θ_s range of zero to fifteen degrees. Note that for a given τ a minimum x_0 cannot be exceeded or else the configuration is unstable.

This problem is common in flying wing aircraft as well. If the washout is taken away, then the only way a flying wing can be both stable and trimmed is to reflex the trailing edge of at least a portion of the wing. There is an alternate solution in the case of the

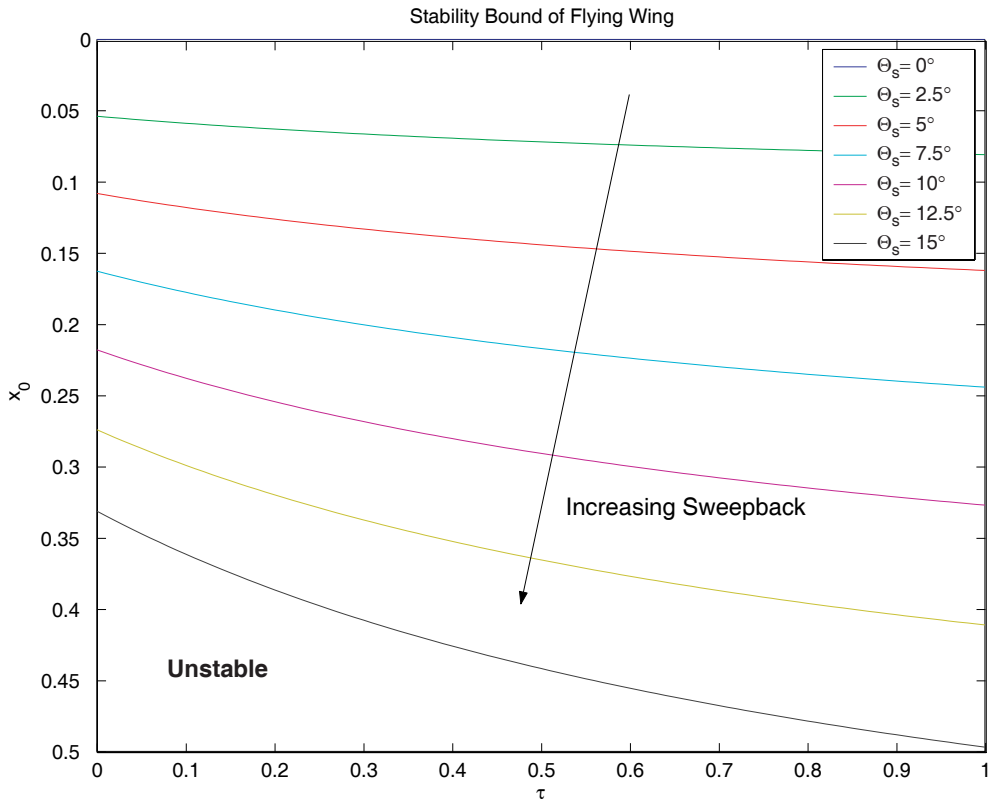

Figure 5-22 Stability bound plot for flying wing configuration with a quarter chord sweep angle, Θ_s, of 15 degrees, showing the relationship between mast pivot offset point, x_0, and taper ratio, τ. For each line, the area below the line is unstable, whereas the area above the line is stable. Both stability and trim can be achieved, but only with a maximum lift coefficient of 0.8. This is approximately the same as a conventional cloth sail. The flying wing configuration, therefore, offers no advantage for its increased weight and complexity.

wing sail which is to use a multi-segment flap and trim part of the flaps in one direction and part in the other. By definition, however, this means dumping lift. Therefore, this cannot be as efficient as any other viable solutions.

SECTION 5.7.5 FREE-FLOATING CANARD

The last configuration to be analyzed is the free-floating canard configuration, pictured in Figure 5-23 below. There are also tri-surface configurations, but these were considered too complex for implementation. The free-floating canard is an unusual

configuration that was first used on the 1942 Curtiss XP55 Ascender, without a great deal of success. Anecdotally, the Ascender suffered from spin recovery problems that caused its test pilots to mispronounce the aircraft's name in a pejorative sense, describing the attitude in which it flew once stalled. Note that after much work, the Ascender's problems were solved, but it remained an army project that never went into production.

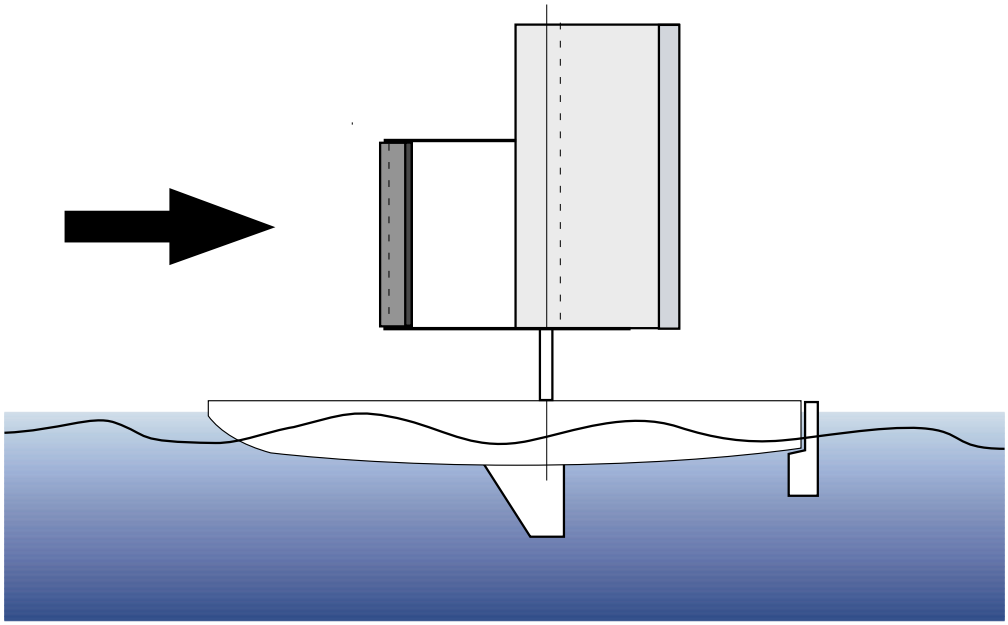

Figure 5-23 The free-floating canard configuration as it would be applied to the Atlantis. The front canard has its own trailing edge flap and is allowed to rotate about a pivot forward of the canard quarter chord line. This, in turn, drags the wing around to a new angle of attack. This system was used on the 1942 Curtiss XP55 Ascender. While it can be made passively stable, the configuration suffers large excursions during stall.

What makes this configuration unique is that the canard itself is allowed to freely pivot in pitch, and is trimmed via a trailing edge flap. Thus, a change in wind direction or a gust causes the canard to rotate into a new trim position, which then, in turn, rotates the wing to its new equilibrium position. The key is that the canard itself is passively stable, and that the entire system retains stability. Forces and moments acting on this system are displayed below in Figure 5-24.

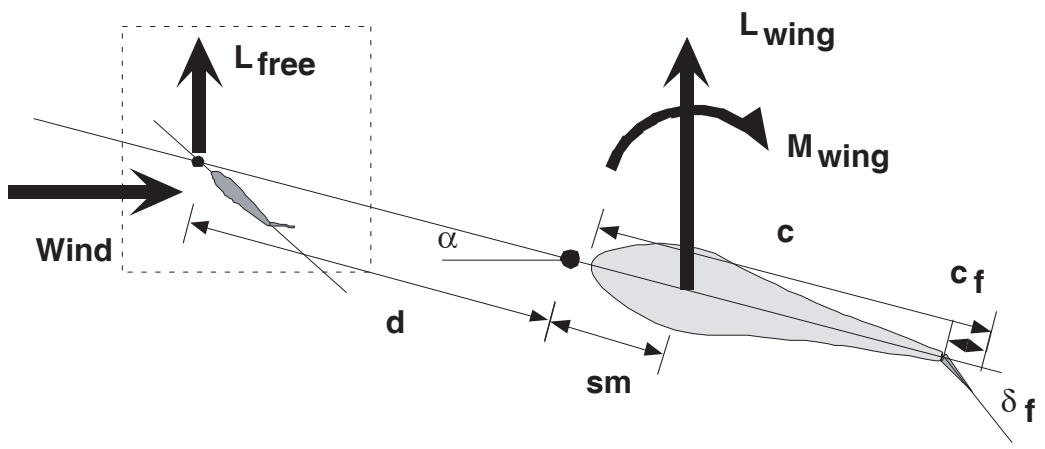

Figure 5-24 Force and moment vectors on the free-floating canard configuration. Since the forward canard is allowed to pivot freely, the force can be taken to act at the pivot point. The forces and moments acting on the forward canard are detailed in Figure 5-25 below. Though this system is complex, it is found to be capable of trim while maintaining passive stability.

The canard itself must be analyzed separately before the analysis of the entire system can be made. Note that the requirements for the canard alone are identical to that of the entire system: the canard be trimmed and stable. Figure 5-25 shows the free-floating canard with its own angle of attack, as well as its own trailing edge flap. Closer inspection of Figure 5-25, along with a comparison to Figure 5-21, will show that the free-floating canard is nothing more that the same flying wing arrangement analyzed in Section 5.7.4 on page 188, with a taper ratio, τ, of one, and a quarter chord sweep angle, Θ_s, of zero. The conclusion of that section was the flying wing configuration could not be stable and trim a high coefficient of lift at the same time. The same is indeed true of the free-floating canard. Based on XFOIL results, the reflexed trailing edge causes the canard to lose 33% of its lifting ability at a Reynolds number of 44,000. Thus the maximum coefficient of lift that the free-floating canard can produce is 0.5, which is much lower than the equivalent conventional tail or canard airfoil.

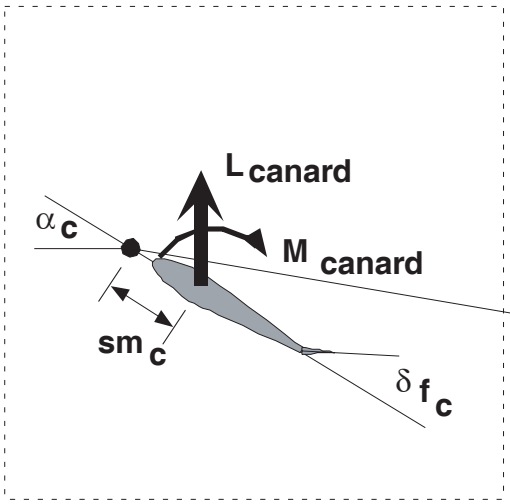

Figure 5-25 Close up view of the forces and moments on the free-floating canard. The free floating canard is essentially the same as the flying wing analysis, except with a taper ratio (τ) of one, and a sweep angle of zero. The conclusion of that section was that stability and trim can only be achieved with a reflexed trailing edge flap. Based on XFOIL results, this reduces the maximum lift coefficient of the free-floating canard to 0.5 versus 0.75 for a conventional tail or canard section.

Restating the requirements, first for the canard:

$$(C_m)_{free} = 0 \tag{EQ 5.51}$$

$$\frac{d}{d\alpha}(C_m)_{free} < 0 \tag{EQ 5.52}$$

With these two equations to guide the analysis, the lift and pitching moment of the free-floating canard can be stated as:

$$L_c = \frac{1}{2}\rho V^2 S_c (C_{L_\alpha})_c (\alpha_c + \alpha_\delta \delta_{f_c}) \tag{EQ 5.53}$$

$$M_c = \frac{1}{2}\rho V^2 S_c c_c (C_{L_\alpha})_c \mu_\delta \delta_{f_c} \tag{EQ 5.54}$$

Formulating the moment balance about the pivot point of the free-floating canard, the so-called free point:

$$\sum (M_\Theta)_{free} = M_c - sm_c \times L_c \tag{EQ 5.55}$$

and setting the results to zero, substituting Equation 5.53 and Equation 5.54,

$$(C_{M_\Theta})_{free} = \frac{(M_\Theta)_{free}}{\frac{1}{2}\rho V^2 S_c c_c} = (C_{L_\alpha})_c[\mu_\delta \delta_{f_c} - \overline{sm}_c(\alpha_c + \alpha_\delta \delta_{f_c})] \quad \text{(EQ 5.56)}$$

The stability criterion, Equation 5.52, is satisfied as long as $sm_c > 0$ (pivot point ahead of the quarter chord). The moment balance, Equation 5.51, results in a fixed relationship between α_c and δ_{fc}:

$$\alpha_c = \left[\frac{\mu_\delta}{\overline{sm}_c} - \alpha_\delta\right]\delta_{f_c} \quad \text{(EQ 5.57)}$$

This relationship is substituted back into Equation 5.53, resulting in a fixed expression for L_{free}:

$$L_{free} = \frac{1}{2}\rho V^2 S_c (C_{L_\alpha})_c \frac{\mu_\delta}{\overline{sm}_c}\delta_{f_c} \quad \text{(EQ 5.58)}$$

The lift generated by the canard is entirely a function of the canard flap angle, δ_{fc}. With this result, the equations relating to the entire free-floating canard and wing system can be analyzed. The trim and stability conditions are found in Equation 5.2 and Equation 5.3. With the forces and moments as shown in Figure 5-24, the entire moment balance of the free-floating canard configuration can be written as:

$$\sum M_\Theta = M_w - sm \times L_w - d \times L_{free} \quad \text{(EQ 5.59)}$$

where L_{free} is defined in Equation 5.58, and L_w and M_w are:

$$L_W = \frac{1}{2}\rho V^2 S_w (C_{L_\alpha})_W(\alpha + \alpha_\delta \delta_f) \quad \text{(EQ 5.60)}$$

$$M_W = \frac{1}{2}\rho V^2 S_w c_w (C_{L_\alpha})_W \mu_\delta \delta_f \quad \text{(EQ 5.61)}$$

Substituting and rearranging:

$$C_{m_\Theta} = \frac{M_\Theta}{\frac{1}{2}\rho V^2 S_w \bar{c}} = (C_{L_\alpha})_W[\mu_\delta - \overline{sm}\alpha_\delta]\delta_f - \{\overline{sm}(C_{L_\alpha})_W\}\alpha - \overline{dS}C_{L_{free}} \quad \text{(EQ 5.62)}$$

where the terms have all been previously defined.

The stability criterion is dependent entirely on the term $\overline{sm}(C_{L_\alpha})_W$. The stability criterion is satisfied for all $\overline{sm} > 0$.

$$\frac{\partial C_{M_\Theta}}{\partial \alpha} = -\overline{sm}(C_{L_\alpha})_W^2 < 0 \quad \text{(EQ 5.63)}$$

The trim condition requires the same corrections for the downwash of the free-floating canard that were used in Section 5.7.3 on page 184:

$$(C_{L_\alpha})_c = 2\pi \frac{AR_c}{AR_c + 2} \quad \text{(EQ 5.64)}$$

$$(C_{L_\alpha})_w = 2\pi \frac{AR_w}{AR_w + 2}\left(1 - \frac{2(C_{L_\alpha})_c \bar{S}}{\pi AR_w}\right) \quad \text{(EQ 5.65)}$$

Using the XFOIL result of a maximum lift coefficient for the free-floating canard with a reflexed trailing edge of 0.5, the set of equations can be solved to yield a stable, trimmed design with a wing lift coefficient of 1.8 at a flap angle of 45 degrees. The resulting normalized tail volume is:

$$\overline{dS} > 0.0107\pi^2 \quad \text{(EQ 5.66)}$$

and the swept radius is:

$$R_s = c_w \times min\left(\left(\bar{d} + \frac{1}{4}\bar{c}\right), \left(\overline{sm} + \frac{3}{4}\right)\right) \quad \text{(EQ 5.67)}$$

The reason for the two terms is due to the possibility that the canard sticks farther forward than the wing extends back.

There remains a quandary about stall into which this simplified analysis did not delve, but which requires some discussion. At stall, the main wing will lose lift, however the free-floating canard will not. The main wing will also gain in nose down pitching moment at stall, due to the separated flow off the back of the main wing section. The loss

of lift on the main wing tends to pitch the nose upwards, increasing the stall. The increase in pitching moment tends to pitch the nose down, decreasing the stall. The canard itself contributes nothing but its steady lift, which tends to increase the stall. Thus, at stall, the main wing must increase its pitching moment faster than the moment generated by the loss of lift and the moment arm to the center of mass. The original design of the Ascender did not achieve this balance and suffered from a divergent stall that pitched the aircraft over onto its back. Eventually, the designers realized the problem and fixed the aircraft. The only fix was in the main wing, as the canard played no role in the transient.

SECTION 5.7.6 RESULTS

After analyzing the four different possible configurations for trim and stability, it can be seen that only two configurations are viable: the conventional tail and the free-floating canard. The normal canard cannot be both stable and trimmed with a trailing edge flap, and the flying wing needs to reflex the trailing edge for stability, thus reducing the attainable maximum lift coefficient.

Of the two configurations that are viable, the relevant equations are repeated here:

$$\overline{dS} > 0.00528\pi^2\Big|_{conventional} \qquad \text{(EQ 5.68)}$$

$$\overline{dS} > 0.0107\pi^2\Big|_{free-floating} \qquad \text{(EQ 5.69)}$$

The tail volume requirement for the free-floating canard is almost double that of the conventional configuration. In addition, the free-floating canard is mechanically much more complicated, with another pivot point and an additional flap on the trimming surface. The main benefit of the free-floating canard is a much reduced swept radius, which might allow the wing to be built inside the original mast's guy wires.

Upon completing the calculations, however, the required tail volume for the free floating canard is such that it, too, will require a free-standing stub-mast, without guy wires. Thus, the only advantage of the free-floating canard is effectively cancelled, and the Atlantis was fitted with a wingsail and conventional tail layout, as shown in Figure 5-16.

SECTION 5.8 STRUCTURAL ANALYSIS

In the analysis of the Atlantis wing sail propulsion system, it was shown that conventional airfoil sections performed poorly at low Reynolds numbers. Therefore, unique sections based on the Reynolds number and other requirements were designed. Following the development of the sections for this application, the configurations were analyzed, and the ideal configuration was chosen. Now it is possible to design the structural elements for the wing and stub-mast itself, validate the design criteria, and build the propulsion system.

In the sections that follow, the methodology for performing the structural analysis will be discussed along with the load design for the spar, the stub mast, and the final structural design.

SECTION 5.8.1 MAXIMUM LOAD DEFINITION

As in the configuration analysis, a simplified model is used for the loading and simple engineering beam theory is used for the structural analysis. The main wing section was shown to be capable of a maximum lift coefficient of 1.8. With this information, and some geometrical parameters, it is possible to formulate the maximum loading of the wing on the catamaran. The maximum flying load will occur when the catamaran is placed on the water and the wind is sufficiently strong to cause an overturning moment about the farthest corner away from the mast center (capsize). However, this will not be the maximum

loading; the maximum load will most likely be the impact load when the capsized wing tip hits the water and will be the result of the dynamic deceleration of the wing along with the momentum of the hulls.

This condition, however, is not amenable to strict analysis, as too many assumptions along the lines of tip over velocity and hydrodynamic deceleration would be required. Rather than attempt a flawed analysis, the maximum *flying* load will be discussed. That is, the load of the wing while interacting with the air comprises the structural analysis of the following sections. Some effort will be made to include safety factors where possible, such that any underestimation of the true maximum load will not result in inadequate structural robustness.

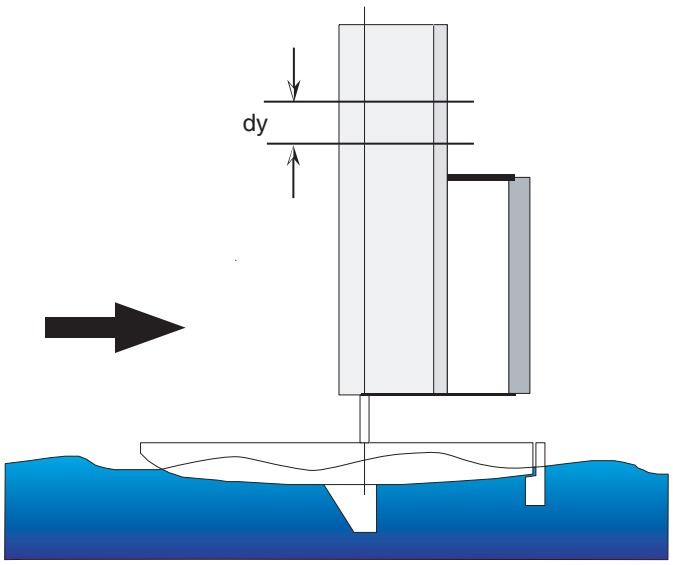

Figure 5-26 Basic configuration of the Atlantis catamaran with conventional wing sail layout. The wind is assumed to be constant along the span of the wing. Strictly speaking, this is a poor assumption, but the error incurred in this analysis is minor. The forces are integrated along the span to find the maximum load that would result in a capsize.

The basic configuration is pictured in Figure 5-26, and the appropriate physical and geometric parameters are given along with their units of measure:

$$b = 5.37 \quad [m] \tag{EQ 5.70}$$

$$c = 1.425 \quad [m] \tag{EQ 5.71}$$

$$S = 7.65 \quad [m^2] \tag{EQ 5.72}$$

$$m = 324 \quad [kg] \tag{EQ 5.73}$$

$$AR = b/c = 3.765 \tag{EQ 5.74}$$

$$C_{L_{max}} = 1.8 \tag{EQ 5.75}$$

$$\rho_{air} = 1.225 \quad \left[\frac{kg}{m^3}\right] \tag{EQ 5.76}$$

$$g = 9.81 \quad \left[\frac{m}{s^2}\right] \tag{EQ 5.77}$$

where b is the semi-span from the bottom of the wing to the top in meters, c is the wing chord in meters, S is the wing area in square meters, and AR is the aspect ratio. The maximum lift coefficient is designated C_{Lmax}, and the parameters for the density of air, ρ, in kilograms per cubic meter and the acceleration of gravity, g, in meters per second squared are included.

Using these parameters, along with the analysis developed in Section 5.7.2 on page 179, the lift (in Newtons, N) for the incremental slice, dy, along the wing can be written as:

$$l = \frac{1}{2}\rho V^2 c C_{l_\alpha}(\alpha + \alpha_\delta \delta_f) = \frac{1}{2}\rho V^2 c C_l \quad [N/m] \tag{EQ 5.78}$$

$$C_l \equiv C_{l_\alpha}(\alpha + \alpha_\delta \delta_f) \tag{EQ 5.79}$$

Thus, integrating the incremental lift along the wing will yield total lift produced by the wing. This analysis assumes no wind gradient for simplicity. The wind gradient is the effect of Earth's boundary layer on surface winds; wind above the surface is faster than

wind at the surface. The same analysis was performed using a simple trapezoidal model for the wind gradient, and the conclusions remain the same. Thus:

$$L = \int_{h}^{h+b} l\,dy = \int_{h}^{h+b} \frac{1}{2}\rho V^2 c C_l\,dy = \frac{1}{2}\rho V^2 c C_l \int_{h}^{h+b} dy = \frac{1}{2}\rho V^2 c C_l y\Big|_{h}^{h+b} = lb \qquad [N] \qquad \text{(EQ 5.80)}$$

Assuming the constant lift load, l, the forces on the catamaran can be summarized in Figure 5-27, where the distances and forces are clearly labeled. From this drawing, it can be seen that there must be a moment balance between the weight of the catamaran and the buoyancy force exerted by the water on the hulls in order to keep from capsizing (in boat design the concept of metacentric height, the virtual height of the center of buoyancy above the center of mass, is frequently used). Furthermore, for the catamaran to not simply accelerate sideways through the water, the centerboards and hulls must generate sufficient side-force to withstand the force of the wing. By assuming a simplified loading, pictured in Figure 5-28, and assuming no acceleration, all forces and moments in each of the directions must be balanced. A very simple load balance ultimately relates the maximum load carried to the moment arm of the hull, d. The equations are simple, and they appear on the following page. This analysis is the same used for a conventional sailboat. The advantage that a catamaran has over monohull designs is the greater moment arm, d, between the center of mass and the buoyancy force, F_{buoy}.

$$\sum F_x = 0 \rightarrow lb = L = R_b \qquad \text{(EQ 5.81)}$$

$$\sum F_y = 0 \rightarrow mg = F_{buoy} \qquad \text{(EQ 5.82)}$$

$$\sum M_{\Theta} = 0 \rightarrow M_b = lb\left[h + \frac{b}{2}\right] \qquad \text{(EQ 5.83)}$$

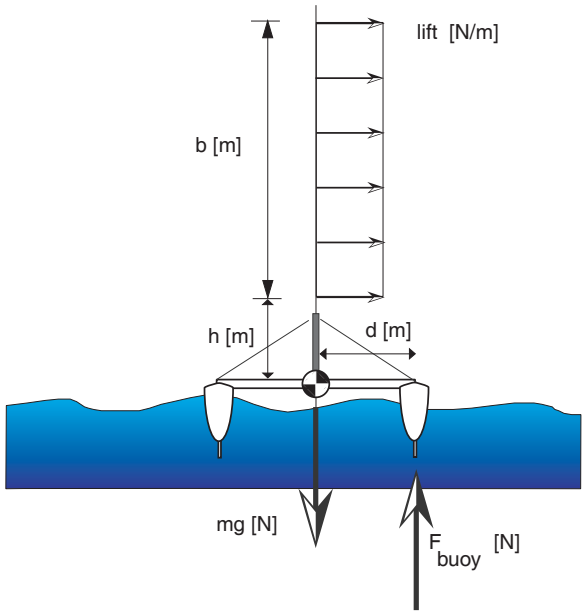

Figure 5-27 Forces and distances for the simplified wing loading as viewed from the bow. The overturning moment that is produced by the net force of the lift acting at the mid-span height of the wing is opposed by the couple between the weight of the vessel and the buoyancy force. The advantage that a catamaran has over a conventional monohull is that the moment arm, *d*, to the buoyancy force is much larger on a catamaran.

where the parameters are all defined in Figure 5-27 and Figure 5-28. Using the above relations, the moment balance can be achieved only by having the buoyancy moment equal and opposite to the overturning moment generated by the lift of the wing:

$$mgd = M_b = lb\left[h + \frac{b}{2}\right] = \frac{1}{2}\rho V^2 S C_l\left[h + \frac{b}{2}\right] \qquad \text{(EQ 5.84)}$$

Where the parameters are previously defined.

With this result, it is possible to set the worst case scenario by maximizing the parameter, *d*, the hull moment arm. In order to overestimate the load, the longest conceivable distance on the hull was used, as pictured in Figure 5-29, where d_{max} is 3.20 meters.

$$d_{max} = \sqrt{\left(\frac{length}{2}\right)^2 + \left(\frac{width}{2}\right)^2} = \sqrt{\left(\frac{5.85}{2}\right)^2 + \left(\frac{2.69}{2}\right)^2} = 3.20 \quad [m] \qquad \text{(EQ 5.85)}$$

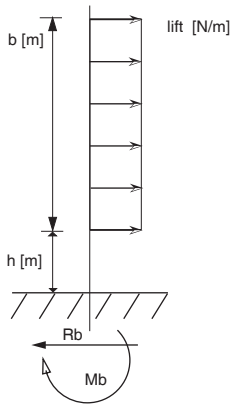

Figure 5-28 Simplified loading diagram from the forces and distances pictured from the bow. The mast is assumed to be cantilevered from a rigid set of hulls. The forces and moments that act on the hulls can be found from simple force and moment balance.

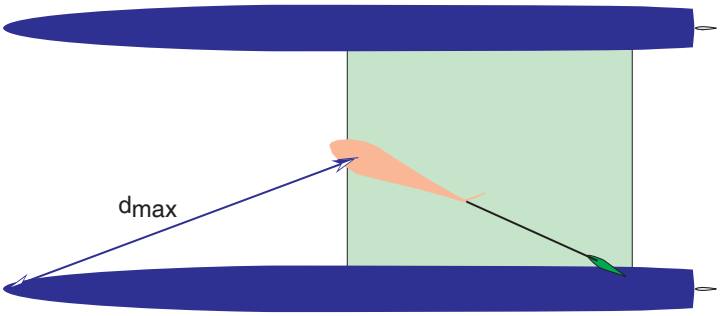

Figure 5-29 Top view of hulls demonstrating the longest possible moment arm, d_{max}. The longest possible distance is used as an artifice to overestimate the maximum loading on the wing and stub-mast. This, in effect, embeds a safety factor into the structural calculations.

Before it is possible to formulate the loading, a height above the deck, h, must be chosen. As the mast will be stayed below this height, a classic trade-off results. Structurally, the most efficient height is for the stay to be at 45 degrees. However, that would cause the overall mast height to be too tall. That is, the lower the bottom of the wing, h, the larger the stay tension. The higher h, the higher the overall overturning loads generated by the wing. This limits the maximum thrust that can be withstood before capsize (there exists an aerodynamic argument for reducing h to zero in order to effectively double the aspect ratio of the wing, but this would be impractical to build). This can be seen from Figure 5-30 denoting the mast as a pin joint, along with the stays.

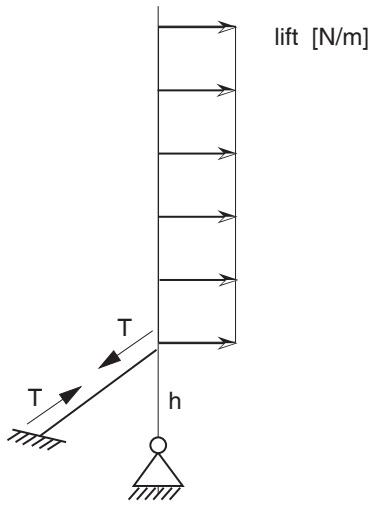

Figure 5-30 Mast as a pin joint for generating the height of the bottom of the wing, *h*. Increasing the mast height lowers the stay tension, *T*, for angles up to 45 degrees from horizontal. Increasing the stay height also raises the height at which the lift acts. This has the direct effect of decreasing the maximum lift that can be obtained before capsizing.

Using simple static equations, and the relations generated earlier, the operative equations become:

$$\sum F_x = 0 \rightarrow R_b = lb - T\frac{d}{\sqrt{h^2 + d^2}} \quad \text{(EQ 5.86)}$$

$$\sum F_y = 0 \rightarrow R_y = T\frac{h}{\sqrt{h^2 + d^2}} \quad \text{(EQ 5.87)}$$

$$\sum M_{\oplus} = 0 \rightarrow \left[\frac{T}{lb}\right] = \frac{\left[h + \frac{b}{2}\right]\sqrt{h^2 + d^2}}{hd} \quad \text{(EQ 5.88)}$$

Note that the distance, *d*, in Equation 5.86 through Equation 5.88 is not the same as d_{max} (from the maximum load scenario), but rather based on hard points where the stays can be attached to the hulls. Thus, the quantity of the tension normalized by total lift is plotted for the two possible values of *d* in Figure 5-31.

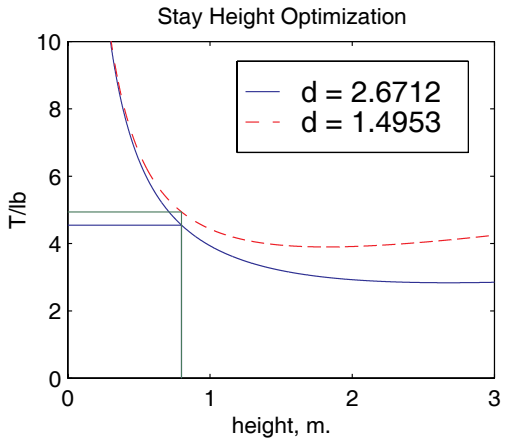

Figure 5-31 Plot of the stay tension normalized by total lift vs. lower wing height. The two values of d are from the geometry to the hard points on the hulls. The stays can be attached only in places where the fiberglass hulls have been reinforced with stainless steel. Note that the shorter stay distance has a minimum tension at approximately 1.85 meters. This is considered too tall for practical implementation.

For both possible stay positions, the height is chosen to be 0.8 meters which is as small as possible without incurring a substantial overturning load increase:

$$h = 0.8 \qquad [m] \tag{EQ 5.89}$$

With Equation 5.89, the maximum loading can be defined. Recall that what is needed is the loading, l, in newtons/meter to define the structural loading. From Equation 5.84, substituting in the values contained in Equation 5.70 to Equation 5.77, and rearranging:

$$mgd = \frac{1}{2}\rho V^2 S C_l \left[h + \frac{b}{2}\right] \rightarrow V^2 C_l = \frac{mgd}{\frac{1}{2}\rho S\left[h + \frac{b}{2}\right]} \tag{EQ 5.90}$$

$$V^2 C_l = 622.86 \qquad \left[\frac{m^2}{s^2}\right] \tag{EQ 5.91}$$

Since the maximum lift coefficient of the wing is 1.8, it is possible to determine the velocity at the maximum lift coefficient such that the maximum load scenario is obtained:

$$C_{L_{max}} = 1.8 \rightarrow V_{max}\Big|_{C_{L_{max}}} = \sqrt{\frac{mgd}{\frac{1}{2}\rho S\left[h + \frac{b}{2}\right]C_{L_{max}}}} \quad \text{(EQ 5.92)}$$

$$V_{max}\Big|_{C_{L_{max}}} = 18.602 \qquad \left[\frac{m}{s}\right] \quad \text{(EQ 5.93)}$$

In other words, setting the wing at maximum lift in a wind speed of 18.6 meters/second or stronger will result in a capsize, no matter which way the wing is oriented. Using this result, Equation 5.93, and:

$$L = lb = \frac{1}{2}\rho V^2 S C_l \quad \text{(EQ 5.94)}$$

yields the value for the total lift, L, and the lift per unit span, l:

$$L_{max} = 2918.51 \qquad [N] \quad \text{(EQ 5.95)}$$

$$l_{max} = \frac{L_{max}}{b} = 543.48 \qquad \left[\frac{N}{m}\right] \quad \text{(EQ 5.96)}$$

These values define the structural loading of the wing and stub mast. Along with an appropriate geometry constraints, both the wing spar and the stub-mast/stays can be designed. The next section will investigate this in more detail.

SECTION 5.8.2 SPAR LOADING

The spar of the wing carries the entire bending moment of the wing. It is secured to the stub mast (which is in turn secured to the catamaran crossbeam) via two bearings which allow the entire wing/tail assembly to rotate freely about the z-axis. There are several design criteria which need be addressed with the structure of the spar. The first is the vertical separation of the bearings. A spherical roller thrust bearing is on top, carrying the entire weight of the wing sail. The lower bearing is a needle roller providing resistance to overturning and holding the wing aligned to the stub-mast. The radial loads of the bearings

decrease as their vertical separation increases. However, since the bottom bearing is fixed at a height h (0.8 meters) above the crossbeam, and the top bearing will be at the top of the stub-mast, the vertical separation of the bearings directly determines stub-mast height.

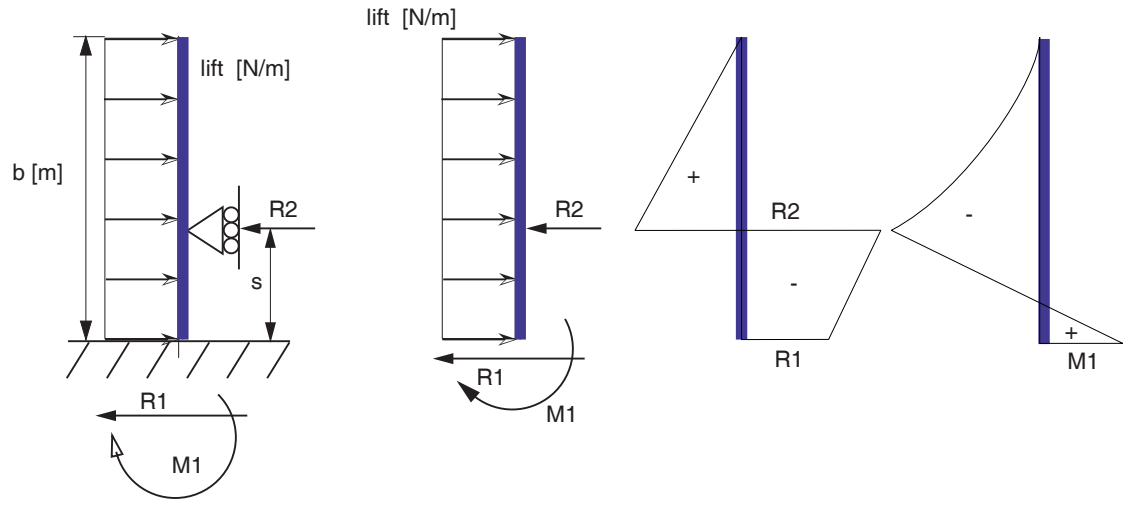

Figure 5-32 Loading of the wing sail spar, with force, sheer, and moment diagrams. The simplified diagram is the representation of the horizontal loads on the stub-mast bearings. The lower bearings is assumed to be able to resist the torque, M_1, and the upper bearing acts as a simple pin joint. Based on these definitions, the loads, sheer, and moment on the wing spar can be computed.

The stub-mast should be as short as possible to allow easy transport. The Atlantis hulls are carried on a trailer to the harbor where the wing sail is assembled in place. The bottoms of the hulls rest upon a trailer that is 0.3 meters high, and the bottom of the stub-mast is more than a meter above the road. Thus, the lower bearing is located close to 2 meters above the road surface. If the vertical separation of the bearings is large, the stub mast will stand too high above the road surface, and it will be difficult to negotiate transportation beneath power lines and overpasses.

The second design criteria is the cross sectional area of the spar caps which will take the bending moment. The spar is a simple “C” construction of spruce caps and marine plywood spar caps. Thus the design of the spar for minimal deflection and safety margin is paramount.

The engineering beam equivalent of the spar is seen in Figure 5-32, where s is the vertical separation of the bearings, and b is the wing semi-span. Using simplified engineering beam theory and Macaulay's method (detailed in [18] and [20]) of solving the successive integration to find the displacement of the beam:

$$M = -\frac{lx^2}{2} + R_2\langle x-\gamma\rangle \quad \text{(EQ 5.97)}$$

$$EIw'' = -M = \frac{lx^2}{2} - R_2\langle x-\gamma\rangle \quad \text{(EQ 5.98)}$$

Integrating Equation 5.98 yields:

$$EIw' = \frac{lx^3}{6} - R_2\frac{\langle x-\gamma\rangle^2}{2} + c_1 \quad \text{(EQ 5.99)}$$

$$EIw = \frac{lx^4}{24} - R_2\frac{\langle x-\gamma\rangle^3}{6} + c_1x + c_2 \quad \text{(EQ 5.100)}$$

The following boundary conditions, defined by the loads in Figure 5-32:

$$w(b) = 0 \quad \text{(EQ 5.101)}$$

$$w'(b) = 0 \quad \text{(EQ 5.102)}$$

$$w(b-s) = 0 \quad \text{(EQ 5.103)}$$

which state that the rigid mast model experiences no deflection or slope at the rigid section (note that the integration proceeds from the free end), and that there is no deflection, but a possible slope at the location of the pin support which models the upper bearing. Applying the boundary conditions yields:

$$R_2 = \frac{l}{8s}[6b^2 - 4bs + s^2] \quad \text{(EQ 5.104)}$$

$$R_1 = lb - R_2 = \frac{l}{8s}[6b^2 + 4bs + s^2] \quad \text{(EQ 5.105)}$$

$$M_1 = -\frac{lb^2}{2} + R_2s = \frac{l}{8}[2b^2 + 4bs + s^2] \quad \text{(EQ 5.106)}$$

Thus the forces and moments on the rigid equivalent wing spar are now completely known. All the factors except the bearing separation distance, s, are known, and thus a

trade-off between total stub mast height, (h+s), and bearing loads can be made. This trade-off study is displayed in Figure 5-33.

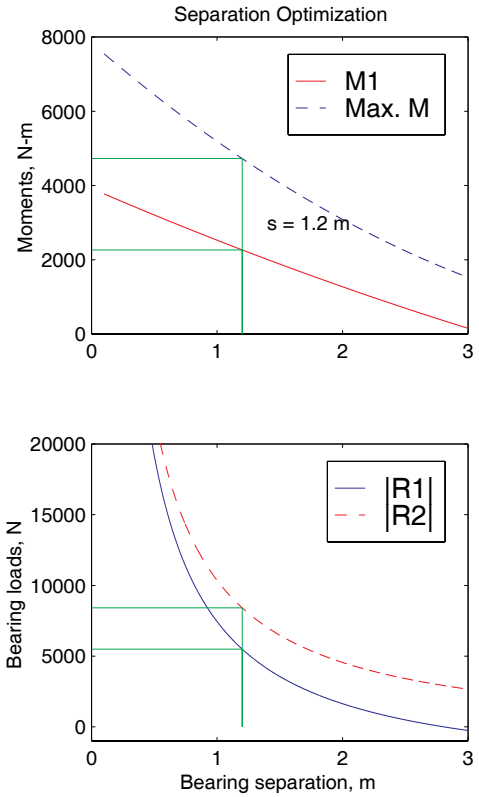

Figure 5-33 Trade-off study for bearing vertical separation distance, s, and bearing loads. The total mast height above the center crossbeam is the sum of the stay height, h, and the bearing vertical separation distance, s. Note that the bending moments are decreasing monotonically up to the distance plotted. The constraint of transporting the Atlantis by trailer drives the height choice.

$$s = 1.2 \qquad [m] \tag{EQ 5.107}$$

Again, the desire is to minimize s while keeping the loads and moments reasonable. From Figure 5-33, s is chosen as 1.2 meters, which brings the total stub-mast height to 2 meters above the crossbeam of the catamaran. Using this bearing separation height, the loading on the rigid equivalent spar is defined for the worst case scenario and displayed in Figure 5-34. The material chosen for the spar is spruce, as wood is light and

inexpensive. With a wing section thickness to chord ratio of 21%, a wooden spar provides adequate strength and rigidity.

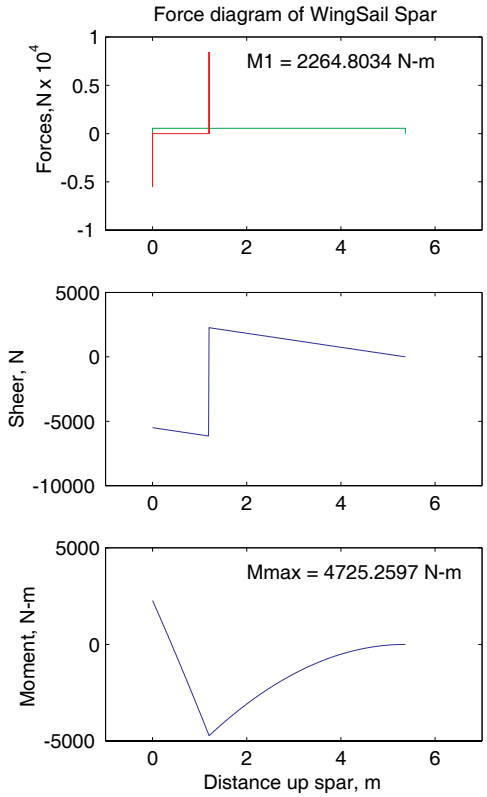

Figure 5-34 The loads, moment and sheer diagrams of the wing sail spar with a bearing separation distance of 1.2 meters and the wing at the maximum loading scenario. The wing spar will fail in bending before it fails in sheer. Thus, the maximum bending moment will drive the choice of materials and design of the wing spar. A wooden spar of sufficient strength is easy to design due to the large internal spacing of the spar caps in a 21% thick airfoil section.

A simple spar that is capable of withstanding the forces and more specifically the maximum moment of 4725.26 Newton-meters is required. A simple model of just the spar caps is used to study the forces and displacements. Spar caps are standard 2 x 8 spruce

planed down to 1 inch thickness, which easily fit within the wing airfoil section. The equation for moment of inertia is:

$$I = \left\{ \frac{bh^3}{12} + hbd^2 \right\} \times 2 \qquad \text{(EQ 5.108)}$$

where I is the moment of inertia, b is the width of the spar cap (~0.15 meters), h is the height of the spar cap (~0.025 meters), and d is the displacement between the two spar caps (~0.13 meters). Substituting the numbers, this yields:

$$I = 1.8534 \times 10^{4} \qquad [m^4] \qquad \text{(EQ 5.109)}$$

Given the ultimate tensile strength of spruce, the maximum moment that this spar should be able to withstand is 57,000 Newton-meters, thus leaving a safety factor of over 12.

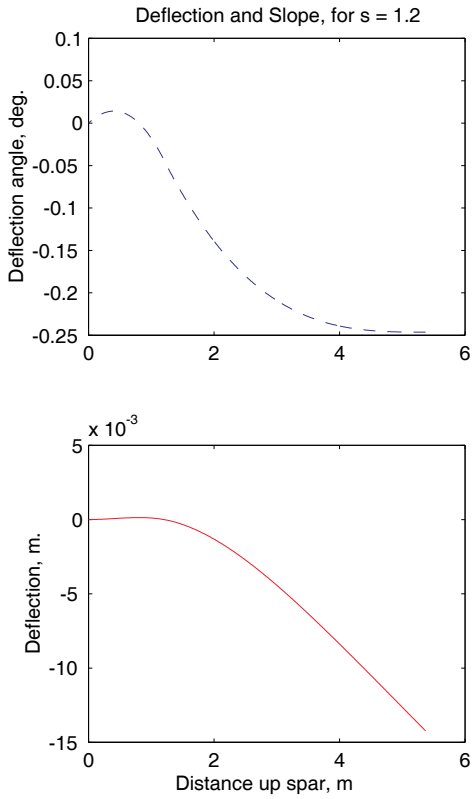

Figure 5-35 Displacement of the wing sail spar for a bearing separation height of 1.2 meters, and a simple spar of two caps of 2.5 x 14 centimeters.

Figure 5-35 shows the displacement of the wing sail spar under the loads and spar configuration previously discussed. As can be seen from the graphs, the spar is extremely rigid, with a displacement of less than 2 centimeters at maximum load, and an angular deflection of less that 1 degree at the tip. In conclusion, the wing sail spar has been found to be feasible, easy to construct using simple materials, and having a very large safety margin.

SECTION 5.8.3 STUB-MAST LOADING

With the structure of the spar determined and analyzed, the stub mast can now be analyzed in full. The stub mast is secured to the cross-beam, using the same ball and socket joint found on the original mast. It is held in place in the vertical plane by four stringers that make up a "spider." Originally, the stub mast was going to be welded to the crossbeam, thus creating a full cantilever joint. Even though this did not happen, the analysis was performed assuming a full cantilever at the intersection of the stub mast and crossbeam. The full engineering model of the stub mast along with the load, sheer, and moment diagrams is shown below in Figure 5-36.

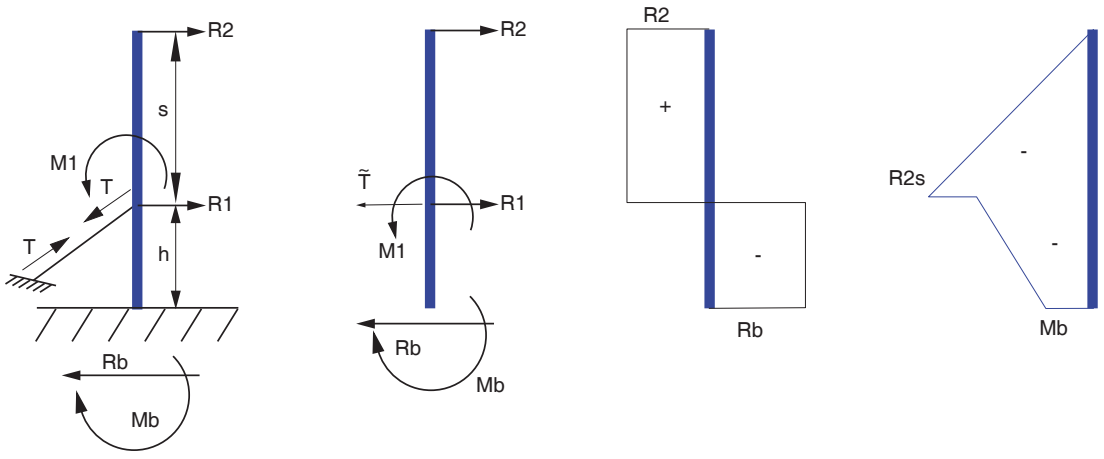

Figure 5-36 Engineering representation of forces and moments on the stub mast. This is a statically indeterminate structure. In order to solve for the forces, the strain of the guy wires are equated to the mast deflection at the point where they join. From this the total force, sheer, and moment diagrams are presented.

Given the nature of this model, the structure is indeterminate. It requires that the boundary conditions for the strain at the juncture between the lower guy wires and the mast deflection be equal. The forces and moments carried by the bearing have been previously defined (Figure 5-36) to be R_1, R_2, and M_1. It is possible to write the equations of equilibrium for the rigid mast structure. Once again, start with the basic loading:

$$\sum F_x = 0 \rightarrow R_b = lb - \tilde{T} \tag{EQ 5.110}$$

$$\sum M_{\Theta} \rightarrow M_b = -\frac{lb^2}{2} - lbh + h\tilde{T} \tag{EQ 5.111}$$

where the quantity, $\tilde{T}$, is the perpendicular tension component of the cable stay. Assuming that the stay makes an angle of Θ with the perpendicular to the stub-mast, then:

$$\tilde{T} = T\cos\Theta \tag{EQ 5.112}$$

Using these relations, along with the results of the previous section and Macaulay's method, the displacement equations are written (integrating from the free end):

$$EIw'' = -M = R_2 x + (R_1 - \tilde{T})\langle x - s\rangle - M_1\langle x - s\rangle^0 \tag{EQ 5.113}$$

$$EIw' = \frac{R_2 x^2}{2} + (R_1 - \tilde{T})\frac{\langle x - s\rangle^2}{2} - M_1\langle x - s\rangle + c_1 \tag{EQ 5.114}$$

$$EIw = \frac{R_2 x^3}{6} + (R_1 - \tilde{T})\frac{\langle x - s\rangle^3}{6} - M_1\frac{\langle x - s\rangle^2}{2} + c_1 x + c_2 \tag{EQ 5.115}$$

and the boundary conditions are:

$$w(s + h) = 0 \tag{EQ 5.116}$$

$$w'(s + h) = 0 \tag{EQ 5.117}$$

$$w(s) = \tilde{\delta} \tag{EQ 5.118}$$

where $\tilde{\delta}$ is the perpendicular displacement of the cable stays, the analog of Equation 5.112 in the displacement domain:

$$\tilde{\delta} = \delta\cos\Theta \tag{EQ 5.119}$$

With these equations, apply the boundary conditions and solve for the displacement at a distance, *s*, from the free end:

$$w|_{x=s} = \frac{1}{EI}\left\{\frac{h^2[3lb^2 + 4lbh - 4h\tilde{T}]}{12}\right\} \quad \text{(EQ 5.120)}$$

Applying the boundary condition in Equation 5.118,

$$w|_{x=s} = \tilde{\delta} = \frac{\tilde{T}L}{AE} \quad \text{(EQ 5.121)}$$

E in Equation 5.120 and Equation 5.121 is the Young's modulus of the material used. In the specific case of the Atlantis, the stub mast is constructed from 6061 Aluminum and the cables from stainless steel stranded cable. Using these material properties, the equations can be solved simultaneously, providing:

$$\tilde{T} = \frac{\frac{lh^2b^2}{4} + \frac{lbh^3}{3}}{\frac{h^3}{3} + L\left(\frac{E_{beam}I_{beam}}{A_{cable}E_{cable}}\right)} \quad \text{(EQ 5.122)}$$

For various cable diameters, the point at which the cable and the stub mast yield simultaneously at the maximum loading can be computed. This results in the simple yield relationships. For the cable stays:

$$\sigma_{yield} = \frac{T}{A_{cable}} = \frac{\tilde{T}}{A_{cable}\sin\Theta} \quad \text{(EQ 5.123)}$$

and for the stub mast:

$$\sigma_{yield} = \frac{M_{max}Y_{max}}{I_{beam}} \quad \text{(EQ 5.124)}$$

In this case the beam is a round tube of 6061 aluminum with a maximum outer diameter of 11.36 centimeters (4 inches). This maximum diameter is set by the inner diameter of the lower needle bearing around the stub-mast. This is the largest standard size needle roller bearing. Specialty larger bearings are available but are very expensive and

beyond the budgetary constraints of the project. Due to the fixed outer diameter both I and Y_{max} are constrained. Y_{max} is set to be the radius of the beam. I is recalculated based on the simplified equation:

$$I_{beam} \cong \pi\{r^3\}t \qquad \text{(EQ 5.125)}$$

where r is the radius of the tube, and t is the wall thickness.

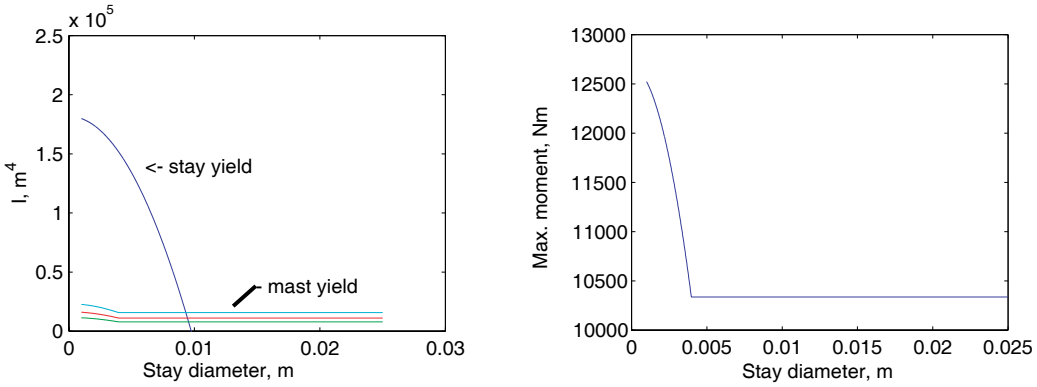

Figure 5-37 Plot showing the yield stress of the cable stay along with the stress in the stub mast for various moments of inertia and cable cross sectional areas. The first frame shows the yield stress line for various moments of inertia and cable stay diameters. The bottom lines show the actual moments of inertial for varying wall thicknesses. The second frame shows the maximum moment in the stub-mast based on the stay diameter.

Figure 5-37 shows the yield limit of the stay versus the stay diameter and moment of inertia for the stub mast. The lower lines show the actual moment of inertia for the stub-mast for various wall thicknesses. It also shows the maximum bending moment in the stub mast versus the stay diameter on the right frame. The conclusion is that beyond a certain size, the stay contributes all it can, and the maximum load experienced by the stub mast occurs above the stays, and thus cannot be remedied by larger stay diameters.

Almost arbitrarily, the stub-mast is chosen to be an 11.43 centimeter diameter tube with a 6 millimeter wall thickness (this corresponds to a standard 4 inch outer diameter pipe with a 0.25 inch wall thickness). The cable stays are chosen to be stainless steel stranded cable with a cross sectional diameter of 9 millimeters (again, this corresponds to

a standard 3/8 inch diameter cable). With these dimensions chosen, the complete maximum loading and deflections may be calculated. Figure 5-38 shows loading of the stub-mast, shear, and moments. Based on these numbers and the previous equations, the stays have a safety factor of eight and the stub mast has a safety factor of 1.35.

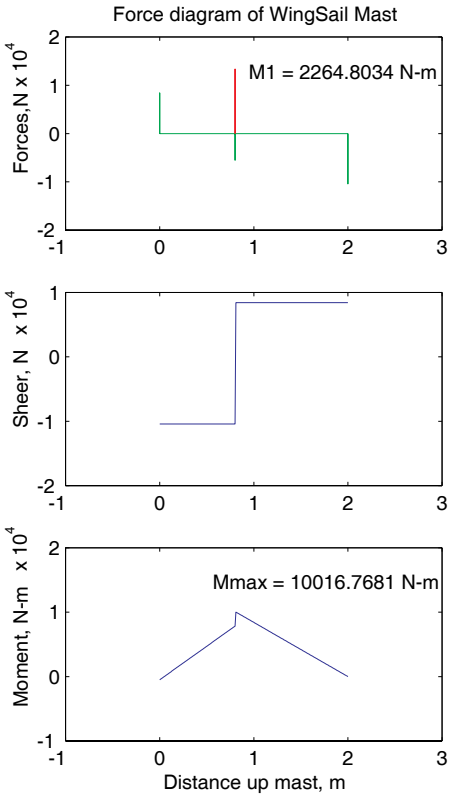

Figure 5-38 Forces, sheer, and moments on the stub mast based on a vertical separation of the bearings of 1.2 meters. The final loads are calculated for the maximum loading condition. The stub-mast is a 6061 aluminum pipe that is11.43 cm (4 inches) in outer diameter and has an internal wall thickness of 9 millimeters (0.25 inches). Based on this shape, the stub-mast can withstand the maximum moment with a safety factor of 1.35.

Finally, with the forces defined and the materials specified, the displacement, *w*, is calculated from Equation 5.115. Using the Young's moduli for both stainless steel and 6061 aluminum, the displacements are computed for the maximum loading and displayed in Figure 5-39. It can be seen that the total displacement at the tip of the stub mast at the maximum loading is less than three centimeters, and that the angular displacement at the

tip is less than two degrees. This is at the worst possible loading conditions, and should correspond roughly to the mast yielding and the spar breaking. The maximum load on the stays is 16 kilo-newtons and well within the capabilities of this size steel cable. In the final analysis, the results show that it is possible to build a structure capable of handling the maximum loading condition.

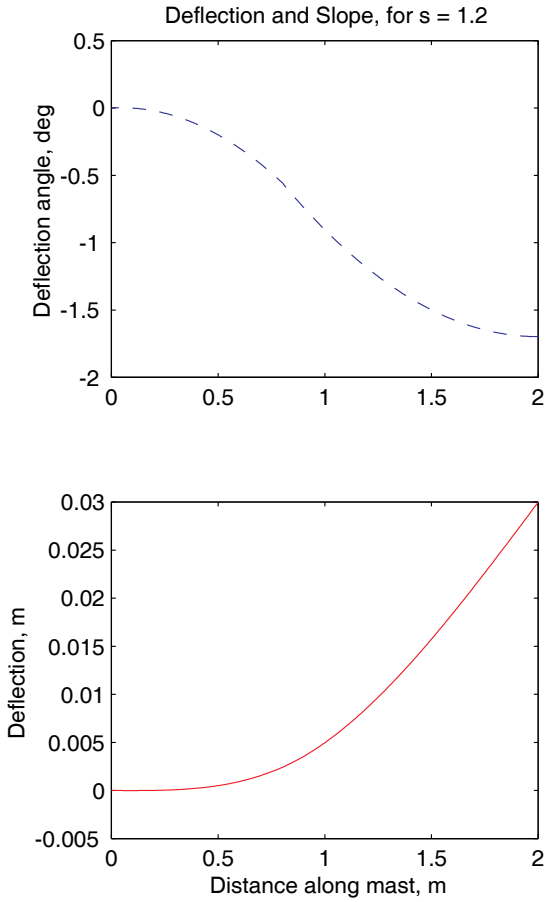

Figure 5-39 Displacements of the stub mast for a 6061 aluminum tube structure of 11.36 cm. in diameter with a 6 millimeters. wall thickness. The final displacement for the stub-mast is 3 cm. at the maximum load condition. The angular displacement of the stub-mast at the end is less than 2 degrees, again at the maximum load condition.

Up until now, this entire analysis assumed a constant wind field that is uniform with altitude. This is not the case, as can be seen in the graph of the wind gradient in [93].

The entire analysis was redone using a simple linear model for wind gradient. The result showed that the current structure remained adequate for the maximum loading. In fact, the safety factor of the stub mast increased slightly while that of the spar decreased somewhat due to the increased loading on the end of the wing.

SECTION 5.8.4 STRUCTURAL DESIGN

The structural analysis presented in the previous section has already demonstrated the type of structure modelled. The actual structural design is very close to that of the design analyzed. The stub mast is secured to the cross-beam through a ball and socket joint, thus rendering the idealized version more complex than the actual one. This was done in order to simplify the attachment process of the mast onto the boat since no welding would be required.

The cable stays were replaced by 6061 aluminum straps that are 2.5 centimeters wide by 9 millimeters thick. This is excessive in terms of strict structural requirements, but they have been repeatedly used as step ladders and hand-holds to maneuver the catamaran while on land and have a very small weight penalty.

The stub mast is standard 6061 aluminum pipe, 11.36 centimeters in diameter with a 9 millimeter wall thickness. Again, this is unnecessarily robust, but the difference in weight was small and since a structural failure would likely have brought all progress to a halt, the decision was made to be conservative. The lower bearing is a simple press fit onto the stub mast; even though the internal diameter of the bearing is the same as the outer diameter of the stub mast, the bearing race required a strong press to slide it into place due to the eccentricity of the stub mast.

Atop the stub mast are the two spherical roller bearings placed to cage the wing onto the stub mast. Also, the Mercotac slip ring is there with four conductors (power,

ground, and the two differential signaling wires) coming out of the stub-mast and looping down into the structure of the lower wing section. The wing section has a pod containing the batteries, ballast, and electronics. This forward pod is used to bring the mass of the entire wing sail and tail assembly in line with the wing quarter chord and bearings.

The wing is built in three sections, each connected by two aluminum tongue and groove joints and secured with stainless steel bolts on either side of the spar caps. While these are sufficiently strong in bending loads across the thickness of the wing, they proved to act as hinges for the in-plane fore and aft loads of the wing. While the prototype was able to sail even with this handicap, future versions will require a better method for joining the wing sections in order to make the entire structure more robust.

The wing and tail are made entirely out of plywood, blue foam, and polyester covering. The wing ribs, spar sheer webs, spar caps, and leading and trailing edge sections of the wing are made out of wood, and the whole thing is covered with polyester cloth that is heat shrunk for a tight fit. The total weight of the complete wing and tail section is 70 kilograms without the ballast weight. The stub-mast and wing spar were tested with a dummy load of 72 kilograms as a point load at the end of the wing and found to withstand that bending load with no damage.

SECTION 5.9 WINGSAIL CONSTRUCTION

Essentially, this section is a pictorial representation of some of the steps taken while constructing the wing. The wing was built by Cris Hawkins Consulting in Santa Rosa, California, over a time period of approximately nine months. This construction included the attachment of the stub mast to the cross beam, creation of the wing sections and tail sections, and the fabrication and installation of the actuators and pushrods.

Figure 5-40 Stub mast, inner bearing surface for needle roller bearing, and stub mast collar for attachment of the aluminum spider. The lower needle roller bearings roll on the surface just above the collar. The collar is used to secure the 6061 aluminum straps (instead of stainless steel guy wires) that support the stub-mast and wing.

Figure 5-40 shows the lower bearing surface and the attachment plate for the top of the 6061 aluminum stringers that replace the stainless steel guy wires on the original construction. These stringers, attached to the mounting plate using stainless steel bolts, are bolted onto hard points of the hulls. The stub mast is shown with the inner part of the needle roller bearing pressed into position above the stub mast collar to which the aluminum straps (spider) attach.

Figure 5-41 shows the portions of the bearings including the spherical roller bearing used for the top of the lower wing section (there are two, and they cage the wing). Figure 5-42 shows the detail of the machined cup used to seat on the original cross beam teflon hemispherical ball joint.

Figure 5-43 shows the stub mast load test, using a dummy point load of 72 kilograms. The stub mast is secured to a replacement cross beam and has two of the six spider

Figure 5-41 Portions of the bearings used to secure the wing to the stub-mast. On the left is the bearing surface for the spherical roller bearings, along with the race in the center. On the right is the needle roller bearing race. The needle roller bearing surface is pressed onto the stub-mast.

Figure 5-42 Stub mast seat that mates with teflon hemisphere on catamaran forward crossbeam. The stub-mast seat was machined out of 6061 aluminum, and is held into the stub-mast by four stainless steel 5/16" machine screws. This allows the easy removal of the stub-mast from the catamaran.

legs attached to the collar. Careful analysis of the load test video showed that the deflection of the stub mast under load test was, in fact, caused by deflection of the wooden building column that was used to secure the cross beam in place. Both of the spherical roller bearings are secured in position on top of the stub mast. This can be seen next to the dummy load's hands. The dummy load was increased to a 153 kilogram point load on the end and the deflection remained undetectable after the deflection of the wooden column

was accounted for in the measurements. This did, however, cause some concern about the stability of the building during the load test, but the roof remained in the appropriate position.

Figure 5-43 Stub mast, two spider legs, and cross beam load tested with a 72 kilogram static dummy load. The stub-mast is attached to a replacement crossbeam that is secured to a wooden column supporting the building. The two spider legs are secured to the crossbeam. Close inspection of the figure shows the two spherical roller bearings at the end of the stub-mast. No deflection occurred in the stub-mast, though the wooden column was deflected under the test load.

Figure 5-44 shows the master wing rib jig. This is milled from a high density plastic using a computer controlled milling machine that is programmed from the points generated from the XFOIL program. This pattern is used to route out all of the main wing ribs and ensure dimensional accuracies are kept throughout the construction. The main wing ribs were cut from marine grade plywood. There are several interesting features of the jig that can be noted in the picture. Circular holes in the front and back of the rib are used to assemble cut ribs onto a jig made from electrical conduit. The eleven smaller holes are for threaded rods to hold the stack of plywood sheets together, thus ensuring uniformity of

fabrication. The notches in the top and bottom are for the spar caps, and the lightening holes in the forward and rear center are to reduce weight.

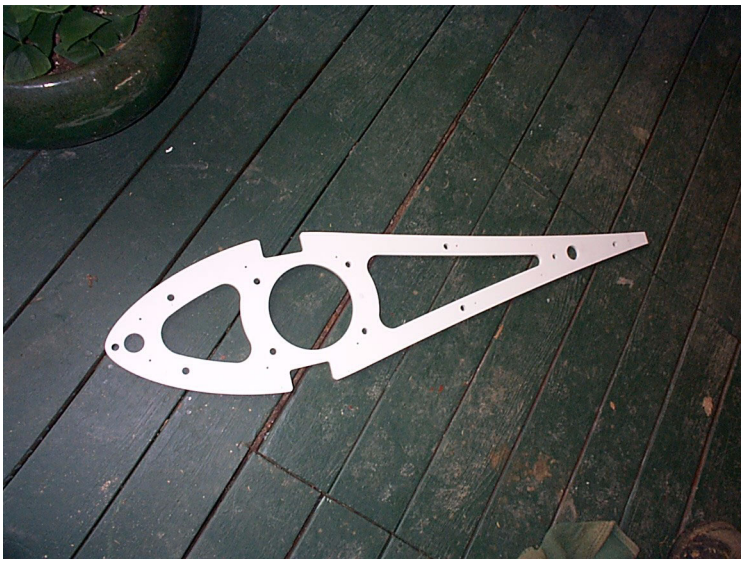

Figure 5-44 Master main wing rib template used to fabricate all wing ribs from marine grade plywood. The large holes are to lighten the ribs. The two medium sized holes forward and back are for a assembly onto a jib made of electrical conduit. The eleven small holes are for threaded rods that secure the stack of plywood together to ensure uniform fabrication.

Figure 5-45 shows the ribs aligned on the jig, with the spar caps glued in place; the large front doubler ribs are for the electronics pod and counter weights, the use of PVC pipe spacers is to ensure the uniform spacing of the wing ribs. In the foreground is the lower wing section. In the back, the center section of the wing can be seen, also with the spar caps gluing in place. All joints are glued with epoxy to ensure maximum joint strength. Epoxy has the added advantage that it will not spontaneously disassemble due to increased moisture or direct immersion in water. Close inspection of Figure 5-45 reveals a hole pattern at the top of the lower wing section spar cap. This is where the two 6061 aluminum 3/8” thick plates will be attached to either side of the spar cap and be used as the slot for a mortise and tenon joint. This joint uses a 5/8” thick 6061 aluminum tongue attached to the bottom of the center section spar caps. This functions not as a draw bar mortise and tenon joint, but rather a plain mortise and tenon in which the wedge is

replaced by two stainless steel bolts on either side of the wing. The inside 6061 aluminum plate is tapped for the right threads. Though overtightening is an issue, as long as the bolts are tightened to the right torque, they will hold the sections together easily.

Figure 5-45 Main wing ribs on jib, spar caps, and pod ribs extending forward on lower wing section. The forward pod ribs are double thickness plywood, and the PVC pipe spacers ensure uniformity in the rib spacing. The spar caps are glued in place with epoxy. The top of the spar caps has been drilled for the aluminum mortise and tenon joint that holds the wing sections together. The three sections will be held together with stainless steel bolts at the joints.

Figure 5-46 shows the leading edge skin being glued onto the ribs to form the front "D" tube assembly. This "D" tube resists the torsional loads imposed by the lift and flap on the wing section and keeps it from twisting. Severe problems with cracking of the leading edge wing skins were encountered while attempting to secure the leading edge skins. Soaking in water only resulted in the outer layers of the marine plywood absorbing water and proved unsuccessful. The solution is to thin the leading edge to half its original thickness at the location of the maximum curvature, and then soak the plywood in water. In retrospect, it would have been wise to reinforce this thinned leading edge with fiberglass and epoxy from the inside before the shear webs were glued between the spar caps. The lead-

ing edge proved to be an extremely delicate area of the finished wing. Great care had to be taken to avoid cracking the leading edge and the wing sections could never be allowed to support their weight on the leading edge.

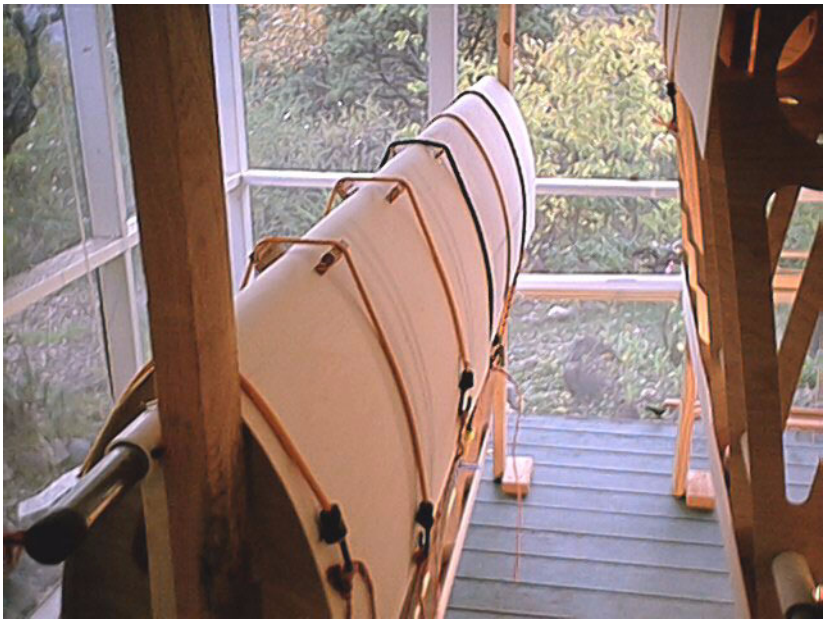

Figure 5-46 Plywood leading edge skin glued to wing ribs. The wing skins suffered severe cracking problems when bent around the leading edge. In order to accommodate the sharp radius of curvature, the wing skins were thinned and soaked in water before gluing them on to the ribs to form the forward "D" tube. This area remained weak and prone to damage in the finished wing structure.

Figure 5-47 shows the sheer webs, looking inside the wing. The sheer webs are made out of the same marine plywood as that of the wing ribs. Lightening holes can be seen cut out of the sheer webs as well as the gap between the spar caps and the sheer web. The gap is required only on the lower section in order to clear the lower bearing and stub mast, which rises up through the center of the hole cut out of the main wing ribs. In order to make up for the distance between the spar caps and the sheer web, the spar cap on the lower wing section is extended back to butt up against the sheer web, and is glued with epoxy and fiberglass to the ribs, sheer web, and wing ribs. This is necessary because the sheer webs were found to buckle, with the center narrowed section of the sheer web twist-

ing into a potato-chip-like shape when the entire wing assembly was subjected to a 72 kilogram dummy point load on the end. The shear webs on the lower section were made solid (no lightening holes) and were increased to 5/8" thickness from the nominal 1/4" plywood that was used on the rest of the sheer webs.

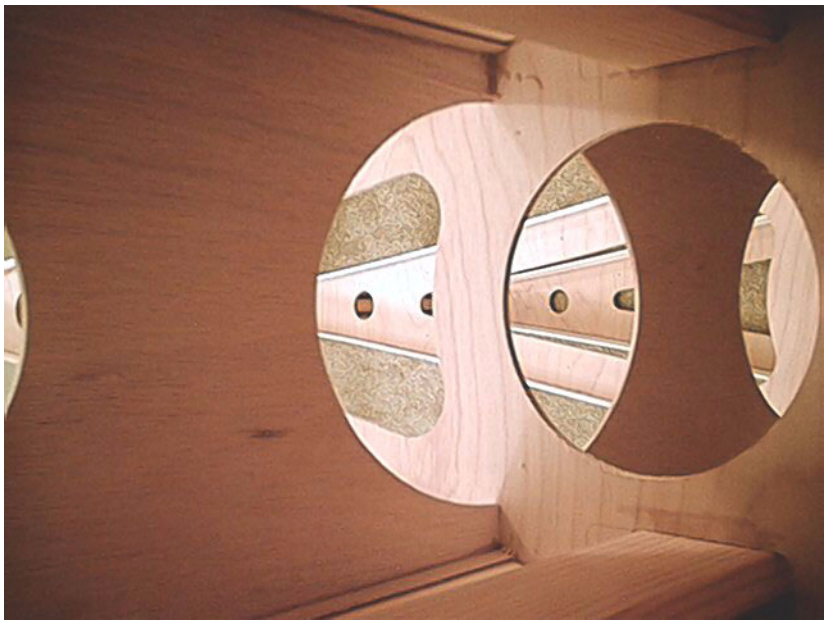

Figure 5-47 Plywood sheer webs join the leading edge skin and upper and lower spar caps. Lightening holes are cut in the sheer webs. Note the distance between the rear of the spar cap and the sheer web. This is because this is the lower section, and the stub-mast will fit just inside the circular opening in the rib. After the load test, the spar cap was extended back to the sheer web and the sheer web reinforced with thicker plywood.

Figure 5-48 shows the three sections of the wing assembled for the final load test before covering. The ladder in the foreground is not actually supporting the wing at all, but is there to prevent the trailing edge from rotating downwards, as the connection between the stub mast and the wing spar is through bearings, and is designed to allow the wing to rotate freely about the axis down the center of the wing spar. At the front of the lower section is the pod for the electronics and counter weight ballast, with the lid removed. The upper flap actuator is visible on the fifth rib down from the top of the wing. The load test was conducted by placing the same 72 kilogram dummy load on the end of

the wing, and resulted in the reinforcement of the bottom section sheer webs. Following the reinforcement, a 72 kilogram dummy load was again placed on the end of the wing, simulating 70% of the maximum loading scenario. This resulted in a 15 centimeter deflection at the end of the wing, though most of this was due to the wooden column support of the building deflecting as well as the cross beam pulling off its mounting. The residual deflection was about 5 centimeters.

Figure 5-48 The final wing assembly setup for load testing, before covering. The ladder is only supporting the rear edge of the wing from rotating downwards (as the wing is attached to the stub-mast via bearings). The diagonal internal brace just above the ladder is the anti-drag bracing. A 72 kg load was suspended from the end of the wing, and the deflection was recorded to be approximately 15 cm. After corrections were made, the residual deflection was 5 cm.

The anti-drag bracing can be seen diagonally bracing the top to third rib. These anti-drag braces give the wing strength when bending in the plane of the wing (in this picture, pulling the top of the wing to the right horizontally). The three sections are pinned together using the mortise and tenon joints, as previously explained. There is no connection of the three sections at the trailing edge. This later proved to be a weakness in the design, as the mortise and tenon joints act as hinges during high velocity pitch motions of the wing (as when crossing through waves). These effectively allow the three sections to

open up like a fan and then come crashing back together, damaging the lower trailing edge structure. A simple method of joining the trailing edge together would mitigate this problem and cause the entire wing to behave in a rigid fashion when pitching front to back.

The wing is covered with “Coverite,” a thick polyester fabric normally used for model airplanes. The fabric is coated on one side with a heat activated glue and with chemical resistant paint on the other. The covering is resistant to water, salt water, oil, alcohol and gasoline.

Figure 5-49 Coverite polyester fabric is used to cover the wooden wing structure. Cris Hawkins of Cris Hawkins consulting applies heat to shrink the fabric onto the ribs. The covering is oil, gasoline, and salt-water resistant. In order to prevent the fabric from pulling off the ribs on the rear section of the airfoil, the fabric was glued down to the ribs during the shrinking process through a process of applying pressure while the coverite was allowed to cool.

Figure 5-49 shows Cris Hawkins of Cris Hawkins Consulting shrinking the covering onto the upper wing section. The concavity of the main wing section requires that the covering be firmly glued onto each rib cap before the final shrinking can take place. Furthermore, great care has to be taken in order to keep the hot, pliable polyester fabric from

detaching from the rib cap while the covering cools into position. This is accomplished by the use of cooling pads that keep the sections of the covering directly above the rib cap from reaching a temperature sufficient to allow the glue bond to lose its strength.

The same covering is used to make the hinges for the trailing edge flaps. These are so called figure eight fabric hinges which allow the flap to deflect through a 180 degree arc without imposing any moment on the surface itself. Another benefit of these hinges is that they effectively seal the gap between the trailing edge of the main wing section and the flap itself. Figure 5-50 shows the lower section trailing edge flap.

Figure 5-50 main wing trailing edge flap with pushrod, control horn, and fabric hinge. The figure 8 hinge is made from the same covering material that covers the wing. The advantage of this kind of hinge is that there is very little hinge friction. Additionally, the hinge seals the gap between the flap and main section, while at the same time allowing a large range of motion.

With the wing sections built, and covered, the next task is to install all of the electronics and wiring, as well as some extra sealed flotation balloons in case of a capsize. With this accomplished, the entire wing sail and tail assembly is very tail heavy. This is to be expected as the entirety of the mass of the booms and tail are very far behind the main wing quarter chord line. In order to bring the center of mass of the entire wing sail and tail assembly in line with the quarter chord, each section is weighed, and the center of gravity position noted relative to a reference at the quarter chord center. This allows the correct ballast position to be computed and the ballast to be added to the electronics pod. In order to correctly balance the wing on the quarter chord, a 25 kilogram battery is placed into the pod, as well as a 12.7 kilogram lead brick. Figure 5-51 shows the interior of the pod.

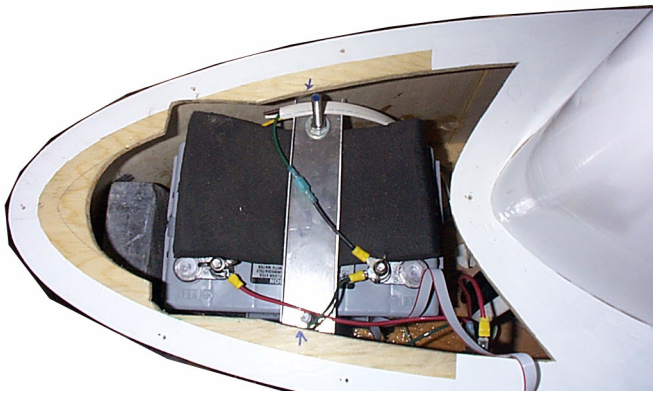

Figure 5-51 Electronics pod, showing the battery and ballast weight. Inside the pod is the main battery, secured by two threaded stainless steel rods. The black material is neoprene for cushioning the electronics. Forward of the battery is a 12.7 kg lead brick that is used to mass balance the wing. The wires lead to the main bus breaker on the side of the electronics pod. The ribbon cable joins the can bus and the anemometer microcontroller which is secured to the underside of the pod lid.

The breakdown for the weight and balance of the wing sections is summarized in Table 5-1 below:

Section	Mass (kg)	Distance aft (cm)
Lower wing section	26.81	1.3
Center wing and tail	29.10	53.6
Upper wing	14.55	15.9
Battery	25.00	-42.6
Ballast	12.70	-59.7
All Sections	108.61	-2.0

Table 5-1: The breakdown for weight and balance of the wingsail sections. The mass of each section was measured with a spring scale and the distances using a two-point suspension method to mark the center of gravity. The net result is a very nearly mass balanced wingsail that exhibits no tendencies to rotate when pitched or rolled.

This leads to a total weight for the wing of 108.61 kilograms and a nominal offset of -2.0 centimeters, slightly nose heavy. This configuration allows the wing to point away from the wind in an upwind heel, reducing lift and stabilizing the sailboat. With the construction of the wing complete, the propulsion system of the Atlantis has been described in detail.

The wing, spider, and hulls can be seen in Figure 5-52 which shows the entire system during a final system check. This is a composite image, and there are no sharp discontinuities in either the wing or the hulls

Figure 5-52 Final Atlantis wing, with spider below and electronics pod. This is a composite image made up of several photographs of the Atlantis taken inside the HEPL high bay entrance. The entire system was assembled inside of a hangar in order to perform a final system check before performing the water trials. The clearance between the top of the wingsail and the roof of the hangar is approximately 12 cm.

6 System Identification and Control Design

The main goal of the Atlantis project is precise control. That is, to minimize the deviation from a desired path in the presence of disturbances such as wind, current, and waves. Precise control is achieved through the development of a mathematical model that accurately responds to external inputs so as to converge upon a predictive model. The process of control is predicated on having a mathematical model of that which is being controlled (referred to as the plant). This is because the controller must be able to predict the trajectory of the plant, and compensate accordingly. Without an accurate mathematical model of the dynamics of the catamaran, precise control is very difficult. System identification techniques are used to extract the required model from experimental data.

SECTION 6.1 SYSTEM IDENTIFICATION OVERVIEW

System identification consists of any number of techniques that attempt to mathematically extract the model of the plant from experimental data. There exists a rather large number of techniques that use the frequency domain response to determine the system model. Typically, these techniques use frequency sweeps or chirps as inputs to the experiment. Other techniques use time domain data to determine the model. Some techniques assume a certain model form and essentially match the parameters of that model to the data. If only a few parameters in the model are unknown, then the extended Kalman filter can be used to perform parameter identification, either as a batch process or as an adaptive process.

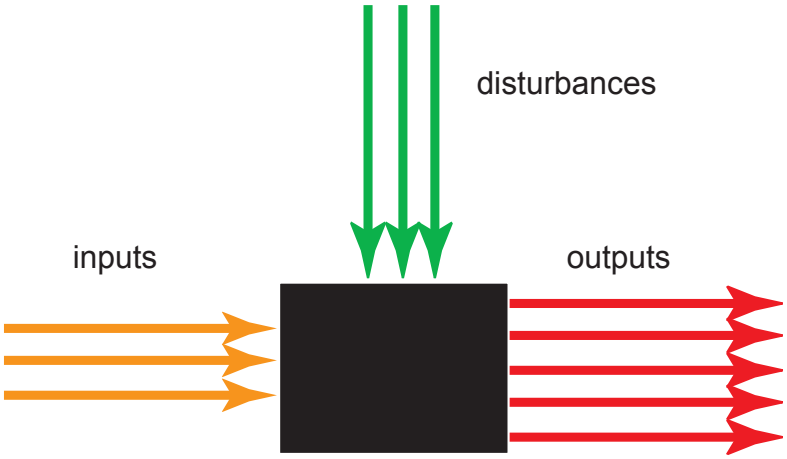

Figure 6-1 The basic structure of the system identification process. Given a "black box" type system, with known inputs, unknown disturbances, and measured outputs, the system identification technique synthesizes a mathematical model for the "black box" under certain assumptions about the character of the disturbances.

The essence of the system identification process is pictured in Figure 6-1, in which a "black box" is shown with known inputs, unknown disturbances generally assumed to be driven by a white noise process, and measured outputs. System identification replaces the "black box" with a mathematical model. Some art is required in the experiment design to

minimize the disturbances and get the best mathematical model (in the sense of the highest fidelity).

There exist alternatives to system identification in order to generate appropriate system models. Generating a model of plant dynamics from first principles and linearizing about an equilibrium state is the classical method of generating an operative mathematical model. Other methods include model free controller design and adaptive control schemes.

SECTION 6.1.1 OBSERVER/KALMAN IDENTIFICATION THEORY

There are a great many ways to approach system identification. The entire process generally starts from experimental test data and uses some mathematical techniques to derive a model of the system. This model is then used to predict the system response for arbitrary inputs. Typically, the experiment consists of exciting the inputs with some rich set of data, such as white noise, and measuring the response at the sample points. The system identification algorithm then produces a differential or difference equation of the system response. The models generated by the system identification techniques are most often used in feedback control systems and aid in the design of Kalman filters and regulators.

An important advantage of the experimental approach is that does not require the tedious and often imperfect modelling of the physical system. This is especially true if large assumptions are made to reduce the model complexity. The primary disadvantage is that, depending on the scheme, any intuition into system behavior is not leverages and is, in fact, lost. Without this intuition, it is difficult to view the internal system description and understand which physical parameters are being modeled. As such, any additional knowledge about the valid range of physical parameters cannot be utilized.

Several methods have been developed over the years for performing the system identification. Ljung [88] provides an excellent introduction to the subject, including the various methodologies that have been developed. Broadly, the methods are either frequency- or time-domain-based, and tend to be grouped into single input single output (SISO) or multiple input multiple output (MIMO) solutions. Recently, a large body of state-space techniques have been aggregated into the subspace methods, presented in [105]. These techniques in general perform well, though one of their deficiencies is that the subspace methods usually produce a biased estimate of the states. In addition, rather than minimize the error of the output data, they minimize the model error. This is due to the fact that the minimization of the output error is non-linear, whereas the modelling error can be reduced to a simple quadratic minimization.

Another time domain technique is the Observer Kalman filter IDentification (OKID) algorithm developed at NASA Langley to model large flexible space structures. The original algorithm was developed and extended by Dr. Juang and his students to include residual whitening and several advances in the model realization algorithms [110]. Also, the OKID algorithm minimizes the error in the observer, which will converge to the true Kalman filter for the data set used, given that the true world process is corrupted by zero-mean white noise.

The method of identifying the plant models, as well as the process and sensor noise statistics, chosen for this project is the observer Kalman filter identification (OKID) method. This method of system identification uses only input and output data to construct a discrete-time state-space realization of the system. Since OKID's development at NASA Langley for the identification of lightly-damped space-structures, many advances on the basic theory have been published [110]. The OKID identification method has several pertinent advantages. First, it assumes that the system in question is a discrete linear time-invariant (LTI) state-space system. Second, it requires only input and output data to for-

mulate the model, no *a priori* knowledge of the system is needed. Third, the OKID method produces a pseudo-Kalman state estimator, which is very useful for control applications. Last, the modal balanced realization of the system model means that truncation errors will be small. Thus, even in the case of model order error the results of that error will be minimal.

The first step in applying the OKID methodology to the experimental data begins with the standard state-space difference equation for an LTI system:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{aligned} \qquad \text{(EQ 6.1)}$$

where k is the index time variable, x is the state vector and has the dimension of *[n x 1]*, y is the output vector and has the dimension of *[m x 1]*, u is the input vector and has the dimension of *[r x 1]*. The *[n x n]* matrix, A, is referred to as the state transition matrix, the *[n x r]* matrix, B, is called the input matrix, the *[m x n]* matrix, C, is the output matrix, and *[m x r]* matrix, D, is sometimes called the feed through or pass through matrix. It has been shown that the triplet, *[A,B,C]* is not unique, but can be transformed through any similarity transformation (i.e., the outputs are unique, but the internal states are not). This can be seen with a simple substitution:

$$x \equiv Tz \qquad \text{(EQ 6.2)}$$

where z is an alternate state vector related to x through a transformation matrix, T. When substituted into Equation 6.1 results in the same set of equations written as:

$$\begin{aligned} z_{k+1} &= T^{-1}ATz_k + T^{-1}Bu_k \\ y_k &= CTz_k + Du_k \end{aligned} \qquad \text{(EQ 6.3)}$$

Thus, the triplet *[A,B,C]* through a similarity transformation becomes $[T^{-1}AT, T^{-1}B, CT]$. However, if the system is assumed to be initially at zero, and the inputs are set to unit values for one sample at time zero, then:

$$\begin{aligned} x_0 &= \lfloor 0 \rfloor \\ u_0 &= \lfloor 1 \rfloor \\ u_k &= \lfloor 0 \rfloor \end{aligned} \qquad \text{(EQ 6.4)}$$

for all k not equal to zero. The convolution of Equation 6.1 results in the unit pulse response of the system:

$$\begin{aligned} y_0 &= D \\ y_1 &= CB \\ y_2 &= CAB \\ y_3 &= CA^2B \\ &\cdots \\ y_k &= CA^{k-1}B \end{aligned} \qquad \text{(EQ 6.5)}$$

which are defined as the Markov parameters of the system, and are invariant under a similarity transformation. In general, the response of the system can be shown to be:

$$\begin{aligned} x_k &= \sum_{i=1} A^{i-1}Bu_{k-i} \\ y_k &= \sum_{i=1}^{k} CA^{i-1}Bu_{k-i} + Du_k \end{aligned} \qquad \text{(EQ 6.6)}$$

By using the conditions set out in Equation 6.4, and assuming that the input, u, is a unit pulse at time zero, then the response reverts to the Markov parameters of the system. These Markov parameters are assembled into a specific form, the generalized Hankel

matrix. Properties of the Hankel matrix can be exploited in order to determine the system order. The generalized Hankel matrix is:

$$H_{k-1} = \begin{bmatrix} Y_k & Y_{k+1} & \cdots & Y_{k+\beta-1} \\ Y_{k+1} & Y_{k+2} & \cdots & Y_{k+\beta} \\ \cdots & \cdots & \cdots & \cdots \\ Y_{k+\alpha-1} & Y_{k+\alpha} & \cdots & Y_{k+\alpha+\beta-2} \end{bmatrix} \quad \text{(EQ 6.7)}$$

This matrix can be decomposed into the observability matrix, a state transition matrix, and the controllability matrix; thus the Hankel matrix (in a noise-free case) will always have rank n, where n is the system order.

$$H_{k-1} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{k-1} \end{bmatrix} \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \ldots & A^{k-1}B \end{bmatrix} \quad \text{(EQ 6.8)}$$

This decomposition is at the center of the OKID process. Because the Hankel matrix should have rank n, then the system order can be determined merely by looking at the rank of the generalized Hankel matrix of the Markov parameters. Later it will be shown how to determine the Markov parameters; assume at this point they are known. Unfortunately, in real data, noise will corrupt the rank deficiency of the Hankel matrix (the Hankel matrix will always be full rank). Thus, the Hankel matrix is truncated using a singular value decomposition (SVD) at an order that sufficiently describes the system. In practice, the singular values of the Hankel matrix are plotted and the singular values will decrease gradually until a sudden drop that indicates the model order. This sudden decrease or "cliff" is the hallmark of the transition between real and noise modes of the system.

This truncated Hankel matrix is then used to reconstruct the triplet *[A,B,C]* in a balanced realization that ensures that the controllability and observability Grammians are equal. This is referred to as the Eigensystem Realization Algorithm (ERA); a modified

version of this algorithm that includes data correlation is used to identify the Atlantis. A more complete treatment of the subject can be found in [110].

Thus, given the Markov parameters, it is possible to determine the order of the system and generate a balanced model that is adequate for control. In real systems, however, the system pulse response cannot be obtained by simply perturbing the system with a pulse input. A pulse with enough power to excite all modes above the noise floor would likely saturate the actuator or respond in a non-linear fashion. The pulse response of the system can, however, be reconstructed from a continuous stream of rich system input and output behavior. Under normal circumstances, there are not enough equations available to solve for all of the Markov parameters. Were the system asymptotically stable, such that $A_k=0$ for some k, then the number of unknowns could be reduced.

Considering the original system in Equation 6.1:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{aligned} \qquad \text{(EQ 6.9)}$$

and adding zero, in the form $(+Gy_k\text{-}Gy_k)$ of an observer:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Gy_k - Gy_k \\ y_k &= Cx_k + Du_k \end{aligned} \qquad \text{(EQ 6.10)}$$

By substituting in the equation for y_k, the system can be rewritten as:

$$\begin{aligned} x_{k+1} &= \bar{A}x_k + \bar{B}v_k \\ y_k &= Cx_k + \bar{D}v_k \end{aligned} \qquad \text{(EQ 6.11)}$$

where:

$$\bar{A} = A + GC \qquad \text{(EQ 6.12)}$$

$$\bar{B} = \begin{bmatrix} B + GD & -G \end{bmatrix} \qquad \text{(EQ 6.13)}$$

$$\bar{D} = \begin{bmatrix} D & 0 \end{bmatrix} \qquad \text{(EQ 6.14)}$$

and:

$$v_k = \begin{bmatrix} u_k \\ y_k \end{bmatrix} \quad \text{(EQ 6.15)}$$

A convolution is performed on the new observer equations, Equation 6.11, and the output at any time can be written as:

$$y_k = Du_k + C\overline{B}v_{k-1} + \ldots + C\overline{A}^{p-1}\overline{B}v_{k-p} + C\overline{A}^{p}\overline{B}x_{k-p} \quad \text{(EQ 6.16)}$$

where p is a time-step integer. In matrix form, Equation 6.16 can be rewritten as:

$$\bar{y} = \overline{Y}\overline{V} + C\overline{A}^{p}X + \varepsilon \quad \text{(EQ 6.17)}$$

where $\bar{y}$ is the collection of output data, $\overline{Y}$ is the matrix of observer Markov parameters, $\overline{V}$ is the stack of input and output data, X is the state, and ε is the mis-modelling noise:

$$\bar{y} = \begin{bmatrix} y_{p+1} & y_{p+2} & \cdots & y_l \end{bmatrix} \quad \text{(EQ 6.18)}$$

$$\overline{Y} = \begin{bmatrix} D & C\overline{B} & C\overline{A}\overline{B} & \ldots & C\overline{A}^{p-1}\overline{B} \end{bmatrix} \quad \text{(EQ 6.19)}$$

$$\overline{V} = \begin{bmatrix} u_{p+1} & u_{p+2} & \cdots & u_l \\ v_p & v_{p+1} & \cdots & v_{l-1} \\ v_{p-1} & v_p & \cdots & v_{l-2} \\ \cdots & \cdots & \cdots & \cdots \\ v_1 & v_2 & \cdots & v_{l-p} \end{bmatrix} \quad \text{(EQ 6.20)}$$

The new observer state transition matrix, $\overline{A}$, has some unique properties that can be exploited for system identification. In a noise-free case, the observer can be made dead-beat and $\overline{A}^p$ will be zero for time-steps greater than the system order, n. In the presence of noise, $\overline{A}$ corresponds to a Kalman filter. In either case, it is asymptotically stable, and there will be a value of p such that the quantity $\overline{A}^p$ is negligible. In the OKID algorithm, p is chosen such that the middle term of Equation 6.17 is negligible. The equation for the observer Markov parameters is solved using the standard least squares solution:

$$\overline{Y} = \bar{y}\overline{V}^T[\overline{V}\overline{V}^T]^{-1} \quad \text{(EQ 6.21)}$$

This minimizes the error term, $\varepsilon^T\varepsilon$, which is the difference between the output obtained using this model (with $\overline{A}^p$ neglected) and the measured values. Then, from Y the Markov parameters of the system are computed and separated out using the ERA/DC algorithm detailed in [76], [77] and [110]. These references go into great detail to show the conditions under which the computer observer will be the unbiased Kalman filter. They also present modifications to the basic OKID algorithm to whiten the residuals in the case that these conditions are not met (usually too little data). That is, the algorithms are modified such that no correlation remains in the post-identification residual errors.

SECTION 6.1.2 SPRING-MASS-DAMPER OKID EXAMPLE

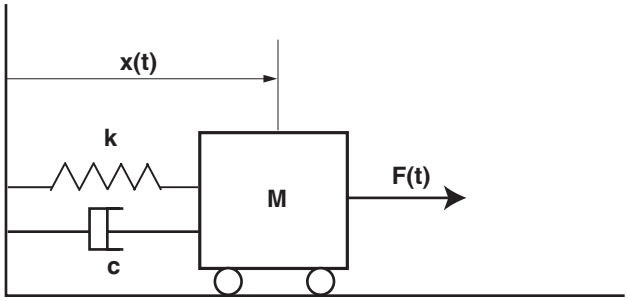

Figure 6-2 Simple system to use as an example for the OKID identification algorithm. The simple mass spring damper system has a damping ratio of 0.05 and a natural frequency of 25 radians/second. The system is driven with a pseudo-random force input on the mass and the displacement output is recorded.

With this basic understanding of the OKID system identification algorithm, a simple example is useful in order to see the performance of the system in the case where the model is known ahead of time. To this end, a simple mass spring damper system (pictured in Figure 6-2) is modelled, corrupted by noise, and then identified. The equations of motion are derived from the first principles of Newton's third law: $F=ma$. A simple free body diagram accounts for the spring and dash-pot terms and they are included in the

equations of motion. The equations of motion for such a system are simple, and in state-space form they are:

$$\begin{bmatrix} \dot{x} \\ \dot{\dot{x}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F \qquad \text{(EQ 6.22)}$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} + [0]F$$

where the input is the force and the output is the measured displacement. The mass of the block is assumed to be 4 kilograms, the spring constant 2500 newtons/meter and the dash-pot damping rate equal to 10 kilograms/second (or 10 newton-seconds/meter). This results in a lightly damped system, with a damping ratio of 0.05, and a natural frequency of 25 radians/second (or just under 4 Hz.).

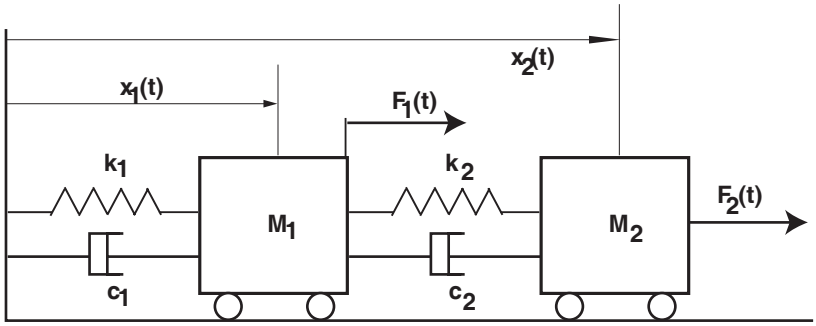

Figure 6-3 A more complex system to demonstrate the ability of the OKID algorithm to correctly identify the system, even in the presence of noise. This second block has very little damping and the only sensor remains the position of the first block. The identification is performed on data modelled using a pseudo-random input on the first mass.

In order to show the power of the OKID algorithm, the algorithm will be demonstrated to identify a more complex system in which there is a hidden and unknown resonance. Take the same system shown in Figure 6-2, and add a second mass block, with another spring and dash-pot between them with very light damping. This example is analogous to a problem of unknown resonance in a structure. The only output that is measured is still the position of the first block. This system is shown in Figure 6-3. The rationale

behind this complex example is that a system may contain an unknown resonance. The input to the system remains the same as the simple example, but the difference in response is due to the second mass responding to the input. Even in this case, the system identification methodology is capable of modelling the full dynamics of the system even though the presence of the second mass is only felt as a disturbance in the position of the first and larger mass. This is a much more challenging task for the system identification process and the presence of noise will greatly affect the outcome.

The equations of motion for the more complex system are:

$$\begin{bmatrix} \dot{x}_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k_1 + k_2}{m_1} & \frac{k_2}{m_1} & -\frac{c_1 + c_2}{m_1} & \frac{c_2}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{c_2}{m_2} & -\frac{c_2}{m_2} \end{bmatrix} \begin{bmatrix} x_1 \\ \dot{x}_1 \\ x_2 \\ \dot{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{m_1} \\ 0 \end{bmatrix} F \qquad \text{(EQ 6.23)}$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \dot{x}_1 \\ x_2 \\ \dot{x}_2 \end{bmatrix} + [0]F$$

where the values for the original mass, spring and dash-pot remain the same, but the additional mass, m_2, is 1 kilogram, the additional spring, k_2, is 800 newtons/meter, and the additional dash-pot, c_2, is 0.5 kilograms/second. This results in a new system having a fourth order response, rather than a second order response. The new system has four poles and two zeros. The natural frequencies are at 20.3 radians/second and 34.8 radians/second, with the slower roots having a damping ratio of 0.0315 and the faster roots having even less damping with a ratio of 0.0265.

The nature of this very lightly damped system is difficult to identify correctly, especially the damping ratios. Figure 6-4 shows the poles and zeros of the simple and complex systems. This is a standard complex plane, with the real axis along the horizontal, and the imaginary axis vertical. Stability is indicated by the poles remaining to the left of the imaginary axis. The horizontal axis has been greatly exaggerated in order to see the poles and zeros as they are very close to the $j\omega$ axis.

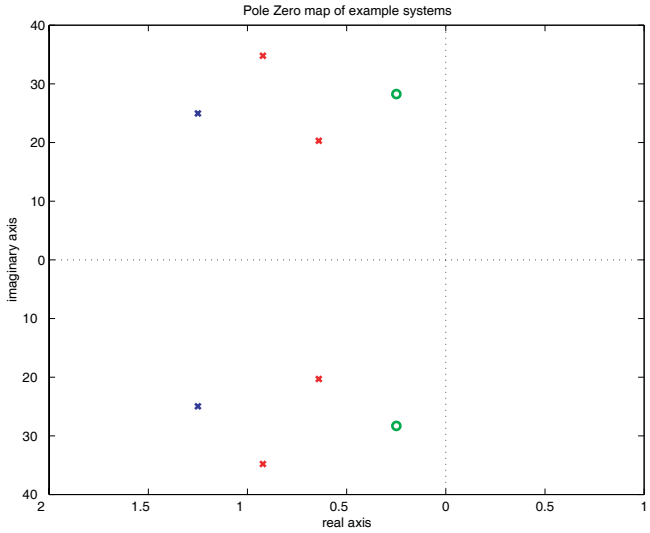

Figure 6-4 The complex plane representation of the poles and zeros of the simple and complex mass-spring-damper examples for the OKID algorithm. The poles for the simple system are shown in blue, and the complex system in red.

Both systems were simulated with a constant random input that had the force set at zero, +100 N, and -100 N. The input is zero mean and has a standard deviation of 81 newtons. A total of six seconds of data is used to excite the system, initially at rest, using this input. The system is sampled at 100 Hz, providing 600 samples for the OKID algorithm. The excursions of the first mass vary as much as 0.2 meters from the nominal position.

In Figure 6-5, the input stream is shown along with the output for the two systems. The simple system is shown is blue, and the complex system is shown in red. In this view, it is very difficult to distinguish the simple from the complex system. Furthermore, the

outputs that are shown are free of noise. The output was then corrupted by noise equal to about 25% of the standard deviation of the actual output ($\sigma_N = \frac{\sigma_Y}{4}$). This was used as the data to run the OKID system identification algorithm. The first step in the OKID process is to arbitrarily choose a maximum system order, p. In both of these examples, the upper bound was set at 25 (this is the value for p, such that $\overline{A}^p = 0$).

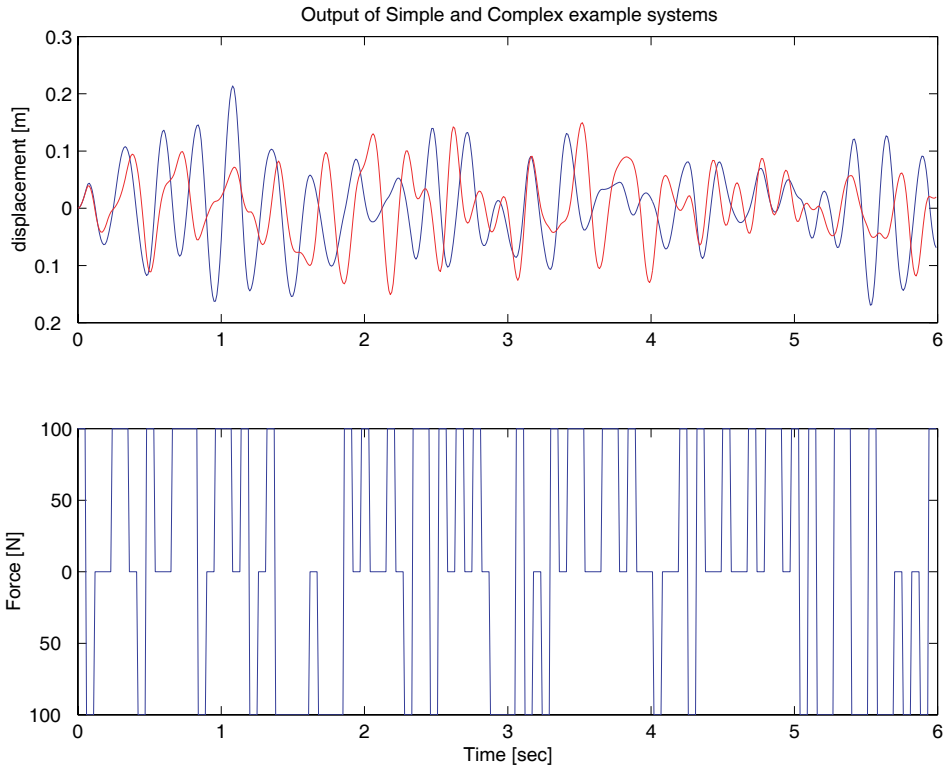

Figure 6-5 The input and output of the example systems. The simple system is shown in blue and the complex system is shown in red. These are the noise free outputs; white noise is added to the outputs before the OKID process.

Once p is defined, the OKID algorithm then sets up the singular value decomposition (SVD) of the Hankel matrix and plots the singular values. In the noise free case, the singular values would drop off radically from n to $n+1$. In the case with noise, the drop will not be as noticeable. Some engineering judgement is required to pick the appropriate order at which to truncate the system. The OKID process shows the modal singular values

of the system model, and at that point the system is truncated (using a balanced realization) to the desired order. Figure 6-6 shows the Hankel and Modal singular values for the simple system. In this case, the drop off is readily apparent at the second mode.

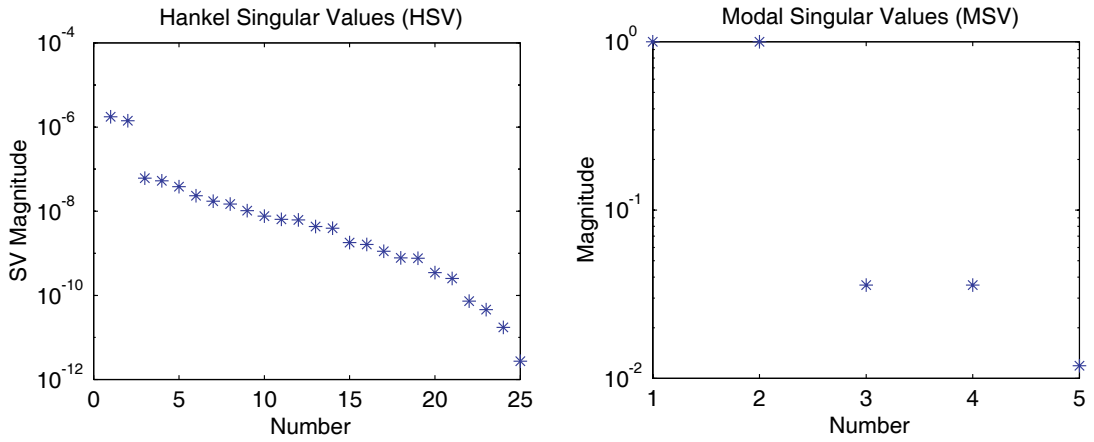

Figure 6-6 Hankel and Modal singular values for the OKID of the simple system. The drop off in the Hankel singular values after the second mode is sharp and noticeable, and is corroborated by the Modal singular values. Note that this includes additive white noise that is 25% of the output levels. This indicates very strongly that the system is second order.

Once the choice of the number of modes is made, the OKID algorithm generates the system model, including the pseudo-Kalman filter. This is extremely useful for control since the optimal state estimator (based on the data) is already provided. All that is required is a controller to regulate the system and the control synthesis is done. In the case of the simple system, the actual and identified systems are very close. This can be seen from the damping ratios and frequencies of the poles, shown in Table 6-1 below. The fre-

quency is exactly matched up and the damping ratio is very close. This included the effect of fairly large noise on the output measurement.

Eigenvalue	Magnitude	Equiv. Damping	Equiv. Freq. (rad/s)
0.957+0.244i	0.988	0.005	2.5
0.957-0.244i	0.988	0.005	2.5
0.956+0.243i	0.987	0.00535	2.5
0.956-0.243i	0.987	0.00535	2.5

Table 6-1: The true and identified system poles. The true system is shown in green, and the OKID system is shown in blue. Note that the magnitude and damping of identified system is virtually identical to the true system. Note that this is with white noise corrupting the outputs to 25% the level of the outputs.

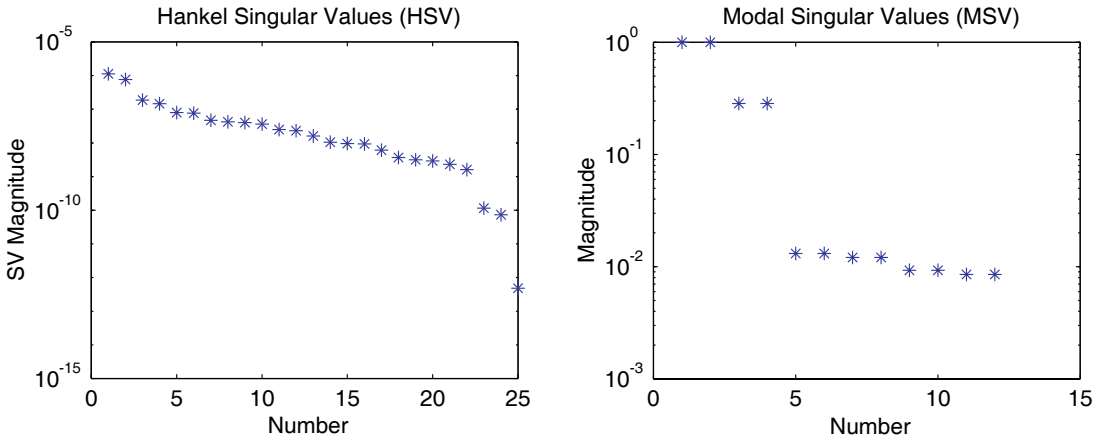

Figure 6-7 The Hankel and Modal singular values for the complex example.The break between the correct order and the noise is much less noticeable. Again, the Modal singular values allow the final system order choice to be made.

The complex system is much more challenging for the OKID algorithm to identify than the simple system. This is due to the lightly damped mode that only shows up as a disturbance on the position of the primary mass. Due to the reduced mass, the motions of the second mass are about the same size as the noise introduced into the sensor. Figure 6-7 shows the Hankel and Modal singular values of the OKID process for the complex example. Here the break in the Hankel modes is much more difficult to see. The Modal singular values show that most of the power is concentrated in the first four modes. It is therefore

safe to use just those modes. Once again, the OKID algorithm at this point generates the model and the Kalman filter. Table 6-2 shows the identified version of the complex system versus the true system, and we can see that the frequencies are excellent, and that the damping ratios are very good as well.

Eigenvalue	Magnitude	Equiv. Damping	Equiv. Freq. (rad/s)
0.973+0.200i	0.994	0.00315	2.03
0.973-0.200i	0.994	0.00315	2.03
0.931+0.338i	0.991	0.00265	3.48
0.931-0.338i	0.991	0.00265	3.48
0.971+0.202i	0.991	0.00417	2.05
0.971-0.202i	0.991	0.00417	2.05
0.917+0.340i	0.978	0.00621	3.55
0.917-0.340i	0.978	0.00621	3.55

Table 6-2: The true and identified system poles of the complex example. The true system is shown in green, and the OKID system is shown in blue. Note that the magnitude and damping of identified system is quite close to the true system. Note that this is with white noise corrupting the outputs to 25% the level of the outputs.

The objective of the system identification exercise is that the predicted and actual outputs be equal. To this end, another set of random input data was used to excite the original example systems, and the same input data was used to excite the identified system. Figure 6-8 shows the comparison of the output of the true versus the identified systems. The lower frame in the figure shows the error in the output between the identified and true systems. Note that the errors are within about 2% of the actual output. This is an excellent match for a fairly short data stream with noisy data.

SECTION 6.2 KINEMATIC MODEL

In order to control the Atlantis, a system model needs to be assembled. While several good modelling techniques exist to model a powered boat travelling through water

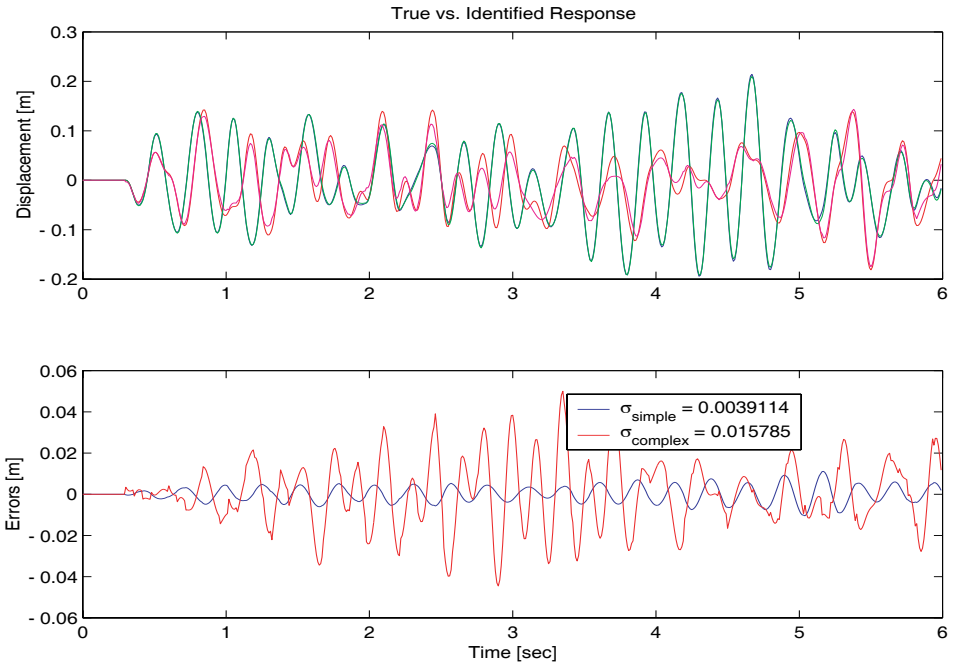

Figure 6-8 Comparison of the output of the true and identified systems. The bottom plot is the errors in the two models (simple and complex) and show very good agreement especially when considering the amount of noise injected into the sensor measurements.

[51], they remain complicated and not easily analyzed. Furthermore, the situation is much more complicated for catamarans. Previous work in England modeling the drag on the hulls has been done and the thrust equations can be based on several parameters of the sails, hulls, and wind effects. These efforts produce the equations of motion for light wind conditions and are detailed in [28]. In order to reduce the model order and obtain a model that would have sufficient fidelity for active control, several different methodologies are attempted.

In order to formulate the equations of motion, the Atlantis is assumed to be travelling upon a straight line, conveniently assumed to be coincident with the x-axis, through water at a constant velocity, V_x. The distance along that line is *X*. The perpendicular dis-

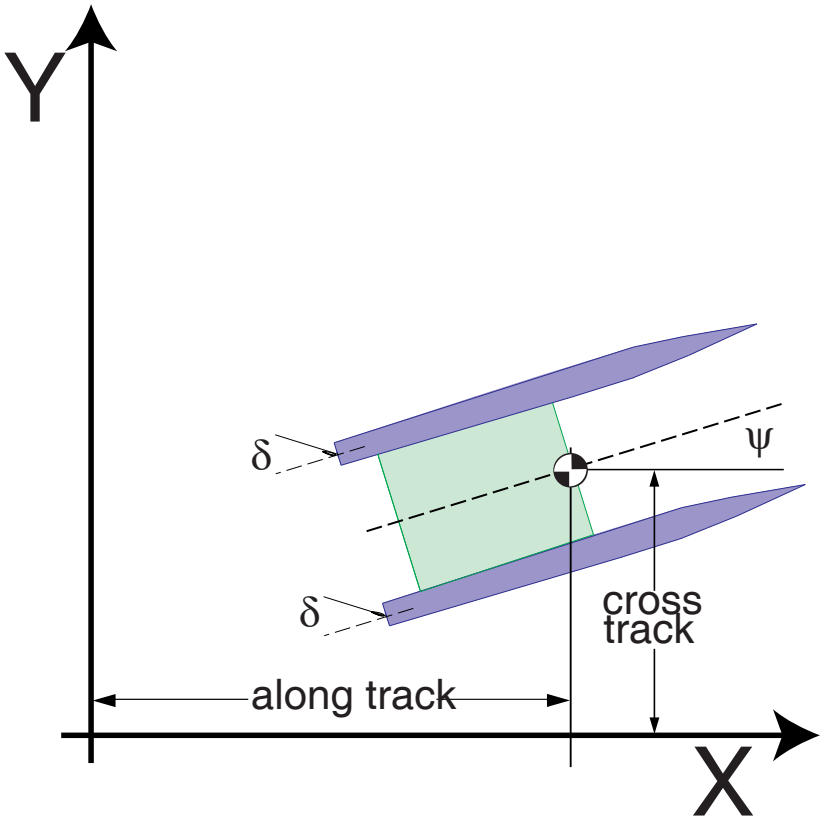

Figure 6-9 Basic diagram of the path of the Atlantis for control system formulation. The Atlantis is assumed to be travelling along a straight line coincident with the x-axis at a constant velocity, with a cross-track error measured as "y," and an azimuth error measured as ψ.

tance to the line is *Y*, the cross-track error, and the angle that the centerline of the Atlantis makes with the x-axis is Ψ, the angular error. Figure 6-9 illustrates the mathematical model of the assumed path of the Atlantis.

The coordinate frame can always be rotated to have the x-axis aligned to the desired path of the Atlantis, and so the assumption that the Atlantis travels down the x-axis is a good one. The assumption of constant velocity, however, is not appropriate since velocity is a function of the wind speed. Wind speed, of course, cannot be controlled and is highly variable.

The first model is a simple kinematic model that assumes that the rudders cannot move sideways through water. This places a kinematic constraint upon the motion of the entire boat, and the linearized analysis produces the following continuous time state-space equations:

$$\begin{bmatrix} \dot{Y} \\ \dot{\psi} \\ \dot{\delta} \end{bmatrix} = \begin{bmatrix} 0 & V_x & 0 \\ 0 & 0 & \frac{V_x}{L} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Y \\ \psi \\ \delta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \qquad \text{(EQ 6.24)}$$

where Y is the cross-track error (in meters), Ψ is the azimuth error (in radians), and δ is the angle of the rudders with respect to the hull centerline (in radians). The distance, L, is from the boat center of mass to the center of pressure of the rudders (in meters), and the input, u, is the slew rate of the rudders (in radians/second).

These simplified equations of motion are insufficient to control the boat to great precision, but are excellent for generating intuition for the system identification process. Equation 6.24, when cast into transfer function form, becomes a triple integrator. As such, it cannot be stabilized by simple proportional control. In addition, the assumption of constant V_x is poor, since unless the wind can be controlled, the velocity will always be dependent on the speed of the wind. Closer inspection of Equation 6.24 shows that the errors in azimuth and cross-track integrate not with time, but rather with distance travelled forward. What this means is that if the boat is sitting still in the water, no amount of rudder deflection will cause the azimuth to change; likewise, when moving very quickly through the water, only very small inputs are required to turn the boat through a considerable angle.

The simple kinematic model is based on prior work on the GPS guided farm tractor done at Stanford University [45]. In this work, it was shown that even with the poor model, the tractor is able to perform the line-tracking task with very high precision. Furthermore, using the extended Kalman filter to perform parameter identification, values for

the critical parameters of V_x and L can be obtained through relatively simple tasks and done in an "on-the-fly" estimation fashion. Extensions of this work presented in [116] show that "on-the-fly" estimation can be greatly simplified by using an extension of the Least Mean Squares (LMS) algorithm to estimate the parameters in real-time. This technique has been validated by experimental data.

Using the kinematic model, a linear quadratic regulator (LQR) controller was designed using the full state feedback of the measured state of azimuth, rudder angle, and cross-track error. Bryson's rule was used to formulate the Q and R matrices, and no cross coupling terms were assumed. Section 6.5 on page 266 describes the entire control design in more detail. The controller was designed such that a maximum rudder deflection rate (approximately 25 degrees/second) would match the maximum cross-track error of 0.1 meters. An entire series of these controllers in which Y_{max}, the maximum cross-track error, was varied from 0.01 meters to 1.0 meters have been designed and tested,

These controllers are tested with the trolling motor and results found to be satisfactory as long as the speed, V_x, remains below $V_x|_{design}$, the design speed of the controller. As soon as $V_x > V_x|_{design}$, the Atlantis started to hunt, eventually going unstable. When V_x was very substantially less than the design velocity, $V_x|_{design}$, the response was sluggish and suffered from large (though bounded) cross-track errors. An integrator state was added, but this only further destabilized the system. As the problem was not one of steady-state error, the integral control did not improve the situation.

Figure 6-10 shows the root locus of the kinematic model with the full state feedback LQR controller as a function of V_x (the along-track velocity). The design velocity, $V_x|_{design}$, is 1.0 meter/second and the model is assumed perfect and noiseless. As the along-track velocity, V_x, is increased, the damping is reduced, until, at V_x = 4.1 meters/second, the system is unstable. This is the best case scenario. When other errors are introduced,

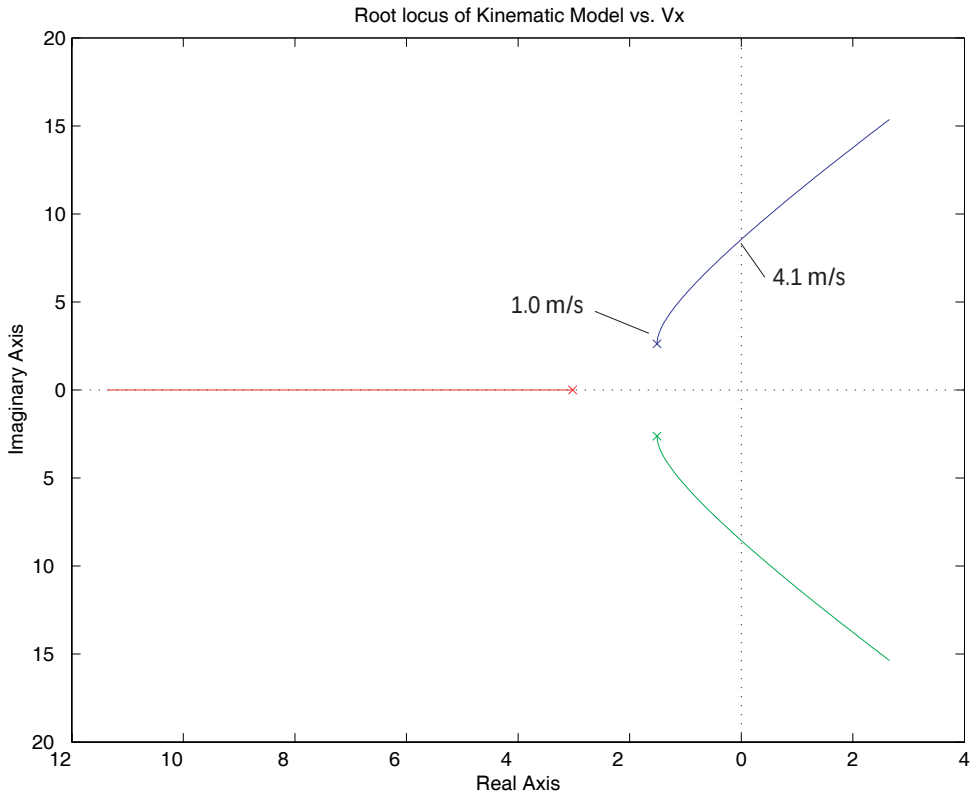

Figure 6-10 Root locus plot of the closed loop kinematic model using an LQR controller as a function of the along-track velocity. Note that the design along-track velocity is 1.0 meters/second, and that the closed loop system goes unstable at an along-track velocity of 4.1 meters/second due to the increased controller authority with increasing along-track velocity.

such as mis-modelling, external disturbances, noise on the sensors, etc., the results are predictably worse. This is evidenced by the poor performance of the kinematic controller in all of the off design points.

SECTION 6.3 VELOCITY INVARIANCE

In the previous section, it was seen that the simple kinematic model suffers greatly as soon as V_x increases above the design point, $V_x|_{design}$. In order to address this shortcoming, an attempt was made to formulate a control algorithm that automatically compensates

for the varying velocity. Several methods for varying the control gains as a function of velocity are possible.

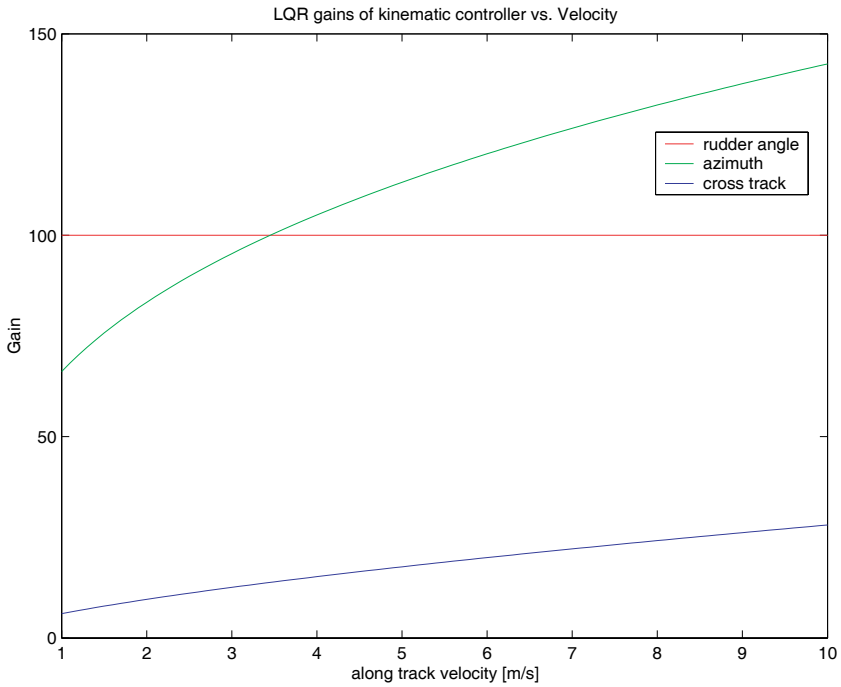

Figure 6-11 LQR gains for kinematic model versus velocity. The rudder angle gains appear to be constant, but the azimuth and cross-track gains are both functions of velocity. This indicates that some form of gain scheduling (as a function of along-track velocity) could be implemented. In this project, a different approach to varying the gains with velocity has been implemented—that of velocity invariance.

Figure 6-11 shows the gains for the LQR controller as a function of V_x. The observed gains on the rudder angle appear to be independent of velocity, but the gains on the azimuth and the cross-track error are strong functions of velocity. The gains are not simple linear or even exponential relationships to velocity, but a gain-scheduled controller could be made to interpolate the gains based on the measured velocity at the current moment.

Inspecting the equations of motion for the simple kinematic model, the relationship:

$$\begin{bmatrix} \dot{Y} \\ \dot{\psi} \\ \dot{\delta} \end{bmatrix} = \begin{bmatrix} 0 & V_x & 0 \\ 0 & 0 & \frac{V_x}{L} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Y \\ \psi \\ \delta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \qquad \text{(EQ 6.25)}$$

which, upon closer inspection, demonstrates that the azimuth and cross-track error do not integrate with time, but rather with distance travelled upon the line. This has great implications, because this is exactly the cause of the instability with increasing velocity present in the simple kinematic model. By introducing two new variables:

$$\tilde{y} \equiv \frac{y}{V_x} \qquad \text{(EQ 6.26)}$$

and

$$\tilde{\psi} \equiv \frac{\psi}{V_x} \qquad \text{(EQ 6.27)}$$

and substituting them back into Equation 6.25, the equations of motion of the kinematic model can be rewritten in the following velocity invariant form:

$$\begin{bmatrix} \dot{\tilde{Y}} \\ \dot{\tilde{\psi}} \\ \dot{\delta} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & \frac{1}{L} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{Y} \\ \tilde{\psi} \\ \delta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \qquad \text{(EQ 6.28)}$$

The state transition matrix no longer has any terms relating to V_x. This formulation allows one single controller to be designed with the sensor input to that controller is scaled by the velocity before the control gains are applied. This has the effect of automatically reducing the sensitivity at high forward velocities and increasing the sensitivity at low velocities. In practice, a lower bound of 1.0 meter/second is used on the velocity measurement, due to the presence of noise.

This velocity invariant formulation was used with great success on the GPS guided farm tractor [45]. A similar technique is used on the Atlantis. This methodology of velocity invariance is fine for the simple kinematic model and addresses the concerns of the instability with increasing velocity. Functionally, it means that a single controller is designed, and that the inputs are scaled before injection into the controller. It does not, however, address any of the issues of mis-modelling or sensor noise. In practice, with the trolling motor, the velocity invariant kinematic model performs well in calm waters but starts to oscillate badly in waves. Furthermore, the controller performance (as measured by y_{max}) has to be relaxed in order for the closed loop system to remain stable. This is, most likely, an artifact of mis-modelling.

Using the input pre-scaling, however, it is possible to normalize any set of input data by the velocity and thus identify the velocity invariant system. Thus, the data collected for system identification is pre-scaled by velocity (with a lower bound of 1.0 meter/second) before applying the system identification techniques. This allows the best velocity invariant model to be identified by the algorithm and contributes to the robustness of the controller.

SECTION 6.4 ATLANTIS SYSTEM IDENTIFICATION

In order to use the OKID method to generate the system model high quality data is required: that means data with the greatest power into the modes to persistently excite the system along with the cleanest sensor data available. Due to the time constraints of the project, the wing was under construction when the off-line identification data was required. Without the wing, little choice was left but to use a trolling motor as the primary source of propulsion of the Atlantis, and to attempt to mimic the wind velocity and direction by rotating the axis of the trolling motor and increasing the voltage from 12 volts to

24 and 36 volts. The Atlantis was ballasted with an additional 75 kilograms of weight (in the form of additional batteries) equivalent to the weight of the wing. All the changes were done so that the data collected on the identification passes would be representative of the dynamics of wing propulsion.

Figure 6-12 The Atlantis on the water with the trolling motor and extra batteries for ballast. This was the configuration of the Atlantis when taking data for OKID identification.

In order to gather data to perform a proper system identification of the Atlantis, a series of open-loop line-following tests were conducted in which a human driver, through the GNC computer, caused the rudders to either slew left or right at the maximum slew rate (~25 degrees/sec.). Also, the driver commanded the rudder slew rate to zero through the rudder actuator in order to track a roughly straight line. This “pseudo”-random input was designed to apply the maximum power to the Atlantis through the controls and produce a rich output that would contain information from all modes of interest.

Several interesting events occurred during the data gathering passes. Once the trolling motor was switched on before Atlantis was placed in the water. Unfortunately, the

impeller was jammed against the dolphin striker and the trolling motor began to overheat. By the time Atlantis was placed in the water, the current through the trolling motor had welded the armature to the housing.

On another occasion in an attempt to gather data while travelling backwards, the trolling motor was turned 180 degrees on its mount. This resulted in the motor mount detaching from Atlantis and the trolling motor submerging into the San Francisco Bay. Luckily, the wires holding the trolling motor to the battery were very stout and the motor was recovered. However, the salt water immersion caused the power circuits to short out rendering the motor unusable. Data taking on that day ended prematurely while the captain and crew rowed back to shore.

Luckily, the Atlantis project was left with one good trolling motor and one good controller board. However, attempts to use more than 12 volts to increase Atlantis' speed through the water resulted in the motor overheating and the controller power mosFETs being destroyed. After once again rowing back to shore, it was concluded that the controller boards were not needed as Atlantis was run only at full throttle or with the trolling motor stopped.

After directly connecting the batteries to the motor through a high capacity switch, more attempts were made to increase the velocity range of the identification passes. It was discovered that in this configuration, the trolling motor could operate at 24 volts, but that at 36 volts, the trolling motor very quickly overheated and lost power. Fortunately, the motor recovered and it was not necessary to once again paddle Atlantis back to shore.

SECTION 6.4.1 TYPICAL IDENTIFICATION PASS

Every attempt was made to fully exercise the dynamics of the Atlantis in order that the system identification algorithm could model the full range of expected responses. Typ-

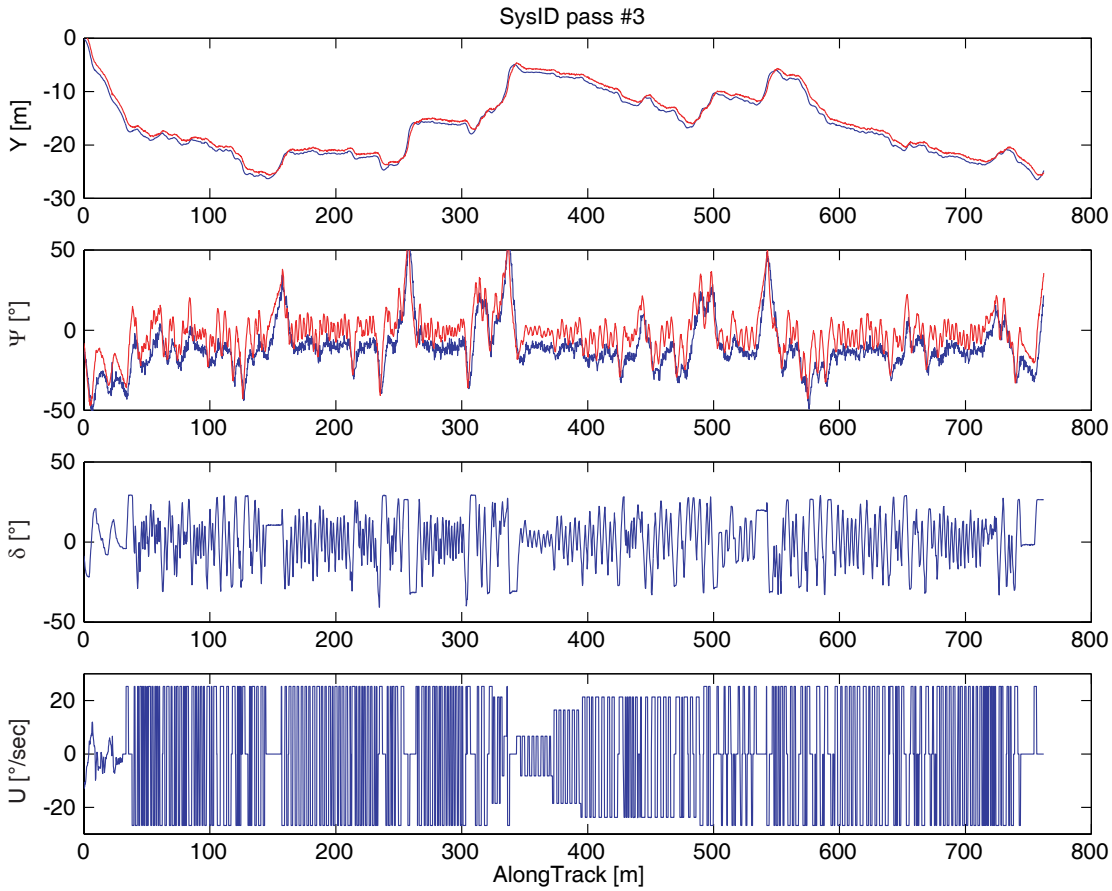

Figure 6-13 Typical system identification pass of the Atlantis using the trolling motor for propulsion and a human pilot to inject a pseudo-random sequence of maximum slew rate into the rudders.

ically, Atlantis lined up at one end of the harbour and started on a straight pass towards a marker at the other end of the harbour. During the pass, the rudders were continually slewed to port and starboard at their maximum rate. A nominally straight line was intended by the human pilot, but excursions in cross-track of several meters did occur.

Eventually, the procedure for gathering the identification data was perfected and a large number of passes were aggregated at different velocities in order to perform the final system identification. Figure 6-13 shows a typical pass, with the rudder slew rate on the bottom, and the rudder angle, azimuth angle, and cross-track deviation above. This pass

was over 700 meters long and had azimuth excursions as high as 45 degrees from the intended path.

SECTION 6.4.2 OKID RESULTS

The system identification experiments resulted in several large electronic files that contain the data from the pseudo-random input and all of the sensor outputs. These files were merged together and the data normalized by the velocity as discussed in Section 6.3 on page 256. Both the velocity-normalized and the non-normalized data were used in the identification process, but the velocity-normalized data produced much better controllers, and is therefore presented in this work.

The first thing required in the OKID method is to choose an upper bound on the possible system order. This is done to reduce the computation time and also to define the blocks of input and output that will be used. For the Atlantis, the maximum system order, p, was chosen to be 80. While it is conceivable that the true system order could be greater than 80 modes (indeed, most real systems have infinite modes), as an engineering approximation, 80 is more than adequate.

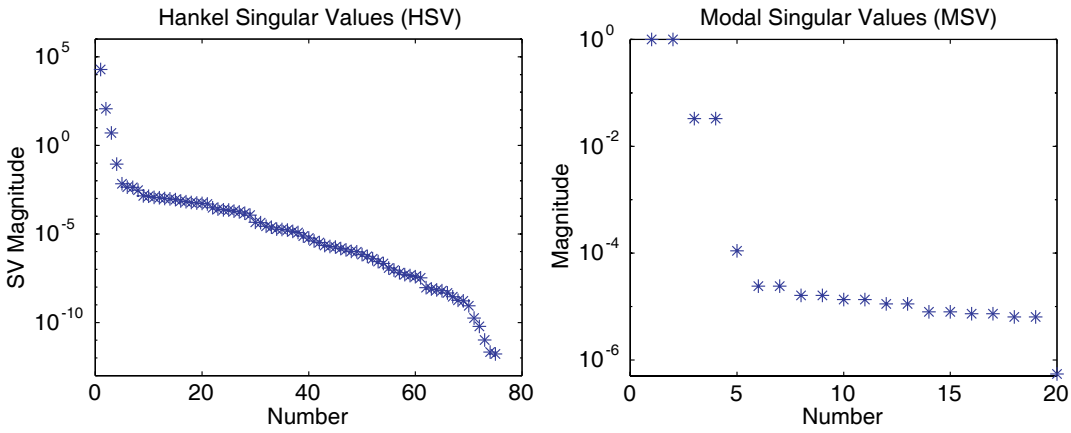

Figure 6-14 Hankel singular values and Modal singular values for the system identification of the Atlantis catamaran. Note the large drop off in the singular values after the fourth one, indicating a system order of four.

Figure 6-14 shows the Hankel singular values and the Modal singular values of the OKID process on the Atlantis. On the real system, the transitions between the real modes and the noise modes are much smoother. The first four singular values of the Hankel matrix do stand out, and this is confirmed by the Modal singular values which demonstrate that even if the higher modes are used, they are virtually indistinguishable from the noise.

The Atlantis, therefore, is assumed to be a fourth order system. Since the kinematic model is a third order system, some speculation is given to what the physical interpretation for the fourth mode should be. It would seem that the fourth mode is a lag term between the actuation of the rudder and the time that the Atlantis begins to rotate about her center of mass. In truth, this is difficult to decipher because the balanced realization form of the state-space model gives no physical intuition. A similarity transformation based on the pseudo-inverse of the output matrix is used to try to discern exactly what the states are in terms of the outputs. The similarity transformation results in the output matrix, C, being a identity matrix. If the number of states are equal to the number of outputs, this maps the states directly into the outputs. Unfortunately, in this case the number of states is greater than the number of outputs, thus the information is not very useful.

SECTION 6.4.3 DATA RECONSTRUCTION

The fourth order model performed well as shown by the matching the output not only of the cross-track error, but simultaneously of the azimuth error and of the rudder angle as well. This, of course, is because the OKID method is perfectly suited to multiple input multiple output (MIMO) systems as well as the traditional single input single output (SISO) systems. By definition, the model that it generates will match the outputs, all of the outputs, in a least squared sense.

Figure 6-15 shows that the model generated was also quite adequate at predicting Atlantis' response to a known input. This particular data was from a system identification

pass which had not been used in the system identification process. This data was reserved to see exactly how well the OKID method would be able to track the output of the actual system. Perfect tracking may not, in fact, be desirable. Consider the case where a sensor has some very high frequency noise on its output: generating a model that tracks that perfectly is not helpful from a control standpoint. Although the high frequency sensor noise might be observable, it is certainly not controllable.

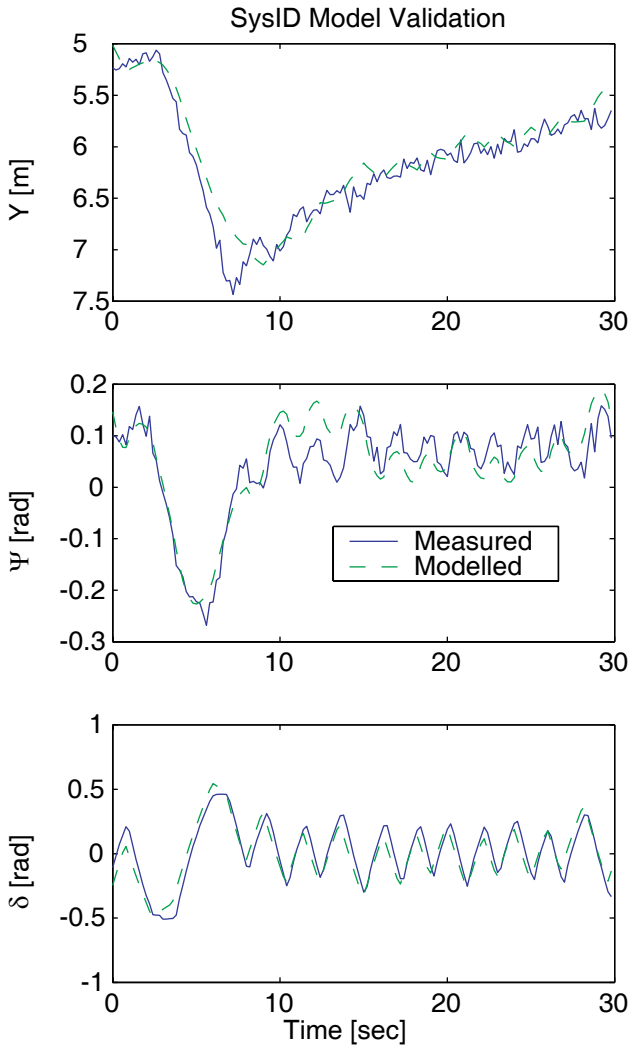

Figure 6-15 The OKID data reconstruction of a system identification pass outputs from the identified model. As can be seen, the agreement is generally very good, with the model tracking the actual outputs.

Figure 6-15 shows the reconstruction of the Atlantis system identification data for approximately 30 seconds. The actual trace goes on much longer, but the detail is lost as the horizontal axis is stretched out. Obviously, with the addition of the observer, the agreement with the data will be much greater. This is due to the filtering of the noisy sensors, but also because the Atlantis itself is a fairly slow moving object. The sampling rate of the control system is such that the motion of the catamaran from one sample time to the next (200 milliseconds) is quite small.

Figure 6-16 shows the errors from the data in Figure 6-15. As can be seen, the errors are quite small. Standard deviations are approximately 0.1 meters for the cross-track mis-modelling error, 1.5 degrees in azimuth, and 3 or 4 degrees in rudder angle mis-modelling error. Once again, this is the open loop model of the catamaran. These errors will become much smaller once the Kalman filter is added to the loop (since the error will then be adjusted to use data from the previous outputs).

It is interesting to note that any identified model will be the controllable and observable subset of the true mode. This is due to the rather simple fact that if it cannot be measured, then it cannot be seen by the algorithm. That is, if the sensor cannot sense (or observe) some state, then as far as the identification algorithm is concerned, it does not exist. Likewise, the same is true of controllability. If the input cannot push on a certain part or mode of the system, then the identification algorithm is blind to that.

SECTION 6.5 ATLANTIS CONTROLLER DESIGN

Based on the previous sections of this chapter, the Atlantis has been identified to the extent that a controllable and observable model capable of predicting the short term behavior of the Atlantis (while under trolling motor propulsion) was developed. In addi-

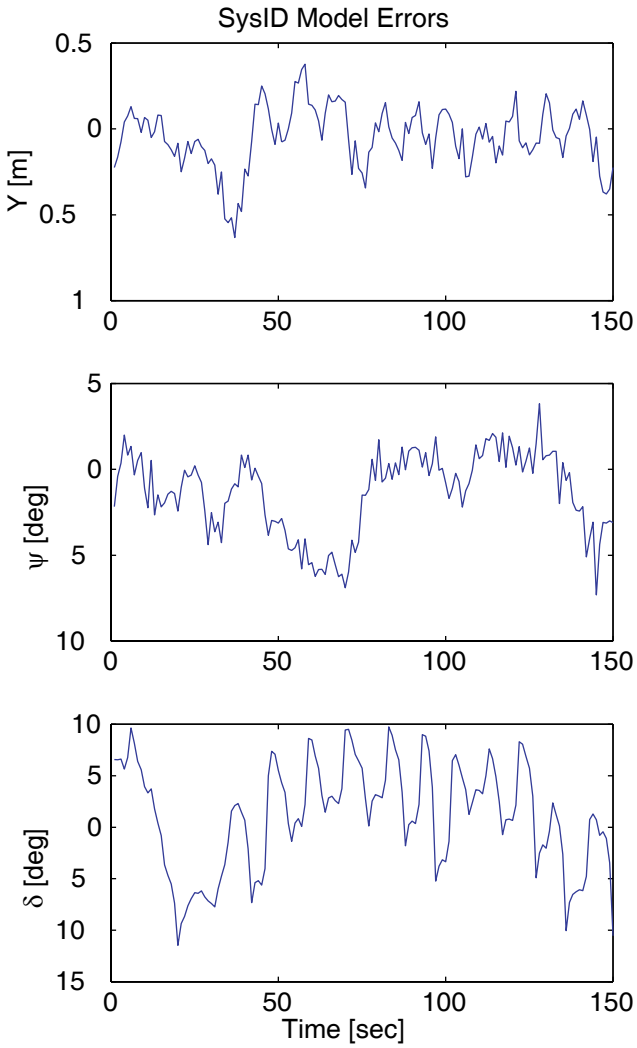

Figure 6-16 OKID data reconstruction errors. This shows the differences between the actual and modelled responses on the previous plot. The standard deviation for the cross-track modelling error is less than 0.1 meters. The standard deviation for the azimuth modelling error is approximately 1.5 degrees, and the standard deviation for the rudder angle modelling error is roughly 3 degrees.

tion, the pseudo-Kalman filter that is the best state estimator has also been obtained from the OKID method.

$$J = \sum_{k=0}^{\infty} \left(\tilde{x}_k^T C^T \begin{bmatrix} \frac{1}{y_{max}^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} C\tilde{x}_k + \tilde{u}_k^T \left[\frac{1}{u_{max}^2} \right] \tilde{u}_k \right) \qquad \text{(EQ 6.29)}$$

With these steps complete, the last step in the process is to design a regulator or controller and close the loop. The control method used throughout this project is the standard Linear Quadratic Regulator (LQR) controller coupled with the identified state estimator that resulted from the OKID algorithm used for identification. The LQR controller is one that minimizes the cost function, J, in Equation 6.29, with the two terms of y_{max} and u_{max} being design parameters. The general controller methodology is to treat this essentially as an output regulator with the cross-track error as the only output of concern.

Equation 6.29 is set up by Bryson's rule to penalize the control at the same time as penalizing cross-track error. There are several excellent techniques to generate the gain matrix, K, such that J is minimized. See, for instance [133], for a very complete coverage of the process. Once the gains are computed, a file is generated that is uploaded to the Atlantis guidance navigation and control computer. A family of controllers are designed at once and tested experimentally to determine which is the best. Figure 6-17 shows the pole-zero map for the open and closed loop kinematic model as well as the open and closed loop identified model. The estimator is not shown in the plot because its poles are much faster than the controller.

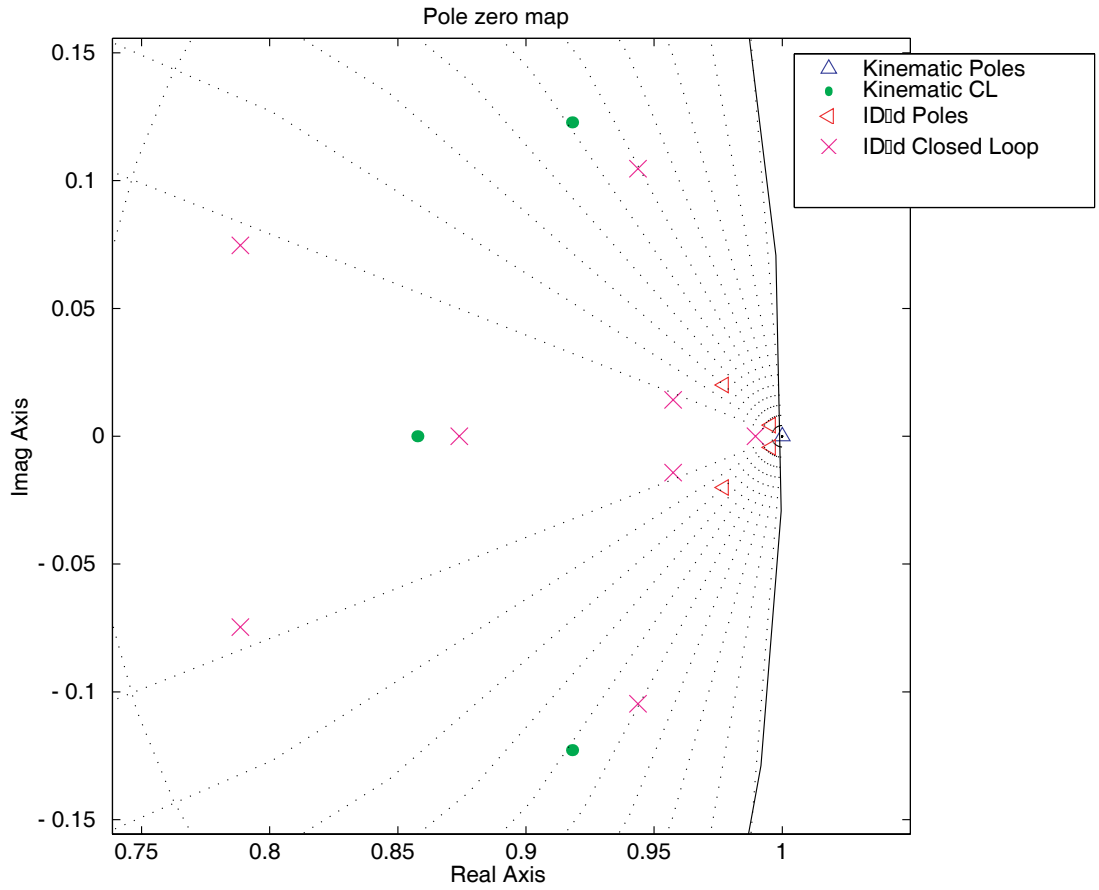

Figure 6-17 Pole-Zero map of the closed loop identified system, as well as the kinematic open and closed loop poles. This shows a section of the unit circle typical discrete time control systems. The grid lines represent constant damping ratios. The large number of fast poles on the identified system is the identified Kalman filter.

SECTION 6.6 CAVEATS AND CONCLUSIONS

Within the system identification and control design issues is a rather large caveat. This is the assumption that the Atlantis powered by the trolling motor is sufficiently similar to the Atlantis while powered by the wing sail such that a control system designed for one will work with the other.

There was, in fact, little basis for this assumption.

There simply was not enough time to wait for the wing to be finished before designing the control system and performing the open loop system identification tasks. Nor was there enough time or funds for the project to delay while the very last bit of performance could be eked out of the system through the exploration of other control topologies. Instead, it was thought that the best strategy would be one that got the Atlantis navigating on its own as soon as possible, even without knowing if there might be more sophisticated control design techniques. In addition, the trolling configuration was deemed much less probable to capsize the catamaran or injure the crew, resulting in a slightly lower risk proposal to develop and test the controllers. However being on a mesh trampoline with four high amperage batteries wired in series while traversing salt water may not be considered “low-risk.” At a minimum, the non-control issues such as sign errors in the computer code that performed the actual control could be dealt with without destroying expensive and irreplaceable hardware.

But most importantly, the assumption proved valid and the controller design worked.

As will be detailed in the next section, the dynamics of the wing sail are so benign that the catamaran hardly heeled. Thus the dynamics were quite well matched from the trolling motor to the sailing tasks. Interestingly enough, the trolling motor was never able to bring the Atlantis up to the speeds at which she sailed. What is remarkable about this is that the identified model/estimator/regulator combination proved sufficiently robust that it performed very well completely outside the range of parameters over which it had been designed.

7 Experimental Results

The goal of the Atlantis project has been, since its inception, to demonstrate precision guidance on a full scale prototype in order to validate the entire unmanned sailing concept. With the completion of the system identification and control design as presented in Chapter 6, the next step in the process was to put the Atlantis in the water and measure the results of the control.

Tests were run in three separate experimental trials. The first trial was performed using the original sails with the differential GPS system providing position and velocity measurements at a rate of once per second. During this trial, the catamaran was piloted by the author and the wind speed was estimated to be approximately 6

meters/second. This was a coarse estimation based on the weather data from buoys in the bay. The purpose of this trial was simply to set a baseline against which to compare the later data acquired from the automatic control experiments. Anecdotally, the catamaran proved to be very difficult to sail on a straight line, and attempts to do so resulted in large excursions in roll. This culminated in a capsize and the loss of the computer recording the data due an overzealous attempt to jibe. The data was, however, recovered, though the laptop was unrepairable.

The second and third trials were performed using automatic control with the trolling motor and the wing sail, respectively. The trolling motor tests were used to validate the control system in advance of the wing manufacture and to test a variety of controllers under experimental conditions.

SECTION 7.1 TACKING, JIBING, AND SELF-TRIMMING

Central to the operation of the Atlantis is the wing sail propulsion system. While the details of design have been discussed in detail in Chapter 5, the general operation of the propulsion system has not been explained. The wing sail is suspended on the stub mast by two bearings that allow the entire wing/tail system to rotate freely in azimuth about the stub mast. Due to the presence of a set of slip rings, there are no wires to wind up, and because the aluminum support spider is below the lower bearing, the wing/tail system can rotate without interference from any other system on the Atlantis.

The operation of the wing is deceptively simple. While the tail has no incidence to the wind (no angle of attack), the wing sail acts as a giant weather-vane, always pointing into the relative wind. Note that the relative wind is the wind that is the vector sum of the

true wind and the Atlantis' velocity; that is, the relative motion of the air mass over the wing sail.

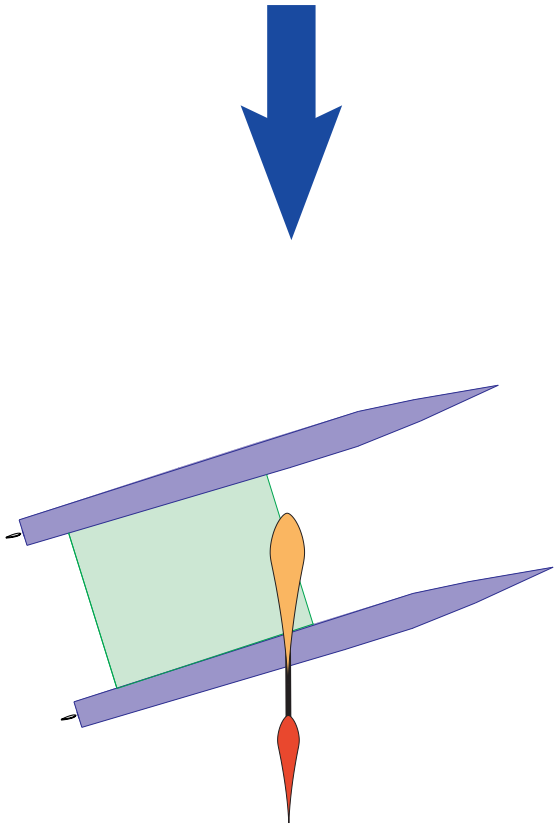

Figure 7-1 The Atlantis sitting still on the water, with the tail set at zero angle of attack. Note that the wing/tail system is acting as a large weather-vane and developing no thrust.

Figure 7-1 shows the Atlantis sitting still on the water with the tail set to zero angle of attack. At this point, the Atlantis will simply drift with the current as the wing is generating no lift and hence no thrust. To get the Atlantis moving, the tail needs to be rotated counter clockwise to give it a negative angle of attack. This in turn will cause the entire wing/tail system to rotate clockwise until a moment balance is achieved between the main wing and the tail. This moment balance will cause the wing to be flying at a positive angle of attack and it will in turn generate lift. The component of lift perpendicular to the hulls will be opposed by the force of the water on the centerboards, preventing the Atlantis from

moving sideways through water. The component of lift pointing along the hulls is the thrust and will accelerate the Atlantis until the drag on the hulls equals the thrust from the wing, at which point the Atlantis will continue to move at a constant velocity. This is shown in Figure 7-2.

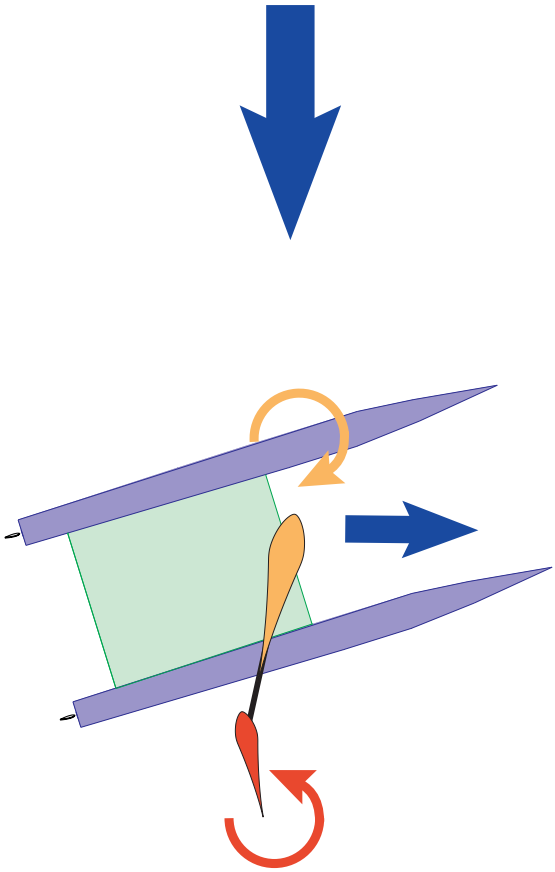

Figure 7-2 The Atlantis after the tail is deflected causing the wing/tail system to rotate into equilibrium with the main wing generating lift in a direction perpendicular to the relative wind. The tail is deflected such that it is at a negative angle of attack. This, in turn, swings the wing/tail system clockwise and causes the main wing to be at a positive angle of attack. With the wing now generating lift, the Atlantis accelerates in the direction of its hulls until the drag on the hulls is in equilibrium with the lift on the wing.

Once the Atlantis is moving at a constant velocity, the wing and tail are in moment equilibrium. The wing lift along the hulls is equal to the hull drag, and the wing lift perpendicular to the hulls is opposed by the centerboards in the water. The centerboards do

fly through the water at a small angle of attack to generate this lift. This angle is usually very small due to the very high density of water (and is referred to as leeway angle). If, however, the wind should suddenly change either in strength or in direction, the wing sail will have to move in order to reach a new equilibrium. Figure 7-3 shows the Atlantis under way.

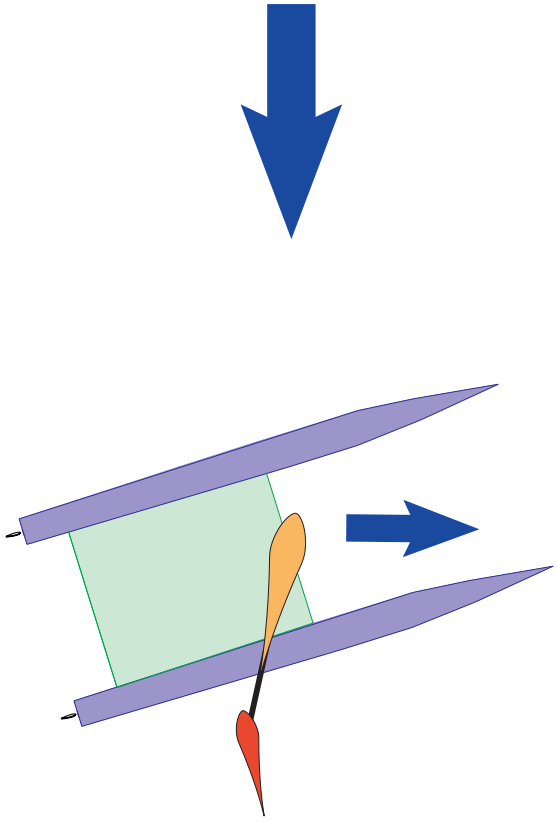

Figure 7-3 The Atlantis underway, with the forces in balance and the wing sail and tail in equilibrium. At this point, any change in wind speed or direction will cause the wingsail to rotate to a new equilibrium position.

After the wind changes direction, the wing sail and tail will have rotated to a new position that is identical relative to the wind. Since the lift can only be generated perpendicular to the relative wind, the lift vector will also rotate to the same position relative to the wind. Thus, the entire wing sail, tail, and lift vector rotate together with the wind as a

rigid unit. This is actually a simplification; in truth there will be a transient behavior before the steady state is reached. However, the transient response is fast and well-damped such that the steady state view is sufficient. Figure 7-4 shows the Atlantis after the wind has changed direction by moving around towards the stern of the catamaran. As can be seen, the wing sail and tail have rotated to the new position as has the lift vector. Thus, no active change is required to re-trim the sail, hence the self-trimming description.

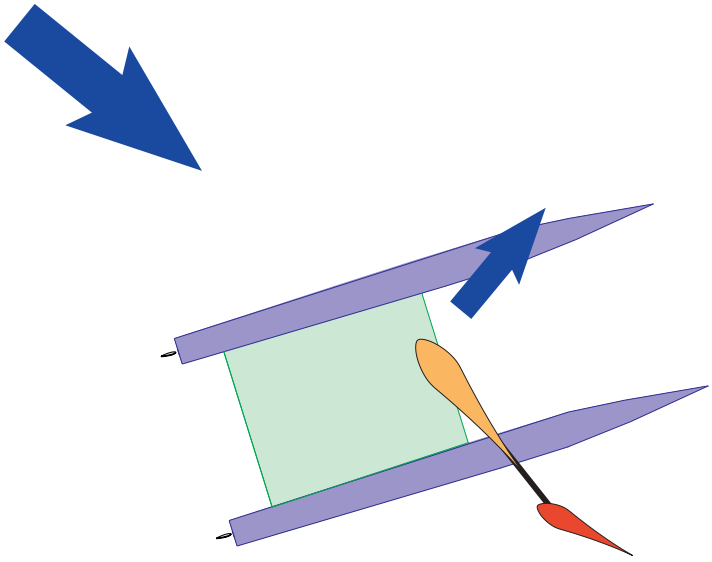

Figure 7-4 The Atlantis after a wind direction change. Due to the self-trimming nature of the wing/ tail system, the wind-wing-tail-lift all move together as a rigid unit.

This is only true if the resulting wind is still on the same half of the catamaran (i.e., from the port or the starboard side). If the new wind has crossed the centerline of the catamaran from the previous wind, then the lift is now in the opposite direction, and the thrust is decelerating the boat and trying to send the catamaran in reverse. Therefore, the wing needs to be tacked or jibed depending on if the wind has crossed the bow or the stern of the catamaran.

By definition, a tack is a sailing maneuver in which the bow of the boat is moved through the wind and the sails are changed to the other side. This allows progress upwind

to be made, as a sailboat can only sail upwind when canted off to one side or the other. Thus tacking allows the sailboat to zig zag across the wind and still make progress into the wind. This is true of all sailing vessels (with the exception of very unusual boats that use air turbines connected to underwater propellers and screw themselves directly upwind). Note that a conventionally sailed sailboat experiences a large aeroelastic instability as the sails enter a regime of low angle of attack. The jib and the mainsail luff, or collapse and rattle around until the tack is completed and they are once again at a fairly high angle of attack.

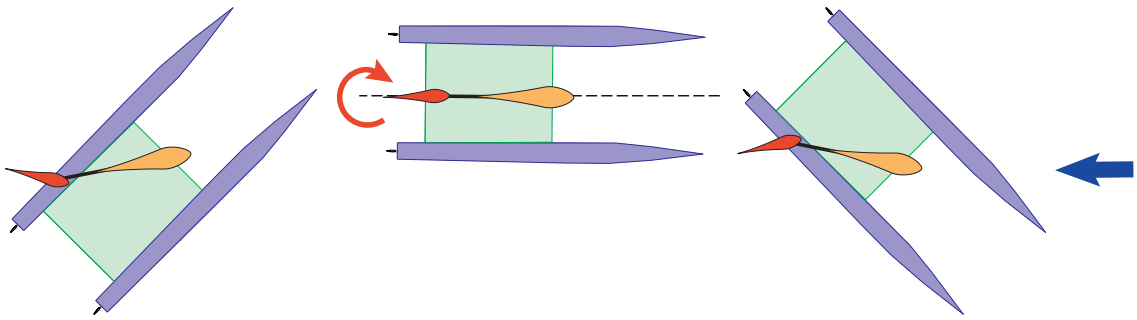

Figure 7-5 The sequence of tacking the Atlantis. The wind is blowing from the right side of the page and the tail is centered as the hulls point into the wind and then is reversed to the mirror image of its previous setting to complete the tack. Tacked in time from start to end, the progress of the tack is sequentially from left to right.

Figure 7-5 above shows the Atlantis tacking from the port side to the starboard side. As the catamaran heads into the tack, the tail is centered to coincide with the hulls pointing directly into the wind. As the momentum of the hulls carries the catamaran through the tack, the tail is reversed to a mirror image position from where it was before and the catamaran accelerates off on the new course. Due to the rigid skin of the wing sail, there will not be an aeroelastic collapse of the wing. In fact, the tack will be very quiet, with none of the usual flapping or large change in angle of heel as the boat is swung through the wind.

The jibe is the exact same maneuver as the tack, but with the hulls pointing away from the wind rather than pointing into the wind. On a conventional sailboat, a jibe results in the main sail swinging violently from one side to the other as the wind catches the back of the sail and forces the boom across. In the case of the Atlantis, the jibe is just as gentle as the tack, with no sudden motions anywhere. It is, in fact, the exact same as Figure 7-5, but with the hulls pointed the other way. Figure 7-6 shows the jibe occurring with the wind blowing from the left of the page.

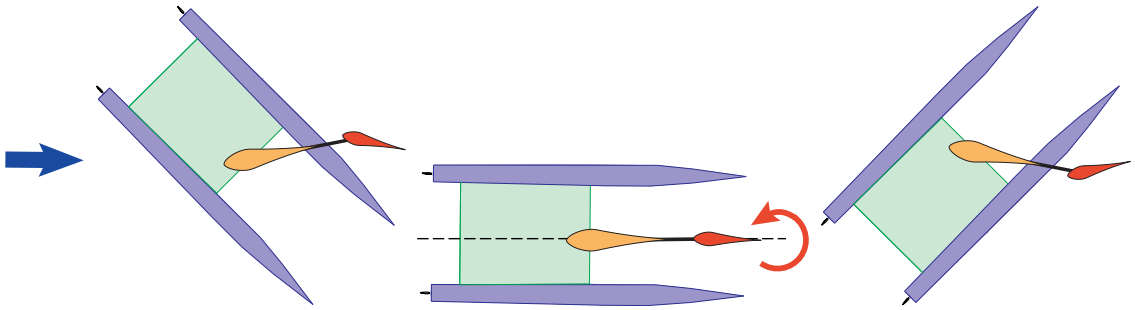

Figure 7-6 The sequence of jibing the Atlantis. The wind is blowing from the left of the page and the tail is centered as the hulls point away from the wind and then is reversed to the mirror image of its previous setting to complete the jibe. Again, note that temporally, the sequence is from left to right.

During the several times that tacks and jibes were performed, the Atlantis remained gentle and fully upright. There never were any occurrences of fast and uncontrolled motions which are so typical of conventionally sailed boats during tack or jibe.

SECTION 7.2 EXPERIMENTAL METHODOLOGY

The basic method of validating the control system performance is to have the Atlantis sail or troll along a straight line. Deviations are measured from that line, along with other parameters. The line is defined from the current waypoint and projected forward along the surface of the water in the desired direction of travel. The latitude and lon-

gitude data are transformed into a local north-east-down coordinate system. This data is, in turn, rotated through a coordinate transformation in order to resolve the north-east-down (NED) components into cross-track, along-track, and below-track coordinates. Note that the altitude, or below-track, components are only used for the lever arm correction to the control point.

The large assumption buried inside this methodology is that the GPS measurements represent "truth." While this is known to be false (based on the data shown in Figure 3-22), the differential GPS position drift has been shown to be roughly Gaussian, with a one sigma position error of 0.36 meters. By forcing the Atlantis control system to follow a line moving slowly through space, an additional disturbance is introduced into the control system. However, this should be within the capability of the system. With that rather large caveat aside, the data presented in the tests are for mean and standard deviations, assuming that the GPS measurement is "truth." Thus, the result reported is the Sailboat Technical Error (STE), the wind-propelled marine vehicle equivalent of the standard Flight Technical Error (FTE). In addition, the navigation system error (NSE) is known to be approximately 0.36 meters, and the simple formula for Total System Error (TSE) is:

$$TSE = STE + NSE \qquad \text{(EQ 7.1)}$$

The trials consist of several straight line segments performed on various points of sail (though this is much less relevant for the trolling motor tests). Due to constraints imposed by the location of the shore, the runs are of different length. However, several good runs are included to give a statistically meaningful estimate of the control system performance. In between the straight line segments, the Atlantis is maneuvered through a simple open loop control consisting of full deflection of the rudder and monitoring the azimuth until the new desired heading is reached. This can certainly be improved upon, and suggestions for future research are included in Section 8.2.1 on page 303.

SECTION 7.3 TROLLING MOTOR EXPERIMENTAL SETUP

The system identification and controller tasks were completed before the wingsail construction. At this point, in order to test out the controllers, a MinKota electric trolling motor was used to simulate the presence of the wingsail and wind. This was done by mounting the trolling motor at the sailboat center of gravity and turning the trolling motor such that its direction of thrust was canted off the centerline by more than 40 degrees. The angle was varied during the course of the straight line segments to simulate changes in wind speed and thus cause a large disturbance to the controlled path from which the control system would have to recover.

Figure 7-7 The Atlantis unmanned trolling motor test run. The trolling motor is canted off the centerline in order to simulate the presence of a wing sail and all steering control is through the control system.

Since the dynamics of the catamaran are greatly affected both by the velocity through the water and the displacement weight of the hulls, the Atlantis was ballasted with an additional 75 kg of lead ballast to bring the total weight of the boat to the same weight

as the boat with the wingsail. Also, in order to test the controllers at various speeds, the MinKota trolling motor was run with 12, 24, and 36 volts at approximately 65 amps. This changes the speed of the boat through water, simulating changes in wind velocity. In order to simulate changes in wing direction, the MinKota trolling motor was turned through various angles while the controller was regulating the path to a line.

In Figure 7-7, the Atlantis with the trolling motor can be seen. The trolling motor is at the center of the boat, and the lead batteries provide the ballast. As pictured, the boat was run unmanned, with the GNC computer providing all navigation. Of note is the fact that the anemometer is located at the front wooden crossbeam. This is only a temporary location and moving the sensor's physical location is very easy due to the bus architecture employed on the Atlantis.

Early in the testing process, the power MosFETs that control the motor velocity failed due to the over-voltage and were removed from the motor control circuit. Subsequently, the trolling motor was able to run only at full throttle, and could be modulated in jumps of 12 volts by adding or removing a battery in series from the power supply. In fact, the motor was not able to cool itself sufficiently with 36 volts to last an entire run, so the usable voltage was effectively limited to somewhere in the vicinity of 28 volts. This was achieved by switching in drained batteries as their charge was consumed during the runs.

SECTION 7.4 TROLLING MOTOR EXPERIMENTAL RESULTS

The trolling motor test data presents an easier task for the automatic control system than does the sailing test. However, there are still disturbances due to wind, waves, and current. The task of regulating the position to a line proved to be a formidable one. In order to fully appreciate the precision with which the control system kept the Atlantis on

line, Figure 7-8 shows a satellite picture of Redwood City harbor, where the experimental trials were performed. In this picture, a series of yellow dots are visible, along with several red lines. The yellow dots are the sailing data mentioned in the introduction to this chapter and are the GPS traces for human pilot and conventional sails. A human pilot is essentially incapable of sailing a truly straight path. This is due to many factors such as variable winds, currents, and a lack of a precise reference to follow. Indeed, if there were a stripe painted across the surface of the water, it might be possible for a human pilot to regulate to it. The red lines are data taken from all of the computer controlled passes and show remarkably straight lines. The obvious point is that the red lines simply appear unnatural and are indeed the product of a machine.

The trace of the trajectories is not intended to be the same. This figure is intended only to give the reader an overview of the system performance. Removing the background in the figure, and zooming in on the trace of the trolling motor test runs, a typical test run is obtained and shown in Figure 7-9.

To the resolution pictured in Figure 7-9, the deviations from a perfect line are very small indeed. This is due to the fact that the figure is shown on a true scale, without distorting the cross-track error. The precision is, however, remarkable in light of the wind, waves, and current that were all injecting disturbances into the system. In order to quantify the performance of the automatic control system, the line is enlarged and rotated. The first part of the line of Figure 7-9 is rotated such that the x-axis corresponds to along-track, the y-axis corresponds to cross-track. Angular deviation from the desired line is also redefined relative to the line. This is presented in Figure 7-10, along with the data on the GPS velocity and the hullspeed sensor.

There are several interesting features of Figure 7-10 that should be noted. The path can be seen to stay quite clearly within the +/- one meter bounds that were set as the goals

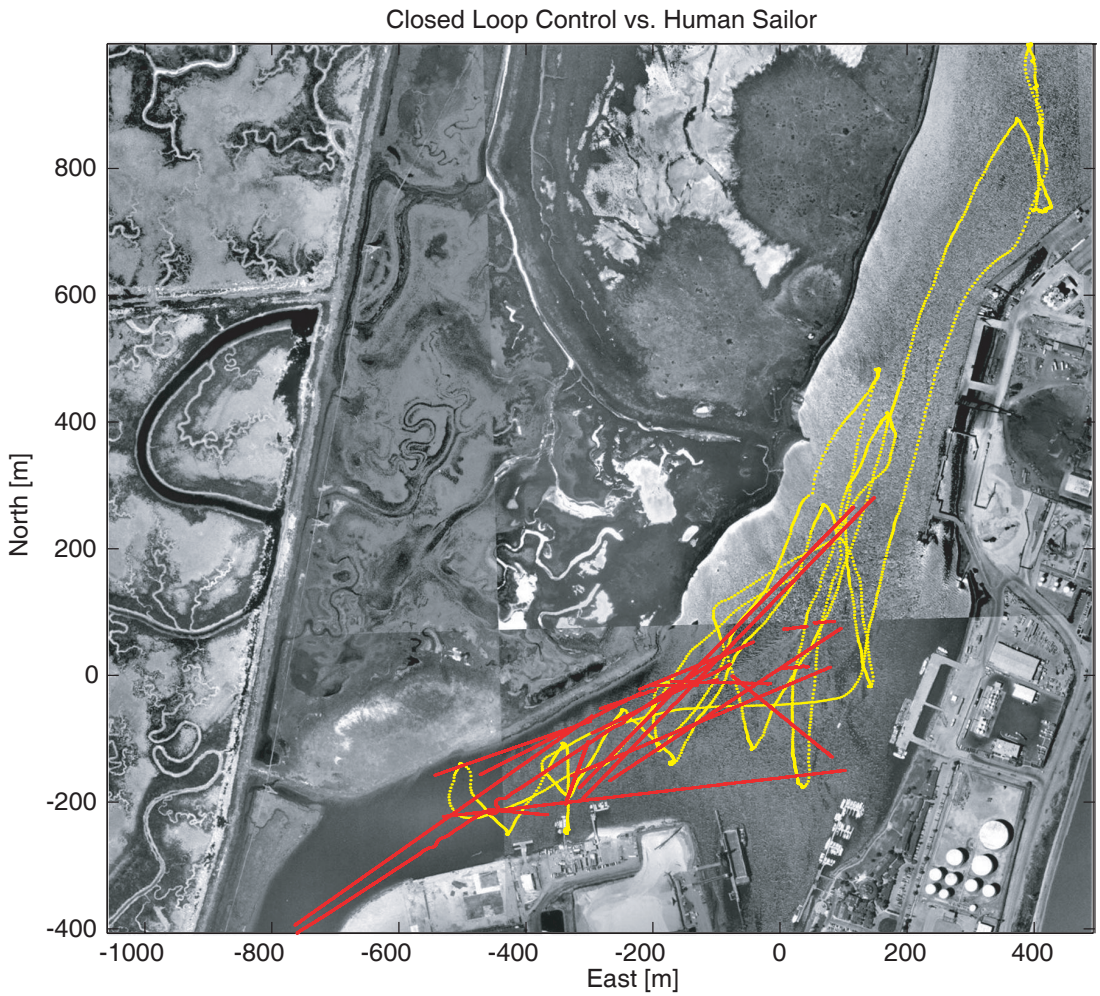

Figure 7-8 Satellite photograph of the Redwood City harbour with both human sailing data (yellow) and computer controlled trolling data (red) superimposed. Note that the paths are not intended to be identical, but rather demonstrate qualitatively the difference between human and computer control.

for the Atlantis. In fact, the mean is less than 0.03 meters and the standard deviation less than 0.10 meters over a course of approximately 600 meters.

The azimuth shows a -20 degree bias due to current for most of the path length of the run pictured in Figure 7-10. This can be verified by looking at the velocity plot at the bottom of Figure 7-10, where the top line is the hullspeed sensor, and the smooth lower

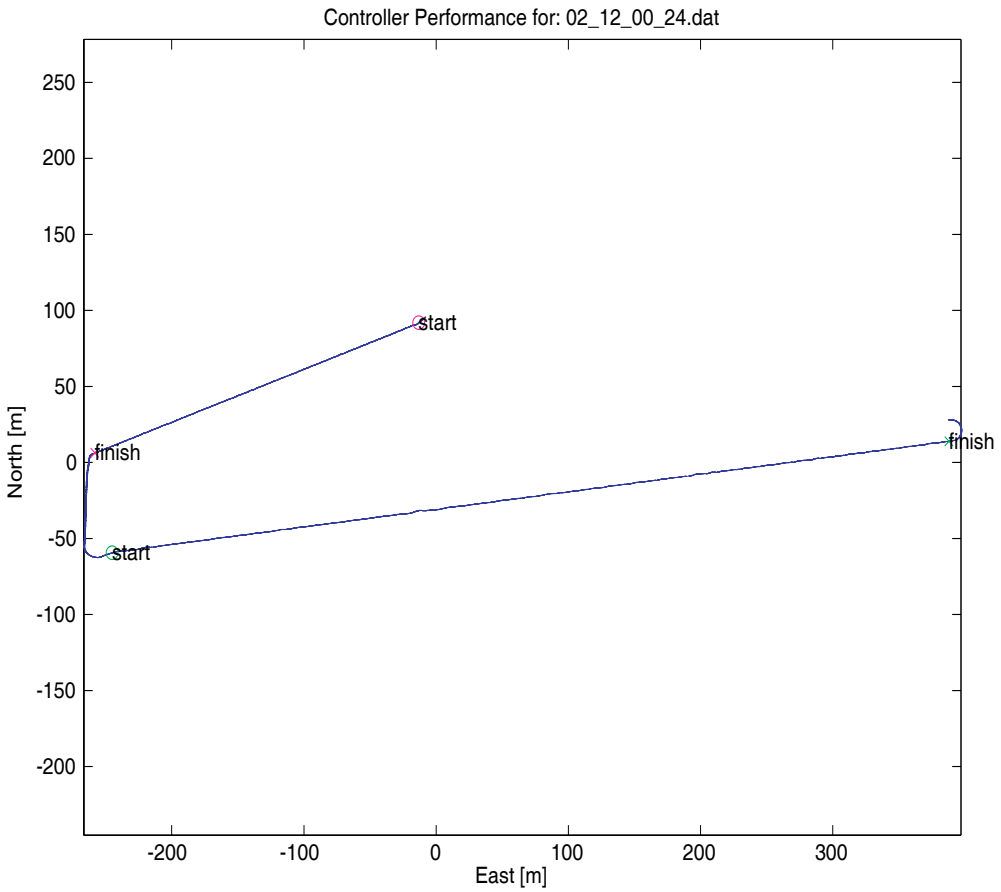

Figure 7-9 Overhead view of one trolling motor pass under automatic control. The automatic control is engaged only between the “start” and “finish” points and the turns between segments are performed open loop or by a human pilot.

line is GPS velocity. The difference in these two is current, and it can be seen in spite of the high frequency noise of the hullspeed sensor.

Figure 7-11 shows the aggregate passes of the identified plant and velocity invariant LQG controller runs. This was the best controller and was used exclusively during the wing-sailed trials. There are four trials superimposed upon one another, in different colors. The difference in path length is due to the shore proximity at the start of each trial and the shape of the Redwood City harbor. The paths vary from 250 to just under 700 meters in

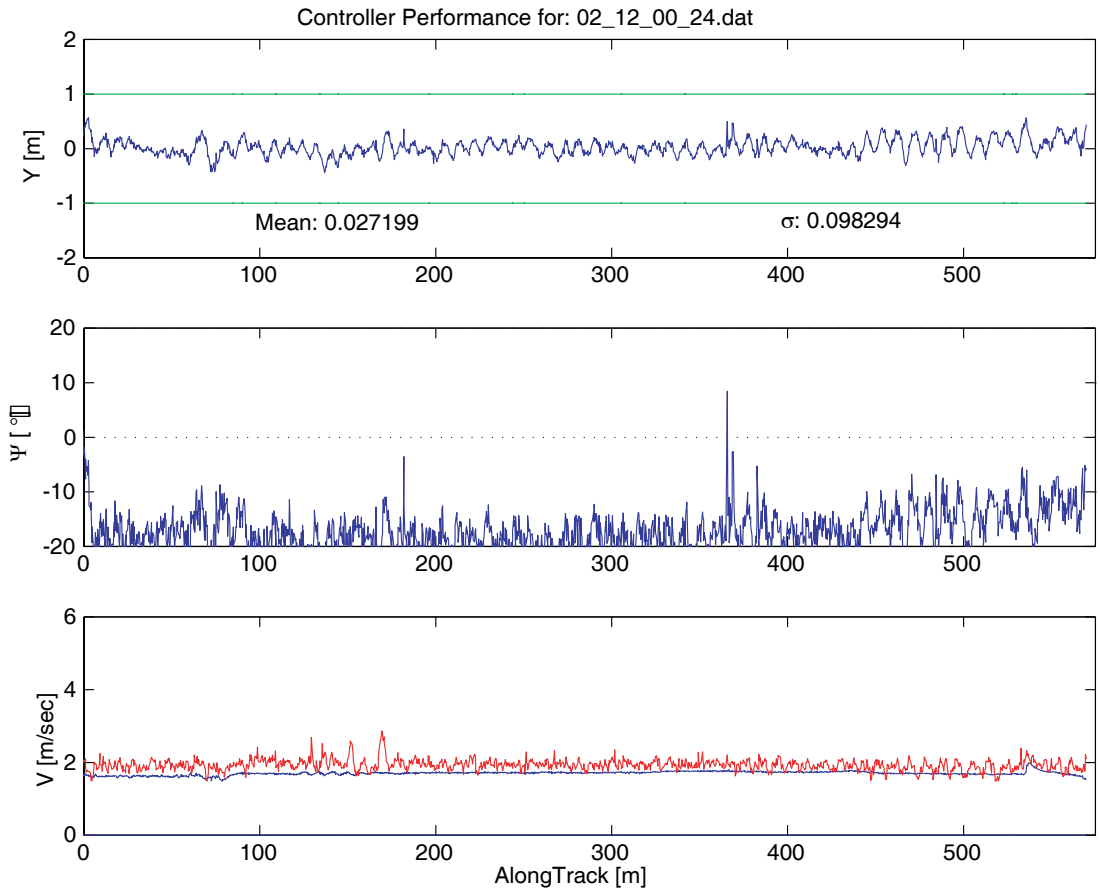

Figure 7-10 Trolling motor automatic control data showing a line following segment that is 600 meters long, with a mean deviation of 0.03 meters, and a standard deviation of 0.10 meters. This is using the identified system and the velocity invariant LQG controller.

length and all very clearly stay within the +/- one meter bound. The means of all of the runs are less than 0.05 meters, and the standard deviations are all less than 0.12 meters.

The spike in the data in Figure 7-10 for the red trace at approximately 275 meters along the path is due to colliding with an underwater pipeline at low tide. This caused the centerboard and rudder on the starboard side to automatically raise and induce a large angular disturbance to the automatic control system. The guidance system was able to recover from this condition and return to the line with only a 0.75 meter deviation, and continued to perform flawlessly once the centerboard and rudder were put back into the

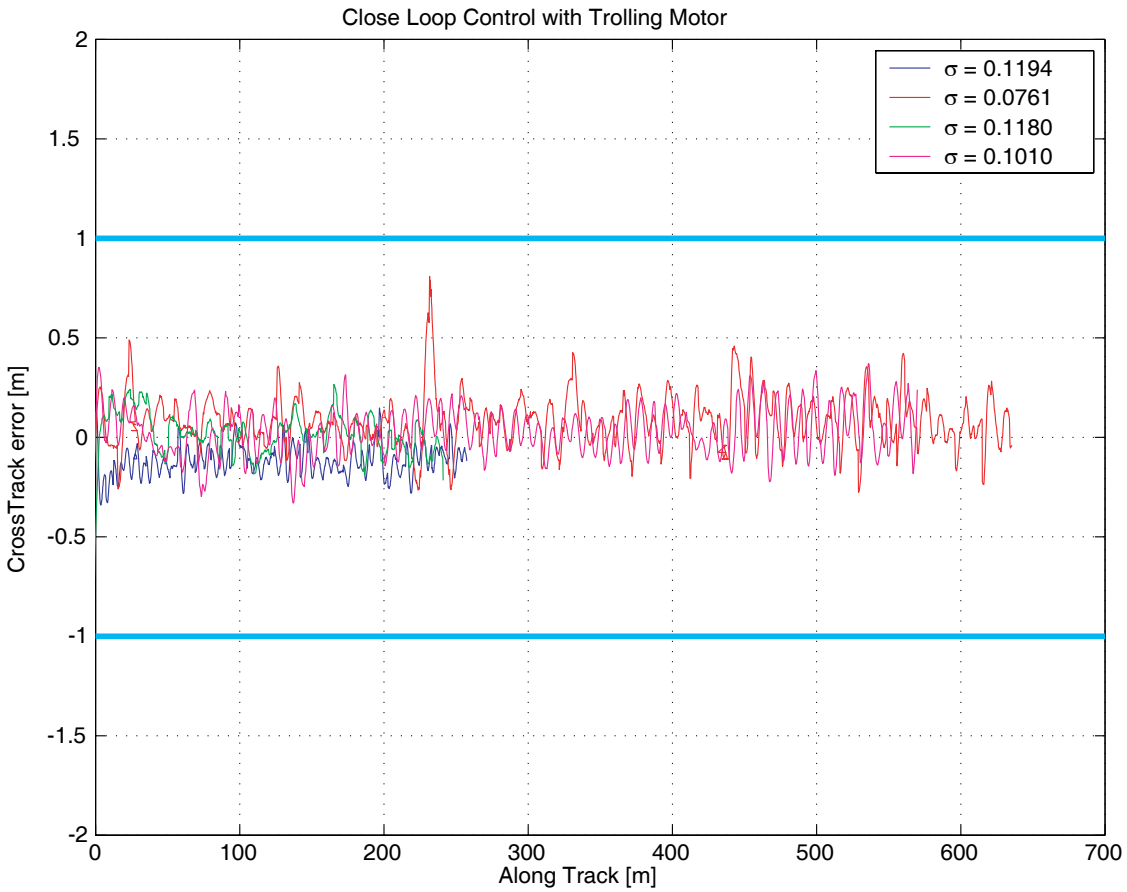

Figure 7-11 Aggregate plot of computer controlled trolling motor runs for the identified plant and velocity invariant LQG controller. The different path lengths were constrained by the location of the shore.

down position. As a point of comparison, the top-of-the-line AutoHelm autopilot system has a quoted system accuracy of 0.05 nautical mile or approximately 92 meters when guided by GPS.

SECTION 7.5 WINGSAIL EXPERIMENTAL SETUP

In order to validate the performance of the controllers and the entire system, closed loop control experiments were performed in Redwood City harbor, California, on 27 Janu-

ary 2001. These tests are intended to verify that the closed loop controllers are capable of precise line following with the increased disturbances due to the wingsail propulsion. No modifications were made to the controller design and the tests were run on a day with approximately 12 knots (or 6 m/s) of wind, with gusts up to the 20 knot (or 10 m/s) range.

The system plant is based on the identified model using the fourth order OKID state space model. The state estimator is the identified Kalman filter from the data, and the LQR controller is the velocity invariant controller intended to regulate the cross-track deviation normalized by velocity. This controller is used to guide the Atlantis along the straight line segments that correspond to the desired trajectory. In order to link the straight line segments, the Atlantis has to be turned via manual control to the new heading. During these turns, if a tack or a jibe is required, then it, too, is performed manually. Details of tacking and jibing, as well as the self-trimming aspect of the wing sail and tail configuration, are found in Section 7.1 on page 272.

Due to the fragile nature of the wing, every effort was made to ensure that a capsize did not occur. It was deemed unlikely that the wing sail could survive an impact with the water and the project did not have sufficient funds to construct a second wing in case the first one was destroyed. To minimize risk, the flaps on the wing were never deflected, resulting in a maximum coefficient of lift of just over one.

Result: the Atlantis sailed admirably with approximately 150 kilograms of live ballast (the author and two crew members, whose sole function was to rapidly move to the high side in the event of a roll angle upset of greater than 15 degrees). The self-trimming nature of the wing sail caused the Atlantis to sail bolt upright with very little roll angle.

Upon encountering a large gust, the Atlantis would give a small shudder and then continue on her course. The only indication that the wind had changed direction was that the wing sail now pointed in a new direction. The wingsail was tufted with yarn at several

points along the wing so that the flow over the surface could be visualized. This allowed the observation that the wing sail would stall (indicated by the tufts standing straight out from the upwind side of the wing) and then quickly rotate to a new angle of attack (at which point the tufts would realign themselves and point straight aft along the wing surface).

Another advantage of the wing sail was that the Atlantis was able to sail very high into the wind. Though exact measurements were not made due to a sensor failure, coarse analysis of video taken during the experimental trials indicated that the Atlantis was able to sail upwind at an angle of roughly 15 to 20 degrees from true wind direction. As a comparison, a mono-hulled sloop sailing in the same harbour during the experimental trials came about while behind the Atlantis and was luffing her jib while at 35 degrees to true wind. This occurred at the same time that the Atlantis was sailing upwind at 20 degrees.

SECTION 7.6 WINGSAIL EXPERIMENTAL RESULTS

The Atlantis was transported to the Redwood City harbour on a trailer, and the wing assembled onto the hulls. The upper flap was damaged in transport, but was quickly repaired in the field. With the aid of several helpers, the wing was assembled onto the catamaran and the catamaran returned to its trailer. A full system check verified that the navigation components were working, but the magnetic sensor that reads the angle between the wing sail and the hulls had failed. A decision was made to proceed with the experimental trials as this sensor is not critical to the operation of the Atlantis.

The wind was measured at approximately 4 to 6 meters/second with gusts up to 9 meters/second. While the Atlantis wing was designed for these wind speeds, it was with some trepidation that she was backed into the water. Among the concerns was the possibil-

ity of capsize, which would most likely destroy the wing and set the project back six months to a year. A chase boat was in operation at all times in order to retrieve the Atlantis should she turn over, and also to provide a platform for video of the experimental trials.

The Atlantis was manually sailed out of the slip into the main harbor area which required two tacks to get out of the narrow area between the two docks. During the first tack, the tail was reversed before the hulls were pointed into the wind, and the thrust effectively reversed. The Atlantis came to a complete halt, reversed course, and slowly drifted downwind until coming to rest against the Ocean Institute's ship, the *Robert Brownlee*. The chase boat gave the Atlantis a tow out from this area. This exact same scenario had happened with the conventionally sailed catamaran before, as it is difficult to tack in small areas. Figure 7-12 shows the Atlantis, the chase boat, and the *Robert Brownlee*.

Figure 7-12 The Atlantis, the chase boat, and the Ocean Institute ship *Robert Brownlee* at the beginning of the experimental trials in Redwood City harbor. The Atlantis had to be towed out past the ship due to a tack in which the tail was reversed prematurely. This resulted in a blown tack and the Atlantis being stuck against the side of the *Robert Brownlee*. This exact same thing had occurred a year before with the conventional sails.

Once clear of the *Robert Brownlee*, the tail was commanded to a position of approximately 12 degrees off of center and the hulls steered manually to a broad reach. The Atlantis quickly and silently accelerated to a velocity that made it difficult for the chase boat to keep up. However attempts to engage the automatic control did not yield any response came from the system. On-the-fly troubleshooting determined that the CAN network was operational, but that the Windows 98 laptop acting as the GNC computer had crashed. Due to the extreme prototype nature of the Atlantis, there was no manual override, and the Atlantis continued to traverse the harbor at a high velocity while the computer was rebooted. Eventually, it was discovered that the keyboard had been splashed and was transmitting keystrokes to the software. Once the keyboard was removed from the system, the control system became active and steered the Atlantis back onto the desired line.

Figure 7-13 above shows a satellite picture of the harbor where both the trolling motor and wingsail tests were performed. The yellow dots are from a previous year, when the Atlantis was conventionally sailed with a sloop rig by a human pilot in as straight lines as possible. The red dots indicate the various closed loop control passes from the recent wing sail propelled tests. Note that the yellow trace has a curving, “human,” look to it, whereas the red trace looks like a perfectly straight line. Qualitatively, the computer control simply looks unnatural. This is very similar to the conclusions presented in the previous section. Again, the trajectories are not intended to be the same, but to give a qualitative comparison.

Figure 7-14 shows an overhead plot of one of the automatic control traces. The controller is again generated by the fourth order OKID system identification, with the Kalman filter and the velocity invariant LQR controller. Note that the disturbances injected into the system through the wing sail propulsion system are much larger than those encountered with the trolling motor. The automatic control is active in each run between

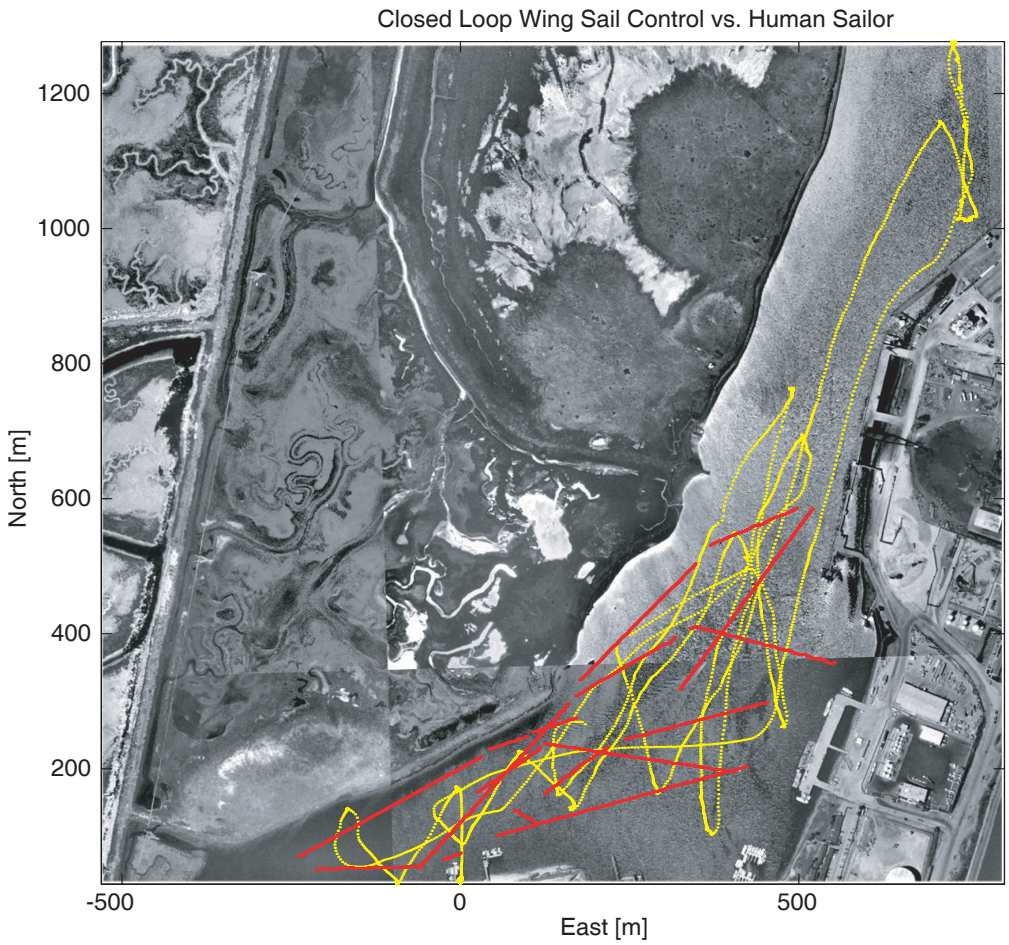

Figure 7-13 Satellite photograph of the Redwood City harbor with both human sailing data (yellow) and computer controlled sailing data (red) superimposed. Note that the paths are not intended to be identical, but rather demonstrate qualitatively the difference between human and computer control.

the "start" and "finish" labels. The maneuvering in between is performed by a human pilot. The wind is from the northwest direction.

There is a large transient in the second segment of the automatic control between the green "start" and "finish" tags. This transient is due to the fact that the wing sail turned out to be rather fragile when stressed fore and aft in the plane of the wing. When this occurred, the pin joints in between the three segments would act as hinges, allowing the wing to open up like a fan at the trailing edge, and come back down with a hard crash.

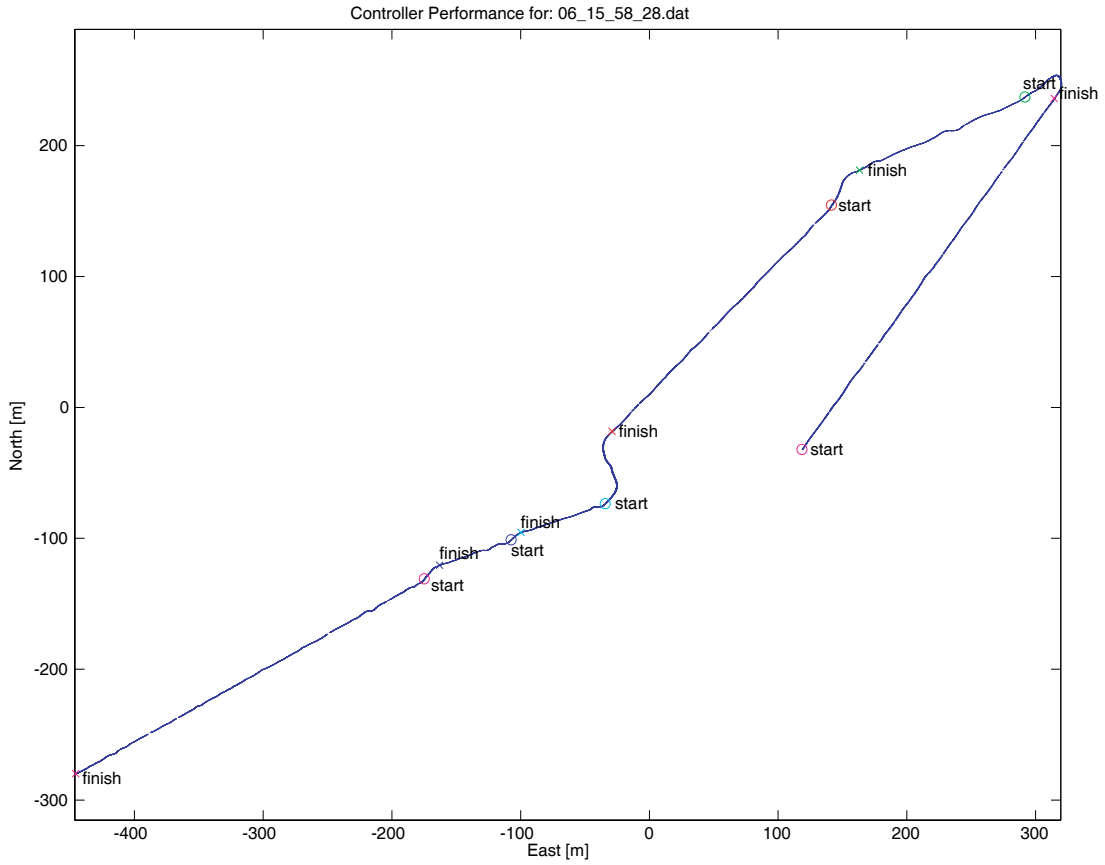

Figure 7-14 Overhead close up of Atlantis automatic control trajectory while under wing sail propulsion. Note that the automatic control is only between the "start" and "finish" points on the line, and that the maneuvering in between is guided by the human pilot.

This was destroying the trailing edge spar and making the structural integrity of the entire wing sail system questionable.

A decision was made to continue with the testing, but to manually force the wing to take any large transients out of the plane of the wing by forcing it perpendicular to the oncoming wave. These waves were generally caused by power boat wakes. In the run in question, a power boat had left a large wave, and the wing was forced 90 degrees to the wave. While the wing was able to cope with the wave with no further damage, the thrust was in fact completely reversed, and the control system wound up exacerbating rather than reducing the errors. As soon as the wing was released and allowed to realign to the relative

wind, the Atlantis locked back onto its trajectory and continued to follow its precise course.

At the resolution provided by Figure 7-14, performance appears to be very good. In order to quantify the errors in precision line following, a close-up look at the first segment is presented in Figure 7-15. This shows the cross-track error, azimuth error, and velocity of the Atlantis.

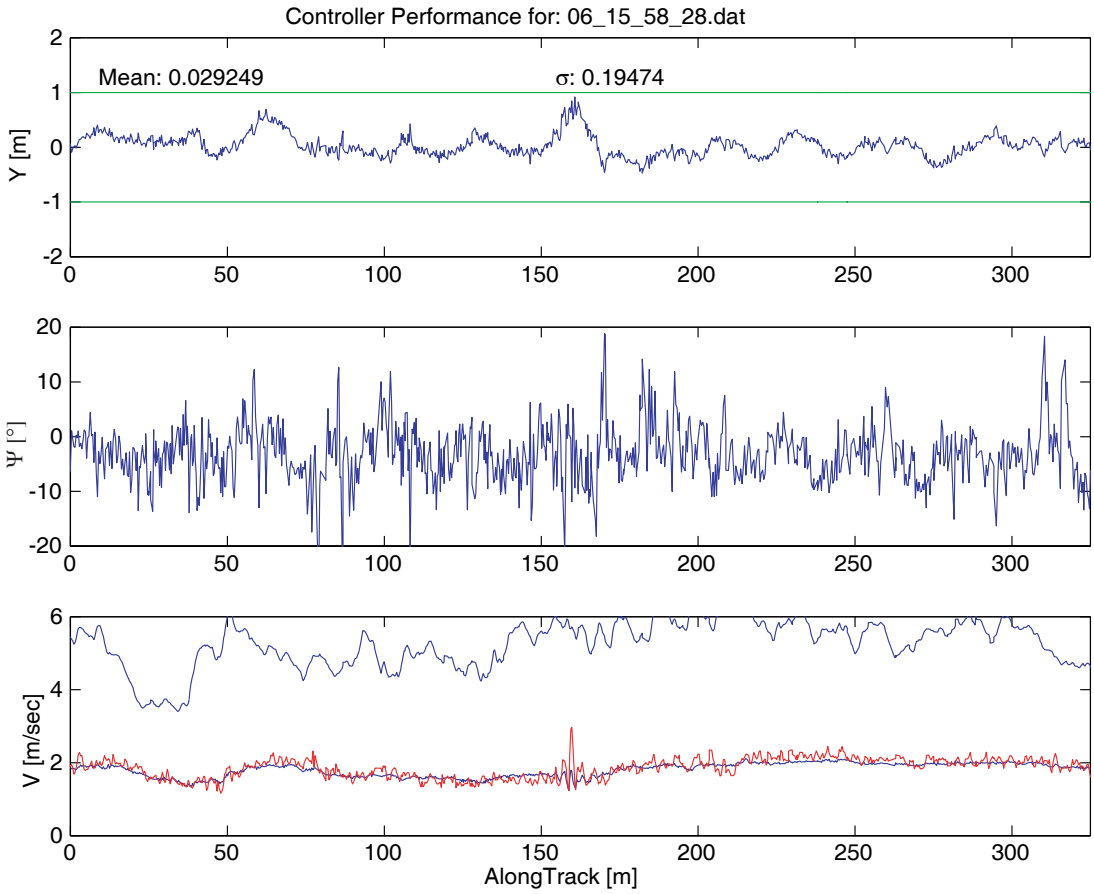

Figure 7-15 Close up view of the automatic control segment under wing sail propulsion. The upper blue line on the lowest plot is the wind velocity. Note that it changes by over 50% during this trial. The mean error of the track is 0.03 meters and the standard deviation is 0.19 meters.

The mean is less than 0.03 meters and the standard deviation is less than 0.2 meters. This, once again, is the Sailboat Technical Error (STE). Note that in this case the azimuth error is much smaller than the example shown for the trolling motor in Figure 7-10. This is due to the fact that on this pass, the current was much smaller, as evidenced by the overlay between the red hull speed sensor velocity data and the blue GPS velocity data on the lower plot. The upper line on the lower plot is the wind speed as recorded by the anemometer. The wind speed varies to well over 50% of its nominal value. This makes the performance of the control system even more remarkable as the driving force on the wing is constantly changing.

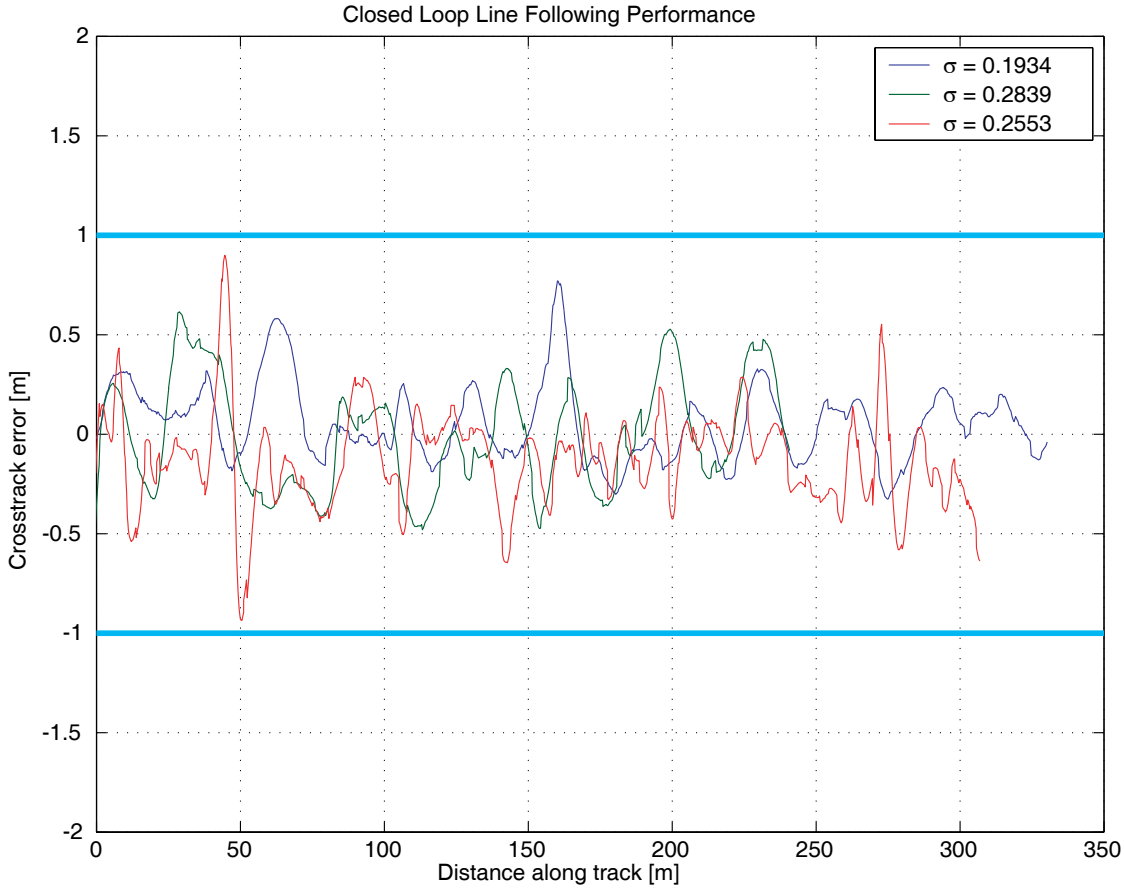

Figure 7-16 Aggregate plot of wing sail propelled automatic control trials using the 4th order OKID and velocity invariant controller. Note that the entire path is spent within the +/- one meter bound, and that the aggregate standard deviations are always less than 0.3 meters.

Figure 7-16 above shows the aggregate plot of three wingsail-propelled automatic control runs. The runs are all approximately 300 meters long, and show much more variation in cross-track than do the trolling motor runs. This was, however, expected due to the disturbances injected to the system by the wingsail propulsion system. The three runs all have means of less than 0.05 meters and standard deviations of less than 0.3 meters. Note that at no point do any of the trajectories leave the +/- one meter bound. Furthermore, the motion has a much lower frequency content than does the data from the trolling motor. This can be explained by the fact that the major disturbance coming into the plant is the variation of wind velocity and strength. Data taken from the anemometer indicates that the wind direction is variable over 40 degrees and that the wind strength varies by 50% off of its baseline value. The wind disturbances are the largest for which the control system has to compensate and have the slowly varying frequency signature of the wind variation.

The net result of these plots is to show that the velocity invariant control system that was designed based on the identified fourth order plant through the OKID algorithm is capable of controlling the Atlantis to a remarkably precise line. The best autopilot currently available commercially claims an accuracy of 0.05 nautical miles, or 92 meters. That the Atlantis can be autonomously controlled to better than 0.3 meters allows a much tighter tolerance on the obstacles that it must avoid.

It has been demonstrated that with the combined advances in GPS technology and the advent of low-cost sensors, an unmanned sailboat can be built that can navigate with unprecedented levels of accuracy. By utilizing a novel wingsail propulsion system, the difficulties of actuating a sail have been overcome and high authority control can be realized. A demonstrated Sailboat Technical Error (STE) in line following of less than 0.3 meters was achieved, in challenging conditions. Combined with a Navigation Sensor Error (NSE) of 0.36 meters, this yields a Total System Error (TSE) of less than 1 meter.

8 Conclusions and Recommendations

Throughout the development of this thesis, it was the goal of the author that the design, development, and implementation process be sufficiently explained to allow the reader to understand and, if desired, replicate the process that resulted in the experimental validation of the Atlantis project. While certain design choices could best be described as arbitrary, a coherent attempt has been made to justify, and thoroughly explore, the various trade-offs required to bring such a complex system from idea to experimentation. Indeed, though this thesis describes the process only for the particular implementation of the Atlantis project, it is hoped that it can act as a guide for the would-be engineer who sets out on a complex system integration task. The goals achieved with the Atlantis

project are of secondary importance to the process required to bring the vision to reality.

SECTION 8.1 CONCLUSIONS

This thesis is a systems work, with contributions in structures, fluid mechanics, and guidance navigation and control areas. The main contributions detailed in this thesis are:

Conceived, designed, built, and experimentally demonstrated an autonomous sailboat capable of precision control to better than 0.3 meters.

Data presented in Chapter 7, specifically in Section 7.4 and Section 7.6, the Atlantis was able to track a straight line defined by GPS to an unprecedented accuracy. The trolling motor tests demonstrated a mean of less than 0.03 meters and a standard deviation of less than 0.15 meters. This was performed entirely in the confines of Redwood City harbor, and represents a relatively benign environment. This is vastly superior to the commercial autopilots whose specifications provide a GPS line following accuracy of 0.05 nautical miles, which is approximately 92 meters. As a caveat, the stated accuracy of the commercial autopilots probably accounts for rough water operation, which the Atlantis never had to deal with. Still, with the trolling motor, the Atlantis achieves remarkable precision when following a line, in spite of current changes and winds, as was amply demonstrated in the results of Chapter 7.

The sailing data is, in many ways, even more impressive. In the case of the trolling motor tests, the wind variation did not feed directly into the disturbances of the catamaran guidance. With the wing sail propelling the Atlantis, variations in wind direction caused the lift vector to rotate an equivalent amount, greatly changing the amount of side force required by the centerboards. The automatic control is, however, able to compensate, yielding a mean of less than 0.10 meters and a standard deviation of less than 0.3 meters.

This was accomplished at speeds significantly higher than those reached with the trolling motor.

Developed methodology to identify robust plant models and controllers for a winged catamaran that are invariant under velocity changes.

The identification procedure and velocity invariance detailed in Chapter 6 consist primarily of adapting an existing system identification algorithm used for flexible space structures to this unique problem. By utilizing some basic modelling techniques from previous work on a GPS-controlled farm tractor, the variables that drive the identification process are noted and separated for the algorithm to operate. Furthermore, the identification process makes no assumption on the structure of the model, but uses a very general state-space structure with the model order decided by the data stream itself.

Great care was taken to exercise all of the modes of the guidance problem in order to inject the identification algorithm with sufficient power for isolating the required variables. In addition, the pseudo-Kalman filter that best estimates the states from the measured input variables was identified. This allowed a simple LQG controller to be highly accurate, as verified experimentally.

The problem of variable velocity was solved in Section 6.3 by recognizing the basic structure of the kinematic model and observing that the errors integrated not with time, but with distance. Thus, the data was pre-scaled by velocity when using the identification algorithm, very similar to a method in a previous paper by the author [45] on tractor implement guidance. By pre-scaling the input to the control algorithm by velocity, the entire system continues to behave well even in the presence of large velocity changes. The control algorithm automatically becomes less sensitive as velocity is increased and more sensitive as velocity is decreased. This allows the greater precision of the automatic control to be retained without forcing the system into instability as velocity is increased.

While there exist other methods to deal with the velocity dependence, the experimental results of this methodology demonstrate that this solution is very powerful indeed.

Described optimization scheme for symmetric wingsail section based on requirements unique to sailing vehicles.

The basis for the propulsion system is a symmetric wingsail that was designed to achieve a high maximum lift coefficient with a simple flap at the Reynolds numbers appropriate to sailing vehicles. The details of this design, laid out in Chapter 5, include the requirements, development of specifications, and full analysis. The wing and tail sections are developed with a flat rooftop pressure distribution and boundary layer transition strip followed by a very gradual pressure recovery.

Using a simple trailing edge flap obviates the need for the more exotic solutions to maximum lift coefficient, such as over-the-top tacking, while preserving the ability to sail on either side of the wind. The optimization of the flap to chord ratio of the trailing edge flap results in the non-intuitive answer of a small (13%) flap with large deflection to retain the high lift to drag ratios at large coefficients of lift.

Due to the very large thickness to chord ratio of the final wing section (21%) the structure of the wing can be made very light. The structure uses simple plywood and aluminum with standard bearings that suspend the wing/tail system on the stub mast. The wing is covered by a thick polyester cloth that is heat shrunk onto the ribs and the hinges are fabric hinges made from the same material. The entire wing/tail system (without ballast) weighs less than 71 kilograms. The ballast is used to bring the center of gravity of the wing/tail system in line with the pivot axis of the stub mast, allowing a de-coupling of pitch and roll motions of the hulls from the aerodynamic angle of attack of the wing/tail system itself.

Developed and experimentally demonstrated novel quaternion-based attitude estimation algorithm from vector observations.

One of the primary inputs to the control system of the Atlantis is the azimuth or yaw orientation of the hulls. While this can be measured either through a magnetic compass, north-seeking gyro, or via the velocity vector derived from GPS, each of these methods has problems associated with it. GPS velocity measurements are very good, but degrade in directional information as the velocity decreases. Furthermore, the velocity measurements are relative to the fixed Earth and thus are skewed by currents in the body of water upon which the Atlantis is navigating.

North-seeking gyros are impressive pieces of equipment, but tend to be large, heavy, and very expensive. Due to budgetary constraints on the project, neither simple north-seeking gyros nor the full inertial navigation (INS) solution could be used. New magnetic sensors that read Earth's magnetic field have greatly reduced the cost of high accuracy heading information. Accurate heading data from these sensors is, however, sensitive to whether they are level in the two orthogonal axes. Due to the fact that the hulls roll and pitch significantly, a solution for attitude was required.

By fusing three orthogonal accelerometers and a three axis magnetometer, a new algorithm was developed that solved Whaba's problem directly in the quaternion domain and is demonstrated to be globally convergent. The new algorithm was simulated and test compared to a commercial navigation grade inertial navigation (INS) unit aboard a maneuvering aircraft. The results compared well with the INS, with pitch and roll errors on the order of 1-2 degrees, and azimuth errors on the order of 3 degrees (during moderate flight conditions, longer term performance is stable, but worse). In the lower disturbance environment of the Atlantis, the attitude errors were on the order of 0.5 degrees in pitch and roll and 1.5 degrees in azimuth. The improvement in accuracy was largely due to the averaging of samples. The sensors in the aircraft were sampled at 100 Hz while those on the boat were averaged down to 5 Hz, yielding a reduction of approximately $\sqrt{20}$ in the errors. The attitude system has proved to be robust and drift free. Though the aircraft flight

test revealed a singularity in the solution when banking at certain angles, this condition is never encountered in the marine application. Note that this singularity could easily be corrected with the use of relatively inexpensive gyros, but the cost of using them was deemed excessive for this project.

Developed and experimentally demonstrated novel method for calibrating any 3-axis sensor that requires no external reference.

In order for the attitude system to work, the absolute value of the magnetic fields needs to be sensed accurately. Conventional calibration of magnetometers requires an external reference and still can only correct for errors in the heading domain. The details of the magnetometer calibration algorithm are given in Section 4.7.

The magnetometer calibration algorithm relies on the fact that the three components of the magnetic field will have a constant vector length and that simple motion of the platform will cause the tip of that vector to trace out a locus on the sphere of the vector length. This algorithm is, in fact, completely general and will work on any three axis sensor that measures a constant vector quantity.

By formulating the problem as a least squares solution, the scale factors and biases on all three axes can be estimated simultaneously. The algorithm essentially takes the aggregate data points collected and, by varying the scale factors and biases, forces the data to lie on the surface of a sphere centered at the origin and of a given radius in a least squares sense. Thus, without any external reference, the magnetometers can be calibrated simply by rotating the magnetometers and continuously recording their output.

Experimental validation of this algorithm showed a very large reduction in the spread of the vector norm of the magnetic field measured by the magnetometers compared to a pre-calibration data set. Furthermore, the simple heading calculation was determined to be within 1.2 degrees of a navigation grade inertial navigation unit. With this calibration

algorithm, the difficulty of precisely aligning the vehicle to a known heading is removed and the calibration task greatly simplified.

SECTION 8.2 FUTURE WORK

Having elaborated on the results of the project, it is appropriate to enumerate the flaws, failures, and things that could have been done better. These are cast in the form of future improvements that would greatly enhance the project and are thus directed to the next researchers to follow along on this work. Many of the suggested improvements are not difficult and only need sufficient time and budget. While myriad improvements are possible, only those deemed sufficiently interesting are included here.

SECTION 8.2.1 IMPLEMENT AUTOMATED TACKING AND JIBING

The current configuration of the Atlantis control system is such that the automated control segments are straight line segments between given waypoints. Turns in between these line segments, and the associated tack or jibe, are currently performed open loop by the human pilot. This process could be automated. An outline of a possible approach follows.

The tack or jibe is in fact the same operation, both require the tail and flap be centered as the wing pivots when the hulls are pointed directly upwind/downwind (regardless whether the boat turns bow or stern into the wind). Afterwards, the position of the flap and tail are then smoothly reversed from the previous tack or reach.

The figures from Section 7.1 have been replicated here, in order to visualize the tacking or jibing process. From Figure 8-1, only the upwind tack is demonstrated. It is

important to note that the jibe is in fact the exact same process, but with the hulls pointed downwind, rather than upwind.

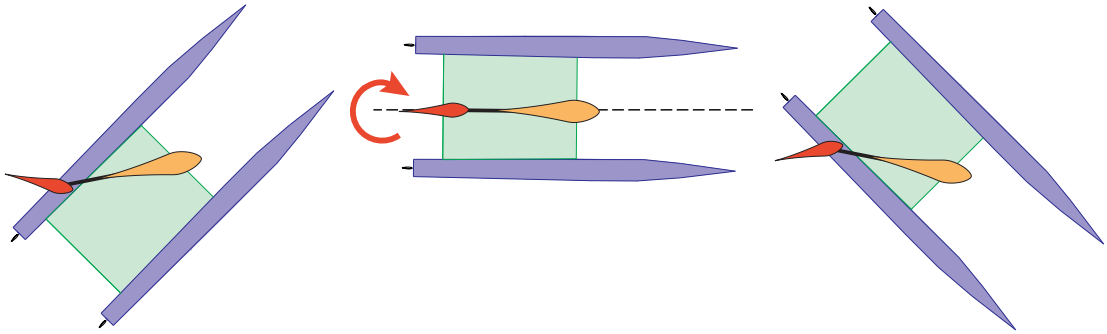

Figure 8-1 The upwind tack, in which the flap and tail are centered, and then reversed as the catamaran heads up through the wind. The tack occurs from left to right and the wind is coming from the right side of the page.

Looking carefully at the sequence, it is clear that the tail and flap must be centered when the hulls are aligned with the wind. If the flap and tail are reversed before the hulls are centered with the wind, then the wing will reverse its lift vector, and propel the catamaran backwards. This has the effect of putting the catamaran in irons, or getting stuck in the center of the tack. By waiting until slightly after the hulls are centered into the wind, the tack will progress smoothly. Evidence from the test sailing shows this was not difficult to do.

Using a second anemometer or weathervane on the hulls of the catamaran, the point at which the wind is straight into the hulls can be measured. By averaging the measurement for a second, the lag of the measurement will induce the required delay. As soon as the averaged measurement is zero or negative, then the flaps and tail should be reversed. The same is true for the jibe, though the timing is much less critical since the windage of the hulls and sail are already pushing the boat through the jibe.

SECTION 8.2.2 IMPLEMENTING REVERSE CONTROL

Though a normal sailboat sails forward on its hulls, there is no real constraint to require this on an automated sailboat such as the Atlantis. Conventionally, a sailboat that is sailed backwards is unstable in azimuth and desires to swap ends with the direction in which it is travelling. Catamarans, in general, are no exception and also have a strong tendency to try and spin about when going backwards. This is exploited in catamaran sailing when a tack is performed poorly and the catamaran is stuck in irons. At that point, the rudders are reversed, and the boom is held out by the crew in order to propel the boat backwards onto the correct tack. At which point the boom is released and the catamaran accelerates on its new course.

The automatic control that is capable of steering the Atlantis along a line to better than 0.3 meters is also fully capable of stabilizing the path while going in reverse. It does require a new batch of data for the catamaran going backwards and a new controller design. This was attempted while collecting data for system identification. However, the trolling motor failed at that point and the crew had to paddle back to shore.

Once a reverse controller exists, it is a simple matter of deciding which direction the boat is travelling, forwards or backwards. This can be accomplished using the azimuth derived from the attitude system, along with the heading information taken from the GPS velocity measurements. If these two measurements are within the same half circle, the boat is moving forwards; otherwise the boat is moving backwards. There is a potential error here due to the possibility that the catamaran will be moving forward through an area with current great enough that the ground velocity will be negative. This can, however, be detected through the hull speed sensor.

With the ability to switch between forward control and reverse control, station keeping becomes a relatively easy task. The Atlantis would be programmed to travel for-

wards from point A to point B, and backwards from point B to point A. As the points are brought close together, a limit cycle is set up in which the Atlantis will oscillate around a given point. Furthermore, the controller can be commanded to bring the GPS velocity to zero, and in the presence of current, this would cause the Atlantis to simply sail into the current at the speed of the current, thus holding position.

There are other ways of implementing station keeping: racetrack patterns, orbits around a point, and a hysteresis loop set up with the point at its center and some nominal offset as the trigger. The ideal method to use should be the subject of future research. Station keeping does not, however, appear to demonstrate any insurmountable difficulties.

SECTION 8.2.3 CURRENT ESTIMATION AND OPTIMAL TRAJECTORIES

The current hull speed sensor is too noisy for high quality current estimation. A better sensor or even better processing would allow for on-the-fly current estimation. This was something that the present project simply did not have time to implement. Several methods could be used to estimate the current in the present location of the Atlantis.

The difference between the GPS velocity vector and the vector of hullspeed projected in the direction of the azimuth can be used to estimate currents perpendicular to the path of the Atlantis. This could be implemented with no change to the existing system and would be useful to plot out the current field of a body of water as the Atlantis traverses it.

Another method is to periodically perform a full circle maneuver and test the velocity through the turn. At this point, the true current vector can be measured, and assuming it is slowly changing, the entire current field can be mapped out. A variation on this method is to periodically stop and drift, again, directly measuring the current from the GPS velocity. The only obvious problem with this method is that the GPS velocity measurement has difficulty below the 0.25 meter/second range due to the internal velocity fil-

tering on the Trimble Ag122 receiver. This filtering can be turned off, but there will remain a lower limit under which the velocity cannot be detected. Using numerical differentiation of position for velocity is a possibility, but when the motion becomes small, the noise in the position measurement will directly corrupt the velocity estimation. (Note that there are other methods to determine GPS velocity such as accumulated delta range or doppler that are more accurate especially at lower velocities)

Once the current field is estimated, then an optimal trajectory can be generated to the desired goal or end point of the trajectory. While this may not be extremely relevant in the case of crossing a small harbor, it can become very relevant if the Atlantis is scaled up and used for transporting freight across the oceans. The advantage of sailing optimal trajectories with a good knowledge of both the prevailing winds and currents, supplemented by current updates and local measurements, can be tremendous.

At present, the only group to use such methods are transoceanic racing sailors, who are quite sophisticated in using satellite data to predict winds and currents. The difference can account for several days in a trans-ocean crossing, where the average speed of the sailing vessel is generally less than 5 meters/second (~10 knots).

SECTION 8.2.4 MAP OR GRADIENT BASED NAVIGATION

Presently, the Atlantis is programmed only to start at a given point and follow a given line on the surface of the water. However, the computers onboard have access to location, wind speed and direction, and at least a coarse estimation of current. One of the obvious next steps with the Atlantis is to build a trajectory generator, or higher level navigator. In many ways, this is very similar to the previously mentioned ideas on optimal trajectories.

Using map- or gradient-based higher level navigation, the Atlantis trajectory generator would simply be given the desired objective (i.e., reach point X in the minimum time). With this objective, the trajectory generator would use maps as well as known up-to-date wind and current conditions to set a series of linked straight line courses to the point. The gradient-based method could be used to ensure that the Atlantis successfully avoids known obstacles such as a structure or the shore.

Unfortunately, in its present incarnation the Atlantis has no way of knowing what other marine vessels are nearby. Thus, without an "other vessel" sensor, the Atlantis is still prone to collision with other vessels. One possibility is to use a radar to flag potential collision hazards. Another is to use a techniques referred to as Automatic Dependent Surveillance (ADS) in which each vessel broadcasts its own GPS-based position on a low-power radio link.

SECTION 8.2.5 DISTRIBUTED CONTROL

Given that the Atlantis uses a number of microcontrollers all networked to each other, there is no particularly good reason to use a centralized Guidance, Navigation, and Control computer. It is quite possible to make the control system extremely fault tolerant without great changes to the current hardware.

The microcontrollers can test each of their respective sensors and actuators, as well as themselves, repeatedly. In the case that a sensor or actuator fails, the control system will drop into a reduced functionality mode in which the faulty sensor or actuator is removed from the system by ignoring the messages generated on the network. Another way of accomplishing this would be to run a bank of Kalman filters, each one missing one specific sensor. When all of the sensors are working, the covariance of the Kalman filters should be very similar.

If a sensor fails, one Kalman filter (whose input does not include that sensor) will continue with the same covariance, while the rest will experience a large rise in their covariance matrices. A simple metric should be able to detect this and use the state estimate from the remaining Kalman filters. Note that these can be cascaded in such a way as to form a logical sensor failure tree. In this case, as each sensor fails, the control system keeps dropping into a further reduced mode while ensuring that control is not lost.

The control system can be made to operate directly from the microcontrollers such that any subset of the microcontrollers can continue to safely operate the Atlantis even after others have failed. There is a large body of research on the applications of distributed control, and the Atlantis makes a very good testbed for these as the sensors and actuators are already hooked together via a fault tolerant high speed serial network (CAN bus).

Other improvements to the system that should be made are in the area of user interface and operating system. The user interface now consists of a series of commands issued through the keyboard of a laptop and plain text onto the screen for debugging purposes. The entire system is run on an interrupt-driven process under *Windows 98*. While underway, the laptop is placed within a sealed watertight container and a remote keyboard is attached.

During one test run, the controls were no longer operating as expected and, with no manual override, it made for an exciting test run. The problem turned out to be a CAPS lock issue with the keyboard. Obviously, a keyboard interface was used because it was quick to construct and very inexpensive. However, a control head microcontroller, with an LCD screen and joystick in a waterproof case with a clear plastic housing, would be easy to construct. This would allow the PC to remain in its waterproof box and be used only to log data to its hard disk for later analysis.

Also, *Windows 98* is notorious for having large latencies in its interrupt processes and really should not be used for real time control. The latencies were finessed by having the main process read a line on the parallel port that was being strobed by the microcontrollers at 100 Hz. Inside the control loop, however, the nominal time was supposed to be 200 milliseconds between control outputs. The measured jitter on this was zero mean and had a standard deviation of 2 milliseconds. Note that the distribution was not gaussian, with occasionally large outliers appearing.

All of this can be circumvented by using a proper Real Time Operating System (RTOS) such as QNX, Lynx, or even LinuxRT with the real time extensions. Since the code is modular and in "C," porting it to an RTOS should not be terribly difficult and would improve the system robustness a great deal.

SECTION 8.2.6 EXPERIMENTALLY MEASURE THE WING SAIL PERFORMANCE

Given the amount of analysis that went into designing the wing sail section, verifying the performance under sail would validate the CFD codes and design methodology. There are a number of ways in which this could be accomplished, either by using strain gauges or by generating high accuracy drag polars of the hulls from towing tests.

Obviously, the entire wing could be placed in a wind tunnel as well, though the costs would most likely be prohibitive. Several methods have been published on how to generate accurate drag polars of the hulls using towing tests [28]. Note that several other attempts to measure the performance of an actual sailing wing have met with much difficulty and little success [8]. Both of these methods required that the measurements of a strain gauge or scale be estimated on-the-fly by a human observer. Modern electronic recording equipment eliminates these obstacles, and better estimation techniques should be able to generate a high confidence estimate of the parameters in question.

The wing itself, as mentioned in Section 5.9, is very light and fragile. Before any bold experiments can be carried out, the wing needs to be made more structurally robust. This can be accomplished by reinforcing the existing structure with fiberglass and epoxy, or by rebuilding the wing with slightly heavier materials. In either case, there needs to be a structural joint between the trailing edge of the three sections to prevent the joints on the main spar from acting like pin hinges.

SECTION 8.2.7 ON BOARD POWER GENERATION

The current incarnation of the Atlantis derives all of its electrical power from a 12 volt marine battery. Since the battery has a 70 amp hour capacity, the Atlantis can run for approximately 15 hours before the battery needs to be recharged, less if the control system is particularly active.

Several methods exist for generating the power onboard the Atlantis. The first is commercially available solar cells used on sailboats today to keep the batteries charged. A typical module of solar cells made by Siemens is capable of an output of 6.7 amps at 12 volts at a cost just over $500.

The other form of power available is a wind turbine. Typical of the many off-the-shelf solutions is one from Industrial capable of sourcing 25 amps at 12 volts in a 12.5 meter/second wind (a rather strong wind). This is more than capable of driving the entire system and recharging the batteries at the same time. Again, the cost is similar, on the order of $600.

The advantage of the wind generation is that while heading upwind, the generator will see a very large increase in wind speed due to the effect of the relative wind. The generator will effectively be operating in the vector sum of the boat speed and the true wind

speed. This will allow the generation of power sufficient to make returning to port unecessary.

SECTION 8.2.8 UNMANNED OCEAN CROSSING

With the improvements to the control system, user interface, wing structural robustness, and on board power generation, the Atlantis becomes capable of self-sufficient crossings of large bodies of water. After several shakedown cruises of longer and longer lengths, it becomes conceivable to attempt a very long crossing, such as the trip between San Francisco and Honolulu. With that crossing, the viability of the concept will truly be established.

References

[1] Amateur Yacht Research Society, *Wingsails*, AYRS Publications #14, AYRS Press, London, 1957.

[2] Analog Devices webpage, <http://products.analog.com/products/info.asp?product=ADXL202>.

[3] Armand, J.L, Marol, P., and St.-Blancat, G., *Aerodynamics of Sail-Assisted Propulsion of Commercial Ships: A Preliminary Study*, The Ancient Interface XII, Proceedings of the Twelfth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Seattle, 1981. Pages 152-163.

[4] Asker, G. C., *Roller-Furled Genoa and Rigid Surface Wingsail, a Flexible Practical Wind-Assist System for Commercial Vessels*, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 20, Amsterdam, 1985. Pages 61-81.

[5] Baker, R. M. L., *Preliminary Mathematical Analysis of a Rigid-Airfoil, Hydrofoil Water Conveyance*, The Ancient Interface II, Proceedings of the Second AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 9, AIAA Los Angeles, 1970. Pages 131-140.

[6] Baker, R.M., and Gallington, R.W., *Preliminary Analysis of the Self-Tending Rigid Airfoils for the Hydrofoil water conveyance*, The Ancient Interface III, Proceedings of the Third AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 10, AIAA Los Angeles, 1971. Pages 178-193.

[7] Baker, R.M, *Four Dimensional Simulation of a Hydrofoil Craft*, The Ancient Interface VIII, Proceedings of the Eigth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Stanford, 1977. Pages 100-118.

[8] Baker, R.M., *Tests of a Rigid-Airfoil Sails usig a shore-based test stand*, The Ancient Interface IX, Proceedings of the Ninth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Pamona, 1979. Pages 25-60.

[9] Bar-Itzhack, I. Y., and Goshen-Meskin, D., *Unified Approach to Intertial Navigation System Error Modeling*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 15, No. 3, May-Jun. 1992, Pages. 648-653.

[10] Bar-Itzhack, I. Y., and Idan, M., *Recursive Attitude Determination from Vector Observations: Euler Angle Estimation*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 10, No. 2, Mar.-Apr. 1987. Pages 152-157.

[11] Bar-Itzhack, I. Y., and Reiner, J., *Estimation of Rodrigues Parameters from Vector Observations*, IEEE Transactions on Aerospace Electronic Systems, Vol. 21, No. 1, April 1996. Pages 578-585.

[12] Bar-Itzhack, I. Y., Montgomery, P. Y., and Garrick, J. C., *Algorithms for Attitude Determination Using the Global Positioning System*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 21, Nov.-Dec. 1998. Pages 846-852.

[13] Bar-Itzhack, I. Y., and Oshman, Y., *Attitude Determination from Vector Observations: Quaternion Estimation*, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-21, No. 1, Jan. 1985. Pages 128-135.

[14] Barkla, H. M., *The behaviour of a Sailing Yacht*, Transactions of the Royal Institution of Naval Architects, Vol. 103, Jan. 1961. Pages 1-23.

[15] Barkla, H. M., *High Speed Sailing*, Transactions of the Royal Institution of Naval Architects, Vol. 93, March 1951. Pages 235-257.

[16] Barkla, H.M., *The Linear Wind/Water-Mill/Propeller*, The Ancient Interface XII, Proceedings of the Twelfth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Seattle, 1981. Pages 88-102.

[17] Barton, C. E., *Revision of International Geomagnetic Reference Field Release*. EOS Transactions, 77(16), April 1996.

[18] Bassmann, L. and Swannell, P., Engineering Statics Study Book, Distance Education Center, Toowoomba, Austrailia, 1996.

[19] Bauer, A. B., *Sailing all points of the compass*, The Ancient Interface III, Proceedings of the Third AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 10, AIAA Los Angeles, 1971. Pages 72-79.

[20] Beer, F. and Johnston, E. R., Statics and Mechanics of Materials, McGraw-Hill, New York, 1992.

[21] Bergeson, L., *Sail Assist Developments 1979-1985*, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 20, Amsterdam, 1985.

[22] Bergeson, L., Clemmer, G., York, J., Bates, A., Mays, J., and Glucksman, M., *Wind Propulsion for Ships of the American Merchant Marine*, MA-RD-940-81034, Norwell, MA 1981.

[23] Bergeson, L., Greenwald, C. K., Hanson, T. F., *Magnus Rotor Test and Evaluation for Auxiliary Propulsion*, The Ancient Interface XIII, Proceedings of the Thirteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA San Diego, 1983. Pages 125-137.

[24] Bergeson, L. and Lord, J., *Current Prospects/Sail-Assist for Ships Above 1000 dwt*, International Conference on Sail-Assisted Commercial Fishing Vessels (ICSACFV), May 1983. Pages 170-190.

[25] Birchill, J., *The Windform Sail*, The Ancient Interface II, Proceedings of the Second AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 9, AIAA Los Angeles, 1970. Pages 59-73.

[26] Quinton, B., Boatek website, <http://www.boatek.demon.co.uk/>

[27] Bowditch, N., The American Practical Navigator, Defense Mapping Agency, Hydrographic/Topographic Center, Bethesda, Maryland, USA, 1995.

[28] Bradfield, W. S., *Predicted and Measured Performance of a Daysailing Catamaran*, Marine Technology, January 1970. Pages 21-37.

[29] Bradfield, W. S., and Menon, S., *A Rig Performance Simulator for Sailboat Preliminary Design Use*, The Ancient Interface XV, Proceedings of the Fifteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 31, AIAA Seattle, 1985. Pages 47-59.

[30] Bryson, A. E., *Kalman Filter Divergence and Aircraft Motion Estimators*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 1, No. 1, Jan.-Feb. 1978. Pages 71-79.

[31] Carl, W. P., *Sails without masts*, Yachting, Vol. 79, No. 2, Feb. 1946, p. 75-77,108.

[32] Caruso, M. J., *Application of Magnetoresistive Sensors in Navigation Systems*, Sensors and Actuators, SAE, 1997. Pages 15 – 21.

[33] Charrier, B., Constans, J., Cousteau, J., and et al, *Foundation Cousteau and Windship Propulsion 1980-1985*, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 20, Amsterdam, 1985. Pages 39-60.

[34] Cornell University RAFT project website, <http://fill in here>.

[35] Creamer, G., *Spacecraft Attitude Determination Using Gyros and Quaternion Measurements*, Journal of the Astronautical Sciences, Vol. 44, No. 3, Jul.-Sept. 1996. Pages 357-371.

[36] Curtiss, H.C., *On the Stability and Control of Sailing Yachts*, The Ancient Interface II, Proceedings of the Second AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 9, AIAA Los Angeles, 1970. Pages 21-32.

[37] Daif, A., Kamoun, B., Mudrey, M., *Evaluation of Wingsail Interactions for Windship Propulsion*, European Journal of Mechanics, B/Fluids. Vol. 10, 1985. Pages 185-206.

[38] Dodd, R.E., *A design analysis of 1974 world's multihull cup competitors*, The Ancient Interface VI, Proceedings of the Sixth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 15, AIAA Los Angeles, 1975. Pages 78-98.

[39] Drela, M., *A new transformation and integration scheme for the compressible boundary layer equations, and solution behavior at separation*, MIT GTL Report #172, MIT Press, May 1983.

[40] Drela, M., *XFOIL: An Analysis and Design System for Low Reynolds Number Airfoils*, Conference on Low Reynolds Number Airfoil Aerodynamics, University of Notre Dame, June 1989.

[41] Drela, M., XFOIL website <http://raphael.mit.edu/xfoil/>

[42] Edmunds, A., and Manners-Spencer, J., *The Wingsail Rig from Aerosystems*, International Conference on Sail-Assisted Commercial Fishing Vessels (ICSACFV), May 1983. Pages 410-424.

[43] Elkaim, G.H., Gebre-Egziabher, D., Powell, J.D., Parkinson, W.B., *A Gyro-Free Quaternion-Based Attitude Determination System Suitable for Implementation using Low-Cost Sensors*, IEEE Position Location and Navigations Symposium (IEEE PLANS), San Diego, 2000. P. 185-192.

[44] Elkaim, G.H., Gebre-Egziabher, D., Powell, J.D., Parkinson, W.B., *A Non-Linear, Two-Step Estimation Algorithm for Calibrating Solid-State Strapdown Magnetometers*, 8th St. Petersburg Conference on Navigation Systems (IEEE/AIAA), St. Petersburg, Russia, May 27-May 31, 2001.

[45] Elkaim, G.H., O'Connor, M.L., Parkinson, W.B., *System Identification and Robust Control of a GPS-guided Farm Tractor*, Proceeding of the Institute of Navigation, ION-GPS, Sept. 16-19, 1997.

[46] Enge, P., Walter, T., Pullen, S., Kee, C., Chao, Y.C. and Tsai, Y.J., *Wide Area Augmentation of the Global Positioning System*, Proceedings of the IEEE, 84(8):1063-1088, August 1996. Paper Number 0018-9219.

[47] Fathy, H. K., Thomson, J. G., and Gorder, P. J., *Direct Determination of Aircraft Quaternion Parameters Using a Combined Inertial-GPS Instrumentation System*, Proceeding of the Institute of Navigation ION-GPS Conference, Nashville, TN, ION-GPS Sept. 14-17, 1999.

[48] Fekete, G. I., Newmann, B. G., *Design and Testing of a Sailboat with Self-Trimming Wingsail*, Canadian Aeronautics and Space Journal, Vol. 29, No. 2, June 1983. Pages 121-130.

[49] Flettner, A., The Story of the Rotor, F.O. Willhoft, New York, 1926.

[50] Forsell, B., Radionavigation Systems. Prentice Hall, New York, New York, 1991.

[51] Fossen, T.I., Guidance and Control of Ocean Vehicles, Wiley and Sons, New York, NY., 1994.

[52] Franklin, G. F., Powell, J. D., and Workman, M. L., Digital Control of Dynamic Systems, Addison Wesley, Menlo Park, CA, 1994.

[53] Friedland, B., *Analysis Strapdown Navigation Using Quaternions*, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-14, No. 5, Sept. 1978. Pages 764-768.

[54] Gai, E., Daly, K., Harrison, J., and Lemos, L., *Star-Sensor-Base Satellite Attitude/Attitude Rate Estimator*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 8, Sept.-Oct. 1985. Pages 560-565.

[55] Gebre-Egziabher, D., Hayward, R., and Powell, J. D., *A Low Cost GPS/Inertial Attitude Heading Refernce System (AHRS) for General Aviaiton Applications*, Proceedings of the IEEE Position Location and Navigation Symposium, PLANS 1998, IEEE 1998. Pages 518-525.

[56] Gebre-Egziabher, D., *Design and Performance Analysis of a Low-cost Dead-reckoning Navigator for General Aviation Applications*, Stanford Ph.D. Thesis, December 2001, Chapter 5.

[57] Gelb, A., Applied Optimal Estimation, MIT Press, Cambridge, MA 1974.

[58] Gifford, E. W. H., *The Tunny Rig, A Variable Chamber Flexible Wing Sail*, International Conference on Sail-Assisted Commercial Fishing Vessels (ICSACFV), May 1983. Pages 170-190.

[59] Glauert, H., The elements of aerofoil and airscrew theory, Cambridge University Press, New York, 1983.

[60] Greenwood, D. T., Principles of Dynamics, 2nd Edition, Prentice Hall, New Jersey, 1988.

[61] Hammitt, A. G., *Optimum Wind Propulsion*, The Ancient Interface, Proceedings of the First AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 8, AIAA Los Angeles, 1969. Pages 133-150.

[62] Haupt, G. T., Kasdin, N. J., Keiser, G. M., and Parkinson, B. W., *Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 19, No. 3, May-Jun. 1996. Pages 643-649.

[63] Hayward, R., Gebre-Egziabher, D., Schwall, M., Powell, J. D., and Wilson, J., *Inertially Aided GPS Based Attitude Heading Reference System (AHRS) for General Aviation Aircraft*. Proceedings of the Institude of Navigation ION-GPS Conference, ION-GPS, Sept. 16-19, 1997. Pages 1415-1424.

[64] Hayward, R., Marchick, A., and Powel, J. D., *Single Baseline GPS-Based Attitude Heading Reference System (AHRS) for Aircraft Applications*, Proceedings of the American Control Conference, San Diego, CA, ACC, June 2-4, 1999.

[65] Hayward, R., Marchick, A., and Powel, J. D., *Two Antenna GPS Attitude and Integer Ambiguity Resolution for Aircraft Applications*, Proceedings of the Institude of Navigation National Technical Meeting ION-NTM, San Diego, CA, ION-NTM, Jan 25-27, 1999.

[66] Herreshoff, H. C., *Hydrodynamics and Aerodynamics of the Sailing Yacht*, Transaction of the Society of Naval Architects and Marine Engineers (SNAME), New York, Vol. 72, 1964. Pages 445-492.

[67] Herreshoff, H. C., and Newman, J. N., *The Study of Sailing Yachts*, Scientific American, Aug. 1966, Pages 60-68.

[68] Hook, C., *Why Sailing Hydrofoils*, The Ancient Interface IV, Proceedings of the Fourth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 11, AIAA Los Angeles, 1973. Pages 82-91.

[69] Hook, C., *A self-tending rig with feedback and compass course*, The Ancient Interface VI, Proceedings of the Sixth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 15, AIAA Los Angeles, 1975. Pages 135-142.

[70] Hook, C., *Large Self-Tending Rig for Ocean Transport*, The Ancient Interface VIII, Proceedings of the Eigth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Stanford, 1977. Pages 141-148.

[71] Hubbard, J. D., and Brescoll, G. P., *Aerodynamic Investigation of a Cup Anemometer*, NACA Technical Note #502, Washington, D.C., 1934.

[72] Huffman, D., *Shadowfax*, The Ancient Interface V, Proceedings of the Fifth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 12, AIAA San Diego, 1974. Pages 107-117.

[73] Idan, M., *Estimation of Rodrigues Parameters from Vector Observations*, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-32, No. 2, Apr. 1996. Pages 578-586.

[74] Irbitis, K., *The Wing Sail*, Canadian Aeronautical Institute Log, March 1955. Pages 12-14.

[75] Johnson, W. J., *A Basic Theory for Wind-Driven Vehicles,* The Ancient Interface II, Proceedings of the First AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 9, AIAA Los Angeles, 1970. Pages 1-8.

[76] Juang, J.-N., Applied System Identification, Prentice Hall, NJ, 1994.

[77] Juang J.-N. Cooper, J.E., and Wright, J.R., *An Eigensystem Realization Algorithm Using Data Correlations (ERA/DC) for Modal Parameter Identification*, Control-Theory and Advanced Technology, Vol. 4, No. 1, 1988.

[78] Kasdin, N. J., New *Guaranteed Positive Time Update for the Two-Step Optimal Estimator*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 23, No. 2, Mar.-Apr. 2000. Pages 215-221.

[79] Kasdin, N. J., and Haupt, G. T., *Second-Order Correction and Numerical Considerations for the Two-Step Optimal Estimator*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 20, No. 2, Mar.-Apr. 1997. Pages 362-369.

[80] Kayton, M. and Fried, W.R., Avionics Navigation Systems. John Wiley and Sons, Inc., Ney York, New York, 2nd edition, 1997.

[81] Kocivar, B., *Rotating Wingsail changes shape to move Cat Fast*, Popular Science, Feb. 1972, Page 104.

[82] Koifman, M., and Bar-Itzhack, I. Y., *Intertial Navigation System Aided by Aircraft Dynamics*, IEEE Transactions on Control Systems Technology, Vol. 7, No. 4, July 1999. Pages 487-493.

[83] Kornfeld, R. P., Hansman, R. J., and Dyest, J. J., *Single Antenna GPS Based Attitude Determination*, Proceedings of the Institude of Navigation National Technical Meeting ION-NTM, Long Beach, CA, Jan. 21-23, 1998.

[84] Kroo, I., PANDA website <http://www.desktopaero.com/PANDACatalogPage.html>

[85] Kroo, I., McGeer, T., *Optimization of Canard Configurations*, 13th ICAS Congress, ICAS-82-6.8.1, 1982

[86] Kuipers, J.B., Quaternions and Rotation Sequences, Princeton University Press, Princeton, NJ, 1999.

[87] Lefferts, E. J., Markley, F. L., and Shuster, M. D., *Kalman Filtering for Spacecraft Attitude Estimation*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 5, Sept.-Oct. 19852. Pages 417-429.

[88] Ljung, L., System identification : theory for the user, 2nd Ed., Prentice Hall, NJ, 1999.

[89] Machine Design Staff, *Airfoil Sailboat*, Machine Design, Vol. 40, Oct 1968. pp. 52.

[90] Mason, P.A., *A complete mathematical model of sailcraft performance*, The Ancient Interface VI, Proceedings of the Sixth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 15, AIAA Los Angeles, 1975. Pages 35-50.

[91] Mason, P.A., *Full Scale Tow and Heel Resistance Test Procedures*, The Ancient Interface VI, Proceedings of the Sixth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 15, AIAA Los Angeles, 1975. Pages 51-57.

[92] Marchaj, C. A., Aero-hydrodynamics of sailing, 3rd Edition, Adlard Coles Nautical, London, 2000.

[93] Marchaj, C. A., Sail performance: techniques to maximize sail power, International Marine, Camden, ME, 1996.

[94] Marchaj, C. A., Sailing Theory and Practice, Dodd and Mead Co., New York, 1964.

[95] Markley, F. L., *Attitude Determination using Vector Observations: A Fast Optimal Matrix Algorithm*, The Journal of the Astronautical Sciences, Vol. 41, No. 2, Apr.-Jun. 1993. Pages 261-280.

[96] Markley, F. L., *Attitude Determination using Vector Observations and the Singular Value Decomposition*, The Journal of the Astronautical Sciences, Vol. 36, No. 3, Jul.-Sep. 1988. Pages 245-258.

[97] McCrary, P. K., *Performance characteristics of Ice-boats and an Ideal Windmill Vehicle*, The Ancient Interface V, Proceedings of the Fifth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 12, AIAA San Diego, 1974. Pages 36-44.

[98] McGeer, T., Kroo, I., *A Fundamental Comparison of Canard and Conventional Configurations*, Journal of Aircraft, Nov. 1983

[99] Montgomery, P., *Carrier Differntial GPS as a Sensor for Automatic Control*, Ph.D. Thesis, Stanford University, Stanford, CA 1996. Pages 158-188.

[100] Murdin, D., *Standby Attitude Estimation*, IEEE Position Location and Navigations Symposium (IEEE PLANS), San Diego, 2000. Pages 429-436.

[101] Myers, H. A. and Krushkov, U. S., *Multihull Dynamics in Wind and Waves*, The Ancient Interface XII, Proceedings of the Twelfth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Seattle, 1981. Pages 152-163.

[102] Nance, C. T., *Outlook for Wind Assistance*, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 20, Amsterdam, 1985. Pages 1-17.

[103] Newman, B. G., and Fekete, G. I., *Analysis and Preliminary Design of of a Sailboat with a Self-Trimming Wingsail*, Marine Technology, Vol. 20, No. 4, October 1983. Pages 370-376.

[104] Ott, M., *Volantis - Hawaii 2000*, HydroWing Corporation business plan. <http://www.hydrowinghawaii.com/CompanyFrameB.htm>

[105] Overschee, P., and d'Moor, B., Subspace identification for linear systems : theory, implementation, applications, Klewar Academic Publishers, Boston, MA 1996.

[106] Parkinson, B. W. and Spilker, J. J., Global Positioning System: Theory and Applications, Vol. 1 and 2. AIAA, Washington, D.C., 1996.

[107] Parkinson, B.W. *GPS Error Analysis*. In Parkinson, B.W. and Spilker, J.J., Global Positioning System: Theory and Applications, Vol. 1, pages 469-483. AIAA, Washington, D.C., 1996.

[108] Payne, F. A., *YATEK - a Concept for a High-Speed Sailboat*, The Ancient Interface XIV, Proceedings of the Fourteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Long Beach, 1984. Pages 111-116.

[109] Pelmatic Windship Corporation, Pelmatic website, <http://www.pelmatic.de/>

[110] Phan, M., Horta, L.G., Juang, J.-N., and Longman, R.W., *Improvement of Observer/Kalman Filter Identification (OKID) by Residual Whitening*, Journal of Vibration and Acoustics, Vol. 117, No. 2, 1995.

[111] Popular Science Staff, *Four Foils to the wind drive a way-out sailboat*, Popular Science, Oct. 1968. Page 95.

[112] Popular Science Staff, *Wingsail Catamaran Undergoes Tests*, Popular Science, Jan. 1976, p. 75.

[113] Priebe, P.D., *The evolution of commercial sailing ship technology*, The Ancient Interface XI, Proceedings of the Eleventh AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Seattle, 1981. Pages 19-42.

[114] Psiaki, M. L., *Attitude Determination Filtering via Extended Quaternion Estimation*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 23, Mar.-Apr. 2000. Pages 206-214.

[115] Psiaki, M. L., Martel, F., and Pal, P. K., *Three-Axis Attitude Determination via Kalman Filtering of Magnetometer Data*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 13, May-June 1990. Pages 506-514.

[116] Rekow, A. K., *System identification, adaptive control and formation driving of farm tractors*, Stanford Ph.D. Thesis, Stanford, 2001.

[117] Rodgers, K. H., *Advanced Sailboat Patents*, The Ancient Interface XIII, Proceedings of the Thirteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA San Diego, 1983. Pages 138-149.

[118] Ross, J. C., *Aerodynamic Design of a Rigid-Wing Sail for a C-Class Catamaran*, The Ancient Interface XVIII, Proceedings of the Eighteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 35, AIAA Stanford, 1989. Pages 17-27.

[119] Savage, P. G., *Strapdown Intertial Navigation Integration Algorithm Design Part 1: Attitude Algorithms*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 21, No. 1, Jan.-Feb. 1998. Pages 19-28.

[120] Savage, P. G., *Strapdown Intertial Navigation Integration Algorithm Design Part 2: Velocity and Position Algorithms*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 21, No. 2, Mar.-Apr. 1998. Pages 208-221.

[121] Schlieben, E.W., *SKAMP - An Amazing Unmanned Sailboat!*, Ocean Industry, Sept. 1969. Pages 38-43.

[122] Selig, M. S., Donovan, J. F., Fraser, D. B., Airfoils at Low Speeds, H.A. Stokley, Virginia Beach, VA 1989.

[123] Shevell, R. S., Fundamentals of Flight, Prentice-Hall, NJ 1983, Page 170.

[124] Shibata, M., *Error Analysis Strapdown Inertial Navigation Using Quaternions*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 9, No. 3, May-Jun. 1986. Pages 379-381.

[125] Shortall, J. W., *World Trends in Sail-Assisted Commercial Fishing Vessels*, The Ancient Interface XIII, Proceedings of the Thirteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA San Diego, 1983. Pages 92-124.

[126] Shuster, M. D., and Oh, S. D., *Three-Axis Attitude Determination from Vector Observations*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 4, No. 1, Jan.-Feb. 1981. Pages 70-77.

[127] Sigurdson, D., and Ott, M., *Irens, Hubbard, Bradfield collaborate on 60' Hydrofoil trimaran*, New England Multihull Association Magazine, NEMA, January 2000, Page 10.

[128] Siouris, G., Aerospace Avionics Systems, Academic Press, San Diego, 1993. Pages 51-69.

[129] Smith, B., *The Aerohydrofoil*, The Ancient Interface II, Proceedings of the Second AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 9, AIAA Los Angeles, 1970. Pages 151-169.

[130] Smith, R., *SKAMP - Robot Boat with Rigid Sails Patrols Ocean Beat*, Popular Science, May 1970. Pages 70-71,142.

[131] Sorensen-Viale, N.E., *Motor and Sailing Bulk Carrier*, The Ancient Interface XI, Proceedings of the Eleventh AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Seattle, 1981. Pages 43-71.

[132] Spens, P. G., DeSaix, P., Brown, P. W., *Some Further Experimental Studies of the Sailing Yacht*, Transaction of the Socienty of Naval Architects and Marine Engineers (SNAME), New York, Vol. 75, 1967. Pages 79-111

[133] Stengel, R.F., Optimal Control and Estimation, Dover Publications, New York, NY, 1994.

[134] Stoeckert, T. M., *Sailboat Wings through Water*, Product Engineering, 15-July 1968, Pages 8.

[135] Thiessen, W. C., *Comprehensive Applications: Modern American Developmental Sail Powered and Sail Assisted Cargo Vessels*, The Ancient Interface XIV, Proceedings of the Fourteenth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Long Beach, 1984. Pages 117-125.

[136] U.S. Coast Guard NAVCEN website, http://www.nis-mirror.com/systems/dgps/index.html, 1999.

[137] University of Furtwangen, Relationship website, <http://www.fh-furtwangen.de/english/index.html>

[138] Wahl, P., *Cruise Aboard This Space-Age Sailing Ship*, Popular Science, June 1971. Pages 48-49.

[139] Walker, J., *Walker Wingsail Systems*, <http://www.walkerwing.com>, 1997.

[140] Walker, J., *A High Performance Automatic Wingsail Auxiliary Propulsion System for Commercial Ships*, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 20, Amsterdam,Sept. 1985. Pages 83-96.

[141] Walker, J., *Wingsail - Auxiliary Ship Propulsion System*, International Conference on Sail-Assisted Commercial Fishing Vessels (ICSACFV), May 1983. Pages 444-458.

[142] Warden, W. H, A *Control System Model for Autonomous Sailboat Navigation*, IEEE, 1991, p. 944-947.

[143] Watanabe, T., Endo, Y., Nakanishi, K. and Takeda, K., *Sail-Equipped Motor Ship Shin Aitoku Maru and Studies on Larger Ship* [sic], The Ancient Interface XII, Proceedings of the Twelfth AIAA Symposium on the Aer/Hydronautics of Sailing, AIAA Lecture Series, Vol. 23, AIAA Seattle, 1982. Pages238-249.

[144] Wahba, G., *A Least Squares Estimate of Satellite Attitude*, SIAM Review, Vol. 7, No. 3, Jul.-1965. Page 409. [Solution given in No. 4, Sept. 1965. Pages 384-386]

[145] Wolfhard, L., CAN System Engineering: From Theory to Practical Application, Springer Verlag, New York, NY, 1997.