



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

TASKSPACE CONTROL WITH A 3D ULTRASOUND POSITION SENSOR

Hans-Wolter Wehn

Department of Electrical Engineering
McGill University, Montréal

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

September 1990

© Hans-Wolter Wehn



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-67629-9

Canada

Abstract

The reduction of the kinematic error for precise trajectory tracking of a robotic manipulator's end effector pose is best achieved by taskspace sensory feedback. To this end, a new taskspace control scheme is introduced which has two main features: Firstly, a hierarchical control structure which feeds back both, joint and taskspace variables and thereby reduces uncertainty better than other schemes. Secondly, a nonlinear decoupling scheme which best linearizes the taskspace loop.

The manipulator's end effector pose is measured by a 3D ultrasound range finder, for which a very detailed deterministic and stochastic model is obtained from experimental data. Some of this information is then used in an extended Kalman filter to compensate for the range finder's imperfections.

Finally, the 3D ultrasound range finder, the extended Kalman filter and the new taskspace control scheme are simulated in concert in a realistic environment, to assess the control system's ability to reduce the kinematic tracking error.

Résumé

La réduction d'erreurs cinématiques pour l'observation précise de trajectoires du poignet du manipulateur robotique est mieux réalisée par l'utilisation d'une rétroaction à l'aide de capteurs dans l'espace Cartésien. Dans cette thèse, nous introduisons un nouveau système commande, qui comporte deux caractéristiques principales:

Premièrement, une structure hiérarchique de control rétroactive sur les articulations et les variables opérationnelles réduisant mieux, de ce fait, les incertitudes que d'autres systèmes. Deuxièmement, un système de découplage non-linéaire qui linéarise mieux la boucle dans l'espace Cartésien.

La position du poignet du manipulateur est mesurée par un télémètre à ultra-son tri-dimensionnel pour lequel un modèle déterministe et stochastique très détaillé est obtenu à partir de données expérimentales. Certaines de ces données sont ensuite utilisées dans un filtre de Kalman afin de compenser les imperfections du télémètre.

Finalement, le télémètre à ultra-son tri-dimensionnel, le filtre de Kalman et ce nouveau système commande dans l'espace Cartésien sont simulés dans un environnement réaliste afin d'estimer la capacité de réduction d'erreurs cinématiques d'observation du système commande.

Acknowledgements

I would like to express my deepest gratitude and appreciation to my research adviser, Prof. Pierre R. Bélanger. His excellent guidance, help and encouragement were essential for the completion of my research.

I would like to thank James Owen, Lin Lin and Le Yi Wang for their help and suggestions regarding those parts of the thesis which deal with H_∞ theory and robust control.

I would like to thank Christian Consol for his assistance with the French translation of the abstract and I would like to thank Meyer Nahon and Greg Gordon for proof-reading the manuscript.

Claim of Originality

This research presents new results regarding robot manipulator taskspace trajectory control, and 3D ultrasound position measurement instrumentation. To the best of my knowledge, the following are original contributions:

- The resolved position loop is introduced as a means to globally decouple and linearize a taskspace loop.
- The double servo loop which feeds back both the joint angles and the taskspace position is recognized to be a superior structure for reducing model uncertainty.
- The combination of the resolved position linearization and the double servo loop is shown to be robustly stable, and robustly performing in the presence of modeling errors including neglected flexible modes.
- A sequential control design procedure for the double loop is presented.
- An experimentally verified analysis of all relevant environmental influences on indoor 3D ultrasound precision ranging for distances between 20 cm and 2 m is given.
- An experimentally verified stochastic model describing the spatial and temporal statistics of the 3D ultrasound measurement noise is given.

- An extended Kalman filter in combination with a reference Kalman filter is proposed to partially compensate for the noisy ultrasound position measurements.
- A simulation backed analysis of the limitations on kinematic error reduction obtainable by a combination of the resolved position linearization, the double servo feedback loop, the 3D ultrasound range finder and the extended Kalman filter is presented.

Contents

Abstract	ii
Résumé	iii
Acknowledgements	iv
Claim of Originality	v
1 Introduction	1
1.1 Reducing the Kinematic Error	1
1.2 Pose Measurement with Ultrasound	5
1.3 Noisy Pose Measurements	8
1.4 Literature Review	11
I The Taskspace Control System	14
2 The Open Loop System	15
2.1 The Plant	15
2.1.1 The Plant Components	15
2.1.2 The Joint Loop	16
2.1.3 The Limiter	20

2.1.4	The Forward Kinematics	21
2.2	Approximate Inverse Kinematics	23
2.2.1	Four Approximate Inverse Kinematics Functions	23
2.2.2	Computational Complexity	26
3	The Closed Loop System	28
3.1	Nonlinear Taskspace Control Schemes	28
3.1.1	Norms, Gains and Definitions	28
3.1.2	The Resolved Position Loop	30
3.1.3	The Rate Linearized Loop	36
3.1.4	The Position Linearized Loop	38
3.2	Linear Compensator Design	41
3.2.1	Hierarchical Control	41
3.2.2	Continuous Time Design	44
3.2.3	Discrete Time Design	49
II	The 3D Ultrasound Position Sensor	53
4	The Hardware	54
4.1	Introduction to Part II	54
4.2	Range Finder Overview	59
4.2.1	A TOF Measurement Scheme	59
4.2.2	Choosing a TOA	63
4.3	Range Finder Limits	65
4.3.1	Pulse Rate Limits	65
4.3.2	Electronic Noise	67
4.3.3	Electronic Delays	68
4.4	Non-ideal Transducers	69

4.4.1	Theoretical Analysis	69
4.4.2	Numerical Analysis	73
5	Deterministic Measurement Corruption	77
5.1	Atmospheric Biases and Drifts	77
5.1.1	Room Temperature Drift	77
5.1.2	Temperature Gradients	79
5.1.3	Wind Biases	88
5.2	Wave-Medium Interaction	88
5.2.1	Diffraction	88
5.2.2	Refraction	89
5.2.3	Ambient Noise	92
5.2.4	Reflection and Reverberation	93
5.2.5	The Doppler Effect	97
6	Turbulences	99
6.1	The Stochastic Turbulence Model	99
6.2	Experimental Evidence	105
6.2.1	Time Correlation	105
6.2.2	Space Correlation	108
III	Taskspace Control with Noisy Measurements	117
7	Noise Compensation Methods	118
7.1	Optimal Microphone-Sender Configurations	118
7.2	The Kalman Filter	121
7.2.1	Overview	121
7.2.2	The Reference Filter	122

7.2.3	The Main Filter	123
7.2.4	Exploiting Spatial Correlation	129
7.3	Calibration	130
8	Simulations	133
8.1	Simulation Parameters	133
8.1.1	Test Trajectories	134
8.1.2	Model Uncertainties	135
8.1.3	Measurement Noise	137
8.2	Noise Free Simulations	140
8.2.1	Step Response	140
8.2.2	Circle Response	144
8.3	The EKF Performance	148
8.3.1	Nonlinear EKF Simulation	149
8.3.2	Covariance Analysis	155
8.3.3	Sensor Fusion	158
8.4	Closed Loop with Noise	160
9	Conclusions	163
9.1	Thesis Summary	163
9.2	Suggestions for Further Research	166
A	Covariance Approximation	174
B	The PUMA 600 Manipulator	176
C	The Ultrasound Transducer	178
D	Circuit Diagrams	180

List of Figures

1.1	Schematic diagram of sender and microphone configuration	6
2.1	The plant	15
2.2	The joint loop	18
2.3	An input-output representation of the closed joint loop	18
3.1	The resolved position loop	31
3.2	The small gain standard form	32
3.3	The rate linearized loop	36
3.4	The position linearized loop.	38
3.5	The double loop taskspace control system	41
3.6	The double loop with a discrete time regulator	50
4.1	Hardware block diagram	60
4.2	A typical received signal	62
4.3	Transducer misalignment	70
4.4	Numerically determined S-R misalignment shift for 7 mm sender.	73
4.5	Center shift function for 10 mm transducer	75
5.1	TOF fluctuations for 1.4m range while heating the small room.	78
5.2	Boundary conditions for model 1.	80
5.3	Vertical temperature for model 1.	81
5.4	Horizontal temperature for model 1.	82
5.5	Boundary conditions, isotherms and streamlines for Model 2.	83

5.6	Experimental vertical temperature profile of small room.	85
5.7	Dynamical behavior of temperature gradient.	86
5.8	Bad case scenario for approximation.	91
6.1	Distribution function of TOF fluctuations.	100
6.2	Time series of TOF fluctuations in small room.	106
6.3	Time series of TOF fluctuations in large room.	106
6.4	Power spectrum of TOF fluctuations in small and large room.	107
6.5	2nd and 3rd order system approximation	107
6.6	Longitudinal Correlation Experiment	108
6.7	Noise standard deviation versus distance for large room.	109
6.8	Noise standard deviation versus distance for small room.	110
6.9	Correlation versus distance ratio for large room.	112
6.10	Correlation versus distance ratio for small room.	112
6.11	Transversal Correlation Experiment	114
6.12	Correlation versus range angle in large room.	114
6.13	Correlation versus range angle in small room.	115
7.1	The optimal S-R configurations for 3 and 4 microphones.	120
7.2	The Kalman filter and the ultrasound system.	124
8.1	Step responses for system with uncompensated inner dynamics.	142
8.2	Step responses for system with compensated inner dynamics.	143
8.3	Control performance for TC1-KP1.	145
8.4	Control performance for TC3-KP2.	146
8.5	Control performance for TC4-KP2.	147
8.6	Compensated and uncompensated measurement noise	150
8.7	EKF performance for TC1-KP1-NC1.	153
8.8	EKF performance for TC2-KP1-NC1.	153
8.9	EKF performance for TC3-KP2-NC1.	154

8.10	EKF performance for TC4-KP2-NC1.	154
8.11	A Bode sketch of the Kalman filter.	156
8.12	Covariance analysis	159
8.13	Tracking error for TC1-KP1-NC1	162
8.14	Tracking error for TC4-KP2-NC1	162
B.1	The PUMA600 robot manipulator	177
C.1	The ultrasound transducer muRata MA40S2R	178
C.2	Ultrasound transducer frequency characteristic.	179
C.3	Ultrasound transducer radiation characteristic.	179
D.1	Block diagramm of the whole ultrasound ranging system.	181
D.2	Diagram of the TIMER circuit.	182
D.3	Diagram of the COUNTER circuit.	183
D.4	Block diagram of the RECEIVER circuit.	184
D.5	Diagram of the AMPLIFIER circuit.	184
D.6	Diagram of the TRIGGER circuit.	185
D.7	Diagram of the TIME WINDOW circuit.	185
D.8	Diagram of the ZERO CROSSING DETECTOR circuit.	186
D.9	Diagram of the SENDER circuit.	187
D.10	Latch circuit	188

List of Tables

2.1	Computational complexities of PUMA600	25
5.1	The effect of reverberation on the SNR.	95
7.1	Optimal microphone positions	120
8.1	Circular test trajectory parameters.	135
8.2	Kinematic error parameters.	137
8.3	Parameter sets for noise realizations.	140
8.4	Results of the nonlinear EKF simulation.	152
D.1	Truth table of latch circuit	188

Chapter 1

Introduction

1.1 Reducing the Kinematic Error

In the early days of robotics, robotic manipulators were mainly used for repetitive tasks in highly predictable and structured environments. Their movements were taught by physically moving the robot's end effector along a desired trajectory and recording the resulting joint angles. The error with which the robot can repeat these taught positions is often referred to as repeatability. Many commercial manipulators, with reasonably large work spaces, have repeatability errors around 0.1mm. This is small enough for most tasks to be performed successfully.

Now, however, many robots are used for more complex tasks in sensor monitored, changing environments. Often, an AI program [1] is used to coordinate the movements of the manipulator with a priori knowledge, multisensor data and the desired performance goals. The robot's movements will therefore depend on the environment and often can not be taught in advance. Therefore, the manipulator must be able to follow numerically specified trajectories.

The difference between the numerically commanded end effector position and the position actually reached is called the kinematic error. Unfortunately, the kinematic

error is much larger than the repeatability error: It can be large as 10mm [2], i.e. hundred times larger than the repeatability. The reason for the kinematic error is an ignorance of the robot's true kinematic function; a function which relates the joint variables to the Cartesian end effector position. Usually, only knowledge of the nominal kinematic function is available. It is derived from the nominal manipulator geometry. The true manipulator geometry, however, is a result of the tolerances allowed in the manufacturing of the robot components, and may differ considerably from the nominal geometry. Other influences include: Joint angle offset, gear backlash, gear transmission errors, joint drive compliance, base motion, shaft wobbling, bending torsion of link structures and temperature dependencies. The relative magnitude of some of these error sources was investigated in [3].

Robots with small kinematic errors have numerous advantages. Predominant among these is the ability to precisely track numerically generated trajectories, thereby helping the manipulator to perform complex tasks in changing environments. Other advantages are mentioned in [2]: A small kinematic error can eliminate most pendant teaching efforts. It also makes robot programs more readily transportable to other machines, and it eases multirobot coordination. Thus the reduction of the kinematic error is an important task.

The goal of this thesis is therefore to reduce the the kinematic error as much as possible. Ideally, the kinematic error is reduced to about the same size as the repeatability error, i.e. 0.1mm. In the remainder of the thesis, this precision will be referred to as the accuracy of interest and effects which are smaller will be considered neglectable.

A popular approach to reduce the kinematic error is to calibrate the manipulator. This is done by measuring several end effector positions and their corresponding sets of joint angles. The collected data are then analyzed to obtain the kinematic function of the particular robot. This approach has had some success in reducing the kinematic

error. For example, the static precision reported in [2] was better than 0.8mm for a PUMA760.

The disadvantages of robot calibration include that the modeling of the nongeometric errors is difficult and robot dependent. Even the usual temperature changes between night and day can yield kinematic errors of up to 0.3 mm [3]. Further, structural compliance effects due to unknown payloads cannot be compensated for. Often, the objective of manipulator control is to track a given trajectory. Evidently, static calibration cannot reduce the part of the tracking error which is due to possible changes of the kinematic function when the robot is moving. Another disadvantage is the computational complexity of calibration. The forward and inverse kinematic functions which incorporate the calibration results are bound to be very complicated. They often require numerical solutions even if the nominal kinematic functions have relatively simple closed forms. Since these functions may be used for real time control purposes, computational simplicity is very desirable.

A better way of reducing the kinematic error is taskspace position feedback: The pose of the manipulator end effector is continuously measured in Cartesian coordinates and this information is used by a suitable feedback scheme to reduce the kinematic error. Evidently, the question of how to achieve taskspace position feedback has two components: How to measure the 3-D end effector pose, and how to use these measurements to reduce the kinematic error.

There are many different technologies for measuring the end effector pose. Among the methods proposed are automated theodolites [4] which are extremely precise ($\pm 0.05\text{mm}$) but also extremely slow. Another method attaches photo diodes to the end effector and monitors their movement with the help of a laser scanner [5] which is not very precise ($\pm 2\text{mm}$). Yet another method attaches LEDs to the end effector and uses cameras to track them [6]. There are also commercially available devices which are based on this idea and which achieve good results: On a two meter range, the older

Watsmart system [61] achieves ca. 2 mm precision, and its successor, the Optotrak system [61] achieves ca. 0.3 mm. Both products allow sampling rates of several hundred Hz. Many of the cited methods require a substantial financial investment.

One objective of this thesis is to explore the advantages and limits of 3-D end effector position measurements based on ultrasound. The idea is to attach ultrasound senders to the end effector of the manipulator and to use a fixed microphone array to track their position. The advantages of such a scheme include:

- *Low cost hardware:* The prototype including the ultrasound transducers, the computer interface and all required analog and digital circuitry was built for under \$200.
- *High speed:* Pulse rates of more than hundred Hz can be achieved.
- *Small computational requirements:* An off-the-shelf single board computer with numeric coprocessor should be sufficient.
- *Good precision:* Depending on the circumstances, good submillimeter precisions can be achieved.

The potential uses for this type of measurement system include low-cost, high precision tracking applications. It may also be useful for sensor fusion applications, where the fast, low-cost ultrasound position measurements supplement slower and more expensive optical measurements. Another application is for environments which are antagonistic to optical methods, like dirty and dusty environments. Moreover, the work presented in this thesis could possibly be extended to applications in underwater robotics.

If used independent of the control system the 3D ultrasound position sensor can be used as a convenient and inexpensive tool for the usual parametric robot calibration. Alternatively, it can be used for nonparametric calibration where the kinematic

error history for a particular robot trajectory is recorded and later used for error compensation along this trajectory.

The new taskspace position feedback scheme introduced in this thesis stands on its own. The scheme is independent of the particular taskspace position measurement device used and therefore is not restricted to the 3D ultrasound range finder used in this theses. Its main features are:

- *A hierarchical servo controller* splits the overall loop uncertainty into a dynamic and a kinematic uncertainty. Hierarchical loops reduce uncertainty more effectively than single loops and are therefore superior to single loop schemes.
- *The resolved position* nonlinear kinematic decoupling scheme is introduced. It decouples and linearizes the kinematics globally and is superior to other methods. It is also very satisfying from a theoretical point of view.

Applications for the new taskspace controller include all situations where high performance sensory pose feedback is required.

1.2 Pose Measurement with Ultrasound

The key to kinematic error reduction, via feedback, are precise position measurements. The purpose of the ultrasound range finder is therefore to measure the position and orientation of the manipulator's end-effector as accurately as possible. The pose is represented with respect to an arbitrary but fixed 3-D reference coordinate system. In the following, a right-handed Cartesian reference coordinate system is assumed. Other choices are possible and may, in certain cases, simplify the kinematic function.

The sender position is measured by attaching an ultrasound transmitter to the end effector of a robot. The transmitter emits carefully timed bursts of sound which are received by distant microphones. Since sound propagates with a finite speed,

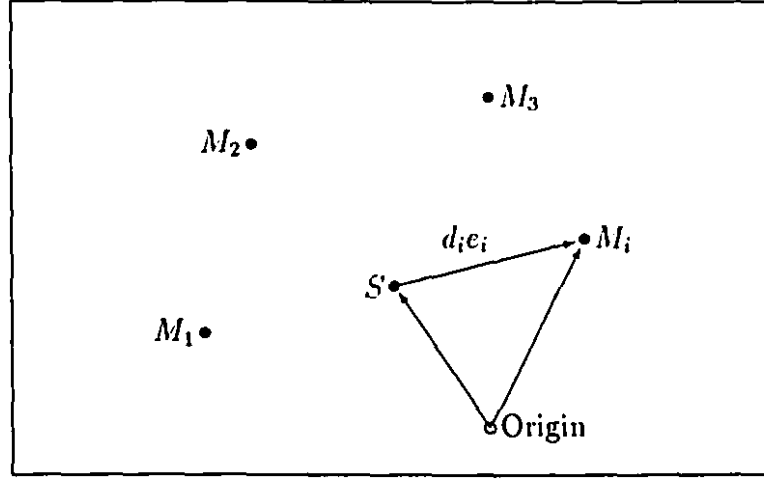


Figure 1.1: Schematic diagram of sender and microphone configuration

there will be a small time delay between the sending and the receiving of the bursts. When the speed of sound c is known, then the scalar distance d_i between the sender and the i -th receiver and the measured time delay tof_i are related as

$$\text{tof}_i = \frac{d_i}{c} \quad (1.1)$$

(During the remainder of this text, the measured time delay tof_i will often synonymously be denoted as “TOF” which is short for “time of flight”.)

In order to measure the sender position in 3-D space using TOF measurements an array of m microphones at well known fixed and distinct positions is needed. The position of the sender can then be calculated from the range measurements by triangulation. The Fig. 1.1 depicts the situation: M_i denotes the position of the i -th microphone, S denotes the position of the sender, and e_i denotes the unit direction from the sender to the i -th receiver. If the M_i are known a-priori and the d_i are measured, then the sender position S can be found by solving a set of nonlinear equations. Since

$$d_i^2 = (M_i - S)^T (M_i - S) \quad ,$$

one has to solve the following equations for S :

$$\begin{pmatrix} d_1^2 - M_1^T M_1 \\ \vdots \\ d_m^2 - M_m^T M_m \end{pmatrix} = \begin{pmatrix} -2M_1^T & 1 \\ \vdots & \vdots \\ -2M_m^T & 1 \end{pmatrix} \begin{pmatrix} S \\ S^T S \end{pmatrix} \quad (1.2)$$

Evidently, for $m = 4$ microphones (1.2) can be solved as four linear equations with four unknowns. Thus provided that the matrix on the r.h.s. has full rank, four microphones do guarantee a unique solution for S . The full rank condition translates into the requirement that not all four microphones should lie on the same plane. Some geometrical insight reveals the other cases: For $m = 3$ linear independent microphones, equation (1.2) yields a unique solution if S lies on the plane spanned by the 3 microphones. It yields two solutions if S lies outside of it. Usually, however, one has some crude a priori knowledge of the sender position and can easily exclude one of the two solutions. Thus, usually three microphones suffice for a unique determination of S . For $m = 2$ distinct microphones the solutions for S are constrained to lie on a circle. Finally, if there is just one microphone the solutions are constrained to lie on the surface of a sphere.

The orientation of the end effector can be calculated from position measurements if at least three senders instead of just one are attached to the end effector. For example, the end effector can be thought of as having a Cartesian, right-handed coordinate system attached to it. The origin of this coordinate system would then collocate with the position of the end effector and the coordinate system's orientation would be identical to the orientation of the end effector. For instance, the position of the origin P_o of the end effector coordinate system could be chosen to lie at the center of the triangle spanned by the 3 sender positions S_1, S_2, S_3 ,

$$P_o = \frac{1}{3}(S_1 + S_2 + S_3) \quad ,$$

and the three unit vectors o_1, o_2 and o_3 of the end effector coordinate system could be chosen to be:

$$\begin{aligned}
o_1 &= \frac{S_1 - S_2}{\|S_1 - S_2\|} \\
o_2 &= \frac{S_2 - (S_2^T o_1) o_1}{\|S_2 - (S_2^T o_1) o_1\|} \\
o_3 &= o_1 \times o_2
\end{aligned}$$

The above description is redundant. It requires 12 numbers to describe position and orientation. However, just 6 numbers suffice if other representations of orientation are used like roll-pitch-yaw, Euler angles, or the equivalent axis representation. However, these descriptions are not unique and can be numerically ill defined. A four number orientation representation like the quaternions could overcome these problems. A detailed study on the representation of orientation for taskspace feedback can be found in [18].

1.3 Noisy Pose Measurements

Unfortunately, the range measurements are rather noisy. This is primarily due to an ignorance of the true sound-speed as a function of time and space. Small variations of the sound-speed can blur the computation of position and orientation from TOF measurements, even if an average sound-speed is known. The uncertainty is in the order of 0.1% of the range. Thus, on a 1 m range one can expect an error in the order of 1 mm.

The reason for the fluctuating nonuniform sound-speed is its strong dependency on temperature and wind, and to a much lesser degree, its weak dependency on pressure and moisture. According to [28], the speed of sound in air can be approximated as

$$c = \sqrt{\left(\frac{\partial p}{\partial \rho}\right)_{\text{adiabatic}}} \quad (1.3)$$

where the partial derivative is evaluated for equilibrium conditions of the air pressure p and the air density ρ .

The most prominent factor which influences the speed of sound is the air temperature. According to [33]:

$$c = c_0 \sqrt{1 + \frac{\vartheta}{273^{\circ}C}} \approx c_0 + 0.6 \frac{m}{sec^{\circ}C} \vartheta \quad (1.4)$$

where $c_0=331.4$ m/sec is the speed of sound at $0^{\circ}C$ and ϑ is the temperature of the air in $^{\circ}C$. At $20^{\circ}C$ room temperature, the speed of sound is $c_{20}=341$ m/sec. It is evident from (1.4) that the sensitivity of c , with respect to the temperature, is about $0.18\%/^{\circ}C$. Thus, on a two meter range, an unaccounted for temperature change of just $0.03^{\circ}C$ causes a measurement error of ca. 317nsec or equivalently ca. 0.1mm which is the same as the accuracy of interest. In particular, the small temperature difference between the ceiling and the floor of a room must be taken into account.

The second major influence on the speed of sound is the presence of air movement. The equation (1.4) can be modified to account for wind:

$$c_i = c_0 \sqrt{1 + \frac{\vartheta}{273^{\circ}C}} + e_i^T w \quad (1.5)$$

where c_i is the speed of sound in the direction e_i^T , and w is the velocity vector of a constant and uniform wind. In the worst case, the wind blows in the direction e_i^T . In this case, the sensitivity of c_i with respect to deviations of the wind speed from zero is $0.29\%/(m/sec)$. Hence, on a 2 m range, an unaccounted for 1.7 cm/sec wind could cause a maximum error of the same magnitude as the accuracy of interest.

There are also several less important factors which influence the speed of sound. According to [33], there is a noticeable dependency on the barometric pressure for overpressures of more than 10bar. For instance, for 25bar overpressure the speed of sound increases by ca. 0.8%. Moreover, if the relative air humidity were to change drastically from 20% to 100%, then the speed of sound would change by 0.15%.

According to [13] the pressure and humidity dependences as well as any dependency on the absorption coefficient or the frequency of the sound can be neglected for normal operation with 40Khz ultrasound (which was used in this study).

The air is a turbulent medium even in a 'quiet' closed room . Convection currents, temperature drifts, drafts, heat conduction etc. are just some of the mechanisms which influence the temperature and wind distribution in a room and therefore the sound-speed. In this thesis, an attempt was made to model the effect of these fluctuations on the TOF measurements as a random process. The model distinguishes between two parts, a 'deterministic' part, which models the very slow and spacially highly correlated influences. and a 'random' part, which models faster and spacially less correlated fluctuations. Since these fluctuations are mainly caused by heat effects, the time constants involved are very large. In fact, the random part fluctuations have a time constant in the order of half a minute. This makes it difficult to reduce them by averaging them out. Hence, the thesis makes an attempt to model the time-space statistics of the fluctuations in an effort to use this information for other means of error reduction.

In order to find a good position estimate, despite the measurement noise, an extended Kalman filter (EKF) is used to track the sender position given the available a priori information and the noisy measurements. In addition, one or more fixed reference senders are employed to exploit the spacial correlation of the turbulences. The closer the moving sender comes to a reference sender the larger is the achievable error reduction. The sender position estimates, obtained from the EKF, are then used in the Cartesian feedback loop instead of the true sender positions. Given the nature of the problem, the low frequency positioning accuracy of the control system is about as good as the EKF position estimates. The latter depends on on the atmospheric conditions of the room, the positions of the reference senders and the trajectory of the moving senders.

1.4 Literature Review

There is a multitude of one dimensional ultrasound ranging methods. They can roughly be divided into two methods: Firstly, there are the continuous wave (CW) methods. They infer the range from sending an uninterrupted ultrasound signal. Secondly, there are gated wave (GW) methods which send out bursts of sound energy and infer range information from the measured time-of-flight (TOF) of the burst. The CW methods include phase shift measurements [8], amplitude modulation [7] and frequency modulation [9]. Unfortunately, CW-methods are extremely sensitive to reflections and perform especially poorly in closed rooms.

Probably, the most prominent representative of the GW-methods is the range finder developed by the Polaroid company [10]. This GW-method features thresholding of the lowpass filtered signal. Another similar scheme can be found in [11]. There are also GW-variations involving sophisticated pulse shape processing methods [12]. Nevertheless, the author of this thesis obtained the best experimental results for a GW scheme, using a combination of thresholding and zero crossing detection. This proved to be a simple but very effective method. This observation was shared by Lamancusa and Figueroa [13] and also by Sasaki, Takano and Akeno [14]. The latter claim an accuracy of 0.06% of the range, limited by air turbulences. This result is in line with our own observations.

Although there is a the wealth of literature on 1-D ranging, the literature on 3-D ultrasound position ranging is scarce. Lamancusa and Figueroa [13] addressed some of the problems connected with 3-D ultrasound ranging. In particular, they measured the position measurement error due to the finite transducer size, for some special situations. However, they did not address the space-time properties of the measurement fluctuations caused by turbulent air. A few of these properties are, at least, mentioned in [15]. Stone [63] used a commercially available 3D sonic digitizer

[62] for robot calibration. In a very controlled environment, after compensating for the sound speed drift and sound speed gradient he obtained an accuracy in the order of 0.1 mm on a 2 m range. This is an excellent result. The work presented in this thesis differs in that it stresses a low cost solution for on-line feedback in not so well controlled environments. Particularly, the effects of atmospheric turbulence were taken into account.

The amount of literature on taskspace feedback is impressive. Early schemes like the resolved rate control [16] use the inverse Jacobian to map the taskspace error into the joint error, other schemes [17] use the transpose Jacobian instead. However, these schemes are not appropriate for fast tracking because they neglect the nonlinear dynamics of the manipulator.

The standard solution to the taskspace control problem is to solve the ‘inverse problem’: Firstly, one writes the dynamic equation of the robot in terms of its taskspace variables instead of its joint variables. Secondly, one compensates for the nonlinear dynamics by nonlinear feed forward. Thirdly, one designs a servo controller for the remaining linear decoupled system feeding back the taskspace variables. Many variations [18]-[23] of this basic theme have been studied.

However, there is little recognition that in the servo part, feedback of both the taskspace variables and the joint variables, is inherently superior to feedback of the taskspace variables alone. Conceptually, one can think of such a feedback scheme as a double loop. For instance, a velocity feedback stabilized inner loop was used in [24]. The taskspace control results presented in this thesis were first reported in [25]. In particular, the resolved position loop was introduced as a kinematic linearization and decoupling scheme.

This thesis is organized in three main parts: In the first part the Cartesian control scheme is studied under the assumption that the end effector position can be measured perfectly. The second part describes the hardware of the 3D ultrasound range finder

and characterizes its imperfections. In particular, it develops a stochastic model of the measurement noise associated with TOF measurements. Finally, the last part simulates the Cartesian feedback control loop when the EKF's end effector position estimates are fed back instead of the true end effector positions. The simulations show what accuracy can be achieved under what circumstances and therefore define the limits to this approach to kinematic error reduction.

Part I

The Taskspace Control System

Chapter 2

The Open Loop System

2.1 The Plant

2.1.1 The Plant Components

The plant which the Cartesian loop is supposed to control can be modeled as the nonlinear dynamic operator

$$p(\cdot) = T \circ L_Q \circ (I + \Delta_m)(\cdot) \quad (2.1)$$

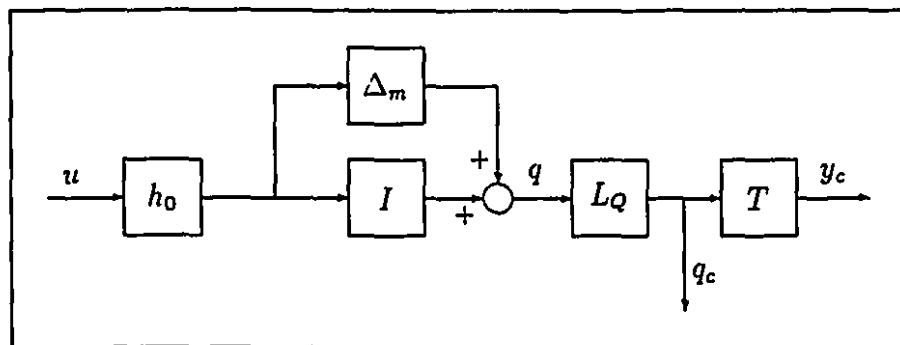


Figure 2.1: The plant

were ' $B \circ A$ ' symbolizes that the operator A is to be applied before the operator B . As shown in Fig. 2.1, the plant has an input u which is the desired joint position, and two outputs: the true Cartesian position y_c and the true joint position q_c . The plant $p(\cdot)$ consists of four parts:

1. The nominal linear decoupled time invariant closed joint loop dynamics $h_0(\cdot)$.
2. A nonlinear dynamic uncertainty $\Delta_m(\cdot)$ centered at unity. It represents the modeling error of the joint loop.
3. A limiter $L_Q(\cdot)$ which represents the range limitations of the joints.
4. The nonlinear, memoryless, true forward kinematic function T which maps joint displacements into taskspace coordinates.

In the following, each of the plant components will be discussed.

2.1.2 The Joint Loop

The most fundamental part of any robot control system is the open loop rigid body joint dynamics of a robot. It can be described as in [40]:

$$\tau = M(q_c)\ddot{q}_c + N(q_c, \dot{q}_c) \quad (2.2)$$

where τ denotes the joint torque vector, M denotes the inertia matrix, N is a nonlinear dynamic vector valued function which represents the centrifugal, Coriolis, gravity and friction torques. Since the joint displacement vector q_c is available as output of the optical shaft encoders, the joint loop can be closed from q_c back to τ . A well known control scheme is the so called computed torque method [40]. This control law has a nonlinear compensation part and a linear servo part: The first part compensates for all nonlinearities such that the nominal manipulator dynamics mimic a unit mass.

The second part is a PD servo controller which is supposed to cope with any remaining uncertainties. The computed torque control law is:

$$\tau = \hat{N}(q_c, \dot{q}_c) + \hat{M}(q_c)[K_v \dot{e} + K_p e] \quad (2.3)$$

where $e = u - q_c$, ($u = q_d$ is the commanded joint displacement), and $\hat{M}$ and $\hat{N}$ are approximations of M and N , respectively. The reasons for the use of approximations are an ignorance of the true manipulator mass properties, the true actuator constants, the correct friction model and the exact manipulator payload.

The ignorance of the true manipulator dynamics grows with the operating frequency. In particular, the rigid body assumptions made in (2.2) break down at a few Hz, and the robot starts to vibrate. According to Daneshmend [43], the first vibrational modes of a PUMA600 were found experimentally to lie around 20-30Hz. The actuators, however, impose an even stronger restriction on the bandwidth of a PUMA. Limits on the available power, joint torques and joint velocities, condemn the PUMA to have a bandwidth even smaller than its structural stability would allow. The gain of the joint loop must therefore roll off before flexible modes or power restrictions take effect. This puts a hard limit on the available bandwidth. A realistic value for the closed loop bandwidth of a PUMA600 was experimentally observed by Daneshmend [43] to be ca. 2Hz.

Nevertheless, a feedforward term like $\hat{M}(q_c)\ddot{u}$ is often added to the r.h.s of (2.3), yielding a two degree of freedom controller with a unity nominal closed loop transfer function. Of course, to enjoy the benefits of feedback requires the bandwidth of the trajectory to be smaller than the closed loop bandwidth of the robot.

The one-degree of freedom loop is sketched in Fig. 2.2, where $1/s^2$ symbolizes that the nominal open loop dynamics after nonlinear compensation and decoupling behaves like n parallel double integrators. Further, R_1 denotes a linear dynamic regulator, and Δ'_m is the nonlinear dynamic operator which represents the modeling

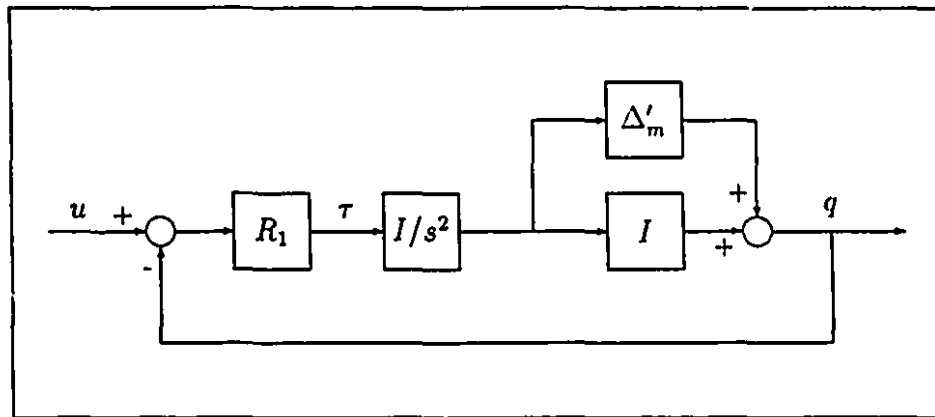


Figure 2.2: The joint loop

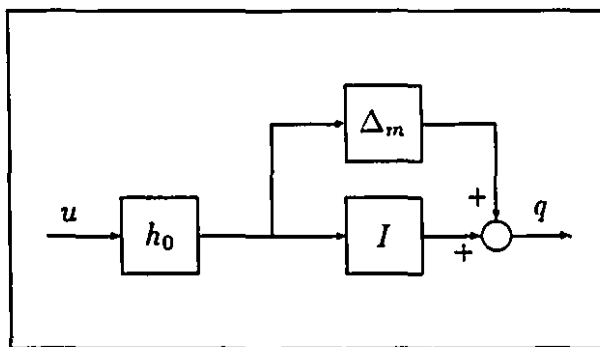


Figure 2.3: An input-output representation of the closed joint loop

errors including the flexible mode dynamics.

Another way to depict the joint loop of Fig. 2.2 is shown in Fig. 2.3, where $h_0(\cdot)$, which has the Laplace transform $H_0(s) = I/s^2 R_1(s)(I + I/s^2 R_1(s))^{-1}$, denotes the decoupled nominal closed loop dynamics. The symbol Δ_m represents a nonlinear dynamic uncertainty centered around unity. The difference between the uncertainties Δ_m and Δ'_m is that Δ_m is reduced by the sensitivity of the loop.

For the sake of argument, consider the loop in Fig. 2.2 to be SISO, linear and time invariant, and let $S_1(s)$ denote the Laplace transform of the true sensitivity function, and $S_{1n}(s)$ the nominal sensitivity function of the loop. The Laplace transform of Δ_m is Δ_M . Now one can write the Laplace transform of q as:

$$\begin{aligned} q(s) &= [1 - S_1(s)] u(s) \\ &= [1 - S_{1n}(s)] [1 + S_1(s)\Delta'_M(s)] u(s) \\ &= H_0(s) [1 + \Delta_M(s)] u(s) \end{aligned}$$

Thus, one obtains the input-output form shown in Fig. 2.3 or Fig. 2.1 with the transformed uncertainty:

$$\Delta_M(s) = S_1(s)\Delta'_M(s) \quad (2.4)$$

The sensitivity function S_1 is small at low frequencies, and approximately unity at high frequencies. Therefore, the effect of the joint loop is to shrink the original uncertainty Δ'_M in the low frequency range. However, at frequencies larger than the closed loop bandwidth, the original uncertainty is left more or less unchanged.

In the following, it is useful to assume that the closed loop uncertainty Δ_m can be represented as the sum of a dominant linear time invariant dynamic uncertainty Δ_l and a nonlinear dynamic uncertainty Δ_n , which is small compared to Δ_l :

$$\Delta_m(\cdot) = \Delta_l(\cdot) + \Delta_n(\cdot) \quad (2.5)$$

The decomposition (2.5) will be used later to limit the closed loop bandwidth of the Cartesian loop.

The representation of the closed joint loop given in Fig. 2.3 or Fig. 2.1 is not restricted to the computed torque method. In fact, any joint control method including adaptive control has a closed loop representation like the one depicted in Fig. 2.3. In many cases, adaptive controllers can shrink the uncertainty more effectively than non-adaptive controllers. If computational simplicity is desired, then the acceleration feedback law [26] is an elegant alternative to the computed torque method. It uses an additional $\ddot{e}$ term in (2.3) and high gain feedback instead of the nonlinear compensation terms. This reduces significantly the computational complexity of the joint controller. However, as before, at low frequencies, the closed loop dynamics can be considered to be approximately linear, time invariant, and decoupled. Moreover, Δ_m is small in this frequency range. At higher frequencies, however, the acceleration control law must roll off its gain before power restrictions and flexible modes become a problem. Hence, as before, at high frequencies Δ_m is large.

2.1.3 The Limiter

It is important for the design of the closed loop to be based on a clear understanding of the domain and the range of the kinematic function. For this reason the open loop plant is modeled to include a limiter function L_Q which is followed by the forward kinematic position function T . This is depicted in Fig. 2.1.

The limiter L_Q restricts the possible joint displacements to a set Q and therefore represents the physical limits of the joints: Prismatic joints have necessarily a finite range; rotational joints, on the other hand, can be unrestricted but are usually confined to ranges of less than 360° . Hence, let the range of the i -th joint be the sector $[a_i, b_i] \subset \mathcal{K}$ and let there be n joints, then $Q \subset \mathcal{R}^n$ is the n -cell

$$Q = \prod_{i=1}^n [a_i, b_i]$$

Occasionally, the set Q is referred to as the joint space. The limiter function can now

be described as

$$L_Q(q) = \begin{cases} q & \text{if } q_c \in Q \\ \arg \min_{q \in Q} \|q - q_c\| & \text{otherwise} \end{cases}$$

Strictly speaking, L_Q is a part of the joint loop. However, here it is modeled as being outside to show the properties of the taskspace loop more clearly.

2.1.4 The Forward Kinematics

The true forward kinematic function T maps the joint space Q onto the Cartesian workspace $\mathcal{X} \subset \mathcal{R}^n$, which is the set of all physically obtainable Cartesian positions and orientations. In other words,

$$\mathcal{X} = T(Q)$$

It is evident, however, that T , with domain Q and range $\mathcal{X}$, is not a 1-1 mapping. In general, there can be many joint positions which correspond to a given Cartesian position. They correspond to different configurations of the robot. For example, many positions of the PUMA 600 wrist can be obtained with four different configurations, namely: 1) left arm - elbow up, 2) left arm - elbow down, 3) right arm - elbow up, 4) right arm - elbow down. A picture of these four configurations can be found in [40].

The complete 6 joint PUMA600 has eight configurations. In general, the exact number of possible configurations depends on the robot architecture, and the degrees of freedom of the robot. The nominal T can be systematically derived with the help of the Denavit-Hartenberg parameters (D-H parameters [35]) as is shown, for instance, in [39].

Besides T , there is also interest in its derivative $J(q_c)$, called the Jacobian matrix of the manipulator, which maps joint rates into Cartesian velocities $\dot{y}_c = J(q_c)\dot{q}_c$,

where:

$$J(q_c) = \left. \frac{\partial T(q)}{\partial q} \right|_{q=q_c \in Q}$$

There are special vectors $q_s \in Q$ which are called the singularities of the manipulator.

They have the following well known properties:

- The Jacobian matrix $J(q_s)$ is singular.
- The manipulator loses a degree of freedom in the Cartesian space.
- Singular joint positions are those for which two joint axes align or two links align.
- Singularities mark the boundary between two configurations. For example, when the PUMA600 arm is stretched out completely it is in a singular position and can be considered to be both in the elbow up and in the elbow down configurations.

Let the set of all singularities be denoted $Q_s \subset Q$.

Let $Q_i \subset Q$ denote the set of all joint position vectors q_c which belong to the i -th configuration and let the manipulator have n_c configurations. Let the configuration sets Q_i include their singularities (i.e include their boundary and make them closed sets). They have the following properties:

$$\mathcal{X}_i = T(Q_i) \text{ is 1-1 and onto.}$$

where $\mathcal{X}_i \subset \mathcal{X}$ is the part of the Cartesian workspace which can be reached while the robot is in the i -th configuration. Note that it is possible that certain Cartesian positions in the workspace cannot be reached in all configurations because of the restrictions on Q . Furthermore:

$$Q = \bigcup_{i=1}^{n_c} Q_i$$

$$Q_s =: \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n_c} (Q_i \cap Q_j)$$

In the previous text, the forward kinematic function T was presented as describing a strictly geometric relation which maps vectors into vectors. It should be kept in mind, however, that this is only a good approximation. The true T has also a weak dynamic component which takes the flexibility of the manipulator into account.

2.2 Approximate Inverse Kinematics

2.2.1 Four Approximate Inverse Kinematics Functions

The forward kinematic position function T is highly nonlinear. In order to apply linear control design methods to the Cartesian loop it is therefore necessary to compensate for T with some kind of inverse kinematics function. Unfortunately, the true T is not known exactly because the D-H parameters of the manipulator's geometry are not known exactly, and because there are kinematic effects like compliance due to gravitational loading which require knowledge in addition to the D-H parameters. Given this ignorance, the best one can do is to use an approximate inverse kinematics function (AIKF). Since the inverse of T can be a very complex function, it is often desirable to trade accuracy for computation speed and use a deliberate simplification as AIKF. The following three AIKFs will be discussed in more detail: The approximate inverse kinematic position function (AIKPF), the approximate inverse kinematic velocity function (AIKVF) and the approximate kinematic force function (AIKFF). In connection with the use of a Kalman filter, there is also interest in generating a nominal trajectory by computing the approximate forward kinematics position function (AFKPF).

The computational complexity of the AIKFs depends on the structure of the manipulator. Most modern manipulators have structures that allow for a relatively simple closed form solution of the nominal inverse kinematics problem. In particular, many robots have the nominal kinematic equivalent of a spherical wrist and neighboring joint axis are nominally oriented at 0° or 90° relative to each other. For these types of manipulators, there exist highly efficient nominal closed form inverse kinematic solutions like the Featherstone method [36]. The Stanford manipulator, and the PUMA are examples of this type of manipulator.

If the manipulator structure does not have a closed form inverse, then one must compute the inverse numerically with convergence methods. There are efficient and numerically reliable algorithms like the one proposed by Angeles in [37] which can accomplish this task. However, numerical solutions are usually much more time consuming than closed form algorithms.

An AIKF must find a balance between accuracy on one side and convenience and computational complexity on the other. The most accurate AIKF would incorporate all available information on the kinematic function of the manipulator. Such information could be obtained through calibration measurements. The disadvantages of this method, however, include the inconvenience and difficulty of precisely measuring a large number of kinematic parameters. They also include that the AIKF would likely not have a closed form even if the nominal manipulator does have an easy closed form inverse kinematics solution. An alternative approach is to use the nominal kinematic functions and let the sensitivity function of the kinematic feedback loop reduce any existing errors.

Let T_i^{-1} denote the true inverse kinematic function for the i -th configuration and let the AIKPF $\hat{T}_i^{-1}$ denote an approximation of T_i^{-1} , i.e.:

$$\begin{aligned} Q_i &= T_i^{-1} \circ T(Q_i) \\ \hat{T}_i^{-1} &\approx T_i^{-1} \end{aligned}$$

	a	m	f
AIKPF	17	15	19
AIKVF	29	48	17
AFKPF	16	21	19
AIKVF+AFKPF	37	55	20

Table 2.1: Computational complexities of the nominal kinematic functions of a PUMA600, using the Featherstone method.

The AIKPF $\hat{T}_i^{-1}$ maps Cartesian positions into joint positions.

An alternative AIKF, the AIKVF, maps velocities instead of positions. This approach can more easily handle arbitrary robot architectures. A prominent AIKVF is $\hat{J}^{-1}(q_c)$, an approximation of the true inverse Jacobian matrix $J^{-1}(q_c)$:

$$\hat{J}^{-1}(q_c) \approx J^{-1}(q_c)$$

The true inverse Jacobian $J^{-1}(q_c)$ does not exist for $q_c \in \mathcal{Q}_s$. The approximate Jacobian $\hat{J}^{-1}(q_c)$, on the other hand, can be made to exist even at singularities. It will, however, be an infinitely bad approximation at those points.

There are at least four established ways to compute

$$\dot{q}_c = \hat{J}^{-1}(q_c)\dot{y}_c$$

The first method is to compute $\hat{J}^{-1}(q_c)$ and then multiply $\dot{y}_c$; the second method is to solve $\hat{J}(q_c)\dot{q}_c = \dot{y}_c$, the third method is given in [44] and is an improved version of the second method, and the fourth method is the Featherstone method which was mentioned earlier. The latter method requires a special robot architecture.

The four methods are progressively more efficient with the fourth method leading by a wide margin. A comparison between them, which was conducted by Hollerbach in [38] for the Stanford manipulator, found the Featherstone method to be about three times as efficient as the second best method.

2.2.2 Computational Complexity

Since most modern manipulators allow the application of the Featherstone method, and since it appears to be the most efficient method available, it will be used in this thesis as a benchmark for computational complexity. Hence, the author of this thesis investigated the complexity of the Featherstone method for a nominal 6 joint PUMA600. The results are displayed in Table 2.1, where 'a', 'm' and 'f' stand for additions, multiplications and transcendental functions, respectively. A similar study can be found in [41].

The closed loop will include an extended Kalman filter which requires a nominal trajectory as one of its inputs. In this context, it is useful to investigate the complexity of computing simultaneously the AIKVF and the AFKPF.

The complexity for the simultaneous computation is smaller than for the separate computations because most transcendental functions and some intermediate results can be utilized by both functions. These translate into substantial savings. The required complexity for the simultaneous computation is just $37a+55m+20f$ (as shown in the last row of Table 2.1) whereas the separate computation of both functions requires $45a+69m+36f$. The simultaneous computation of AIKPF and AFKPF, however, does not permit savings of this nature. The savings are even more pronounced, if the computed torque methods are employed for the joint loop, because then the intermediate results can also be shared by the dynamic feedforward equations. Some aspects of this idea are discussed in [42].

A look at Table 2.1 suggests that the most time consuming factor is the computation of the occurring transcendental functions. Hence, in order to further reduce the computational complexity, one could replace the occurring transcendental functions by simpler functions, lookup tables, functions over a simpler number field, dedicated VLSI circuits or a combination of these methods.

Moreover, the AIKFs will be used in a real time closed loop environment. In particular, suppose the resolved rate Cartesian loop [16] is used, but the multiplication of the error vector by the inverse Jacobian matrix is replaced by Featherstones' method. The transcendental functions of the AIKVF are not functions of the Cartesian velocity or perturbation, but strictly functions of the current joint angles. This suggests the possibility of reducing the computational load by computing the transcendental functions at a lower sampling rate than the rest of the AIKVF.

If the manipulator does not have an efficient closed form inverse solution, one could compute the vectors $\hat{T}_i^{-1}(y)$ off-line by numerical methods for some $y \in \mathcal{X}_i$, and store them. One could then extrapolate between the stored values in real time for any given Cartesian position in $\mathcal{X}_i$. An alternative is to use an AIKFF, the approximate transpose Jacobian matrix $\hat{J}^T(q_c)$ as AIKF. It maps Cartesian forces into joint torques. The complexity of this method depends only on the efficiency of the computation of the Jacobian. Several methods are compared in [45]. Unfortunately, AIKFFs perform poorly in trajectory control loops.

Chapter 3

The Closed Loop System

3.1 Nonlinear Taskspace Control Schemes

In the following, three nonlinear closed loop control schemes will be presented: The resolved position loop, the rate linearized loop and the position linearized loop. These schemes linearize and decouple the Cartesian feedback loop. When they are implemented, the remaining control design problem reduces to a linear decoupled servo problem. All three methods have in common that they operate with an embedded joint loop of the kind introduced in Sec. 2.1.2.

The methods differ in that the resolved position loop compensates for the kinematic function globally, while the other two schemes compensate only locally. Consequently, there are trajectories and disturbances which cause the local schemes to perform poorly or even to be unstable. The resolved position loop, on the other hand, can maintain stability and performance for all trajectories and disturbances.

3.1.1 Norms, Gains and Definitions

Definitions: Let $L_2^n[0, \infty)$ be the usual vector valued continuous time function space

with norm

$$\|x(t)\|_{L_2^n}^2 = \int_0^\infty x(t)^T x(t) dt < \infty$$

For any fixed t , $x(t) \in \mathcal{R}^n$. As in [46], let $(\cdot)_\tau$ denote the truncation operator, i.e.

$$(x(t))_\tau = \begin{cases} x(t) & \text{if } t < \tau \\ 0 & \text{otherwise} \end{cases}$$

The extended space $L_{2e}^n[0, \infty)$ is the space consisting of all those functions whose truncations lie in $L_2^n[0, \infty)$. Furthermore, let $\mathcal{B}$ be the class of those operators $f(\cdot)$ on L_{2e}^n having the property that the zero element, denoted 0, lies in the domain of $f(\cdot)$ and $f(0) = 0$.

Let the operator gains for $f(\cdot) \in \mathcal{B}$ be:

$$\begin{aligned} \bar{g}(f(\cdot)) &= \sup_{\substack{\forall x \in L_{2e}^n \\ x \neq 0}} \frac{\|(f(x))_\tau\|_{L_2^n}}{\|(x)_\tau\|_{L_2^n}} \\ \tilde{g}(f(\cdot)) &= \sup_{\substack{\forall x, y \in L_{2e}^n \\ x - y \neq 0}} \frac{\|(f(x) - f(y))_\tau\|_{L_2^n}}{\|(x - y)_\tau\|_{L_2^n}} \\ \underline{g}(f(\cdot)) &= \inf_{\substack{\forall x \in L_{2e}^n \\ x \neq 0}} \frac{\|(f(x))_\tau\|_{L_2^n}}{\|(x)_\tau\|_{L_2^n}} \end{aligned}$$

The operator gains defined above have the following three properties:

Property 1: If $h(\cdot) \in \mathcal{B}$ is linear then $\tilde{g} = \bar{g}$. If, in addition, it is also stable and has a proper Laplace transfer function $H(s)$, then

$$\bar{g}(h(\cdot)) = \sup_{\forall \omega} \bar{\sigma}(H(j\omega))$$

$$\underline{g}(h(\cdot)) = \inf_{\forall \omega} \underline{\sigma}(H(j\omega))$$

where $\bar{\sigma}(\cdot)$ and $\underline{\sigma}(\cdot)$ denote respectively the maximum and minimum singular value of a matrix.

Property 2: If $h(\cdot) \in \mathcal{B}$ is memoryless, then the gains can be defined over the simpler space $\mathcal{R}^n$ rather than L_{2e}^n . Hence, $(\cdot)_r$ can be ignored and any references to L_2^n or L_{2e}^n can be replaced by $\mathcal{R}^n$ in the definitions of $\bar{g}$, $\underline{g}$ and $\hat{g}$.

Property 3: Let $h(\cdot), f(\cdot) \in \mathcal{B}$ and let $I(\cdot) \in \mathcal{B}$ be the identity operator, then

$$\bar{g}(f(\cdot) + h(\cdot)) \leq \bar{g}(f(\cdot)) + \bar{g}(h(\cdot))$$

$$\underline{g}(I(\cdot) + h(\cdot)) \geq 1 - \bar{g}(h(\cdot))$$

$$\bar{g}(f \circ h(\cdot)) \leq \bar{g}(f(\cdot))\bar{g}(h(\cdot))$$

$$\underline{g}(f \circ h(\cdot)) \geq \underline{g}(f(\cdot))\underline{g}(h(\cdot))$$

In the following it will be assumed that $\Delta_m \in \mathcal{B}$. This can be justified from (2.4) because the joint loop sensitivity is zero or at least very small at $\omega = 0$. However, this assumption could be removed at the price of a more tedious discussion.

3.1.2 The Resolved Position Loop

The resolved position loop is shown in Fig. 3.1. The box R_2 refers to a linear time invariant regulator which has the transfer function $R_2(s)$. The other parts of the loop have already been introduced in Sec. 2.1. The limiter L_Q has been omitted from Fig. 3.1 for the sake of simplicity.

The point of the resolved position loop is that it uses all available kinematic information to linearize and decouple the nonlinear taskspace loop. It causes therefore the smallest possible uncertainty and can for this reason be considered superior to

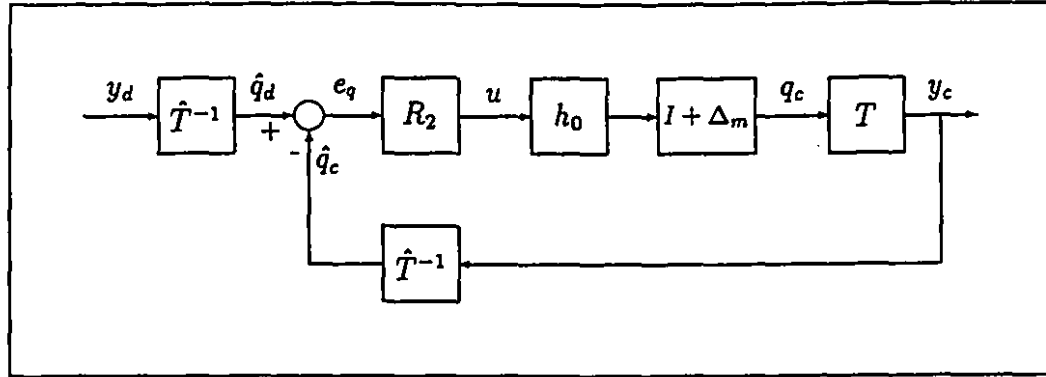


Figure 3.1: The resolved position loop

other schemes. The resolved position loop was first reported in [25]. In the following, the stability and the performance of this control scheme will be investigated.

Preconditions for Stability

Before a more formal discussion can be undertaken, some preconditions must be stated. If these are violated, the resolved position loop will be rendered unstable. The first of these conditions is that it must be ensured that in Fig. 3.1 $q_c \in Q_i$, because it is possible to choose trajectories y_d for which $q_c \notin Q_i$. Assume that q_c is very close to a boundary point of Q_i . This can lead to two types of instabilities:

Type 1: If the boundary point of Q_i is also a boundary point of Q and a disturbance tries to move q_c beyond the confines of Q , then the limiter L_Q confines q_c to Q . If for instance, the set point of the loop $\hat{q}_d$ happens to lie outside Q and the loop contains an integrator, then the limiter action will lead to integrator windup. This is because the integrator would continue to integrate an error which the limiter would prevent from being reduced.

Type 2: If the boundary point of Q_i is not a boundary point of Q , but a boundary point of Q_j , which is a neighboring inverse kinematic solution set, then a disturbance

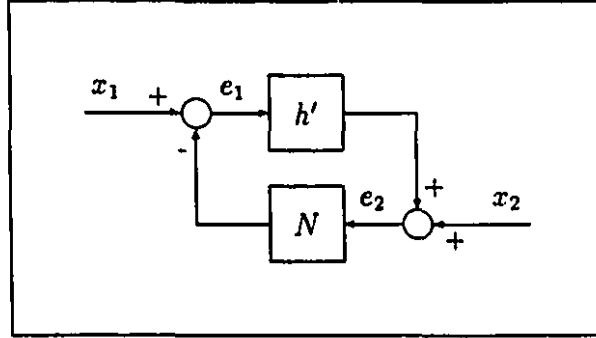


Figure 3.2: The small gain standard form

can cause q_c to cross through the singularity, i.e. leave Q_i and enter Q_j . This is equivalent to the manipulator changing its configuration. Even for a perfect approximation, i.e. $\hat{T}^{-1} = T_i^{-1}$, the operator $\hat{T}^{-1}T(\cdot)$ in Fig. 3.1 would map elements of Q_j into elements of Q_i if $q_c \in Q_j$. This would certainly cause some joints to move in the wrong direction. In other words, it would cause the loop's nonlinearity to be indefinite and thereby provoke instability.

The easiest remedy, against both types of instabilities, is to choose only those trajectories $y_d(t)$ which ensure that $q_c \in Q_i$ and $\hat{q}_d \in Q$.

Robust Stability

When the loop transformation theorem [46] is applied to the loop in Fig. 3.1, then the standard form shown in Fig. 3.2 is obtained, where $h' \in \mathcal{B}$ denotes the nominal closed loop dynamics with the associated transfer function matrix:

$$H'(s) = H_0(s)R_2(s)(I + H_0(s)R_2(s))^{-1} \quad (3.1)$$

Furthermore, N in Fig. 3.2 denotes a nonlinear dynamic operator which represents the multiplicative uncertainty of the loop:

$$N(e_2) = \hat{T}^{-1} \circ T \circ (I + \Delta_m)(e_2) - \hat{T}^{-1} \circ T(0) - e_2 \quad (3.2)$$

Note that $N(\cdot) \in \mathcal{B}$. The inputs x_1 and x_2 in Fig. 3.2 are:

$$\begin{aligned} x_1 &= \hat{T}^{-1}(y_d) - \hat{T}^{-1} \circ T(0) \\ x_2 &= 0 \end{aligned}$$

For simplicity, it was assumed above that $q_c = 0$ is in the domain of $\hat{T}^{-1} \circ T(\cdot)$. This assumption could be removed by choosing slightly different N , x_1 and x_2 . Moreover, let

$$\hat{T}^{-1} = T_i^{-1} + \Delta \hat{T}^{-1}$$

and let

$$N_0(\cdot) = \Delta \hat{T}^{-1} \circ T(\cdot) - \Delta \hat{T}^{-1} \circ T(0)$$

Note that $N_0 \in \mathcal{B}$. Now one can write (3.2) as:

$$N(e_2) = (I + N_0) \circ (I + \Delta_m)(e_2) - e_2 \quad (3.3)$$

According to the small gain theorem [46], the loop depicted Fig. 3.2 is BIBO stable if

1. h' is stable.
2. $\bar{g}(h')\bar{g}(N) < 1$

If, in addition, the second condition holds when $\bar{g}$ is replaced by $\tilde{g}$, then the loop is input-output stable, which is BIBO stable and continuous.

The first condition is easily met because R_2 is chosen such that the nominal Cartesian loop h' is stable. An easy check as to whether the second condition is met, is obtained by roughly decomposing $\bar{g}(N_0)$ into its linear, dynamic part and its nonlinear part. Now it is possible to check the second condition via property 1 and property 2 of gains. From (3.3):

$$N(\cdot) = N_0 \circ (I + \Delta_m)(\cdot) + \Delta_m(\cdot) \quad (3.4)$$

taking gains on both sides of the equation above and using (2.5) yields

$$\begin{aligned}\bar{g}(N) &\leq \bar{g}(N_0)(I + \bar{g}(\Delta_m)) + \bar{g}(\Delta_m) \\ &\leq \bar{g}(N_0)(I + \bar{g}(\Delta_l) + \bar{g}(\Delta_n)) + \bar{g}(\Delta_l) + \bar{g}(\Delta_n)\end{aligned}$$

Suppose the linear time invariant unstructured uncertainty Δ_l has a Laplace transform $\Delta_L(s)$ and suppose one knows a scalar function $l_m(\omega)$ such that for all ω

$$\bar{\sigma}(\Delta_L(j\omega)) \leq l_m(\omega)$$

Under the assumptions made in (2.5) suppose that for all ω it is true that $\bar{g}(\Delta_n) \ll \bar{\sigma}(\Delta_l(j\omega))$.

Thus similar to [48], if for all ω

$$\bar{\sigma}(H'(j\omega)) < \frac{1}{\bar{g}(N_0)(1 + l_m(\omega) + \bar{g}(\Delta_n)) + l_m(\omega) + \bar{g}(\Delta_n)} \quad (3.5)$$

we can conclude that the system is BIBO stable. Moreover, if $\bar{g}(N_0)$ and $\bar{g}(\Delta_n)$ are replaced by $\hat{g}(N_0)$ and $\hat{g}(\Delta_n)$ and the inequality still holds, then the system is input-output stable.

The inequality (3.5) constitutes a hard limit for the nominal closed loop bandwidth. The limit is enforced by the loop uncertainties. At low frequencies, $l_m(\omega) \approx 0$, and therefore, $1/\bar{g}(N_0)$ restricts the magnitude of peaks of the nominal closed loop frequency response. This is the MIMO equivalent to the avoidance of a critical disk in the Nyquist plot [47]. At higher frequencies $l_m(\omega)$ is large and therefore H' is primarily restricted by $(\bar{g}(N_0)l_m(\omega) + l_m(\omega))^{-1}$, which fixes the bandwidth and the required roll-off at high frequencies. This concludes the discussion on robust stability of the resolved position loop.

Robust Performance

Let $l_0(\cdot)$ be the nominal open loop transfer operator with Laplace transfer function $L_0(s) = H_0(s)R_2(s)$, and let $S^{-1}(\cdot)$ denote the inverse of the true sensitivity function

of the outer loop:

$$S^{-1}(\cdot) = I(\cdot) + (I + N_0) \circ (I + \Delta_m) \circ l_0(\cdot) \quad (3.6)$$

This can be manipulated as follows

$$\begin{aligned} \underline{g}(S^{-1}) &= \underline{g}(S^{-1} \circ (I + l_0)^{-1} \circ (I + l_0)) \\ &\geq \underline{g}(S^{-1} \circ (I + l_0)^{-1}) \underline{g}((I + l_0)) \\ &\approx \underline{g}((I + N_0) \circ (I + \Delta_m)) \underline{g}(l_0) \\ &\geq [1 - \bar{g}(N_0)][1 - \bar{g}(\Delta_m)] \underline{g}(l_0) \end{aligned}$$

The approximation step in the derivation above is valid for $\underline{g}(l_0) \gg 1$. This is only feasible of course if we compute $\underline{g}(l_0)$ via property 1 and restrict the ω over which the sup is taken to the lower frequency range. As in [48], we want the minimum gain of the true sensitivity of the system to be smaller than a scalar performance function $1/ps(\omega)$. We can conclude that the system is robustly performing in the presence of the uncertainties, if

$$\underline{\sigma}(L_0(j\omega)) \geq \frac{ps(\omega)}{[1 - \bar{g}(N_0)][1 - l_m(\omega) - \bar{g}(\Delta_n)]} \quad (3.7)$$

The above inequality is of course only valid in the low frequency range, when $l_m(\omega) < 1$ and $\underline{\sigma}(L_0(j\omega)) \gg 1$.

It shows what gain the nominal open loop must at least have in order to meet the performance requirements $ps(\omega)$ in the presence of uncertainties. The nominal open loop gain must be chosen to be larger than would be necessary without uncertainties. For a fixed gain-bandwidth product, this means a reduction of bandwidth. This concludes the discussion on robust performance.

Error Equivalence

When we say that the loop is robustly performing, then we refer to the reduction of the joint space error $\|e_q\| = \|\hat{q}_d - \hat{q}_c\|$. The purpose of the Cartesian loop, however, is

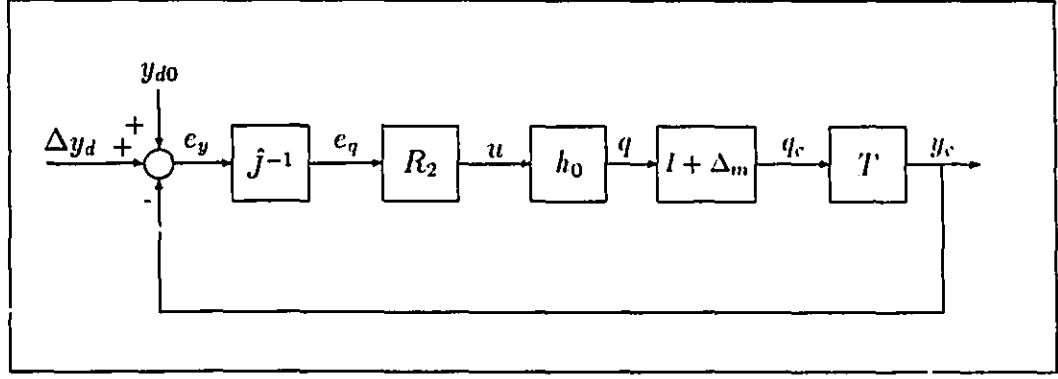


Figure 3.3: The rate linearized loop

to reduce the error $\|e_y\| = \|y_d - y_c\|$. Let $\hat{T}^{-1} : \mathcal{X}_i \rightarrow \mathcal{Q}_i$ be a function that is 1-1 and onto. It is evident from the definition of a function that $(x_d = x_c) \implies (\hat{T}^{-1}(x_d) = \hat{T}^{-1}(x_c))$, or $\|e_y\| = 0 \implies \|e_q\| = 0$. Further, it is evident from the definition of bijective mappings that $(x_d = x_c) \iff (\hat{T}^{-1}(x_d) = \hat{T}^{-1}(x_c))$, or

$$\|e_y\| = 0 \iff \|e_q\| = 0,$$

The previous result is not satisfying. There is a need to establish that convergence in the joint space implies convergence in the Cartesian space. Hence, if in addition to being 1-1 and onto, $\hat{T}^{-1}$ is continuous and $\mathcal{X}_i, \mathcal{Q}_i$ are compact, then by definition of continuity: $(\|e_y\| \rightarrow 0) \implies (\|e_q\| \rightarrow 0)$. The converse is established via Theorem 4.17 in [53], which states that under the above conditions the inverse of $\hat{T}^{-1}$ is also continuous. We can therefore conclude that:

$$(\|e_y\| \rightarrow 0) \iff (\|e_q\| \rightarrow 0)$$

3.1.3 The Rate Linearized Loop

An alternative scheme to the resolved position loop is the rate linearized loop. It is shown in Fig. 3.3. The use of the approximate inverse Jacobian in a taskspace loop was first proposed by Whitney in [16] and is referred to as the resolved rate loop. The

loop in Fig. 3.3 differs from [16] in featuring an embedded closed joint loop which improves the overall dynamic performance.

The approximate inverse Jacobian $\hat{J}^{-1}$ can, of course, be replaced by any of the AIKVs discussed in Sec. 2.2. The vector $y_{d0} = T(q_{d0})$ is constant and Δy_d is a small desired deviation from y_{d0} . The rate linearized loop can be analyzed with essentially the same tools which were used for the resolved position loop. The nonlinearity of the rate linearized loop can be written as:

$$N(q_{d0}, q) = \hat{J}^{-1}(q_{d0})[T(q_{d0}) - T \circ (I + \Delta_m)(q)]$$

when the above equation is linearized around $q = q_{d0}$ and the Laplace transform is taken one obtains

$$N'_{q_{d0}}(s)E_1(s) = \hat{J}^{-1}(q_{d0})J(q_{d0})(I + \Delta_M(q_{d0}, s))E_1(s) \quad (3.8)$$

where $E_1(s)$ is the Laplace transform of the perturbation vector $e_1(t) = q_{d0} - q(t)$ and $\Delta_M(q_{d0}, s)$ is the Laplace transfer function matrix of the derivative of $\Delta_m(\cdot)$ at q_{d0} . A good approximation is $\Delta_M(q_{d0}, s) \approx \Delta_L(s)$. Let

$$\begin{aligned} \hat{J}^{-1} &= J^{-1} + \Delta J^{-1} \\ N_0(q_{d0}) &= \Delta J^{-1}(q_{d0})J(q_{d0}) \end{aligned}$$

then from (3.8) one obtains for the loop uncertainty with good approximation

$$N''_{q_{d0}}(s) = (I + N_0)(I + \Delta_L) - I$$

this is similar to (3.4).

Consider the standard small gain form of Fig. 3.2 with H' and N exchanged. Now, H' is as in (3.1) and $N = N''_{q_{d0}}$ and

$$\begin{aligned} x_1 &= q_{d0} \\ x_2 &= q_{d0} + \hat{J}^{-1}(q_{d0})\Delta y_d \end{aligned}$$

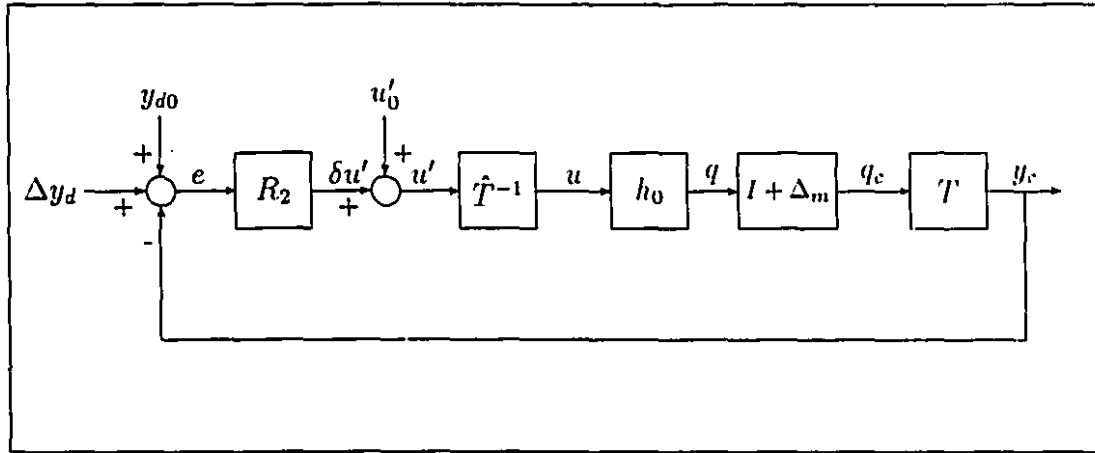


Figure 3.4: The position linearized loop.

One would like to find a controller R_2 such that (3.5) and (3.7) are satisfied for all setpoints q_{d0} in the workspace (with the possible exception of small neighborhoods around singular values). Therefore, when computing gains of N_0 the sup and inf must be taken over all desired set points q_{d0} .

If the equilibrium point q_{d0} is not constant, but is changing more slowly than the largest time constant of the closed loop, then the above stability analysis would still remain valid. On the other hand, no guarantees can be given for faster trajectories. However, it can be argued that faster trajectories would not be very practical in the first place because the use of feedback has no beneficial effects for frequency ranges where the sensitivity is larger than 1.

The preceding analysis requires that q_c be in the domain of T . This can most easily be accomplished by choosing R_2 to include an integrator with appropriate initial conditions.

3.1.4 The Position Linearized Loop

The position linearized loop is essentially the integral form of the rate linearized loop and can be analyzed in an analogous manner. Alternatively, the analysis can be

conducted in the Cartesian space. Consider the nonlinearity which is a part of the closed Cartesian loop shown in Fig. 3.4:

$$y_c = T \circ (I + \Delta_m) \circ h_0 \circ \hat{T}^{-1}(\delta u' + u'_0) \quad (3.9)$$

where u'_0 is such that $y_{d0} = T \circ \hat{T}^{-1}(u'_0)$, where y_{d0} is a constant vector. In practice, if R_2 contains an integrator, then u'_0 are its initial conditions.

If one linearizes the nonlinearity about u'_0 , takes the Laplace transform and approximates Δ_M one obtains

$$\delta Y_c(s) = J(q_0)(I + \Delta_L(s))H_0(s)\hat{J}^{-1}(q'_0)\delta U'(s)$$

where $q'_0 = T_i^{-1}(u'_0)$ and $q_0 = \hat{T}^{-1}(u'_0)$ and $\delta U'(s)$ and $\delta Y_c(s)$ are the Laplace transforms of $\delta u'(t)$ and $\delta y_c(t) = y_c(t) - y_{d0}$. Letting $H_0(s) = h'_0(s)I$, where $h'_0(s)$ is a scalar function, one obtains the loop uncertainty

$$N(s) = (J(q_0)\hat{J}^{-1}(q'_0) - I) + J(q_0)\Delta_L(s)\hat{J}^{-1}(q'_0)$$

The linearized loop can be manipulated into the form given in Fig. 3.2. As before, H' is the nominal closed loop dynamics and the loop uncertainty is N . The comments made in Sec. 3.1.3 on slowly changing y_{d0} , and on taking gains, are also valid for the position linearized loop. Thus taking gains yields

$$\bar{g}(N(\omega)) \leq \bar{g}(J(u_0)\hat{J}^{-1}(u_0) - I) + \bar{g}(J(u_0))\bar{g}(\hat{J}^{-1}(u_0))l_m(\omega)$$

As before, the robust stability condition is:

$$\bar{\sigma}(H'(j\omega)) < \frac{1}{\bar{g}(N(\omega))} \quad (3.10)$$

which fixes the bandwidth of the linearized loop.

Comparison between the three loops

Precision: The resolved position loop achieves the best use of kinematic information, i.e the true $\bar{g}(N)$ is smaller than for the two other methods. Hence, it can achieve better trajectory tracking.

Allowed Trajectories: Both the rate linearized and the position linearized loop method are local methods and require small loop errors to work properly. This restricts the allowed input trajectories to those that do produce small errors. The resolved position loop, on the other hand, has no such trajectory restriction. It guarantees global stability.

Need for a Trajectory Generator: Because they require a small loop error, the rate linearized and the position linearized loop method need a trajectory generator which specifies a point of the desired trajectory for each sampling instant. Often one is only interested in the end position, but not in the exact intermediate path. Since the resolved position loop can handle large loop errors, it allows one to simply specify the end position. The resolved position loop will automatically generate a smooth trajectory leading to any given end position.

Complexity: For a PUMA robot employing the Featherstone method and without considering the Kalman filter, the position linearized loop is the least complex method, requiring only $17a+15m+19f$ for the inverse kinematics, it is followed by the rate linearized loop with $29a+48m+17f$ and the resolved position loop with $2(17a+15m+19f)$, which is the most complex. If the trigonometric functions of the rate linearized loop are not computed at every sample instance, then the resolved rate loop is the least complex of the three methods.

One can conclude that the resolved position loop has considerable advantages over the other two methods. The price for these benefits is an increased complexity. The rate linearized loop, on the other hand, can offer the smallest numerical complexity.

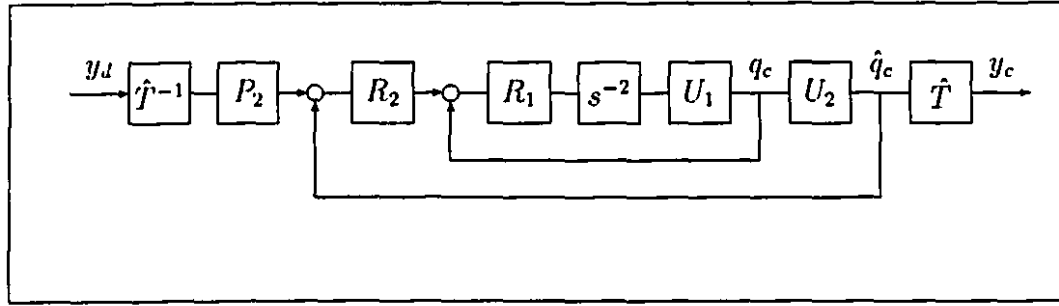


Figure 3.5: The double loop taskspace control system

3.2 Linear Compensator Design

3.2.1 Hierarchical Control

The design philosophy used in this thesis can best be illustrated with the help of Fig. 3.5, which represents a form of the resolved position loop Fig. 3.1 with Fig. 2.2 inserted. As can be seen, the overall system features a double loop: An inner or joint loop and an outer or kinematic loop. The inner loop feeds back the joint angle measurements q_c of the optical shaft encoders and is identical to the joint loop shown in Fig. 2.2. The outer loop feeds back the Cartesian position measurements y_c as was shown in Fig. 3.1. However, to formally obtain a unit feedback system, the outer loop in Fig. 3.5 is shown to feed back $\hat{q}_c = \hat{T}^{-1}(y_c)$.

Consequently, U_1 and U_2 are defined as:

$$\begin{aligned} U_1 &= (I + \Delta'_m)(\cdot) \\ U_2 &= (I + \Delta \hat{T}^{-1}T)(\cdot) \end{aligned}$$

For convenience's sake, U_1 and U_2 will be referred to as uncertainties, keeping in mind that only $\Delta'_m(\cdot)$ and $\Delta \hat{T}^{-1}T(\cdot)$ truly deserve that name. As mentioned before, inaccurately known manipulator mass properties, actuator constants, friction etc. do not allow a complete compensation of the nonlinear dynamics and give rise to the dynamic uncertainty U_1 . Similarly, the kinematic function is not completely

known and cannot entirely be compensated for by a known inverse kinematic function. Hence, the kinematic uncertainty is included in Fig. 3.5 as U_2 . The reasons for the existence of U_2 include inaccurately known robot geometry, gravitational loading, gear train transmission errors etc. Furthermore, as before, R_1 and R_2 refer to linear dynamic compensators and P_2 denotes a linear prefilter, yielding a 2-degree of freedom kinematic loop regulator.

The acoustic transmission delay was neglected in Fig. 3.5. However, it could be taken into account by regarding it to be part of U_2 . Alternatively, it could be regarded as a part of the plant which is not contained in the inner loop. A control design for the outer loop would then have to account for the delay explicitly.

The main point of using feedback control as opposed to feedforward control is the ability of feedback to reduce uncertainty. The success of a feedback controller can therefore be measured by the amount to which the original uncertainty is reduced. It is for this reason that hierarchically organized control structures are potentially better than single loop designs. Hierarchical structures break the overall uncertainty into many smaller parts, each of which is reduced by a local control loop. Supervisory controllers then coordinate the interaction of these local loops. This idea can be made precise in a H_∞ setting. In a very rudimentary way, the idea of hierarchical control design can be applied to the double loop in Fig. 3.5.

When faced with the task of finding suitable controllers R_1 and R_2 , for the double loop of Fig. 3.5, one has three ways to go about it:

1. *Single loop design:* One cuts the feedback branch of the inner loop and sets $R_1 = I$, i.e one uses no feedback of q_c at all. Then one designs R_2 for the resulting big loop.
2. *Sequential loop design:* One first designs R_1 for the inner loop ignoring the outer loop. Then one designs R_2 for the outer loop with the inner loop replaced by

the Fig. 2.3.

3. *Simultaneous loop design*: Design R_1 and R_2 simultaneously.

The bulk of the literature uses the single loop design method. However, this is clearly not the best method. The uncertainty for this case is $U_{sl} = U_2 \circ U_1$, which is larger than either U_1 or U_2 individually. Thus, a single loop control design has to reduce a larger uncertainty than necessary for the other two methods and is inferior for this reason.

On the other hand, in the sequential loop design method the outer loop has to cope with a much smaller uncertainty than U_{sl} because U_1 was reduced by the sensitivity function of the inner loop as was shown in Fig. 2.3, yielding

$$U_{sq} = U_2 \circ (I + S_1 \circ (U_1 - I))$$

The smaller uncertainty U_{sq} makes the sequential design method potentially better than the single loop method. If the design is formulated as a H_∞ problem then one can quantify the advantage which the second design has over the first design by comparing the H_∞ norm of the optimal achievable sensitivity for both cases. In this sense, the sequential H_∞ design is guaranteed to be better than the single loop H_∞ design.

Finally, the simultaneous loop design method is potentially the best of the three options because it can take all cross coupling between the loops into account. Unfortunately, at present, there is no H_∞ control design theory which would yield the simultaneously optimal controllers R_1 and R_2 . It is for this reason that in the following the next best method, the sequential loop design method, was used for the design of the taskspace servo loop.

3.2.2 Continuous Time Design

Taskspace and Joint Space Errors

The overall design goal of the double loop manipulator control system is to make it track as closely as possible a desired taskspace trajectory. In other words, one wants the taskspace error e_y to be small, where

$$e_y = y_d - y_c$$

The loop design for Fig. 3.5, however, will be carried out in the joint space. The corresponding error in the joint space is

$$e_{\hat{q}} = \hat{T}^{-1}(y_d) - \hat{q}_c$$

Hence, small taskspace and joint space errors are related by:

$$e_{\hat{q}} = \hat{J}^{-1}e_y$$

This equation can be used to map taskspace performance specifications into joint space specifications. However, often one can formulate the performance requirements as a desired percentage reduction of e_y . In this case, the same requirements apply for taskspace and joint space.

In staying consistent with the sequential loop design philosophy, one can break the overall design task down into three subtasks:

1. The design of the inner loop regulator R_1 to reduce U_1 .
2. The design of the outer loop regulator R_2 to reduce U_{sq} .
3. The design of the prefilter P_2 to prescribe the tracking error behavior of the nominal Cartesian loop.

In the following, each of the above items will be addressed.

1st Order Design

The reduction of the dynamic uncertainty U_1 is the prime objective of the joint loop. The servo part R_1 of the computed torque method is designed to achieve this goal. If the servo part is chosen to be a simple decoupled PD-controller

$$R_1(s) = (sK_v + K_p)I \quad (3.11)$$

then the sensitivity of the nominal joint loop exhibits a type-2 uncertainty reduction and the nominal closed joint loop transfer function matrix H_0 is:

$$H_0(s) = \frac{sK_v + K_p}{s^2 + sK_v + K_p} I \quad (3.12)$$

The scalar parameters K_v and K_p are chosen to ensure stability of the system in the presence of unstructured uncertainties, in particular flexible modes and neglected time delays.

The second subtask is the reduction of the kinematic uncertainty U_{sq} and is the primary concern of this thesis. In principle, the design of the MIMO compensator $R_2(s)$ for the Cartesian loop can be done by reformulating the conditions (3.5) and (3.7) as a H_∞ optimization problem [49], [50]. However, it is probably simpler to reduce the design of the MIMO compensator $R_2(s)$ to the design of a classical SISO loop: If the regulator is chosen to have the form $R_2(s) = r_2(s)I$, where $r_2(s)$ is a scalar transfer function, then it is evident from the decoupled nature of $H_0(s)$ and from (3.1) that the nominal Cartesian closed loop transfer function $H'(s)$ and the nominal Cartesian open loop transfer function $L_0(s)$ are decoupled, and have the forms $H'(s) = h'(s)I$ and $L_0(s) = l_0(s)I$ respectively, where $h'(s)$ and $l_0(s)$ are scalar transfer functions. Hence, neglecting Δ_n , the input-output stability and performance conditions (3.5) and (3.7) simplify to:

High frequency conditions:

$$|h'(j\omega)| < \frac{1}{\tilde{g}(N_0)(1 + l_m(\omega)) + l_m(\omega)} \quad (3.13)$$

Low frequency conditions:

$$|l_0(j\omega)| \geq \frac{ps(\omega)}{[1 - \hat{g}(N_0)][1 - l_m(\omega)]} \quad (3.14)$$

Therefore, the objective of the MIMO loop design simplifies to finding a scalar transfer function $r_1(s)$ which satisfies the above conditions.

Thus, a simple type-1 controller can be synthesized as:

$$R_2(s) = \frac{K_i}{s} h_0^{-1}(s) I \quad (3.15)$$

Since h_0 is second order, minimum phase and has just one pole excess, $R_2(s)$ can be easily realized. The sensitivity of the nominal Cartesian loop is therefore

$$S_0(s) = \frac{sK_i}{s + K_i} I \quad (3.16)$$

The scalar K_i is chosen such that the bandwidth of the Cartesian loop does not exceed the bandwidth of the joint loop to avoid stability problems because of the unstructured uncertainty of the inner loop. More precisely, the condition (3.13) must be obeyed.

An additional design problem is to prescribe the desired tracking error behavior of the nominal loop. The nominal tracking error e_{yn} can be defined as:

$$e_{yn}(s) = [I - (1 - S_0(s))P(s)]y_d(s)$$

If one chooses

$$P_2(s) = \frac{s}{K_i} I + I$$

then the nominal tracking error is zero. Naturally, in order to reap the benefits of feedback, the bandwidth of $y_d(s)$ must be smaller than the bandwidth of the loop. Further, it is of interest to examine the homogeneous solution of e_{yn} . The dynamics of e_{yn} can be written as:

$$\dot{e}_{yn} + K_i e_{yn} = 0$$

Hence, any initial error will be damped out without overshoot. This is an important consideration in robotics because it makes it easier to control the interaction of the manipulator with its environment.

2nd Order Design

If the 20dB/dec error reduction of the type-1 system is not fast enough, and one rather wants 40dB/dec, then one can design a type-2 controller analogously to the type-1 controller:

$$R_2(s) = \frac{K_i}{s^2}(s + K_c)h_0^{-1}(s)I$$

The sensitivity of the nominal Cartesian loop is therefore

$$S_0(s) = \frac{s^2}{s^2 + K_i s + K_i K_c} I \quad (3.17)$$

As before, the parameters K_i and K_c are chosen such that the bandwidth of the Cartesian loop does not exceed the bandwidth of the joint loop. The nominal tracking error can be made zero if one chooses the prefilter to be

$$P(s) = \frac{s^2 + K_i s + K_i K_c}{K_i s + K_c} I \quad (3.18)$$

The nominal tracking error dynamics are therefore:

$$\ddot{e}_{tn} + K_i \dot{e}_{tn} + K_i K_c e_{tn} = 0 \quad (3.19)$$

Thus, to avoid overshoot but at the same time make the error response not too slow, one should choose K_i and K_c to yield critically damped error dynamics.

A Design Example for the PUMA 600

The desired performance $ps(\omega)$ of the Cartesian loop depends on the desired reduction of the the Cartesian error. In [2] the approximate average size of the Cartesian error was characterized to be 5-6 mm. To reduce this error to less than 0.1 mm, (the

accuracy of interest), requires a loop gain of 30-40dB at low frequencies and likely somewhat less at higher frequencies, in the case that the end effector is restricted to a subregion of the workspace.

In addition to this performance requirement, the kinematic loop must be stable and performing in the presence of uncertainty. It was found in simulations that $\bar{g}(N_0) < 0.06$ is a reasonable guess for trajectories which avoid singularities by more than the distance of the robot's kinematic error. Other nominal trajectories are not very useful anyways because there is not even a guarantee that they lie in the workspace of the robot. Hence, if we only consider the contribution of $\bar{g}(N_0)$ and neglect Δ_m , then the system must be designed to be robust with respect to changes of maximal 6% of the loop gain or 0.5dB. This is so small that it does not have to be considered explicitly in the design.

A realistic choice for the closed loop bandwidth of a PUMA600 was suggested by Daneshmend [43] to be about 2 Hz. Thus, if we fix the bandwidth for the type-1 kinematic loop at 2 Hz, then the kinematic error can be neglected for frequencies below 0.2Hz-0.02Hz.

The overall trajectory error, however, may be larger than expected from the considerations above. This is because the the joint loop contributes uncertainty.

A type-1 loop design requires the determination of the three constants K_v , K_p and K_i . For a 2 Hz bandwidth, one can choose $K_v = 8\pi$ and $K_p = 16\pi^2$ which places both poles of the closed joint loop at 2 Hz. The Cartesian loop design constant K_i is simply chosen to be $K_i = 4\pi$, placing the only pole at 2 Hz.

A type-2 loop design requires the determination of the four constants K_v , K_p , K_i and K_c . For a ca. 2Hz bandwidth, one can choose K_v and K_p as before, and simply choose $K_i = K_v$ and $K_c = K_p/K_i$. Then, both the inner and the outer loop have two poles at 2Hz.

3.2.3 Discrete Time Design

Robot control is often done by computers. This necessitates the design of a discrete rather than a continuous time controller. Most of the continuous time design considerations, which were mentioned in the previous section, remain valid and will not be repeated here. However, there are some additional considerations. In particular, the inner loop must include a sample and hold device and a computational delay. Depending on the sampling rate, these additions can lead to a substantially altered high frequency behavior. For instance, there will be additional phase lag at the cross-over frequency and a reduction of the high frequency roll-off. In general, this means a less stable, less robust system.

The discrete time control problem is illustrated in Fig. 3.6, which essentially represents the discrete time equivalent of Fig. 3.5. The dashed box contains the discrete regulator. Its linear part consists of three identical, decoupled sets of four SISO filters. Alternatively, it could be described to consist of three identical, decoupled linear MISO systems, each having two inputs and one output. Since the controller is the same for each of the three dimensions of space, only one of the three channels will be discussed.

The nonlinear part of the discrete regulator consists of two $\hat{T}^{-1}$ transformations. Further, it is assumed that there is a computational delay of one sampling period. This is accounted for by the z^{-1} block. The "plant" is shown outside of the dashed box in Fig. 3.6. Notice the zero order hold circuit (ZOH) preceding the nominal plant s^{-2} . Moreover, the z^{-1} block explicitly accounts for the acoustic delay.

Perhaps, the most common method of discrete loop design is to discretize the regulator obtained by a previous continuous time design. The main disadvantage of this method is that it requires high sampling rates. This is not desirable in this case because of the unnecessarily high computer load and the decreased receiver performance due to high pulse rates (see Sec. 4.3.1).

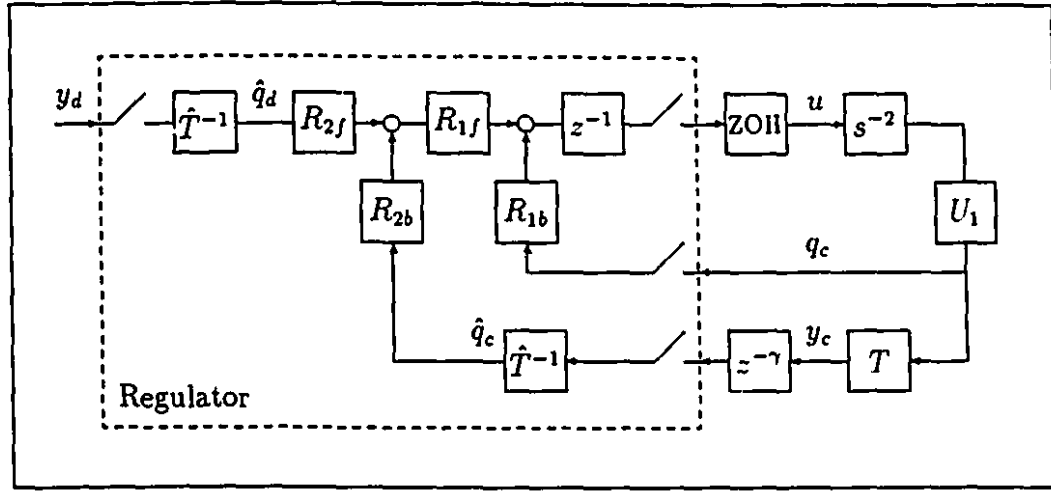


Figure 3.6: The double loop with a discrete time regulator

There are many alternative discrete regulator design methods. One way of approaching the problem is to design the controller explicitly as a discrete time controller. For the joint loop, for instance, it was shown in [51] that this can lead to a discrete time lag-lead design which shows good performance, stability and robustness.

Here, a standard transfer function synthesis method was used. The method is well explained in [52].

First, the z -transform of the computational delay in series with the ZOH and the double integrator plant is obtained:

$$H(z) = \frac{T_s^2}{2} \frac{z+1}{z(z-1)^2}$$

where T_s is the sampling period. Then a desired closed loop pulse transfer function is specified. For instance:

$$H_m(z) = \frac{\frac{1}{2}(1+p_1+p_0)(z+1)}{z(z^2+p_1z+p_0)}$$

where p_1 and p_2 are chosen s.t. the poles are critically damped for a natural frequency of $f_n = 2\text{Hz}$. It is good practise not to attempt to cancel unstable or lightly damped

plant zeros with regulator poles. Hence, the marginally stable LHP plant zero at -1 is carried through from $H(z)$ and becomes a zero of $H_m(z)$, too.

The design method also requires the specification of a characteristic observer polynomial $A_o(z)$. This allows one to influence the loop's sensitivity function by specifying some of its poles.

For a 50 Hz sampling rate, an observer polynomial which yields an acceptable sensitivity function was found to be:

$$A_o(z) = z(z - e^{-8/50});$$

The compensators R_{1f} and R_{1b} can now be found by solving a diophantine equation:

$$\begin{aligned} R_{1b}(z) &= \frac{296.89z^2 - 278.63z}{z^2 - 0.4077z + 0.0557} \\ R_{1f}(z) &= \frac{123.47z^2 - 105.21z}{z^2 - 0.4077z + 0.0557} \end{aligned}$$

The design for the outer loop is analogous, provided that the delay $z^{-\gamma}$ and the nonlinearity $\hat{T}^{-1}T(\cdot)$ are neglected. The inner loop's $H_m(z)$ becomes both the outer loop's plant and the outer loop's desired closed loop transfer function $H_m(z)$. To ensure high gain at low frequencies and zero steady state error, the outer feedback regulator must include at least one integrator. With this constraint on the regulator, the observer polynomial which yielded an acceptable outer loop sensitivity function was found to be:

$$A_o(z) = z^2(z - e^{-20/50});$$

The resulting compensators R_{2f} and R_{2b} are:

$$\begin{aligned} R_{2b}(z) &= \frac{1000(0.8242z^3 - 1.2821z^2 + 0.4986z)}{123.47z^3 - 82.763z^2 - 20.352z - 20.352} \\ R_{2f}(z) &= \frac{123.47z^3 - 82.763z^2}{123.47z^3 - 82.763z^2 - 20.352z - 20.352} \end{aligned}$$

It is evident from Fig. 3.5 that the nominal tracking error can be represented as:

$$e = y_d - \hat{T} \circ h_m \circ \hat{T}^{-1}(y_d)$$

where $h_m(\cdot)$ represents the time domain dynamics of the linear filter $H_m(z)I$. It is clear that there is nonlinear dynamic cross-coupling between the three Cartesian space channels because h_m is sandwiched between the two nonlinearities. As a result, the nominal tracking errors for the three directions of space are not independent. It would be better to have:

$$e = y_d - h_m(y_d)$$

which has independent channels.

To achieve this form requires canceling all nominal inner dynamics and prefiltering y_d . In other words, in Fig. 3.6, R_{2f} is replaced by R'_{2f} :

$$R'_{2f} = R_{2f}H_m^{-1}(z)$$

and y_d is replaced by y'_d :

$$y'_d = H_m(z)y_d$$

There is a potential problem of the system not being causal, particularly, if the prefilter for y_d is omitted. However, even if the system is uncausal this usually constitutes no serious problem because, in practice, y_d will be known at least a few sampling periods in advance; if not, the prefilter must ensure causality.

A more serious problem is the fact that the H_m -zero at -1 cannot be canceled. Instead, one can choose:

$$R'_{2f} = R_{2f}H_m^{-1}(z)\frac{(z+1)^2}{z}$$

which does not cancel the zero. The resulting inner dynamics are $\frac{(z+1)^2}{z}$ which resemble unity quite closely, particularly at frequencies which are small compared to the sampling rate.

The preceding design can be slightly modified to account for the acoustic delay.

Part II

The 3D Ultrasound Position Sensor

Chapter 4

The Hardware

4.1 Introduction to Part II

The second part of the thesis introduces the ultrasonic position measurement device and analyzes its strong and weak points. In particular, a stochastic model is presented which describes how the fluctuating medium corrupts the position measurements. The factors which influence the TOF measurements include:

The **electronic hardware** of the ultrasound system:

- *Non-ideal Transducers* prevent the acoustical and the geometrical center of a transducer from coinciding: When the sender and receiver transducers face each other at an angle, then the transducers cannot longer be appropriately modeled as mathematical points.
- *Electronic noise* influences the measurement in the form of added signal processing delays, and the noise of the signal amplification and detection electronics.

The **wave nature** of sound:

- *Reflections* can lead to multipath arrival of the signal, making it difficult to determine the exact time of flight.

- *Reverberation* is the multiple reflection of sound in a closed room. Thus, an earlier sound burst can still corrupt the current TOF measurement.
- *Ambient noise* can corrupt the measurements when other ultrasound sources besides the sender are present in the room and emit sound energy at frequencies to which the receivers are sensitive.
- *Refraction* is caused by the presence of a temperature gradient in the room. The ray path is bent and the measured TOF differs from the case without a gradient.
- *Doppler* can corrupt the TOF measurements when the sender is moving or when there is a strong wind.
- *Diffraction* and scattering have only negligible influence.

The fluctuating medium in which the sound propagates:

- *Temperature drifts* and *temperature gradient drifts* cause slow global changes in the speed of sound and the speed gradient.
- *Turbulences* due to convection currents, wind gusts etc. lead to relatively fast local changes of the sound-speed.

It was decided to model the TOF measurement fluctuations, irrespective of their origin, as a stochastic process primarily because this form can readily be used by a Kalman filter for estimating the sender position in the Cartesian feedback loop. Hence, the measurement model is not intended to be a theory of sound propagation in turbulent air, but is simply a means to improve the precision of the Cartesian feedback loop. Given this intended use, the emphasis here is on a rough and simple model which nevertheless captures the dominant statistical features.

In the following, all stochastic processes are understood to be defined with respect to an underlying well defined probability space $(\Omega, \mathcal{F}, P)$ [60], where, Ω denotes the fundamental sample space, $\mathcal{F}$ is the underlying σ -algebra and P is the probability function.

It will be shown in Sec. 5.2.1 and Sec. 5.2.2 that the acoustic ray theory is valid and that the sound ray path is well approximated by a straight line which connects the sender with the receiver. Mathematically, the straight-line ray path connecting the sender with a receiver can be parameterized as:

$$\Gamma_i(t, \alpha) = S(t) + \tilde{e}_i(t)\alpha \quad \alpha \in [0, d_i(t)] \quad (4.1)$$

as before, $d_i(t)$ is the distance and $\tilde{e}_i(t)$ is the unit direction from the sender at $S(t)$ to the receiver at M_i . The time t in (4.1) denotes the instant when the sound departs from the sender.

The properties of the medium along the ray path are often described by modeling the refractive index of sound along the path as a random process. However, for our purposes it is slightly more convenient to use the “inverse sound-speed” n instead of the refractive index. It is defined as the inner product

$$n(\bar{\omega}, t, \Gamma_i(t, \alpha), \tilde{e}_i(t)) = \left[\frac{1}{c_x}, \frac{1}{c_y}, \frac{1}{c_z} \right] \tilde{e}_i \quad \bar{\omega} \in \Omega \quad (4.2)$$

where c_x, c_y, c_z denote the speed of sound in x-, y- and z-direction respectively. They are random space-time functions themselves. Thus, the inverse sound-speed depends on

1. $\bar{\omega}$ because of the statistical nature of the inverse sound-speed.
2. t because the inverse sound-speed evolves with time.
3. $\Gamma_i(t, \alpha)$ because the inverse sound-speed is a function of space.
4. $\tilde{e}_i(t)$ because of the effects of air movements and sound-speed gradients.

Furthermore, it is convenient to distinguish between a “deterministic” part Dn and a “random” part Rn of inverse sound-speed such that:

$$n(\bar{\omega}, t, \Gamma_i(t, \alpha), \tilde{e}_i(t)) = ^Dn(t, \Gamma_i(t, \alpha), \tilde{e}_i(t)) + ^Rn(\bar{\omega}, t, \Gamma_i(t, \alpha), \tilde{e}_i(t)) \quad (4.3)$$

where $^Dn = E\{n\}$ is thought to represent the very slow, spacially highly correlated fluctuations which can be regarded to be deterministic. The slowly changing average room temperature and average room temperature gradient or a slowly changing average wind velocity are examples of mechanisms which give rise to Dn . On the other hand, Rn represents the faster and spacially less correlated zero-average fluctuations i.e. $E\{^Rn\} = 0$; the fluctuations described by Rn are produced by air turbulences, convection etc..

Similar to inverse sound-speed, one can model the TOF measurements themselves as a random process. For this purpose it is convenient to distinguish among four parts of the TOF:

$$\text{tof}_i(\bar{\omega}, t) = ^MD\text{tof}_i(t) + ^MR\text{tof}_i(\bar{\omega}, t) + ^ID\text{tof}_i(t) + ^IR\text{tof}_i(\bar{\omega}, t) \quad (4.4)$$

where:

1. MD means ‘medium dependent deterministic’ and models the influence of the average inverse sound-speed, the average inverse sound-speed gradient and the average wind velocity.
2. MR means ‘medium dependent random’ and models the influence of atmospheric turbulences.
3. ID means ‘medium independent deterministic’ and models the influence of electronic delays, clock errors, effective transducer size and other medium independent biases.

4. IR means ‘medium independent random’ and models the influence of the receiver SNR etc.

The two random parts are defined to have zero mean. Each of the 4 parts above will be discussed in detail. However, the by far most important of the 4 parts are the two medium dependent parts MD and MR . The dependence between the medium dependent part of the TOF and the inverse sound-speed can be expressed as (in [57]) as a spacial random walk:

$$^{MR}_{\text{tof}_i(\bar{\omega}, t)} = \int_0^{d_i(t)} R_n(\bar{\omega}, t, \Gamma_i(t, \alpha), \bar{e}_i(t)) d\alpha \quad (4.5)$$

$$^{MD}_{\text{tof}_i(t)} = \int_0^{d_i(t)} D_n(t, \Gamma_i(t, \alpha), \bar{e}_i(t)) d\alpha \quad (4.6)$$

Note that the integrals above are taken over space only. This is a very good approximation since the largest possible TOF is much smaller than the smallest time constant of the turbulences. The latter can, therefore, be considered ‘frozen’ during sound emission.

In the following, the experimental results presented were conducted in two distinct locations: A ‘small room’ and a ‘large room’. The small room had a floor size of 4m×4m and a height of 2.2m. It had a quiet and well controlled environment. The floor size of the large room was ca. 11m×9m and its height was 3.5m. The large room was a busy electronics and robotics lab, many people were moving about, there were many fans and heat emitting equipment, the doors were being constantly opened and closed.

4.2 Range Finder Overview

4.2.1 A TOF Measurement Scheme

Background

The hardware has the task of generating a suitable ultrasound ranging signal and then precisely measuring the time delay between the sending and the receiving of the signal. The main design problem is to accurately acknowledge the arrival of the signal. The difficulties stem from the small bandwidth of the transducers used which does not permit impulse-like signals to be sent. Rather, only a relatively slow rise of the signal magnitude can be achieved. The problem is further aggravated by the large wavelength ($\lambda = 8.5mm$) of the ultrasound used: The desired precision requires a detection resolution of about a hundredth of a wavelength. Moreover, reflections and reverberation may influence the signal phase and shape; one has to guard against them.

Given these problems, the method used to determine the time of arrival become crucial. In the following, the abbreviation TOA for 'time of arrival' is meant to describe the time delay between the first reception of sound energy belonging to the direct path signal and the time at which this signal is actually acknowledged by the receiver electronics as having arrived. On the other hand, the abbreviation TOF for 'time of flight' is meant to refer to the time delay between the time when the first bit of signal energy is sent and the time when the first bit of direct path signal energy is received.

Many ultrasound ranging systems use a gated continuous wave (CW) signal and a voltage trigger. When the amplitude of the received signal exceeds the trigger level, then the signal is regarded as having arrived. The problem with this method is that it is critically dependent on the magnitude of the signal. Thus, amplitude variations

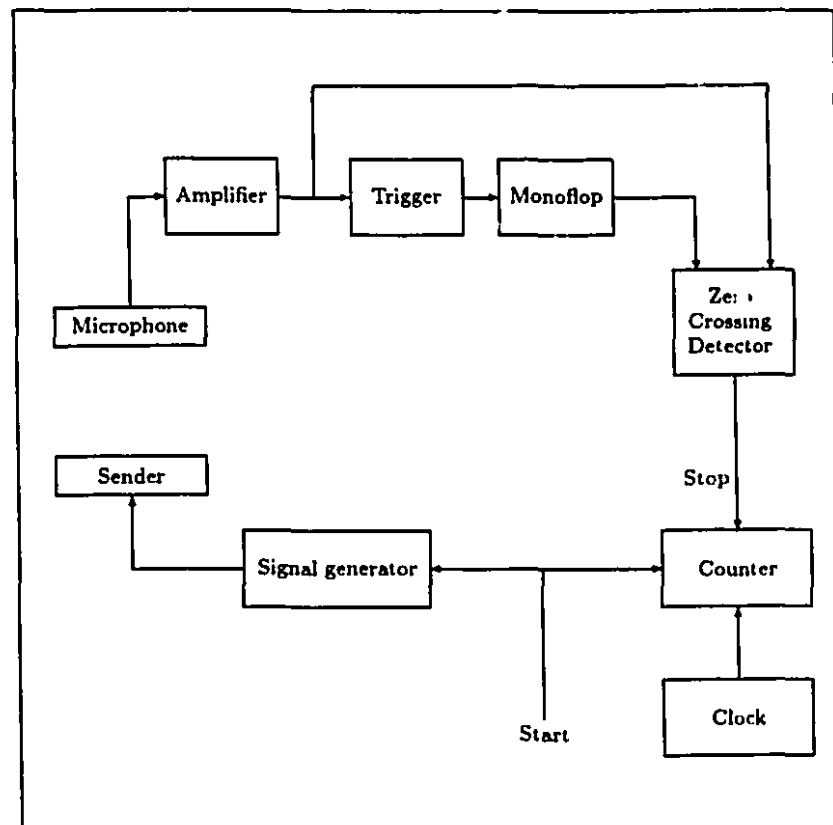


Figure 4.1: Hardware block diagram

due to distance changes, transducer reorientation or due to turbulences in the air will greatly influence the TOA. This may be perfectly acceptable for many applications, but it appears to be not sufficient for high precision applications.

The other obvious method is to use a one tone continuous wave and measure the "arrival time" by measuring the phase shift between sent wave and received wave. The resulting ambiguity in terms of multiples of the wavelength could be reduced by using a two-tone scheme [7] where the beat frequency is used to help resolve the ambiguity. CW methods have the potential of being extremely precise. Unfortunately, because of unavoidable reflections, CW methods are useless in a real life environment. One CW method was tried, using a frequency-shift keying modulated signal and a phase locked loop receiver with voltage trigger at the loop output to determine the time of arrival. However, experiments with this method showed that even this scheme is too prone to reflections.

There are also more exotic methods, for example, spread spectrum. This scheme is widely used for precise ranging. It has the potential to be insensitive to reflections, reverberation and ambient noise. However, one disadvantage of this method is its complexity and the associated costs. Another problem in our context is the small available bandwidth of the chosen transducers.

The Prototype

The experimental ranging system that was used for this work is in some respect a combination of a one tone CW method and a gated CW method. It has proved to be superior to all other tested methods. A block diagram of the scheme is shown in Fig. 4.1: A gated 40Khz CW signal is sent and simultaneously, a timer is started. When the signal is picked up by a microphone it is first amplified. If the amplified signal is smaller than a preset trigger level, then the signal is disregarded. On the other hand, if it is larger than the trigger level, then the signal is regarded to have

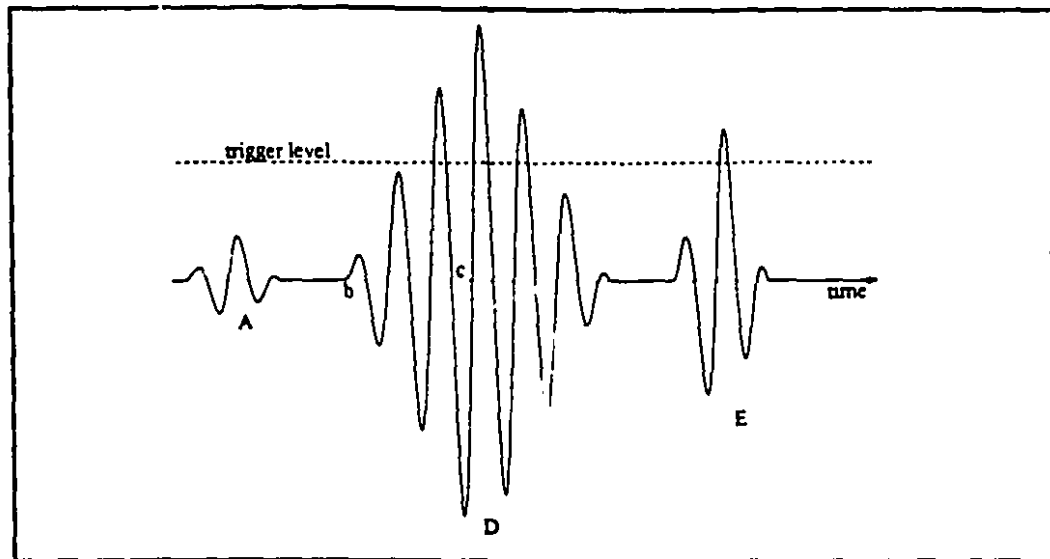


Figure 4.2: A typical received signal

“roughly” arrived and a strobe pulse triggers a monoflop which in turn enables a zero crossing detector to send a stop pulse to the timer at precisely the time of the next zero crossing of the signal. Thus, the first zero crossing after the trigger is regarded as the exact instance of the arrival of the signal.

The Fig. 4.2 shows a pseudo oscilloscope view of the received signal. The wave package ‘D’ represents the signal whose instance of arrival one desires to determine. The label ‘b’ points to the true beginning of this wave package and the label ‘c’ points to the first zero crossing after the trigger. This is the moment with respect to which the time of arrival is measured. Hence, the delay between ‘b’ and ‘c’ is the TOA. The wave package ‘E’ represents a reflected signal. The reflected package arrives later than the main package ‘D’ because it has a longer path. Often, the reflected signal has a smaller amplitude than the main signal because its path is longer and because not all of the signal energy is reflected. The directivity of the ultrasound transducers, however, makes it possible that the amplitude of the reflected signal can vastly exceed the amplitude of the direct path signal. This can occur when sender

and receiver face each other at a large angle and the reflected signal arrives at a more favourable angle than the direct path signal. The wave package 'A' represents a reflected signal associated with a previous main signal or, it represents the reflection of a reflected signal whose main signal was received many sampling instances before. These multi-reflection signals arise in closed rooms mainly due to the high reflectivity of the walls.

The first task of the trigger is to distinguish between the desired wave package 'D' and the undesirable signals 'A' and 'E'. In order to be able to do this, it is necessary that those 'A' packages, which are received immediately before 'D', have a substantially smaller amplitude than 'D' itself. 'A' packages which arrive much earlier are of no concern in our case, because a priori knowledge of the time of arrival of the signal can be used to define a reception window during which the receiver is solely sensitive: The trigger mechanism is enabled only shortly before the estimated arrival of the signal and closed immediately after signal arrival was detected. As long as the trigger level is larger than the largest amplitude of the 'A' package which appears in the reception window, but smaller than the largest amplitude of the 'D' package, the 'D' package will be successfully distinguished from the other signals. In particular, the amplitude of the reflected wave is immaterial because once 'D' is detected, 'E' is outside of the reception window.

4.2.2 Choosing a TOA

It is not shown in Fig. 4.2 that 'D', 'E' and 'A' can overlap. This reality introduces undesired phase shifts and signal distortions into 'D', which in turn change the TOA. Therefore, once the correct wave package is identified, the second task of the trigger is to perform a tradeoff between the "SNR(TOA)" and the "reflection robustness" of the receiver. The term "reflection robustness" refers to the smallest time delay between a direct path and a (once) reflected path which has no influence on the TOA. The

symbol "SNR(TOA)" stands for the ratio between the amplitude of the direct path signal and the amplitude of the multi-reflected signals as a function of the TOA.

The best possible reflection robustness is achieved when the time of arrival is measured at 'b' instead of 'c'. At 'b', no reflection path, however close to the direct path can influence the measurement. Unfortunately, the point 'b' has the worst SNR(TOA). Therefore, $\text{SNR}(0)=0$, since at 'b' the signal amplitude is zero. The best SNR(TOA), on the other hand, occurs for a TOA where the amplitude of 'D' has reached its maximum value. If, however, the trigger level is set so high that the SNR(TOA) is maximized, then the reflection robustness is low. For a typical signal, the maximum amplitude is reached after ca. ten periods. Thus, any reflection path which is less than 8.5cm longer than the direct path will probably render the measurement useless. Another compelling reason for not choosing a very high trigger level is, of course, the danger that a small fluctuation of the signal amplitude may cause the trigger to miss 'D' altogether. There are also other considerations for the choice of trigger level and TOA. In particular, it is desirable that the pulse envelope has reached a plateau when the zero crossing is measured. The reason for this is an improved TOF bias reduction. This is discussed in Sec. 4.4.

In practice, some kind of compromise must be found. The situation can be improved by increasing the available bandwidth of the transducers by electronic means. This yields a faster rise time and the trigger can more easily detect the correct period. Now one can delay the zero crossing detection until after the pulse reaches its plateau. This yields maximum SNR(TOF) and best TOF bias reduction and at the same time, because of the faster rise time, offers a reasonable reflection robustness.

Once one has decided on a nominal TOA, one would like it to be maintained independent of the signal amplitude. This is particularly important because the TOF is computed by subtracting the TOA from the signal delay measurements. Therefore, when the signal amplitude changes due to locomotion or reorientation of the sender,

the amplifier gain must be adapted to leave the TOA invariant. Alternatively, the trigger level could be changed. These types of adaptation can be achieved with relative ease since position and orientation estimates are available from the Kalman filter. This information could be used to send signals to simple D-A converters, which in turn, would set the trigger level or the amplifier gain such that the desired TOA is maintained. A simpler method is to estimate the correct trigger level from past measurements of the signal strength at the receiver. This could be done locally in the receiver with analog electronics.

The above introduced experimental ultrasound position measurement system is completely controlled by a special interface card for an IBM-AT compatible computer. The computer sets the card parameters once and then retrieves the data from the card by means of a C interface program, which is listed in Appendix E. The board is capable of handling four receivers and two senders. It features a single 10 Mhz crystal clock that was used to derive the sampling period, the carrier period, the signal lengths of the first and second sender, and the delay between the first and second sender. It also acted as a clock for the TOF timer chips. The details can be found in the Appendix D.

4.3 Range Finder Limits

4.3.1 Pulse Rate Limits

An important problem in designing digital control systems is the choice of the sampling rate. The range finder imposes an upper limit on the sampling rate because its maximum pulse rate is limited. A first limit is reached for a pulse rate of ca. 170 Hz. This limit is imposed by the 5.8 msec flight time of the signal for the maximum range (2 m). A higher pulse rate than 170 Hz causes the measured TOF to be ambiguous. However, one has very good estimates (better than 1 cm) and can easily resolve many

ambiguities. The theoretical limit for the pulse rate depends on the a priori knowledge one has of the range. If the range is known within an error of say ± 10 cm, then even a pulse rate of 3.4 kHz would cause no ambiguities. However, such a high pulse (sample) rate may not improve the performance of the feedback control system. The 5.8 msec delay imposed by the travel time of the sound for the maximum range effectively limits the bandwidth of the closed loop system to less than ca. 170 Hz. This limit is inherent in the use of ultrasound for position feedback over 2-meter ranges.

There are other factors besides a priori knowledge which limit the pulse rate. One factor is the bandwidth of the transducers which poses an upper limit on the pulse rate because pulses must be well separated from each other to avoid signal corruption. If the 5 kHz transducer bandwidth is not increased by electronic means, then the pulse rate is effectively limited to less than, say, 500 Hz. However, the factor that most restricts the pulse rate are reflections. For high pulse rates it becomes increasingly difficult to separate the signal from the reflections in the room which have no time to die down before the next pulse. Experiments seem to indicate that, say, a 300Hz pulse rate is still bearable in a regular lab. Of course, the situation improves in larger rooms and in rooms with more sound absorbing material.

When only position measurements are needed, then the maximum sampling rate equals the maximum pulse rate. However, when both orientation and position measurements are needed, then three senders instead of one are required. If the communication channel is time multiplexed, then the maximum sampling frequency is just a third of the maximum pulse rate. The use of additional senders for the purpose of continuous calibration does not reduce the maximum pose sampling rate in the same linear manner because the calibration process requires only a sampling rate of about 4 Hz. Another solution is to use both time and frequency multiplexing. For instance, there exist transducers produced by Murata Inc., which have very similar dimensions and radiation characteristics to the transducers used in this thesis, but they have two

resonance frequencies instead of one. Hence, the maximum sampling frequency for pose measurements would be about half the maximum pulse rate.

4.3.2 Electronic Noise

Quantization Noise

Perhaps the most obvious electronic signal corruption is the quantization noise due to the finite clock speed. A 25 μm resolution at a 2 m range one would require ca. 13.5 MHz clock speed and a 17 bit counter. Our experimental ultrasound system used just 10 Mhz 16 bit counter chips because of the lower prize and the ready availability of these chips. Hence, the maximal resolution is 0.1 μsec . If one subtracts half a clock period from all TOF measurements, one obtains a symmetrical resolution of $\pm 0.05 \mu\text{sec}$, which corresponds to ca. $\pm 0.018 \text{ mm}$. If a higher resolution is desired, this can be easily achieved by increasing the clock speed to 16 MHz or above and the use the appropriate counters.

Receiver Background Noise

The sensor signal is quite weak (just a few mV). The receiver noise floor constitutes a problem particularly when the distance between sender and receiver is large or the sender and receiver axis are misaligned by a large angle. Thus, the amplifier stages of the receiver must be well shielded and designed. Nevertheless, the influence of circuit noise cannot be avoided completely. Fortunately, such disturbances can be modeled as white noise and can be averaged out easily. It has also been observed, that the microphone provides a small 40 kHz signal even if no sound source is present in the room. Thus, the receiver background noise consists of the amplifier thermal noise and the amplified microphone self-noise. Table 5.1 contains information on the relative strength of the receiver background noise and on the TOF measurement errors it can

cause. Table 5.1 is discussed in the section on reverberation noise.

4.3.3 Electronic Delays

The analog part of the sender, the ultrasound sender and microphone as well as the receiver amplifier, all distort the received signal somewhat and add a small lags. In particular, the zero crossing detector has a small hysteresis and its output pulse has a finite rise time. Moreover, the first layer of logic gates of the digital parts of sender and receivers add a small delay. Since all senders and receivers are nominally built the same way and the receivers trigger at the same zero crossing thanks to an automatic signal gain adjustment, it is reasonable to assume that the total effect of these delays and lags is approximately the same for all channels.

It is estimated that the total delay due to all sources is in the order of $0.1\mu\text{sec}$ for the quality of electronic components used in the prototype. The electronic delay may change slightly over time with temperature, and with ageing of the parts, but given the smallness of the delay and its conceivable changes, it probably can be considered a constant for all practical purposes, especially when high quality electronic parts are used. The electronic delay can be estimated at the time when the system is calibrated. The estimated delay would then be added to the TOA and the sum would be subtracted from the counter readings, yielding the TOF measurements. In the following, the sum of the electronic delay and the TOA will be referred to as the hardware delay (HD). In fact, at calibration one would rather estimate HD than its two components because the two cannot easily be distinguished and have the same effect on the measurements. Alternatively, HD could continuously be recalibrated by a reference sender along with the speed of sound.

It has been observed, that the amplitude of the signal can fluctuate very violently under the influence of wind gusts; so much, in fact, that sometimes the receiver triggers at a later period of the signal than the one corresponding to the desired

TOA. Even an automatic gain adjustment circuit is not always able to avoid this. In these cases the “true” TOF can be approximately recovered from the counter readings CR provided that one has an estimate TOF_e of TOF which has an error better than $\pm T_c/2$, where T_c denotes the period of the 40 kHz carrier. A recovery is possible because the TOA is known to be approximately an integer multiple of the carrier period T_c . Hence, given the nominal hardware delay HD, the estimate TOF_e and the counter reading CR one can compute the TOF as:

$$\text{TOF} = \text{CR} - \text{HD} - [(\text{CR} - \text{TOF}_e) \bmod T_c]T_c$$

4.4 Non-ideal Transducers

4.4.1 Theoretical Analysis

For the purpose of 3D ranging, an ideal ultrasound transducer would be a perfect point source or point sink and have a perfectly spherical radiation characteristic. These transducer properties are desirable because, in the first place, one wants to guarantee that all the receivers receive the sender at all times, independent of the relative orientation of sender and the receivers. Secondly, the triangulation formula assumes the sender and the receivers to be mathematical points without spacial extensions. In other words, the geometrical center and the acoustical center of the transducer are assumed to coincide.

Another desired characteristic of an ideal transducer would be that it operates at a small wavelength because this would improve ranging accuracy. However, the heavy damping of high frequency ultrasound transmission seriously limits the frequencies one can use and still operate at the desired range of two meters. Moreover, to get near spherical radiation characteristics, the transducer diameter must be in the order of the wavelength. A small diameter, however, makes it difficult to transmit a sufficient

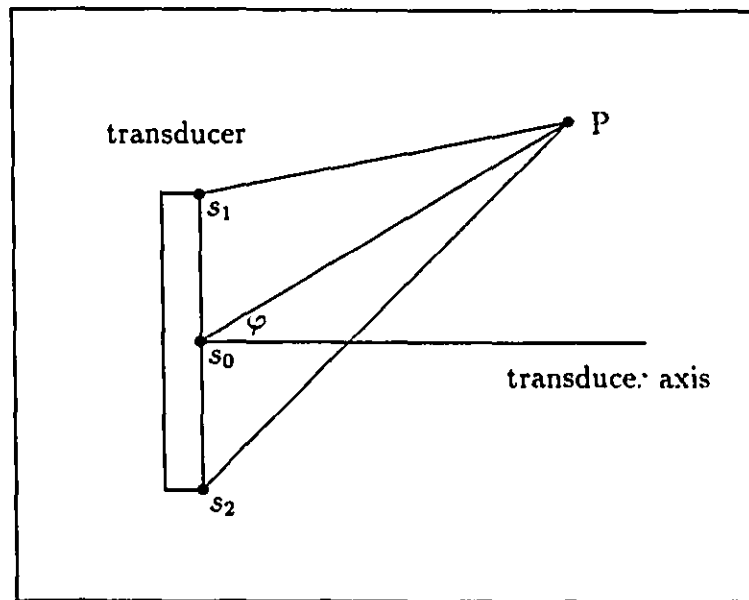


Figure 4.3: Transducer misalignment

amount of power for reliable communication. Thus, 3D high frequency ultrasound ranging is difficult to achieve.

Unfortunately, the ideal transducer does not exist. The acoustic center and the geometrical center of real transducers does not always coincide. The best compromise we were able to find was a 40 kHz ceramic ultrasound transducer with a wide opening angle and sub-wavelength size. Even at 2 m distance and a more than 90° angle between the sender and the receiver axis, the signal magnitude was sufficient for reliable signal detection. At smaller distances up to 180° were achieved. The radiation characteristics of the transducer which was used for this study is shown in Appendix C. The device used approaches the ideal transducer in so far as its membrane is slightly smaller than the wavelength. The wavelength is about 8.5 mm and the membrane has a diameter of about 7 mm.

Suppose the transmitter transducer has the form of a disk with radius r . Assume further that the transmitter membrane moves uniformly. One can then make a simple

plausibility argument in order to understand the effects which a non-ideal sound source has on the measured TOF: Consider only three pressure waves p_0 , p_1 and p_2 generated by the uniform movement of the three points s_0 , s_1 and s_2 , which are on the membrane of the sender transducer. As shown in Fig. 4.3, s_0 is at the center of the transducer and s_1 and s_2 are symmetrical with respect to the transducer axis and lie on the very rim of the transducer. We are interested in the phase relationship at a point P in the far field of the transducer between the p_0 wave and the composite wave of p_0 , p_1 and p_2 .

Taking damping and geometrical dispersion into account the pressure at a distance d from a point source can be described by:

$$p(d, t) = A \frac{\exp(-\gamma d)}{d} f(t - d/c) \sin(\omega(t - d/c))$$

where A is a constant, $\gamma \approx 0.0219$ neper/m is the damping constant for 40 kHz ultrasound in air and $f(t)$ represents the (positive) envelope of the pulses sent. The composite pressure p_c at P caused by the movement of s_0 , s_1 and s_2 can be computed by adding p_0 , p_1 and p_2 . For small r/d one can use $\Delta = r \sin(\varphi)$ to compute p_c :

$$\begin{aligned} p_c(d, t, \varphi) &= p(d, t) + p(d - \Delta, t) + p(d + \Delta, t) \\ &= A_c(t) \sin(\omega(t - d/c + B_c(t))) \end{aligned}$$

The time interval $B_c(t)$ is the bias which corrupts the TOF because the detection relevant zero crossing is shifted in time by $-B_c$ with respect to the nominal TOF d/c .

The bias $B_c(t)$ can be readily computed: After some algebraic manipulations and simplifications using $\frac{\Delta}{d} \rightarrow 1$ one obtains:

$$B_c(t) \approx \frac{1}{\omega} \arctan \left(\frac{[f_+ e_+ - f_- e_- + \frac{\Delta}{d} (f_+ e_+ + f_- e_-)] \sin(k\Delta)}{1 + [f_+ e_+ + f_- e_-] \cos(k\Delta)} \right)$$

where $k = 2\pi/\lambda$ is the wave number, $e_+ = \exp(\gamma\Delta)$, $e_- = \exp(-\gamma\Delta)$ and

$$f_+ = \frac{f(t - d/c - \Delta/c)}{f(t - d/c)}, \quad f_- = \frac{f(t - d/c + \Delta/c)}{f(t - d/c)}$$

The formula above indicates that B_c increases with Δ . Therefore, for large φ the TOF bias is large. Further, B_c will decrease with increasing range d . Moreover, the pulse shape and the TOA influence B_c . If damping is ignored, and the zero crossing detection occurs during the rising edge of the pulse, then the term $(f_+ - f_-)$ in the formula above is a positive number contributing to the TOF bias. Thus, keeping the influence of the pulse shape small requires a choice of TOA such that the derivative of the pulse shape is small at the time the zero crossing is measured. Note that zero crossing detection during the raising edge or the plateau of the pulse implies that B_c is non-negative and thus the measured TOF will always be smaller or equal to the nominal TOF which is d/c .

The preceding analysis of the bias B_c does only consider the spacial extension of the sender. The receiver, however, is not a mathematical point either but has the same dimensions as the sender. To take this into account, suppose there are 3 points m_0 , m_1 and m_2 on the receiver membrane, positioned in the same manner as s_0 , s_1 and s_2 . Further assume that m_0 and P in Fig. 4.3 coincide and that the receiver transducer is oriented such that the transducer axis coincides with the vector pointing from m_0 to s_0 . Furthermore, assume that the receiver transducer acts as a perfect spacial pressure integrator. Now, if the range d is very large compared to the transducer radius r , then all three points on the receiver membrane will experience approximately the same pressure and therefore B_c will be the same as if the receiver was a mathematical point at P.

On the other hand, suppose the receiver axis is tilted by an angle φ' away from the m_0 - s_0 -line, then one can use the reciprocity of this situation to the one previously discussed for the sender transducer. One can argue that the point s_0 on the sender membrane plays the role of P, the angle φ' that of φ and m_0 , m_1 and m_2 play the role of their sender counterparts. From this setup one can compute another TOF bias B'_c . Hence, the overall TOF delay could then approximately be computed as the sum of

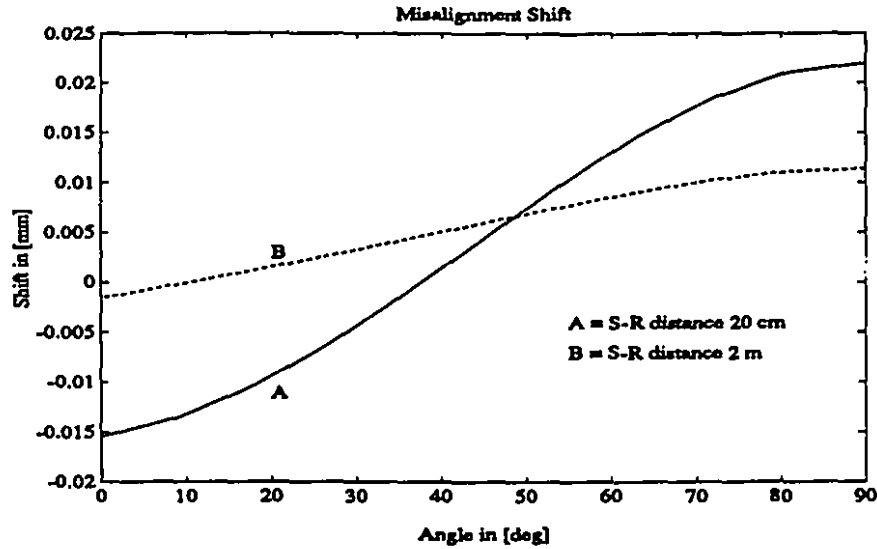


Figure 4.4: Numerically determined S-R misalignment shift for 7 mm sender.

B_c and B'_c .

4.4.2 Numerical Analysis

The situation depicted in Fig. 4.3 can also be analyzed numerically. A quantitative analysis of the TOF bias perceived at point P requires a computation of the value of the pressure field at P by numerically integrating the contribution of every sender surface point. The exact form of the double integral to be solved can be found in most text book on acoustics, for instance [34, page 549]. The calculation made in this thesis, however, also includes damping effects, which are neglected in [34]. For steady state CW conditions Fig. 4.4 shows the misalignment shift G as a function of the angle φ , where

$$p_c(d, t, \varphi) = A \sin(\omega(t - d/c + G/c))$$

and A is a constant.

Contrary to the approximation discussed before, there is a negative bias (delay) for $\varphi = 0$. When φ increases, so does the bias. Finally, the misalignment shift becomes

positive. As can be seen in Fig. 4.4, the misalignment shift G can be neglected. Even when d is as small as 20 cm the shift G varies only by 0.037mm when φ sweeps from 0° to 90° . As expected, the shift is even smaller when d is larger. For comparison, Fig. 4.4 shows also the case when $d=2m$. The results in Fig. 4.4 were obtained for a 40Khz ultrasound transducer with 7mm diameter, the type used throughout this thesis. The shift G is very sensitive to the transducer diameter. For a transducer with 10 mm diameter and $d=20cm$, one obtains a change of 0.2 mm instead of just 0.037 mm when φ sweeps from 0° to 90° . Thus, if 10 mm transducers were used, one must compensate for the misalignment shift.

A complete quantitative analysis of the TOF bias requires a computation of the pressure field at each point on the receiver surface by integrating the contributions of every sender surface point. Next, one would have to integrate the pressure field over the entire surface of the receiver. Unfortunately, integrals of this kind do not have closed form solutions and are computationally very demanding. However, there can be little doubt that the sender and receiver separation principle outlined in Sec. 4.4.1 in combination with Fig. 4.4 allows at least a rough estimate of the misalignment shift. For real time misalignment shift compensation some good approximate closed form solution may certainly be found.

Even this approach has some problems. Firstly, the transducer membranes may not move uniformly, or may not act as perfect spacial integrators. This is true in particular for transducers with large diameters. The incorporation of these effects into the compensation formula is very tedious. Secondly, it is difficult to calibrate the orientation of sender and receiver transducers. Nevertheless, there can be little doubt that a computed compensation will greatly reduce the bias.

A possible alternative method for the numerical compensation of the misalignment bias could be to use acoustical lenses or waveguides [55], [56] to focus the sound and thereby reduce the effective transducer size. Moreover, some additional measures can

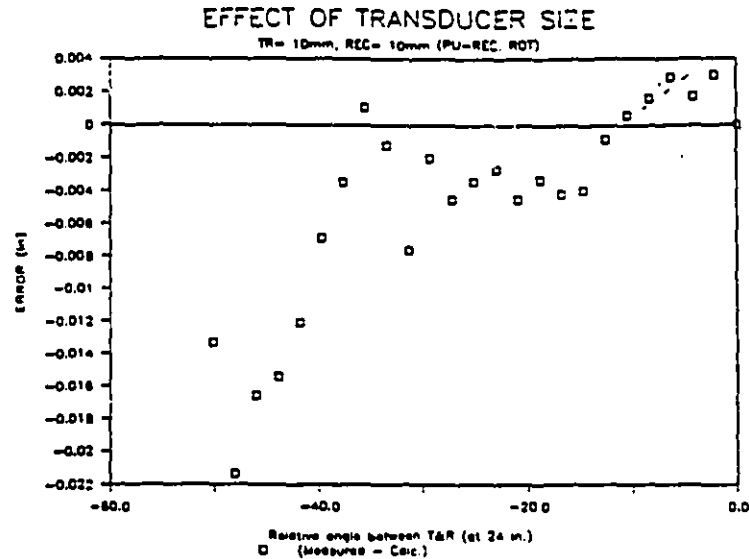


Figure 4.5: Center shift function for 10 mm transducer

be taken to reduce the bias. For instance, one could have many redundant receivers and choose those for the measurements which have the best alignment angle at the sampling instant. Also, when the transmitter is moving, especially when it involves a reorientation of the transmitter, some of the bias can be averaged out. This effect is enhanced if three or more senders are mounted on the manipulator hand with different orientations.

Given the complexity of the problem, it is useful to employ experimental methods to determine the misalignment shift. Experimental data about the misalignment shift can be found in [13] for 40 kHz transducers with 10 mm diameter and a 24 inc (61 cm) distance between the center points of the sender and the receiver transducer. The results were obtained by measuring the phase shift under steady state CW conditions. The receiver was oriented such that its axis formed an angle φ with the sender's axis. Now, the two axis do not align anymore, the sender staying put. The results are reproduced in Fig. 4.5 which shows the experimentally obtained misalignment shift (error) as a function of the relative angle φ between sender and receiver. According

to the authors of [13], the data in Fig. 4.5 have a maximal measurement error of 0.097 mm, which may explain the large data variance in Fig. 4.5. Nevertheless, the misalignment shift in Fig. 4.5 is very much larger than the earlier mentioned numerically obtained result of 0.2 mm. Even at $\varphi = 50^\circ$ and 24inc distance the shift reported in Fig. 4.5 is ca. 0.5 mm. The authors of [13] claim that this result is in accordance with the theory. The theory they refer to, however, is never explicitly stated in [13]. To resolve this dispute would likely require some further experimental verification, which, however, was not done at the present time.

Chapter 5

Deterministic Measurement Corruption

In this chapter, the the medium dependent deterministic disturbances mentioned in Sec. 4.1 are analyzed. Thus, atmospheric conditions like temperature drifts and temperature gradients, and wave-medium interactions like diffraction, refraction, reflection and doppler are discussed.

5.1 Atmospheric Biases and Drifts

5.1.1 Room Temperature Drift

Even a small change of the room temperature can seriously degrade the measurement precision: An unaccounted tiny change of just $0.2\text{ }^{\circ}\text{C}$ could cause an error in the order of 1mm on a 2 m measurement distance. This is unacceptable and therefore temperature drifts must be compensated for.

The graph 'A' in Fig. 5.1 shows typical TOF changes for a fixed sender-receiver distance in the small room. The range was 1.4 m and the sampling frequency was

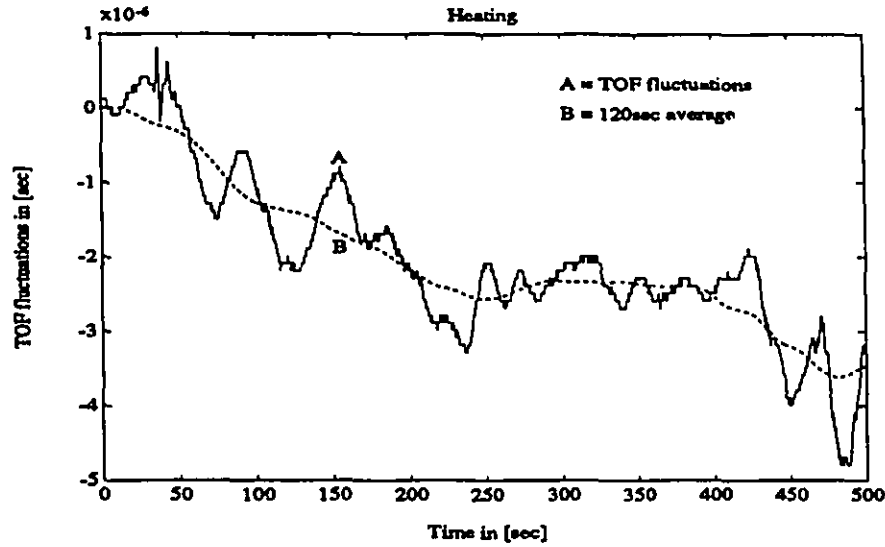


Figure 5.1: TOF fluctuations for 1.4m range while heating the small room.

4 Hz. The graph 'B' in Fig. 5.1 was obtained by filtering the measurements with a zero-lag non-causal 2 min moving average filter. Clearly, it is useful to separate the effects of the slow average temperature drift from the effects of the much faster turbulences. Thus, both will be modeled separately as outlined in Sec. 4.1: One as 'deterministic' and the other as 'random'.

Without turbulent air movement, the change of room temperature would be very slow. Suppose $L=5\text{m}$ is the characteristic length of the room and $\gamma=0.2\text{cm}^2/\text{sec}$ is the thermal diffusivity of the room. If only molecular diffusion is considered, then, according to [29], the scale of the characteristic heating time is $L^2/\gamma=100\text{h}$. This is *very* slow, indeed. If, however, there are convection currents and the largest eddy has an average velocity of say $u=5\text{cm}/\text{sec}$ and a characteristic size equal to the room ($L=5\text{m}$), then, according to [29], the scale of the heating time constant is $L/u=100\text{sec}$. This is indeed closer to the order of magnitude that was experimentally observed for heating and cooling of a room.

It follows from this that the magnitude and time constant of the room temperature

changes are closely correlated to the magnitude and time constants of the convection currents. Indeed, in the small room it was observed that a steep change of the room temperature was always accompanied by turbulences with large magnitude. Conversely, it is possible to speed up heating by increasing the turbulences in the room. For example, operating a heater with a strong fan can warm a room very rapidly (and very turbulently). Large rooms, on the other hand, generally exhibit larger heating time constants than small rooms. They are much less subject to fast changes in temperature. For example, in the large room, no appreciable average temperature drift was detected over any 500 sec interval. This should be compared with Fig. 5.1 for the small room. In general, the measurement system should be located as far away from heat sources or heat sinks as possible.

5.1.2 Temperature Gradients

Overview

Unfortunately, the temperature is not evenly distributed throughout the entire room. There is a slowly changing or, virtually constant temperature gradient. It is mainly caused by the natural convection of the air in the room. Usually, the air near the ceiling tends to be warmer than the air near the floor. Even though these temperature differences are quite small, usually less than 1°C , they do seriously affect the TOF measurements: If one does not compensate for the gradient, then in the worst case a constant gradient of say $0.5^{\circ}\text{C}/\text{m}$ over a measurement distance of 2 m would cause an error of about 2 mm. This is about an order of magnitude larger than the precision of interest.

Hence, one has to compensate for the temperature gradient. Unfortunately, there is no easy way short of simulation to predict the exact temperature and velocity distribution for a completely general architectural enclosure. Most rooms, however,

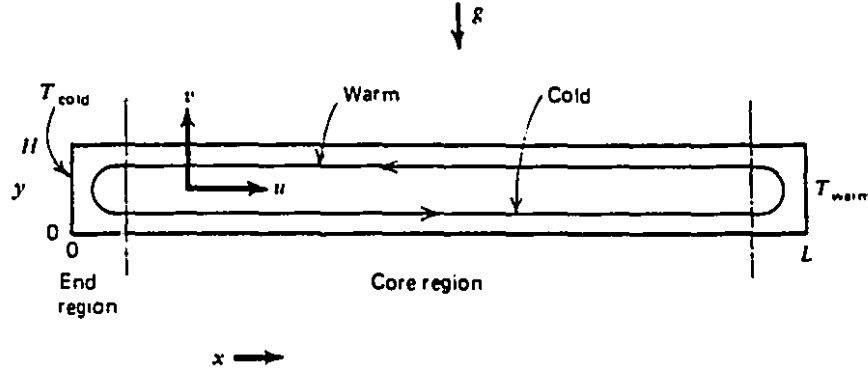


Figure 5.2: Boundary conditions for model 1.

have a rectangular shape, mostly shallow, and possess walls that may have different temperatures or may allow different heat flux. For these idealized situations there exist some simple solutions for the core region of such rooms. The solutions for regions close to the walls are generally more complicated, and are of no central interest to the purpose of this study.

First Convection Model

The first convection model assumes that two opposite vertical walls of a shallow room are kept uniformly at different constant temperatures T_{cold} and T_{warm} , where $\Delta T = T_{warm} - T_{cold} > 0$. Further, it is assumed that the remaining walls do not allow any heat flux or material diffusion through their surfaces. The situation is depicted in Fig. 5.2. Following reference [30], which should be consulted for details, one can define the following dimensionless quantities: $T_c = (T - T_{cold})/\Delta T$, $x_c = x/L$, $y_c = y/H$ and u_c and v_c are the properly normalized horizontal and vertical air velocities. In terms of these quantities, the core solution for the temperature is given as a fifth order polynomial in y_c and is linear in x_c and the horizontal velocity is given as a

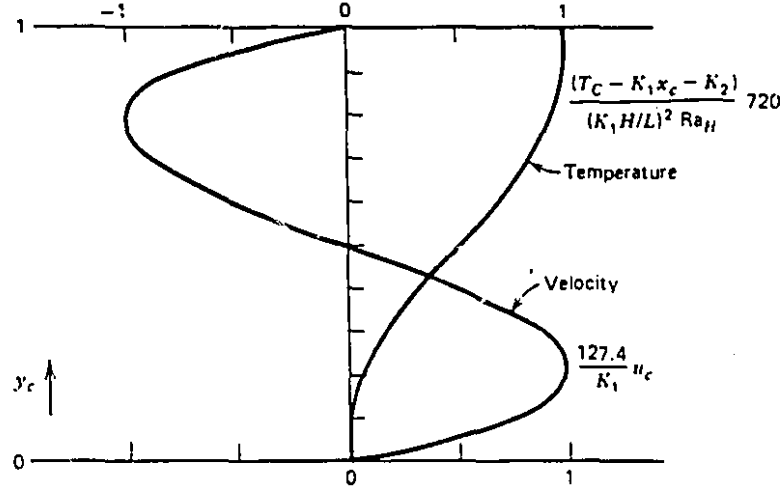


Figure 5.3: Vertical temperature for model 1.

third order polynomial in y_c :

$$T_c(x_c, y_c) = k_2 + k_1 x_c + k_1^2 \left(\frac{H}{L}\right)^2 Ra_H \left(\frac{y_c^5}{120} - \frac{y_c^4}{48} + \frac{y_c^3}{72}\right) \quad (5.1)$$

$$u_c(y_c) = k_1 \left(\frac{y_c^3}{6} - \frac{y_c^2}{4} + \frac{y_c}{12}\right), \quad v_c = 0 \quad (5.2)$$

where the Rayleigh number is

$$Ra_H = g c_1 \Delta T H^3$$

and the k_1 , k_2 and c_1 are known constants.

The vertical distribution of the temperature and the velocity as described in the above formulas are reproduced from [30] in Fig. 5.3 and the horizontal temperature distribution is reproduced in Fig. 5.4. As can be seen, the temperature distribution is symmetrical with respect to the point $(x, y) = (L/2, H/2)$ and the velocity distribution is a counter flow and symmetrical with respect to a plane through $y = H/2$ which is parallel to the floor. It is clear from Fig. 5.4, that for very shallow rooms the

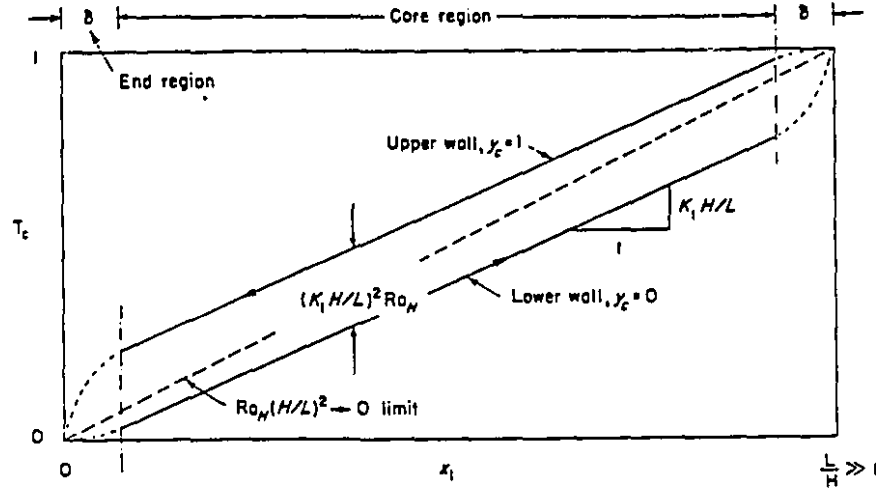


Figure 5.4: Horizontal temperature for model 1.

gradient will be small in the vertical direction and large in the horizontal direction. As the factor $(H/L)^2 Ra_H$ increases and the room becomes less shallow, the vertical temperature gradient becomes predominant.

Second Convection Model

The second model does not assume that the vertical walls at $x = 0$ and $x = L$ are isotherms; rather, it is assumed that the heat flux through the vertical walls is uniform. The temperature distribution of the wall will then be a result of a heat flux from the outside, rather than be a given and the wall temperature will increase with altitude. This assumption is believed to model the actual situation in many buildings more closely than the first model does. The problem is illustrated in Fig. 5.5, which is reproduced from [30]. In this problem the vertical temperature gradient in the core region is a constant and the horizontal gradient is zero; the horizontal and the vertical

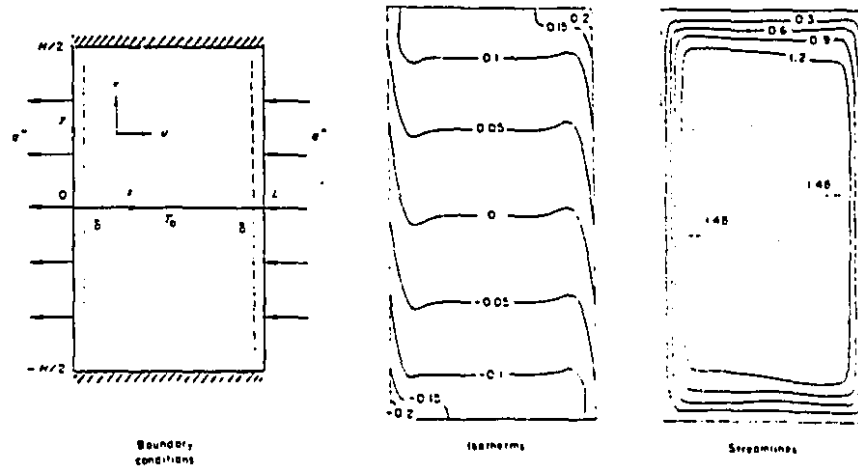


Figure 5.18 Rectangular enclosure with uniform heat flux from the side walls: isotherms and streamlines for $Ra_{a,H} = 1.5 \times 10^4$, $Pr = 7$, $H/L = 2$.

Figure 5.5: Boundary conditions, isotherms and streamlines for Model 2.

velocity are also zero:

$$\frac{\partial T}{\partial y} = \frac{c_2}{H^4} \left(\frac{H}{L} \right)^{4/9} Ra_{a,H}^{8/9}, \quad \frac{\partial T}{\partial x} = 0 \quad (5.3)$$

$$u = 0, \quad v = 0 \quad (5.4)$$

where the Rayleigh number is

$$Ra_{a,H} = g c_3 q H^4$$

and q is the heat flux and c_2 and c_3 are known constants. The Fig. 5.5 depicts the numerically simulated isotherm contours for the above model.

There are, of course, many other theoretical models dealing with different room shapes and different thermal boundary conditions. Most of these results, however, are only in the form of numerical simulations. In particular, much work has been done for the case in which the floor is warm and the ceiling is cold instead of the vertical walls, see for example [31]. This situation may arise when a room is uniformly heated

through the floor. In this case, other than for rooms that are heated from the side, there has to be a minimal ΔT before convection starts and then it evolves in the form of Bénard cells. For large ΔT very complicated stable patterns can arise.

Experimental Temperature Profile

In order to test the two convection models, and to get a practical picture of typical temperature distributions in a real room, it was decided to make some measurements which are hoped to be in some way representative.

To do this, it was necessary to measure temperature differences with a resolution better than a tenth of a degree Celsius. Since such a thermometer was not readily available, it was decided to use the TOF measurement of the ultrasound system itself as a temperature measurement device. This has the additional advantage that one measures directly the variable which is relevant for the position estimation and treats temperature as a quantity which is derived from the TOF measurements via the equations (1.1) and (1.4).

The measurements were done with two sender-receiver (S-R) pairs. As precisely as possible with a ruler, each receiver was placed 30 cm away from its sender. The two S-R pairs were placed at the same height, parallel to and flush with each other at a sender-sender distance of ca. 10 cm, which for mechanical reasons was the smallest possible. Then the TOF, for each pair, was measured with a sampling rate of 4 Hz, and the difference was averaged over 4 minutes. Then the approximate 30 cm distance of one of the two S-R pairs was adjusted until the absolute value of the averaged TOF difference was reduced to a very small value ϵ_0 , which then was subtracted from all future TOF difference measurements. The two S-R pairs were fixed onto two boards. If one pair is now moved away from the other, then any change in the averaged TOF difference must be caused by spacial sound speed differences.

The two S-R pair method has the advantage to make the measurements robust

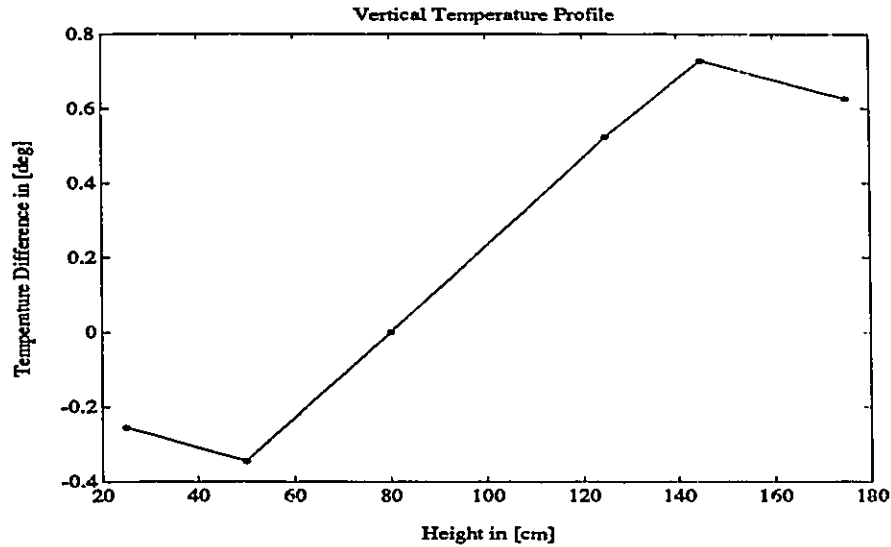


Figure 5.6: Experimental vertical temperature profile of small room.

with respect to temperature drifts during the long measurement process. For example, when the average room temperature increases slightly between measurements, say by 1°C , then the averaged TOF difference decreases only by ca. $\epsilon_0/290$. Without the reference pair, it would be difficult to distinguish between a temperature drift and a temperature gradient.

The Fig. 5.6 shows the vertical temperature profile in the middle of the small room as measured with the method explained above. The maximum vertical temperature difference was 1.1°C , which is large enough to cause position measurement errors in the millimeter range if not compensated for. The temperature gradient is essentially a constant $1^{\circ}\text{C}/\text{m}$ in the midsection of the room height, but sharply changes its value and even its sign in the regions close to the ceiling and close to the floor. Such a behavior can be reconciled with the theoretical convection models: The assumption made in the theoretical models, that the ceiling and the floor do not permit heat transfer, appears not to be a very good approximation of the situation in this particular room; rather, the room seem to loose some heat through the ceiling and gain

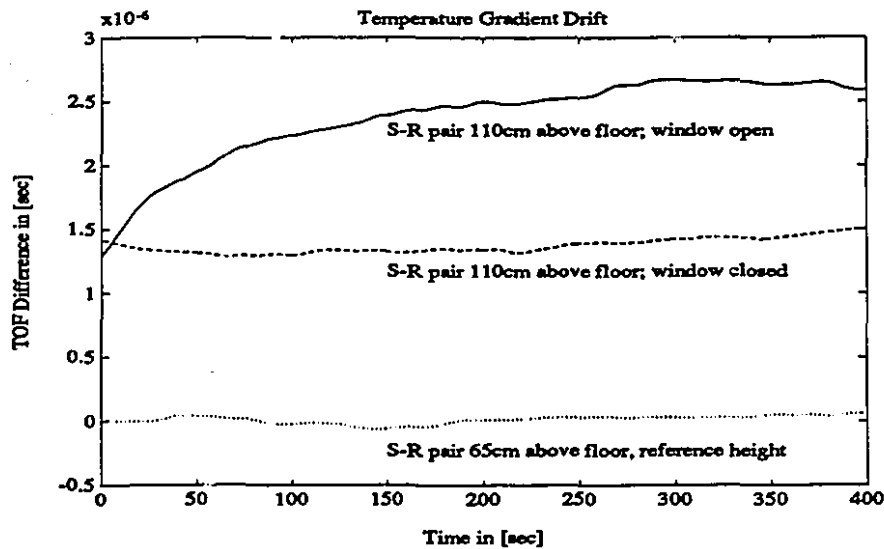


Figure 5.7: Dynamical behavior of temperature gradient.

some heat through the floor. This could account for the sharp gradient change in the end regions.

Measurements were also done in order to determine possible gradients in the two horizontal directions. However, in the core region of the room, no conclusive results were obtained: All “gradients” were in the order of the measurement error. The absence of a significant horizontal gradient leads one to the conclusion that the second convection model is indeed a better description for regular rooms. Of course, this does not mean that the horizontal gradients can be neglected in all rooms.

Experimental Temperature Gradient Drift

The Fig. 5.7 shows the dynamic behaviour of the temperature gradient. The graphs were obtained by filtering the TOF differences with a 100 sec moving average filter. The dotted graph depicts the time behavior of the averaged TOF differences when the two pairs are at the same height. As expected, there are only very small deviations from the zero line.

The dashed graph in Fig. 5.7 shows the time behavior of the averaged TOF differences when one of the pairs was moved to a position that was 45cm higher than the position of the reference pair, which was located 65cm above the floor. It can be seen that the averaged TOF differences do not change much with time. The average TOF difference stays at about $1.4\mu\text{sec}$ for this height difference. The corresponding temperature difference is ca. 0.8°C .

On the other hand, one would expect the TOF difference to increase when either the temperature difference ΔT or the heat flux q increases, depending on what theoretical model one chooses. This can indeed be observed. The solid graph in Fig. 5.7 shows the time behavior of the averaged TOF differences for the same setup as before, but now the window of the room is slightly open, thereby increasing both, the heat flux q and ΔT . The open window decreases the average temperature, but increased the temperature gradient. As can be seen, the TOF difference increases steeply, and then levels off at a value of ca. $2.5\mu\text{sec}$, corresponding to a 80% increase of the gradient.

It should be remarked, that in practice, a change of the temperature gradient would hardly be achieved without a change in temperature. Considering the theoretical models, one would reasonably expect that the time constants for temperature changes are similar to the time constants for temperature gradient changes.

The above experiments were not repeated for the large room. It can be expected, however, that the temperature gradient will be less pronounced in the large room. If the second theoretical model holds, then one would expect the vertical gradient to be inversely proportional to the room height. Thus, a reasonable estimate of the vertical gradient for the large room would be about $0.5^\circ\text{C}/\text{m}$.

5.1.3 Wind Biases

Air movement is a serious problem for the ultrasound system. In the worst case, when wind direction and measurement direction are aligned, an uncompensated 20 cm/sec wind would cause an error in the order of about 1 mm for a 2 m measurement distance. Moreover, wind was observed to be always accompanied by a high level of spacially weakly correlated turbulences. As will be shown later, even if the average wind is compensated for, a high level of spacially weakly correlated turbulences may severely limit the precision of the 3D ultrasound range finder. In other words, an environment where one must compensate for an average wind is not a good place for the operation of a high precision ultrasound range finder.

Fortunately, the effects of constant winds on the measurements is usually not of as much concern as is temperature. One reason is that, other than temperature biases, wind can often be avoided by moving or shielding the measurement system from the wind source, or by removing the wind source itself. Moreover, furniture and other obstacles in the room may act as natural wind barriers.

5.2 Wave-Medium Interaction

5.2.1 Diffraction

The equations which describe the propagation of sound [28] can be considerably simplified if they can be replaced by ray theory. It is well known [58], that this approximation is valid if the size of the first Fresnel zone $\sqrt{\lambda L} \ll a$, where $\lambda=8.5\text{mm}$ is the wavelength, $L=2\text{m}$ is the maximal range of interest and a is the spacial correlation distance, which is shown in Sec. 6 to be in the order of 25-90cm. Thus, ray theory is a valid approximation for this case and diffraction can usually be neglected. A special case may arise when small objects are blocking the line of sight between sender and

receiver and the sound signal diffracts around them.

5.2.2 Refraction

It is well known that a sound-speed gradient will cause the sound rays to follow a curved path in space instead of a straight line. This is referred to as refraction. In the following, it will be shown that refraction can be neglected for the type of sound gradients encountered in normal rooms. In this context, two explicit forms of (4.6) are presented: The velocity form and the inverse velocity form.

The Velocity Form

Given the equations (1.1) and (1.4) and in the results of Sec. 5.1.2 one can well approximate the speed of sound at a point ζ as:

$$c(\zeta) = c_r + c_\nabla(\zeta - S_r)_z$$

where c_r is the speed of sound at a reference position S_r , and c_∇ is the scalar vertical velocity gradient. Further, it was found in Sec. 5.1.2 that often the temperature gradient $\vartheta_\nabla < 1^\circ\text{C}/\text{m}$, which corresponds to $c_\nabla < 0.6 \text{ sec}^{-1}$.

Hence, the ray path will be curved because of the non-vanishing sound-speed gradient. For ranging purposes one is not so much interested in the shape of the ray path, but rather one wants to know the time the sound needs to travel from the sender at S to the receiver at M_i . In the absence of any other factors, the travel time can be computed from (4.6) as:

$$\text{tof}'_i = \int_0^{d'_i} \frac{ds}{c_0 + c_\nabla(\Gamma'_i(s) - S_r)_z} \quad (5.5)$$

where d'_i denotes the length of the parameterized curved ray path $\Gamma'_i(s)$. If one made the approximation $\Gamma'_i = \Gamma_i$, where Γ_i is a straight line as defined as in (4.1),

then one obtained:

$$\text{tof}_i = \int_0^{d_i} \frac{ds}{c_s + c_\nabla(\vec{e}_i)_z s} \quad (5.6)$$

where $c_s = c_r + c_\nabla(S - S_r)_z$ denotes the speed of sound at the sender S . For $c_\nabla \neq 0$ the above integral has the solution:

$$\text{tof}_i = \frac{\ln(|1 + \frac{c_\nabla}{c_s}(\vec{e}_i)_z d_i|)}{c_\nabla} \quad (5.7)$$

It is evident from ray theory that if $\nabla\Gamma'_i(0)$ is vertical i.e. has the same direction as the sound-speed gradient, then (5.7) is exact and the ray path $\Gamma'_i = \Gamma(i)$ is a straight line of length $d'_i = d_i$ connecting S and M_i . The more $\nabla\Gamma(0)_i(0)$ differs from a vertical, the more the ray path differs from a straight line and the less valid is the straight line approximation (5.7). Therefore, if (5.7) is a good approximation even for nearly horizontal cases, then it must be a good approximation for other angles as well. (If $c_\nabla(\vec{e}_i)_z = 0$, then (5.7) must be replaced by $\text{tof}_i = d_i/c_s$).

It is well known [28], that for a constant vertical velocity gradient and a horizontal initial angle the ray path is a circle segment with radius $R = c_s/c_\nabla$. Since the circle segment connects S and M_i which are d_i apart, the angle which spans the circle segment is $\phi_i = 2\arcsin(\frac{d_i}{2R})$. Moreover, $\nabla c^T e_i = -R(1 - \cos(\phi_i))/d_i$. Now, for $c_\nabla \neq 0$ we obtain from (5.5):

$$\text{tof}'_i = \frac{1}{c_\nabla} \int_0^{\phi_i} \frac{d\varphi}{\cos(\varphi)} = \frac{\ln(|\tan(\pi/4 + \phi_i/2)|)}{c_\nabla} \quad (5.8)$$

In order to compare (5.7) and (5.8) the following worst case scenario was assumed: $c_\nabla = 1\text{sec}^{-1}$, $c_s = 331.6\text{m/sec}$ and $d_i = 2\text{m}$. Under these conditions, $d'_i - d_i = 0.003\text{mm}$ and the approximation error is $\text{tof}'_i - \text{tof}_i = 9.1\text{e-}9\text{ sec}$. Clearly, the approximation error can be neglected because it is by about a factor 30 smaller than the desired precision. Thus, the bending of the rays due to refraction can be neglected and (5.5) can always be replaced by (5.7).

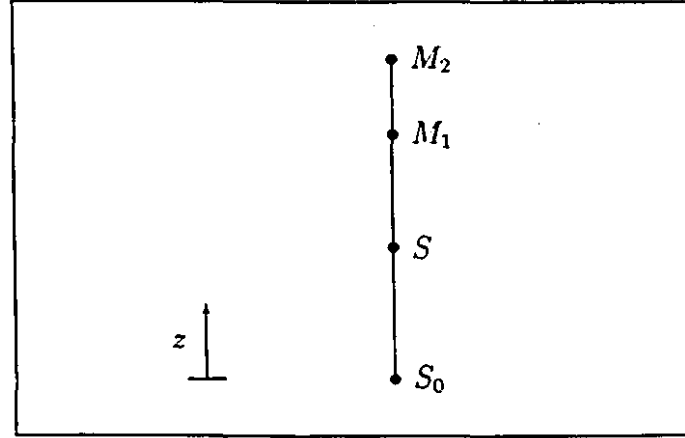


Figure 5.8: Bad case scenario for approximation.

The Inverse Velocity Form

Unfortunately, (5.7) is quite complex and may not be ideal for real time applications. Moreover there is a problem with a potential division by zero. Thus, the inverse-speed or index of refraction form of (4.6) appears to be more advantageous:

$$\text{tof}_i = n_r d_i + n_\nabla (S - S_r)_z d_i + \frac{1}{2} n_\nabla (\vec{e}_i)_z d_i^2 \quad (5.9)$$

where n_r and n_∇ are chosen such that (5.9) is exact for two vertically distinct sender or microphone positions. An advantage of (5.9) over (5.7) is that the former is linear with respect to the parameters n_r and n_∇ while the latter is non-linear with respect to c_r and c_∇ . Linearity is important because these quantities must be estimated by a Kalman filter. The linear measurement equation makes it possible to use a computationally advantageous constant gain filter.

Clearly, (5.9) is easier to compute in real time than (5.7), it does not have the problem of a potential division by zero like (5.7) and it permits the atmospheric variables n_0 and n_∇ to be estimated by a constant gain linear filter. Given the desirability of (5.9) the question arises as to how good an approximation it is for (5.7), which now is considered to describe the underlying physical reality correctly.

Clearly, the size of the factor $n_{\nabla}(S - S_r)_z$ in (5.9) is a measure for the approximation error. Therefore, one would expect the error to be large for large n_{∇} and large vertical distances between the sender position S and the reference position S_r . To get an idea of the size of this error a 'bad case' situation was considered as shown in Fig. 5.8. The microphone M_1 is positioned 3.5 m, M_2 is 3.2 m and the reference sender S_r is positioned 2 m above the main sender S . The following is assumed: A vertical velocity gradient of 1 sec^{-1} and a speed of sound of 343m/sec at S_r .

For this scenario, the norm of the error $\|e\|$ between the 'true' TOFs computed with (5.7) and the approximated TOFs computed with (5.9) was determined. The parameter n_r and n_{∇} in (5.9) were calculated from tof_1 and tof_2 for $S = S_r$. The approximation error for S as in Fig. 5.8. was then found to be $\|e\| = 4.1 \times 10^{-8} \text{ sec}$ which corresponds to ca. 0.014mm. Even for this extreme case the approximation error remains well below the precision of interest. Hence, in any normal situation it is possible to use (5.9) instead of (5.5) or (5.7). The three formulations are equivalent within the specified precision.

5.2.3 Ambient Noise

The ceramic ultrasound transducers which were used in the experimental system behave like a bandpass filter with center frequency at 40 kHz and a bandwidth of ca. 5 kHz. The receiver amplifier has a similar frequency characteristics. Thus, only sound with a frequency which lies in a narrow band around the carrier frequency has an appreciable influence on the measured TOF. Even if the ultrasound sender is switched off, there will probably be some residual 40 kHz sound in the air. This disturbance may come from many possible sound sources. See [27] for a study of the spectral composition and strength of typical industrial ultrasound emitting processes.

The study found that beyond 100 kHz the noise is negligible for ranging with two exceptions: Aerodynamic noise and laser etching. The situation is worse for 40 kHz.

It appears that most industrial processes have some kind of 40 kHz component. Much of this type of noise is impossible to quantify a priori and therefore must be measured in each individual case. Since it cannot be removed from the signal, the foreign carrier sources must be silenced or be removed from the room if the SNR(TOA) becomes too large.

If the sender is moving, which can be expected when it is attached to a manipulator, or if the foreign carrier source has a phase and amplitude which change fast compared to the sampling frequency, then of course the ambient noise would not cause a measurement bias, because it can be approximately averaged out. If, on the other hand, the sender does not move, and the ambient noise has a steady state harmonic component at 40kHz, then a fixed bias of the measured TOF may result.

If the just described scenario ever contributed a significant bias to the overall measurement error, then it would be easy to prevent it. One could change very slightly the pulse period each time a signal is sent. If this is done in a pseudo random manner, with the mean at the nominal pulse period and the standard deviation of, say, a sixth of a carrier period, then a steady state foreign carrier would behave as if it were white noise with respect to the 'random' phase of the signal carrier.

5.2.4 Reflection and Reverberation

The worst effect which reflections can have, occurs when the reflected signal has the same or even a larger magnitude than the direct path signal and the TOA is so large that the reflection has a full impact on the TOF. In practice this may occur for once reflected signals, whose path is not much longer than the direct path. Therefore, TOF measurements corrupted by short path reflections are useless. As mentioned earlier, the TOA must be chosen small enough to provide reflection robustness.

Once sent, the carrier can remain present in the room for a long time. The sound partially reflects from walls, furniture, people etc., until all its energy is converted into

heat. The sound can reflect many times and finally the room is filled with a diffuse background sound level. This is usually called reverberation noise. The reverberation constant is a measure for the time necessary to reduce the noise level of a room by 60 dB after the sound source is switched off. The reverberation constant can be computed as a function of the room dimensions and the reflection coefficients of the reflecting surfaces [28].

In our case, it appears that only the first couple of reflections from the walls have an appreciable effect on the measurements. In fact, in this report the expressions 'reverberation', 'multi-reflections' or 'long path reflections' are used synonymously and do describe all at least once reflected signals whose path is so long, that they arrive at the receiver after, together with, or immediately before the next direct path signal.

Provided that there is no movement in the room, these multi-reflection signals from the regularly pulsing sender appear on the oscilloscope as regular patterns with little or no apparent phase or amplitude changes. However, the patterns seem to be highly susceptible to small changes in the room. Even a slight repositioning of a person's head was observed to change the patterns. Thus, in a normal environment, where persons and objects are moving, and in particular in a robotics context, where the ultrasound sender is moving most of the time, the reverberation noise can almost certainly be approximately averaged out and will not cause biases.

Nevertheless, a regularly pulsing static sound source like the transmitter, when used in a static room, could conceivably produce measurement biases due to the standing wave patterns of the multi-reflection signals. A remedy against this possibility would be the same as in the case of ambient noise. The sampling period could be changed pseudo randomly after each sampling instance. This way, a standing pattern could be avoided.

pulse rate	0° alignm.		90° alignm.	
	5th	8th	5rd	8th
5Hz	56dB	61dB	42dB	47dB
20Hz	56dB	61dB	42dB	47dB
40Hz	36dB	41dB	22dB	27dB
120Hz	31dB	36dB	17dB	22dB

Table 5.1: The effect of reverberation on the SNR for 30cm S-R distance in a small room.

Reverberation

Let us assume that $\text{SNR} \geq 1$, otherwise the measurement is pretty useless anyways and should be discarded. Under this assumption the worst case occurs when $\text{SNR} = 1$ and the reverberation noise is 90° out of phase with the signal at the time where the zero crossing is measured. The maximal induced TOF error Δt_r can be computed as a function of the SNR:

$$\Delta t_r(\text{SNR}) = \frac{c_0 \lambda}{2\pi} \arctan\left(\frac{1}{\text{SNR}}\right) \quad (5.10)$$

where λ is the wave length of the ultrasound. Thus, in the worst case Δt_r corresponds to a position error of ca. 1 mm.

The SNR depends on the pulse rate, the reverberation constant, the signal magnitude at the receiver and the signal length. When the sampling frequency is increased, then the reflected signals which arrive at the receiver together with the next direct path signal have larger magnitude because their path is shorter than before. When the signal length is increased, the energy available for reflections is increased. Similarly, when the reverberation time constant of the room increases, then the noise amplitude increases as well. Since the signal magnitude remains constant in all those cases, the SNR decreases. On the other hand, when the distance or the misalignment angle is increased, then the signal magnitude decreases but the noise magnitude stays constant. Thus, again, the SNR decreases.

In order to gain some idea of the order of magnitude of the reverberation error

for different operating conditions, the SNR was estimated from oscilloscope readings for different sampling frequencies and misalignment angles. The noise level was read immediately before the arrival of the ranging signal. The experiments were conducted for the small room and for the large room. Table 5.1 shows the estimated SNR in dB for the signal amplitudes measured at the 5th and at the 8th crest of the received wave package, and for the misalignment angles of 0 degree and 90 degree. The values in Table 5.1 were obtained in the small room, a sender-receiver distance of 30 cm, and a transmitter input signal length of 6 carrier periods. The results for the large room are about same except that for pulse rates higher than 20 Hz there appears to be a 6 dB improvement over Table 5.1.

As can be seen from Table 5.1, the SNR is approximately the same for 5 Hz and for 20 Hz and even for smaller sampling frequencies, which are not shown in Table 5.1. This suggests, that this disturbance is due to the receiver background noise, i.e. the transducer self-noise and the amplifier noise floor, rather than to reverberation. This explanation was supported by the persistence of this noise level even after acoustic insulation of the microphone. A future receiver may be able to reduce the background noise further.

Fortunately, other than in the case of reverberation noise, the receiver background noise was observed on the oscilloscope to undergo fast random phase shifts. This implies that it can be averaged out even if the sender and its environment were static. Moreover, the effect of the background noise on the measurement error is rather small. It could only become a noticeable factor when the received signal is extremely weak. For instance, the maximal measurement error caused by the receiver noise floor for a 2 m sender-receiver distance, a sender-receiver misalignment angle of 90° and a TOA at the 8th period is about $67 \mu\text{m}$. (This corresponds to 26 dB SNR.)

For both rooms, the reverberation dominated the receiver background noise when the pulse frequency was increased beyond ca. 25 Hz. As shown in Table 5.1, for

40 Hz, the noise level is increased by about 20 dB and even more for 100 Hz. Thus, at 40 Hz the maximal measurement error for the above extreme case problem would be a remarkable 0.62 mm (corresponding to 6 dB SNR) instead of just 67 μm for 20 Hz.

Extrapolations like the one above can be obtained from Table 5.1 by recalling that the noise level does not change when the distance or the misalignment angle are changed. These changes affect only the signal level. The signal level, however, is inversely proportional to the distance and its dependence on the alignment angle can be obtained from the radiation characteristic chart of the ultrasound transducer which is reproduced in Appendix C. For instance, a 90° misalignment error will cause a ca. 14 dB damping of the signal. Indeed, this is the observed factor in Table 5.1 between the SNRs of the 0° misalignment column and the 90° column. Thus, Table 5.1 enables one to predict the noise strength.

5.2.5 The Doppler Effect

It is well known that the Doppler effect changes the perceived frequency of a tone when the sound source or the air moves with respect to the receiver. This effect does not influence the measured TOF directly, but only indirectly by changing the TOA. Hence, if there is no compensation for the Doppler effect, then the computed TOF for the i -th microphone will have a Doppler induced error:

$$\Delta\text{tof}_i = - \left(\frac{k}{f_0} \right) \left(\frac{v^T e_i}{c + w^T e_i} \right)$$

where $f_0=40$ kHz, k is an integer such that the TOA is measured at the zero crossing after the k -th period, v denotes the velocity of the sender and w is the velocity of the wind.

Suppose that the TOA is measured after the 5th period, i.e $k=8$, the sender moves with a speed of 0.34 m/sec in the direction of the i -th microphone, there is

no wind and the speed of sound is 340 m/sec. In this case, the error will be 5/1000-th of a period, corresponding to 0.043 mm. This is a small and possibly negligible error. Theoretically, one could compensate for the Doppler effect, since one does have estimates of v , e_i , k and possibly even w . On the other hand, the dynamic control error for fast maneuvers will be much larger than the Doppler error. Only for slow movements will the controller need high precision measurements, but then the Doppler effect is negligible.

Alternatively, instead of trying to compensate for the Doppler effect, it could be used for measuring the velocity of the sender and thereby improving the tracking precision of the control system. However, this would require the ability to measure the received carrier period of the ultrasound signal precisely. At the present time, this was not attempted.

Chapter 6

Turbulences

6.1 The Stochastic Turbulence Model

In (4.4) the TOF measurements were divided into four components. This chapter concerns itself with the medium dependent random component $^{MR}\text{tof}_i(\bar{\omega}, t)$ of the TOF. For notational convenience, however, the prescript MR will be dropped throughout this chapter. Thus, $\text{tof}_i(\bar{\omega}, t)$ denotes the scalar random process which models the random fluctuations of the i -th. channel. These fluctuations are caused by small and short lived regions of space which contain air with slightly different temperature or wind velocity than their surroundings. These regions move in space and also change their form and content over time. This changes the distribution of the refractive index in the VOI. In a room, space-time fluctuations of the refractive index are generated by heat convection and conduction, turbulences, drafts, moving people etc.. It is evident that sound traveling through this fluctuating medium will experience slightly different traveling times depending on the time of departure and the regions of space which were crossed.

One can readily extent the scalar random process $\text{tof}_i(\bar{\omega}, t)$ to a random vector

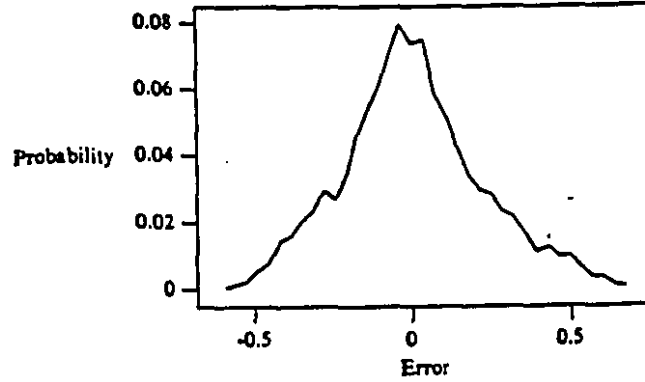


Figure 6.1: Distribution function of TOF fluctuations.

process $\text{tof}(\bar{\omega}, t)$:

$$\text{tof}(\bar{\omega}, t) = [\text{tof}_1(\bar{\omega}, t), \text{tof}_2(\bar{\omega}, t), \dots, \text{tof}_m(\bar{\omega}, t)]^T \quad (6.1)$$

which models all m measurement channels. In the following, the $\bar{\omega}$ -argument will sometimes be omitted when referring to stochastic processes.

The objective of the modeling effort is to characterize the random vector process $\text{tof}(\bar{\omega}, t)$, i.e one wants to know the joint probability distribution function F of $\text{tof}(\bar{\omega}, t)$:

$$F = F_{\text{tof}(t_1), \text{tof}(t_2), \text{tof}(t_3), \dots}(\vec{\xi}_1, \vec{\xi}_2, \vec{\xi}_3, \dots)$$

Note that F is implicitly a function of the microphone positions and the sender trajectory $S(t)$.

In the following we assume that F is jointly gaussian. Besides the central limit theorem, there is an other argument to support this assumption: A typical scalar distribution density function for one fixed sender and one microphone is given in [15] and is reproduced in Fig. 6.1. It approximates the shape of a gaussian distribution

density function. The result was experimentally obtained from the time series of TOF measurements. Hence, under the assumption of strict stationarity and ergodicity Fig. 6.1 supports the claim of a gaussian distribution function. Moreover, considering the underlying physical reality, it appears reasonable to assume that the whole random process tof is jointly gaussian for any combination of sender and receivers.

If the TOF is modeled as a gaussian random vector process, then F is completely characterized by its first and second moment. Since the low frequency biases are already subtracted, the first moment is the zero vector at all times. The second moment is determined by

$$P_{tt}(t_1, t_2) = E\{\text{tof}(t_1)\text{tof}^T(t_2)\} \quad (6.2)$$

where t_1 and t_2 are arbitrary points in time. The covariance kernel P_{tt} is an implicit function of the geometry of the ultrasound system. In the following, the nature of this relation will be analyzed.

The dependency of the TOF on the ultrasound system's geometry can be made explicit by modeling the fluctuating underlying inverse sound-speed itself as a random space-time process R_n as was shown in (4.2). Then the TOF can be linked to R_n and the ultrasound system's geometry via equation (4.5). Since R_n and the TOF are by definition zero average processes, the TOF is completely defined by its second moment:

$$\begin{aligned} (P_{tt}(t_1, t_2))_{i,j} &= E \left\{ \int_0^{d_i(t_1)} R_n(\bar{\omega}, t_1, \Gamma_i(t_1, \alpha_1), \tilde{e}_i(t_1)) d\alpha_1 \cdot \right. \\ &\quad \left. \int_0^{d_j(t_2)} R_n^T(\bar{\omega}, t_2, \Gamma_j(t_2, \alpha_2), \tilde{e}_j(t_2)) d\alpha_2 \right\} \\ &= \int_0^{d_i(t_1)} \int_0^{d_j(t_2)} P_{nn}(t_1, t_2, \Gamma_i(t_1, \alpha_1), \Gamma_j(t_2, \alpha_2), \tilde{e}_i(t_1), \tilde{e}_j(t_2)) d\alpha_2 d\alpha_1 \end{aligned} \quad (6.3)$$

where $(P_{tt}(t_1, t_2))_{i,j}$ denotes the (i,j) -th element of the covariance kernel P_{tt} , which

represents the covariance function between the i -th and the j -th communication channel. The scalar function P_{nn} is the covariance kernel of the inverse sound-speed. Thus, from (6.3) we can conclude that knowledge of P_{nn} and of all the ray paths Γ are equivalent to having knowledge of P_{tt} .

The key question is therefore the shape of P_{nn} . Unfortunately, there appears to be no literature that addresses itself directly to the problem of modeling turbulent sound propagation in a subregion of a closed room. On the other hand, there is a wealth of literature on sound propagation in a random ocean and on general random communication channels. Thus, the strategy used in this thesis is to apply the general theoretical methods to the problem and then determine the missing parameter values experimentally.

In order to simplify P_{nn} it was assumed that the process R_n is stationary and homogeneous, i.e. its statistics do not depend on absolute time and absolute position. This appears to be not too unreasonable an assumption since the process statistics are likely to change only very slowly and there is no pressing reason to assume that there are positions within the relatively small VOI (diameter 2 m) which have significantly different statistics from their neighbors. Furthermore, if one wants to have any hope of obtaining results of manageable complexity, one has to assume that in addition to being stationary and homogeneous, the statistics of R_n are also isotropic, i.e. independent of direction. The problem with this assumption is that it seems to imply a poor model for the effects of air movement. However, the main applications of the stochastic model is for cases where two ranges whose correlation one wants to find have similar directions. Moreover, experimental evidence suggests that the assumption is adequate for the coarse statistical model of tof that one can best hope to obtain.

With the assumptions made above, the covariance kernel P_{nn} is only a function of time delay and spatial shift and simplifies to:

$$P_{nn} = P_{nn}(\tau, \zeta) \quad (6.4)$$

where $\tau = |t_1 - t_2|$ and $\zeta = \|p_1 - p_2\|$ and t_1, t_2, p_1 and p_2 are arbitrary points in time and space, respectively.

Unfortunately, even the true shape of the strongly simplified covariance kernel $P_{nn}(\tau, \zeta)$ is unknown in the literature. Looking at the basic Physics of air turbulences in a room, one can, however, make the following qualitative remarks on the likely shape of P_{nn} : P_{nn} is bound to be small for large distances $\|p_1 - p_2\|$ and delays $|t_1 - t_2|$. Moreover, the turbulences which extend over a large region of space are the ones which are likely to have the longest life time.

In the absence of theoretical results, one can only follow the advice given in [57] and guess a general parametric form for P_{nn} and then choose the parameters to best fit the experimental data. First, P_{nn} is further simplified by assuming that one can split it into a time component and a space component:

$$P_{nn}(\tau, \zeta) = P_{nn}(0, \zeta)P_{nn}(\tau, 0) \quad (6.5)$$

The decomposition above has the drawback of not modeling the empirical observation that low-frequency turbulences are higher correlated than high-frequency turbulences.

Given the decomposition (6.5), there are two popular guesses for its space component:

$$P_{nn}(0, \zeta) = c_1 \exp(-|\zeta|/D_1) \quad (6.6)$$

$$P_{nn}(0, \zeta) = c_2 \exp(-(|\zeta|/D_2)^2) \quad (6.7)$$

The parameters c_1, D_1 or c_2, D_2 are unknown and must be determined experimentally.

To characterize the time component $P_{nn}(\tau, 0)$ it is convenient to switch to a discrete time description, where the sample instances $t_k = kT_s$ occur at integer multiples of the sampling period T_s . Suppose that the discrete time behavior of R_n is well modeled as the output of a discrete time-invariant linear filter (A, B, C) with state ρ of dimension r , whose input is white Gaussian noise with covariance Q . In other words, R_n is a r -th order Gauss-Markov process.

Thus, one can find the state covariance kernel $P_{\rho\rho} = E\{\rho\rho^T\}$ as the solution of the algebraic Lyapunov equation:

$$P_{\rho\rho} = AP_{\rho\rho}A^T + BQB^T$$

With $P_{\rho\rho}$, one can characterize the discrete time behavior of P_{nn} as:

$$P_{nn}(t_k - t_l, 0) = CA^{|k-l|}P_{\rho\rho}C^T \quad (6.8)$$

The parameter matrices A , C and BQB^T must be determined experimentally.

Under the assumptions made about the process, substituting (6.5) and (6.8) into (6.3) yields the desired covariance kernel P_{tt} ,

$$(P_{tt}(t_k, t_l))_{i,j} = CA^{|k-l|}P_{\rho\rho}C^T \int_0^{d_i(t_k)} \int_0^{d_j(t_l)} P_{nn}(0, \|\Gamma_i(t_k, \alpha_1) - \Gamma_j(t_l, \alpha_2)\|) d\alpha_2 d\alpha_1 \quad (6.9)$$

where the integrand of the double integral is given either by (6.6) or by (6.7). In the following it is often convenient to drop the time dependency of the double integral. One can then simply think of it as representing the covariance between two arbitrary static ranges:

$$P_{II}(\Gamma_1, \Gamma_2) = \int_0^{d_1} \int_0^{d_2} P_{nn}(0, \|\Gamma_1(\alpha_1) - \Gamma_2(\alpha_2)\|) d\alpha_2 d\alpha_1 \quad (6.10)$$

Except for special cases, (6.10) has no closed form solution when $P_{nn}(0, \zeta)$ is defined by either (6.7) or (6.6). However, an approximate solution if $P_{nn}(0, \zeta)$ is defined by (6.6) can be found in Appendix A.

In the following, the missing parameters in (6.9), are experimentally determined and the formula (6.9) is experimentally validated.

6.2 Experimental Evidence

6.2.1 Time Correlation

The Fig. 6.2 shows an example of a typical time series of TOF fluctuations for the small room for a 1m range. Similarly, the Fig. 6.3 shows typical TOF fluctuations for the large room for a 70 cm range. The sampling rate was 4 Hz. As can be seen, both time series are quite similar. Both are dominated by low frequency noise, which is probably due to the room's convection currents.

The power spectrum normalized by its largest value is shown in Fig. 6.4 for both time series. The Fig. 6.4 was obtained by smoothing the autocorrelation function of the time series with a Hamming window. It can be seen that both spectra have a very similar low-pass characteristic. The large room exhibits a slightly larger bandwidth and a somewhat less steep roll-off which is not surprising given the more turbulent environment of the large room. The bandwidth of the noise in Fig. 6.4 is ca 0.03-0.05 Hz.

It is shown in Fig. 6.5 that a 3rd order linear system with poles at $-0.03 \pm j0.031$ Hz and -0.73 Hz and zeros at -0.081 Hz and -1.5 Hz, produces a very good fit to the estimated spectrum of the large room. The estimated power spectrum is the same as in Fig. 6.4, however, it is less smoothed. The best second order approximation is also shown. It is not as good as the 3rd order fit. A best first order system approximations was not shown in Fig. 6.5 because its spectral fit was unacceptable. The identification results in Fig. 6.5 can be used to obtain the (A,B,C) state space parameterization used in (6.9).

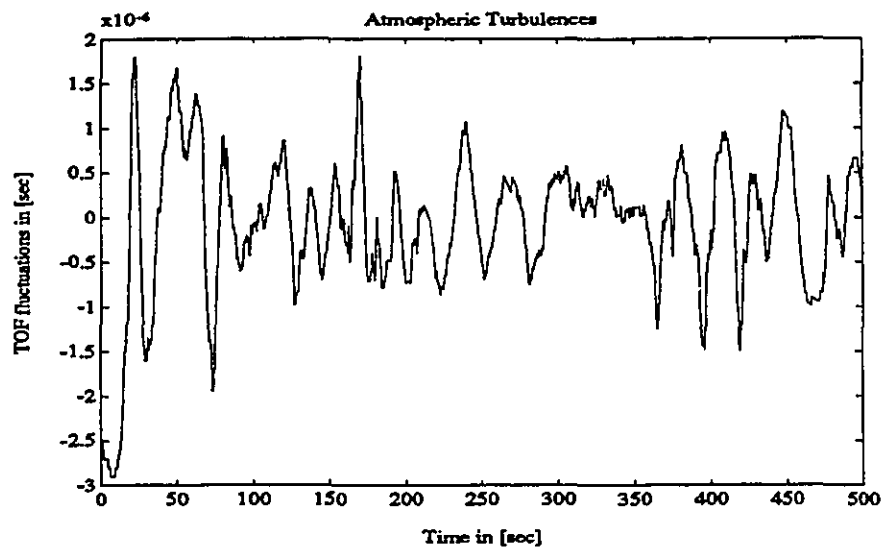


Figure 6.2: Time series of TOF fluctuations in small room.

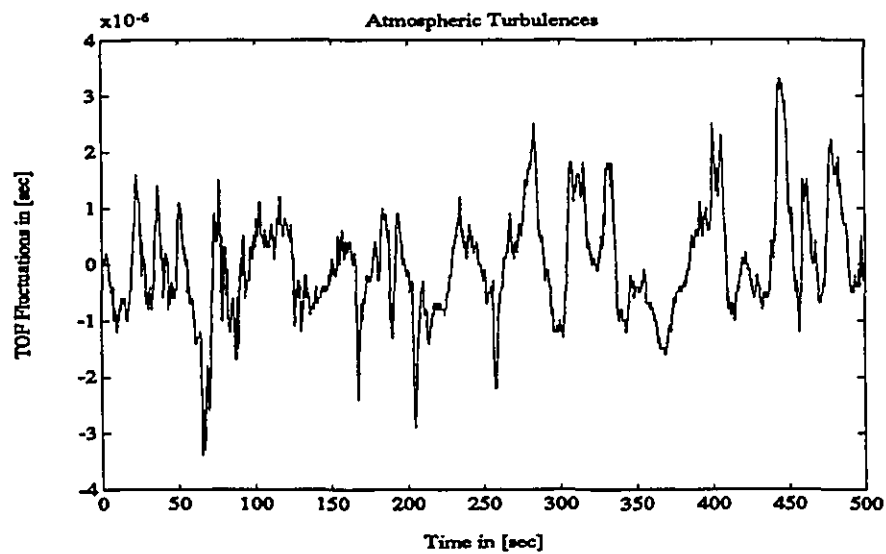


Figure 6.3: Time series of TOF fluctuations in large room.

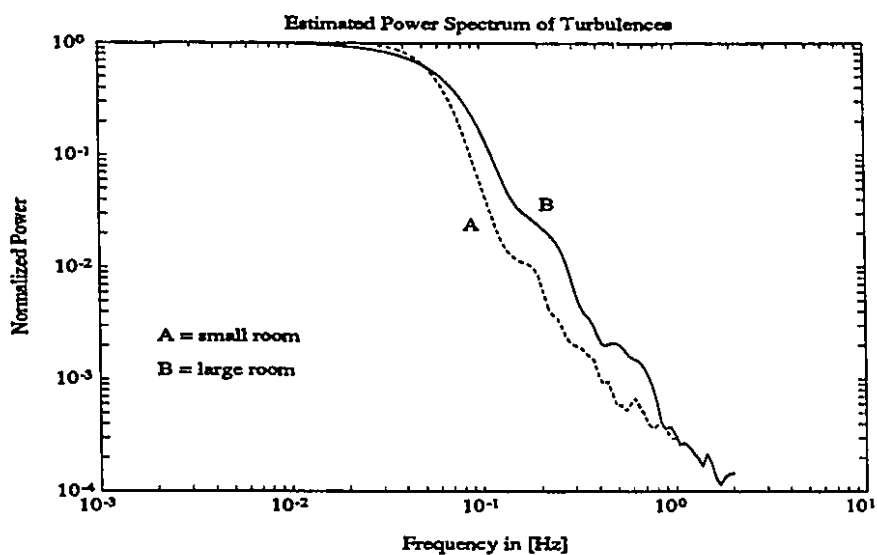


Figure 6.4: Power spectrum of TOF fluctuations in small and large room.

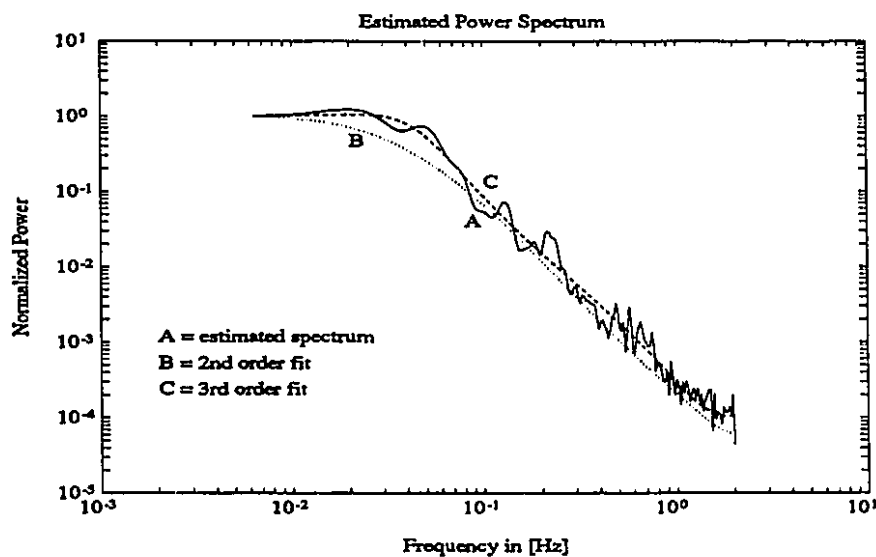


Figure 6.5: 2nd and 3rd order system approximation of TOF fluctuations in small room.

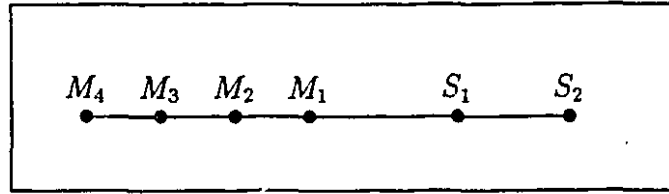


Figure 6.6: Longitudinal Correlation Experiment

6.2.2 Space Correlation

Noise Strength

P_{tt} is a complex non-stationary matrix valued function. In order to make the task of characterizing P_{tt} easier, consider first the most simple case, where one has only one fixed sender and one fixed microphone facing each other at a distance $d_i \in [20, 200]$ cm.

For our purposes, the most important statistical quantity is the standard deviation (STD) of tof_i . A wide range of STDs was observed for the TOF fluctuations in the small room: Very quiet air with very little convection or wind was observed to cause STDs as small as $0.06 \mu\text{sec}$ or 0.02 mm for 1 m ranges. Such low values were achieved by closing door and windows and disabling heaters and fans. On the other hand, very turbulent air, such as it is produced by opening door and window of the small room, on a cold and windy day, was observed to cause STDs even in excess of $2 \mu\text{sec}$ (or 0.68 mm) for 1 m ranges. The usual STD values for the normally heated small room were around 0.1 mm on 1 m ranges. The STDs in the large room were more stable: Despite of many operating fans, moving people etc., the STDs for 1 m ranges were about $0.5 \mu\text{sec}$ (or 0.17 mm) when the air conditioning system was off and about $0.9 \mu\text{sec}$ (or 0.3 mm) when it was operating. These values were quite stable in the large room despite moving people and computer fans. Thus, under reasonably calm atmospheric room conditions, one can expect the STDs of the corresponding distance measurement error for a 1 m range to fall somewhere between 0.02 mm and 0.7 mm with typical values around $0.1\text{-}0.2 \text{ mm}$.

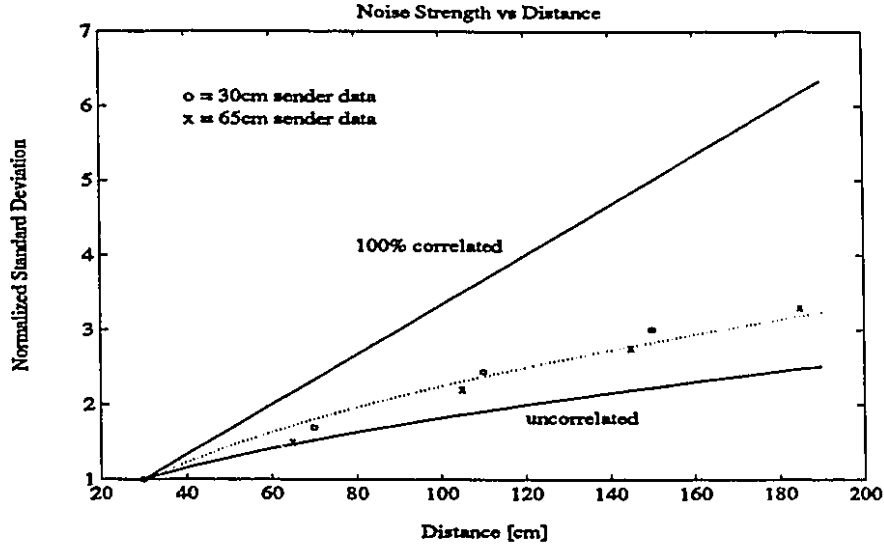


Figure 6.7: Noise standard deviation versus distance for large room.

Since the ultrasound system has to accommodate different ranges, one is interested to know $\sigma(d_i)$, which denotes the STD of tof_i as a function of the distance d_i between sender and microphone. Clearly, there are two limiting cases: If R_n is completely correlated in space, then from (6.10) and (6.6) one obtains $P_{nn}(0, \zeta) = c_1$ and:

$$\sigma^2(d_i) = c_1 d_i^2 \quad (6.11)$$

On the other hand, if R_n is completely uncorrelated, then $P_{nn}(0, \zeta) = c'_1 \delta(\zeta)$, where $\delta(\zeta)$ is a dirac function. Hence, from (6.3) one obtains:

$$\sigma^2(d_i) = c'_1 d_i \quad (6.12)$$

The true $\sigma(d_i)$ lies somewhere between these two extremes: Using (6.6) in (6.10), one obtains the form:

$$\sigma^2(d_i) = 2c_1 D_1^2 (d_i/D_1 + \exp(-d_i/D_1) - 1) \quad (6.13)$$

There is no closed form if (6.7) is used instead of (6.6).

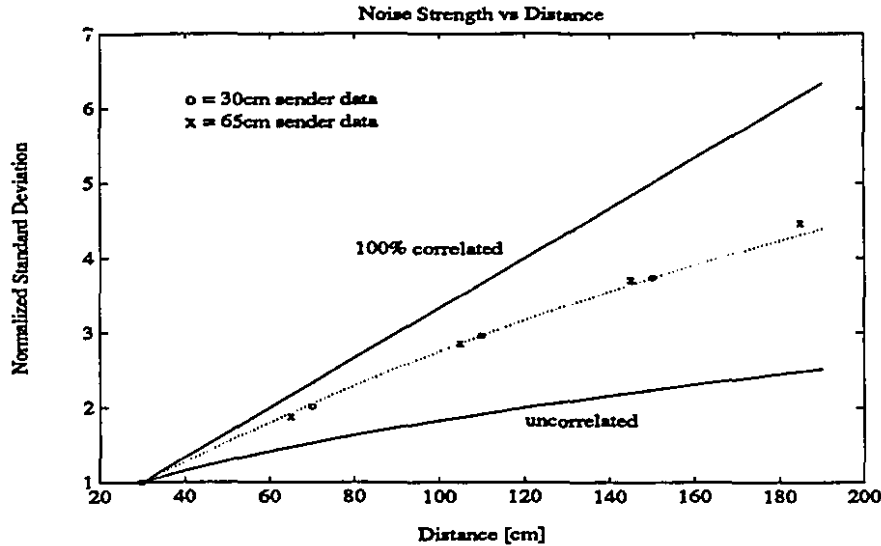


Figure 6.8: Noise standard deviation versus distance for small room.

An attempt was made to verify experimentally the validity of (6.10) for changing sender-receiver distances. For this purpose an experiment was set up as shown in Fig. 6.6: 4 microphones and 2 senders were placed in a row so that any two adjacent microphones were 40 cm apart and the distance between S_1 and M_1 was 30 cm and between S_2 and M_1 it was 65 cm. Great care was taken to ensure that distances between receivers and senders were as specified above. However, the measurement error may be as large as ± 1 cm. All statistical results were based on sample sizes of 2000 time samples for each of the 8 ranges and the data were obtained with a 2 Hz sampling rate.

For the large room Fig. 6.7 shows the STD of the 8 ranges as a function of the range. The STD of the 4 ranges associated with S_1 are marked by an 'o' and the 4 ranges associated with S_2 are marked by an 'x'. All STDs were normalized by the STD value of the S_1 - M_1 range. This was done to be able to compare (6.11), (6.12) and (6.10) with the experimental data. Each of these curves was normalized by its STD value for a 30 cm S_1 - M_1 range to eliminate the influence of the constants c_1 and

c_1' . In effect, Fig. 6.7 shows a comparison of the range dependency of the STD for totally correlated noise, totally uncorrelated noise, the actual experimental data and exponentially correlated noise which fits the data. To compare them one forces all to have the same STD for a 30cm range.

As expected, the experimental data are neither completely correlated nor completely uncorrelated but are sandwiched between the two limiting functions. The curve defined by (6.10) and (6.7) fits best to the data when $D_2=38\text{cm}$. This is shown in Fig. 6.7. As can be seen, the curve fits quite nicely. When (6.7) is replaced by (6.6) in (6.10) the best fit is obtained for $D_1=24\text{cm}$. The obtained graph fits equally well for this case, however, it is not shown in Fig. 6.7 to avoid confusion.

A similarly good fit was obtained when the experiment was repeated in the small room. The equivalent of Fig. 6.7 for the small room is Fig. 6.8. Clearly, the noise is much more spatially correlated in the smaller room. There, the correlation distances for the best fit to the experimental data were $D_2=90\text{cm}$ and $D_1=90\text{cm}$, more than double the correlation distances for the large room.

Longitudinal Correlation

In order to further verify the accuracy of the model (6.10), exactly the same ranging data were used as before. This time, however, instead of computing the STD versus distance, one computes the correlation coefficient between two ranges, as a function of the distance ratio $\gamma \geq 1$ between the two ranges. The results are shown in Fig. 6.9. The marker 'x' represents correlation coefficients which were obtained by correlating the TOF measurements for the 1st microphone and the 65 cm distant sender with the measurements obtained for the same sender and the 2nd, 3rd and 4th microphone. Similarly, the symbol 'o' represents correlation coefficients which were obtained by correlating the TOF measurements for the 1st microphone and the 30 cm distant sender with the measurements obtained for the same sender and the 2nd, 3rd and

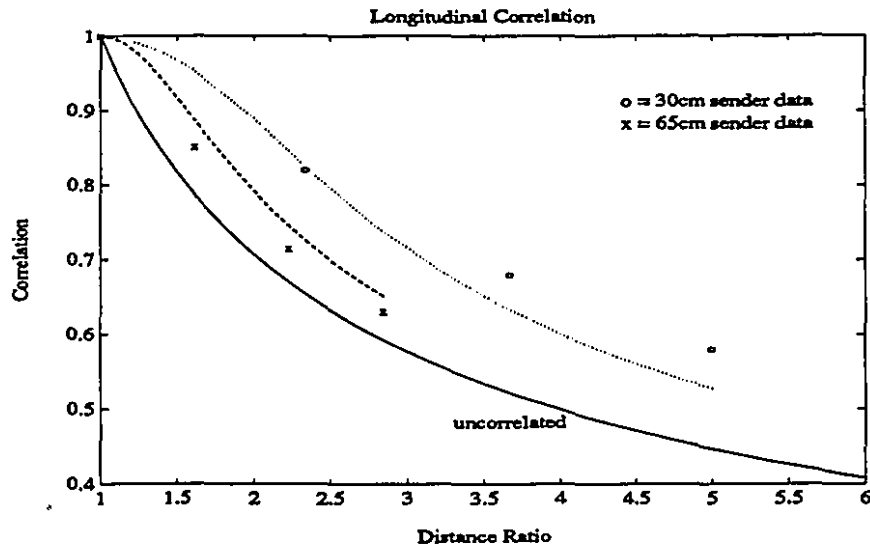


Figure 6.9: Correlation versus distance ratio for large room.

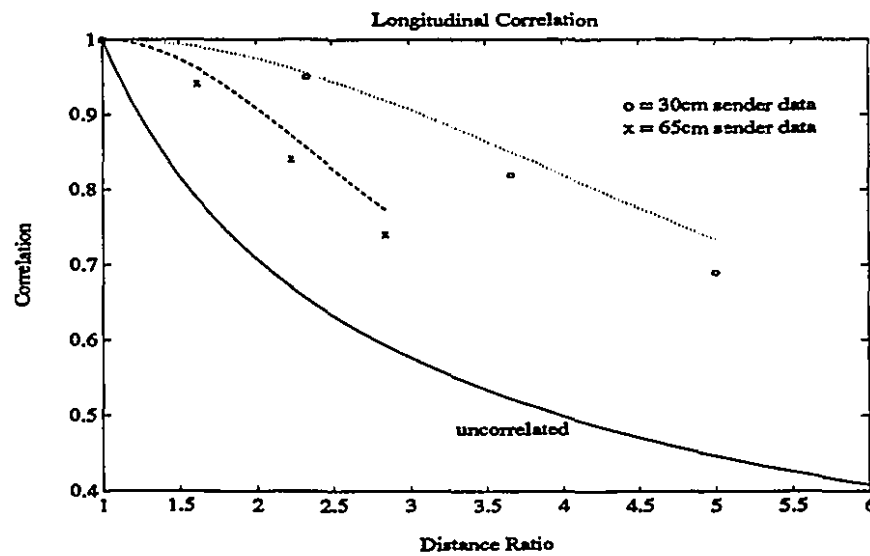


Figure 6.10: Correlation versus distance ratio for small room.

4th microphone.

As before, there are theoretical limits for completely correlated and completely uncorrelated R_n . The completely correlated case is obviously independent of γ :

$$C = 1$$

where C denotes the correlation. Of course, from (6.10) one can readily calculate the correlation between any two ranges Γ_i and Γ_j as:

$$C = \frac{P_{II}(\Gamma_i, \Gamma_j)}{\sqrt{P_{II}(\Gamma_i, \Gamma_i)P_{II}(\Gamma_j, \Gamma_j)}}$$

The completely uncorrelated case yields:

$$C = \sqrt{\frac{1}{\gamma}}$$

For exponentially correlated noise one obtains from (6.6) and (6.10) the equation:

$$C = \frac{[2d_i/D_1 + \exp(-d_i/D_1) + \exp(-\gamma d_i/D_1) + \exp(-(\gamma-1)d_i/D_1) - 1]}{2\sqrt{[d_i/D_1 + \exp(-d_i/D_1) - 1][\gamma d_i/D_1 + \exp(-\gamma d_i/D_1) - 1]}} \quad (6.14)$$

As can be seen from Fig. 6.9, the function (6.10) with (6.7) and $D_2=38\text{cm}$ does fit the data sufficiently well to prove that (6.10) is consistent for computations of the STD and the longitudinal correlation between ranges: The parameter value $D_2=38\text{cm}$ is valid for both computations. The reason for having two graphs to fit to the data and not just one like Fig. 6.7 is that the d is different for the two senders. Again, if (6.6) is used instead of (6.7) in (6.10) a similarly good fit is obtained for $D_1=24\text{cm}$. As before, this graph is not shown. Notice from the solid graph in Fig. 6.9, that even if R_n is completely uncorrelated, the longitudinal correlation between two ranges does not vanish.

The small room equivalent of Fig. 6.9 is Fig. 6.10. Clearly, the ranges are more space correlated. As before, the graphs show good fit for $D_2=90\text{cm}$ or $D_1=90\text{cm}$, respectively, proving the consistency of (6.10).

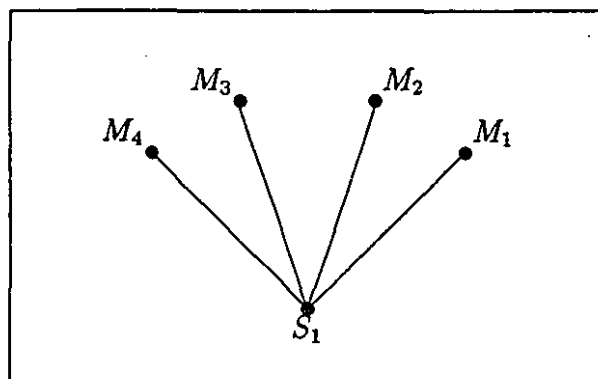


Figure 6.11: Transversal Correlation Experiment

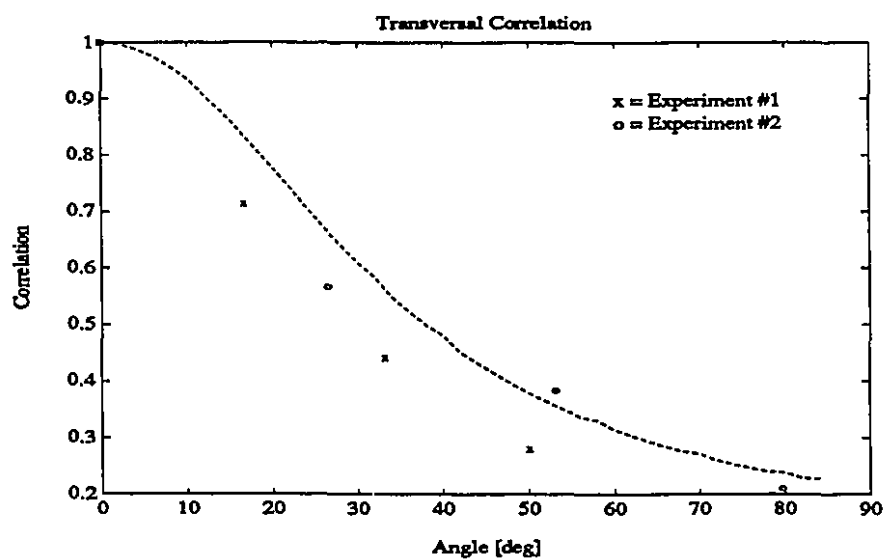


Figure 6.12: Correlation versus range angle in large room.

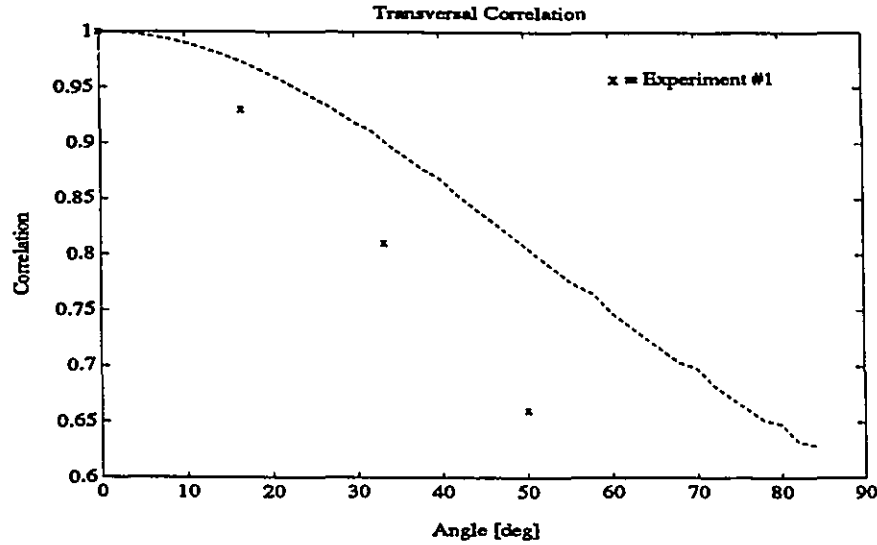


Figure 6.13: Correlation versus range angle in small room.

Transversal Correlation

Another pair of experiments was designed to verify (6.10) as a function of the range angle for $\gamma = 1$. As shown in Fig. 6.11 the microphones were placed on the circumference of a circle with the sender in the center. Each pair of neighboring microphones had the same fixed distance. The experimental results for the large room are shown in Fig. 6.12. The sample size and the sampling frequency are as for the previous experiments. All 6 cross-correlations were computed from the raw TOF data and averages were computed for cross-correlations corresponding to the same angle, yielding 3 cross-correlations. The Fig. 6.12 shows the results of two consecutive experiments conducted in the large room. For the first experiment, the distance between neighboring receivers was 30 cm and the distance to the sender was 1 m. The results of this experiment are marked with a 'o' in Fig. 6.12. The second experiment had the same receiver distance, but the distance to the sender was 1.40 m. Its results are marked by 'x'. Moreover, the theoretical curve obtained from the equation pair (6.7) and (6.10) for $D_2=38$ cm is also shown in Fig. 6.12. It can be seen, the curve does

not fit that well to the data obtained. The same ill fit occurs when (6.7) is replaced by (6.6) with $D_1=24$ cm (not shown). Similar problems occur for the small room equivalent of Fig. 6.12 which is Fig. 6.13.

The reason for the problem could be that the atmospheric conditions changed between the longitudinal experiments and the transversal experiments. The setup and the conduction of the experiments takes a considerable amount of time. An other explanation could be that the neglected effects of air movements decorrelate the noise faster than otherwise expected. In other words, this may be the price one has to pay for the isotropy assumption made in (6.9).

A further look at the spatial correlation of the TOF reveals that the longitudinal correlation is higher than the transversal correlation and that smaller ranges lead to higher correlations. Moreover, it should be said that it is intuitively clear that high frequency fluctuations are spatially less correlated than slower fluctuations. This is particularly important for the compensation of the low frequency noise with a reference filter. This aspect is not modeled in (6.7), rather, all frequencies have the same spatial correlation. The main reason for not modeling this aspect is the desire for simplicity.

It can be concluded that the experimental results confirm that (6.9) is self-consistent and provides an adequate although not perfect model of the spatial correlation of the measurement noise.

Part III

Taskspace Control with Noisy Measurements

Chapter 7

Noise Compensation Methods

7.1 Optimal Microphone-Sender Configurations

The relative position of the microphones and senders play an important role in the measurement noise reduction and the numerical conditioning of the system. For a small perturbation δS from a nominal sender position S_n , the measurement equation (1.1) can be written as:

$$\text{tof}_i - \frac{1}{c} \|S_n - M_i\| = \frac{1}{c} \frac{(S_n - M_i)^T}{\|S_n - M_i\|} \delta S + \eta_i \quad i = 1, 2, \dots, m \quad (7.1)$$

Where tof_i , c , and M_i are defined as before and η_i represents the measurement noise for the i -th receiver. The above equation can be written in vector form:

$$\delta \text{tof} = \frac{1}{c} H \delta S + \eta \quad (7.2)$$

where δtof is the measurement perturbation vector, H is a matrix whose rows are the unit directions from S_n to the microphones, and η is the vector of the gaussian measurement noise where $E\{\eta\} = (0, 0, 0)^T$ and $E\{\eta\eta^T\} = Q_n$.

The linear minimum-variance unbiased estimate $\hat{\delta S}$ of δS is well known [59]:

$$\hat{\delta S} = (H^T Q_n^{-1} H)^{-1} H^T Q_n^{-1} c \delta \text{tof} \quad (7.3)$$

and the corresponding error covariance is:

$$P_{ss} = (H^T Q_n^{-1} H)^{-1} \quad (7.4)$$

Clearly, one is interested in reducing the covariance of the estimation error. Thus, it seems reasonable to define the "optimal" microphone positions H_{opt} to be those which minimize a norm on P_{ss} . Unfortunately, any such optimization of microphone positions is valid for one nominal sender position S_n only. Any movement of the sender away from S_n will result in a suboptimal configuration.

For simplicity, it is first assumed that all microphones are at an equal distance to S_n and that the measurement noise is uncorrelated, i.e. $Q_n = \sigma_n^2 I$, where σ_n^2 is the variance of the measurement noise for a single channel and I is the identity matrix. Without loss of generality we may set $S_n = (0, 0, 0)^T$ and choose the first microphone to lie in the direction of the z-axis of a right-handed Cartesian coordinate system with origin at S_n . Further, w.l.g., the second microphone can be restricted to lie on the x-z-plane. Moreover, using spherical coordinates to describe the microphone positions, one can explicitly fix the radius component of each channel to be unity.

With these assumptions, the problem becomes one of finding the vectors φ and ϑ which minimize the trace of $P_{ss}(\varphi, \vartheta)$, where φ and ϑ denote vectors of the angles of the microphone directions expressed in spherical coordinates. This is done in such a way, that the i-th row (φ_i, ϑ_i) of the spherical angle vector pair denotes the x-axis if its value is $(0^\circ, 0^\circ)$, the z-axis if it is $(0^\circ, 90^\circ)$ and the y-axis if it is $(90^\circ, 0^\circ)$. Now, one seeks to find φ and ϑ such that the cost J is minimized:

$$J_{opt} = \min_{\varphi, \vartheta} \|P_{ss}(\varphi, \vartheta)\|_F^2 = \min_{\varphi, \vartheta} \text{trace}(P_{ss}). \quad (7.5)$$

The Matlab Nelder-Mead algorithm was used to find the optimal (φ, ϑ) for 3, 4 and 5 microphones respectively. The optimal configurations are given in Table 7.1. A sketch of the optimal configurations for 3 and 4 microphones is given in Fig. 7.1.

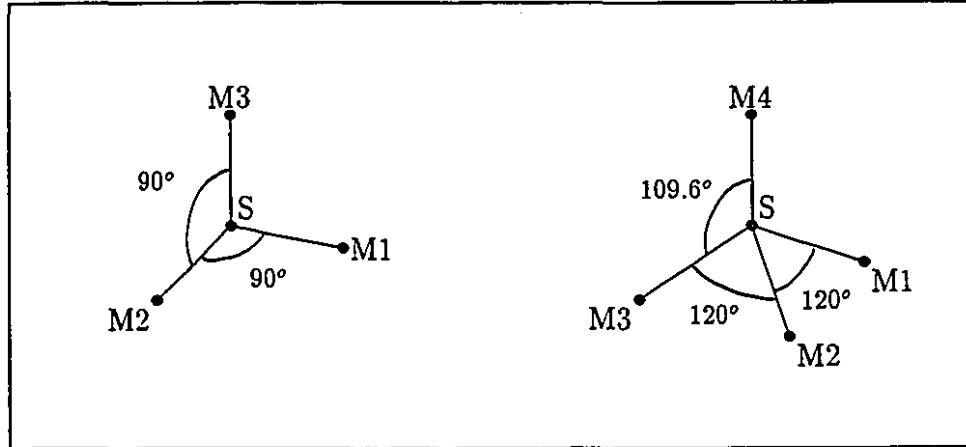


Figure 7.1: The optimal S-R configurations for 3 and 4 microphones.

	3 microphones		4 microphones		5 microphones		6 microphones	
	J=3.00		J=2.25		J=1.83		J=1.50	
	φ	ϑ	φ	ϑ	φ	ϑ	φ	ϑ
M1	0°	90°	0°	90°	0°	90°	0°	90°
M2	0°	0°	0°	-19.6°	0°	-19.6°	0°	0°
M3	90°	0°	120°	-19.6°	120°	-19.6°	90°	0°
M4	—	—	-120°	-19.6°	-120°	-19.6°	-90°	0°
M5	—	—	—	—	-120°	-19.6°	0°	-90°
M6	—	—	—	—	—	—	180°	0°

Table 7.1: Optimal microphone positions

Besides being optimal in the previously defined sense, the configurations in Table 7.1 with more than 3 microphones have the additional property that any additive disturbance which is common to all channels has no influence on the sender position estimation error. This is important since one would like $\delta\hat{S}$ to be insensitive to global changes of the speed of sound and to additive delays caused by the electronics and a nonzero TOA. Both goals can be readily satisfied if all microphones have the same distance to the sender and if one chooses H_{opt} such that the vector $(1, 1, \dots, 1)^T$ is an element of the nullspace of $H_{opt}^T Q_n^{-1}$. The configurations in Table 7.1 have this property and are therefore completely insensitive to common mode noise. The only exception is the 3 microphone case where the nullspace of $H_{opt}^T Q_n^{-1}$ is trivial. If the microphones are not equidistant to the sender, then additive delays and global changes of the speed of sound cannot be compensated for simultaneously.

The optimal configurations in Table 7.1 were obtained for spatially uncorrelated noise and in the absence of sound-speed gradients. If both these factors are considered, then the optimal configurations differ slightly from the results given in Table 7.1.

7.2 The Kalman Filter

7.2.1 Overview

The purpose of the Kalman filter is to estimate the wrist position y_c given the noisy range measurements. The size of the estimation error depends to a large degree on its internal model. The question as to what constitutes a good model is best answered by looking at the variables which influence the TOF. Thus, using (5.9) one can write:

$$\text{tof}_i = \|M_i - y_c\| \left[n_r + \frac{1}{2}(M_i + y_c - 2S_r)_z n_\nabla \right] + \eta_i + w_i \quad (7.6)$$

where η_i is the time correlated noise caused by air turbulences as discussed in Sec. 6 and w_i is a small white noise caused by the receiver noise floor and the room reverberation as discussed in Sec. 5.2.4. Thus, from (7.6) one can see that the states of a TOF measurement model fall into three categories:

1. The main states which model the wrist position y_c .
2. The parameter states n_r and n_∇ which were defined in Sec. 5.2.2.
3. The noise states η_i and w_i which model the measurement noise.

The second category of states are by definition spatially highly correlated. This suggests that one could decompose the overall state into two smaller ones by adding a fixed reference sender. This solution is preferable because it increases the estimation precision and reduces the overall system complexity. Moreover, if the parameter states were actually estimated by the main filter, then full observability would require that the main sender is received by 5 microphones at all times, which may be difficult to achieve. The other solution requires only 3 microphones for the main sender. The third group of states, the noise states, are not as highly spatially correlated as the parameter states and can therefore only partially be transferred to the reference filter.

7.2.2 The Reference Filter

The main purpose of the reference Kalman filter is to estimate the parameters n_r and n_∇ . Another purpose is to estimate $\hat{\eta}_{r,i}$ which is the noise in the i -th measurement channel of the reference system. The estimate is later used to reduce $\hat{\eta}_i$. To accomplish these objectives, a reference sender is installed at a known fixed position y_r in the workspace of the robot. Since y_r and M_i are constants, the measurement equation of the reference Kalman filter is linear with respect to its states $\hat{\eta}_{r,i}$, $\hat{n}_r$ and $\hat{n}_\nabla$. From

(7.6) one obtains for $S_r = y_r = y_c$:

$$\text{tof}_{ri} = \|M_i - y_r\| [\hat{n}_r + \frac{1}{2}(M_i - y_r)_z \hat{n}_\nabla] + \hat{\eta}_{ri} + w_{ri} \quad (7.7)$$

- where the index 'ri' refers to the i-th communication channel of the reference system.

Given the analysis in Sec. 5.2.2, it is clear that $\hat{n}_r$ and $\hat{n}_\nabla$ are nearly constant in a controlled environment. Thus, they can be modeled as random walks with very small pseudo noise. Similarly, the dynamics of $\hat{\eta}_{ri}$ can be modeled by a linear, time invariant filter (A,B,C). This was discussed in Sec. 6.2.1. Therefore, the reference Kalman filter is time invariant. It can be implemented as a fixed gain filter, which is computationally advantageous.

Since the fastest time constant the reference system is concerned with is about 30 sec, one can reduce the sampling rate for the reference Kalman filter. A 2 Hz reference filter sampling rate, for instance, would be enough. This makes the reference filter computationally very inexpensive. Hence, one can easily install more than one reference system to cover different regions of the workspace. This is particularly important because in general one can not guarantee that all receivers are reached by a single reference sender.

An additional task of the reference system could be to provide the statistics of the atmosphere. In particular, the spatial correlation distance, the noise strength at 1 m distance, and the correlation time. These three numbers suffice to characterize the noise. This could be the task of some identification scheme which would run in parallel with the reference system.

7.2.3 The Main Filter

The Measurement Equation

The purpose of the main Kalman filter is to estimate the Cartesian position y_c of the wrist given noisy TOF measurements and nominal wrist position y_n . In other words,

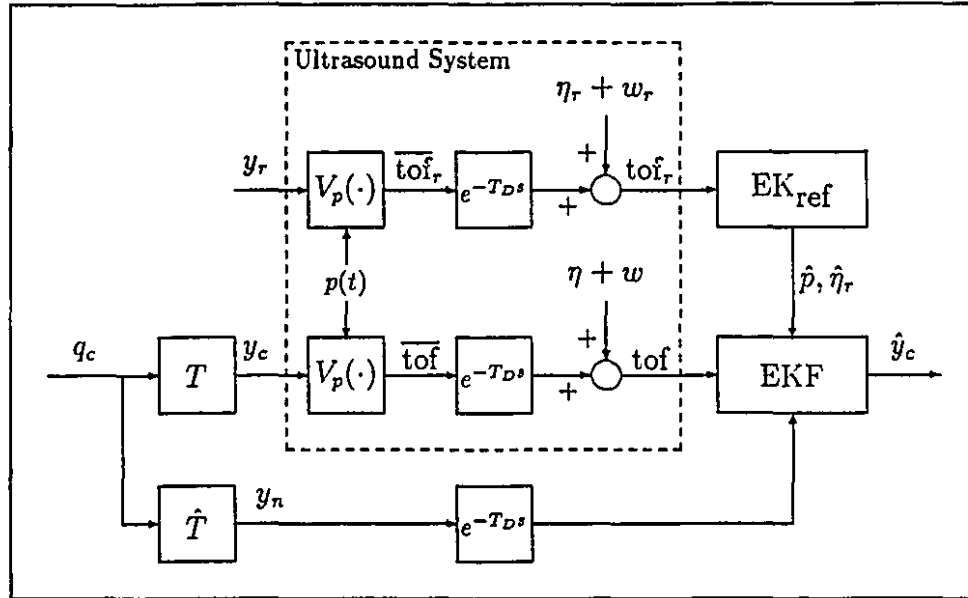


Figure 7.2: The Kalman filter and the ultrasound system.

the true forward kinematic function T which maps the true joint positions q_c into y_c is replaced by an arrangement shown in Fig. 7.2. There, q_c is mapped into $\hat{y}_c$, where $\hat{y}_c$ is an estimate of y_c . The estimates $\hat{y}_c$ will be used in lieu of the true positions y_c in the kinematic feedback loop. Hence, the size of the overall Cartesian tracking error depends critically on the size of the estimation error e_f , which is defined as $e_f = \hat{y}_c - y_c$. The accuracy of the position estimation scheme is therefore measured in terms of e_f . Since one knows a nominal position y_n , it is advantageous to estimate the kinematic error $e_c = y_c - y_n$ instead of y_c .

The diagram shown in Fig. 7.2 depicts how the robot, the ultrasound system (indicated by the dashed box) and the Kalman filter are connected. The robot's wrist position y_c can be considered to be an input of the ultrasound system. Conceptually, the wrist position is then mapped by a nonlinear, memoryless vector valued function $V_p(\cdot)$ into the uncorrupted TOF delays $\overline{\text{tof}}$. The function $V_p(\cdot)$ depends on the slowly

time varying parameters $p(t)$. They are: The average inverse sound-speed n_r measured at the reference point y_r , the average vertical inverse sound-speed gradient n_∇ (and possibly the average wind velocity W). In the following, however, W will not be modeled. The mapping V_p was discussed in detail in Sec. 5.2.2. Explicitly, the i -th component of $V_p(y_c)$ is the deterministic part of (7.6).

The signal is also delayed. The time delay T_D is quite small, namely ca. 5.8 msec, and represents the time the sound needs to travel the longest allowed range (2m). The processing of the data begins only after this time. For small pulse rates, T_D can be neglected with respect to the pulse period. A much more serious corruption is the measurement noise η . This was discussed in detail in the second part of this thesis. The so corrupted signal, denoted tof in Fig. 7.2, can be measured and serves as an input to the extended Kalman filter EKF.

In order to improve the estimation precision of the EKF a reference Kalman filter KF_{ref} was used in parallel with the EKF. This is reflected in Fig. 7.2 by the parallel signal paths for the main and the reference filter. The notable difference between the two is that the reference sender position y_r remains fixed for all times while the main sender position y_c moves with the manipulator. The fixed sender position allows KF_{ref} to estimate the parameter states and the noise states more easily than the main filter. It then passes its estimates $\hat{n}_r$, $\hat{n}_\nabla$ and $\hat{\eta}_r$ on to the EKF.

Another important system component which improves the estimation accuracy of the EKF is the approximate forward kinematics function $\hat{T}$ which maps q_c into the nominal wrist position y_n . This was discussed in Sect. 2.1.4. The nominal trajectory allows the EKF to track only the kinematic error and not the whole trajectory. It is necessary to delay y_n by the acoustic transmission delay T_D before feeding it into the EKF. When this is not done an unnecessary error due to the delay is introduced.

The measurement equation for the main filter can be derived from (7.6):

$$\text{tof}_i = \|M_i - \hat{y}_c\| \left[\hat{n}_r + \frac{1}{2}(M_i + \hat{y}_c - 2y_r)_z \hat{n}_\nabla \right] + \hat{\eta}_i + w_i \quad (7.8)$$

The parameter estimates $\hat{n}_r$ and $\hat{n}_v$ are obtained from the reference Kalman filter. The estimates $\hat{y}_c$ and $\hat{\eta}_i$ belong of the main filter. However, instead of using $\hat{y}_c$ as a state, it is advantageous to use the kinematic error $\hat{e}_c = \hat{y}_c - y_n$.

The State Propagation Model

It is difficult to characterize the kinematic error e_c spectrally. In fact, without additional a priori knowledge of the error parameters and the trajectory there is no information to model e_c at all.

One attempt to deal with this problem is to simply use a double integrator as model for the spectrum of the kinematic error. Thus, the filter state is $[\hat{e}_c, \dot{\hat{e}}_c]$ and the time propagation of the state is:

$$\begin{bmatrix} \hat{e}_c(t_{k+1}) \\ \dot{\hat{e}}_c(t_{k+1}) \end{bmatrix} = \begin{bmatrix} I & T_s I \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{e}_c(t_k) \\ \dot{\hat{e}}_c(t_k) \end{bmatrix} \quad (7.9)$$

where T_s is the sampling period, and the associated covariance matrix is:

$$Q = \begin{bmatrix} IT_s^3/3 & IT_s^2/2 \\ IT_s^2/2 & IT_s \end{bmatrix} q_a \quad (7.10)$$

where q_a is a factor that can be associated with the variance of the kinematic error acceleration $\ddot{\hat{e}}_c$.

Polynomial filtering, similar to the double integrator model proposed above was successfully used in other applications: For instance, there are many parallels between the ultrasound ranging problem and the NAVSTAR/GPS satellite navigation problem described in [54]. There, the range and position information of four satellites was used to determine the position of an airplane. The main states were position, velocity and acceleration of the airplane. The ultrasound ranging problem is quite similar at the first look. However, there is one important difference: The noise that corrupts the measurements is of very low frequency and often coincides spectrally with the kinematic error.

Thus, even if the true kinematic error model was known and both the measurement noise states and the kinematic error states were included in the internal process model of the EKF, the estimation error could still be quite large. The reason is that if e_c has about the same bandwidth as η , then the states are poorly observable and the EKF's performance is quite low. A covariance analysis was conducted which supports this observation.

Usually, however, the spectral behavior of e_c is not known. If the double integrator model is used instead then it is not necessarily advantageous to include the measurement noise states into the internal process model of the EKF. In fact, for some circular test trajectories (narrow band kinematic errors) the performance with the noise states included was worse than without them.

An additional problem with the use of the noise states is that their dynamic behavior depends on the entire trajectory of the sender. In a sense, they fail to have a finite dimensional state space representation. This becomes a serious problem for fast trajectories which cross large regions of the workspace. It is further discussed in the next chapter.

Therefore, a better solution is not to use any noise states at all. Essentially, this means that the high frequency part of the measurement noise is averaged out, but the powerful 30 sec noise cannot be reduced much unless an extremely slow trajectory is used, or the robot stands still for several minutes. This form of measurement noise reduction can be advantageously used for calibration purposes, but is inadequate for trajectory tracking.

Hence, without noise states one would expect the estimation error to be about as good as the noise itself. In the large room (without the air conditioning running), the std of the noise at 1 m sender-receiver distance was about 0.15 mm. Taking the 3 std value, one can expect a maximum estimation error of ca. ± 0.5 mm. This is still better than the ± 0.8 mm reported in [2] for the static accuracy of a PUMA750 after

calibration. Moreover, the estimation error is the same for fast and slow trajectories because the averaging of the high frequency noise does not result in a significant error reduction.

When the double integrator model is used then the only thing one must know about the kinematic error e_c is some bound on its current acceleration in order to choose a good value for the state noise strength q_a . Such a bound can be obtained by writing $\tilde{e}_c$ as:

$$\tilde{e}_c = \Delta J J^{-1} \tilde{y}_c + \frac{d(\Delta J J^{-1})}{dt} \dot{y}_c$$

where $\Delta J = J - \hat{J}$. Since $y_d \approx y_c$ one can replace y_c by y_d in the above equation. Thus a bound on $\tilde{e}_c$ is:

$$\begin{aligned} |\tilde{e}_c| &\leq |\Delta J J^{-1}| |\tilde{y}_d| + \left| \frac{d(\Delta J J^{-1})}{dt} \right| |\dot{y}_d| \\ &\leq a_2 |\tilde{y}_d| + a_1 |\dot{y}_d| \\ &\leq a_2 |\tilde{y}_d| + a_0 |\dot{y}_d|^2 \end{aligned}$$

The problem of finding a bound is now reduced to finding those positive factors a_2 and a_0 (or a_2 and a_1) which are not too conservative. Clearly, near a singularity J^{-1} can become arbitrarily large and thereby cause a_2 and a_0 to be very large. However, for the trajectories and kinematic errors simulated in this thesis it turned out that the choice $a_2=0.005$ and $a_0=0.005$ yields good results when $q_a = |\tilde{e}_c|^2$.

Concluding, it can be remarked that the double integrator model is a very simple and computationally inexpensive model, and it requires very little information about the kinematic error. The advantage of having an extended Kalman filter, as opposed to simply solving sets of nonlinear equations, is that it smoothes the kinematic error estimate, uses times of stand-still or slow-down to reduce the noise, provides a convenient framework for sensor fusion, and ensures operation even during short periods of partial unobservability like blockage of a microphone or even a sensor failure.

7.2.4 Exploiting Spatial Correlation

In order to improve the estimation accuracy one can use the reference filter to reduce the time correlated measurement noise of the main filter. It is evident from (7.7) that the reference filter produces estimates $\hat{\eta}_{ri}$ of its time correlated measurement noise η_{ri} . Similarly, it is clear from (7.6) that the main filter has the time correlated measurement noise η_i . The two scalar random processes η_{ri} and η_i are highly correlated when the moving sender and the reference sender are close to each other. This fact can be exploited to reduce the variance of the noise entering the main Kalman filter.

In general, suppose there are N_r reference senders and associated reference Kalman filters and the j -th filter produces an estimate $\hat{\eta}_{ri,j}$ of its measurement noise for the i -th receiver. A vector v_i of all N_r estimates can now be formed:

$$v_i = (\hat{\eta}_{ri,1}, \dots, \hat{\eta}_{ri,N_r})^T$$

Now one can subtract a linear combination of the elements of v_i from the measured tof_i in (7.6). In effect, one replaces the old η_i by a new scalar random process ζ_i :

$$\zeta_i = \eta_i - K v_i$$

where K is a row vector of length N_r . The vector K should be chosen s.t. $E\{\zeta_i \zeta_i^T\}$ is minimized. The optimal estimator K is well known to be:

$$K = E\{\eta_i v_i^T\} E\{v_i v_i^T\}^{-1}$$

The minimum achievable variance for ζ_i is:

$$\begin{aligned} E\{\zeta_i \zeta_i^T\} &= E\{\eta_i \eta_i^T\} - E\{\eta_i v_i^T\} E\{v_i v_i^T\}^{-1} E\{\eta_i v_i^T\} \\ &= E\{\eta_i \eta_i^T\} (1 - E\{\eta_i \eta_i^T\}^{-1} E\{\eta_i v_i^T\} E\{v_i v_i^T\}^{-1} E\{\eta_i v_i^T\}) \\ &= E\{\eta_i \eta_i^T\} (1 - C_r^2) \end{aligned}$$

For $N_r = 1$ (the scalar case) C_r is the correlation factor between η_i and η_{ri} .

The expected values in the formulas above can be computed from the formula given in Appendix A for $\hat{\eta}_r = \eta_r$. This is a complicated formula which may not be well suited for real time applications. Thus, it may be recommendable to use a simplification instead.

Suppose that y_r and y_c have the same distance to a microphone. Their noise variances are therefore also the same. Under these conditions one would require a correlation of $C_r = 0.9$ to reduce the maximal estimation error by 56%. Incidentally, this is about the same correlation that the measurement noise was found to have if y_r and y_c are separated by the spatial correlation distance b . Recall that b was observed to be ca. 24 cm in the large room and 52-91 cm in the small room. Thus, the main sender traveling in the large room on a path which has always a distance of 20 cm to the reference sender would reduce the ± 0.5 mm maximal estimation error to ca. ± 0.2 mm. An even higher reduction (up to 100%) can of course be achieved if the sender comes even closer to the reference sender.

Thus, the region around a reference sender can be viewed to be a “high precision region”. So, the reference sender could be placed in regions where high accuracy is needed. It is even imaginable to have more than one reference sender. This raises the question as to the ability of the system to “hold” the high accuracy obtained in a region of high precision while the robot ventures outside of that region. This has to be answered by simulations.

7.3 Calibration

An important aspect of noise reduction is robot and microphone calibration. Before the 3D ultrasound position sender can be used it is essential to know the precise position of the microphones. The best way to obtain this knowledge is to move the sender to several well known positions in the convex hull spanned by the microphones.

This can be done automatically with the use of a programmable positioning table. At any given location the measured TOF for each microphone is averaged over 2-3 minutes and then recorded. Preferably, the room is atmospherically well controlled during the calibration. Finally, the averaged TOF values are used to solve a linear set of equations similar to (1.2) but with S and M changing roles. The minimum number of sender positions required for a linear solution is four. However, it is preferable to use more sender positions in order to reduce the microphone calibration error even further.

One problem with the method above is that it neglects electronic delays. However, it is easy to take delays into account and add more sender calibration positions. Another problem is that the method requires that there are no drifts of the speed of sound or the sound speed gradient. This is difficult to achieve even for a well controlled room.

Another application of the 3D ultrasound position sensor is as a convenient tool for automatic parametric calibration of the robot. The robot is commanded to move to a specified number of Cartesian positions. For each commanded position the true position and the joint displacements are recorded. To increase the precision with which the true position can be measured, one must stand still for a while to allow the EKF to average out the measurement noise. Finally, the true Davenit-Hartenberg parameters can be computed from the obtained data set.

One disadvantage of this method is that the inverse kinematic function can be very complex and even fail to have a closed form. A remedy for this problem is to use nonparametric calibration instead. The idea is to command the robot to execute a given trajectory. However, the speed with which the trajectory is executed is drastically reduced, allowing the EKF to average out the measurement noise and determine the kinematic error along the trajectory. Now, whenever this trajectory is executed at its regular speed, the kinematic error can be compensated. This can work

well for a highly structured environment, where the same task must be performed over and over again. It is also applicable to robots in environments, where the trajectory can change slightly from run to run, or when there is a set of trajectories to choose from.

Chapter 8

Simulations

8.1 Simulation Parameters

This chapter reports on several simulation experiments that were conducted to demonstrate the properties of the Cartesian feedback loop and the Kalman filter. The simulations are meant to be tools to gain new insight into the system's behavior and to illustrate and support claims about the system's performance which were derived theoretically. They are not meant to be proofs in any strict sense. Therefore, only those simulation results are presented which achieve the stated objective.

To this end it is important to use a good truth model which resembles as closely as possible the real life environmental factors which influence the results. On the other hand, the computational requirements for the simulations should stay within reasonable limits. This may lead to a less complex model which neglects minor effects. The decision as to which effects are minor and which are not was based on the analysis presented in chapters 4-6.

The following environmental truth model was used for the simulations: The *refraction* and *diffraction* of sound were neglected, so was any *ambient noise*. As well, strong *reflections* of sound were also not modeled because they would render the

measurements useless. On the other hand, *sound reverberation* was modeled as white noise with a small STD (standard deviation) which is proportional to the S-R distance. Its dependency on the alignment angle between sender and receiver was not modeled. The *electronic noise* was modeled as white noise with a small STD. The *turbulent noise* was modeled as a time- and space-correlated random vector process. Atmospheric and electronic *biases* were neglected, thereby assuming that the system is well calibrated. The potential biases due to the finite size of the ultrasound transducers were also neglected. The *Doppler effect* was not modeled because it can be easily compensated. The inverse *sound speed* was set to $n_r = (1/342)$ sec/m. The inverse *sound speed gradient* was set to $n_{\nabla} = 4.4 \times 10^{-6}$ sec/m². No *drift* of n_r or n_{∇} was modeled.

8.1.1 Test Trajectories

In an effort to evaluate the performance of the control system, several test trajectories were used.

The closed loop i/o behavior of the system was tested with a 20 cm step function in the z-direction:

$$y_d(t) = \begin{cases} (400, 149, 0)^T & \text{if } t < 0.4 \text{ sec} \\ (400, 149, 200)^T & \text{if } t \geq 0.4 \text{ sec} \end{cases}$$

This trajectory will be referred to as TS1.

The ability of the loop to reduce the kinematic error is tested by several circular trajectories, all of which lie entirely in the x-y plane. The center of the circle is at c and its radius is r . The trajectory accelerates from zero velocity at an initial angle φ_0 with a constant circular acceleration α until it reaches the angular velocity ω_0 . Then

	c [mm]	r [mm]	f_0 [Hz]	T_a [sec]	φ_0 [deg]
TC1	$(500,149,0)^T$	200	0.02	1	0
TC2	$(500,149,0)^T$	200	0.004	1	0
TC3	$(319,149,0)^T$	200	0.02	1	0
TC4	$(500,149,0)^T$	200	0.1	1	0

Table 8.1: Circular test trajectory parameters.

it continues with a constant ω_0 . Thus, the test trajectory prototype is:

$$y_d(t) = \begin{cases} c + r(\cos(\alpha t^2 + \varphi_0), \sin(\alpha t^2 + \varphi_0), 0)^T & \text{if } 0 \leq t < T_a \\ c + r(\cos(\omega_0 t + \varphi), \sin(\omega_0 t + \varphi), 0)^T & \text{if } t \geq T_a \end{cases}$$

where $\alpha = \omega_0/T_a$, and $\omega_0 = 2\pi f_0$. The trajectory parameters c , r , f_0 , T_a and φ_0 which were used in the simulations are shown in Table 8.1. The trajectory TC1 is so slow that the dynamic error is small compared with the kinematic error. Therefore, TC1 demonstrates the ability of the system to reduce a typical kinematic error. The trajectory TC2 is very slow and tests the low frequency performance of the EKF. On the other hand, the trajectory TC3 passes a nominally singular point at ca. 3 mm distance. Thus it tests the system's ability to operate in ill-conditioned regions of the workspace. Finally, TC4 is very fast and is used to show the interaction between the control system and the EKF.

The simulations used 3 microphones placed at optimum locations: 1 m away from the trajectory center in the x,y, and z directions, respectively.

8.1.2 Model Uncertainties

The loop uncertainties U_1 and U_2 from Fig. 3.5 must be included in the simulation. The multiplicative inner loop uncertainty U_1 was chosen to be linear, having the

Laplace transform:

$$U1(s) = \begin{pmatrix} f & g & g \\ g & f & g \\ g & g & f \end{pmatrix} \frac{a_0}{s^2 + a_1 s + a_0}$$

where a_0 and a_1 were chosen such that the second order system has a resonance peak of 10dB at 20 Hz, in order to account for the first flexible mode of the robot. The matrix in the equation above accounts for an imperfectly known inertia matrix. It was assumed that the diagonal elements are known with a 20% error i.e. $f=1.2$. Any residual dynamic cross coupling is represented by g . For circular test trajectories, 10% cross coupling was assumed i.e. $g=0.1$. Certain system properties are easier to demonstrate if there is only kinematic cross coupling. It was therefore assumed that $g = 0$ for simulations with the TS1 trajectory.

The outer loop uncertainty U_2 was defined previously as:

$$U_2 = \hat{T}_i^{-1} \circ T(\cdot)$$

This uncertainty causes the kinematic error

$$e_{kin} = \hat{T}(q_c) - T(q_c)$$

Suppose $\hat{T}$ represents the nominal forward kinematic function. This yields results in the smallest kinematic error. For this choice of $\hat{T}$, Chen [2] calculated the kinematic error of a PUMA780 by using Cartesian position data obtained from measurements with three theodolites. The mean kinematic error over 80 positions was reported to be 5.9 mm and the maximum error was 10mm. For simulations, a realistic T should therefore be chosen such that the associated kinematic error has a similar size than the one reported in [2].

Table 8.2 shows two sets of values, KP1 and KP2, each representing differences between the nominal and the true kinematic parameters. In particular, p_1 denotes

	p_1	p_2	p_3	p_4	p_5	p_6
<i>KP1</i>	2.0 mm	2.0 mm	2.0 mm	0.3°	0.2°	0.1°
<i>KP2</i>	-2.5 mm	2.0 mm	2.5 mm	0.2°	-0.4°	-0.2°

Table 8.2: Kinematic error parameters.

the distance which was added to the PUMA600's nominal shoulder length, and p_2 and p_3 denote the distances added to the nominal lengths of the upper and lower arm. Furthermore, p_4 , p_5 and p_6 denote offset angles added to the angles which nominally are measured by the shaft encoders. Both *KP1* and *KP2* have the property that the resulting "true" forward kinematic functions have associated kinematic errors which are similar in size to the ones described in Chen [2]. Otherwise they are arbitrary.

8.1.3 Measurement Noise

When the sender is static, then the simulation of the measurement noise is straightforward: It can be realized by applying a sequence of independent white noise vectors to the input of a linear time-invariant MIMO filter.

However, when the sender is moving, then the properties of the TOF fluctuations as perceived by the microphones change in accordance with the trajectory of the sender. A moving sender produces TOF fluctuations which are more time-decorrelated than those of a static sender would be because the space-decorrelation due to a locomotion of the sender is perceived by the receivers as a time-decorrelation. For instance, a sender traveling slowly along a small circle produces TOF fluctuations which are more time-correlated than those associated with a sender which travels quickly along a large circle.

In order to reproduce this observation with a computer simulation while keeping the computational requirements within reasonable limits, the following steps were taken:

- 1) The measurement noise was computed for the desired trajectory instead of the

true trajectory. It was felt that very little, if any accuracy is lost since for closed loop control, the difference between the two trajectories can be expected to be smaller than the precision of the stochastic model used.

2) The desired sender path was sampled yielding a set $S_z = \{S_1, \dots, S_{N_z}\}$ of N_z path points. It is then imagined that a virtual, fixed sender is located at each of the N_z path points. There are also N_r fixed reference senders and N_m microphones. Now, similar to (6.1), a long vector $\nu(\bar{\omega}, t)$ can be formed, representing the TOF fluctuations of all $N_{ch} = (N_z + N_r)N_m$ imagined measurement channels. Each virtual channel produces its own noise, which is the same as that which would be obtained by a static sender at that location. The effect of a moving sender on the measurement noise can now be realized by switching between the N_z virtual static senders while moving along the trajectory. This reduces the original dynamic noise realization task to a static problem.

3) From (6.5), for static channels, the space correlation and the time correlation are independent. Furthermore, from (6.8), there are Markov processes (A,B,C,D) which describe the TOF fluctuations' time correlation for a single static channel. Thus, to realize the time correlation of ν , it is sufficient to simulate N_{ch} decoupled SISO Markov processes (A,B,C,D), one process for each element of ν . The i -th process is given as

$$\begin{aligned} x_i(k+1) &= Ax_i(k) + Bu_i(k) \\ y_i(k) &= Cx_i(k) + Du_i(k) \end{aligned}$$

where the input u_i to the i -th filter is independent white gaussian noise with unit variance. The outputs of the N_{ch} processes form a vector $y = [y_1, \dots, y_{N_{ch}}]^T$.

4) The spatial correlation between channels of ν can be realized by first computing the spatial covariance matrix $P_{\nu\nu} = E\{\nu\nu^T\}$. This can be accomplished by solving (6.10) for each element of $P_{\nu\nu}$. Further, using a Cholesky decomposition, a positive

definite matrix X_{cor} can be found such that $P_{\nu\nu} = X_{cor}X_{cor}^T$. Finally, the time and space correlated process $\nu(k)$ is realized as:

$$\nu(k) = X_{cor}y(k)$$

Now, as desired, the covariance kernel of ν is:

$$E\{\nu(k)\nu(j)^T\} = P_{\nu\nu}C \exp(A|k-j|)P_{xx}C^T$$

where P_{xx} is the solution of the Lyapunov equation for any of the N_{ch} identical Markov processes.

5) Faced with computational limitations, it was decided to first realize the measurement noise on a coarse space-time grid and then linearly interpolate between the coarse grid points to obtain a realization on a finer space-time grid. First, the sender path was parameterized. Here, the parameter was chosen to be the distance d traveled by the movable sender from the beginning of the path. Therefore, the location of a path sample $S_i \in S_z$ can also be characterized by its associated distance d_i .

Suppose the movable sender has traveled a distance d along its path at time kT_{sr} , where $d_i < d < d_{i+1}$, and T_{sr} is the rough sampling period. Let $\nu_j(k)$ and $\nu_{j+1}(k)$ be the realizations at time kT_{sr} of those two communication channels which are associated with S_i and S_{i+1} and the microphone M_l . The coarsely sampled measurement noise realization at time kT_{sr} for receiver l , denoted $\eta_l(k)$, is now obtained as a linear interpolation between $\nu_j(k)$ and $\nu_{j+1}(k)$:

$$\eta_l(k) = \nu_j(k) + (\nu_{j+1}(k) - \nu_j(k)) \frac{d - d_i}{d_{i+1} - d_i}$$

The final noise realization is found by resampling η_l at a higher sampling rate. The new sample points were generated with a lowpass interpolation, using the MATLAB function 'interp'.

6) All simulations for circular test trajectories featured $N_z=12$ equally spaced virtual senders on the circle, and the coarse sampling rate was $T_{sc}=5\text{Hz}$.

	n_1	n_2	n_3	n_4	n_5
NC1	0	0.02 μsec	0.02 μsec	0.29 μsec	600 mm
NC2	2e9	0.02 μsec	0.02 μsec	0.29 μsec	600 mm

Table 8.3: Parameter sets for noise realizations.

Aside from the noise caused by turbulence in the air, there are also two less important noise sources which were realized. The reverberation noise is independent white noise whose standard deviation increases linearly with the sender-receiver distance. The receiver background noise is also independent white noise, however, its variance is not a function of the sender-receiver distance.

The noise parameters used in the simulations are listed in Table 8.3, where n_1 is the seed for the MATLAB random number generator, n_2 is the STD of the receiver background noise, n_3 is the STD of the reverberation noise for 1 m S-R distance, n_4 is the STD of the turbulence noise at 1 m S-R distance, and n_5 is the spatial correlation constant. Thus, NC1 and NC2 represent different realizations of the same random process.

8.2 Noise Free Simulations

8.2.1 Step Response

In order to demonstrate certain features of the control loop, the measurement noise was neglected. Consider the loop shown in Fig. 3.6. First, it was subjected to a 20 cm input step in z-direction (test trajectory TS1). The step response of the system is shown in the upper left hand plot in Fig. 8.1. The curve labeled 'B' shows the z-component of y_c for the nominal system. As expected, it displays a no-overshoot critically damped behavior with a one sample delay. The curve denoted by 'C' shows the z-component of y_c when both the dynamic and kinematic uncertainties are included. There is a slight distortion of the nominal step response, in particular, there

is a very small overshoot. These distortions are mainly due to the neglected acoustic transmission delay and the other dynamic uncertainties. The inclusion or omission of the kinematic uncertainties has no visible effect on the step response.

The x-component of y_c for the same experiment is shown in the upper right hand plot of Fig. 8.1. Again, there is a slight distortion between the nominal response, 'A', and the response with the uncertainties included, 'B'. Again, the inclusion or omission of the kinematic uncertainties has no visible effect on 'B'. Notice that the nominal x-component of y_c is not a constant 400 mm as commanded by y_d . Rather, it deviates from this value by a maximum of about 12 cm. Consequently, the nominal response trajectory of the robot to a step is not a straight line, but rather an arc. The reason for this behavior is dynamic nonlinear cross coupling. This is evident from Fig. 3.5 by writing the nominal tracking error as :

$$e = y_d - \hat{T} \circ h_m \circ \hat{T}^{-1}(y_d)$$

where h_m represents the nominal closed loop dynamics expressed in joint coordinates. Clearly, if h_m is different from identity, then the nonlinear dynamic cross coupling will increase the control error. The effect is particularly strong when the bandwidth of y_d exceeds the bandwidth of h_m . One can avoid this form of cross coupling by making h_m as close to identity as possible.

The y-component of y_c is shown in the lower right hand plot of Fig. 8.1. Here, curve 'A' represents the step response for both the nominal system and the system with added dynamic uncertainties. The curve 'B' on the other hand, shows the response when kinematic uncertainty is also added. Here, there is no nonlinear cross coupling except for the kinematic uncertainty. The reason for this lies in the special geometry of the PUMA600 and the particular test trajectory chosen. For TS1, the first joint is the only one that can change the end-effector's y-coordinate, thereby decoupling it from the other two joints. The cross coupling is therefore caused by the kinematic

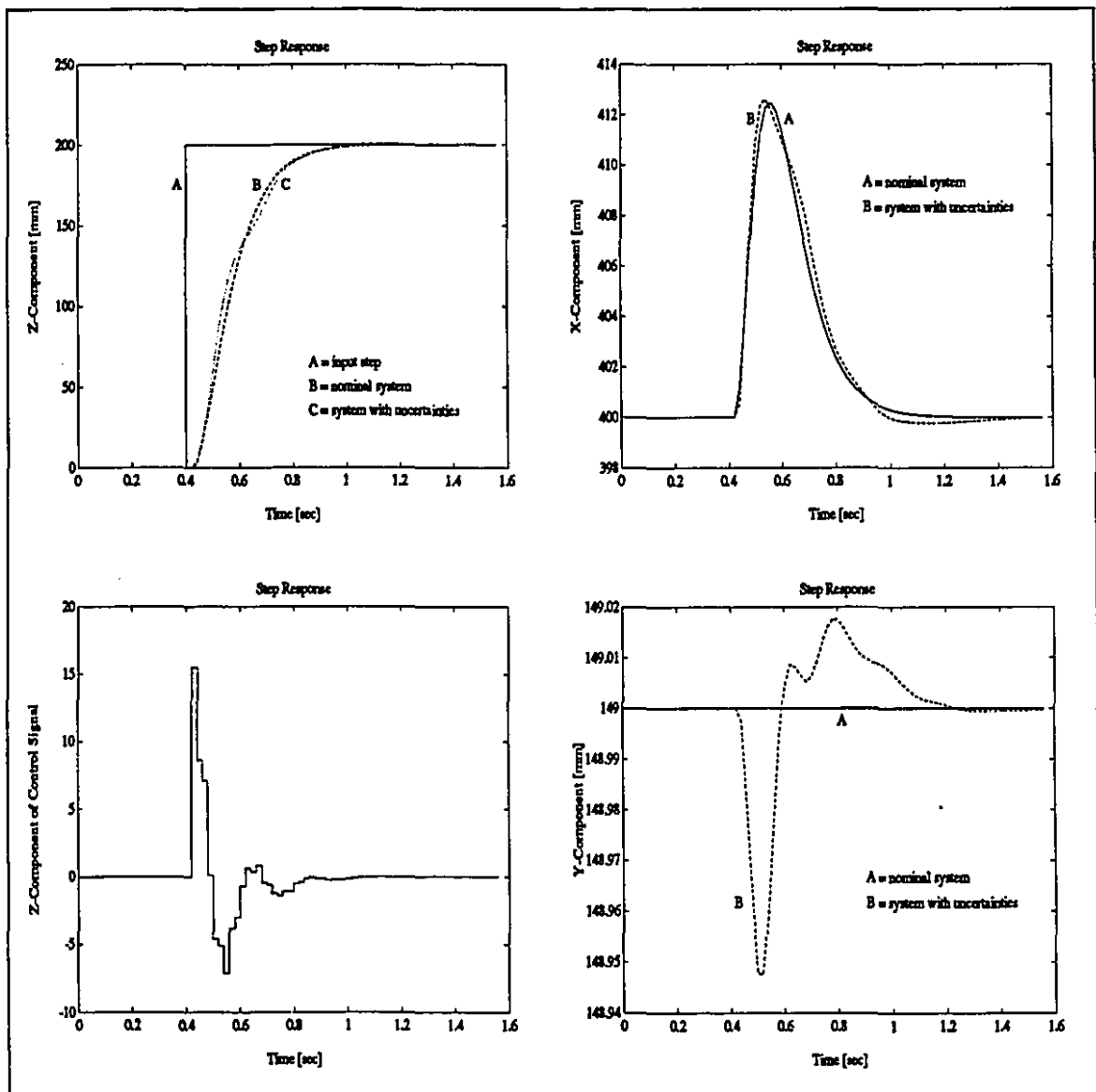


Figure 8.1: Step responses for system with uncompensated inner dynamics.

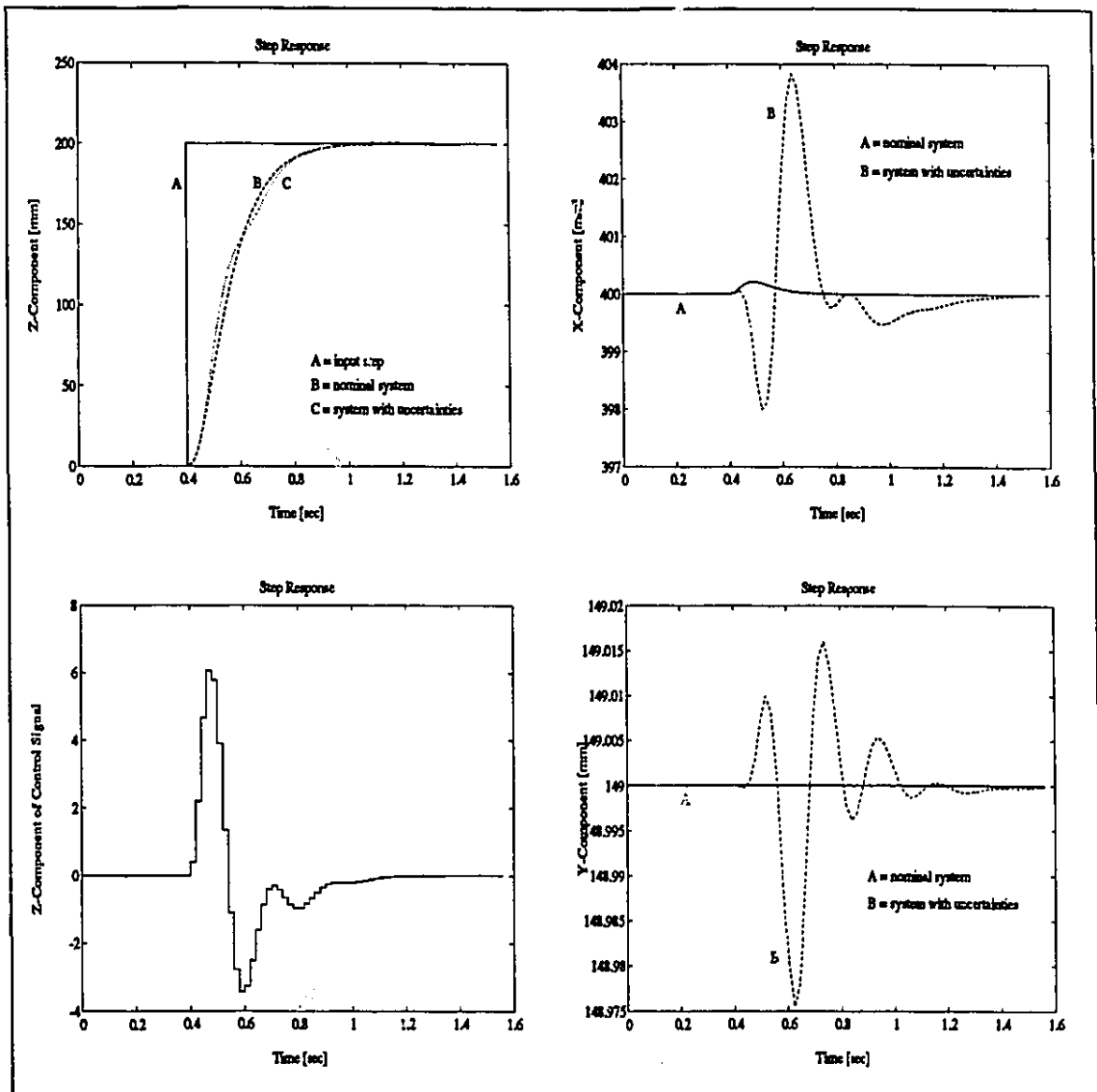


Figure 8.2: Step responses for system with compensated inner dynamics.

uncertainty, only.

Finally, it may be of interest to view the shape of the control signal u . The z -component of u is shown in the lower left plot of Fig. 8.1 for the case where all uncertainties are included.

To improve the performance, the nominal closed loop dynamics from q_d to q_c were approximately canceled and y_d was prefiltered. This method to reduce the nonlinear dynamic cross coupling was described in detail in Sec. 3.2.3. The performance of the improved system is shown in Fig. 8.2 which should be compared to the results shown in Fig. 8.1: There is little change in the z -component of the step response. The x -component, however, shows a large reduction of its nominal error due to the reduced nonlinear dynamic cross coupling. The response with added system uncertainties is also significantly reduced compared to Fig. 8.1. A further reduction could be achieved if the acoustic transmission delay was accounted for in the regulator design. The size of the y -component remains approximately unchanged because it is not subject to nonlinear dynamic cross coupling. Note that the control signal does not show excessive oscillatory behavior as would have been expected, had the nominal closed loop dynamics been canceled completely.

8.2.2 Circle Response

In order to test the system's tracking ability, it was commanded to follow circular test trajectories. One problem of circle tracking is lag, particularly if a no-overshoot requirement is imposed. The lag has the effect that only extremely slow trajectories can be tracked with submillimeter error.

To partially overcome this restriction, the closed loop dynamics were compensated as discussed in Sec. 3.2.3. However, other than the case discussed in Sec. 3.2.3, no pre-filter for y_d was used, and consequently, the system's nominal transfer characteristic from y_d to y_c was approximately unity. (The acoustic delay was not simulated.) Care

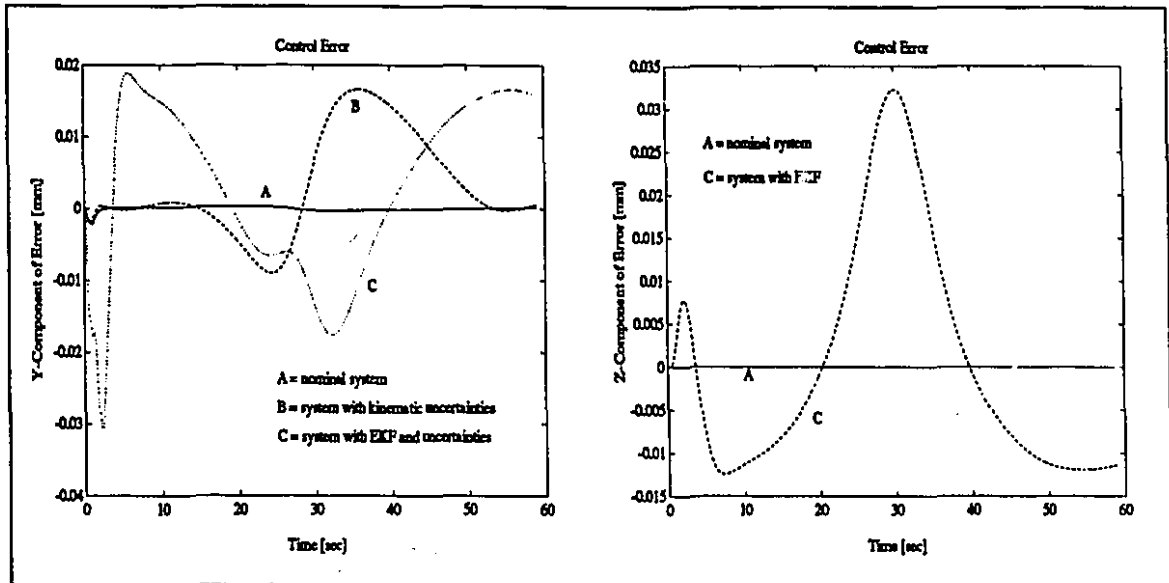


Figure 8.3: Control performance for TC1-KP1.

must be taken to ensure that the commanded trajectory has no high frequency spectral components which may excite flexible modes or cause the actuators to operate outside of their linear range.

The system's response to the test trajectory TC1 (0.02Hz) is shown in Fig. 8.3. The y-component of the control error $e = y_d - y_c$ is depicted on the left. Curve 'A' represents the nominal system response. It can be seen that the dynamic compensation works well because 'A' exhibits only a very small tracking error during the angular acceleration phase of TC1. The system's response when the kinematic uncertainty KP1 is added is depicted by the curve 'B'. The maximum tracking error is approximately 0.02 mm. This is about what one would expect: The kinematic error for TC1-KP1 is shown in Fig. 8.7. It has sinusoidal shape with an amplitude of 1.5 mm and a frequency of 0.02 Hz. The crossover frequency of the type-1 outer loop control system is about 2 Hz. Hence, one would expect ca. 40 dB error reduction at 0.02 Hz. Thus, the error would be expected to be about 0.015 mm.

Curve 'C' represents the system's response to TC1-KP1 when the EKF and the

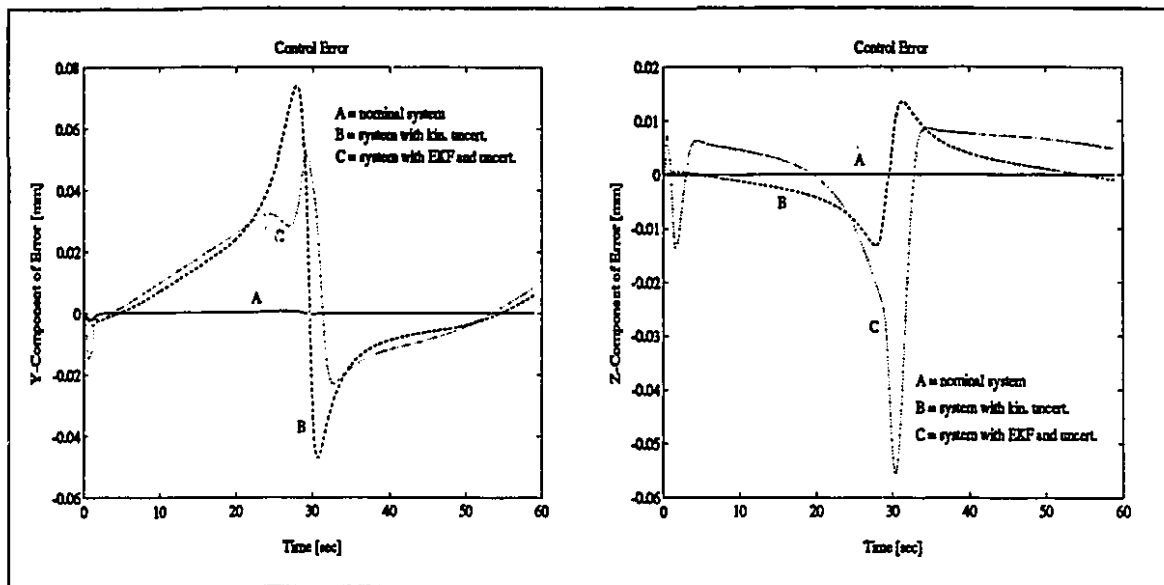


Figure 8.4: Control performance for TC3-KP2.

dynamic uncertainties are added to the loop. The control error for the system with EKF but without added dynamic uncertainties is almost identical to 'C' and is therefore not shown in Fig. 8.3. The EKF parameters were chosen in the same manner as if there was NC1 measurement noise. The computation of the EKF parameters from knowledge of TC1, KP1 and NC1 is explained in more detail later. The EKF-induced error is not significantly larger than the error due to pure kinematic uncertainty. The exception is the acceleration phase of TC1, where the EKF has a more pronounced impact on the control error.

A further aspect of the EKF is best illustrated with the right hand plot in Fig. 8.3. The curve 'A' represents both the z-component of the nominal control error and the control error when the kinematic uncertainties KP1 are added. Both curves are indistinguishable for the scaling given. The reason for this similarity is that the z-component of the open loop kinematic error for TC1-KP1 is very small. Thus, it is surprising to find that the z-component of the control error is much larger when the EKF is added. This is shown by curve 'C'. The reason for this, of course, is that the

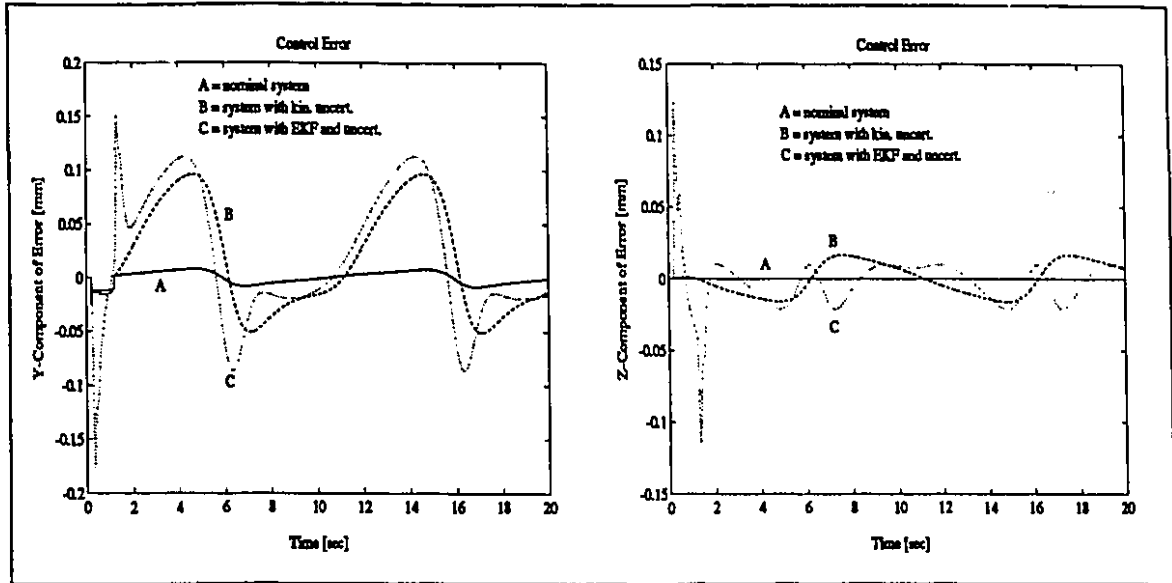


Figure 8.5: Control performance for TC4-KP2.

EKF dynamics introduce nonlinear dynamic cross coupling similar to that discussed previously. This can be seen by writing the “estimation” error in the joint space:

$$e = q_c - \hat{T}^{-1} \circ h_f \circ T(q_c)$$

where h_f denotes the EKF dynamics. Clearly, the estimation error e will exhibit significant nonlinear dynamic cross coupling if the bandwidth of h_f is smaller than the bandwidth of q_c . The estimation error e , in turn, will affect the control error in the manner illustrated by curve 'C'.

The x-component of the control error was not shown in Fig. 8.3 because no new insight about the control system can be gained from it. The same argument applies to the x-components of the control error for Fig. 8.4 and Fig. 8.5.

Next, the control system was subjected to the trajectory TC3 and the kinematic uncertainty KP2 to demonstrate the control system's ability to operate well even in the vicinity of singularities. The results of the simulation are shown in Fig. 8.4 where they are presented in the same manner as in Fig. 8.3. The singularity is passed

by the sender after traveling for 30 sec. It can be seen in Fig. 8.4 that passing by the singularity introduces peaks and near-discontinuities into the error response. The maximum error is also increased. In this example it is about 0.07 mm for the y-component of the error. Curve 'C' in the left hand plot of Fig. 8.4 shows the y-component of the error when the EKF is added to the system. In this particular case, this error is smaller than the error without EKF. However, the z-component of the control error for the system with EKF is much larger than for the system without EKF, particularly near the singularity, as shown in the right hand picture of Fig. 8.4. Again, the peak is largely due to nonlinear dynamic cross coupling.

Finally, the system was subjected to the test trajectory TC4 (0.1 Hz). The frequency of this trajectory is closer to the crossover frequency (2 Hz) of the outer loop than any other test trajectory. Since the loop sensitivity is larger in the vicinity of the crossover frequency than at lower frequencies, one would expect the control error to increase. This is indeed the case as can be seen in Fig. 8.5. Note that even the nominal error response shown in the left hand plot of Fig. 8.5 shows a small deviation from zero. The control error for the system with EKF included shows a particularly large error (-0.17 mm) during the initial acceleration phase.

The trajectory TC2 was not used to test the noise-free system performance because it is just a five times slower version of TC1. The only insight which can be gained from it is that the control error for TC2 is even smaller than for TC1. However, TC2 will be used in the next section.

8.3 The EKF Performance

The EKF is the key element of the loop because its ability to reduce the measurement noise almost entirely determines the achievable tracking precision. To establish its performance, two methods were used: Nonlinear simulations and covariance analysis

of a linearized system. The two analyses support each other.

8.3.1 Nonlinear EKF Simulation

It is desirable to test the performance of the EKF independently of the control loop in which it usually operates. To this end, the EKF was tested outside the loop as shown in Fig. 7.2, where $q_c = T_i^{-1}(y_d)$ is determined by the desired trajectory rather than the true trajectory y_c . Moreover, a reference sender was placed in the center of the circular test trajectories.

The main parameters of the EKF are the measurement noise strength and the state noise strength.

1) The measurement noise strength for a S-R distance of 1 m was set to the values associated with the test noise NC1 or NC2, respectively. No corrections were made for the fact that the measurement noise is not white. While the sender progressed along the trajectory, the noise strength values were continuously adapted to reflect the changing strength in each measurement channel. When the reference sender was used, the noise strength values were appropriately reduced.

2) The state noise strength was set to the RMS value of the acceleration of the kinematic error. This value was chosen because it is easy to compute and appears to be an almost optimal choice for the type of trajectories considered. Only minor performance improvements can be achieved by choosing other (larger) values.

The measurement noise seen by the first receiver for the test trajectory TC1 and the noise realization NC1 is shown in Fig. 8.6. In accordance with NC1, the nominal measurement noise STD in mm for a static sender receiver pair with 1 m distance is 0.1 mm. The uncompensated noise for the moving sender displayed in the left hand plot of Fig. 8.6 has a STD of about 0.1 mm, too. If the reference sender's optimal estimate of this noise is used for compensation, then the situation is improved. The right hand plot of Fig. 8.6 displays the first receiver's measurement

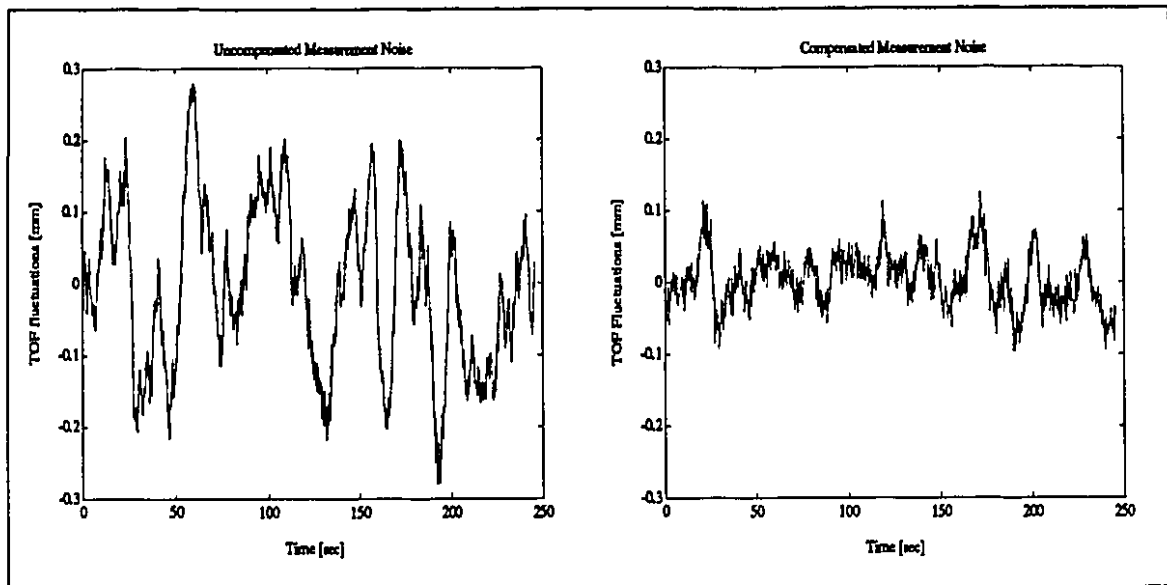


Figure 8.6: Compensated and uncompensated measurement noise for receiver # 1

noise after compensation. Clearly, the compensated noise has a STD that is smaller by a factor 2-3 than the uncompensated noise. Moreover, it resembles white noise more closely than the uncompensated noise does, thereby improving the validity of the assumption that all inputs to the Kalman filter are white noise.

The left hand plot of Fig. 8.7 shows both the y-component of the KP1 kinematic error and the y-component of the EKF's estimate of the kinematic error using the TC1 (0.02Hz) test trajectory, KP1 kinematic uncertainty, and the NC1 noise parameters. It can be seen that the kinematic error for this case closely resembles a sinusoidal signal. Its frequency is the same as the frequency of the test signal and its amplitude is ca. 1.5 mm, varying between 3 mm and 6 mm. The EKF tracks this signal in the presence of uncompensated measurement noise with minor errors. The curve 'B' shows the EKF estimates, while curve 'A' shows the kinematic error.

The y-component of the estimation error can be seen more clearly as the curve 'A' in the right-hand plot of Fig. 8.7. The error is as large as 0.3 mm and there is not much reduction over and above the size of the measurement noise. On the

other hand, when the reference sender is employed to reduce the measurement error, the estimation error is about 2-3 times smaller and does not exceed 0.1 mm. The y-component of this error is depicted by curve 'B'. The standard deviations and the maximum absolute values of the estimation errors for all three dimensions of space are displayed in Table 8.4 for both the compensated and the uncompensated cases. Moreover, Table 8.4 displays the respective values for another measurement realization (NC2). The two noise realizations show about the same result.

In order to see how the estimation error improves when the trajectory is even slower than TC1, the above simulation was repeated for the test trajectory TC2 which is five times slower (9.004Hz). The kinematic error parameters and the measurement noise parameters remained unchanged (KP1-NC1). The results are displayed in Fig. 8.8 which is presented in an analogous manner to Fig. 8.7. As before, the maximum values and the standard deviations for all directions of space as well as for the NC2 noise realization can be found in Table 8.4. From this evidence it is clear that only a very modest improvement, if any, was achieved by using a trajectory which was five times slower. However, the estimation error displayed in the right-hand plot of Fig. 8.8 is much smoother than the error depicted in Fig. 8.7. This was expected because the Kalman filter has a reduced bandwidth for the slower trajectory.

A more demanding kinematic error than TC2-KP1 is TC3-KP2. It is shown in the left-hand plot of Fig. 8.9. The spikes and near-discontinuities exhibited by the kinematic error are consequences of the closeness of the trajectory TC3 to a singular point. Moreover, a different kinematic parameter set (KP2) was used than that in the last two simulations. The main difference with respect to the estimation error achieved is an increase of the maximum error, at least for NC1. This is shown in the right-hand plot of Fig. 8.9, where large errors coincide with the spikes in the kinematic error. However, the result could almost certainly be improved by opening the Kalman filter wider. The simulation results for NC2 and the other dimensions of

TC1-KP1		NC1			NC2		
		x	y	z	x	y	z
max	reg	0.31	0.30	0.38	0.31	0.26	0.31
	cmp	0.10	0.11	0.13	0.11	0.12	0.14
STD	reg	0.10	0.12	0.13	0.11	0.091	0.13
	cmp	0.034	0.038	0.052	0.038	0.043	0.052
TC2-KP1		NC1			NC2		
		x	y	z	x	y	z
max	reg	0.24	0.31	0.24	0.24	0.16	0.24
	cmp	0.14	0.088	0.11	0.14	0.10	0.14
STD	reg	0.11	0.11	0.10	0.090	0.078	0.092
	cmp	0.047	0.033	0.044	0.045	0.039	0.050
TC3-KP2		NC1			NC2		
		x	y	z	x	y	z
max	reg	0.40	0.38	0.36	0.23	0.29	0.31
	cmp	0.12	0.15	0.13	0.11	0.13	0.11
STD	reg	0.10	0.13	0.13	0.089	0.094	0.12
	cmp	0.036	0.040	0.046	0.032	0.040	0.042
TC4-KP2		NC1			NC2		
		x	y	z	x	y	z
max	reg	0.30	0.33	0.34	-	-	-
	cmp	0.19	0.12	0.14	-	-	-
STD	reg	0.10	0.12	0.12	-	-	-
	cmp	0.04	0.031	0.048	-	-	-

Table 8.4: Results of the nonlinear EKF simulation.

space can be found in Table 8.4.

Finally, a very fast kinematic error TC4-KP2 (0.1 Hz) was tested. The simulation results can be found in Fig. 8.10 and Table 8.4. It is clear from this evidence, that even for a trajectory 25 faster trajectory than TC2, there is no serious deterioration of the estimation error.

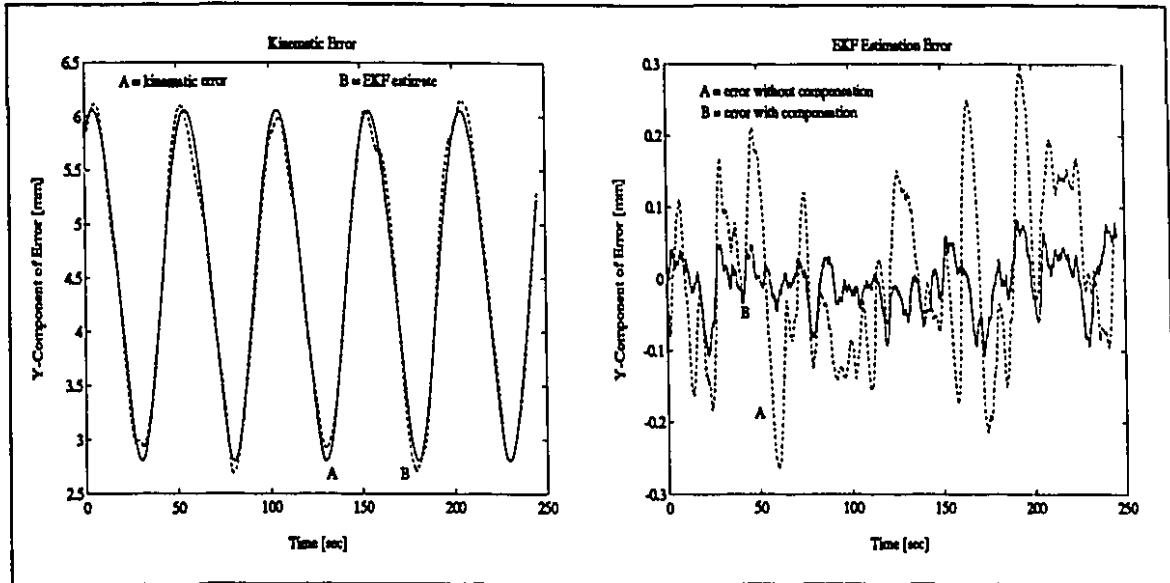


Figure 8.7: EKF performance for TC1-KP1-NC1.

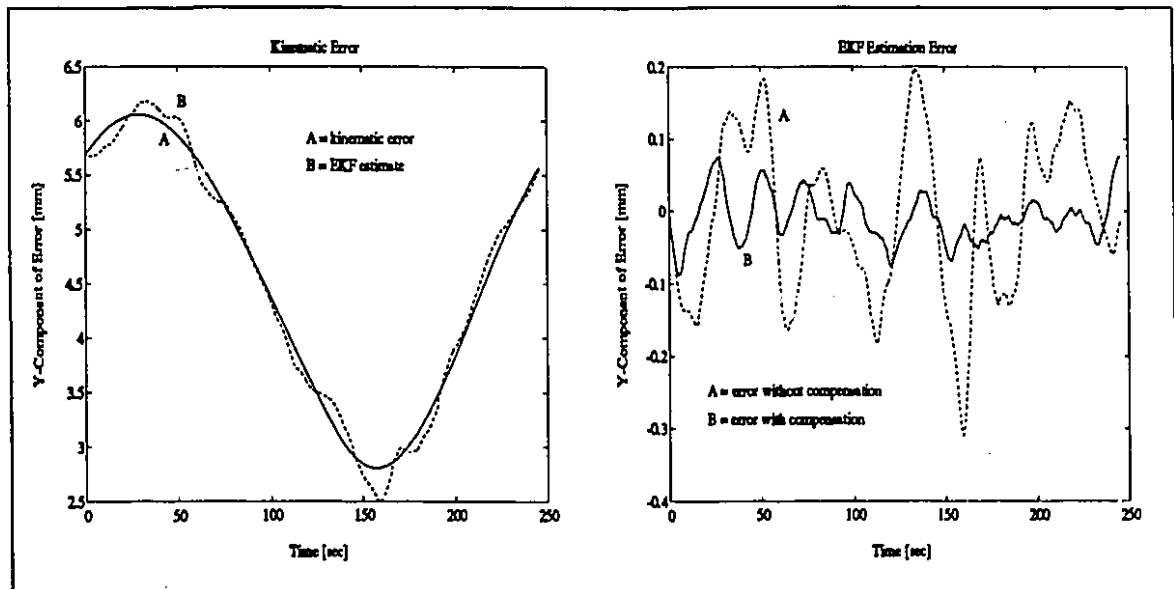


Figure 8.8: EKF performance for TC2-KP1-NC1.

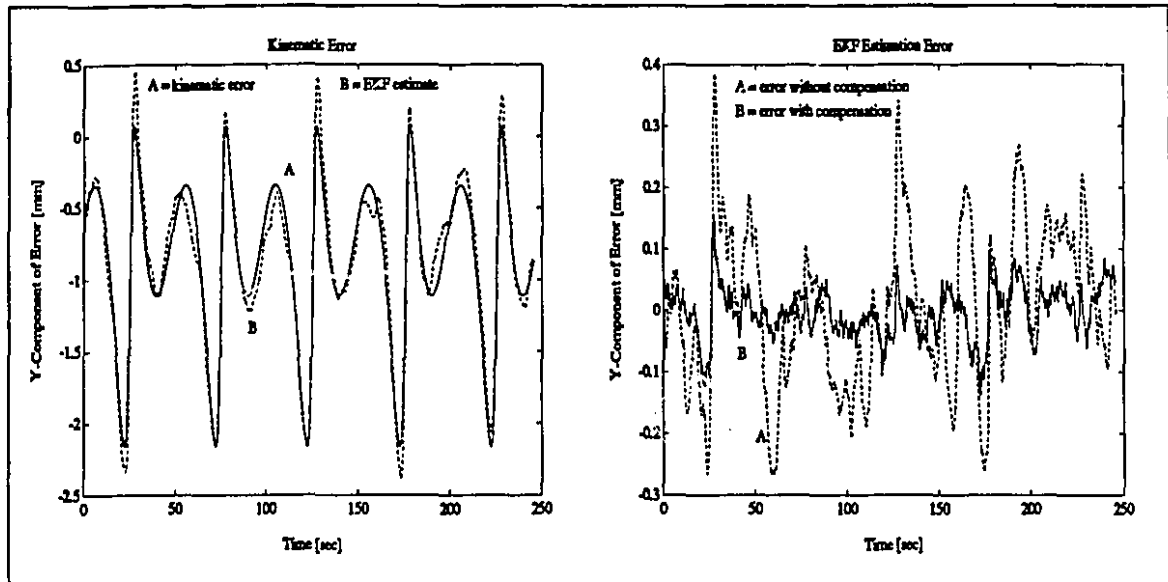


Figure 8.9: EKF performance for TC3-KP2-NC1.

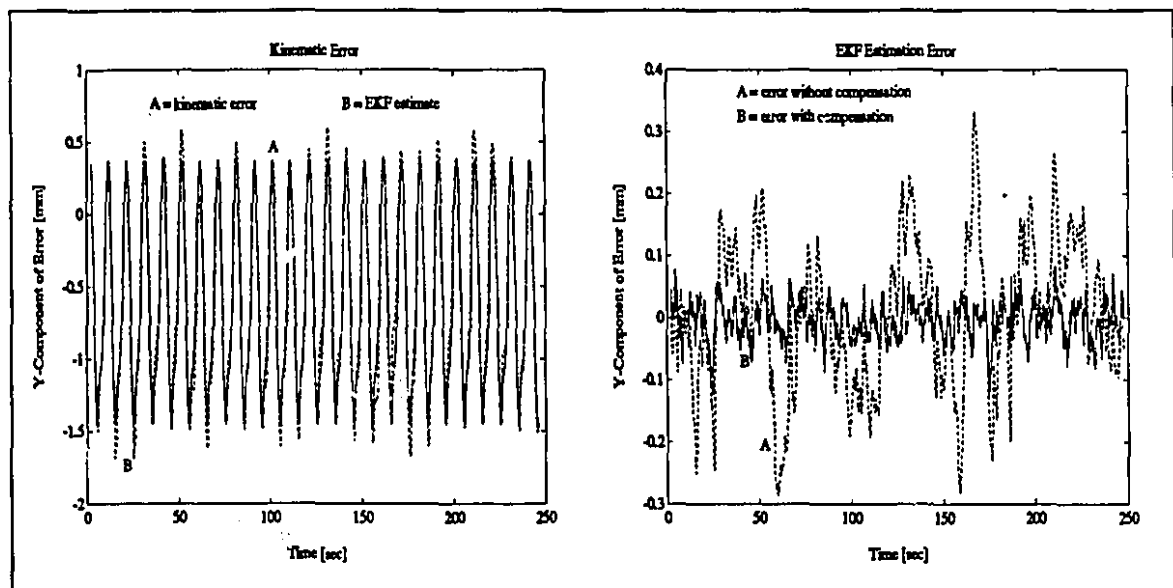


Figure 8.10: EKF performance for TC4-KP2-NC1.

8.3.2 Covariance Analysis

The EKF simulation results can be supplemented by an approximate covariance analysis which provides additional insight into the EKF's performance and can support the validity of the previously obtained results. A covariance analysis becomes possible when the system's nonlinearities are ignored. Usually, the Kalman filter's internal model is a replica of the entire stochastic process. In this case a covariance analysis requires the solution of a Riccati equation. Here, however, the Kalman filter's internal model consists of just three decoupled double integrators. A covariance analysis for this system comprises three steps:

- 1) Choose the EKF gain K_f as previously, except that now the sender is static.
- 2) Augment all process dynamics to the closed loop Kalman filter dynamics.
- 3) Solve the Lyapunov equation corresponding to the augmented process.

The augmented dynamics consisted of the dynamics of the measurement noise and the dynamics of the kinematic error acceleration.

1) The dynamics of the measurement noise were the previously mentioned, experimentally obtained 3rd order noise dynamics for a static sender. In other words, the influences on the noise which stem from the movement of the sender were ignored. Further, it was assumed that the sender is located at the center of the circular test trajectories. This ensures the measurement noise to be of some average strength over the circular path.

2) The dynamics of the kinematic error acceleration are modeled as a white noise driven 2nd order system with a 30dB resonance peak at f_s , where f_s is the frequency of the trajectory. The output standard deviation of the 2nd order system was chosen to be the RMS value for the acceleration of a sinusoidal signal with frequency f_s and amplitude 2 mm. This type of kinematic error model works well for the circular trajectories and the kinematic error parameters chosen. This is confirmed by the previous nonlinear EKF simulations. For instance, the kinematic error in Fig. 8.7 is

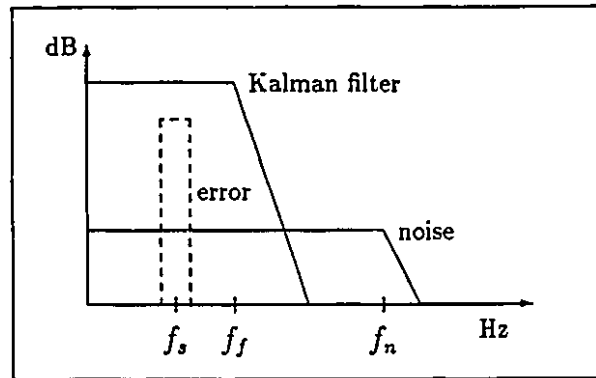


Figure 8.11: A Bode sketch of the noise, the kinematic error and the Kalman filter. (Not to scale.)

approximately sinusoidal with frequency f_s and has an amplitude of ca. 1.5 mm.

The frequency domain relationship between the measurement noise, the kinematic error and the closed loop Kalman filter is sketched in Fig. 8.11. The task of the Kalman filter is to let the kinematic error pass but reject the measurement noise. To do this requires that the cut-off frequency f_f of the closed loop Kalman filter be larger than the center frequency f_s of the kinematic error. On the other hand, if there is any noise reduction to be achieved then f_f must be smaller than the cut-off frequency f_n of the measurement noise.

Unfortunately, the bandwidth of measurement noise is very small ($f_n \approx 0.03$ Hz). Hence for a test trajectory like TC1 (0.02 Hz), there is very little that the KF can do to separate the kinematic error from the measurement noise. The covariance analysis yields a standard deviation of 0.11 mm in any direction of space. This is well in line with the simulation result shown in Table 8.4. Suppose the KF is told that the kinematic error acceleration is just a tenth of the true acceleration for TC1. In other words f_f is artificially forced to be smaller than f_s . The resulting STD in any direction of space is 0.29 mm, which is substantially larger than before. On the other hand, if the filter is opened wider by telling the KF that the kinematic error acceleration is ten times larger than the true value, then the result even improves slightly: The STD

is 0.10 mm, which is the same as the STD of the measurement noise and the best one can do for the given KF structure. Thus, in general, one can recommend that it is better to overestimate f_f than to underestimate it.

For trajectories for which $f_s > f_n$, the filter is so wide open that it lets all measurement noise pass. The test trajectory TC4 (0.1 Hz) is such a case. The corresponding covariance analysis STD is 0.11 mm. Only if the trajectory is substantially slower than f_n can the Kalman filter reduce the noise. However, even for the much slower trajectory TC2 (0.004 Hz), there was not much reduction. The STD was 0.095 mm. The result could possibly be improved by using a steeper roll-off for the Kalman filter, i.e. triple integrators instead of double integrators.

A further question which can be approximately answered by a covariance analysis is the influence of the measurement noise strength on the estimation error. Given the previous analysis it is not surprising to find that the estimation error is approximately proportional to the measurement error. For instance, if the measurement noise is increased from 0.1 mm to 0.2 mm, then the STD for TC1 increases from 0.11 mm to 0.23 mm.

Finally, it is of interest to know how much could be gained if one used a full state optimum Kalman filter instead of just a double integrator. In theory this question could satisfactorily be answered by solving a Riccati equation for a system model which is composed of the measurement noise dynamics and the dynamic model of the kinematic error. The problem is that the minimum estimation error covariance obtained by such an analysis is highly sensitive to the error model used. In particular, the size of the resonance peak of the kinematic error model greatly influences the result. The previous analysis for the suboptimum KF, however, was not sensitive to the exact shape of the dynamic error model. This is a further argument for the use of a suboptimum estimator because, in reality, the exact kinematic error model is unknown.

Nevertheless, if the peak of the 2nd order kinematic error model is fixed at 30 dB, then for TC1 (0.02 Hz), an estimation error STD of 0.067 mm is obtained as compared to 0.11 mm for the suboptimum estimator. This is an improvement of about 40% . If the resonance peak is increased the improvement diminishes. A more substantial saving was achieved for the slower trajectory TC2 (0.004 Hz): The optimum STD was 0.022 mm compared to 0.095 mm for the suboptimum.

Concluding, it can be remarked that the covariance analysis confirmed the results of the nonlinear simulation. The main result of both types of analyses was that the estimation error is about the same size as the measurement noise. A reduction beyond that requires either very slow trajectories or the compensation of the measurement noise with data obtained by reference senders. A third method is the use of sensor fusion as explained in the next section.

8.3.3 Sensor Fusion

An important aspect of robot sensor technology is sensor fusion. In the context of the 3D ultrasound position sensor it is of particular interest to study how the position information obtained by an optical system can be fused with the data obtained by the ultrasound sensor. Often, vision data are processed slowly and are computationally expensive. However, optical methods have the potential of being quite precise. Ultrasound data, on the other hand, can be processed quickly and are computationally inexpensive. However, they tend to be quite noisy. It is therefore useful to investigate how the two sensor technologies can be merged to obtain position data which are fast, computationally inexpensive, and precise.

The obvious answer is to supply the zero-variance position measurements of the vision system at a low rate and use the ultrasound measurements in the intervening time. When a new vision position measurement becomes available then the EKF state is set to the new value and the corresponding elements of the EKF covariance

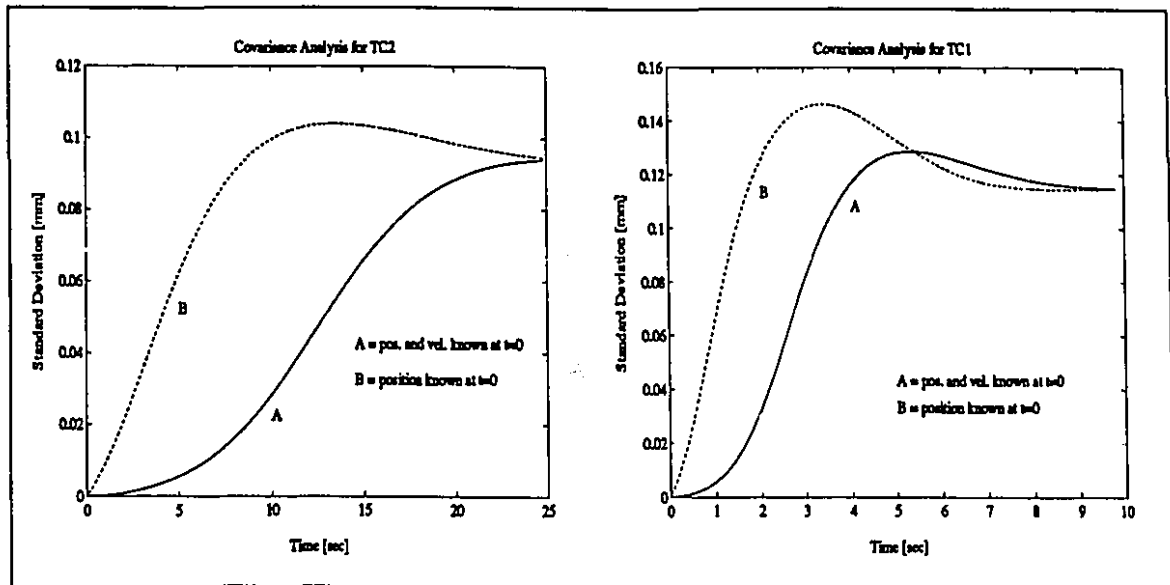


Figure 8.12: Covariance analysis of any one component of the estimation error for the trajectories TC1 and TC2. It was assumed that perfect measurements are available at the time $t=0$.

matrix are set to zero. For a while the EKF will then run with a reduced estimation error until the error again reaches its normal level.

In practice, there is the additional difficulty that, being slow, the vision position measurements may be delayed by several sampling periods. Thus, to optimally incorporate this information into the EKF algorithm requires a modification which allows for data smoothing to occur. The EKF performance would, of course, be diminished compared to the undelayed case.

Assuming the undelayed case, a covariance analysis was conducted for this problem. The results for the test trajectories TC1 and TC2 are shown in Fig. 8.12. The figure was computed by first computing the steady-state values for both the KF covariance matrix and the covariance matrix of the whole process given the steady state KF gain. Those elements of the two covariance matrices which involve position states were then set to zero. These new covariance matrices were then used as initial covariance matrices in the covariance analysis. At each time step the KF gain was updated,

changing the system matrix of the stochastic process. The two curves labeled 'B' show the resulting time histories of the standard deviations of the estimation errors for any of the three directions of space for TC1 and TC2, respectively. The corresponding 'A' curves were obtained analogously to the previous case, but it was assumed that not only the position but also the velocity of the kinematic error are known perfectly.

As expected, it can be seen from Fig. 8.12 that for a given trajectory both the 'A' and the 'B' curve converge to the steady state value for the test trajectory. The 'A' curves, however, remain at smaller STDs for a longer time than the 'B' curves. This is not surprising since the processes generating the 'A' curves were given more information.

If one was to demand that the estimation error for any one dimension of space be less than 0.1 mm with a 99.7% probability, then one would like the STD of the estimation error to be less than 0.033 mm. For TC2, the time required to first exceed this margin after perfect measurement of position and velocity is about 10 sec. If only the position is known perfectly then the time required is about 2.5 sec. On the other hand, for TC1, which is five times faster than TC2, the times are 2 sec and 0.5 sec, respectively. Thus, the time to reach the margin is approximately inversely proportional to the angular velocity of the trajectory (or proportional to the root of the acceleration of the kinematic error). The margin time does not depend strongly on the STD of the measurement noise. However, as mentioned previously, the steady-state value of the estimation error's STD is approximately proportional to the measurement noise.

8.4 Closed Loop with Noise

Finally, the closed taskspace loop was simulated with the measurement noise. The results for the TC1 test trajectory, the KP1 kinematic error parameters and the NC1

noise realization are shown in Fig. 8.13. The left-hand plot in Fig. 8.13 shows the y-component of the tracking error when the measurement noise is not compensated by the reference sender. The right-hand plot of Fig. 8.13 shows the compensated case.

It is not surprising that the tracking error displayed in Fig. 8.13 is practically identical to the EKF estimation error shown in Fig. 8.7. The loop gain is large at the frequency of the kinematic error TC1-KP1. Thus, the tracking error without measurement noise is very small as was shown in Fig. 8.3. Consequently, the control error is identical to measurement noise. For slower trajectories like TC2, this is even truer. Thus, the preceding analysis of the the EKF estimation error applies to the tracking error as well.

On the other hand, one would expect different results for kinematic errors whose spectra are close to the crossover frequency of the outer loop. To some extent, this is the case for the kinematic error TC4-KP2. The simulation results for this case are shown in Fig. 8.14 in an analogous manner to Fig. 8.13. The results are still very similar to the EKF estimation error which is displayed in Fig. 8.10. However, during the initial acceleration phase of TC4, the control error is quite large. This was documented in Fig. 8.5. This little peak can also be seen in the right-hand plot of Fig. 8.14. Moreover, the tracking error for the system with compensated noise exhibits a 0.1 Hz component, which is the frequency of the noise-free control error shown in Fig. 8.14 for TC4-KP2. It can be expected that eventually, for faster trajectories, the (noise-free) control error will come to dominate over the measurement noise.

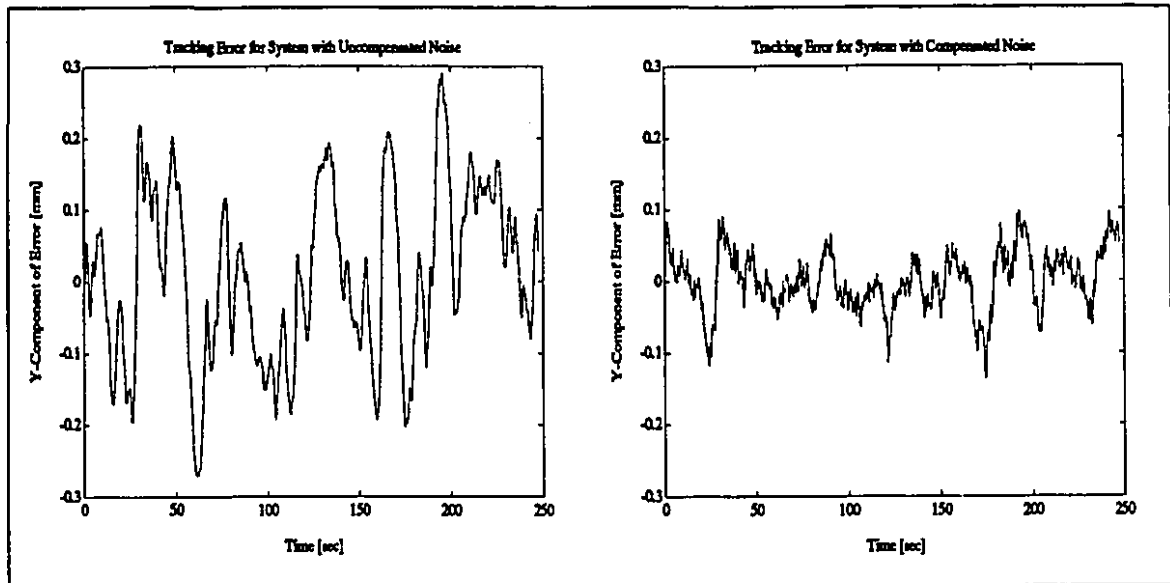


Figure 8.13: Tracking error for the complete system with compensated and uncompensated measurement noise.(Test TC1-KP1-NC1.)

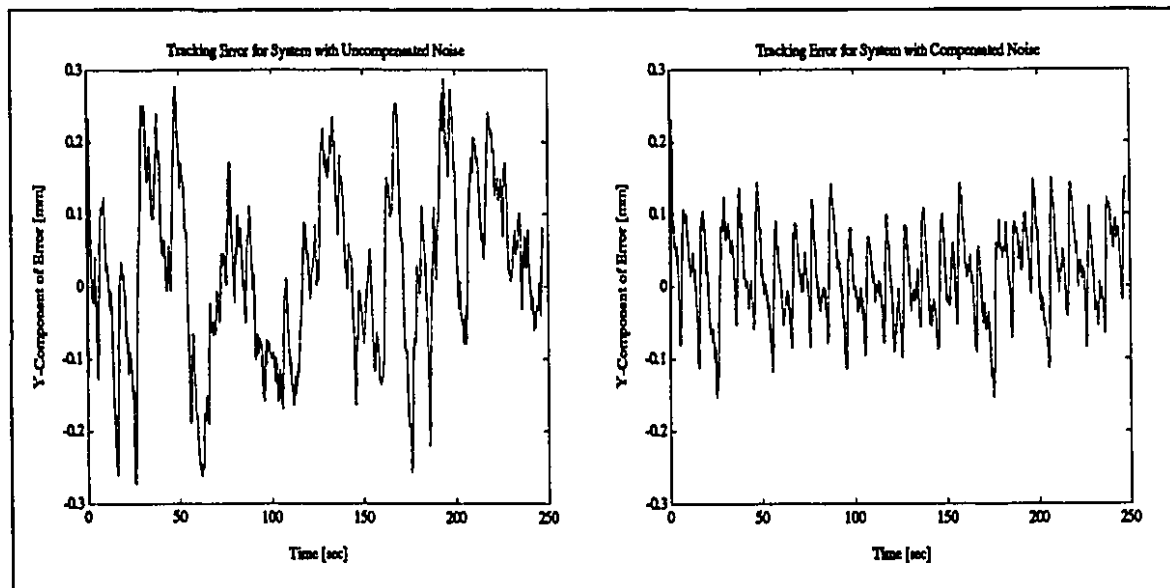


Figure 8.14: Tracking error for the complete system with compensated and uncompensated measurement noise.(Test TC4-KP2-NC1.)

Chapter 9

Conclusions

9.1 Thesis Summary

The precise end effector tracking of taskspace trajectories is an important problem in the control of manipulators. However, ignorance of the true kinematic and dynamic properties of the robot often precludes high-precision tracking. This thesis proposes a solution to overcome this restriction by measuring the end-effector pose in both the joint space and the task space, and using the two measurements in a feedback scheme to reduce the tracking error. It was shown that this double loop design is superior to a single loop design because it can reduce the uncertainties more effectively.

Another important feature of the control system is a new nonlinear kinematic decoupling scheme. It is superior to the traditional Jacobian-based methods because it does not involve a local linearization of the robot kinematics. Rather, it compensates for the nonlinearities globally. This has several advantages which include a reduced loop uncertainty and the ability to handle large disturbances or step trajectories. It also offers a clean proof of robust global stability and robust global performance.

The second issue studied in this thesis was the problem of actually measuring the end-effector position in taskspace coordinates. Many methods have been proposed by

other researchers. This study explored the properties of using ultrasound for 3D end effector position measurements.

One aspect of this task was the design of the ultrasound sender-receiver circuitry. The key problem was to determine the exact time-of-flight of the ultrasound signal in the presence of background noise and reflected signals. The problem was solved by a combination of thresholding and zero crossing detection. Using this method, an inexpensive prototype was built and used as a data acquisition tool. The prototype consisted of four analog receivers, two senders and a counter/signal generator card that plugs into any IBM-AT. The time-of-flight data acquired by the prototype were used to verify the proposed measurement noise model experimentally.

Another key hardware problem was the choice of the ultrasound transducers. Those transducers which are large compared to the wave length cause a mismatch between the acoustical center and the geometrical center. Depending on the effective transducer size, it may or may not be necessary to compensate for this mismatch.

To compensate for biases and drifts a deterministic error model was developed. Many aspects of sound transmission in air were studied. However, the main factors are the room temperature, the room temperature gradient and drafts. The former two can relatively easily be compensated for by a reference sender. Strong air movements, on the other hand, should be avoided. Another important factor of sound transmission is reverberation whose main effect is to limit the achievable pulse rate.

To compensate for time-of-flight fluctuations, a stochastic error model was developed. The measurement fluctuations were shown to be correlated in time and space. The correlation time constant is about 20-30 sec and the correlation distance constant is about 30-90 cm, depending on the atmospheric conditions in the room. The model is capable of describing the joint space-time statistics of the time-of-flight fluctuations of an arbitrary number of sender-receiver pairs. The stochastic model was verified with experimentally obtained data.

An extended Kalman filter was then used to best estimate the kinematic error given the knowledge of the previously mentioned environmental error models. The key problem of this method was the difficulty of characterizing the (unknown) kinematic error spectrally. The problem was solved by modeling the dynamic behavior of the kinematic error as a double integrator. The noise states were not included in the EKF model.

To further reduce the estimation error, a measurement noise compensation scheme was introduced. This idea exploits the known spacial correlation properties of the noise: One or more reference senders, which are placed at strategic points, are used to optimally compensate for the measurement noise of the main sender.

The system behavior was simulated to understand how certain factors contribute to the overall system tracking error. First the system was simulated without measurement noise. It was shown that the main influence of the extended Kalman filter dynamics on the system behavior was through nonlinear dynamic cross coupling and sensitivity to acceleration.

Next, the performance of the EKF was tested outside the loop. It turned out that the EKF estimation error is about the same size as the measurement noise. To improve that requires any of the following: Better knowledge of the spectral properties of the kinematic error, very slow trajectories, the use of the above mentioned measurement noise compensation method, or sensor fusion with a vision system.

Finally, the entire system including the measurement noise was simulated. It turned out that the overall tracking error is identical to the estimation error, except for very fast trajectories where the dynamic control error also contributes to the tracking error. For atmospherically calm, small to medium sized rooms the maximum tracking error can be expected to be about 0.4 mm in each direction of space. When reference senders are used to compensate for the noise, one can expect this value to be reduced to about 0.15 mm in the vicinity of a reference sender. With some notable exceptions,

the speed of the trajectory plays only a minor role in the precision achievable.

9.2 Suggestions for Further Research

Any future continuation of the work presented in this thesis should include an experimental setup to test the overall performance of the system in a real robot environment. Associated with an experimental setup should be an investigation into the adaptive identification of environmental factors such as the correlation time constant, the correlation distance constant, the inverse sound speed and the inverse sound speed gradient.

Further, it would be important to investigate which practical tasks and configurations preclude complete observability. Are there microphone-sender configurations which ensure observability for all common tasks? In connection with this question, it would be interesting to study the best strategy to pursue when partial unobservability does occur.

Another aspect requiring investigation is the reduction of the numerical complexity of the control algorithm. In particular, it would be desirable to find computationally inexpensive approximations for the spatial covariance kernel and the inverse kinematic function.

Bibliography

- [1] T.Lozano-Perez, *Robot Programming*, Proc. of the IEEE, Vol.71, No.7., July 1983.
- [2] J.Chen and L.M.Chao, *Positioning Error Analysis for Robot Manipulators with All Rotary Joints*, Journal of Robotics and Automation, Vol.RA-3, No.6, Dec. 1987, pp.539-544.
- [3] C.P.Day, *Robot Accuracy Issues and Methods of Improvement* 17th. ISIR, pp. 5.1-5.24, 1987.
- [4] R.Gottwald, *Kernspace: An Automated Non-Contact 3-D Measuring System for Industrial Applications*, Proc. 7th. Int. Conf. on Rob. Vision and Sensory Contr.
- [5] B.Sorensen, G.Ben Yang, R.Starr, *Performance of the Minnesota Scanner: A Prototype for 3-D tracking of moving body segments*, IEEE Conf. on Robotics and Automation, pp.793-800, 1987.
- [6] H.El-Zorkany, R.Liscano, B.Tondu, G.Sawatzky, *Sensor-Based Location and Trajectory Specification and Correction in Robot Programming*, 16th. ISIR, pp.613-655, 1986.
- [7] J.S. Schoenwald, C.V.Simith, *Twotone CW Acoustic Ranging Techniques for Robotic Control*, IEEE Ultrasonics Symposium, pp.469-474, 1984.

- [8] J.L.Rose, P.A.Meyer, *Ultrasonic Signal Processing Concept for Measuring the Thickness of Thin Layers*, Material Evaluation, December 1974.
- [9] G.Mauris, L.Foulloy, H.Clereot, *A Microprocessor Based Sonar Sensor*, Int. Conf. on Advanced Robotics, Versailles, Oct. 1987.
- [10] C.Biber, S.Ellin, E.Shenk and J.Stempeck, *The Polaroid Ultrasonic Ranging System*, 67th. Convention, Audio Eng. Society, Technical Paper 1696(A-8), N.Y., Oct. 1980.
- [11] S.Ciraria, *An Ultrasonic Ranging System*, Byte Magazine, pp.113-123, Oct. 1984
- [12] P.M.Gammell, *Improved Ultrasonic Detection Using the Analytic Signal Magnitude*, Ultrasonics, pp.73-76, March 1981.
- [13] J.S.Lamancusa and F.Figueroa. *An Ultrasonic Ranging System for End-Effector Position Feedback in Robotics*, Final Report, National Science Foundation, Grand# ECS-8504464, Dept. of Mech. Eng., Pennsylvania State University. Dec. 1987.
- [14] K.Sasaki, M.Takano and K.Akeno, *Ultrasonic Range Sensor Assists 6-DOF Manipulator by Locating Objects in 3-Dimensional Space*, 18th. ISIR, pp.211-220.
- [15] M.K.Brown, *On Ultrasonic Detection of Surface Features*, 1986 IEEE Robotics and Automation Conf., pp.1785-1790
- [16] D.E.Whitney, *The Mathematics of Coordinated Control of Prosthetic Arms and Manipulators*, Journal of Dynamic Systems, Measurement, and Control, Dec. 1972, pp.303-309.
- [17] F.Miyazaki and S.Arimoto, *Sensory Feedback for Robot Manipulators*, Journal of Robotic Systems, Vol.2, 1985, pp.53-71.

- [18] J.Y.S.Luh, M.W.Walker, and R.P.C.Paul, *Resolved-Acceleration Control of Mechanical Manipulators*, IEEE Transactions on Automatic Control, Vol. AC-25, No. 3, June 1980, pp.468-473.
- [19] O.Khatib, *A Unified Approach for Motion and Force Control of Robot Manipulators: The Operational Space Formulation*, IEEE Journal of Robotics and Automation, Vol.Ra-1, No.1, Feb. 1987, pp.43-53.
- [20] E.G.Gilbert and I.J.Ha, *An Approach to Nonlinear Feedback Control with Applications to Robotics*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-14, No. 6, Nov. 1984, pp.879-884.
- [21] K.Furuta, K.Kosuge, and N.Mukai, *Trajectory Tracking Control of Articulated Robot Arm with Laser Beam Tracking System*, IECON'85, pp.307-312.
- [22] W.Khalil and C.Chevallereau, *An Efficient Algorithm for the Dynamic Control of Robots in the Cartesian Space*, Proc. 26th. Conf. on Decision and Control, Dec. 1987, pp.582-588.
- [23] E.Freund, *Fast Nonlinear Control with Arbitrary Pole-Placement for Industrial Robots and Manipulators*, The International Journal of Robotic Research, Vol. 1, No. 1, 1982, pp.65-78.
- [24] J.K.Mills and A.A.Goldenberg, *Robust Control of Robotic Manipulators with Task Space Feedback*, Proc. of 25th Conf. on Decision and Control, Dec. 1986, pp.423-428.
- [25] H.W.Wehn and P.R.Belanger, *On Approximate Inverse Kinematic Operators for Cartesian Space Feedback*, Proc. 26th. Conf. on Decision and Control, Dec. 1987, pp.647-650;

- [26] J.Studenny, *Control of Robotic Manipulators Using Acceleration Feedback*, Ph.D. thesis, McGill University, Jan. 1987.
- [27] H.E.Bass and L.N.Bolen. *Ultrasonic Background Noise in Industrial Environments*, J. Acoust. Soc. Am. 78 (6), Dec.1985, pp.2013-2016.
- [28] L.E.Kinsler and A.R.Frey. *Fundamentals of Acoustics*, second edition, J.Wiley&Sons, 1962
- [29] H.Tennekes and J.L.Lumley. *A First Course in Turbulence*, The MIT Press, 1972.
- [30] A.Bejan. *Convection Heat Transfer*, J.Wiley&Sons, 1984, pp.159-196.
- [31] H.Oertel, Jr., *Three Dimensional Convection Within Rectangular Boxes*, Natural Convection in Enclosures, Heat Transfer Division, ASME, HTD-VOL.8, 1980, pp.11-16.
- [32] A.S.Monin and A.M.Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence*, Vol.2, MIT Press, 1975
- [33] E.Philippow (editor), *Taschenbuch der Elektrotechnik*, Vol.2, VEB Verlag Technik Berlin, 1977, p.747
- [34] D.D.Reynolds, *Engineering Principles of Acoustics*. Allyn and Bacon Inc, Boston, 1981
- [35] J.Davenit and R.S. Hartenberg, *A Kinematic Notation for Lower-Pair Mechanisms Based on Matrices*, ASME Journal of Applied Mechanics, June 1955, pp.215-221
- [36] R.Featherstone, *Position and Velocity Transformations between Robot End Effector Coordinates and Joint Angles*, The Int. Journal of Robotics Research, Vol.2, No.2, Summer 1983, pp.35-45

- [37] J. Angeles, *On the Numerical Solution of the Inverse Kinematic Problem*, The Int. Journal of Robotics Research, Vol.4, No.2, Summer 1985, pp.21-36
- [38] J.M. Hollerbach and Gideon Sahar, *Wrist-Partitioned, Inverse Kinematic Acceleration and Manipulator Dynamics*, The Int. Journal of Robotics Research, Vol.2, No.2, Winter 1983, pp.61-75
- [39] R.P. Paul, *Robot Manipulators: Mathematics, Programming and Control*, MIT Press, 1981
- [40] J. Craig, *Introduction to Robotics, Mechanics and Control*, Addison Wesley, 1986.
- [41] S. Elgazzar, *Efficient Kinematic Transformations for the PUMA560 Robot*, IEEE Journal of Robotics and Automation, Vol. RA-1, No.3, Sept. 1985, pp.142-151
- [42] C.H. Liu and Y.M. Chen, *Multi-Microprocessor-Based Cartesian Space Control Techniques for a Mechanical Manipulator*, IEEE Journal of Robotics and Automation, Vol. RA-2, No.2, June 1986.
- [43] Prof. L. Daneshmand, McGill University, Montreal, *personal communication*, March 1990.
- [44] R.P. Paul, B. Shimano, and G.E. Meyer, *Differential Kinematic Control for Simple Manipulators*, IEEE Trans. Syst., Man, Cybern., Vol. SMC-11, No.6, 1981, pp.449-455
- [45] D.E. Orin, W.W. Schrader, *Efficient Computation of the Jacobian for Robot Manipulators*, The International Journal of Robotics Research, Vol.3, No.4, Winter 1984, pp.66-75.
- [46] G. Zames, *On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems Part I: Conditions Derived Using Concepts of Loop Gain, Conicity, and*

- Positivity*, IEEE Transactions on Automatic Control, Vol. AC-11, No.2, 1966, pp.228-237
- [47] G.Zames, *On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems Part II: Conditions Involving Circles in the Frequency Plane and Sector Nonlinearities*, IEEE Transactions on Automatic Control, Vol. AC-11, No.3, 1966, pp.465-475
- [48] J.C.Doyle and G.Stein, *Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis*, IEEE Transactions on Automatic Control, Vol.AC-26, No.1, 1981, pp.4-16
- [49] B.A.Francis, *A Course in H_∞ Control Theory*, Springer Verlag, 1987
- [50] J.C.Doyle, K.Glover, P.Khargonekar and B.Francis, *State Space Solution to Standard H_2 and H_∞ Control Problems*, Proc. American Contr. Conf., Atlanta, June 1988.
- [51] Y.Chen, *Replacing a PID Controller by a Lag-Lead Compensator for a Robot - A Frequency Response Approach*, IEEE Trans. on Robotics and Automation, Vol.5, No.2, April 1989
- [52] K.J.Astrom, B.Wittenmark, *Computer Controlled Systems*, Prentice-Hall, 1984.
- [53] W.Rudin, *Principles of Mathematical Analysis*, 3rd Edition, McGraw-Hill, 1976, p.90
- [54] M.J.Dyment and D.F.Liang, *Dynamic Performance Analysis of Navstar/GPS Navigation Filters*, AGARD Conference Proceedings No.298 on Precision Positioning and Inertial Guidance Sensors Technology and Operational Aspects, pp15.1-15.19

- [55] T.Tarnoczy, *Sound Focussing Lenses and Waveguides*, Ultrasonics, July-Sept.1965, pp.115-127
- [56] D.L.Folds, *Status of Ultrasonic Lens Development*, Underwater Acoustics and Signal Processing, Proc. NATO Adv. Study Inst. Copenhagen, 1980, pp.263-279
- [57] L.A.Chernov, *Wave Propagation in a Random Medium*, McGraw-Hill, 1960
- [58] R.Laval and Y.Labasque, *Spacial and Temporal Processing* Underwater Acoustics and Signal Processing, Proc. NATO Adv. Study Inst. Copenhagen, 1980, p.49
- [59] D.G.Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, Inc.
- [60] P.S.Maybeck, *Stochastic Models, Estimation, and Control*, Academic Press, Vol.1-3, 1979
- [61] Watsmart and Optotrak product literature, Northern Digital Inc., Waterloo, Ontario, Canada.
- [62] *GP-8-3D Sonic Digitizer Operator's Manual*, Science Accessories Corporation, Southport, CN, 1985.
- [63] H.W.Stone, *Kinematic Modeling, Identification, and Control of Robotic Manipulators*, Kluwer Academic Publishers, 1987
- [64] J.Lloyd and V.Hayward, *Kinematics of common industrial robots*, Robotics, Vol.4, 1988.

Appendix A

Covariance Approximation

It is desirable to find an approximate closed form solution for the spatial covariance equation (6.10) because the numerical evaluation of the double integral is computationally expensive. This is particularly important for the case where the two measurement channels share the same sender and/or the same receiver. The reason is that the evaluation must be done in real time if a reference sender is to be used for optimal noise reduction.

For this case, assume that the two ray-paths Γ_1 and Γ_2 share a sender or a receiver, and let the angle between the two paths be denoted by φ . Further, without loss of generality, let $\gamma > 1$ be such that $d_j = \gamma d_i$. Furthermore, let $\alpha = 1/D_1$.

Unfortunately, if $\|\cdot\| = \|\cdot\|_2$ in (6.10) is taken to denote the 2-norm, then (6.10) does not have a closed form solution. On the other hand, if the 2-norm is approximated by the average of the 1-norm and the ∞ -norm, i.e. $\|\cdot\|_2 \approx (\|\cdot\|_1 + \|\cdot\|_\infty)/2$, then (6.10) with (6.6) does have a closed form solution:

If $|\varphi| \neq \arctan(2)$:

$$P_{II} = \begin{cases} F_0(\gamma) & \text{if } \varphi = 0^\circ \\ 2F(\frac{3}{2}|s|) - 4F(|s|) + F(b_{p1}) + F(b_{m1}) \exp(-\alpha\gamma d_i) & \text{if } 0^\circ < |\varphi| \leq 45^\circ \\ F(\frac{3}{2}|s|) - 4F(|s|) + 2F(b_{p2}) + F(b_{m1}) \exp(-\alpha\gamma d_i) & \text{if } 45^\circ < |\varphi| \leq 90^\circ \\ F(\frac{3}{2}|s|) - 2F(b_{p2}) + F(b_{p1}) \exp(-\alpha\gamma d_i) & \text{if } 90^\circ < |\varphi| \leq 135^\circ \\ F(b_{p1})(\exp(-\alpha\gamma d_i) - 1) & \text{if } 135^\circ < |\varphi| \leq 180^\circ \end{cases}$$

if $|\varphi| = \arctan(2)$:

$$P_{II} = F(\frac{3}{2}|s|) - 4F(|s|) + 2F(b_{p2}) - \alpha d_i \exp(-\alpha\gamma d_i) c_1 D_1^2$$

where

$$F_0(\gamma) = c_1 D_1^2 (2d_i/D_1 + \exp(-d_i/D_1) + \exp(-\gamma d_i/D_1) + \exp(-(\gamma - 1)d_i/D_1) - 1)$$

$$F(\cdot) = c_1 D_1^2 \frac{(\exp(-\alpha d_i(\cdot)) - 1)}{(\cdot)}$$

$$|s| = |\sin(\varphi)|, \quad |c| = |\cos(\varphi)|$$

$$b_{p1} = |s|/2 + |c|, \quad b_{p2} = |s| + |c|/2, \quad b_{m1} = |s|/2 - |c|$$

Appendix B

The PUMA 600 Manipulator

The kinematic structure of the PUMA600 is shown in Fig. B.1. The following nominal values were used for the relevant kinematic parameters: $b_1=0\text{mm}$, $b_2=149\text{mm}$, $a_3=432\text{mm}$, $a_4=20\text{mm}$, $b_4=432\text{mm}$. The moving sensor was located at the origin of the 4th frame.

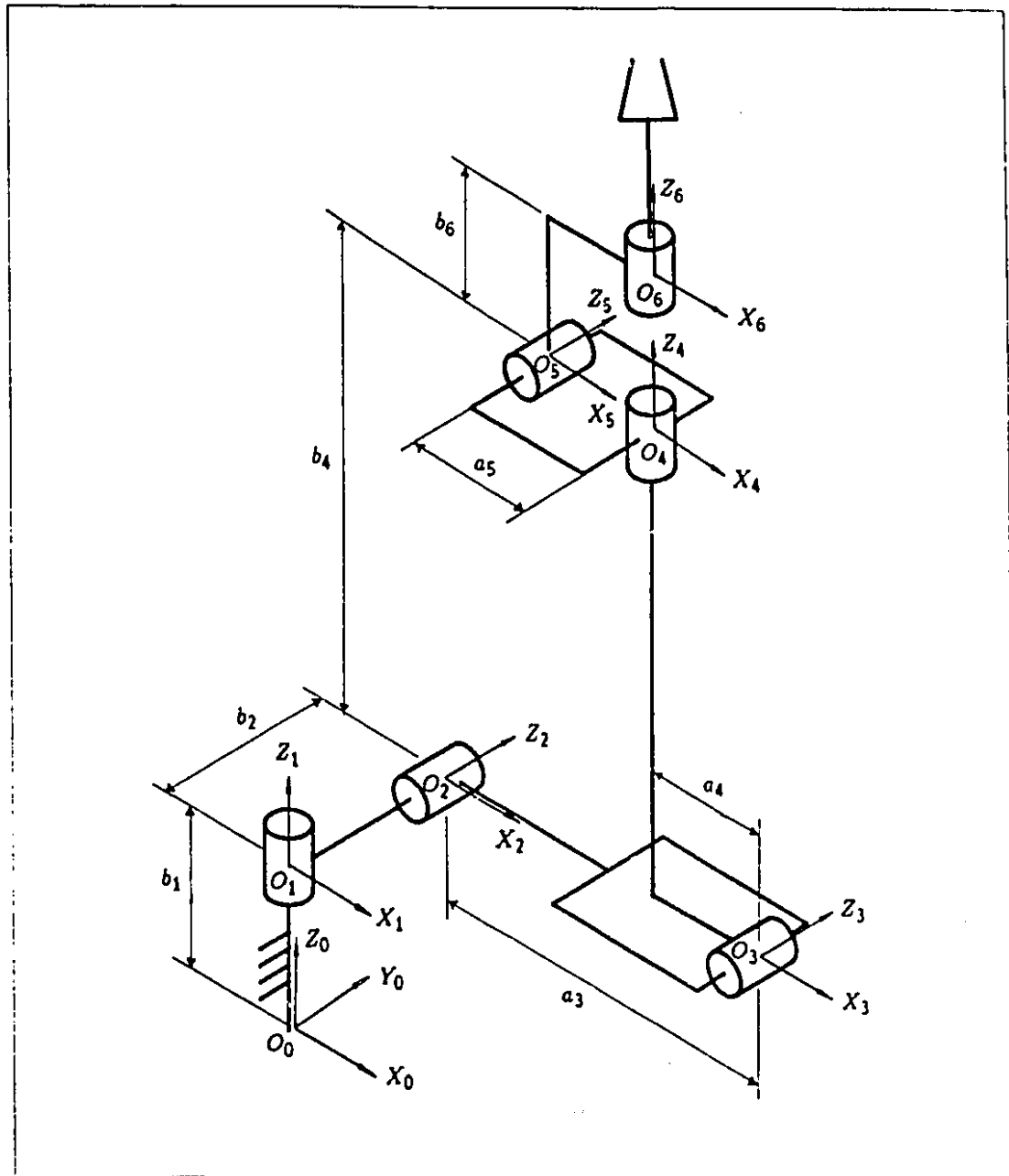


Figure B.1: The PUMA600 robot manipulator

Appendix C

The Ultrasound Transducer

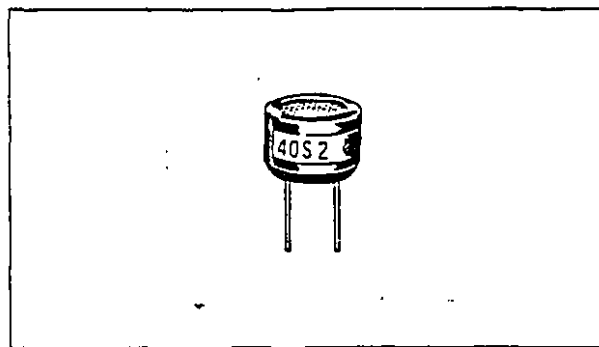


Figure C.1: The ultrasound transducer muRata MA40S2R

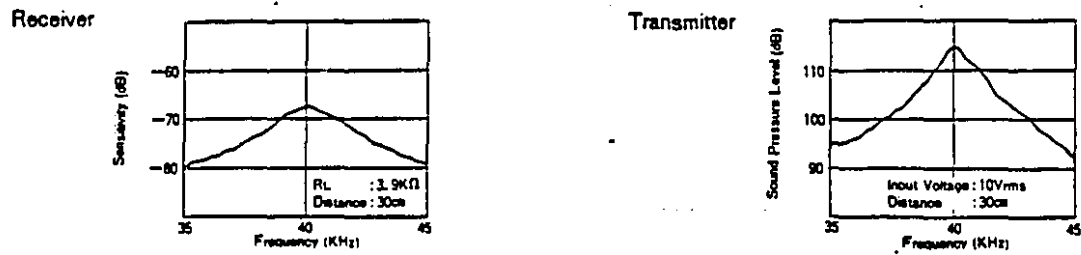


Figure C.2: Ultrasound transducer frequency characteristic.

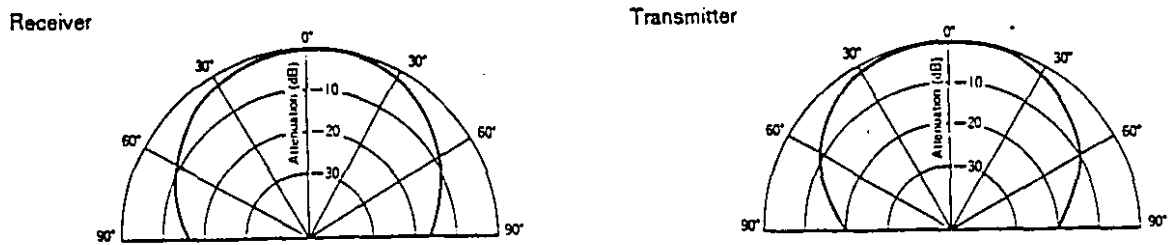


Figure C.3: Ultrasound transducer radiation characteristic.

Appendix D

Circuit Diagrams

For the two senders, the gated CW signals were generated by 5/3 Intel 8254-2 16 bit counter-timer chips and fed directly to drivers for the two ultrasonic sender transducers. The sender transducer and the drivers were connected by shielded cables. The parameters for the gated CW signal like burst length, sampling period and delay between the first and the second sender burst are defined by the controlling C-program, which in turn does the setup for the Intel 8254-2. An additional 4/3 Intel 8254-2 chips were used to count the (TOF+TOA) for the four receiver channels with 10 Mhz clock speed. This brings the total number to nine 16 bit counter-timers on three chips. Counting is enabled when a ranging signal is sent and is disabled when the analog part of the channel's receiver had detected a zerocrossing of the signal after a preceding trigger. The analog receiver was located as closely to the transducer as possible in order to improve the noise problems with the very weak transducer signal. The whole receiver electronic for one transducer fitted comfortably on a small copperplated board of dimensions $2.5\text{cm} \times 6.5\text{cm}$. Shielded cables connected the computer with the four analog receivers. The cables carried the count disable TTL output signal from the analog board and the carefully filtered powerbus from the computer. Circuit-, and block-diagrams of the receiver and interface card can be found below.

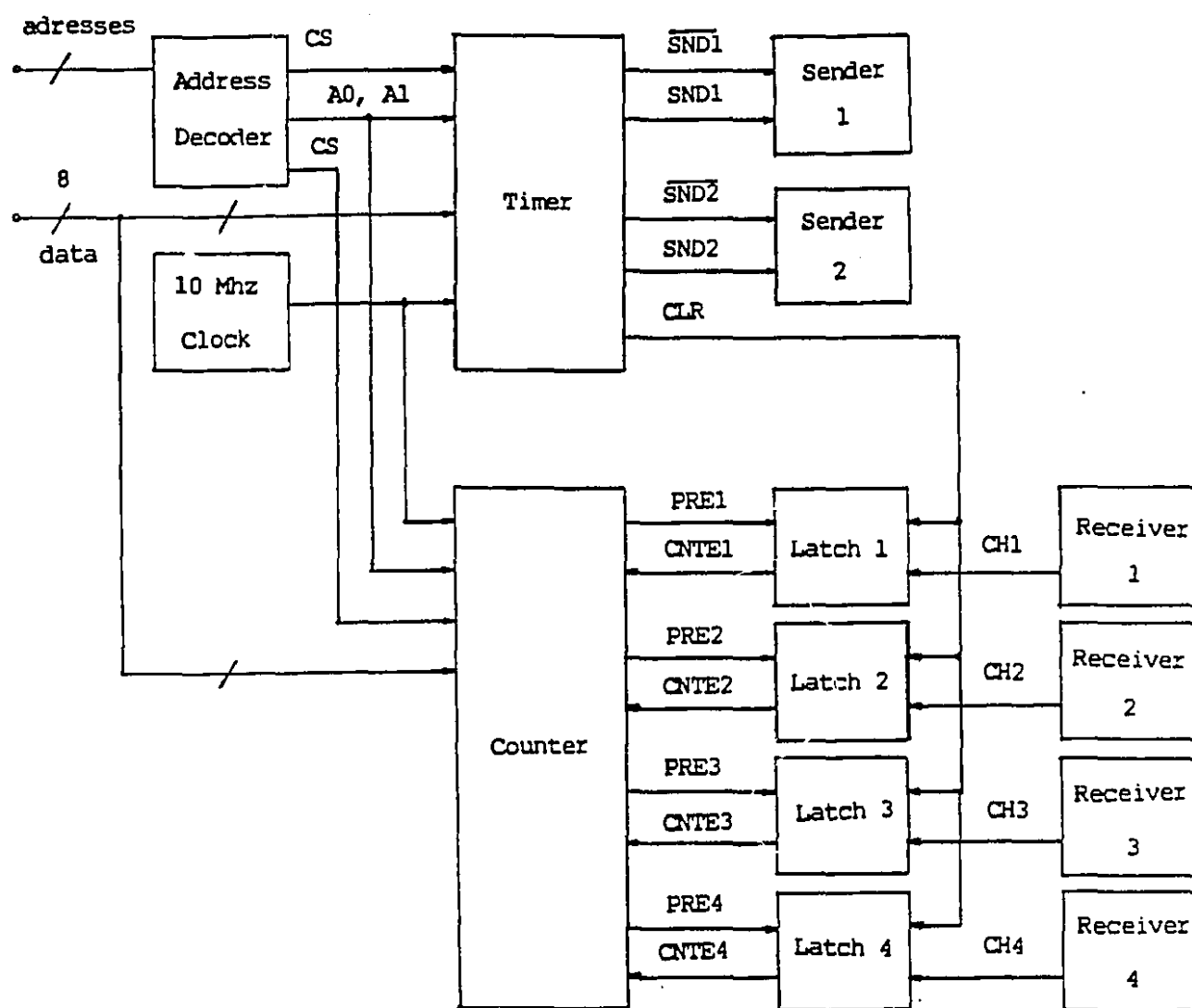


Figure D.1: Block diagram of the whole ultrasound ranging system.

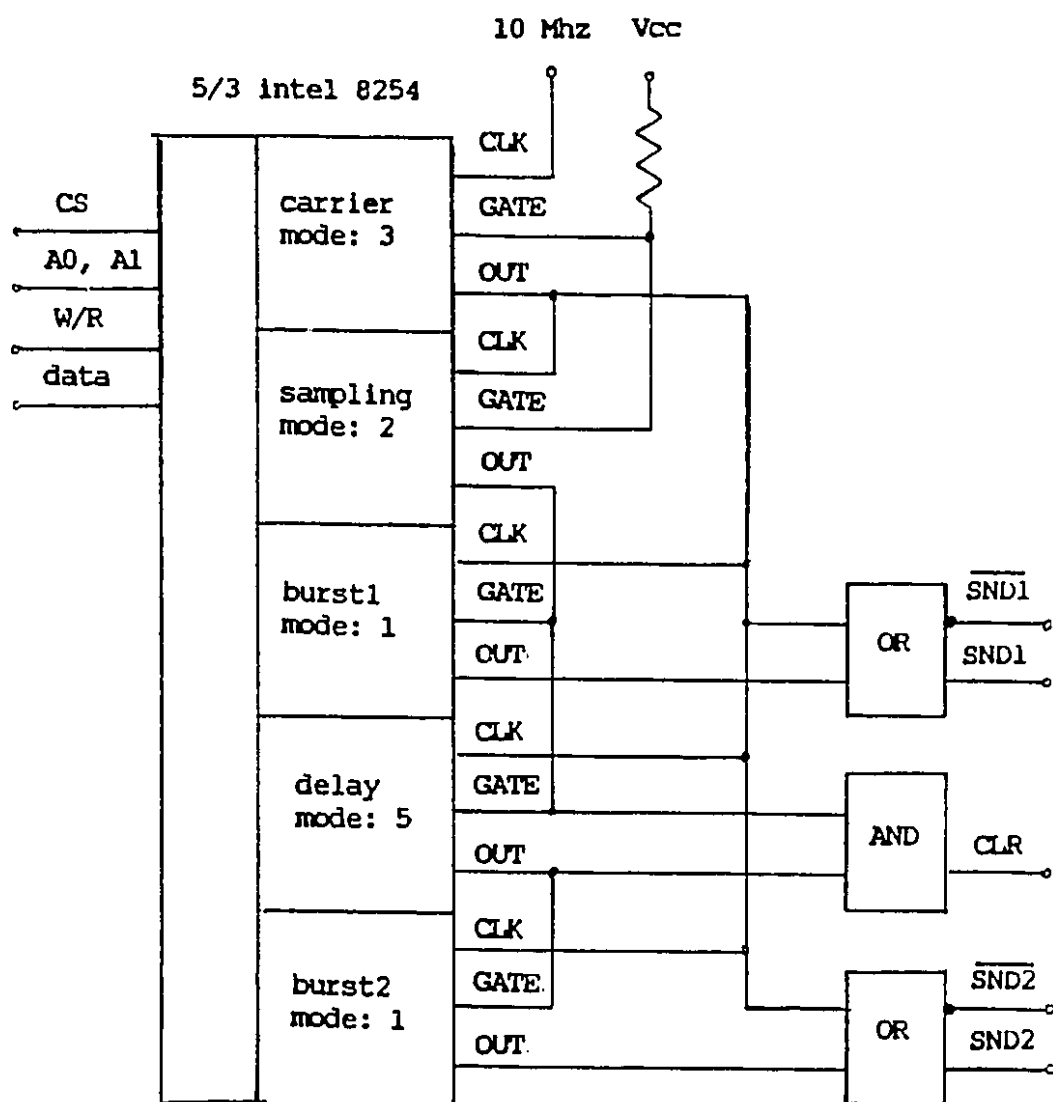


Figure D.2: Diagram of the TIMER circuit.

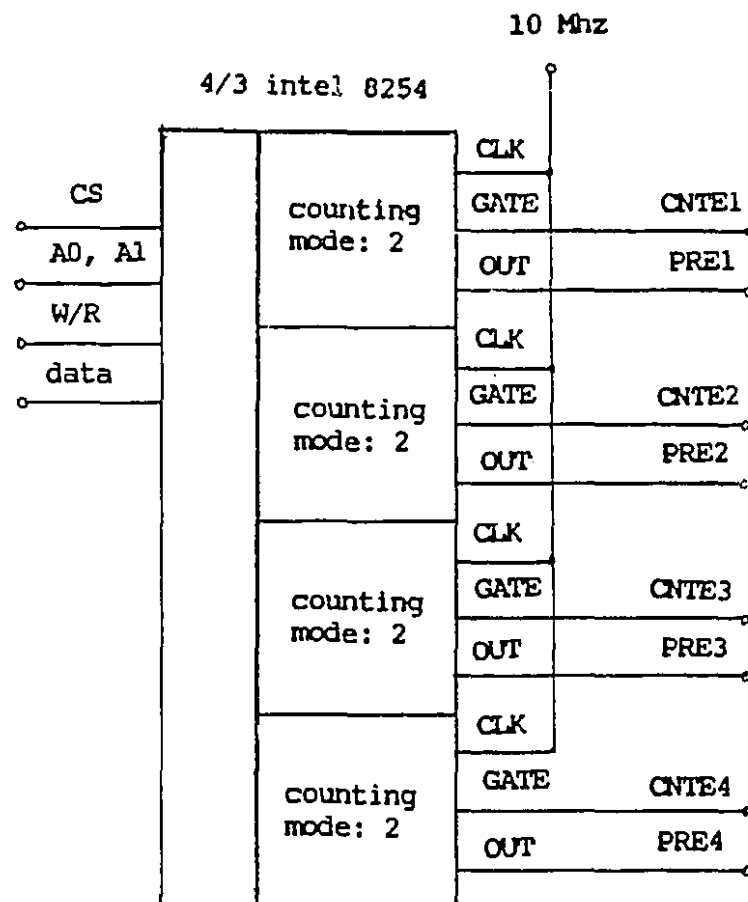


Figure D.3: Diagram of the COUNTER circuit.

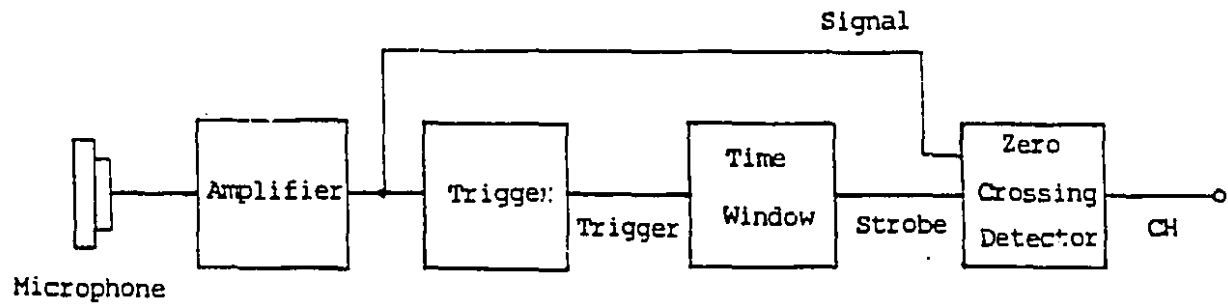


Figure D.4: Block diagram of the RECEIVER circuit.

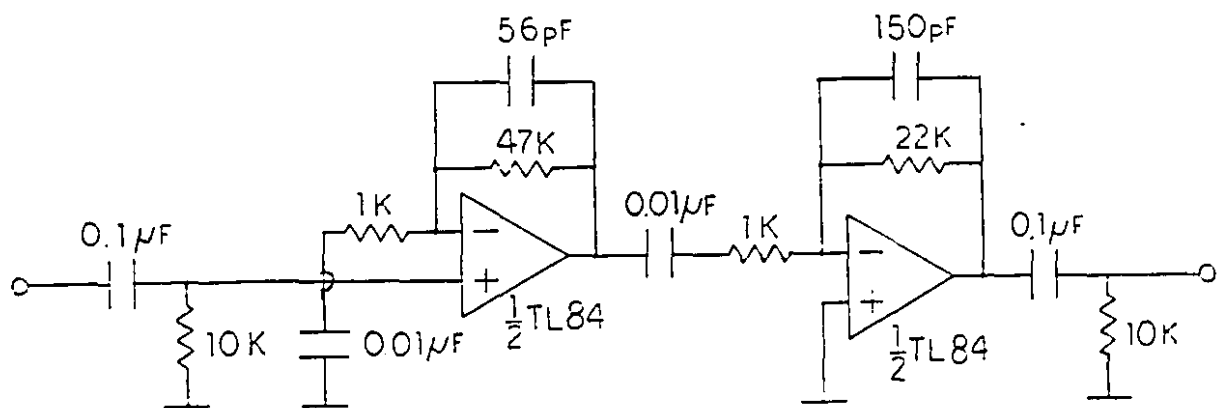


Figure D.5: Diagram of the AMPLIFIER circuit.

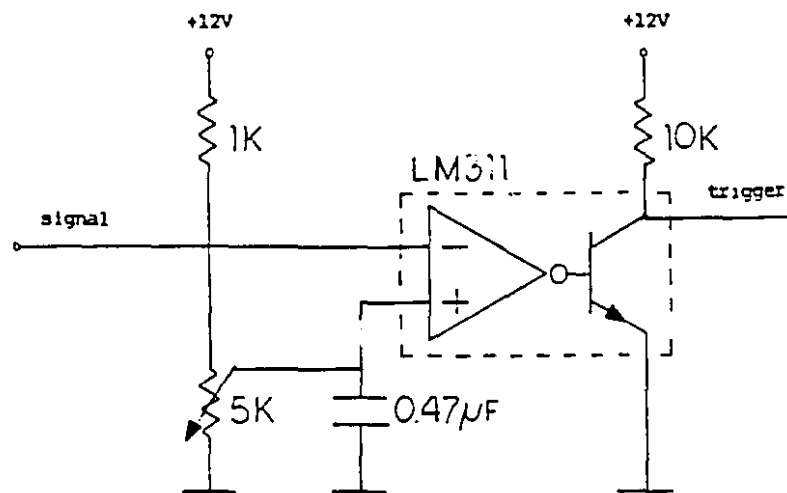


Figure D.6: Diagram of the TRIGGER circuit.

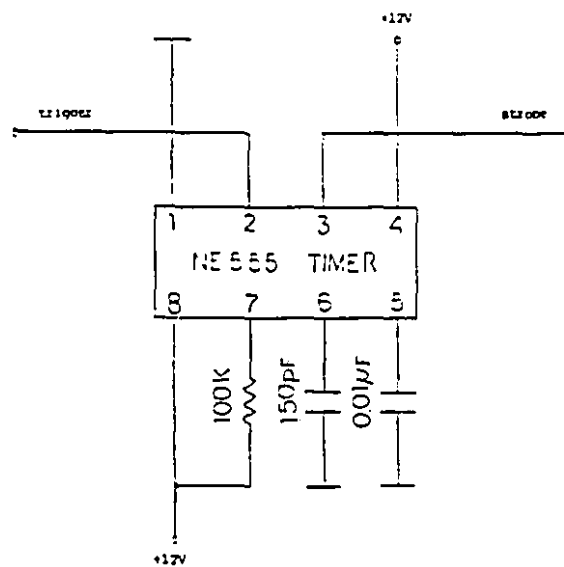


Figure D.7: Diagram of the TIME WINDOW circuit.

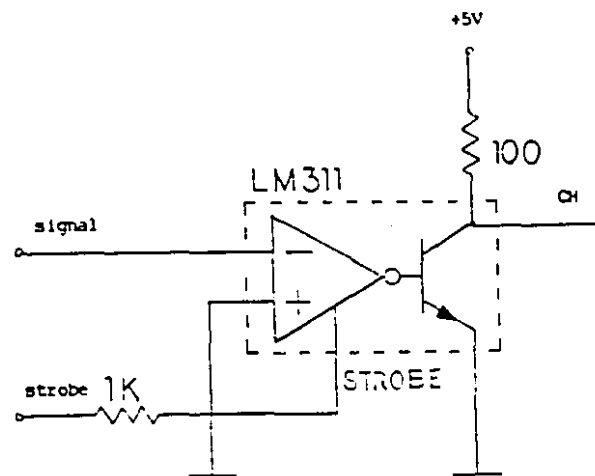


Figure D.8: Diagram of the ZERO CROSSING DETECTOR circuit.

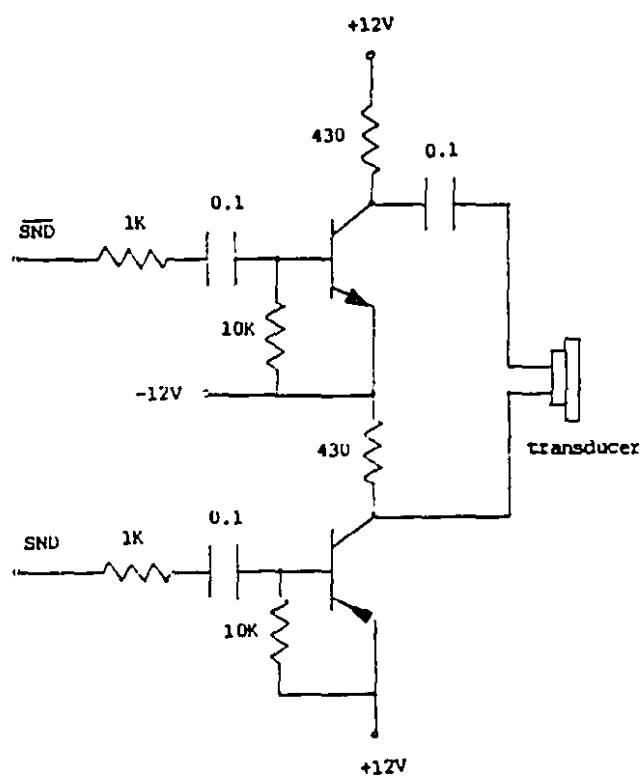


Figure D.9: Diagram of the SENDER circuit.

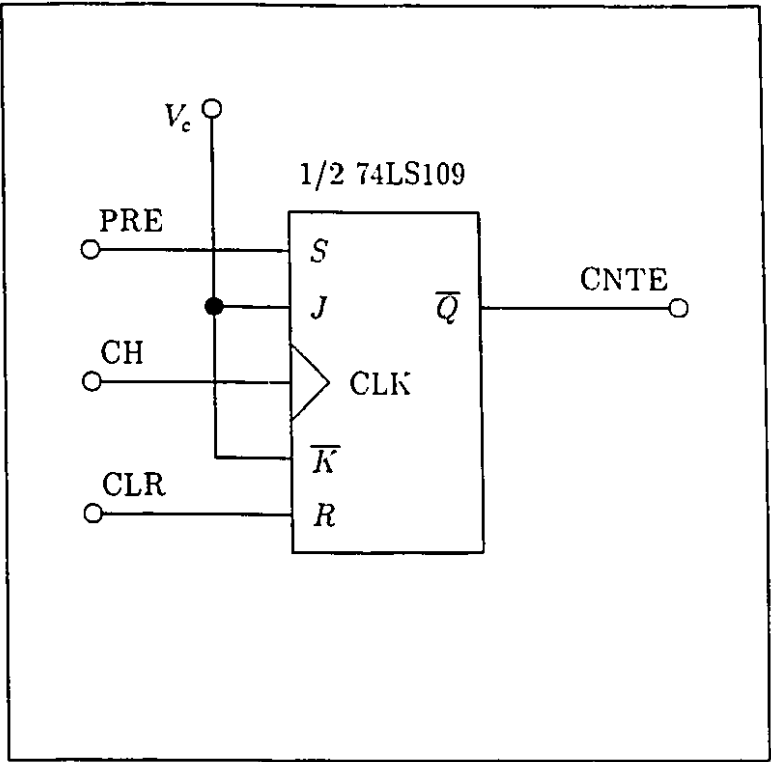


Figure D.10: Latch circuit

PRE	CLR	CH	CNTE
H	L	X	H
H	H	↑	L
L	H	X	L

Table D.1: Truth table of latch circuit

Appendix E

C-Program

```
/*                      ultsound.h
/*
    This program contains the address and command
    definitions for the Intel 8254-2 counter chips.
*/

/* Macros for converting bytes into short integers and vice versa */

#define MAKE_SHORT(msb,lsb)  msb*0x0100+lsb
#define MSB(x)  _rotr(x&0xFF00,8)
#define LSB(x)  x&0x00FF

/* Address definitions for the 8254-2 chips on the ultrasound card */

#define CHIP1_CO  768
#define CHIP1_C1  769
#define CHIP1_C2  770
#define CHIP1_CR  771

#define CHIP2_CO  772
#define CHIP2_C1  773
#define CHIP2_C2  774
#define CHIP2_CR  775

#define CHIP3_CO  776
#define CHIP3_C1  777
#define CHIP3_C2  778
#define CHIP3_CR  779
```

```

/* Programming commands for the 8254-2 chips          */
/* (first LSB then MSB for read/write; binary 16bit counters) */

#define CO_MODE0 0x0030
#define CO_MODE1 0x0032
#define CO_MODE2 0x0034
#define CO_MODE3 0x0036
#define CO_MODE4 0x0038
#define CO_MODE5 0x003A

#define C1_MODE0 0x0070
#define C1_MODE1 0x0072
#define C1_MODE2 0x0074
#define C1_MODE3 0x0076
#define C1_MODE4 0x0078
#define C1_MODE5 0x007A

#define C2_MODE0 0x00B0
#define C2_MODE1 0x00B2
#define C2_MODE2 0x00B4
#define C2_MODE3 0x00B6
#define C2_MODE4 0x00B8
#define C2_MODE5 0x00BA

/* Latch commands for reading values from the 8254-2 chips */

#define LATCH_CO 0x0000
#define LATCH_C1 0x0040
#define LATCH_C2 0x0080/

/*
                                exp_read.c

This program reads the counter values of the receiver channels
for the ultrasound card and writes the obtained values into a
data file for further processing. It also sets up the signal
generators for the two ultrasound sender which are supported
by the card. The device uses intel 8454-2 chips for both counting
and signal generation.
*/

```

```

#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
- #include "ultrasound.h"

#define CARRIER_PERIOD 250

main()
{
    unsigned int input(), clean(), x, count, end_count;
    unsigned int read1_time, read2_time;
    unsigned int in1_1, in1_2, in1_3, in1_4;
    unsigned int in2_1, in2_2, in2_3, in2_4;
    unsigned int old1_1=0, old1_2=0, old1_3=0, old1_4=0;
    unsigned int old2_1=0, old2_2=0, old2_3=0, old2_4=0;

    unsigned int burst_length, sampling_period, delay_sender;
    char filename_a[41], filename_b[41];

    FILE *fptr_a, *fptr_b;

    /* parameter input */

    printf("Number of samples to be taken ? ");
    scanf("%u",&end_count);
    printf("Burst length ? ");
    scanf("%u",&burst_length);
    printf("Delay of sender #2 or switch #2 off ? ");
    scanf("%u",&delay_sender);
    printf("Sampling period ? ");
    scanf("%u",&sampling_period);
    printf("Data filename for sender #1 ? ");
    scanf("%s",filename_a);
    if (delay_sender!=0)
        { printf("Data filename for sender #2 ? ");
        scanf("%s",filename_b);}

    /* set carrier period */

    outp(CHIP1_CR, CO_MODE3);

```



```
    outp(CHIP1_CO, LSB( CARRIER_PERIOD ));
    outp(CHIP1_CO, MSB( CARRIER_PERIOD ));

    /* set the sampling period */

    outp(CHIP1_CR, C1_MODE2);
    outp(CHIP1_C1, LSB( sampling_period ));
    outp(CHIP1_C1, MSB( sampling_period ));

    /* set the burst length for sender #1 */

    outp(CHIP1_CR, C2_MODE1);
    outp(CHIP1_C2, LSB( burst_length ));
    outp(CHIP1_C2, MSB( burst_length ));

    /* set the delay between sender #1 burst and sender #2 burst */

    outp(CHIP2_CR, C1_MODE5);
    outp(CHIP2_C1, LSB( delay_sender ));
    outp(CHIP2_C1, MSB( delay_sender ));

    /* set the burst length for sender #2 */

    outp(CHIP2_CR, C2_MODE1);
    outp(CHIP2_C2, LSB( burst_length ));
    outp(CHIP2_C2, MSB( burst_length ));

    /* setup receiver #1 */

    outp(CHIP3_CR, CO_MODE2);
    outp(CHIP3_CO, 0xFFFF);
    outp(CHIP3_CO, 0xFFFF);

    /* setup receiver #2 */

    outp(CHIP3_CR, C1_MODE2);
    outp(CHIP3_C1, 0xFFFF);
    outp(CHIP3_C1, 0xFFFF);

    /* setup receiver #3 */
```

```

    outp(CHIP3_CR, C2_MODE2);
    outp(CHIP3_C2, 0xFFFF);
    outp(CHIP3_C2, 0xFFFF);

/* setup receiver #4 */

    outp(CHIP2_CR, C0_MODE2);
    outp(CHIP2_C0, 0xFFFF);
    outp(CHIP2_C0, 0xFFFF);

/* start read of receivers */

    fptr_a=fopen(filename_a,"w");
    fptr_b=fopen(filename_b,"w");
    read1_time = sampling_period - 300;
    read2_time = sampling_period - delay_sender - 300;

    for (count=1; count<=end_count+2; count++)
    {

        /* input for sender #1 emission */

        while ( input( CHIP1_CR, CHIP1_C1, LATCH_C1) >= read1_time )
        {;}
        in1_1 = clean(0xFFFF-input(CHIP3_CR,CHIP3_C0,LATCH_C0), &old1_1, count);
        in1_2 = clean(0xFFFF-input(CHIP3_CR,CHIP3_C1,LATCH_C1), &old1_2, count);
        in1_3 = clean(0xFFFF-input(CHIP3_CR,CHIP3_C2,LATCH_C2), &old1_3, count);
        in1_4 = clean(0xFFFF-input(CHIP2_CR,CHIP2_C0,LATCH_C0), &old1_4, count);

        /* input for sender #2 emission */

        while ( input( CHIP1_CR, CHIP1_C1, LATCH_C1) >= read2_time )
        {;}
        in2_1 = clean(0xFFFF-input(CHIP3_CR,CHIP3_C0,LATCH_C0), &old2_1, count);
        in2_2 = clean(0xFFFF-input(CHIP3_CR,CHIP3_C1,LATCH_C1), &old2_2, count);
        in2_3 = clean(0xFFFF-input(CHIP3_CR,CHIP3_C2,LATCH_C2), &old2_3, count);
        in2_4 = clean(0xFFFF-input(CHIP2_CR,CHIP2_C0,LATCH_C0), &old2_4, count);

        /* write input to file */

```

```

        if (count>2 && delay_sender!=0)
{ fprintf(fptra_a,"%7u%7u%7u%7u\n", in1_1, in1_2, in1_3, in1_4);
  fprintf(fptra_b,"%7u%7u%7u%7u\n", in2_1, in2_2, in2_3, in2_4);}

-      if (count>2 && delay_sender==0)
        fprintf(fptra_a,"%7u%7u%7u%7u\n", in1_1, in1_2, in1_3, in1_4);

        /* wait until the end of the sampling period */

        while ( input( CHIP1_CR, CHIP1_C1, LATCH_C1) < read2_time )
        {;}

    }
    close(fptra_a);
    close(fptra_b);
}

unsigned int input( chip_command, chip_counter, latch_counter )
unsigned int chip_command, chip_counter, latch_counter;
/*
    the function "input" returns the integer count of the selected
    counter at the latch instance.
*/
{
    , unsigned int lx, mx;
    outp( chip_command, latch_counter );
    lx=inp( chip_counter );
    mx=inp( chip_counter );
    return( MAKE_SHORT( mx, lx ));
}

unsigned int clean( in, old, count )
unsigned int in, *old, count;
/*
    the function "clean" has the purpose to correct errors that
    are caused when the burst arrival trigger misses a period
    because of amplitude fluctuation. These errors are multiples
    of the carrier period and can therefore be corrected.
*/
{

```

```
int x;
div_t mod;
x = (*old) - in;

- if ( count < 3 )
    { (*old) = in; }

    else if ( x > 200 )
    {
mod = div( x - 200, CARRIER_PERIOD );
    (*old) = in + (mod.quot + 1)*CARRIER_PERIOD;
    }

    else if ( x < -200 )
    {
mod = div( x + 200, CARRIER_PERIOD );
    (*old) = in + (mod.quot - 1)*CARRIER_PERIOD;
    }

    else
    { (*old) = in; }

    return((*old));
}
```