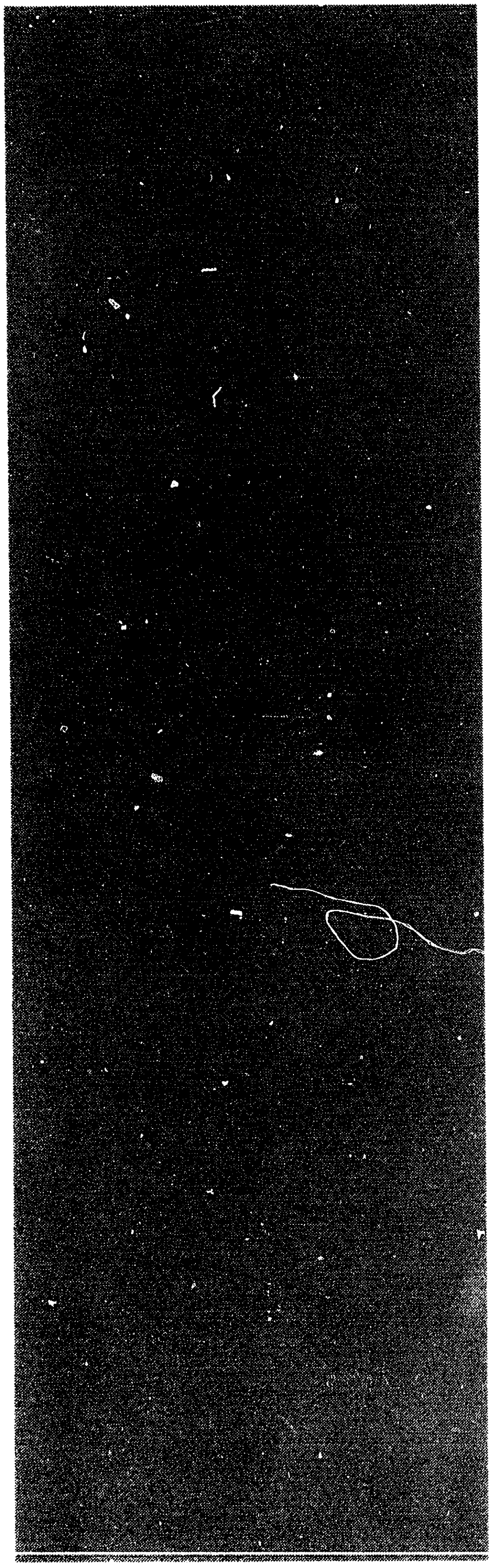
LA-12202-T
Thesis



A New Search for Conversion of Muonium to Antimuonium

Los Alamos

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

This thesis was accepted by the Department of Physics, Yale University, New Haven, Connecticut, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. It is the independent work of the author and has not been edited by the IS-11 Writing and Editing staff.

This work was supported by the US Department of Energy, Office of Energy Research.

LA--12202-T

DE92 002129

A New Search for Conversion of Muonium to Antimuonium

Björn Eckart Matthias

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Contents

List of Figures

List of Tables

Acknowledgements

A great many people have contributed in varied ways to make this experiment possible and successful. I am particularly indebted to all members of the collaboration which carried this experiment from inception to completion. They are: Hyo Ahn, Andreas Badertscher, Frank Chmely, Morton Eckhause, Vernon Hughes, Klaus Jungmann, John Kane, Steve Kettell, Yunan Kuang, Hans-Jürgen Mundinger, Benwen Ni, Herbert Orth, Gisbert zu Putlitz, Reiner Schaefer, Michael Witkowski, and Kim Woodle. All are appreciated as fellow physicists and as friends.

I thank my adviser, Vernon Hughes, for providing the framework for my studies, for the opportunity to conduct them at Los Alamos, and for his confidence in my abilities when they were called into question.

Without the energy and planning of Hans-Jürgen Mundinger, this experiment would not have been carried out promptly. His leadership is as highly valued in the success of the experiment as his optimism was in maintaining my good spirits and endurance when the going was tough.

Reiner Schaefer has been involved in all stages of this work from writing the proposal to helping finish the data reduction. I appreciate highly his steady participation in and enthusiasm about this experiment and its analysis. His encouragement and guidance helped me survive the ordeal of the graduate student.

I am grateful to Steve Kettell for his intuition and talent to ask the right question, both of which contributed to taking the proper experimental steps and to focusing on the relevant ingredients to the data analysis. He shared both in the trials and the excitement of bringing about a functioning apparatus.

Working with Kim Woodle has been a privilege. Among the many things he has taught me, most valued are the recognition that the cleanest approach to a problem is always patient, analytic thought and the uncompromising attitude that there is no substitute for doing things "right."

Herbert Orth's suggestion that one seek to detect the atomic positron of the antimuonium atom was formative to the design of the apparatus. During the first few years of my studies it was his leadership style that defined a productive, fair, and humane research group atmosphere.

I am indebted to Gisbert zu Putlitz, Klaus Jungmann and the rest of the muonium group at the Universität Heidelberg for their patience in holding a position for me while I finished this work.

Benwen Ni carried out calculations of the knock-on e^- background, of the required dimensions for the $NaI(Tl)$ crystal, and of the first version of the matrix trigger which were very helpful to the design of the apparatus.

I thank Hyo Ahn for pointing out to me the failings of the least-squares fitting procedure when seeding the relaxation fit.

The opportunity to conduct my studies at LAMPF allowed me to become integrated into the LAMPF "family," which was a source of professional and personal support over the years.

The experiment would not have been set up without the technical expertise of the secondary beam-lines support group, MP-7. I thank James Abernathy, Steve Cushing, Scott Dick, Chip Johnson, Neville June, Clyde Kain, Glen Camp, Ted Newlin, Chandra Pillai, Bobby Rector, Dick Werbeck, and especially Joe Ivie.

Engineering support was provided by MP-8 and I thank Steve Archuletta, Ed Bush, Arlene Lopez, Mike Mays, Pat McClellan, Louie Morrison, Bobby Poe, Seth Rislove, Gus Roybal, and Joe Salas. Though the power supplies of the SMC preferred to develop difficulties in the middle of the night, Leo Brock, Steve Sleeper, and Ted Stewart of the power supply section were quick to respond in effecting repairs.

Thanks are due MP-4 technicians Art Gonzales and Fred Montoya for manufacturing exotic light guides for the muon counter and Val Armijo for helping repair and assemble the MWPCs. Jim Sena helped arrange the loan of the DC supplies to power the MWPC readout electronics.

All people of the operations group, MP-6, provided varied forms of support besides operating the accelerator. The staff of the electronics equipment pool (LEEP), including Joanne Herrera, Dan Holterman, and Dick Sena, has been a pleasure to interact with. I thank especially Ron Richardson whose repair wizardry got the programmable LeCroy HV supplies back online quickly. The computer maintenance staff is appreciated for their quick response to all computer-related problems. Consulting Mike Oothoudt helped solve many problems that appeared both in the online software and in the replay code. The data analysis center (DAC) at LAMPF provides ideal conditions for analyzing the large volume of data generated by this experiment, and I particularly thank Jim Harrison and Mike Potter for help with many little problems. Martha Hoehn, MP-6, was a pleasure to work with in scheduling the beam times for this experiment.

Cy Hoffman, Martin Cooper, and Dick Mischke of MP-4 taught me much of what I needed to construct the track fitting code, to carry out the maximum likelihood fit and to correctly interpret the results. I am deeply indebted to them for frequently taking the time to discuss these problems and explain the solutions.

Advice from Dave Lee, MP-10, was helpful when the MWPCs used in this experiment were first tested. Bernhard Brinkmöller and Dick Boudrie were helpful with discussions early in the development of the track fitting algorithm.

I am grateful to Chris Francis for her professional preparation of several of the figures in this dissertation. Without Bruce Baca, Louis Cutler, Kathy Garduno, Roberta Marinuzzi, Lupe Neudecker, Patsy O'Brien, Pol Padilla, Bernadette Pesenti, and Chuck Woodard many things often taken for granted at LAMPF would not happen. This experiment enjoyed the confidence and support of Lew Agnew, Jim Bradbury, and Don Hagerman of the MP division office, and of LAMPF director

Gerry Garvey.

Peter Herczeg of the theory group T-5 has been very generous with his time in guiding me through the difficulties of the theoretical questions relevant to this experiment. His enthusiasm about the subject of this work contributed to the drive necessary to complete it.

The tungsten foils used in the slow positron source were vacuum annealed by Rick Honnell of MST-4.

I gratefully acknowledge the loan of the large $NaI(Tl)$ crystal used in this experiment from Byron Dieterli of the University of New Mexico and of the phototubes and bases for this detector from Barrie Hughes of Stanford University.

At Yale, Sara Batter and Jean Belfonte were most supportive and always ready to help solve the frequent administrative impasses. I also thank Pat Fleming, Debbie McGraw, Harley Pretty, and Barbara Voiges for their mastery over the group's business and administrative matters. Bob Broughton was essential for his skill to repair damaged MWPC readout electronics as fast as we could destroy them and send them to him. Larry Trudell saved us from embarrassment by repairing the damaged MWPC planes in time for the first summer of running. I am particularly grateful to Janice Pinkard in the Office of the Adviser to Foreign Students and Scholars at Yale for her expedient handling of all visa matters that were required for my stay in the US.

A NEW SEARCH FOR CONVERSION OF MUONIUM TO ANTIMUONIUM

By
Björn Eckart Matthias

ABSTRACT

To search for conversion of muonium ($M \equiv \mu^+e^-$) to antimuonium ($\overline{M} \equiv \mu^-e^+$) with very low background, a new signature was implemented that required the time-coincident detection of the decay $e^-(\leq 53MeV)$ with the atomic $e^+(\sim 13\ eV)$ from decay of an $\overline{M}$ atom. A 20 MeV/c μ^+ beam was stopped in a 9 mg/cm^2 SiO_2 powder target. Muonium, formed in the powder, diffused into a vacuum region at thermal velocities and was observed for a coincidence of $\overline{M}$ decay products. Any decay e^- was charge and momentum analyzed in a dipole magnet and tracked by an array of MWPCs; any atomic e^+ was electrostatically collected, accelerated to 5.7 keV, and magnetically transported to a microchannel plate detector. To calibrate the signature, M was observed for the first time by the coincidence of its decay e^+ and its atomic e^-. A maximum likelihood analysis of the position distribution of decay origins finds no $\overline{M}$ events and less than 2 at 90% confidence. This places an upper limit on the conversion probability per atom of $S_{\overline{M}} < 6.5 \times 10^{-7}$ (90% C.L.), which corresponds to an upper limit of $G_{M\overline{M}} < 0.16\ G_F$ (90% C.L.) on the effective coupling constant for a $(V - A)$ conversion coupling. In a class of left-right symmetric models, the value of $G_{M\overline{M}}$ may be in this range.

Chapter 1

Introduction

One of the simplest bound systems in which the fundamental interactions among leptons may be studied is muonium. This atom is the Coulomb bound state of a positive muon and an electron and is symbolized by $M \equiv \mu^+e^-$. The structure of this system is presently believed to be well-described by the standard electroweak theory. Therefore, the experimental study of M is well-suited to check this model and to search for any departures from it in the observations made.

Muonium was first formed in 1960 [Hug60,Hug70] by stopping a μ^+ beam in a Ar gas target. Its presence was unmistakably proven by the observation of the Larmor precession signal, determined by the gyromagnetic ratio of 1.4 $MHz/Gauss$, characteristic of the M atom, and the directional anisotropy of the muon decay distribution which is a consequence of parity violation. Further study of the formation process focused on M production in noble gases [Tho69,Sta74] and on obtaining M in vacuum by the beam-foil method [Bol81] and by thermal emission from hot W-foil targets [Mil86] and from silica powder targets at room temperature [Bee86,Woo88].

The study of the hyperfine interval of the M $1^2S_{1/2}$ ground state and the Zeeman effect in its two hyperfine levels at weak and strong magnetic fields has been enthusiastically pursued over the years [Pre61,Zio62,Cle64,dV70,Fav71,Cas77,Mar82a], since the theory here does not suffer from the finite structure effects of the proton, as in the case of hydrogen. This allowed one to focus explicitly on checking quantum electrodynamic predictions.

Spectroscopic experiments on the M $n = 2$ state were made possible with the availability of M in vacuum, where previously any excited state of M in a gas environment would suffer very rapid de-excitation to the ground state by collisions with the gas atoms. The Lamb shift in the $n = 2$ state has been measured by forming $2S$ metastable M with the beam-foil method [Bad84,Ora84] and directly inducing the Lamb shift transition. Recently, the $2^2S_{1/2} - 2^2P_{3/2}$ fine structure interval was measured for the first time [Ket90b]. Finally, using the silica powder source of thermal M in vacuum, the $1S - 2S$ interval has been observed with Doppler-free two-photon laser spectroscopy [Chu88].

This dissertation reports in detail on the most recent experiment addressing a different aspect of the physics of M: the possibility of a spontaneous conversion of M to its antiatom, *antimuonium*, which is symbolized by $\overline{M} \equiv \mu^- e^+$. This process was suggested in analogy to the weak interaction mixing of K^0 and $\overline{K}^0$ by Pontecorvo in 1957 [Pon57]. The $M \longrightarrow \overline{M}$ conversion was thought of as an allowed second-order weak process, proceeding *via* an intermediate state of two neutrinos, where it must be remarked that different neutrino flavors were unknown at the time. It is clear that a direct reaction, if allowed, might proceed at a stronger rate. Within the modern minimal standard model of electroweak interactions, however, there is no channel available for this process. It has been forbidden, along with other lepton-number violating interactions, by the imposition of a discrete symmetry inherent in the grouping of leptons into distinct families between which no direct couplings are included. The search for $M \longrightarrow \overline{M}$ is, therefore, a test of the validity of this postulate and complements spectroscopic experiments on M in challenging the present understanding of the interactions between muons and electrons.

Several earlier experiments have searched for the spontaneous $M \longrightarrow \overline{M}$ conversion [Ama68a,Bar69,Mar82b,Ni87,Hub90]. With the exception of one that searched for the cross-channel reaction $e^- e^- \rightarrow \mu^- \mu^-$ [Bar69], these experiments have all relied upon the radiations induced in matter by μ^- from $\overline{M}$ breakup for a detection signature. This necessarily entailed sensitivity to the $\overline{M}$ admixture in the M, $\overline{M}$ system at only one point in time – when the system struck the material that served as the detection medium in which the μ^- was to induce the signature.

The experiment described in this dissertation is the first to be sensitive to the $\overline{M}$ admixture to the system over its entire lifetime. This was made possible by the use of a silica powder target to produce thermal M, which remained quite localized up until its decay. Furthermore, the signature used required the coincident observation of the energetic decay e^- and the "slow" atomic e^+ , thus including both detectable decay products of $\overline{M}$ disintegration. Establishing the viability of this signature necessitated the first observation of M by the coincident observation of the decay e^+ and the atomic e^- produced in its breakup.

Energetic $e^{\pm}$ were observed in this experiment in an array of four position-sensitive multiwire proportional chambers (MWPCs), where charge identification was possible through separation in a dipole magnetic field between the second and third MWPCs. This arrangement was used to detect particle tracks due to decay $e^{\pm}$ and to trace their trajectories back to their points of origin near the M production target. The corresponding atomic $e^{\mp}$ were detected by electrostatic acceleration and magnetic transport to a microchannel plate (MCP). The most important data characterizing the events recorded were the location of the decay point observed – it was required to lie in the region of the vacuum near the downstream surface of the target, the curvature of the track in the MWPCs to distinguish energetic e^- from e^+ , and the time difference between the track in the MWPCs and a pulse on the MCP – it had to occur in a well-defined time window to be considered due to an atomic $e^{\mp}$ correlated to the decay $e^{\pm}$ track observed. To search for the $M \longrightarrow \overline{M}$ conversion, as many M atoms as possible are produced and one searches for any events fulfilling the $\overline{M}$ signature. The total number of M atoms formed must be known, so that a value or a limit on the probability of $M \longrightarrow \overline{M}$ per atom may be given.

The next chapter places the study of the $M \longrightarrow \overline{M}$ conversion into the context of other searches for lepton-number violation, discusses the effects of external fields on the conversion, and establishes the quantities necessary to interpret the results of the experiment. Chapter 3 surveys previous experiments and describes in detail the apparatus used in this one. The analysis procedures and results are presented in

Chapter 4. Finally, Chapter 5 discusses the implications of the result of this experiment. Several appendices detail the derivations of results used in both theoretical and data analysis chapters and describe some of the data analysis procedures that were too lengthy to fit comfortably into the relevant chapter.

Chapter 2

The Theory of the Conversion

Since the possible coupling of muonium to antimuonium involves new processes among elementary particles, but is also affected by the atomic properties of these bound states, the physics of this conversion is particularly rich. It is, therefore, appropriate to present the theoretical description of the $M \longrightarrow \overline{M}$ conversion in some detail. Also, I will try to place the study of this process into the proper perspective among the many other efforts that seek to illuminate the origin and possible limits of lepton family number conservation.

This chapter will briefly discuss the leptonic portion of the minimal standard model of electroweak interactions; its predictions for processes involving muons will be presented. Next is a survey of the theoretical work to date that attempts to extend the standard model to incorporate leptonic interactions that are forbidden in its minimal version. The process of $M \longrightarrow \overline{M}$ conversion will be focused upon in more detail. Different possibilities of mediating the coupling will be discussed and the implications of the conversion mechanism on the atomic properties of the coupled M, $\overline{M}$ system will be noted. Finally, the connection to our experimental situation and the measureable quantities will be established.

2.1 The Leptonic Sector of the Standard Model

The arguably most successful physical theory to date is the unification of weak interactions with quantum electrodynamics (QED) into a single gauge theory, referred to as the *standard model* [Gla61,Wei67,Sal68]. Its exact, local gauge group is $SU(2)_L \otimes U(1)$, where the 3-parameter special unitary group describes the symmetry properties of *weak isospin* and the 1-parameter unitary group incorporates the gauge invariance generated by *weak hypercharge*. The leptons in the model are grouped *via* weak isospin and lepton flavor as

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$$

$$(e)_R \qquad (\mu)_R \qquad (\tau)_R \tag{2.1}$$

where the doublets have an isospin eigenvalue of $t = \frac{1}{2}$ and a hypercharge eigenvalue of $y = -1$ and the singlets have $t = 0$ and $y = -2$. The subscripts L and R denote left- and right-handed components of the fermion fields in the chiral representation. Maximal parity violation is evident from the absence of right-handed neutrinos. The gauge bosons, which give rise to the physical vector bosons, $W^{\pm}$, Z^0, and the photon, enter through the minimal coupling prescription in the covariant derivative for the gauge model. Because bare mass terms for the fermion fields would break gauge invariance, they cannot be included in the theory. However, adding a complex scalar doublet of elementary fields with explicit couplings to the fermion fields causes a *spontaneous* symmetry breakdown when the neutral component is fixed to a non-zero vacuum expectation value, a transformation which yields terms in the resulting Lagrangian that may be interpreted as mass terms for the fermion and gauge boson fields and leaves a single, physical neutral scalar field, the Higgs boson. Further technical details of the standard model may be found in the numerous reviews in the literature [Lon81,Sal72,Beg74].

The main prediction of the standard model pertinent to the search for $M \longrightarrow \overline{M}$ is summarized in Table 2.1. The $M \longrightarrow \overline{M}$ conversion is forbidden! By construction, the number of leptons from each family is required by the standard model to be

Additive Lepton Number Assignment									
	e^-	ν_e	e^+	$\overline{\nu_e}$	μ^-	ν_μ	μ^+	$\overline{\nu_\mu}$	all others
L_e	+1	+1	−1	−1	0	0	0	0	0
L_μ	0	0	0	0	+1	+1	−1	−1	0

Additive Lepton Number Conservation Laws:

$$\Sigma L_e = \text{const.}$$
$$\Sigma L_\mu = \text{const.}$$

Forbidden	Allowed
$\mu \longrightarrow e\,\gamma$	$\mu^+ \longrightarrow e^+\,\nu_e\,\overline{\nu_\mu}$
$\mu \longrightarrow e\,\gamma\,\gamma$	$\mu^+ \longrightarrow e^+\,\nu_e\,\overline{\nu_\mu}\,\gamma$
$\mu \longrightarrow 3e$	$\mu^+ \longrightarrow e^+\,\nu_e\,\overline{\nu_\mu}\,e^+\,e^-$
$\mu^-(Z,N) \longrightarrow e^-(Z,N)$	
$\mu^-(Z,N) \longrightarrow e^+(Z-2,N)$	
$\mu^+ \longrightarrow e^+\,\overline{\nu_e}\,\nu_\mu$	
$\mu^+e^- \longrightarrow \mu^-e^+$	
$e^-e^- \longrightarrow \mu^-\mu^-$	

Table 2.1: Additive lepton number conservation.

separately, additively conserved in any reaction. Thus, the experiment described in this dissertation is effectively a search for physics beyond the scope of the standard model. The $M \longrightarrow \overline{M}$ conversion violates the conservation of electron and of muon number by two units each.

2.2 Survey of Lepton Number Non-Conservation

To date, all experimental evidence supports the postulate that lepton family numbers are subject to an exact, separately additive conservation law in any physical process. There is, however, significant interest in improving the experimental limits on a possible violation of this symmetry, or, alternatively, in establishing the actual strength of lepton family number non-conserving couplings, should they turn out to occur at some level. Improving the observational sensitivity to rare muonic processes helps to confirm the standard model as the appropriate effective theory to higher energy scales. Actually observing such a rare process would reveal a breakdown of this effective theory and should provide insight into an underlying, more general structure.

The rare muonic processes relevant in this context are not all strictly free decays of the muon into forbidden channels. In addition to searches for $\mu \rightarrow e\gamma$, $\mu \rightarrow 3e$, $\mu \rightarrow e$ conversion on nuclei, and μ decay into e and forbidden neutrinos, the study of the $M \longrightarrow \overline{M}$ conversion provides yet another perspective in probing the limits of the standard model.

Since its discovery [And36,Ned37], the muon has motivated efforts to illuminate the principles of fundamental interactions among elementary particles. Although it was first mistaken as the proposed Yukawa meson mediating the strong interaction, its decay proved to be closely related to the process of nuclear beta-decay [Kon53].

Even before the identification of a muon neutrino, distinct from the electron neutrino, interest in the decay of the muon motivated conjectures on the possibility of the neutrinoless decay $\mu \rightarrow e\gamma$ and attempts to forbid this unobserved mode by proposing a conserved lepton number [Kon53,Fei61a]. With the discovery of the muon neutrino [Dan62], the introduction of the intermediate (gauge) bosons as the mediators of the weak interaction [Sch57,Fey58], and the advent of a unified description of the electromagnetic and weak interactions among leptons [Gla61,Wei67,Sal68], there seemed to be a natural scheme forbidding unobserved muon decay modes by introducing an additively conserved quantum number for each lepton family involved in an interaction.

But with the incorporation of the quark sector into the description of the weak interaction [Gla70], in the framework of today's standard model, new questions have been raised. Is the classification of quarks and leptons into left-handed doublets and right-handed singlets with respect to weak isospin artificial? Should a more fundamental description be left-right symmetric? Is there a leptonic analogue to the Cabbibo-Kobayashi-Maskawa mixing [Kob73] of weak interaction eigenstates of the quarks to give their mass eigenstates? Is there then also a leptonic version of GIM cancellation [Gla70] to ensure that the as yet unobserved muon decay modes are not predicted above current experimental limits? Further investigation of rare muon processes should, then, lead to improved understanding of the weak interaction and may reveal new physics.

2.2.1 Early Theoretical Work

The idea of an intermediate boson to carry the weak interaction dates back to at least 1957, to a remarkable paper by Schwinger [Sch57] that also addresses the concepts of an internal symmetry space of particles – later to be formalized in a *gauge group* – and of a neutral scalar field to generate masses for leptons and hadrons without disturbing their "internal symmetries" – the idea of a Higgs field whose non-vanishing vacuum expectation value generates particle masses while maintaining the gauge invariance of the Lagrangian density of the interacting fields involved. Further development of the concept that the weak interaction is carried by an "intermediate charged vector meson of high mass" was presented by Feynman and Gell-Mann [Fey58] and, in particular regard to the decay mode $\mu \rightarrow e\gamma$, by Feinberg [Fei58] in 1958. In Feinberg's approach, the $M \longrightarrow \overline{M}$ conversion would proceed as a second-order weak process, with a $\nu\,\overline{\nu}$ virtual intermediate state (see Fig. 2.1).

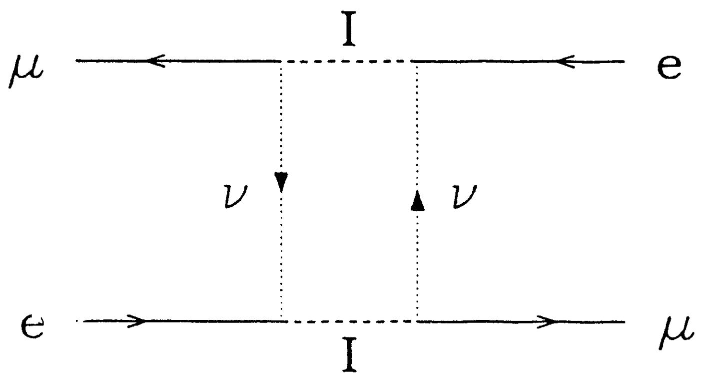


Figure 2.1: The $M \longrightarrow \overline{M}$ conversion in the Feinberg model.

But, since Feinberg found a tree-level contribution of 10^{-4} to the branching ratio for $\mu \rightarrow e\gamma$, at variance with the best experimental limit on this process then available [Lok55] of 2×10^{-5}, the intermediate boson hypothesis was thought to be incompatible with observations. Pontecorvo pointed out [Pon60] that two different types of neutrinos would rule out the process $\mu \rightarrow e\gamma$ via such an intermediate boson. The identification of the muon neutrino [Dan62] as distinct from the electron neutrino vindicated this viewpoint, but barred the $M \longrightarrow \overline{M}$ conversion from occurring as in Fig. 2.1.

The relevant conservation laws that follow from assuming that the electron and muon neutrinos remain distinct in all interactions can be formalized as in Table 2.1. These separate, additive conservation laws for each lepton family number may easily be extended to include further generations of leptons, *i.e.* the tau generation, τ and ν_τ, and are a direct prediction of the minimal standard model.

An alternative to this additive conservation scheme was proposed by Feinberg and Weinberg [Fei59,Fei61a] and by Cabbibo and Gatto [Cab60,Cab61]. Without deeper foundation in a gauge model, they postulated a multiplicatively conserved *muon parity* in addition to an additively conserved lepton number that distinguished leptons and their antiparticles. The relevant muon parity and lepton number assignments and their consequences are summarized in Table 2.2. The difficulty with this scheme is that it cannot easily be extended to include further lepton generations, but, since it allows $M \longrightarrow \overline{M}$ while forbidding $\mu \rightarrow e\gamma$, it directed attention to the $M \longrightarrow \overline{M}$ conversion.

In a subsequent paper [Fei61b], Feinberg and Weinberg more carefully considered the possibility of the spontaneous conversion of muonium to antimuonium, supposing an effective interaction Hamiltonian density of $V - A$ form, characteristic of low-energy weak processes:

$$\mathcal{H} = \frac{G_{M\overline{M}}}{\sqrt{2}} \overline{\psi_\mu} \gamma^\lambda (1 + \gamma^5) \psi_e \overline{\psi_\mu} \gamma_\lambda (1 + \gamma^5) \psi_e \; + \; H.c. \qquad (2.2)$$

The consequences of this assumption will be investigated in sections 2.4 and 2.5.

As the minimal standard model has been extremely successful in describing elementary particle interactions and in predicting effects that were later experimentally

Multiplicative Lepton Number Assignment									
	e^-	ν_e	e^+	$\overline{\nu_e}$	μ^-	ν_μ	μ^+	$\overline{\nu_\mu}$	all others
L	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1	0
P_μ	$+1$	$+1$	$+1$	$+1$	-1	-1	-1	-1	$+1$

Multiplicative Lepton Number Conservation Laws:

$$\Sigma L = \text{const.}$$
$$\Pi P_\mu = \text{const.}$$

Forbidden	Allowed
$\mu \longrightarrow e\gamma$	$\mu^+ \longrightarrow e^+ \nu_e \overline{\nu_\mu}$
$\mu \longrightarrow e\gamma\gamma$	$\mu^+ \longrightarrow e^+ \overline{\nu_e} \nu_\mu$
$\mu \longrightarrow 3e$	$\mu^+ e^- \longrightarrow \mu^- e^+$
$\mu^-(Z,N) \longrightarrow e^-(Z,N)$	$e^- e^- \longrightarrow \mu^- \mu^-$
$\mu^-(Z,N) \longrightarrow e^+(Z-2,N)$	

Table 2.2: Multiplicative lepton number conservation.

confirmed, it is now widely believed that any extension to it should reduce to this model in the appropriate limits. Rumors of the observation of $\mu \rightarrow e\gamma$ at SIN in 1977 motivated a flurry of papers presenting models in this class that might allow this process at an adequately suppressed level.

In order to allow for small but nonvanishing lepton number non-conserving currents, it is necessary to arrange for the leptonic mass eigenstates of a theory to differ from the eigenstates of weak isospin. This requires that lepton number violating couplings be introduced into the theory to give small mixings between the weak eigenstates of the leptons. Such couplings may be added by extending the standard model with, for example, massive neutrinos [Che77a,Wil77], additional Higgs fields [Bjo77], or heavy leptons [Che77b,Lee77,Mar77].

One class of left-right isospin symmetric theories is particularly appealing and will be discussed in more detail in the next section.

2.3 Mechanisms Mediating $M \longrightarrow \overline{M}$

The M to $\overline{M}$ conversion is allowed in many extensions of the standard model [Ver86], among which left-right symmetric theories [Pat74,Moh75b,Moh75a] are an appealing class. Of particular current interest is the left-right symmetric extension of Mohapatra and Senjanović [Moh81a,Moh81b,Moh80]. They propose an additional Higgs triplet as the minimal extension for achieving left-right symmetry. In this model, parity violation becomes a low-energy symmetry breaking effect by giving appropriate vacuum expectation values (VEV) to the extended Higgs sector. This choice of VEVs gives the right-handed gauge boson a much larger mass than the left-handed one, thus suppressing right-handed currents. Another effect is the appearance of massive Majorana neutrinos.

The new Higgs bosons have couplings to leptons that can induce lepton-number violating processes. The doubly charged member of this triplet, Δ^{++}, mediates the M to $\overline{M}$ conversion as shown in Fig. 2.2. For a scalar coupling of this sort, the M

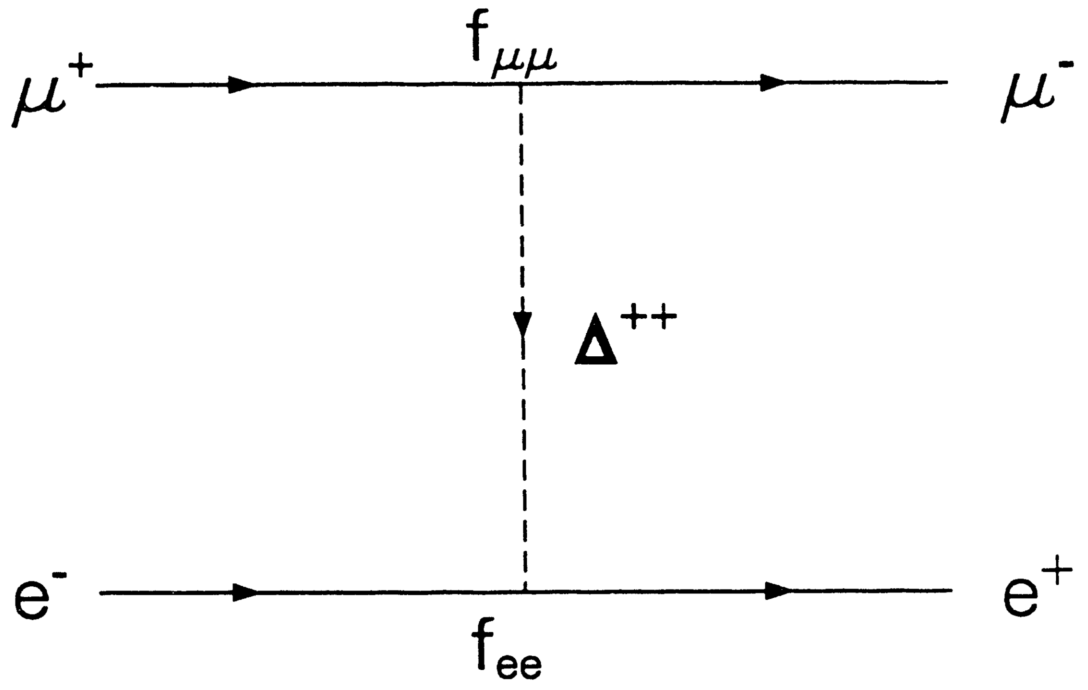


Figure 2.2: Diagram of the $M \rightarrow \overline{M}$ conversion mediated by a doubly charged Higgs boson.

to $\overline{M}$ effective low-energy conversion Hamiltonian has the form [Her89]

$$\mathcal{H}_{M\overline{M}} = \frac{f_{ee} f^*_{\mu\mu}}{4m^2_{++}} \; \overline{e^c}(1 - \gamma_5)e \; \overline{\mu}(1 + \gamma_5)\mu^c \quad + \quad H.c. \quad , \tag{2.3}$$

where f_{ee} and $f_{\mu\mu}$ are Yukawa couplings and m_{++} is the mass of the mediating Δ^{++}. A Fierz transformation on this Hamiltonian can be shown [Her89,Swa89] to give a Hamiltonian of the form of Eq. 2.2 if one identifies the effective coupling constant as

$$G_{M\overline{M}} \equiv \sqrt{2} \; \frac{f_{ee} f^*_{\mu\mu}}{8m^2_{++}} . \tag{2.4}$$

Thus, all experimental results quoted as limits on the coupling constant of Eq. 2.2 are also valid as limits on the effective coupling in the case of M to $\overline{M}$ conversion proceeding via Δ^{++} exchange.

As a result of astrophysical constraints on this model and assuming that the muon neutrino is no heavier than 65 eV, it places a *lower* limit on the conversion of M to $\overline{M}$ of $G_{M\overline{M}} > 4 \times 10^{-4}\ G_F$ [Her]. The upper limit within this model is given simply by the present experimental limit on $G_{M\overline{M}}$.

The massive Majorana neutrinos may also mediate an $M \longrightarrow \overline{M}$ conversion, as depicted in Fig. 2.3. This diagram has, however, been estimated [Hal82] to give a contribution no larger than $G_{M\overline{M}} \sim 10^{-5}\ G_F$, using experimental limits from searches for neutrinoless double-beta decay.

Though it has been ruled out by its predictions of semi-leptonic, lepton-number violating decays of the B meson, the 1979 model of Derman [Der79] was interesting, as it allowed $M \longrightarrow \overline{M}$ to proceed *via* exchange of a neutral Higgs boson. The relevant Higgs entered the theory as a Goldstone boson after breaking of the group symmetry of $SU(2)_L \otimes U(1) \otimes S_3$, where the permutation group operates on lepton flavor. This model inspired the calculation of the neutral scalar $M \longrightarrow \overline{M}$ matrix element given in Appendix A.

As there is active speculation about small, but nonvanishing neutrino mass in the context of the problem of "missing" solar neutrinos, the possibility of the $M \longrightarrow \overline{M}$ conversion proceeding by massive Dirac neutrinos must be mentioned. The relevant diagram is shown in Fig. 2.4. This process has been considered by Swartz in analogy

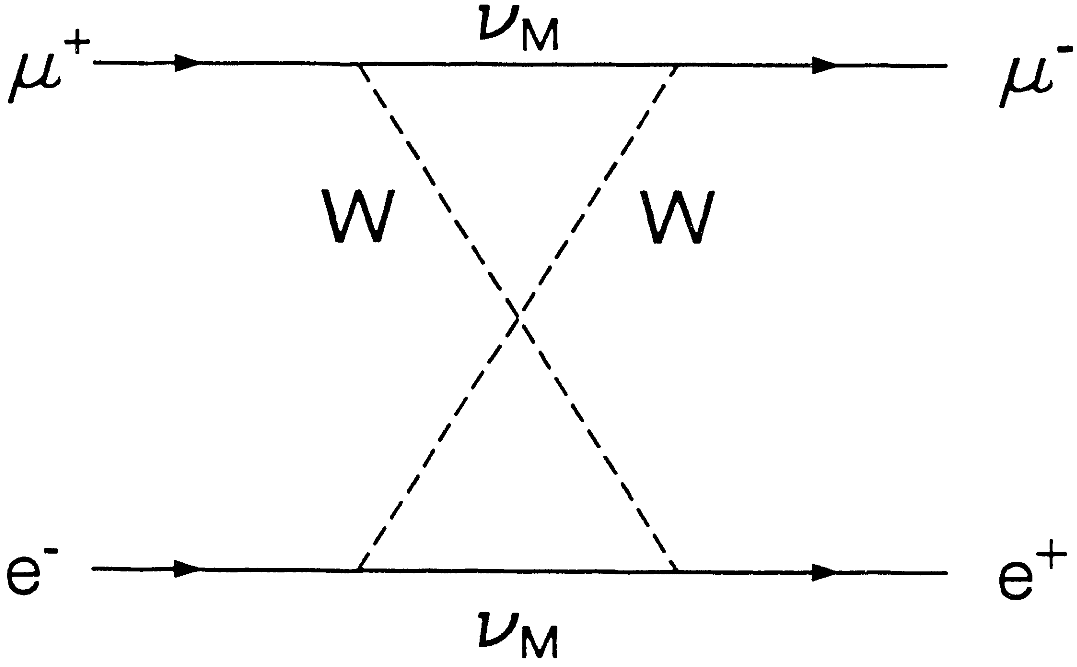


Figure 2.3: The $M \longrightarrow \overline{M}$ conversion mediated by massive Majorana neutrinos.

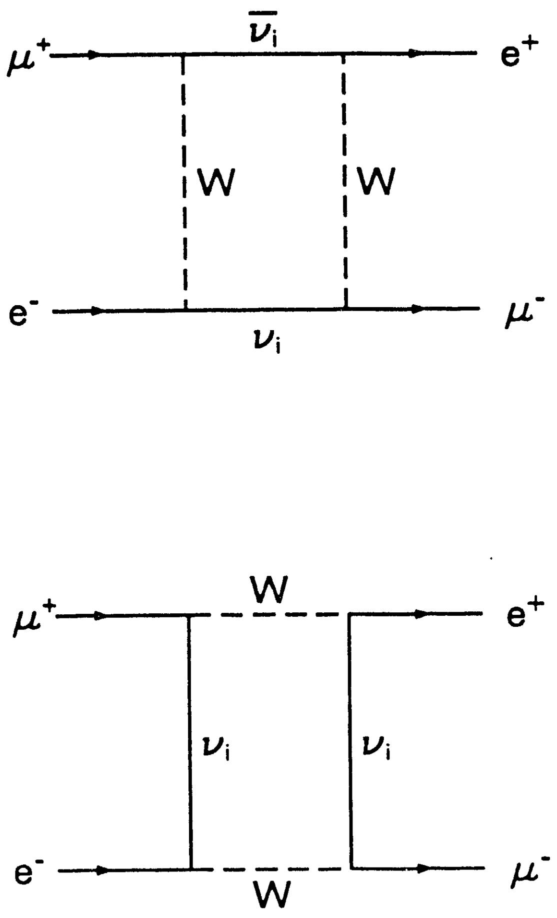


Figure 2.4: The $M \longrightarrow \overline{M}$ conversion mediated by massive Dirac neutrinos.

to $B^0 - \overline{B}^0$ mixing [Swa89]. He gives the effective Hamiltonian as

$$\begin{aligned}\mathcal{H}_{M\overline{M}} &= \frac{G_A}{\sqrt{2}}\overline{\psi}_\mu\gamma^\alpha(1+\gamma_5)\psi_e\overline{\psi}_\mu\gamma_\alpha(1+\gamma_5)\psi_e \\ &\quad + \frac{G_B}{\sqrt{2}}\overline{\psi}_\mu(1-\gamma_5)\psi_e\overline{\psi}_\mu(1-\gamma_5)\psi_e \;+\; H.c. \;, \qquad (2.5)\end{aligned}$$

where G_A and G_B are functions of lepton masses, neutrino masses, and mixing angles. Swartz estimates a conservative upper bound on the coupling constants at

$$G_A, G_B < \frac{G_F\sqrt{2}}{8\pi^2}m_\mu^2 = 7\times 10^{-9}\,G_F\,. \qquad (2.6)$$

This is far below the present limits of observability.

2.4 The Atomic Physics of $M \longrightarrow \overline{M}$

The matrix elements of the $M \longrightarrow \overline{M}$ conversion place specific conditions on the spin projections of the particles involved in the interaction by way of the coupling operator's action on the helicities of the particles involved. The significance of this will not be fully apparent until the effects of external fields on the initial and final states are carefully investigated. Under the assumption of a $(V - A)$ conversion coupling, Appendix B gives the relevant derivations of the $n = 1$ M and $\overline{M}$ Breit-Rabi levels, the eigenstates of the M, $\overline{M}$ system coupled by the $M \longrightarrow \overline{M}$ conversion, and the probabilities of finding the system in an $\overline{M}$ state after preparing it in an initial state of M. The resulting level diagrams for M and $\overline{M}$ are shown in Figs. 2.5 and 2.6, respectively.

The spatial structure of the coupled system is determined by the kinetic energy and central-field Coulomb terms in the Hamiltonian, in complete analogy to hydrogen. The reduced mass of the muonium atom differs by less than 0.5% from that of hydrogen, so the spatial wave function of the M, $\overline{M}$ system is essentially that of hydrogen. The electron and muon spins couple by the Fermi contact interaction, giving rise to the hyperfine splitting of $a = 1.85\times 10^{-5}eV$, and these spins interact separately *via* the Zeeman effect with any external magnetic field. Because the system is assumed to be in the $n = 1$ spatial state, there is no spin-orbit coupling

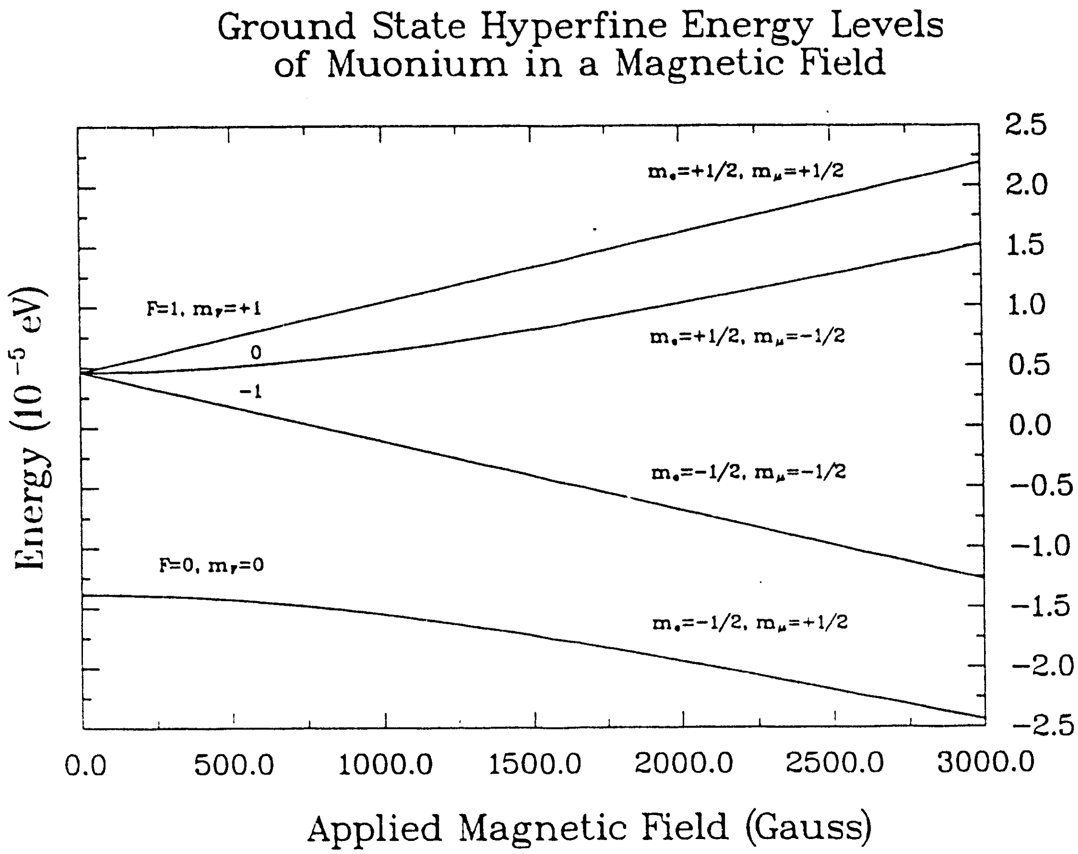


Figure 2.5: Energy levels of ground state muonium in an external magnetic field.

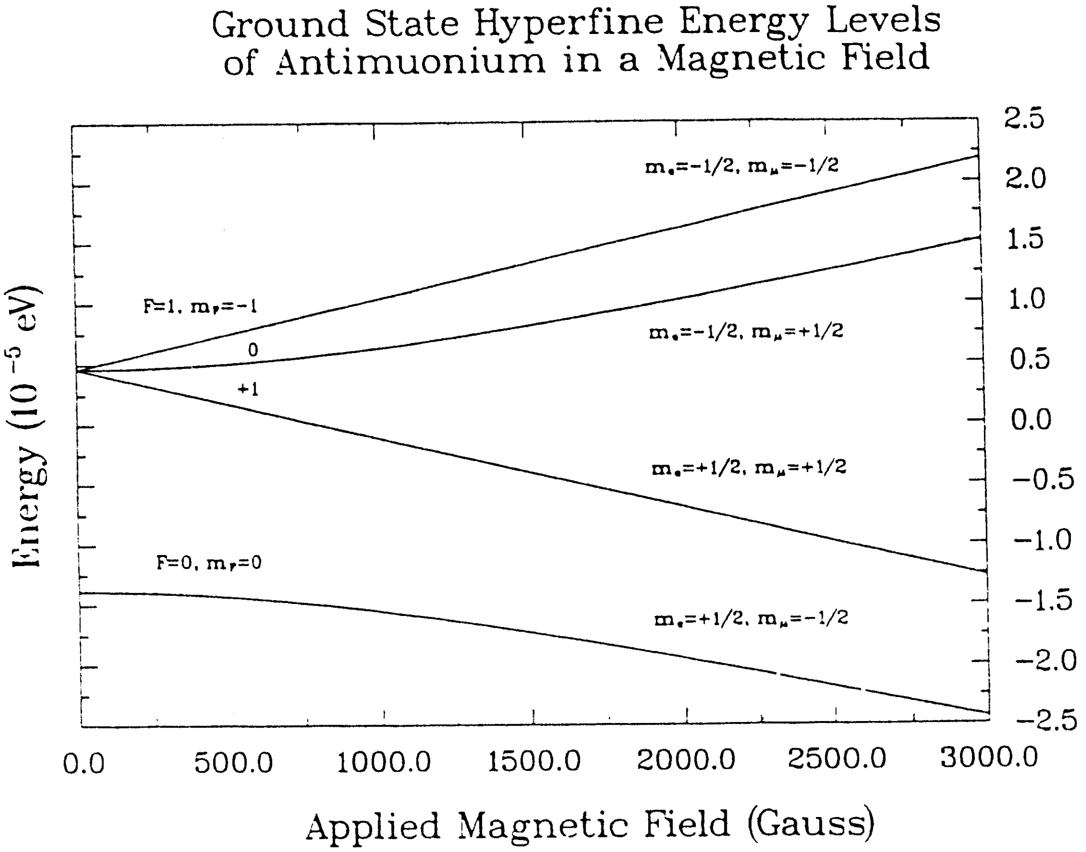


Figure 2.6: Energy levels of ground state antimuonium in an external magnetic field.

to consider. Finally, the conversion Hamiltonian can change the charge state of the electron and muon from that in M to that in $\overline{M}$ or *vice versa*, but only with proper regard to conserving angular momentum by its specific sensitivity to the spin states of the reacting particles. To solve for the eigenstates and -energies of the coupled M, $\overline{M}$ system, one diagonalizes the full Hamiltonian matrix, most conveniently in the basis of uncoupled spins for M and $\overline{M}$. The approximations made are to work in the non-relativistic limit and to use a truncated basis set, one that includes only the $n = 1$ states of the system. Since the energy difference to the first excited level in M, $\overline{M}$ is very large compared to the effects listed above, this is believed to be a valid simplification. One finds, interestingly, that the energy levels in the coupled system have been shifted and split by the conversion coupling (see Fig. 2.7). The value of the splitting in the unpolarized levels is given by twice the magnitude of the $M \longrightarrow \overline{M}$ conversion matrix element. One might consider the precision spectroscopy of the hyperfine interval of the ground state or of the $1S - 2S$ interval as an alternative way to observe the effects of an $M \longrightarrow \overline{M}$ mixing. This is not thought to be a promising approach, however, since the development with time of the $\overline{M}$ admixture in the wave function of the system during the interaction time with the radiation field leads to a time-dependent transition frequency. As the splitting of the levels of the M, $\overline{M}$ system is already known to be small compared to the decay width of the levels, this complicates the interpretation of a measured transition frequency for the extraction of a limit on the $M \longrightarrow \overline{M}$ coupling.

Because the leptons constituting the $\overline{M}$ atom are the antiparticles of those making up M, the splitting of the hyperfine levels in the ground states of M and $\overline{M}$ under the influence of an external magnetic field will be relatively opposite in energy. This lifts the zero-field degeneracy of the levels participating in the $M \longrightarrow \overline{M}$ conversion and will increasingly hinder the conversion with rising field values. For the case of a coupling of $(V - A)$ form, a simple interpretation will be given of the magnetic field value that suppresses the conversion probability between the polarized levels of M and $\overline{M}$ ($|\lambda_1\rangle$ and $|\lambda_1\rangle$; see Appendix B) by 50% and that which suppresses the conversion probability between the unpolarized levels of M and $\overline{M}$ ($|\lambda_2\rangle$ and $|\lambda_3\rangle$) by 50%. The derivations of some of the equations used are given in

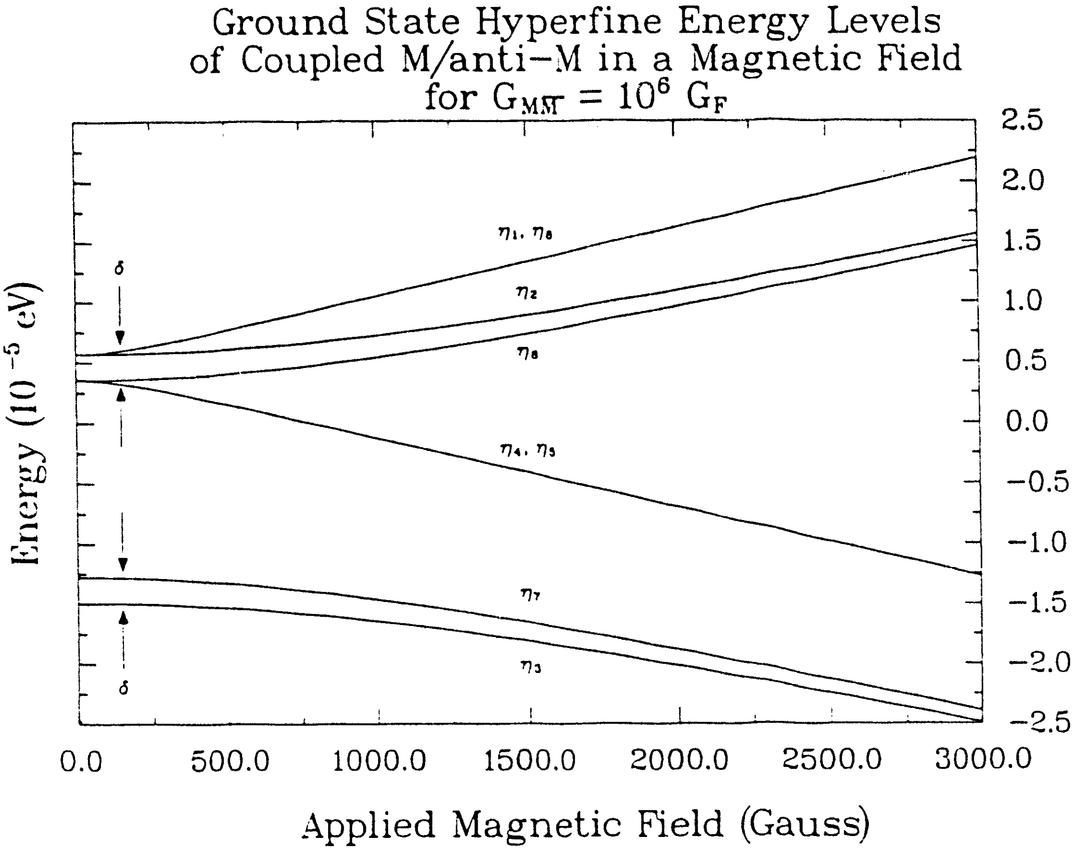


Figure 2.7: Energy levels of the coupled M, $\overline{M}$ system in an external magnetic field. The conversion parameter δ has been strongly exaggerated for effect.

Sec. B.4.

2.4.1 Conversion Between Polarized Levels

For the conversion probability between the polarized states of M and $\overline{M}$, integrated over all observation times of the decay of the system, one has the expression

$$S'_{\overline{M}} = \frac{2\left(\frac{\delta}{2}\right)^2}{\hbar^2\gamma^2 + a^2\left[\left(\frac{\delta}{a}\right)^2 + Y'^2\right]} \tag{2.7}$$

from the first term of Eq. B.85 (multiplied by 2), where

$$Y' = \frac{\mu_B B}{a}\left(g_e - \frac{m_e}{m_\mu}g_\mu\right) , \tag{2.8}$$

γ is the muon decay rate, a is the ground state hyperfine interval, μ_B is the Bohr magneton, g_e and g_μ are the absolute values of the g-factors, and m_e and m_μ are the masses of the electron and the muon, respectively. The expression for $S'_{\overline{M}}$ shows that the conversion is unsuppressed at zero field and that any applied field will diminish the conversion probability. Thus, we seek a solution to the rescaled field $Y' = Y'_0$ that reduces $S'_{\overline{M}}(Y')$ to half of its zero-field value:

$$S'_{\overline{M}}(Y'_0) = \frac{1}{2}\, S'_{\overline{M}}(Y' = 0) . \tag{2.9}$$

The result is

$$Y'_0 = \frac{\hbar\gamma}{a}\sqrt{1 + \left(\frac{\delta}{\hbar\gamma}\right)^2}, \tag{2.10}$$

which leads to

$$B'_0 = \frac{\hbar\gamma\sqrt{1 + \left(\frac{\delta}{\hbar\gamma}\right)^2}}{\mu_B\left(g_e - \frac{m_e}{m_\mu}g_\mu\right)} \tag{2.11}$$

that can be evaluated to give $B'_0 = 26\ mG$. It is appropriate to make a few approximations:

$$\frac{m_e}{m_\mu} \approx \frac{1}{206} \approx 0 , \tag{2.12}$$

$$g_e \approx g_\mu \approx 2 , \tag{2.13}$$

and, assuming $0 \leq G_{M\overline{M}} \leq G_F$,

$$1 \leq \sqrt{1 + \left(\frac{\delta}{\hbar\gamma}\right)^2} \leq 1.0000245 \tag{2.14}$$

so that

$$\sqrt{1 + \left(\frac{\delta}{\hbar\gamma}\right)^2} \approx 1 \ , \tag{2.15}$$

and

$$Y_0' \approx \frac{\hbar\gamma}{a} \ . \tag{2.16}$$

This simplifies the expression for B_0':

$$B_0' \approx \frac{\hbar\gamma}{2\mu_B} \ . \tag{2.17}$$

To interpret this result physically, we note that the levels of M and $\overline{M}$ can be considered to be spread by the finite lifetime of the system, which is governed by the muon decay rate. So the conversion between the polarized states begins to be inhibited when the degeneracy breaking between the levels of M and $\overline{M}$ due to an applied magnetic field becomes comparable to the natural level width, as depicted in Fig. 2.8. The overlap area of the two Lorentzians shown can be interpreted as a relative measure of the conversion probability with respect to the zero-field conversion probability, where this overlap area is unity for normalized Lorentzians. To complete the argument, consider a Lorentzian of the form

$$L(x) = \frac{1}{\pi} \frac{\left(\frac{\gamma}{2}\right)}{x^2 + \left(\frac{\gamma}{2}\right)^2} , \tag{2.18}$$

which is normalized according to

$$\int_{-\infty}^{+\infty} L(x)\, dx = 1 \ . \tag{2.19}$$

Figure 2.9 shows that γ is the FWHM. Referring to the case indicated in Fig. 2.8, it is clear that the overlap area of the two Lorentzians is given by

$$A = \int_{-\infty}^{-\frac{\gamma}{2}} L(x)\, dx + \int_{+\frac{\gamma}{2}}^{+\infty} L(x)\, dx = 1 - \int_{-\frac{\gamma}{2}}^{+\frac{\gamma}{2}} L(x)\, dx , \tag{2.20}$$

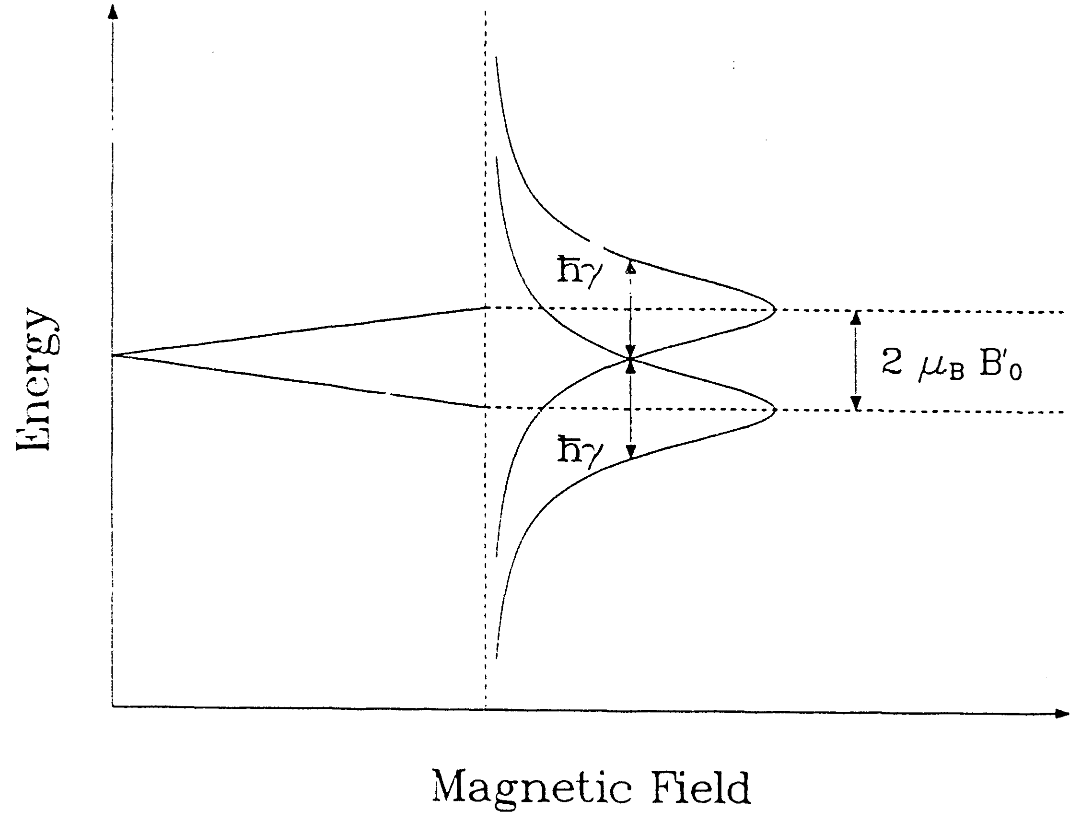


Figure 2.8: Level diagram of polarized states versus external magnetic field with a Lorentzian spread due to the finite lifetime of the system.

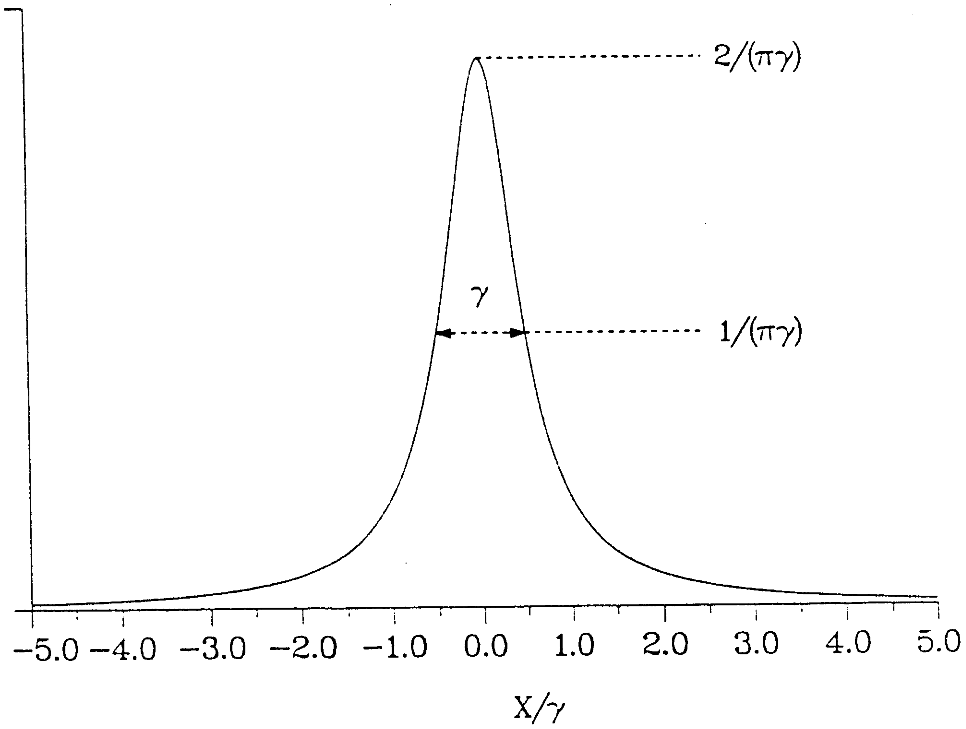


Figure 2.9: Lorentzian distribution with area normalized to unity.

and A turns out to be

$$A = 1 - \frac{\gamma}{2\pi}\int_{-\frac{\gamma}{2}}^{+\frac{\gamma}{2}} \frac{dx}{x^2 + \left(\frac{\gamma}{2}\right)^2} = 1 - \frac{\gamma}{2\pi}\left[\frac{2}{\gamma}\,\mathrm{Atan}\left(\frac{2x}{\gamma}\right)\right]_{-\frac{\gamma}{2}}^{+\frac{\gamma}{2}} = \frac{1}{2}\ . \tag{2.21}$$

Thus, if the degeneracy splitting between the M and the $\overline{M}$ levels involved is such that

$$2\mu_B B_0' = \hbar\gamma, \tag{2.22}$$

then the overlap of the levels is reduced by 50% and the conversion probability is reduced to half of its zero-field value.

2.4.2 Conversion Between Unpolarized Levels

After applying the approximation that $\frac{\delta}{a} \ll 1$, it is possible to express the conversion probability between the unpolarized states of M and $\overline{M}$, again integrated over all decay observation times, as in the second term of Eq. B.87 (multiplied by 2):

$$S_{\overline{M}} = \frac{\left(\frac{\delta}{\hbar\gamma}\right)^2 + \left(\frac{\delta}{a}\right)^2}{2\left[\left(\frac{\hbar\gamma}{a}\right)^2 + (1 + X^2)\right]} \tag{2.23}$$

where

$$X = \frac{\mu_B B}{a}\left(g_e + \frac{m_e}{m_\mu} g_\mu\right)\ . \tag{2.24}$$

Again, the conversion probability is unsuppressed at zero field and decreases with increasing applied field. Here, too, we want a solution to

$$S_{\overline{M}}(X_0) = \frac{1}{2}\, S_{\overline{M}}(X = 0)\ . \tag{2.25}$$

The result is

$$X_0 = \sqrt{1 + \left(\frac{\hbar\gamma}{a}\right)^2} \approx 1\ , \tag{2.26}$$

or

$$B_0 = \frac{a\sqrt{1 + \left(\frac{\hbar\gamma}{a}\right)^2}}{\mu_B\left(g_e + \frac{m_e}{m_\mu} g_\mu\right)}\ , \tag{2.27}$$

which gives $B_0 \approx 1.6\ kG$. With the same approximations as made earlier for the first case and the additional realization that

$$\frac{\hbar\gamma}{a} = 1.6 \times 10^{-5} \ll 1 \ , \tag{2.28}$$

we get a simpler expression for B_0:

$$B_0 = \frac{a}{2\mu_B} \ . \tag{2.29}$$

This result is a little trickier to interpret. The general, magnetic field dependent Breit-Rabi eigenstates of M and of $\overline{M}$ (in the absence of a conversion) can be expressed as a superposition of the uncoupled basis states, with the admixture of each being field dependent. Designating the middle level in the triplet by T and the singlet level by S we can write:

$$\begin{aligned}
|M;T\rangle &= \frac{1}{N}\left[\left(X+\sqrt{1+X^2}\right)|M;\uparrow\downarrow\rangle + |M;\downarrow\uparrow\rangle\right] \\
|M;S\rangle &= \frac{1}{N}\left[|M;\uparrow\downarrow\rangle - \left(X+\sqrt{1+X^2}\right)|M;\downarrow\uparrow\rangle\right] \\
|\overline{M};T\rangle &= \frac{1}{N}\left[|\overline{M};\uparrow\downarrow\rangle + \left(X+\sqrt{1+X^2}\right)|\overline{M};\downarrow\uparrow\rangle\right] \\
|\overline{M};S\rangle &= \frac{1}{N}\left[-\left(X+\sqrt{1+X^2}\right)|\overline{M};\uparrow\downarrow\rangle + |\overline{M};\downarrow\uparrow\rangle\right] \ .
\end{aligned} \tag{2.30}$$

Here, the spin projections in the uncoupled basis are represented by ↑ for $+\frac{1}{2}$ and by ↓ for $-\frac{1}{2}$. The first spin is that of the electron and the second is that of the muon. The charge and spatial states of the particles are indicated by M or by $\overline{M}$. The normalization, N, is given by

$$N = \left[1 + \left(X+\sqrt{1+X^2}\right)^2\right]^{\frac{1}{2}} \ . \tag{2.31}$$

It is easy to check that the limits of these expressions are correct. For $X \rightarrow 0$, that is, as $B \rightarrow 0$:

$$\begin{aligned}
|M;T\rangle &\rightarrow \frac{1}{\sqrt{2}}\left[\,|M;\uparrow\downarrow\rangle + |M;\downarrow\uparrow\rangle\,\right] = |M;F=1,m_F=0\rangle \\
|M;S\rangle &\rightarrow \frac{1}{\sqrt{2}}\left[\,|M;\uparrow\downarrow\rangle - |M;\downarrow\uparrow\rangle\,\right] = |M;F=0,m_F=0\rangle \\
|\overline{M};T\rangle &\rightarrow \frac{1}{\sqrt{2}}\left[|\overline{M};\uparrow\downarrow\rangle + |\overline{M};\downarrow\uparrow\rangle\right] = |\overline{M};F=1,m_F=0\rangle \\
|\overline{M};S\rangle &\rightarrow \frac{1}{\sqrt{2}}\left[-|\overline{M};\uparrow\downarrow\rangle + |\overline{M};\downarrow\uparrow\rangle\right] = |\overline{M};F=0,m_F=0\rangle \ .
\end{aligned} \tag{2.32}$$

For $X \to \infty$, that is, as $B \to \infty$:

$$\begin{aligned} |M;T\rangle &\to |M;\uparrow\downarrow\rangle \\ |M;S\rangle &\to -|M;\downarrow\uparrow\rangle \\ |\overline{M};T\rangle &\to |\overline{M};\downarrow\uparrow\rangle \\ |\overline{M};S\rangle &\to -|\overline{M};\uparrow\downarrow\rangle \ . \end{aligned} \tag{2.33}$$

More rigorously, this limit is valid for the range

$$1 \ll X \ll \frac{1}{2a}\alpha^2 m_e c^2 \ , \tag{2.34}$$

which corresponds to about

$$10^3 \ G \ll B \ll 10^9 \ G \ , \tag{2.35}$$

so that one ensures the perturbative nature of the interaction with the external field. The case of most interest here, however, is $X = 1$:

$$\begin{aligned} |M;T\rangle &= \frac{1}{\sqrt{2}\sqrt{2+\sqrt{2}}}\left[\left(1+\sqrt{2}\right)|M;\uparrow\downarrow\rangle + |M;\downarrow\uparrow\rangle\right] \\ |M;S\rangle &= \frac{1}{\sqrt{2}\sqrt{2+\sqrt{2}}}\left[|M;\uparrow\downarrow\rangle - \left(1+\sqrt{2}\right)|M;\downarrow\uparrow\rangle\right] \\ |\overline{M};T\rangle &= \frac{1}{\sqrt{2}\sqrt{2+\sqrt{2}}}\left[|\overline{M};\uparrow\downarrow\rangle + \left(1+\sqrt{2}\right)|\overline{M};\downarrow\uparrow\rangle\right] \\ |\overline{M};S\rangle &= \frac{1}{\sqrt{2}\sqrt{2+\sqrt{2}}}\left[-\left(1+\sqrt{2}\right)|\overline{M};\uparrow\downarrow\rangle + |\overline{M};\downarrow\uparrow\rangle\right] \ . \end{aligned} \tag{2.36}$$

It is evident that, for general X,

$$\langle\overline{M};S|H_{M\overline{M}}|M;T\rangle = \left(\frac{\delta}{2}\right)\frac{1-\left(X+\sqrt{1+X^2}\right)^2}{1+\left(X+\sqrt{1+X^2}\right)^2} \tag{2.37}$$

and

$$\langle\overline{M};T|H_{M\overline{M}}|M;S\rangle = \left(\frac{\delta}{2}\right)\frac{1-\left(X+\sqrt{1+X^2}\right)^2}{1+\left(X+\sqrt{1+X^2}\right)^2} \ , \tag{2.38}$$

but one expects that such transitions will be strongly suppressed by the large energy difference between the bra and ket states. This energy difference is at least as large as the hyperfine interval, a, a spacing that is much larger than the natural level width. In fact, $\frac{\hbar\gamma}{a} = 1.6 \times 10^{-5}$. So the matrix elements to consider are

$$\langle\overline{M};T|H_{M\overline{M}}|M;T\rangle = \left(\frac{\delta}{2}\right)\frac{2\left(X+\sqrt{1+X^2}\right)}{1+\left(X+\sqrt{1+X^2}\right)^2} \tag{2.39}$$

and

$$\langle\overline{M};S|H_{M\overline{M}}|M;S\rangle = \left(\frac{\delta}{2}\right)\frac{-2\left(X+\sqrt{1+X^2}\right)}{1+\left(X+\sqrt{1+X^2}\right)^2}\ . \tag{2.40}$$

The cases of interest are:

$$\begin{aligned}\langle\overline{M};T|H_{M\overline{M}}|M;T\rangle &\longrightarrow \left(\frac{\delta}{2}\right) \text{ as } X \longrightarrow 0\\ &\longrightarrow 0 \text{ as } X \longrightarrow \infty\\ &= \frac{1}{\sqrt{2}}\left(\frac{\delta}{2}\right) \text{ for } X = 1\end{aligned} \tag{2.41}$$

and

$$\begin{aligned}\langle\overline{M};S|H_{M\overline{M}}|M;S\rangle &\longrightarrow -\left(\frac{\delta}{2}\right) \text{ as } X \longrightarrow 0\\ &\longrightarrow 0 \text{ as } X \longrightarrow \infty\\ &= -\frac{1}{\sqrt{2}}\left(\frac{\delta}{2}\right) \text{ for } X = 1\ .\end{aligned} \tag{2.42}$$

In words, for a value of the rescaled field of $X = 1$, the admixture of the spin state in the $\overline{M}$ final state that mates with the spin state of the initial state of M has dropped to $\frac{1}{\sqrt{2}}$ of its zero-field amplitude. This means that the conversion probability, which scales as the square of this matrix element, has dropped to $\frac{1}{2}$ of its zero-field value for $X = 1$. The external magnetic field has begun to uncouple the spins which at zero field may be imagined to precess around one another. The individual spins are partially polarized in the external field and the magnetic moment of the electron begins to dominate over that of the muon in determining the orientation of the atom. It is the unpolarized part of the T and S states that can participate in the $M \longrightarrow \overline{M}$ conversion.

2.4.3 Field Dependence of Scalar Coupling

The action of the conversion Hamiltonian on the spins, as discussed above, is a direct consequence of the given $(V - A)$ matrix element of the conversion Hamiltonian. It seems reasonable that there be at least one other possibility, which may arise from a neutral scalar coupling, as discussed in Appendix A:

$$\langle \overline{M}; m'_{S_e}, m'_{S_\mu} | H_{M\overline{M}} | M; m_{S_e}, m_{S_\mu} \rangle = \left(\frac{\delta}{2} \right) \delta_{m'_{S_e}, m_{S_\mu}} \delta_{m'_{S_\mu}, m_{S_e}} \ , \tag{2.43}$$

where translating the helicity δ-functions to conditions on the spin quantum numbers is the most difficult point. This matrix element would entail the same low-field dependence of the conversion probability as above, but would give no suppression of the conversion at high fields at all. The conversion probability would never fall below 50% of its zero-field value. This is very interesting! It seems that, if a conversion were ever observed, one would check its strength at low and high values of the magnetic field to see whether or not the field affects it. This might, then, determine the rank of the operator involved in the conversion.

In order to give a complete treatment, one should, in principle, calculate the conversion for an arbitrary linear combination of scalar, pseudoscalar, vector, axial vector, and tensor bilinear forms of a low energy effective conversion coupling. The important conclusion, however, is that statements about the $M \longrightarrow \overline{M}$ coupling constant (limits, etc.) as well as the specific behavior of the conversion suppression in a magnetic field are model-dependent predictions.

2.5 The Connection to Experiment

The influence of an external magnetic field on the $M \longrightarrow \overline{M}$ conversion was discussed at length above, since magnetic fields have the greatest aptitude for suppressing the coupling. For a static, uniform electric field applied to the 1S ground states of M and $\overline{M}$, the Stark effect lifts their degeneracy in energy in third order, and a simple estimate indicates that fields on the order of 10^6 V/cm are required to split the M and $\overline{M}$ levels by an amount comparable to the decay width of the levels.

Due to the different reaction channels available to M and $\overline{M}$ atoms in collisions with the atoms of a host medium, the $M \longrightarrow \overline{M}$ conversion is strongly inhibited in a material environment [Mor73,Mor70,Mor66,Fei61b]. For M formed in gases, the conversion is suppressed by a factor of 10^5 per atmosphere of gas pressure [Ama68b]. It is, therefore, very advantageous to search for $M \longrightarrow \overline{M}$ in a vacuum environment, where the process may occur uninhibited by anything but residual magnetic fields, which are the most difficult to control.

There are two established methods for producing M in vacuum. The beam-foil method [Bol81,Bol82] has been successfully used in many muonium experiments [Woo90,Kua87,Kua89,Ket90a,Ahn90] including a search for $M \longrightarrow \overline{M}$ [Ni87,Ni88a]. As M produced by this technique has kinetic energies of several tens of keV, searching for an $\overline{M}$ component with this source of M necessarily implies a small solid angle of detection. That is where the production of thermal M from SiO_2 powders [Bee86,Woo88] offers a distinct advantage. Muonium at thermal energies with $T \approx 300K$ remains confined within a space of only a few cm extent throughout its entire lifetime. This not only enables an increase in detection acceptance, it also allows, for the first time, the observation of the breakup of the M, $\overline{M}$ system in vacuum at *any* time after its formation. By contrast, all searches for $M \longrightarrow \overline{M}$ in vacuum to date relied on the $\overline{M}$ atoms striking a specially chosen material to induce the relevant signature. This experiment is not the first to search for $M \longrightarrow \overline{M}$ using thermal M [Hub90], but it is the only one so far that searches for a decay from the $\overline{M}$ component over the entire natural lifetime of the M, $\overline{M}$ system.

To relate the theoretical probabilities for $M \longrightarrow \overline{M}$ discussed above to the quantities measured in this experiment, we recall that the definition of $S_{\overline{M}}$ was the probability of observing the system, initially prepared as M, decay from the $\overline{M}$ state. The experiment yields the number of M decays, N_M, and the number of $\overline{M}$ decays seen, $N_{\overline{M}}$, corrected to equal acceptance. The probability of the $M \longrightarrow \overline{M}$ conversion is, then, the number of $\overline{M}$ decays divided by the total number of atoms decaying from either the M or the $\overline{M}$ state:

$$S_{\overline{M}} = \frac{N_{\overline{M}}}{N_M + N_{\overline{M}}} \ . \tag{2.44}$$

Since the $M \longrightarrow \overline{M}$ reaction is known to be at best very weak, the condition $N_{\overline{M}} \ll N_M$ will hold. Thus, one may write

$$S_{\overline{M}} = \frac{N_{\overline{M}}}{N_M} \tag{2.45}$$

as an excellent approximation. Giving the results of an experiment searching for $M \longrightarrow \overline{M}$ as a value or an upper limit on $S_{\overline{M}}$ is, then, a model-independent way of stating the result. This should be the number used in comparing the limits on this process to those on other lepton-number violating decays to get a feeling for the relative "sensitivities" achieved in different rare decay searches.

In the case of this experiment, any $M \longrightarrow \overline{M}$ conversion suffered a 50% suppression due the external field of around 10 G. This, in effect, means that only half of the observed M atoms, those in an unpolarized state, were available for a conversion coupling to act upon. This field suppression has not been shown to be model-independent, so one must allow now that it may be. It is conjectured that any model-dependence enters only in the high-field behavior of the $M \longrightarrow \overline{M}$ probability, in which case the present experiment could still be interpreted in a model-independent way. The above expression must be modified within this context to read

$$S_{\overline{M}} \approx 2\,\frac{N_{\overline{M}}}{N_M}\ , \tag{2.46}$$

if we are to use the zero-field form of the conversion probability, $S_{\overline{M}}$ as given in Eq. B.90, to obtain a result for the $M \longrightarrow \overline{M}$ coupling constant. Into this equation we will insert the measured number of M atoms observable during this search for $M \longrightarrow \overline{M}$, $N_{\overline{M}}$ will be the taken as the most probable number of $\overline{M}$ events seen or as an upper limit thereon, and the factor of 2 accounts for the 50% suppression of the conversion due to the ambient field in the apparatus.

Alternatively, one may use Eq. 2.45 and include the magnetic field dependence in the theoretical half of the consideration. To relate the observed $S_{\overline{M}}$ to a value for the $G_{M\overline{M}}$ coupling constant as it appears in the $(V - A)$ Hamiltonian in Eq. 2.2, we refer to Eq. B.87,

$$S_{\overline{M}}(\infty) = \frac{\left(\frac{\delta}{a}\right)^2}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + Y^2\right]} + \frac{\left(\frac{\delta}{\hbar\gamma}\right)^2 + \left(\frac{\delta}{a}\right)^2}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + (1 + X^2)\right]}\ , \tag{2.47}$$

derived in Sec. B.4 and evaluate it for the measured magnetic field of about 10 G that was present in the region of the apparatus where a $M \longrightarrow \overline{M}$ conversion was searched for. The rescaled fields, X and Y, are about 3×10^{-3} at 10 G. The effect of this field is to completely suppress any conversion between the polarized levels of M and $\overline{M}$, but to leave the conversion between the unpolarized levels essentially undisturbed. The result is

$$S_{\overline{M}} = \left(1.28 \times 10^{-5}\right) \left(\frac{G_{M\overline{M}}}{G_F}\right)^2 , \tag{2.48}$$

where G_F is the Fermi coupling constant. Given a limit on $S_{\overline{M}}$ from the experiment, a limit on the coupling constant, $G_{M\overline{M}}$, may then be calculated.

The result in either case is

$$\frac{N_{\overline{M}}}{N_M} = \left(1.28 \times 10^{-5}\right) \left(\frac{G_{M\overline{M}}}{G_F}\right)^2 . \tag{2.49}$$

Then, the value or limit for $S_{\overline{M}}$ must be given according to Eq. 2.48 or, equivalently, Eq. 2.46, but not Eq. 2.45, since we are not allowed to include those M atoms in the normalization which were not available to conversion at any strength because of suppression by the magnetic field.

It has been customary to give the results of searches for $M \longrightarrow \overline{M}$ as upper limits on $G_{M\overline{M}}$ in units of the Fermi constant, assuming a $(V - A)$ conversion Hamiltonian density. Lately, the limits have also been given on $S_{\overline{M}}$ and this inclination will be followed here.

Chapter 3

The Experimental Search for $M \longrightarrow \overline{M}$

This is not the first experiment to search for the conversion of muonium to antimuonium. As is often the case with experiments seeking to improve on existing results, the present measurement is to be seen as the natural consequence of such previous work, which has built a body of experience indicating which avenues are worth pursuing and which are likely to be fruitless in the development of an improved approach.

To put this experiment in the proper context among previous searches for the $M \longrightarrow \overline{M}$ conversion, the other experiments will be briefly surveyed. Further, this chapter will present in detail the method and apparatus employed in this work.

3.1 Previous Searches for $M \longrightarrow \overline{M}$

The last few years have been particularly active ones in the experimental search for the $M \longrightarrow \overline{M}$ conversion, mainly as a consequence of new developments in experimental techniques of M formation. Interest in searching for this process, however, has been present for a much longer time. Early experiments did not achieve very high sensitivities for the coupling, but were important in developing the experience necessary in designing improved approaches.

After the first observation [Hug60,Hug70] of M, it was not long before the first experimental search [Ama68a] for the $M \longrightarrow \overline{M}$ conversion was carried out. A beam of μ^+ produced at the Columbia University Nevis cyclotron laboratory was directed into a gaseous target of 1 atm of purified Ar, where stopping μ^+ captured electrons from the Ar atoms to form M. As in the work confirming M production, the presence of polarized M was verified by the observation of its characteristic spin-precession signal. In the event of an $M \longrightarrow \overline{M}$ conversion, subsequent collisions of the $\overline{M}$ atom with Ar gas atoms would form $\mu^- Ar$ with high efficiency [Mor66]. The event signature for the presence of $\overline{M}$ was then the observation of the 643 keV $2p - 1s$ (K_α) Ar muonic X-ray, detected in either of two $NaI(Tl)$ scintillation counters after a μ^+ stopped and was not followed by a decay e^+. For 5.2×10^7 μ^+ stopping in the Ar target, 4.2×10^6 M atoms were formed and no $\overline{M}$ events were observed. Because of the severe suppression of the $M \longrightarrow \overline{M}$ conversion in gases due to collisions, the resulting limit on the coupling constant of $G_{M\overline{M}} < 5800\ G_F$ (95 % C.L.) was not very stringent.

An investigation of a cross-channel to the $M \longrightarrow \overline{M}$ conversion, the process $e^-e^- \rightarrow \mu^-\mu^-$, was conducted [Bar69] at the Princeton-Stanford electron storage rings by colliding two 525 MeV electron beams. The Stanford Mark III linear accelerator supplied 300 MeV electrons that were stacked and raised to 525 MeV in the storage rings. Collision products in the transverse direction were analyzed by an arrangement of spark chambers, absorbers, and scintillation counters to detect and distinguish electrons from the more penetrating muons and to veto cosmic ray induced counts. Only events compatible with Möller scattering $(e^-\ e^- \rightarrow e^-\ e^-)$ were observed and an upper limit on the cross-section for $e^-e^- \rightarrow \mu^-\mu^-$ of $\sigma_T < 0.67 \times 10^{-32} cm^2$ (95 % C.L.) was reported for the center-of-mass energy used. This implies an upper limit of $G_{M\overline{M}} < 610\ G_F$ (95 % C.L.) on the muonium-antimuonium coupling.

The first search for $M \longrightarrow \overline{M}$ using M in vacuum [Mar82b] was carried out at the Tri-University Meson Facility (TRIUMF) in 1982. Muonium was formed by stopping a 29 MeV/c μ^+ *surface* beam in a layered silica powder target on CaO substrates, taking advantage of the diffusion of M from within the SiO_2 grains to the surrounding voids. A fraction of any $\overline{M}$ atoms that would have stopped in the CaO coating of the support foils would have induced Ca $2p - 1s$ muonic X-rays. These were searched for using two Ge detectors. For 2.32×10^{10} stopping μ^+, no evidence of a conversion signal was seen and background limitation resulted in an upper limit of $S_{\overline{M}} < 0.04$ (95% C.L.) on the conversion probability per atom. This translates to an upper limit of $G_{M\overline{M}} < 42\ G_F$ (95% C.L.) on the muonium-antimuonium coupling constant. A reanalysis [Bee86] with more complete understanding of the M formation and diffusion process in fumed SiO_2 powders led to an improvement of this limit to $G_{M\overline{M}} < 20\ G_F$ (90% C.L.).

After its employ in searches [Bol88] for $\mu \rightarrow e\gamma$, $\mu \rightarrow e\gamma\gamma$, and $\mu \rightarrow 3e$, the Crystal Box detector [San85], a large solid-angle detector of 4 faces of 9×10 $NaI(Tl)$ crystal modules covered by a plastic scintillator hodoscope, was modified for use in another search for the $M \longrightarrow \overline{M}$ conversion [Ni87,Ni88a,Ni88b]. Muonium in vacuum with kinetic energies of several tens of keV was formed *via* the beam-foil method [Bol81] using a 10 MeV/c sub-surface μ^+ beam from the LAMPF stopped muon channel (SMC) incident on a 0.7 μm Al foil. While charged particles were magnetically swept out of the beam-line, M drifted further through a magnetically shielded region of 280 cm length and then stopped in a 1 μm thick Bi target of 40 cm diameter coated with 7.5 nm of MgO. If an $\overline{M}$ atom struck this target, its impact would have liberated an average of 5 secondary electrons from the MgO coating and caused its breakup. The μ^- would have had sufficient energy to penetrate into the Bi substrate with high probability. There it would be captured on a Bi atom, giving a cascade of $\mu^- Bi$ X-rays, including the transition quanta $3d - 2p$ (L_α) at 2.5 MeV and $2p - 1s$ (K_α) at 6.0 MeV. The secondary electrons from the target were electrostatically collected and counted on a 7.5 cm diameter microchannel plate (MCP) detector, viewing the target from the upstream side. The three-fold coincident observation of the Bi muonic X-rays, K_α and L_α, in the Crystal Box detector and the secondary electron burst on the MCP detector was required to signal an $\overline{M}$ atom striking the target. A maximum likelihood analysis of the two-dimensional photon energy spectrum cut on the presence of a MCP count in the proper time window yielded no events attributable to the $\overline{M}$ signal distribution. The result for the upper limit on the $M \longrightarrow \overline{M}$ coupling constant was $G_{M\overline{M}} < 6.9\ G_F$ (90% C.L.).

At TRIUMF, another search [Hub90,Hub88] for the $M \longrightarrow \overline{M}$ conversion was conducted using a rather original method. A 28.5 MeV/c surface μ beam from the M15 channel was partially stopped in a SiO_2 powder target. Thermal muonium was formed in the powder and diffused out into the vacuum region downstream of it. There, it drifted freely and a possible $\overline{M}$ component was allowed to develop until the system struck the UHV grade surface of a W "catcher" foil, where any μ^- would create ^{184}Ta by nuclear capture on W. After typically 12 hours of beam exposure, the foil was removed, its surface layer chemically removed and placed in a low-background germanium counter. The signature for the ^{184}Ta isotope is a β^- in coincidence with a γ and a delayed γ. In 2 runs totaling 625 hours of data taking, 2.7×10^{12} μ^+ were accepted and the single event passing the $\overline{M}$ signature was attributed to the background process of ^{72}Ge disintegration. The upper limit given on the conversion probability per atom is $S_{\overline{M}} < 2.1 \times 10^{-6}$ at 90 % confidence. This corresponds to a limit of $G_{M\overline{M}} < 0.29\ G_F$ on the coupling constant of a $V - A$ type mediating interaction.

The experiment reported in this dissertation also uses a thermal M source, but detection of M and $\overline{M}$ are afforded through a new and more direct method. Details of the apparatus used and the signatures exploited follow in the next sections.

3.2 The Experimental Apparatus

The limitation on searching for a conversion of M to $\overline{M}$ has always been given by backgrounds, how well these were understood, and by low detection efficiencies. This experiment is, of course, not exempt from the same concerns, but new experimental developments [Bee86,Woo88] and a novel signature have paved the way for a better method to detect $\overline{M}$ atoms with very low background.

3.2.1 The Tools for a New Approach

Producing M in vacuum avoids the large suppression of the coupling to $\overline{M}$ that is due to collisions in gases or solids. The beam-foil interaction [Ber77] applied to μ^+ passing through thin foils was the first source of vacuum M. The charge-exchange reactions of μ^+ incident on a thin Al foil produced M at kinetic energies up to several tens of keV [Bol81,Bol82]. At the corresponding velocities, the M atoms traveled a significant distance during a time interval allowed for a conversion to $\overline{M}$ to take place. As an example, the experiment in ref. [Ni87] used beam-foil M at an average kinetic energy of 11.7 keV, which corresponds to an mean velocity of 0.34 cm/ns, traveling over a flight path of 340 cm in an average of 980 ns. It was during this time that an $M \longrightarrow \overline{M}$ conversion was allowed to take place. Because M produced by the beam-foil interaction is not fully forward-directed [Ahn91], this source of M implied a small solid angle for detection of any $\overline{M}$ component [Ni87,Ni88a,Ni88b]. The development of sources of thermal M [Bee86,Mil86] enabled the detection of an admixture of $\overline{M}$ with a larger solid angle, since M at thermal energies moves only about 1.5 cm during its average lifetime.

While the M formation fraction per incident μ^+ from gaseous and solid targets can exceed 50% [Sta74,Kie79], the beam-foil method has achieved yields up to 12% [Ket90b] and the SiO_2 powder method has reached yields up to 11% [Woo88]. However, the benefits of having M in vacuum far outweigh the disadvantage of the reduction in formation fraction.

Thus, observing the thermal M "cloud" downstream of an SiO_2 powder target over its entire lifetime by looking at decay e^+ and searching for any admixture of Michel-distributed e^- promised a new way to search for an $M \longrightarrow \overline{M}$ mixing.

Since thermal M is detected by observing the decay e^+ originating in the vacuum, two or more multiwire proportional chambers (MWPCs) are needed to reconstruct the decay origins of Michel e^+ through a thin vacuum window. If one is to search for $\overline{M}$ in the same way, a dipole magnet is necessary to separate Michel e^+ from Michel e^- and further MWPCs are required to allow observation of the direction of deflection and the curvature of the tracks in the field.

The background for such a measurement will be dominated by knock-on electrons, caused by Bhabha scattering of incident e^+ on the e^- in the materials in the vacuum window, the space between the MWPCs (air or He bags), and the MWPCs themselves. A simple estimate shows that it may be difficult to improve significantly on the results of $M \longrightarrow \overline{M}$ searches to date in this way.

Assuming that the material present in the MWPCs gives us the typical knock-on probability per decay e^+ (of energy above 30 MeV) of 10^{-5}, there would be one background count per 10^5 observed M decays. According to Eq. 2.48, this corresponds to background limiting beginning at $G_{M\overline{M}} \sim 1\ G_F$, and so would not improve upon the best published limit [Hub90]. Given a high statistics measurement of the background and a distribution representing the conversion signal, one could carry out a maximum likelihood analysis to find any admixture of a conversion in the data. This method would then not be limited directly by the number of counts acquired, but rather by the statistics of the background determination. Thorough understanding of the background distribution(s) is the central difficulty in this approach, though one cannot *a priori* rule out its viability.

A detection method for $\overline{M}$ is needed that is much less susceptible to background and that would be convincing in the case of the observation of a conversion. A signature containing more information characteristic of the $\overline{M}$ system than only the decay e^- would meet this requirement. So the question is: What *else* about $\overline{M}$ can be observed? *The atomic* e^+ *!* An $\overline{M}$ atom disintegrates when the μ^- in it decays and one expects to find an atomic e^+ with about 1 Ry of kinetic energy left behind – an energy that should allow high efficiency detection (see Sec. 3.2.9).

The full detection signature for $\overline{M}$ is then chosen as the coincident observation of the decay e^- with the atomic e^+ . Detecting M atoms in the analogous way – observing time coincident decay e^+ and atomic e^- – can serve as a calibration.

This idea was implemented recently at the Los Alamos Clinton P. Anderson Meson Physics Facility (LAMPF). The objective of this experiment [Sch87] was to observe thermal M from a SiO_2 powder target by detecting for the first time the coincidence of the decay e^+ and the atomic e^- from M atom decays and to search for any $\overline{M}$ admixture by looking for coincident decay e^- and atomic e^+. After the proposal was submitted and approved in 1987, a test run in 1988 demonstrated the feasibility of the experimental method. During the final beam time of more than 800 hours in 1989, about two thirds of the time was required for testing and calibrating the apparatus and the balance was spent searching for the $M \longrightarrow \overline{M}$ conversion.

3.2.2 The LAMPF Accelerator

The heart of LAMPF is a half-mile long, pulsed LINAC capable of accelerating protons (H^+), negative hydrogen ions (H^-), and polarized negative hydrogen ions (P^-) simultaneously to kinetic energies of up to 800 MeV at very high beam intensity. A duoplasmatron source at the high-potential side of an 750 kV Cockroft-Walton accelerator injects H^+ into an Alvarez drift-tube LINAC driven by 201 MHz RF, which boosts the energy up to 200 MeV. The final and longest stage of the machine is the "side-coupled cavity" section, resonant at 805 MHz, which accelerates the beam up to an energy of 800 MeV. The H^- ions are produced by passing the H^+ beam from a duoplasmatron source through a charge-transfer channel filled with H_2 at low pressure. These ions are then accelerated during the reverse phase of the RF cycle in the machine. Focusing of the beams between waveguide sections and in drift regions is accomplished by quadrupole magnets. The beams then enter the "switchyard" where they are distributed to the various experimental areas that include a proton storage ring (PSR), pion channels (LEP, EPICS, P^3), a neutron scattering facility (LANSCE), a neutron time-of-flight spectrometer (NTOF), proton scattering spectrometers (HRS, MRS), a beam-stop neutrino area, and a muon channel (SMC).

The macrostructure of the typical operating conditions of the LAMPF H^+ beam are an average current of 1 mA at a duty factor of 6.4%, where beam bursts are of 800 μs length and 80 Hz repetition rate. The macropulse of the beam has the additional structure of 250 ps long bunches every 5 ns.

3.2.3 The Stopped Muon Channel

The LAMPF Stopped Muon Channel (SMC) (see Fig. 3.1) views the A2 target

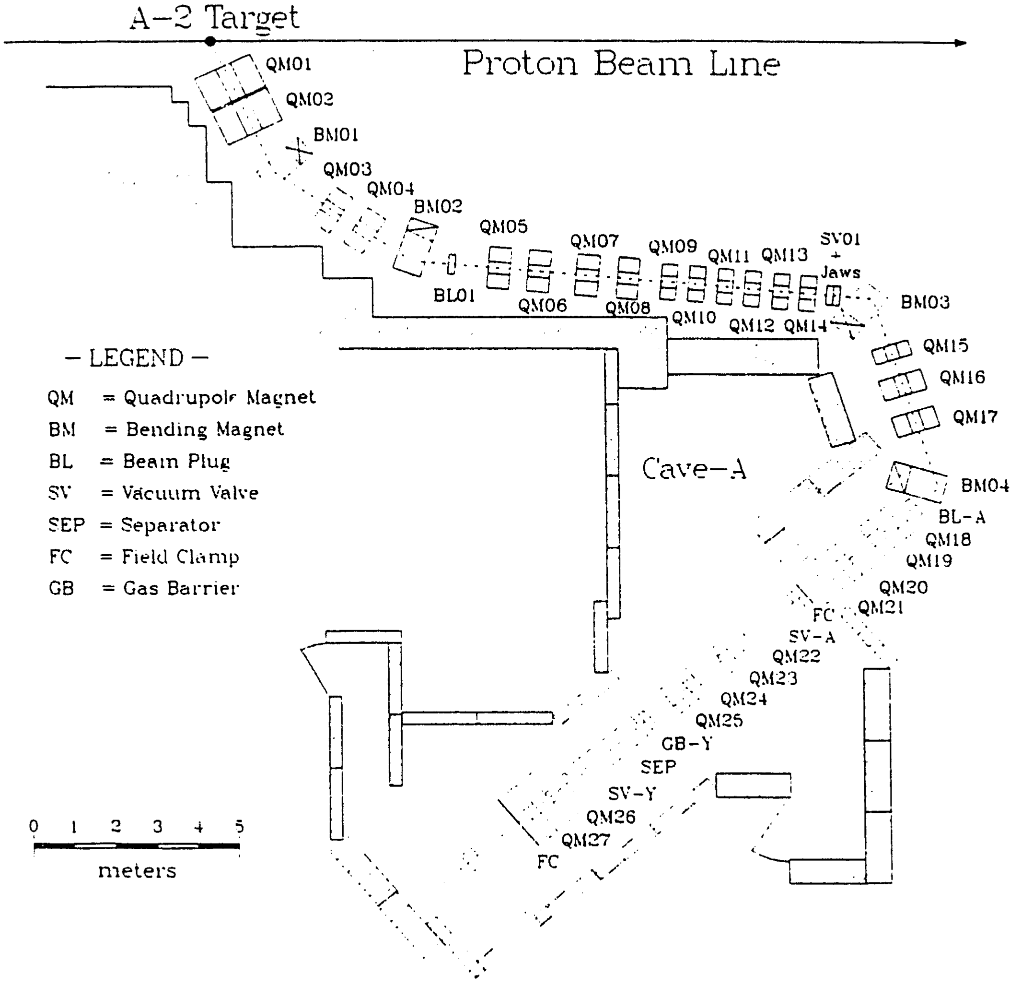


Figure 3.1: The LAMPF Stopped Muon Channel.

wheel in the main proton beam line at an angle of 65° to the proton beam. This channel consists of 4 bending magnets for momentum selection and 21 quadrupole magnets for focusing control. In addition, this experiment used a removable extension of 3 quadrupole doublets and a static-field beam separator to allow locating the apparatus in a semi-permanent fashion, out of the way of other users of the channel.

The channel can accept and transport charged pions that are produced by the interaction of protons with the 4 *cm* thick A2 graphite target, allow these pions to decay in flight and deliver a muon beam at the channel ports [Tho79]. These muon beams (μ^+ or μ^-) are called "decay beams" and are highly polarized. The high degree of polarization is a direct result of the full polarization of the muons from pion decay. Since a small fraction of these decay muons moving into the acceptance of the SMC are transported, the resulting beam has a well-defined polarization. An alternative mode for the SMC is to accept muons from pion decays in the pion "cloud" around the A2 target ("cloud beam"). These beams have rather low polarization. Finally, the SMC can accept positive muons resulting from π^+ decays at rest near the surface of the A2 target ("surface beam") [Pif76,Rei78], or even from π^+ decaying *within* the A2 target ("subsurface beam") [Bad85]. The surface and subsurface μ^+ beams are nearly 100% polarized. Because the nuclear capture rate of π^- is large compared to the $\pi^- \longrightarrow \mu^- \ \overline{\nu_\mu}$ decay rate, the surface and subsurface muon beams are possible only for μ^+ . Since the $\pi^\pm$ free decay lifetime is 26 *ns*, the microstructure of the H^+ beam is washed out and the time-structure of $\mu^\pm$ beams follows that of the macrostructure of the proton beam.

In this experiment, the channel was operated in the subsurface μ^+ mode at about 20 MeV/c beam momentum. Because of the profusion of positrons due to μ^+ decays and pair production from the γ's from π^0 decays at the A2 target and due to μ^+ decays in flight in the channel, the e^+ contamination in the beam would be about [Bad85] $e^+/\mu^+ \sim 10$ at 28 MeV/c. Compared to the μ^+ subsurface rate, the rate of e^+ varies only slowly with the channel momentum, so we can estimate that $e^+/\mu^+ \sim 30$ by scaling the muon rate, R_μ, to 20 MeV/c using the relation [Pif76]

$$R_\mu \propto p^{3.5} \, e^{-\frac{m_\mu \ell}{p \tau_\mu}} \ , \tag{3.1}$$

where p is the momentum, ℓ is the length of the channel (about 40 m), and m_μ and τ_μ are the muon mass and lifetime, respectively. The first factor arises from the range-momentum dependence for muons in this momentum region [Tro66] and the second is a correction due to the decays of μ^+ in flight down the channel. Because these e^+ cause background in some of the detectors of this experiment, it is desirable to remove them from the beam as well as possible. At subsurface momenta, the in-channel degrader (0.3 mm CH_2), that can otherwise be used to separate muons from positrons in the beam by their different energy losses, must be kept out of the beam, so as not to stop it. At 28 MeV/c, this method has been shown to reduce the e^+ contamination by a factor of 10 with slight loss of μ^+ rate [Bad85].

3.2.4 The Wien Filter Beam Separator

Instead, the subsurface μ^+ beam in this experiment is freed of e^+ contamination by passing it through a Wien filter – a section with mutually perpendicular, static $\vec{E}$ and $\vec{B}$ fields, both transverse to the beam. The velocity selected by it may be found by simply requiring the Lorentz force on the particles to vanish:

$$v \propto \frac{|\vec{E}|}{|\vec{B}|}, \tag{3.2}$$

where the proportionality factor may differ from unity due to the slightly different lengths and fringe behaviors of the electric and magnetic field. Another contribution to this factor may arise if the centroid of the transverse positional distribution of the beam were not properly positioned on the beam-line axis. Since the channel has been tuned for a particular momentum and the separator acts as a velocity-filter, it selects on particle mass, effectively on particle type. Thus, it is set to transmit μ^+ with negligible loss while reducing the e^+ fraction by a factor of $\approx 10^1$ [Bad85], leaving a beam positron contamination of roughly $e^+/\mu^+ \sim 0.003$.

The separator used has electric field plates of 152 cm length and 20 cm width spaced by a 10 cm gap and a magnet with effective field length of 146 cm. For this experiment, the high voltage was around -50 kV ($+50$ kV) on the top (bottom) plate leading to an electric field of 10 kV/cm. A vacuum of $\sim 10^{-7}$ $torr$ was maintained in the separator by a CTI-Cryogenics Cryo-Torr 8 cryopump to allow these voltages to be held by the plates. The typical magnetic field was generated by 166 A giving 144 G. As particular example, a tune at 20.5 MeV/c used $E = 9.42\ kV/cm$ and $B = 144\ G$. In units of c_0, the ratio of these fields is $\beta = 0.22$, whereas the nominal momentum value gives $\beta = 0.19$. The discrepancy is due in part to the correction factor in Eq. 3.2 and in part due to an unknown calibration error in relating the channel settings to the transported central momentum.

3.2.5 Radioactive Gas in the SMC

The intense proton beam on the A2 target – heating the A2 target to about 400 K at 650 μA – will produce a variety of short-lived isotopes that may diffuse from the A2 location down the entire length of the SMC and give rise to background in the detectors of experiments[Don83]. For proton currents exceeding 250 μA there is a rapid onset of the production of these spallation products. The gaseous ones are dominantly ^{12}N and ^{6}He with half-lives of 0.011 s and 0.805 s, respectively. With such lifetimes, these isotopes can cover the $\sim$ 40 m distance from the A2 target to the experiment's detectors unless countermeasures are taken. At subsurface momenta, the 25 μm Kapton gas barrier in the channel cannot be used, as it would stop an appreciable fraction of the μ^+ beam. Instead, a 1.5 μm Mylar gas barrier is placed just upstream of the beam separator. By itself, it can stop the ^{12}N, but the ^{6}He would mostly pass through it. The important effect of this gas barrier is to separate the vacuum region of the experimental apparatus, with a pressure of $< 10^{-6}$ $torr$, from the vacuum in the channel, which settled at $\sim 10^{-3}$ $torr$ when thus decoupled from the pumps on the separator and the apparatus. This rather poor vacuum retarded the diffusion of the gaseous spallation products so that most of them decayed before reaching the experimental apparatus.

3.2.6 The Muonium Formation Process

The formation of thermal M by stopping μ^+ in a SiO_2 powder target has become a well-established technique [Bee86,Woo88,Jan90]. In this experiment, the subsurface μ^+ beam, which had a relative momentum width of about 10%, was partially stopped in a SiO_2 powder target of 9 mg/cm^2 typical thickness. As the target was always inclined at 50° to the horizontal, this thickness translates to about 12 mg/cm^2 projected along the beam axis. The stopping μ^+ could then capture an e^- and diffuse from the powder grains into the vacuum region downstream of the target. The M atoms thus produced have been shown to have thermal energies.

The fraction of μ^+ forming M upon stopping in SiO_2 powder has been measured by the spin-rotation technique to be $(61 \pm 3)\%$ [Kie79]. Of the M atoms thus produced, $(97 \pm 1)\%$ emerge from the silica grains into the surrounding vacuum regions [Mar78]. This is determined by introducing a few *torr* of O_2 into the system and observing that the M precession signal is almost entirely quenched. The reason is the high cross-section for depolarizing spin-exchange collisions of M with O_2, which can, of course, only be taking place in the space between the powder grains.

Of the remaining $(59 \pm 3)\%$, some M atoms will diffuse to the surface of the powder layer and escape into the vacuum region beyond the powder target. Marshall has estimated this rate using a diffusion constant obtained by modeling the powder as a uniform distribution of SiO_2 grains [Mar81]. This diffusion coefficient is [Jan90]

$$D = \frac{4}{9} v\, r \left(\frac{\rho}{\rho'}\right) , \tag{3.3}$$

where v is the mean speed of the M atoms, r is the radius of the spherical SiO_2 grains, $\rho = 2.2\ g/cm^3$ is the bulk density of silica, and $\rho' = 0.032\ g/cm^3$ is the density of the silica powder. For the fumed silica used [Cab88], the radius of the particle grains may be estimated to be 7 nm, as determined from N_2 adsorption measurements. For thermal M at $\sim 300\ K$, one obtains $D \approx 8\ cm^2/s$ *via* $v \approx 0.75\ cm/\mu s$.

The diffusion equation for this problem may be written as

$$\frac{\partial \rho}{\partial t} = D \vec{\nabla}^2 \rho - \gamma \rho , \tag{3.4}$$

where ρ is the number density of M atoms in the powder and γ is the muon decay rate. This diffusion equation follows from Fick's Law

$$\vec{j} = -D\vec{\nabla}\rho \tag{3.5}$$

and the equation of continuity

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{j} = -\gamma\rho \ , \tag{3.6}$$

where a "sink" term representing muon decay has been added on the right-hand side compared to the sourceless continuity equation. It is possible to isolate the muon decay time-dependence by the definition of a new variable η, implicitly given by

$$\rho(\vec{r},t) = \eta(\vec{r},t)\, e^{-\gamma t} \ , \tag{3.7}$$

where the space- and time-dependence of both ρ and η has been explicitly indicated. The density η can be interpreted as a number density which does not suffer an effective sink from decay. With the concurrent substitution of

$$\vec{j} = \vec{k}e^{-\gamma t} \tag{3.8}$$

one may write

$$\frac{\partial\eta}{\partial t} + \vec{\nabla}\cdot\vec{k} = 0 \tag{3.9}$$

and

$$\vec{k} = -D\vec{\nabla}\eta \ . \tag{3.10}$$

Using this change of variable, the diffusion equation becomes

$$\frac{\partial\eta}{\partial t} = D\,\vec{\nabla}^{\,2}\eta \ . \tag{3.11}$$

The normalization conditions on the M number densities are

$$\begin{aligned} \int \eta(\vec{r},t)\, d^3\vec{r} &= N_M(t=0) \\ \int \rho(\vec{r},t)\, d^3\vec{r} &= N_M(t=0)\, e^{-\gamma t} \ , \end{aligned} \tag{3.12}$$

where $N_M(t = 0)$ is the number of M atoms that have migrated from within the powder grains to the surrounding empty space.

Though the problem does not have exact azimuthal symmetry because of the angular placement of the target and the finite-sized, elliptical μ^+ beam spot, we will make the useful approximation that the x- and z-dependences in the density η may be integrated out:

$$\eta'(y,t) = \int\int \eta(\vec{r},t)\,dx\,dz \ , \tag{3.13}$$

where the y-direction is that perpendicular to the powder-vacuum interface surface. Dropping the prime, the diffusion equation becomes

$$\frac{\partial \eta}{\partial t} = D\,\frac{\partial^2 \eta}{\partial y^2} \ . \tag{3.14}$$

The boundary conditions relevant here are that the particle density vanish at the powder-vacuum interface, as the escape of M into the vacuum acts like an infinite sink, and also that the density vanish at the Al-mylar foil that holds the target powder, assuming that the M atoms that hit this surface will be destroyed. What one wishes to solve for is the current of M atoms passing through the powder-vacuum surface. Such a solution would allow us to estimate the expected M formation fraction, given a stopping distribution of μ^+ , a diffusion constant, and the target thickness.

Equation 3.14 may be solved either by a Fourier series in the position variable, y, or by an initial-condition Green's function method. In the first solution method, the boundary conditions are intrinsically satisfied by the choice of expansion function, whereas in the second approach one must make use of the method of images to fulfill them. We let the holder foil-powder interface be located at $y = 0$ and the powder-vacuum interface at $y = a$. Then we define the stopping distribution of μ^+ in the powder by $S(y)$, where

$$\int_0^a S(y)\,dy = N_{\mu^+} \ . \tag{3.15}$$

Here, N_{μ^+} is the total number of μ^+ stopping in the target. This may be related to the number of M atoms by

$$N_M = f\,N_{\mu^+} \tag{3.16}$$

and

$$\eta(y, t=0) = f\, S(y) \ , \tag{3.17}$$

where f is the formation fraction of M that reaches the space between the powder grains. From above, $f = (59 \pm 3)\%$.

The Fourier series solution is given by

$$\eta(y,t) = \sum_n C_n e^{-\lambda_n t} \sin\left(\frac{n\pi y}{a}\right) \ , \tag{3.18}$$

where

$$\lambda_n = \left(\frac{n\pi}{a}\right)^2 D \tag{3.19}$$

and the coefficients, C_n, are to be determined by the distribution at $t = 0$, *i.e.* the stopping distribution, according to

$$C_m = \frac{2}{a} f \int_0^a S(y) \sin\left(\frac{n\pi y}{a}\right) dy \ . \tag{3.20}$$

The solution may be checked by assuming a flat stopping distribution, calculating the particle current j_y at the powder-vacuum interface, integrating over all times, and taking the limit $\gamma \to 0$. The sum over Fourier terms may be carried out using a Watson transform to convert the sum into a calculable contour integral. As expected for this case, one finds that exactly half of the M atoms formed emerge at the vacuum surface.

The full solution from the Green's function approach, including the infinite number of images required to satisfy the boundary conditions, is

$$\begin{aligned}\rho(y,t) \quad &= \quad \frac{e^{-\gamma t}}{2\sqrt{\pi D t}} f \int_0^a S(y')\, dy' \\ &\quad \times \sum_{n=-\infty}^{\infty} \left[e^{-(y+2na-y')^2/(4Dt)} \right. \\ &\quad \left. - \ e^{-(y+2na+y')^2/(4Dt)} \right] \ . \end{aligned} \tag{3.21}$$

The current of particles crossing the $y = a$ surface is then given by

$$j_y(y=a,t) \quad = \quad -D \left.\frac{\partial \rho}{\partial y}\right|_{y=a}$$

$$\begin{aligned} &= \frac{e^{-\gamma t}}{2\sqrt{\pi D t^3}}\, f \int_0^a S(y')\, dy' \\ &\times \sum_{n=1,3,\ldots}^{\infty} \Big[(na - y')\, e^{-(na-y')^2/(4Dt)} \\ &- (na + y')\, e^{-(na+y')^2/(4Dt)} \Big] \ . \end{aligned} \tag{3.22}$$

The total number of M atoms crossing this surface, integrated over all times, is then

$$N_{vac} = \int_0^\infty j_y(y = a, t)\, dt \ . \tag{3.23}$$

This integral is a little tricky, but with a wise change of variable and sectioning of the integration region, it can be carried out. The result is a series that may be easily summed to give

$$N_{vac} = f\, \frac{2e^{-\sqrt{\frac{\gamma}{D}}a}}{\left(1 - e^{-2\sqrt{\frac{\gamma}{D}}a}\right)} \int_0^a S(y') \sinh\left(\sqrt{\frac{\gamma}{D}}y'\right)\, dy' \ . \tag{3.24}$$

This expression has also passed the check described above for the case of a flat stopping distribution and a vanishing muon decay rate.

For reference, we give the solution for flat stopping distribution

$$\begin{aligned} S(y) &= \frac{N_{\mu^+}}{a} \quad \text{for } 0 < y < a \\ &= 0 \quad \text{for } y < 0 \ \text{ or } \ y > a \ . \end{aligned} \tag{3.25}$$

The result is

$$N_{vac} = f\, \frac{N_{\mu^+}}{a} \sqrt{\frac{D}{\gamma}} \tanh\left(\sqrt{\frac{\gamma}{D}}\frac{a}{2}\right) \ . \tag{3.26}$$

The two methods are, of course, completely equivalent and related by a Watson transform. The difference lies in varying computational convenience depending upon the case being considered.

The estimate of the M formation fraction of thermal M in vacuum per μ^+ stopping in the powder can now be written down as

$$\begin{aligned} f_{M,vac} &= \frac{N_{vac}}{N_{\mu^+}} \\ &= \frac{f}{a} \sqrt{\frac{D}{\gamma}} \tanh\left(\sqrt{\frac{\gamma}{D}}\frac{a}{2}\right) \ , \end{aligned} \tag{3.27}$$

where f gives the fraction of stopping μ^+ that capture an electron and escape from the SiO_2 particles; the rest of the expression describes the efficiency of the diffusion process in delivering M atoms to the surface of the powder, starting from an initially flat position distribution.

Inserting $f = 59\%$ and $D = 8\ cm^2/s$ and assuming that the target thickness is $a = 3\ mm$ gives $f_{M,vac} = 1.4\%$. Actually, an example of a measured M yield is 18% per stopped μ^+ [Woo88]. This result corresponds to a diffusion constant of about 1300 cm^2/s ! The reason for this very large discrepancy in the estimated and the indirectly measured diffusion coefficients is presumed to be due to the structure of the SiO_2 powder. Electron micrographs [Cab88] show that the powder particles are not at all distributed uniformly. Rather, they tend to form chains with large spaces between, drastically increasing the mean free path that a M atom experiences in moving through this powder and leading to a much larger diffusion constant.

Once a M atom reaches the vacuum, it travels unimpeded until it decays. It is during this travel time that a component of $\overline{M}$ could develop. The speed is assumed to be Maxwell-Boltzmann distributed at the temperature of the powder, but the directionality has been found to be more likely distributed according to $\cos\theta$, where θ is the angle to the normal on the target surface [Woo88,Jan90], than isotropically. This is also what one might expect from a simple geometrical argument. To escape from the powder into the vacuum, a M atom will have to pass through a "hole" in the target surface. The probability of passing through this hole is proportional to its projected area along the direction of motion of the atom. If the hole area is A, then the projected area seen by the M atom is $A\cos\theta$, where θ is the angle between the normal to the area element and the direction of incidence of the atom. This gives rise to the $\cos\theta$ angular distribution of the M atoms escaping from the powder into the vacuum region beyond it.

3.2.7 Experimental Details of M Formation

To maximize the production of M, one must arrange the most favorable stopping distribution of μ^+ within the powder target. This has two aspects that are coupled. First, the momentum of the incident μ^+ beam must be chosen roughly to maximize the fraction stopping in the powder. Secondly, since the formation of the M atoms which diffuse into the vacuum takes place at the surface of the target and down to a distance within it that is characterized by the diffusion length, the stopping distribution should be as narrow as possible to maximize the number of μ^+ stopping in this region of the powder. This requires a beam of narrow momentum spread. As arranging this often comes at the expense of absolute rate from conventional muon channels, a suitable compromise must be chosen.

In the case of this experiment, the tune for the SMC had a FWHM momentum spread of about 10% while the full subsurface channel rate that was expected at the momenta used was obtained. With fully open channel slits, this rate at 20 MeV/c channel momentum was about 10^6 s^{-1} (average) at 6.4% duty factor with about 800 μA in the primary proton beam.

This μ^+ beam was counted by transmission through a NE104 plastic scintillator of 150 μm thickness mounted at 50° to the horizontal. Its projected thickness along the beam axis was then 196 μm. This scintillator was viewed by two Amperex XP2020 phototubes (5 cm diameter) mounted on light guides that acted also as vacuum feed-throughs for the scintillation light. The deposition of approximately 600 keV in the scintillator for each incident μ^+ gave several thousand scintillation photons ($\sim$ 180 eV/photon). This allowed nearly 100% detection efficiency for the μ^+ , as long as the instantaneous rate in the beam was not so high as to cause pulse pileup in the tubes. At the full channel rate this was the case, and the measure taken was to lower the operating voltage on the tubes. This cured the pileup problem, possibly at the expense of some of the detection efficiency. The reduction in efficiency was not determined, as the absolute number of counts from the beam scintillator was not needed in this case. Even if the efficiency was reduced, it served well as a relative rate monitor to check the constancy of the beam rate and to guide the tuning of the channel magnets for maximum μ^+ rate.

The scintillator also acted as a degrader that worked in conjunction with the powder target to slow the μ^+ from initially about 1.9 MeV of kinetic energy to a stop in the powder. The balance of the energy was lost in the 6 μm Al-mylar holder foil of the target assembly and the approximately 3 mm thick powder layer placed on it.

The target powder was deposited on the 6 μm Al-mylar foil, which was held in a U-shaped frame with the open end toward the vacuum window and the spectrometer. The powders used were Cab-O-Sil grades M-5, EH5, and PTG [Cab88]. Before making a target, the powder containers were shaken to fluff up the powder in case storage had caused it to collapse and compress. This sometimes gave a surprising doubling of the volume of the powder. It was then sifted onto the target foil through a fine mesh, taking care to distribute it evenly across the entire area of the target. The reflectivity of the aluminized mylar holder foil helped judge the uniformity of the powder layer. The target thickness was determined by weighing the target frame before and after adding the powder layer to it and measuring the area covered by powder. This area was 177.4 cm^2. The powder portion of the typical target weighed about 1.6 g and therefore had an area density of 9 mg/cm^2. This corresponds to a thickness of roughly 3 mm, since the density of the fumed silica powder is 32 mg/cm^3. As this procedure did not always produce targets of equal thickness, the beam momentum had to be tuned for optimal M production for each new target.

The pressure in the vacuum system was kept at $< 10^{-6}$ $torr$ by a 1000 l/s Cryo-Torr 7 cryopump (CTI-Cryogenics) placed below the target region, a 600 l/s Turbo-Torr Model 3131 turbomolecular pump (Sargent-Welch) located just upstream of the target chamber, and two Balzers turbopumps – TPU200 (200 l/s) and TPU330 (300 l/s) – evenly spaced on the downstream section of the apparatus. Good vacuum is necessary to ensure that M escaping the powder target is not destroyed by collisions with residual gas molecules. At pressures $< 10^{-6}$ $torr$, the mean free path is estimated to be $> 10^4$ cm. The nature of the remaining gas in the vacuum system was monitored by a residual gas analyzer (Leybold-Inficon). The greatest impairment to M formation came from the adsorption of pump oil fragments on the SiO_2. This was only a problem when the oil level on the roughing pumps (Sargent-Welch) ran low, causing excessive backstreaming of pump oil vapors. The remedy is refilling the pump oil to the required level and, necessarily, installing a fresh SiO_2 target.

New targets were usually installed once every three or four days, as we did observe a steady decline in the M formation fraction over this time-scale, typically by a factor of two in this interval. The reason for the decrease in M is not clear. One possibility is contamination of the powder surface with hydrocarbon fragments from the pump oil backstreaming. Another may be that the microstructure of the powder is somehow affected by pumping out the adsorbed gases and water from the large surface area it presents. Finally, simple mechanical settling of the powder grains over time, assisted by the slight but unavoidable vibrations in the apparatus, may have the effect of reducing the open volume in the powder layer and thus diminishing the diffusion constant the M atoms experience. An observation which may be related is that when old targets were removed and their powder shaken off into the waste basket, the powder layer fell off in loosely connected flakes. This is to be contrasted to an almost liquid behavior of the fresh powder used to cover a new target. Due to scheduling constraints, it was not possible to study explicitly the time dependence of the formation fraction and the effects that may be involved.

The velocity distribution of the M atoms in the vacuum folded with the decay time dependence of the muon gives rise to a characteristic position distribution of the points where the M atoms decay. Since the thermal velocity at 300 K is about 0.75 $cm/\mu s$ for M, the major part of this position distribution is well-contained in a region of only a few cm in size. This region may be completely viewed by an arrangement of MWPCs that seeks to detect the e^+ resulting from the M atom decays. To facilitate the passage of decay e^+ from this region into the MWPCs, which will be described in the next section, a 15 cm diameter 100 μm thick heat-tempered Al vacuum window was installed on the side of the beam-line facing the MWPCs.

3.2.8 The Decay Electron Spectrometer

Central to the experiment was the ability to detect the decay e^+ from M and μ^+ decay as well as any decay e^- that might signify an $M \longrightarrow \overline{M}$ conversion. The momentum distribution of these decay $e^\pm$ from $\mu^\pm$ disintegrations follow the well-known Michel distribution [Mic50]. This distribution rises monotonically up to

52.8 MeV/c with very little of its area below 15 MeV/c. Averaging over final state helicities, neglecting the electron mass, and evaluating at an observation angle of 90° relative to the incident muon polarization, this distribution (see Fig. 3.2) becomes

$$M(\epsilon) = 2\epsilon^2 (3 - 2\epsilon) \ , \tag{3.28}$$

where ϵ is the decay $e^{\pm}$ energy in units of half the muon mass:

$$\epsilon = \frac{2E_e}{m_\mu c^2} \leq 1 \ . \tag{3.29}$$

Thus, these e^+ and e^- lend themselves well to high-efficiency detection in an array of position sensitive multiwire proportional chambers (MWPCs). The position sensitivity allows recognition of $e^{\pm}$ tracks defined as a correlated set of hits in these chambers and also enables tracing the trajectory back to find the position distribution of the decay points. It is in this distribution that evidence for $\overline{M}$ decays is sought.

Four MWPCs are placed on an axis (z) at a right angle to the beam-line with their planes perpendicular to this axis, as shown in Fig. 3.3. The locations chosen for the MWPCs along the z-axis were dictated by the requirement for maximum solid angle subtended at the SiO_2 target in the beam-line. Space was left only to accommodate the read-out electronics for the chambers, the fixtures that held them in place, and the magnet between the second and third MWPC.

The active areas of the first two chambers were 32 cm × 32 cm, covered by 160 wires spaced by 2 mm along both the x- and y-planes. The x-axis has been chosen to lie along the projected plane of the SiO_2 target so that the y-direction is the axis perpendicular to the target. Since the powder target was mounted at an angle of about 50° to the horizontal, the MWPCs were also mounted with this rotation angle. The reason for this will be explained when the trigger is discussed.

The third and fourth MWPCs were larger, having active areas of 89.6 cm × 32 cm, where the x-dimension is the first and the y-dimension the second. Each of these MWPCs had 448 wires sensing the x-coordinate and 160 wires along the y-dimension, all spaced by 2 mm. In addition, these chambers had a plane oriented at 45° to the others with 416 wires, spaced by 2.02 mm. In the following, these

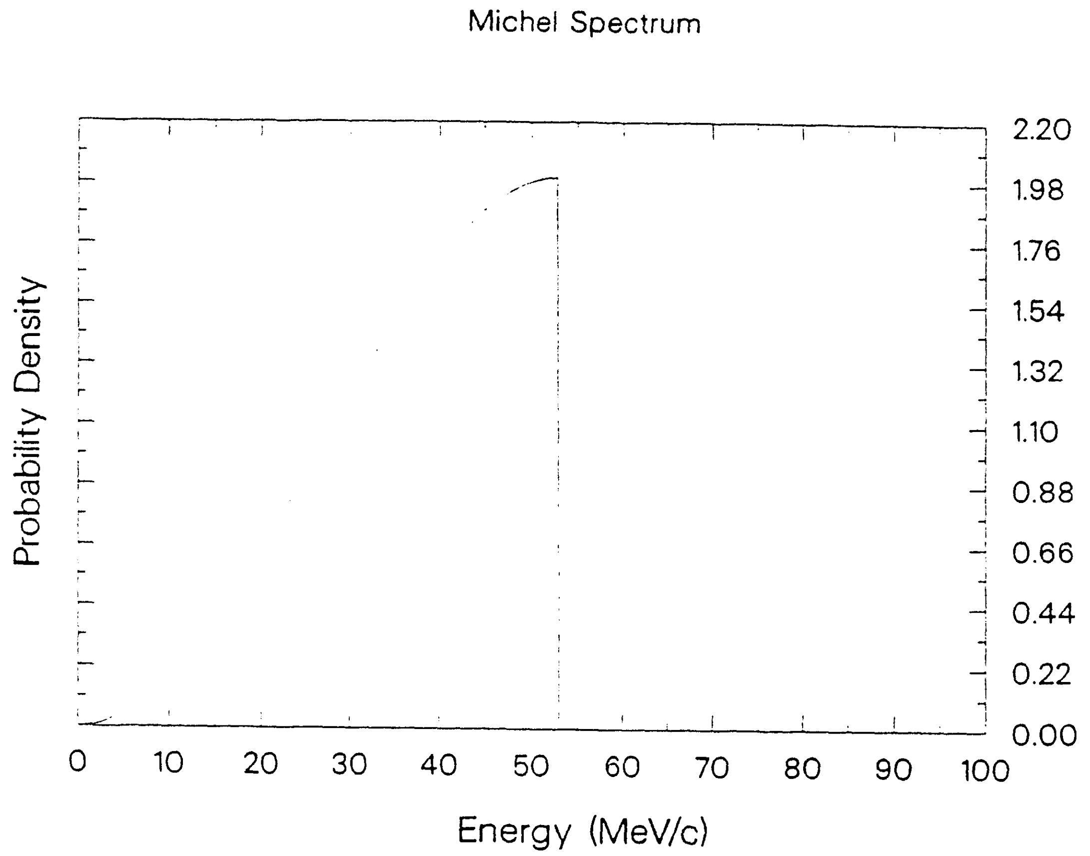


Figure 3.2: Energy distribution of e^+ from μ^+ decays.

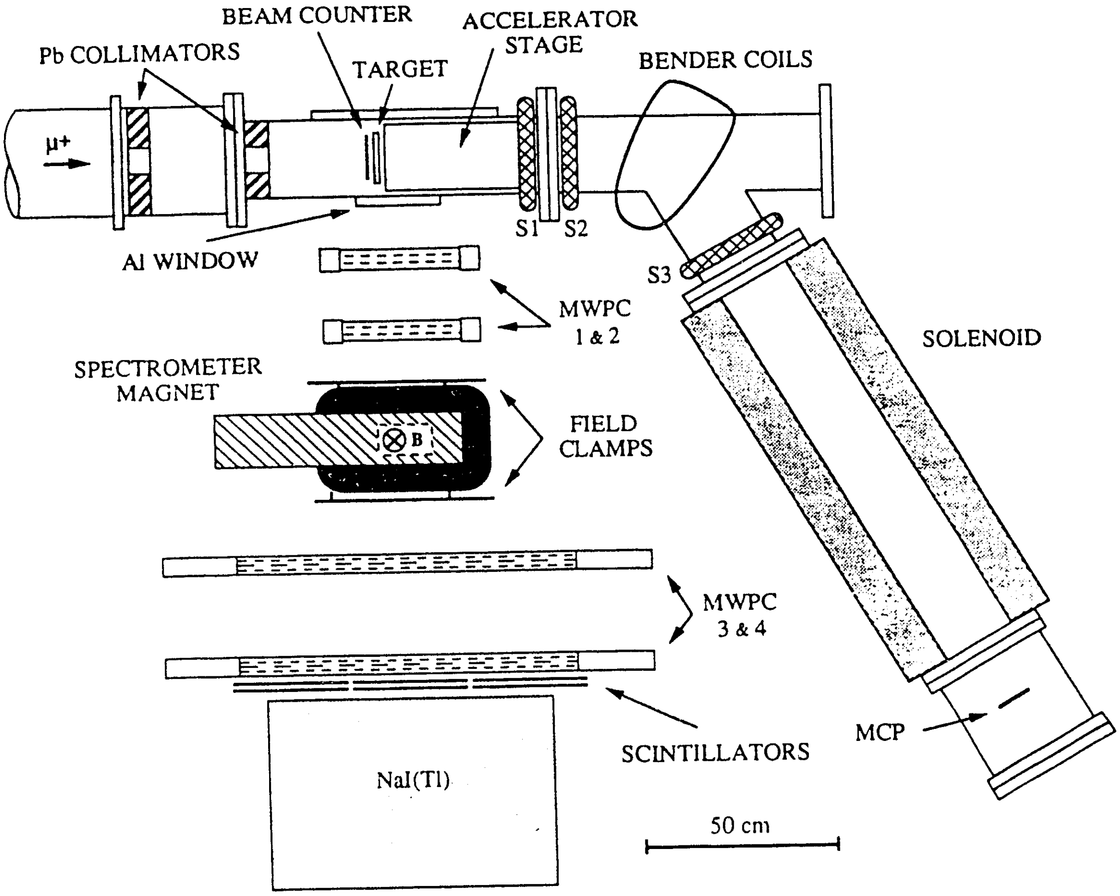


Figure 3.3: Schematic of the apparatus used to search $M \longrightarrow \overline{M}$. View is from the top for the vacuum apparatus and at 50° to the vertical for the spectrometer elements.

planes will be referred to as the "u-planes." The information from this plane is not purely redundant, as it can help pair up the correlated x- and y-hits when there are multiple hits in these chambers. It may also be used to restore a missing hit in x or y, when one of these planes does not respond to a particle passing through it. The relations needed are

$$\begin{aligned} u(mm) &= -\frac{1}{\sqrt{2}}\left[x(mm) + y(mm)\right] + 846.1\ mm \\ x(mm) &= -\sqrt{2}\left[(2.02) \times u(\text{wire no.}) - 846.1\ mm\right] - y(mm)\ , \end{aligned} \qquad (3.30)$$

where u designates the 45° plane coordinate. The large MWPCs are schematically shown in Fig. 3.4. A summary of MWPC dimension data is presented in Table 3.1.

All wires in the MWPCs were 20 μm diameter gold-plated tungsten strung at a tension of 50 g onto G10 epoxy-fiberglass frames. The guard wires at the edges of the wire planes were thicker and mounted with higher tension in order to hold the electrostatic force from the neighboring wires and to control sparking from this last wire to the high-voltage planes of the MWPCs. See Tables 3.2 and 3.3 for a summary of the sequence of windows, high-voltage planes, and wire planes constituting the MWPCs and of the materials and thicknesses used for these. To estimate the effects of multiple scattering on the particle trajectories, the track fitting algorithm will require the radiation lengths in these chambers, so these have been estimated to be 7773 cm for the small MWPCs and 2215 cm for the large MWPCs, where the chambers are thought of as homogeneous scattering media in this presentation. The calculation actually added separately estimated contributions from each of the layers in the chambers to then arrive at the results stated for the net effect due to the entire MWPC. A final number to characterize the behavior of the MWPCs is the average energy loss suffered by a 35 MeV/c $e^{\pm}$ passing through them. This energy loss is about 30 keV for the small chambers and about 255 keV for the large ones.

The gas mixture used for operation throughout was determined by minimizing the idle current drawn by the MWPCs at the plateau voltage operating point while maximizing the efficiency reached on this plateau. The gas used was a mixture of

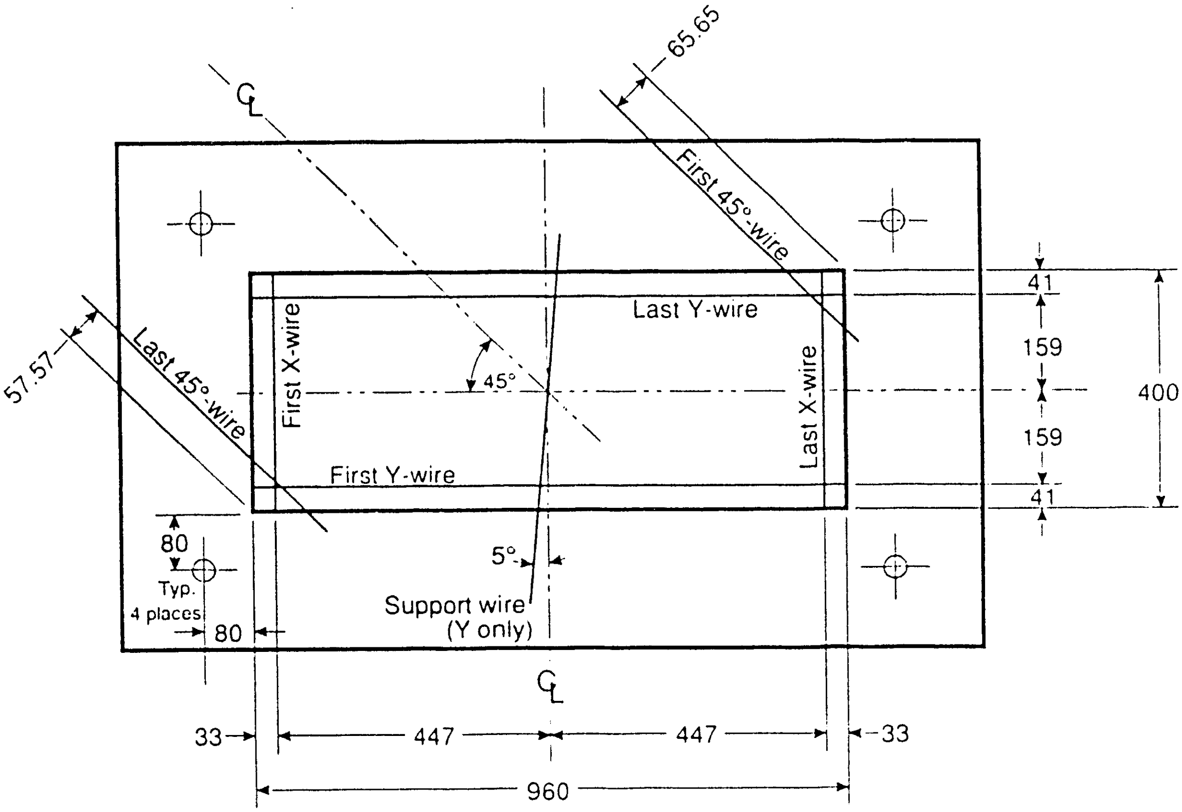


Figure 3.4: Schematic of large MWPCs with dimensions.

	MWPC1	MWPC2	MWPC3	MWPC4
x-dimension	320 mm	320 mm	896 mm	896 mm
no. of x-wires	160	160	448	448
y-dimension	320 mm	320 mm	320 mm	320 mm
no. of y-wires	160	160	160	160
u-dimension	–	–	840.44 mm	840.44 mm
no. of u-wires	–	–	416	416
z-location	326 mm	493 mm	1369 mm	1802 mm

Table 3.1: Overview of dimensions and locations of MWPCs in spectrometer. The z-locations given are measured between the center of the MWPC bodies and the beam-line axis.

Thicknesses of planes in small MWPCs		
type of plane	material	thickness
window	mylar	13 μm
gap	MWPC gas	0.64 cm
HV plane	Al-mylar	6.4 μm
gap	MWPC gas	0.64 cm
wire plane	Au/W wires, 2 mm spacing	20 μm
gap	MWPC gas	0.64 cm
HV plane	Al-mylar	6.4 μm
gap	MWPC gas	0.64 cm
wire plane	Au/W wires, 2 mm spacing	20 μm
gap	MWPC gas	0.64 cm
HV plane	Al-mylar	6.4 μm
gap	MWPC gas	0.64 cm
window	mylar	13 μm

Table 3.2: Material composition of small MWPCs.

Thicknesses of planes in large MWPCs		
type of plane	material	thickness
window	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$1.76\ cm$
HV plane	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$0.48\ cm$
wire plane	Au/W wires, 2 mm spacing	$20\ \mu m$
gap	MWPC gas	$0.48\ cm$
HV plane	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$0.48\ cm$
HV plane	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$0.48\ cm$
wire plane	Au/W wires, 2 mm spacing	$20\ \mu m$
gap	MWPC gas	$0.48\ cm$
HV plane	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$0.48\ cm$
HV plane	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$0.48\ cm$
wire plane	Au/W wires, 2 mm spacing	$20\ \mu m$
gap	MWPC gas	$0.48\ cm$
HV plane	Al mylar	$8.9\ \mu m$ $76\ \mu m$
gap	MWPC gas	$1.76\ cm$
window	Al mylar	$8.9\ \mu m$ $76\ \mu m$

Table 3.3: Material composition of large MWPCs.

75% Ar, 25% isobutane (C_4H_{10}), a trace (0.09%) of Freon 13B1 ($C\,Br\,F_3$) and a trace of 2-propanol ($C_3H_7\,OH$), which was added by bubbling the mixture of the first three gases through a refrigerated flask of 2-propanol. The fractions of each gas in the mixture were adjusted by pressure regulators followed by needle valves and calibrated flow-meters before combining them in a ballast tank for mixing. From this reservoir, the flow rates for the MWPCs were maintained at 100 cm^3/min for the large chambers and at 75 cm^3/min for the small chambers by a passive gas-metering system of needle-valves and flow-meters that measured the supplied and returned gas flow. This allowed detection of any leaks in the MWPCs which might admit air into them and degrade their performance. The overpressure of operating gas in the chambers was about 1.8 $torr$ above the ambient atmospheric pressure of an average of 585 $torr$ in Los Alamos (at an elevation of about 2100 m above sea-level). In a test, the idle MWPCs held this overpressure without loss over a period of at least several hours, so they were assumed to be leak-tight for their purpose. The operating voltages were chosen to be $-3700\ V$ for the small chambers, which had a 0.64 cm half-gap, and $-2650\ V$ for the larger ones, with a half-gap of 0.48 cm. By "half-gap" one refers to the perpendicular distance from a plane of wires to either of their neighboring HV planes. The idle currents at these conditions were less than 10 nA for all planes. With the full beam rate present in the experimental cave, these values rose to several hundred nA, with more current being drawn in the chambers closer to the beam-line. Over the course of the experiment, all MWPC planes maintained efficiencies for Michel e^+ of around 95%. As an example, the efficiency of the MWPC3X plane plotted against the run number in the experiment is shown in Fig. 3.5.

A C-yoke dipole magnet with a 30 cm gap was placed between the second and third MWPCs to deflect Michel e^+ and e^- in opposite directions. The pole tips were rectangular, 30 cm wide and 14 cm long with respect to the spectrometer axis, and the field at the center of the gap was 522 G. Field clamps below and above the gap on either side of the magnet sought to control the fringe fields. This magnet was also mounted at the canonical 50° angle, so that the main component of the field, B_y, was responsible for deflecting the $e^\pm$ along the x-axis. The mapping of

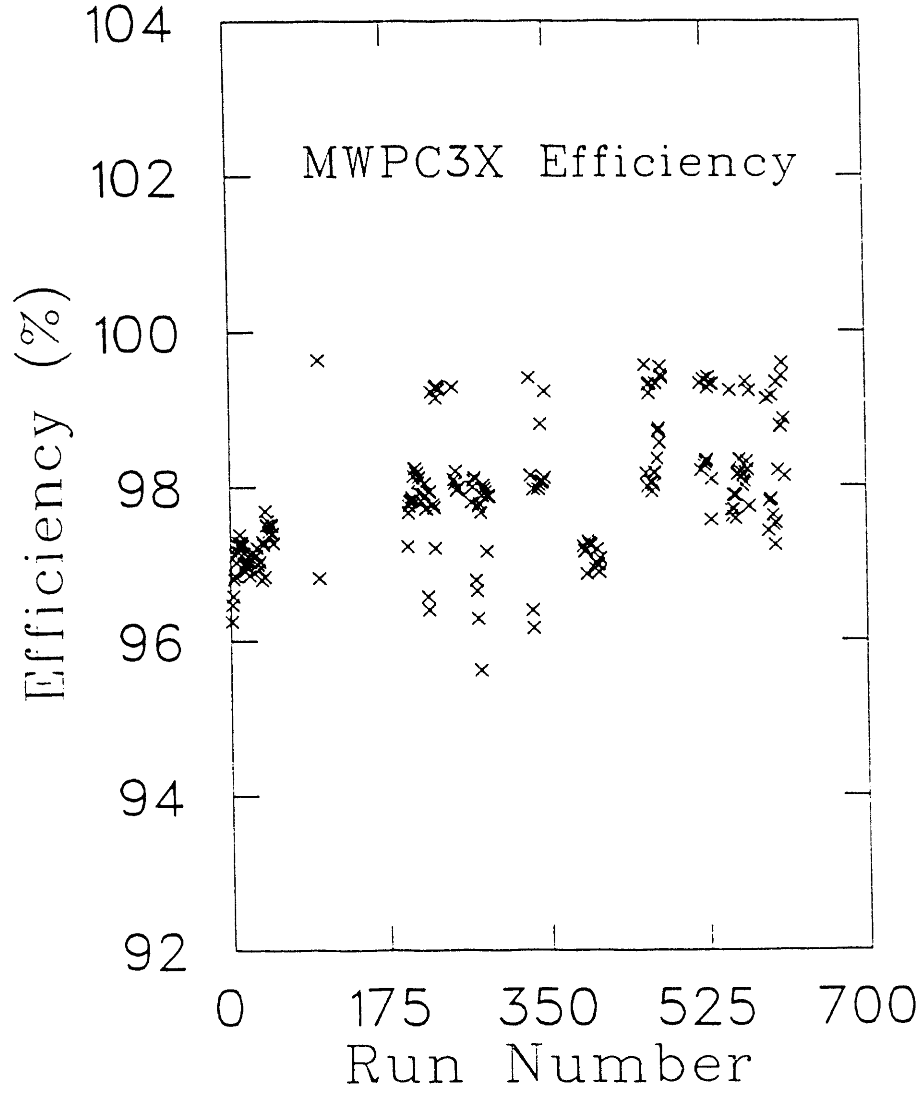


Figure 3.5: The detection efficiency of the MWPC 3X plane over the course of the experiment.

the field produced by this magnet in described in Sec. 4.1.

Two planes of plastic scintillator (NE104) were located behind MWPC4 at $z = 2020$ *mm* and $z = 2070$ *mm*. Their dimensions were 991 *mm* along x and 356 *mm* along y, and so completely covered the acceptance of MWPC4. Each plane was actually composed of three separate scintillators mounted adjacently, but their signals were logically added to effectively make two large scintillator planes. As MWPC signals, with rise-times of $\sim$ 50 *ns*, are not well-suited for accurate timing applications, these scintillator planes in coincidence provided a signal that marked the time of passage of a particle through the spectrometer.

A large, cylindrical $NaI(Tl)$ crystal [Hug72] placed behind the plastic scintillators served to measure the Michel $e^{\pm}$ energy when the spectrometer field was turned off for calibration purposes and to provide a backup energy measurement to compare to momentum information deduced from the track in the MWPCs. Also, it was essential to the online analysis, since the thorough treatment of track information was then not yet available. This crystal was 76 *cm* in diameter and 51 *cm* deep. Scintillation light from it was viewed by ten 5 *in* diameter phototubes (RCA8055). The tube bases had passive gain-balancing potentiometers to adjust the voltage on the first few dynodes. These acted more as a timing adjustment than as a gain adjustment, so the criterion used to set them was for time coincident signals when the crystal was tested with the 4.4 *MeV* γ's from a $PuBe$ source. This ensured coherence of the signals from all 10 tubes when each responded partially to the energy deposited by a Michel e^+. When the analog signals from each of the tubes were then added, the resulting pulse-height distribtion had a resolution of about 20% at the high-energy edge of the Michel distribution (53 *MeV*).

For data taken to tune the momentum of the beam to maximize the production of M, the trigger used to cause a spectrometer track to be latched and read out by the data acquisition system required MWPC planes 1X and 4X to respond in addition to hits in all MWPC Y-planes and pulses in the scintillator planes (see Fig. 3.6). This trigger was only useful when the incident μ^+ rate was reduced, using a variable slit aperture in the channel ("jaws"), to the order of $3 \times 10^1\ s^{-1}$ (average) to prevent excessive dead time in the data acquisition system.

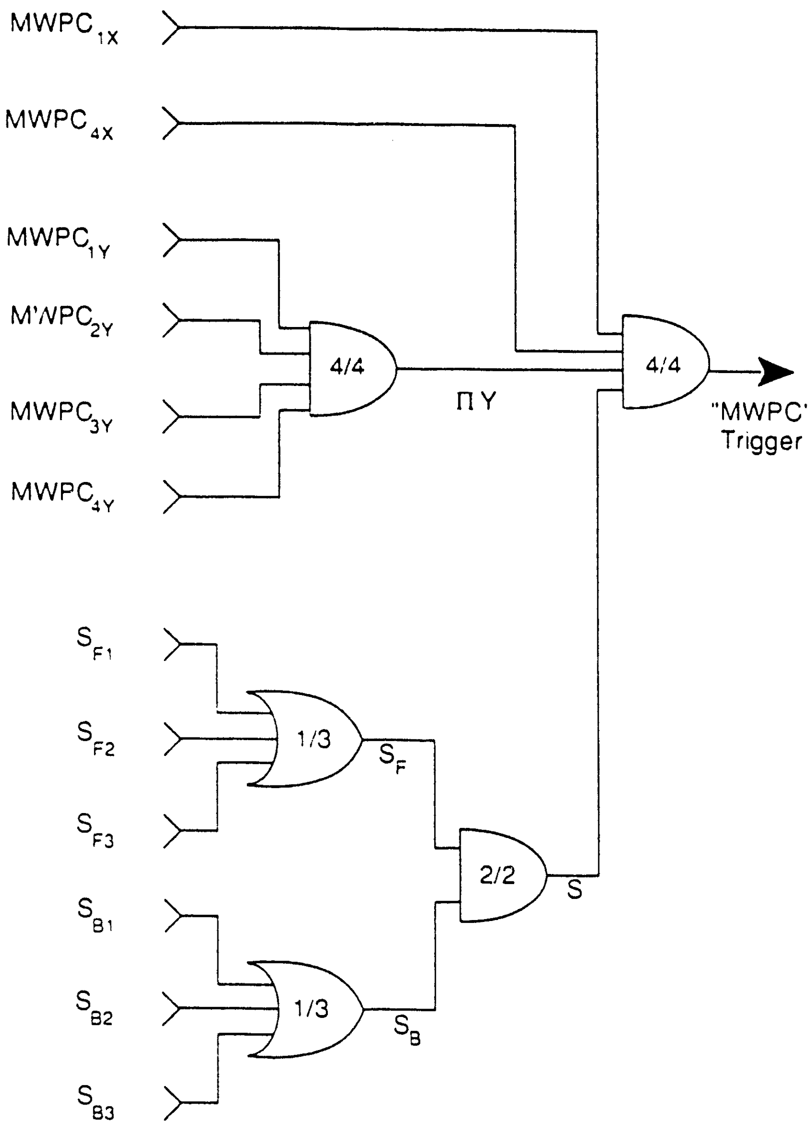


Figure 3.6: Trigger used for tuning momentum for optimal M production.

When searching for the $M \longrightarrow \overline{M}$ conversion, the full channel rate of about 10^6 μ^+/s was admitted into the apparatus. Since the rate of decay e^+ from ordinary μ^+ and M decays into the spectrometer was then so high that triggering on each one would have caused excessive dead time, a trigger with reduced acceptance for e^+ was developed. To accomplish this, the X-planes of all MWPCs were coarsely segmented into 6 groups of 24 wires each (32 wires for the first and last segment) for MWPCs 1 and 2 and into 10 groups of 40 wires each (64 wires for the first and last segment) for MWPCs 3 and 4. This allowed preselection in the trigger logic for tracks with e^--like curvature in the field by combining signals from the appropriate segment combinations in scores of coincidence gates. First, combinations of MWPC 1,2 and MWPC 3,4 segments were made in discrete logic gates. The resulting signals were then used as inputs to two "matrix logic" units, which are nothing more than a large array of coincidences between a set of horizontal and a set of vertical inputs with the topology of a matrix. At each row/column pair, the specified coincidence may be selected or deselected by a simple DIP-switch. This trigger, which will be referred to as the "matrix trigger", is then also the reason for tilting the MWPCs and the C-magnet to the same angle as the target: to decouple the measurement of the y-coordinate (used to determine the position of decay along the axis normal to the target) from the x-coordinate, along which the $e^\pm$ were magnetically deflected.

This matrix trigger is summarized in Table 3.4, which gives the segment combinations used, and in Fig. 3.7, which shows the inclusion of the other logic signals required. The segments in the X-planes are numbered from positive to negative x as mXn, where m is the chamber number and n is the segment number on that chamber. The matrix trigger accepts about 66% of the e^- and 1.1% of the e^+ passing through the spectrometer. The determination of these acceptances will be discussed in the next chapter.

3.2.9 Detection of the Atomic Positron

As the atomic e^+ and e^- start out at rather low energies, their detection can be arranged with modest static electric and magnetic fields. At the same time, care is necessary to control the influence of stray magnetic fields from unrelated elements of

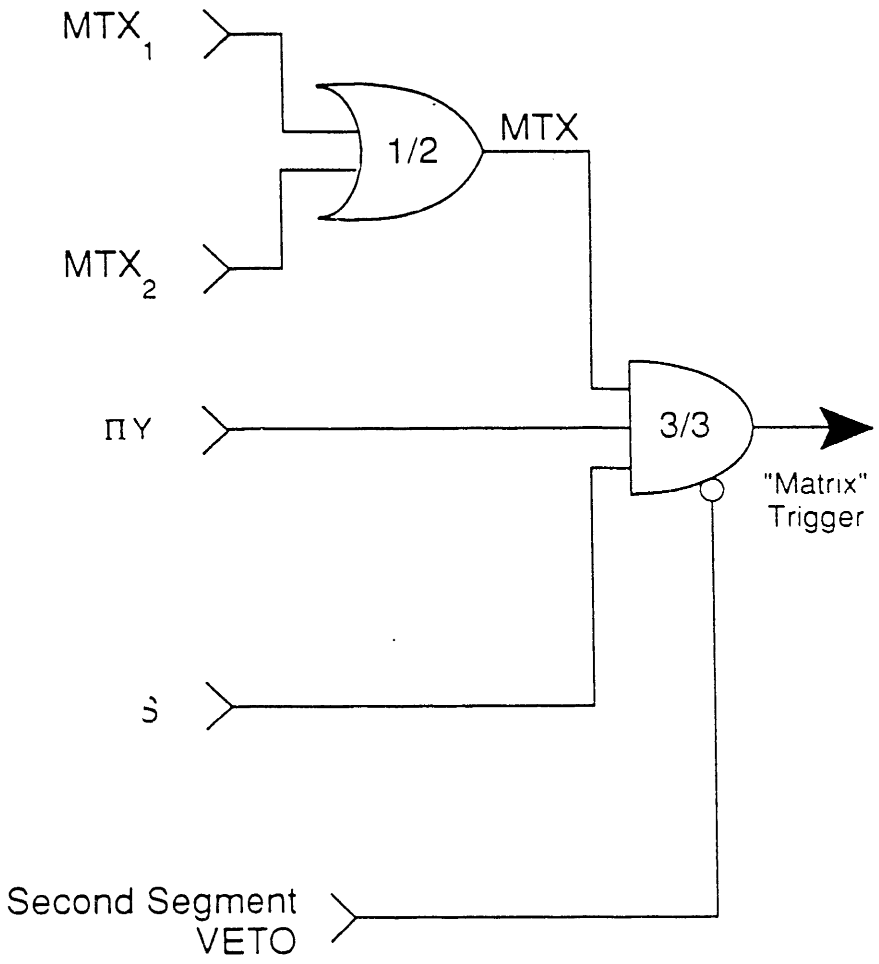


Figure 3.7: Trigger used to search for $M \longrightarrow \overline{M}$ taking the full channel rate.

	MWPC 1.2 combinations									
MWPC 3,4 combinations	1X1 2X1	1X2 2X1	1X2 2X2	1X2 2X3	1X3 2X1	1X3 2X2	1X3 2X3	1X3 2X4	1X4 2X3	1X4 2X4
3X2 4X1					✓					
3X2 4X2	✓	✓			✓					
3X2 4X3	✓	✓								
3X3 4X2		✓								
3X3 4X3	✓	✓								
3X3 4X4	✓	✓	✓			✓				
3X4 4X4	✓	✓	✓			✓				
3X4 4X5	✓	✓	✓			✓	✓			
3X4 4X6	✓	✓	✓							
3X5 4X5		✓	✓			✓			✓	
3X5 4X6	✓	✓	✓			✓	✓		✓	
3X5 4X7	✓	✓	✓			✓	✓			
3X6 4X6			✓			✓			✓	
3X6 4X7	✓	✓	✓	✓		✓	✓		✓	
3X6 4X8	✓	✓	✓	✓		✓	✓		✓	✓
3X6 4X9	✓	✓		✓						
3X7 4X8	✓					✓	✓	✓	✓	
3X7 4X9	✓		✓	✓		✓	✓	✓	✓	✓
3X7 4X10	✓		✓				✓			
3X8 4X9							✓	✓		✓
3X8 4X10			✓	✓			✓	✓		✓
3X9 4X10										✓

Table 3.4: Matrix trigger segment combinations used to select for e^- and against e^+ in the spectrometer.

the apparatus, such as the C-magnet or the last quadrupole magnet of the channel. In any case, the detection of these slow $e^{\pm}$ is simple in principle, but was non-trivial to develop into a reliable method.

To estimate the energy of the atomic e^- remaining after the decay of a M atom, we start with the 1S spatial wavefunction of M in relative coordinates:

$$\begin{aligned} \psi_{1S}(r,\theta,\phi) &= \langle \vec{r}\,|\psi_{1S}\rangle \\ &= \frac{1}{\sqrt{\pi a_0^3}}\, e^{-r/a_0} \;, \end{aligned} \tag{3.31}$$

where a_0 is the Bohr radius for muonium. This wavefunction is then transformed into momentum space by

$$\begin{aligned} \hat{\psi}_{1S}(\vec{p}\,) &= \langle \vec{p}\,|\psi_{1S}\rangle \\ &= \int d^3\vec{r}\, e^{-i\vec{p}\cdot\vec{r}/\hbar}\, \psi_{1S}(\vec{r}\,) \;, \end{aligned} \tag{3.32}$$

which may be shown to give

$$\hat{\psi}_{1S}(\vec{p}\,) = \frac{8\sqrt{\pi a_0^3}}{\left[1+\left(\frac{a_0 p}{\hbar}\right)^2\right]^2} \;. \tag{3.33}$$

To derive the momentum distribution of the atomic e^- one then calculates the overlap with momentum space plane waves

$$\begin{aligned} \phi_{\vec{p}_0}(\vec{r}\,) &= \langle \vec{r}\,|\phi_{\vec{p}_0}\rangle \\ &= \frac{1}{\sqrt{V}} e^{i\vec{p}_0\cdot\vec{r}} \;, \end{aligned} \tag{3.34}$$

or

$$\begin{aligned} \hat{\phi}_{\vec{p}_0}(\vec{p}\,) &= \langle \vec{p}\,|\phi_{\vec{p}_0}\rangle \\ &= \frac{(2\pi\hbar)^3}{\sqrt{V}}\, \delta^{(3)}\left(\vec{p}-\vec{p}_0\right) \;, \end{aligned} \tag{3.35}$$

where V is a normalization volume for the wave function. It will later cancel out. The overlap is then

$$\begin{aligned} D(\vec{p}_0) &= \langle \phi_{\vec{p}_0}|\psi_{1S}\rangle \\ &= \frac{1}{\sqrt{V}}\psi_{1S}(\vec{p}_0) \;. \end{aligned} \tag{3.36}$$

This is an example of the *sudden approximation* and is thought to be quite applicable to this case, where the μ^+ that was at the center of the M atom quite suddenly disappears. To obtain a probability distribution from the overlap, $D(\vec{p})$, we calculate

$$\begin{aligned} P(p)\,p^2\,dp &= \int_{4\pi} d\Omega\,|D(\vec{p})|^2 \\ &= \frac{1}{V}\,\frac{256\pi^2 a_0^3 p^2\,dp}{\left[1+\left(\frac{a_0 p}{\hbar}\right)^2\right]^4}\,, \end{aligned} \tag{3.37}$$

where the integral over any angular dependence (the raw distribution is isotropic!) has been taken, leaving a distribution purely in the magnitude of the momentum. This distribution may be translated into an energy distribution according to

$$\mathcal{P}(E)\,dE = P(p)\,p^2\,dp\,, \tag{3.38}$$

which results in

$$\mathcal{P}(E) = \frac{1}{V}\,\frac{256\pi^2 a_0^3 m_e\sqrt{2m_e E}}{\left[1+\left(1+\frac{m_e}{m_\mu}\right)^2\frac{E}{R_\infty}\right]^4}\,, \tag{3.39}$$

where the non-relativistic relation between energy and momentum has been assumed and R_∞ is the Rydberg constant for infinite nuclear mass. This distribution is very interesting. It has a peak at

$$\begin{aligned} E_0 &= \frac{1}{7}\,\frac{R_\infty}{\left(1+\frac{m_e}{m_\mu}\right)^2} \\ &\approx 1.9\ eV\,, \end{aligned} \tag{3.40}$$

but has an expectation value of

$$\begin{aligned} \langle E\rangle &= \frac{\int_0^\infty E\,\mathcal{P}(E)\,dE}{\int_0^\infty \mathcal{P}(E)\,dE} \\ &= \frac{R_\infty}{\left(1+\frac{m_e}{m_\mu}\right)^2} \\ &= 7\ E_0 \\ &= 13.5\ eV\,, \end{aligned} \tag{3.41}$$

as might be expected. A plot of the normalized distribution

$$\frac{\mathcal{P}(E)\,dE}{\int_0^\infty \mathcal{P}(E)\,dE} = \frac{16}{\pi}\,\frac{\sqrt{\left(1+\frac{m_e}{m_\mu}\right)^2 \frac{E}{R_\infty}}\left(1+\frac{m_e}{m_\mu}\right)^2 \frac{dE}{R_\infty}}{\left[1+\left(1+\frac{m_e}{m_\mu}\right)^2 \frac{E}{R_\infty}\right]^4} \tag{3.42}$$

is shown in Fig. 3.8.

The atomic e^+ , thus distributed in energy and distributed in space as the $\overline{M}$ atoms would be, must be extracted with maximum efficiency. This is accomplished by an arrangement of electrodes that collect, focus, and accelerate these e^+ up to a kinetic energy of 5.7 *keV*. A vertical section through this structure is shown in Fig. 3.9. For orientation, the beam counter and target are included. After being thus accelerated, the e^+ would further be transported and focused by a series of coils wrapped axially onto the beam pipe, deflected by 60° in an iron-free dipole magnet, and finally focused by a solenoid onto a 75 *mm* diameter chevron-pair microchannel plate detector (MCP).

The geometry for the 11 electrodes was designed by running Monte Carlo simulations of the transport of electrons through the electric field imposed by various choices of potentials on several possible electrode arrangements. The basic spirit of the design was to find the optimum geometry for a three-stage device: collection, shaping and acceleration. The best results for the collection efficiency of the atomic e^+ from the point of $\overline{M}$ decay to the exit grid of the electrode structure were around 60%.

The further magnetic transport was simulated to give a maximum overall geometrical acceptance of the atomic e^+ detection of 35%. The actual acceptance measured in the experiment included the detection efficiency of the MCP as a factor. Direct verification of the simulated results of low-energy e^+ transport is not possible, as the literature gives widely scattered values for the MCP efficiency for $e^\pm$ at a few *keV* of kinetic energy. Determination of the experimental acceptance gives about 16% and will be discussed in the next chapter. One may suggest, however, that a detection efficiency for the MCP of the order of 50% is not unreasonable. Stray magnetic fields in the apparatus, mainly from the C-magnet, are not included in the simulation, however, so there is freedom for the actual transport efficiency to

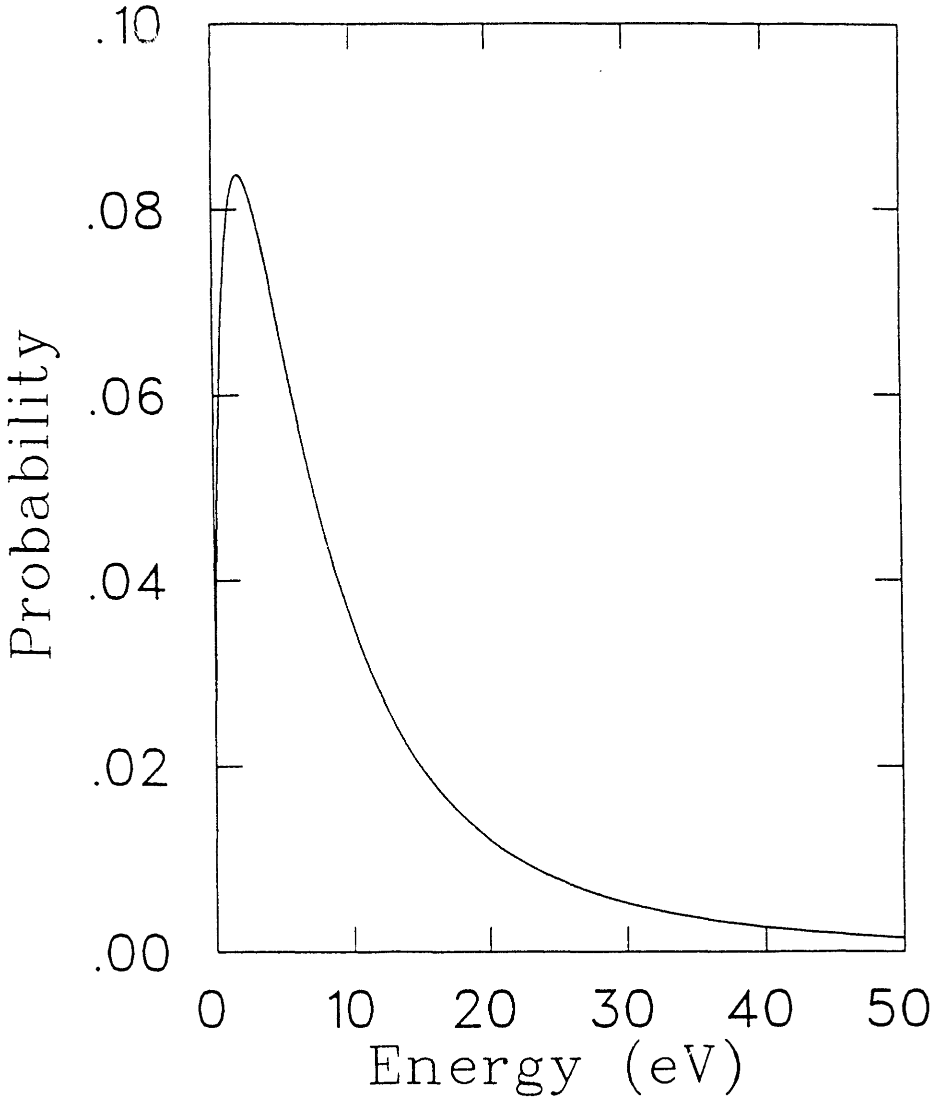


Figure 3.8: Energy distribution of the atomic e^+ (e^-) after the decay of μ^- (μ^+) in M ($\overline{M}$).

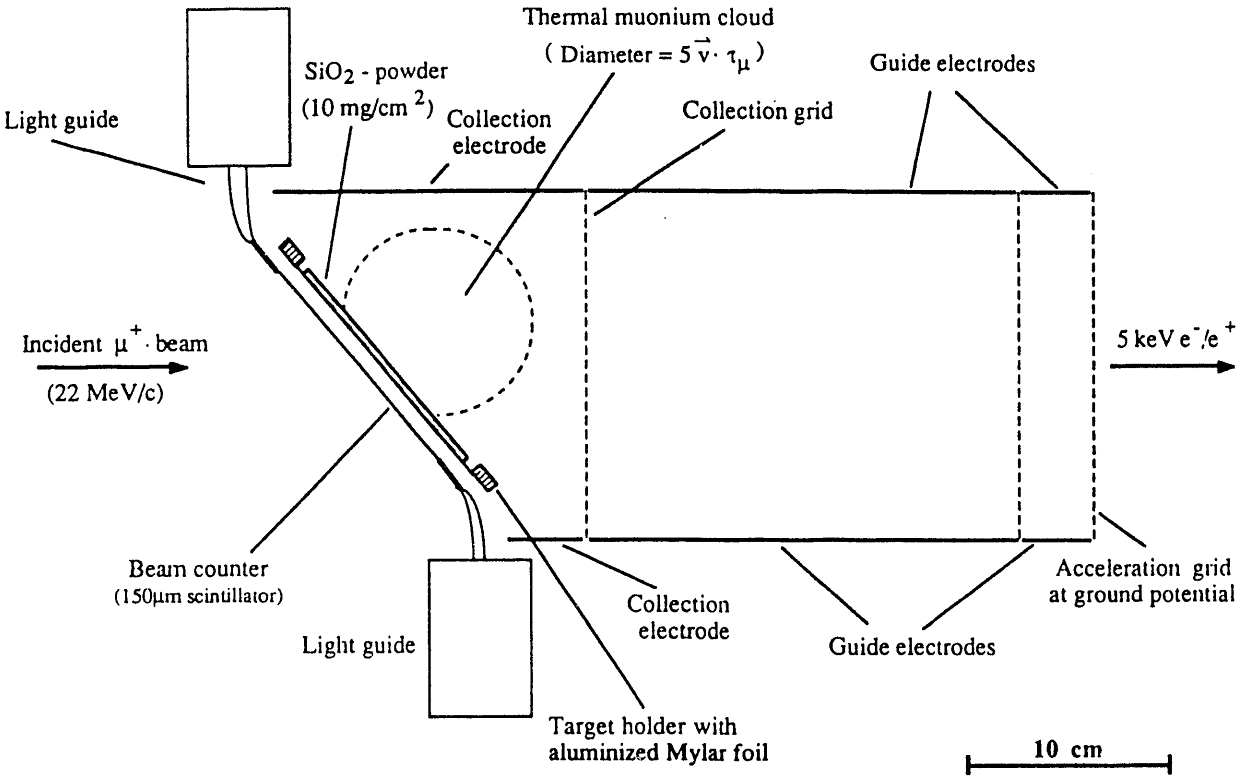


Figure 3.9: Vertical section through the electrode structure for extraction of the atomic e^+ of $\overline{M}$ decay or, when polarity-reversed, of the atomic e^- from M decay.

lie below the calculated 35% if the MCP efficiency is larger than 50%.

The first stage of the electrode structure was to collect as many of the widely distributed atomic e^+ as possible toward the second stage, while trying to compress their transverse spread. The voltage drop from the target to the vertical wire grid plane delimiting this region was from $-5.7\ kV$ to $-3.2\ kV$ over an average of 10 cm (the distance from the center of the target to the first grid). The next stage, of 20 cm length, served as a shaping region in which the potentials applied from the sides controlled the transverse envelope of the distribution of e^+ and only 900 V were dropped in this region. The final stage was only 3 cm long, but the potential falls from 2.3 kV on the second grid to ground at the exit grid, and the purpose of this region was the final acceleration of the e^+ to the kinetic energy of 5.7 kV. The transverse geometry of this small accelerator was square with interior dimensions of 15.3 cm.

The structure was made up of rectangular pieces of Cu-clad G10, where the Cu coating had been milled off over a margin at the edges to separate the conductive surfaces of neighboring electrodes. The in-line grids were strung with 20 μm Au-plated W wires at 2 mm spacing, giving a 99% transmissivity per grid. The potentials on the various electrodes were under computer control *via* a CAMAC LeCroy 2132 interface to the LeCroy 4032A programmable high-voltage supply bin. The target was maintained at 5.7 kV, where a 12 line-per-inch electroformed mesh placed on the surface of the SiO_2 target served to define the potential on this otherwise insulating surface.

The magnetic elements of the transport system included steering coils with axes perpendicular to the beam-line placed around the M production region and axial focusing coils at three points after the exit grid of the accelerator. These were capable of producing fields up to typically $10-20\ G$. The bending magnet consisted of wedge-shaped windings mounted in a non-magnetic Al frame and were operated at about 16 A, which gave a vertical field of about 15 G at its center, bending the e^+ by 60° in the direction of the solenoid. The absence of a yoke gave large fringe fields whose effects were compensated by two of the focusing coils, placed before and after the bender. The aperture presented by the vacuum pipe in this region

gap	MWPC gas	1.76 cm
window	Al	8.9 μm
	mylar	76 μm

Table 3.3: Material composition of large MWPCs.

was rectangular, 20 cm wide and 15 cm high. The last element was a 117 cm long, 30 cm diameter yoke-less solenoid that generated an axial field of 11 G with 2.22 A in the windings (462 turns/m). The fringe fields were shortened by wrapping two layers of 0.1 mm thick, 38 cm wide CONETIC shielding foil around the ends of this solenoid. The shortened fringe field at its exit had the effect of focusing slow e^+ onto the MCP, placed about 25 cm downstream of the last turn. Figure 3.10 shows the longitudinal field of this solenoid as a function of position along the solenoid axis.

With beam in the apparatus, we only had atomic e^- available for testing this transport system. It was with atomic e^- that the overall acceptance was measured, with polarities of all elements reversed from those used for detecting e^+ . But, in order to verify that the system was capable of transporting slow e^+ , a source of slow e^+ was constructed.

The e^+ from a sealed, in-vacuum ^{22}Na β^+ source of about 3 mCi, placed above the location of the target, were directed onto a "Venetian blind" arrangement of W foil strips, mounted in a frame located in the same target holder that was used for the SiO_2 targets. The effect utilized is the reemission of a fraction of these e^+ at energies on the order of an eV [Che85,Ver83,Can82,Lyn80,Dal80,Lyn77].

For best e^+ reemission efficiency with reasonable effort, the 25 μm thick and 5 mm wide W foils were cleaned by degreasing in a Freon ultrasound bath for 15 minutes, then etching in a 1:1:1 mixture of 30% H_2O_2, 1-molar $NaOH$, and 1-molar NH_4OH for about 10 minutes, and finally thorough rinsing in distilled H_2O. After thus removing impurities from the surface of the foils – which showed a matte finish after this procedure – they were annealed at 2100°C for two hours in a vacuum of better than 10^{-5} $torr$. The effect of this heat treatment is to enlarge the single-crystal regions ("grains") of the polycrystalline foil from sizes of less than 1 μm to as large as 20 μm, as verified by photomicrographs.

Large e^+ reemission yields (10^{-3}) are seen with thin (100 μm) single-crystal W foil moderators, but require UHV conditions to ensure surface cleanliness. For our purposes, the grain-enlarged polycrystalline W served quite well, as the vacuum conditions were not better than 10^{-7} $torr$. Though it was not measured, the

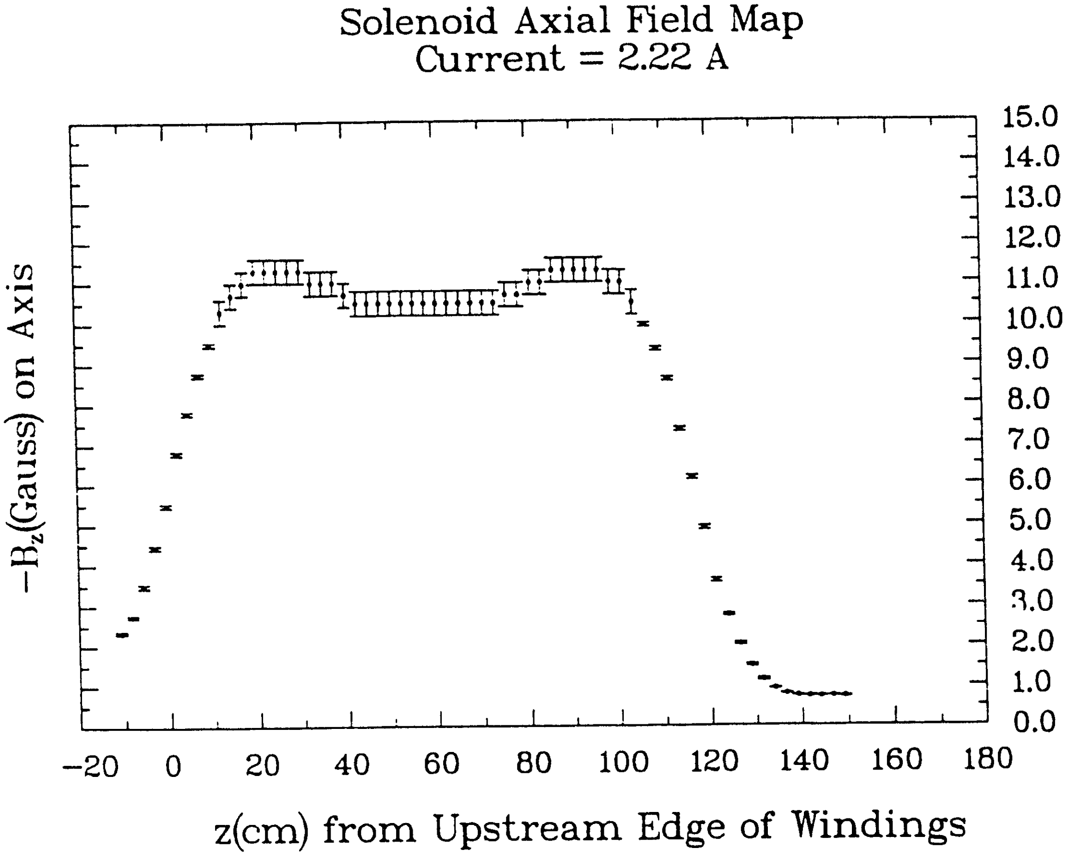


Figure 3.10: Axial Field of solenoid used for transport and focusing of slow $e^{\pm}$ onto the MCP. The effect of field clamps at the ends is apparent.

backward reemission yield for our foils was likely on the order of 10^{-5} per incident e^+.

After much adjustment of the electrostatic and magnetic elements of the transport system, a fully tunable rate of 300 s^{-1} of e^+ was successfully transported from the W foil source to the MCP. Most importantly, the tune developed for the system also transported secondary e^- *optimally* when the polarities of all voltages and currents were simply reversed. The secondary e^- are due to the impact of primary e^+ from the source on the W foils.

As the tuning curves of the slow e^+ for all elements of the transport system agreed with those for e^-., the conclusion drawn was that any atomic e^+ left after $\overline{M}$ decay would be transported with the same efficiency as that which may be measured by the detection of the atomic e^- of M.

3.2.10 The Data Acquisition System

Data in this experiment were recorded through a CAMAC interface by a VAXStation II/GPX mini-computer equipped with 5 MB of memory, two disk drives (71 MB and 120 MB), and two Kennedy 9100 9-track tape drives. The online software was based on the LAMPF-standard "Q-system" [Shl74], with extensions specific to this experiment. The capabilities included:

- Cumulating counts of detector singles rates and coincidence signals in the trigger logic to offer rate information used to monitor the performance of the apparatus.

- Controlling the voltages applied to the electrode structure that accelerates slow $e^{\pm}$.

- Offering histograms of the wires struck in the MWPCs and the projected decay origins obtained from these, of the time-of-flight (TOF) between tracks in the spectrometer and the MCP, and of the pulse-height distribution in the $NaI(Tl)$ crystal. Also, histograms with some simple cuts were available for focusing on the effects needed to tune beam momentum and slow $e^{\pm}$ collection.

- Sampling individual events from the data stream and displaying their attributes, including the spectrometer hits, TOF value, and the $NaI(Tl)$ pulse height in a diagram of the apparatus.
- Reading the wires struck in the MWPCs upon receipt of the appropriate trigger. The trigger also caused the readout of the TOF and the $NaI(Tl)$ pulse-height correlated with the triggering MWPC track. This trigger will be discussed in more detail below.

The principal triggers used were presented in the previous section, however without reference to the details of ensuing response of the data acquisition system that must read out the information in the MWPCs. Each wire on every chamber plane is equipped with a preamplifier, discriminator, and a monostable multivibrator ("one-shot") that is triggered by a threshold-crossing pulse on the wire. The one-shots held a logic-true level after such a triggering for ~ 50 ns. These signals were also promptly available for other trigger logic uses and were nothing more than the direct outputs from the wire segments discussed above ("segment FAST-ORs"). Additional one-shots ("MONO-OR") in the chamber plane control electronics held true levels for ~ 400 ns after any wire in that plane was hit. During this time a trigger decision was made in external logic.

Thus, to decide whether or not to read out the wires struck in the MWPCs for a given particle passage through the chambers, the FAST-OR outputs were used together with pulses in the plastic scintillator planes to decide whether or not the event was worth recording. For the MWPC trigger, this simply meant that something must have passed through all chambers. In the case of the MATRIX trigger, the additional requirement was for a magnetic deflection in the direction consistent with negative charge on the particle, as discussed above. The resulting trigger signal was provided to the chamber plane controller electronics at the trailing edge of the MONO-OR to cause the present wire hits to go into an indefinite holding state. No further wire hits could be recorded until the present event data had been read out. The trigger signal also started the chamber readout handshaking between the chamber plane controllers and the word buffers in the CAMAC crate and set an external latch indicating a busy state of the MWPC readout. This busy signal disabled any further triggers until it was reset. After the last bit of MWPC information was sent to the CAMAC crate (in a serial, bi-phase code) and the TDC (TOF) and ADC ($NaI(Tl)$) were read, a signal indicating this completion cleared the busy latch, the chamber plane latches, and the CAMAC word buffers for the next event.

3.2.11 Summary of Online Data Taking

The data stream from the CAMAC crate was buffered in a Bi-Ra Systems Microprogrammable Branch Driver (MBD), which then dumped its data to the MicroVAX when time was available. From there, the event data was both taped directly and partially processed for online presentation. The online analysis guided progress through different steps in the calibration and data taking:

1. Tuning the beam momentum on a new target for optimum M production per incident μ^+ . The number of spectrometer tracks originating from the vacuum region downstream of the powder target normalized to the counts in the beam scintillator was used to judge the rate of M formation. To somewhat suppress the effects of multiple scattering on the reconstruction, a cut on the pulse height in the $NaI(Tl)$ crystal eliminated all events with less than 25 MeV energy. The mean-square scattering deflection diminishes with increasing energy. The requirement that the trajectory passed through the vacuum window was also imposed, to eliminate any decay e^+ that scatter heavily in passing through the thicker walls of the vacuum chamber or the flange holding the thin window.

2. To determine the absolute acceptance of the spectrometer, the incident beam momentum was lowered until the reconstruction showed that all of the incident μ^+ were stopping in the beam scintillator. It was then concluded that none were passing through the scintillator and target. In other words, all of the incident μ^+ counted by the beam scintillator were stopped within the field of view of the spectrometer. Thus, the total number of tracks seen in it divided

by the number of beam scintillator counts gives the measured spectrometer acceptance.

3. The acceptance for detecting atomic $e^{\pm}$ was found by comparing the number of M counts in the reconstruction of the decay origin as used for momentum tuning to the same histogram with the additional requirement that there be a TOF count in the time window placed around the observed peak. The ratio of the latter to the former is a direct measure of the TOF acceptance if the background in the TOF distribution is small compared to the number of correlated counts.

4. It was necessary in the deduction of the result for the conversion probability per atom to know the acceptance of the MATRIX trigger relative to the MWPC trigger for both e^+ and e^- . To cause the copious e^+ to act as e^- in the spectrometer, the field in the C-magnet was reversed in polarity. Since the magnet was operated far below saturation of the yoke, a completely reversed field was obtained by simply changing the direction of the current and maintaining the magnitude of the shunt reading.

The most important point to make about procedure is that the polarities of the C-magnet, of the voltages on all acceleration electrodes for the slow $e^{\pm}$, and of all magnets in their further transport to the MCP were always reversed together when switching from tuning of M production and measurement of atomic $e^{\pm}$ acceptance (in which the important tracks in the spectrometer are those of the Michel e^+) to the conditions in which evidence for $\overline{M}$ was sought (where e^- tracks in the spectrometer were of interest). This was to ensure that the atomic e^+ from $\overline{M}$ decays would be detected with the same efficiency as that with which the atomic e^- from M decays were seen. Any fringe field effects from the C-magnet on the transport of the atomic $e^{\pm}$ then always contributed with the same polarity. Also, any small transverse misplacements of the MWPCs in the spectrometer, which would render the acceptance for e^+ and e^- different for a given C-magnet polarity, did not have an effect when the polarity of the C-magnet was always adjusted to give the same direction of deflection for the tracks of interest.

The next chapter presents in detail the data taken for each of the points enumerated above and describes how the relevant information was extracted.

Chapter 4

Data and Analysis

Though this experiment is simple in principle, extracting a result for the conversion probability per atom for $M \longrightarrow \overline{M}$ proved to be rather involved. In this chapter, I will describe in detail what data were taken to establish the acceptances of the detectors, to verify the production of thermal M, to calibrate the acceptance of the $\overline{M}$ signature and measure the background to it, and to search for the $M \longrightarrow \overline{M}$ conversion.

All stages of the data analysis depended on the development of an effective algorithm to fit the particle tracks through the spectrometer MWPCs. For the alignment corrections to the coordinates of the MWPC hits, the tracks were measured with the spectrometer magnet shut off and thus fit to straight lines. During further data taking, the spectrometer field was maintained at a constant level suitable for the separation of Michel distributed decay e^+ from decay e^- with simultaneous broad-range momentum acceptance. Tracks recorded during this phase of the experiment were fit to a track model based upon the measured field-map of the spectrometer magnet and a fourth-order Runge-Kutta forward integration of the equations of motion of a charged particle in this field. This algorithm was used to identify decay e^+ tracks from the M atom decays as well as to search for decay e^- tracks from $\overline{M}$ atom decays.

The final presentation of the data sample in which evidence of $\overline{M}$ decays was sought is a one-dimensional position histogram displaying the origin of a spectrometer track as projected onto an axis perpendicular to the powder target at its center. In this histogram, one identifies decays of μ^+ from the beam-counter and the powder target and M decays from the target and the vacuum region downstream of it. In the event of a non-zero $M \longrightarrow \overline{M}$ conversion coupling, one would also expect to see decays of $\overline{M}$ atoms in the vacuum. A maximum likelihood fit was developed to obtain the contribution of each of these processes to these data.

4.1 Field Map of the Spectrometer Magnet

Critical to this experiment is the ability to discriminate between tracks of positively and negatively charged particles observed in the spectrometer. Placing a dipole magnetic field between the second and third of four MWPCs is the "classical" arrangement for a spectrometer in rectangular coordinates. It allows for the determination of the particle charge by simply comparing the incident and outgoing asymptotic directions of its trajectory and for the reconstruction of a two-dimensional projection of the origin of the particle track. With more effort, one may estimate the probability that a given set of hits in the MWPCs represents a track due to the passage of a single particle. This means one must fit the observed MWPC hits to a model of a single particle track, a procedure that yields the momentum of the particle as one of the determining parameters. Clearly, reliable knowledge of the field of the spectrometer magnet is central to an accurate track model.

4.1.1 Field Map Data

The three Cartesian components of the magnetic field of the C-yoke dipole magnet were mapped on a rectilinear grid of $17 \times 10 \times 81$ position points (see Fig. 4.1). There were 17 points with 2.54 *cm* spacing along the x-axis (the direction along which the deflection of the charged particles is expected), 10 points with 2.54 *cm* spacing along the y-axis (the direction between the pole pieces of the magnet), and 81 points with 1.27 *cm* spacing along the z-axis (the axis of the spectrometer). Under the control

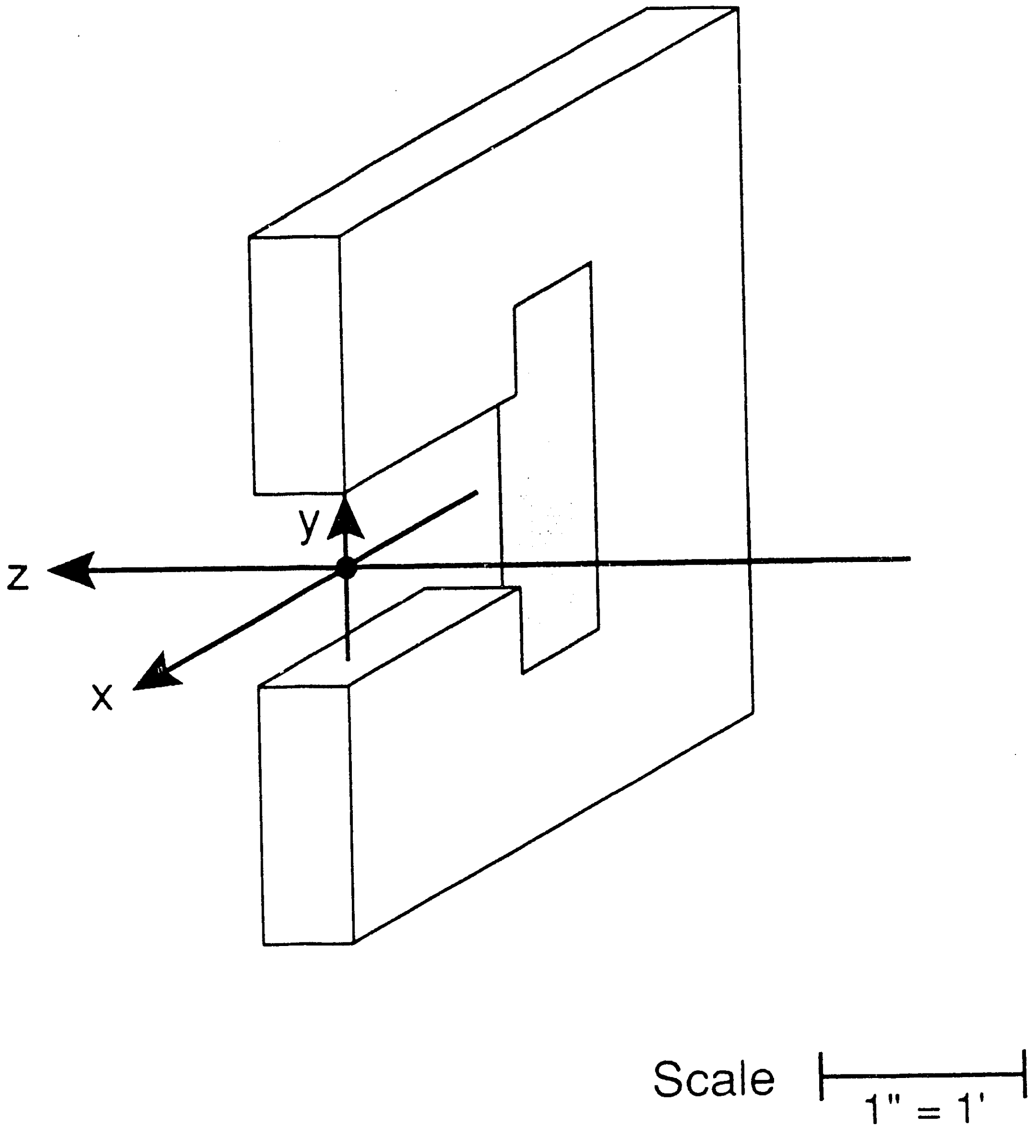


Figure 4.1: Spectrometer C-magnet with mapping coordinate system.

of a MicroVAX II computer, a Kinetic Systems CAMAC interface, and three Bausch & Lomb servo drives with 0.0025 *cm* position resolution, three Group-3 Hall probes were moved over this three-dimensional grid, sampling each field component with an absolute measurement error of less than 0.5 G at each of the 13770 grid points. A current-regulated power supply maintained a constant 200.00 A in the magnet throughout this 1-week procedure. To verify the actual current in the magnet at the time of measurement at each grid point, the voltage drop across a precision shunt (100 $mV/1500$ A, 0.25%) was recorded through CAMAC together with the position readout and the three field components.

To correct for various experimental idiosyncrasies, like left-handed coordinate systems and different coordinate systems chosen for the position grid and the field components, and to inspect the quality of the data, a replay system for the raw field map data was developed. Its capabilities include the graphical presentation of any field component as a function of one or two position coordinates, the ability to list any subset of the raw or interpolated field map data, and the possibility to check Maxwell's equations around any interior grid point in the map. Figure 4.2 shows the y-component of the field as a function of the x- and z-position with respect to the center-gap point. This field component is chiefly responsible for the magnetic deflection of charged particles in the spectrometer. Its value is 522 G at the center of the magnet gap. The behavior of the components B_x and B_z is shown in Figs. 4.3 and 4.4.

4.1.2 Taylor Expansion of the Magnetic Field

It is apparent from Fig. 4.2 that we cannot assume the field to be constant in any relevant region, so no simplifying assumptions regarding its shape can be made. Since the track model that is fit to observed particle tracks in the spectrometer relies on the ability to retrieve the value of all field components at any spatial point within the boundaries of the mapping region, it is necessary to determine the field accurately between the actual points where it was recorded. To accomplish this, one may, for example, fit a functional form motivated by the boundary conditions to the measured field data. This was judged to be an unnecessarily complex approach.

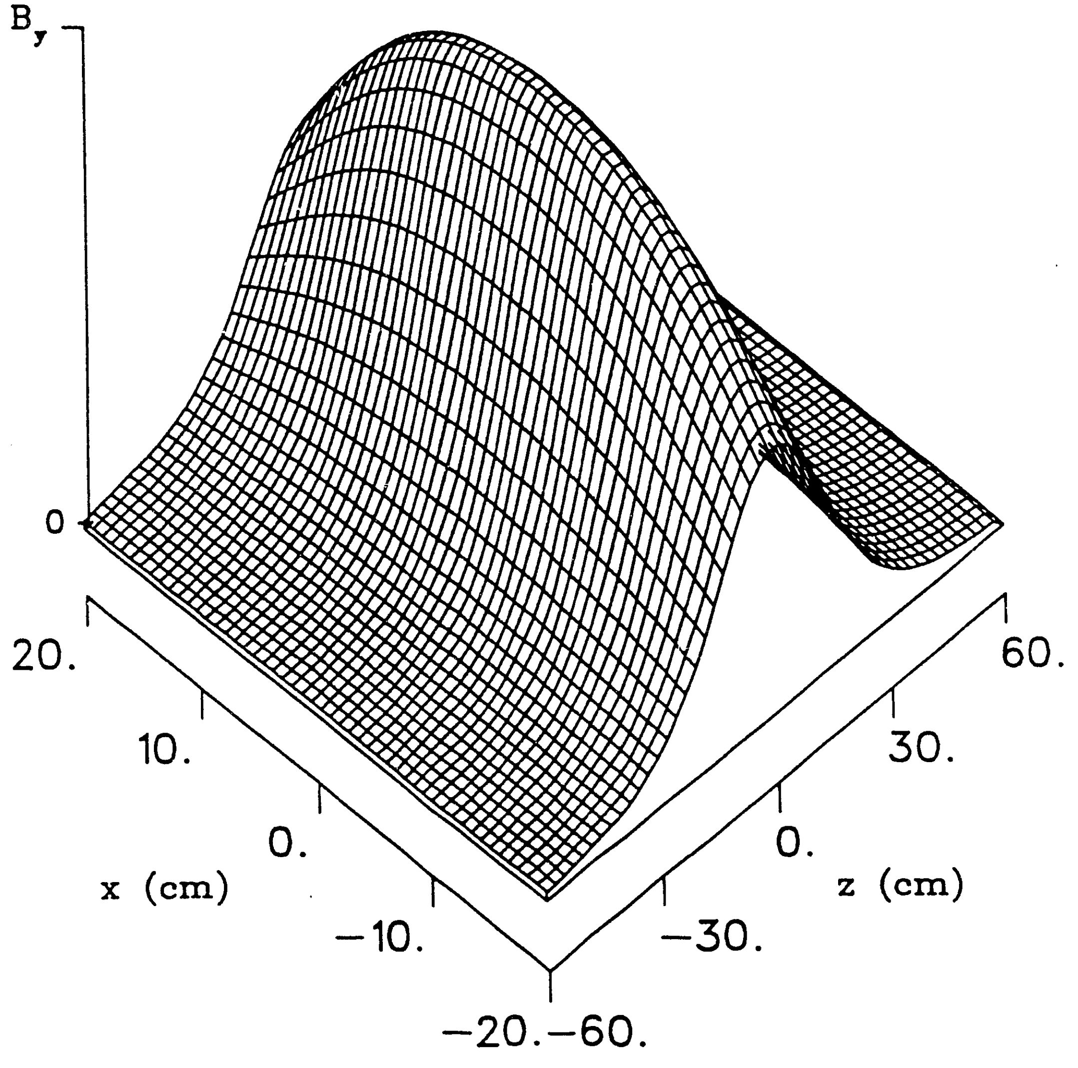


Figure 4.2: Spectrometer magnet field component B_y as a function of x- and z-position relative to the center-gap point in the mid-plane. ($B_y = 522$ G at the center of the gap.)

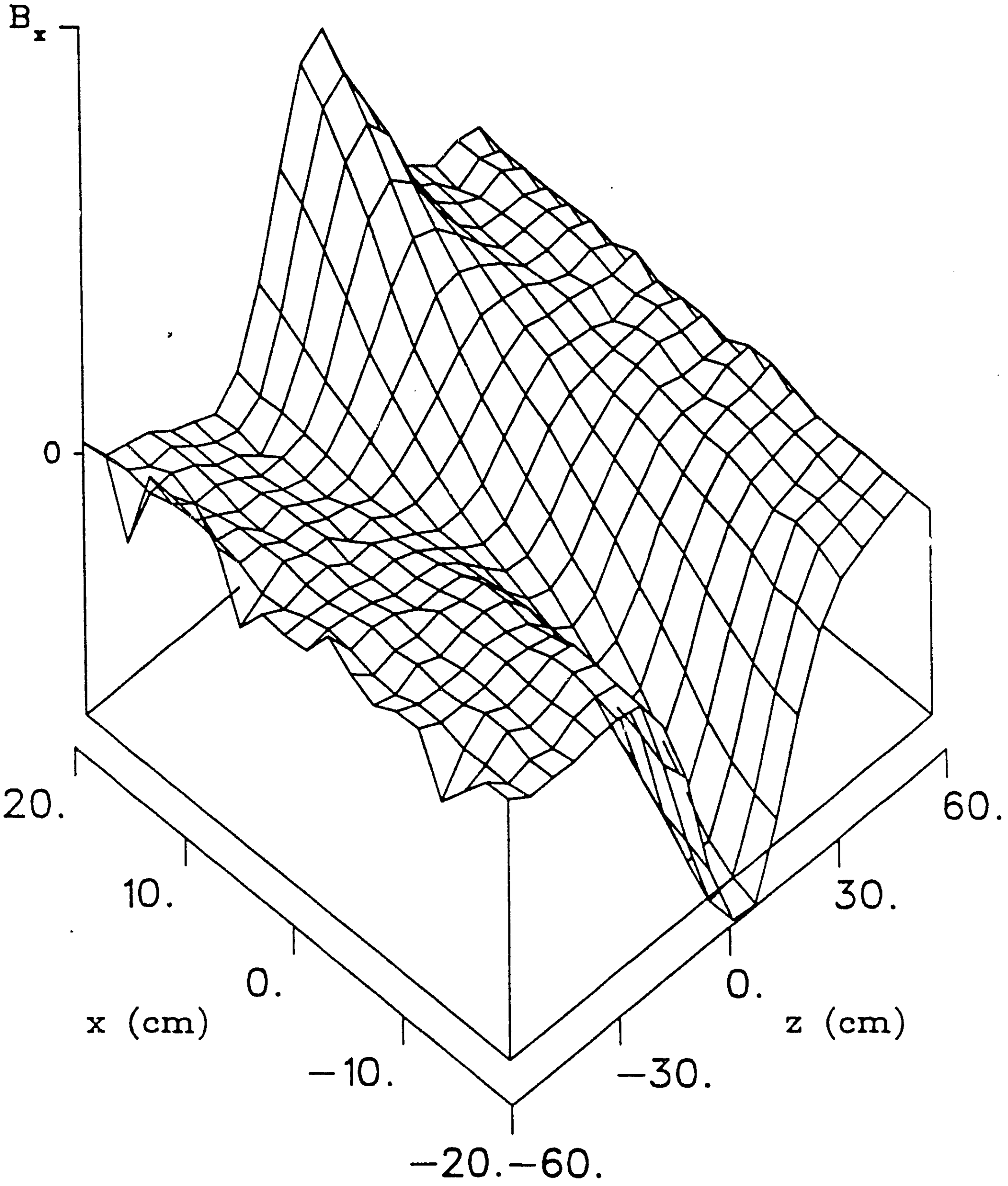


Figure 4.3: Spectrometer magnet field component B_x as a function of x- and z-position relative to the center-gap point in the mid-plane. (The magnitude of B_x is ≤ 1.7 G in this plane.)

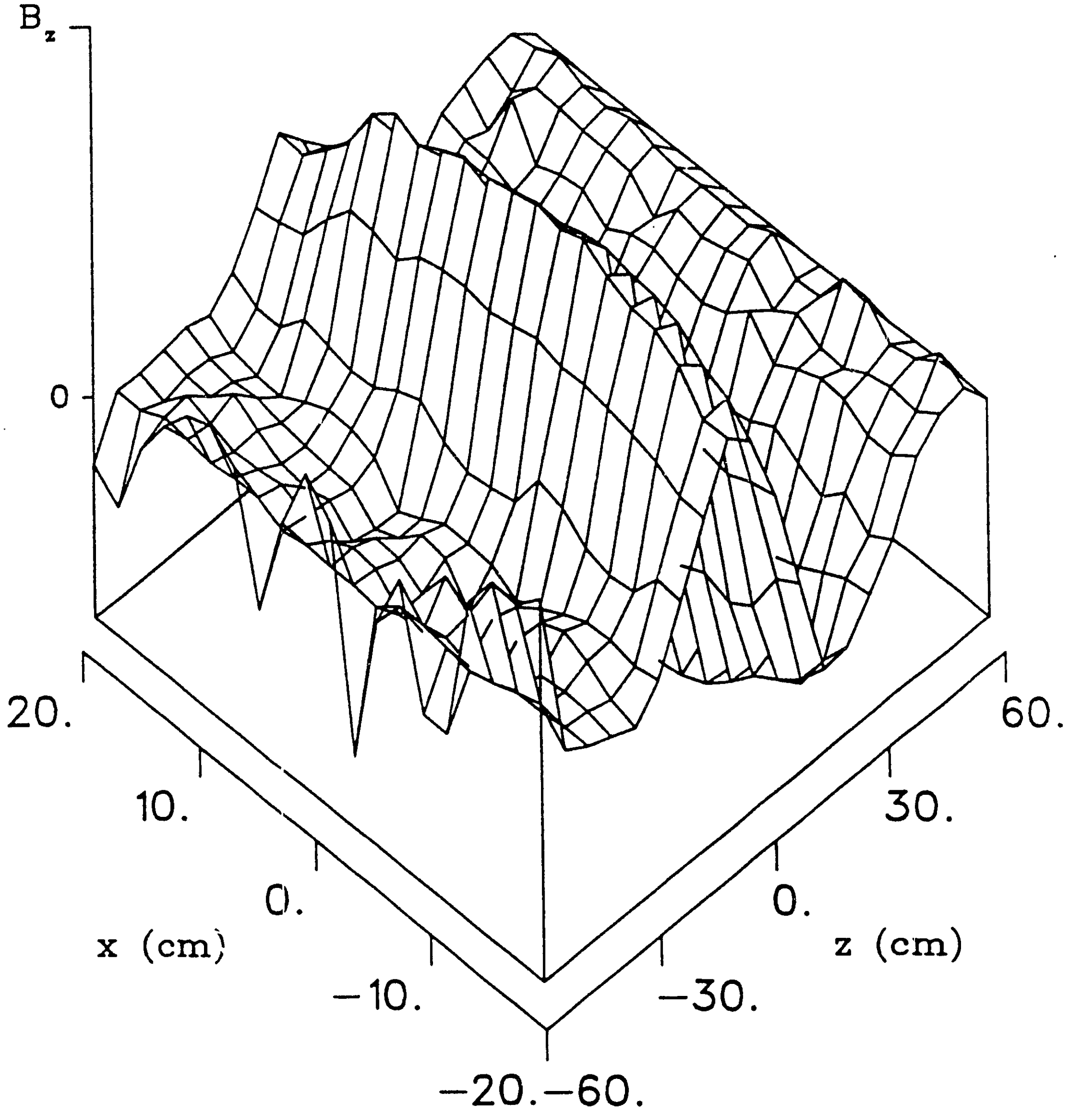


Figure 4.4: Spectrometer magnet field component B_z as a function of x- and z-position relative to the center-gap point in the mid-plane. (The magnitude of B_z is ≤ 1.7 G in this plane.)

Fitting an *ad hoc* empirical choice of function was deemed hopeless if the range of validity was to be the entire mapping grid. Rather, the most attractive choice is the local expansion of the magnetic field in a second order Taylor series for each field component about each grid point of the map.

Appendix C details the derivation of the formulae involved in the expansion of each field component in a second-order Taylor series in the three spatial coordinates around each point in the field map. The field map data at the central point of the expansion and at all nearest and several next-nearest neighbor points in the grid is used to determine the expansion coefficients. Since the magnetic field must obey Maxwell's equations, one finds constraints relating the parameters of the Taylor expansion. But, as the field map components all suffer from a measurement error of about 0.5 G and are not all simultaneously large enough to keep the relative errors negligible, small deviations from Maxwell's equations were found. Another source of error in the field map data may be slight nonorthogonalities in the positioning of the Hall probes. This is not evident from the field data, however, as the largest component, B_y, does not appear to mix in to the other two components near the center gap point of the magnet, where such an effect would be expected to be largest. One might devise a method by which the field measurements may be corrected in a globally consistent and convergent fashion to accord to Maxwell's equations. The particles of interest in the spectrometer are the Michel-distributed e^+ and e^- at several tens of MeV/c. Since the small fields and the fluctuations in their measured values that cause the discrepancies with Maxwell's equations lead only to negligible deflections of these particles, this correction was not attempted.

Rather, the full complement of 30 expansion coefficients for each field component around a given field map grid point was retained. The field at any space point is then obtainable from the three-dimensional second-order Taylor series in the coordinates using the expansion coefficients calculated for the grid point nearest to this point. This amounts to a parabolic interpolation in the space coordinates. One might like to be able to *extrapolate* beyond the limits of the field map, say, further toward the pole pieces than the closest grid points. Since the magnetic field at the pole piece surfaces and edges and near the field clamp edges is expected to vary strongly with position, such an extrapolation is probably invalid. Therefore, this is not permitted and the edges of the field map constitute an effective limit on the "aperture" of the spectrometer in the y-direction.

4.2 Fitting the Spectrometer Tracks

With a complete and reliable field map, one is prepared to construct a model of the tracks observed in the spectrometer. It is this model that constitutes the "fit function" that is used to extract the determining parameters for a measured track *via* an appropriate fit algorithm.

Before attempting to fit a track, however, it is prudent to carry out some simple tests on the MWPC hits that are purported to be due to a single particle. These checks will intercept events with MWPC hits that under no reasonable circumstances can constitute a track due to the passage of a single particle. Such events may then be skipped without expending much computing time, after which one would only find that the χ^2 for this "track" is very large. These simple ***pattern recognition*** criteria are:

1. The event may have no more than one missing hit in the x-dimension and no more than one missing hit in the y-dimension. This is to ensure sufficient degrees of freedom so the track fit does not become under-constrained. A certain amount of recovery is possible, so in the event of up to one missing hit in either dimension, this is attempted. For the small MWPCs (1 and 2) this is accomplished by assuming a virtual point on the track to lie in the center of the vacuum window and extrapolating to the plane(s) for which a missing hit is to be recovered. For the large MWPCs (3 and 4), the u-planes are used to restore a missing x- or y-hit for a given chamber (but not both!) and to ensure that the x, y hit pair is consistent with the u-hit.

2. If the product of the hit multiplicity per plane over all chamber planes is larger than 2, the event is deemed too complex to analyze and is discarded. This means that if there are more than 2 possible combinations of hits to

characterize a track the event is skipped.

3. The straight-line projection in the yz-plane from y-hits in MWPCs 1 and 2 onto MWPC3 must lie somewhere on the active y-area of this chamber.

4. Similarly, the straight-line projection in the yz-plane from the y-hits in MWPCs 3 and 4 onto the z-location of MWPC2 must give a y-coordinate that lies on the active area of this chamber.

5. The straight-line asymptotes in the xz-plane of the track before and after the magnetic field region must cross in this field region when projected toward the center of the magnet. The asymptote before the magnet is approximated by assuming a line through the x-hits in MWPCs 1 and 2 and that after the field by placing a line through the x-hits in MWPCs 3 and 4.

For up to two possible combinations of chamber hits, the above criteria are tested. If neither passes them all, the entire event is skipped. If one passes, then it is delivered to the fitting routine. If both pass, then both are fitted and the better result (the fit with the smaller value of χ^2) is chosen for the track.

4.2.1 The Track Model

Any track in the spectrometer will be fully specified by five parameters: the momentum, the initial x- and y-positions, and the initial direction cosines along the x- and y-axes. It is possible, under certain circumstances, to develop analytic expressions that take the positional hits in the MWPCs as arguments and give as a result, for example, the momentum of the particle that produced the track [Win74]. These methods use the track model to "train" the parameters of the analytic expressions toward the best values for the most general validity. Because for such simulated tracks the determining parameters are known, the coefficients in the expressions for them may be found. This approach was attempted for our spectrometer without success. The reason appears to be that the phase space acceptance of the spectrometer is too large to be described by any tractable analytic expression or series expansion. This conclusion is indicated after having obtained a successful expression for the track momentum in terms of the MWPC hits when a restricted phase space of incident particles was assumed. All attempts to generalize this to the full acceptance of the spectrometer failed. The trouble ought not to be due to the inhomogeneous magnetic field in the spectrometer, as dealing with this complication is claimed to be the design objective of the method [Lec69].

The track model for this experiment is based directly on the equations of motion for a charged particle in a magnetic field. These equations are numerically solved by a 4th order Runge-Kutta method which uses the interpolated field map information described in the previous section. Details of this algorithm are presented in Appendix D. The model was tested by applying it to the transport of electrons through a uniform field region, for which the motion may be analytically determined. The step parameter in the Runge-Kutta integration was adjusted to the largest value that still gave positional agreement with the analytically obtained track to better than 0.1 *mm* over its entire length.

4.2.2 Least-Squares Algorithm

Armed with a working track model, a suitable algorithm must be chosen that yields the desired track parameters for a given event in the spectrometer. Assuming that the measurement errors of the MWPC hits are Gaussian distributed, one is led to adopt a suitably defined χ^2 as the estimator of track quality. The strategy, then, is to estimate the track parameters for a given event, optimize these parameter values using a least-squares minimization algorithm, and conclude that the parameter values at this optimum are the best possible characterization of the particle track for this event.

The actual measurements of x and y in MWPC1 were used to estimate the starting point of the particle. Since the field of the C-magnet between MWPC1 and MWPC2 was about 15 G on average and only insignificant deflection of the particle is expected in this region, a straight line between the hits in these chambers was used to give the initial guess at the x- and y-direction cosines of its track:

$$\begin{aligned} c_x &= \frac{(x_2 - x_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \\ c_y &= \frac{(y_2 - y_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} . \end{aligned} \tag{4.1}$$

The third direction cosine, c_z, is, of course, fixed by the condition

$$c_x^2 + c_y^2 + c_z^2 = 1 . \tag{4.2}$$

To determine the charge of the particle, the in- and out-going directions at the field region are compared. The charge is simply the sign of the difference of the x-direction cosines before and after the field region. There is no danger of charge misidentification since the decay $e^{\pm}$ at the top edge of the Michel distribution (52.83 MeV/c) will still be separated by about 10 cm at MWPC3 and about 18 cm at MWPC4. Multiple scattering by a total of about 4° in the direction opposite that of its magnetic deflection for a Michel e^+ is necessary for it to appear to have negative charge. Futhermore, this scattering deflection must occur mostly in a region on the trajectory that would feign the magnetic deflection of an electron. Comparing this to the approximate r.m.s. multiple scattering angle of the order of 0.4° for the portion of the spectrometer with non-negligible magnetic field, one concludes that this is a very improbable event which will be characterized by larger than average χ^2 if it is even successfully fitted.

Estimating the momentum for the track from the measured hits is a little more difficult. The calculation is detailed in Sec. E.1 in the context of obtaining a momentum estimate at which to evaluate the weight matrix to be used in the fit to a given track. This approximate momentum is also the one used as the initial guess to the momentum parameter in the track fit. It is given by

$$p_{est} = e\, B_{eff}\, \frac{\Delta x^2 + \Delta z^2}{2(\Delta z \sin\phi_1 - \Delta x \cos\phi_1)} , \tag{4.3}$$

where Δz is the effective length of the field of the magnet, B_{eff} is the approximate field integral along the trajectory divided by Δz, Δx is the magnetic deflection of the track over the field length, and ϕ_1 is the incident x-direction cosine of the track onto the field region. The effective field, B_{eff}, was estimated by calculating

$$B_{eff} = \frac{1}{\Delta z} \int B_y \, dz \tag{4.4}$$

at x and y chosen as the averages between the transverse coordinates of the point of entry into and the point of exit from the field region. Figure 4.5 gives the geometrical details for the momentum estimate.

Thus, given initial values for all track parameters, the least-squares algorithm described in Sec. E.4 is iterated to convergence. In brief, the χ^2 for a given track is expanded to second order in the determining parameters. This effectively fits a paraboloid surface to the local dependence of χ^2 on the parameters. The minimum of this surface gives improved parameter values for this event. As there may be deviations from parabolic dependence in the parameters, the procedure is repeated, taking the minimum of the previous step as the starting point for the next, until χ^2 changes by less than 10% from one iteration to the next. The parameter values thus obtained are considered to be the fit results for this track. It may seem that the 10% convergence margin is rather sloppy, but the fitting of some data with a 1% upper limit on the change in χ^2 to define convergence gave nearly identical results. This is borne out by the observation that the great majority of track fits require only a single iteration of the least-squares algorithm to converge under either condition. The reason for this is, of course, that the dependence of χ^2 on each parameter accurately follows a parabola in a domain including the initial estimate as well as the minimum for almost all measured tracks. In addition, the correlations between the parameters in the region near the χ^2 minimum are negligible.

In evaluating the χ^2 the correlations between the MWPC hits can, however, not be ignored. To see this, one may imagine that a particle passing through the chambers scatters between MWPCs 1 and 2. As a result, the deviations of this particle from an ideal, unscattered track at each of the subsequent MWPCs are correlated – they are caused by the same stochastic event. The effects of multiple scattering and of intrinsic measurement error in giving a deviation from an ideal track may be formalized by characterizing these processes by random variables

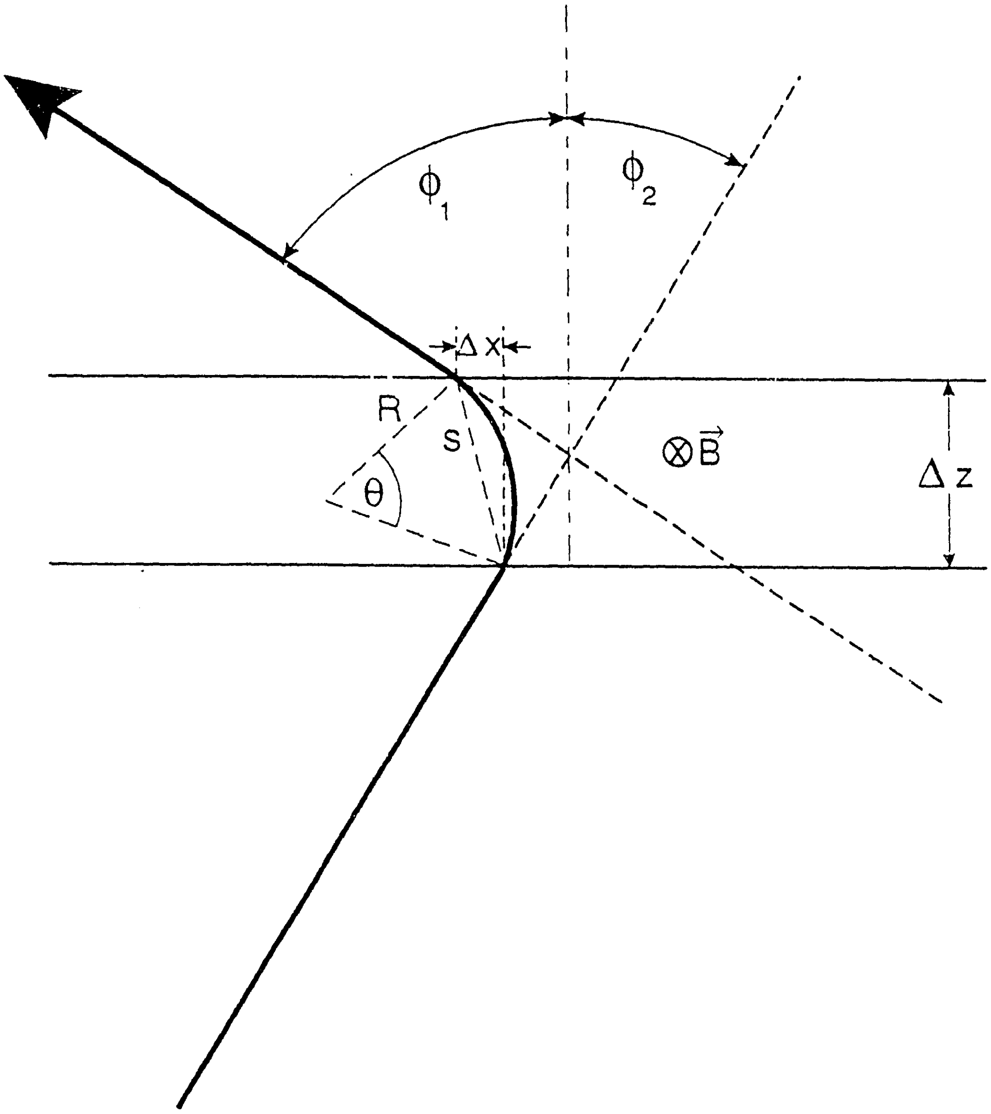


Figure 4.5: Geometry for estimating the track momentum.

representing scattering angles and measurement uncertainty. The most convenient choice is one that decouples the stochastic effects in the xz- and yz-planes. Still, the χ^2 definition that must be used may seem unusual, as it involves the cross-terms between the x-hits in different chambers and, separately, between the y-hits in different chambers. This definition is

$$\chi^2 = \left(\vec{f} - \vec{h}\right)^T \mathcal{W} \left(\vec{f} - \vec{h}\right) , \tag{4.5}$$

where $\vec{h}$ is the vector containing the 8 MWPC hits (4 in x and 4 in y) as measured, $\vec{f}$ is the vector of the corresponding fitted coordinates on the track, and $\mathcal{W}$ is the weight matrix. By construction, $\mathcal{W}$ is block-diagonal, as there are no xy cross-correlations in the choice made for random variable representation. The full derivation of the weight matrix, that correctly accounts for the hit correlations is given in Sec. E.1.

4.3 Corrections to MWPC Alignment

Reliable track information derives from individual MWPC plane hits that are as accurate as possible. There are three principal sources of uncertainty here. Firstly, there is the intrinsic position resolution of the MWPC planes which is determined mainly by the spacing of the individual wires. This spacing is 2 *mm* for all MWPC x- and y-planes in this experiment. Then, there is the effect of multiple scattering that may cause a particle to depart from the "ideal" track that it would follow in a vacuum. As this is a probabilistic process, its contribution cannot be corrected for in an individual track, though it may contribute to the weighting of the hits, as described in Sec. E.1. Finally, if the MWPCs do not share the same transverse coordinate system, that is, if the origins of their x- and y-axes do not coincide or if any MWPCs are relatively shifted along or rotated about any of the coordinate axes, the raw hits from different MWPCs are not compatible in the sense that they have been measured in different coordinate systems. Such misalignments remain even after careful placement of the detectors.

To offer the most reliable values for the coordinates of the hits in the MWPC planes for a given particle track, it is necessary to correct the raw values of these data according to those relative shifts and rotations of the MWPC planes which persist even after careful adjustment of the MWPC placements in the apparatus. This is most easily accomplished by analysis of data taken with the spectrometer magnet turned off, since the track model for this case simplifies to a straight-line. Such data were acquired at various times spaced throughout the entire run of the experiment to allow the determination not only of the best-fit values to the misalignments between the MWPCs, but also of their constancy in time and their possible changes during maintenance on the apparatus. To prevent overall traveling of the spectrometer coordinate system, the corrections to the transverse placements of MWPCs 2, 3, and 4 were made with respect to MWPC1, which was kept fixed as the "reference chamber." The details of the procedure are described in Sec. E.2.

Since a direct fit to rotational misalignments would be significantly more difficult than the fit to shift misalignments is, a pedestrian approach to the correction for these is chosen. Using the procedure described in Appendix E, an initial determination of the shift corrections is made. With these adjustments in place, one may make a first attempt at determining the relative misrotations of the MWPCs. To judge when two chambers have been rotationally aligned, one reconstructs a one-dimensional view of the target along the direction perpendicular to its surface (y). When the two chambers being considered are relatively parallel, the y-width of the target image is minimized, as any misrotation between them causes a degradation of the position resolution.

Thus, each of MWPCs 2, 3, and 4 are paired up in turn with MWPC1 to reconstruct a y-view of the target. Different angular corrections of chambers 2, 3, and 4 relative to MWPC1 (which is again kept fixed) are attempted and the full-width at half-maximum (FWHM) of the target is recorded for each of these. The resulting dependences of the target resolution *vs.* angular correction around the spectrometer axis are shown in Fig. 4.6 for MWPC2. The resolution of the decay position reconstruction is much less sensitive to misrotations around the axes transverse to the spectrometer axis (z), so such misalignments were left uncorrected. As the fit to shift misalignments assumes that there are no misrotations around the

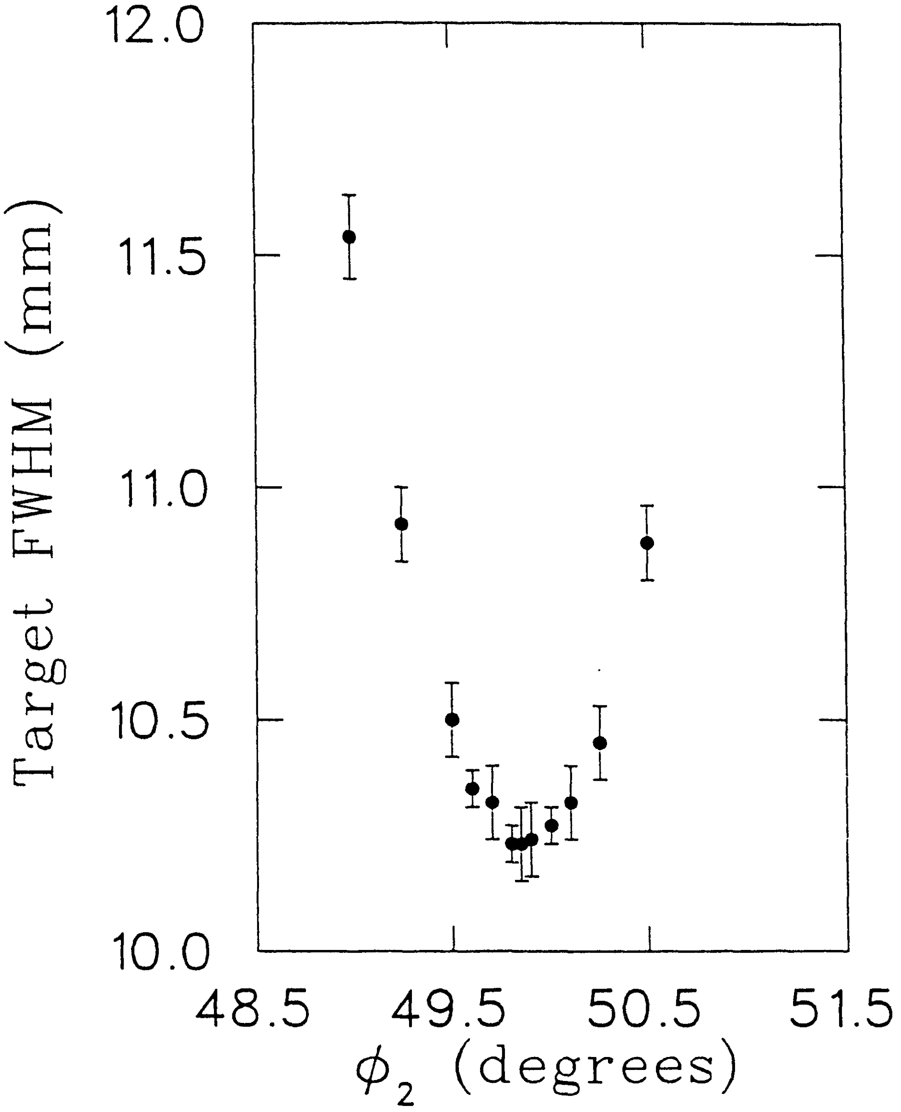


Figure 4.6: Reconstructed target FWHM as a function of angular alignment corrections around spectrometer axis. Here, MWPCs 1 and 2 were used in the reconstruction.

x- and y-axes, the presence of such small rotational misalignments is manifested by the need for more than a single iteration in the shift correction procedure. From the relative magnitude of these corrections from one iteration to the next (the third was always of the order of its error) and from measurements taken on the apparatus it may be inferred that the angular positioning errors around the x- and y-axes is less than 0.5° for all planes.

Though the C-magnet was off during the acquisition of alignment data, one might object that there is a remanent field of the yoke iron that can spoil the alignment data. This remanent field was unfortunately not measured, but it is reasonable to assume that it is less than 10 G. Thus, if we estimate that the e^+ passing through the spectrometer in the alignment runs experience a field along the y-axis of 10 G over a distance of 30 cm around the center of the magnet, the deflections due to the field at MWPCs 3 and 4 may be estimated. The radius of curvature of 35 MeV/c e^+ in a 10 G field is 117 m. This leads to a transverse deflection of $\Delta x \approx 0.04$ cm over the assumed 30 cm field length. This projects up to deflections of 0.13 cm at MWPC3 and 0.25 cm at MWPC4. Since this is of the order of the wire spacing in the MWPC planes, it is not taken to be a significant problem and is left untreated.

4.3.1 Alignment of Target to MWPCs

To be able to project the reconstructed decay origin onto one coordinate axis (y), it is necessary to have the target plane arranged perpendicularly to this axis. Care was taken to satisfy this requirement as closely as possible in the mounting of the MWPCs at a 50° inclination angle to the horizontal to match the nominal mounting angle of the target, but there are again residual discrepancies to this goal. Thus, it is necessary to correct the reconstructed coordinates of the decay origin, x_0 and y_0, for a remaining misrotation by $(\theta_T - \theta_{MWPC1})$ according to

$$\begin{aligned} x_0' &= x_0 \cos(\theta_T - \theta_{MWPC1}) + y_0 \sin(\theta_T - \theta_{MWPC1}) \\ y_0' &= -x_0 \sin(\theta_T - \theta_{MWPC1}) + y_0 \cos(\theta_T - \theta_{MWPC1}) \ , \end{aligned} \tag{4.6}$$

where θ_T and θ_{MWPC1} are the angles of the target and of the MWPC1 x-plane to the horizontal, respectively. The reason that MWPC1 appears here explicitly is that it is used throughout as the reference for alignment procedures, as stated above in the context of mutual alignment of the MWPCs.

To obtain the correction angle $(\theta_T - \theta_{MWPC1})$, data taken with the MWPC trigger are used to reconstruct the decay origin (x_0, y_0) and record it in a scatterplot. Then, assuming that the distribution of decays from the target follows a Gaussian in both x' and y', though with different widths, a fit is performed that includes the correction angle as a fit parameter. The width of the Gaussian in x' is determined by the beam spot on the target and the width in y' is given mostly by the finite resolution of the reconstruction, as the physical width of the target is small. The required correction angle was found to be 3.9°, meaning that the angle of the target to the horizontal was actually nearly 54°.

This fit is performed with data at various points during the experiment to ensure that changing targets and maintenance on the apparatus has not introduced additional misalignments. Such deviations were found to be less than 10 *mrad*, where the fit error was usually about 2 or 3 *mrad*. A rotational misalignment of 10 *mrad* causes an additional y-width of the reconstructed target of 0.2 *mm*. Since its reconstructed FWHM is about 10 *mm*, these deviations of the target angle throughout the experiment represent fluctuations in the reconstructed target width of only a few percent.

A by-product of the fit described is that we may locate the coordinate system of the spectrometer on the center of the vacuum window. This is accomplished from measurements of the apparatus that relate the placement of the target to the center of the vacuum window and the peak position of the target obtained from the fit. Overall shifts to the spectrometer coordinates are implemented so that $(x_0', y_0') = (0, 0)$ refers to the center of the vacuum window. This alignment will make it simple to place the condition on events that the track pass through the vacuum window, a cut that will be described in more detail below.

4.4 The y-Position Histogram

With a working track fitting algorithm and carefully aligned MWPCs, one may trace these trajectories back to the vertical (xy) plane through the beam-line axis. Since the magnetic field up to the second MWPC is negligible, one may use a simple straight-line extension from the track-fitted hits in MWPCs 1 and 2 back to the $z = 0$ plane to obtain the distribution of the points of origin, (x_0, y_0), of the particles observed in the spectrometer:

$$\begin{aligned} x_0 &= x_1 - \left(\frac{x_2 - x1}{z_2 - z_1}\right)(z_1 - z_0) \\ y_0 &= y_1 - \left(\frac{y_2 - y1}{z_2 - z_1}\right)(z_1 - z_0) \ . \end{aligned} \tag{4.7}$$

A two-dimensional contour plot of points (x_0, y_0) is shown in Fig. 4.7. The contours represent the logarithm of the number of counts from bins of 2 $mm \times 2$ mm size. Decays from the powder target are the dominant feature, but the enhancement due to thermal M decays from the vacuum is visible. A few decays from μ^+ stopping in the beam scintillator are also recognizable. The image plane has been rotated, as described above, so that the plane of the target is horizontal.

It is instructive to show the same contour plot for the case of a μ^- beam incident into the apparatus at the same momentum. This distribution is shown in Fig. 4.8. In contrast to the case of an incident μ^+ beam, μ^- do not form M; this is apparent in the comparison of Figs. 4.7 and 4.8.

The most important histogram in the analysis is the y-projection of Fig. 4.7. This is shown in Fig. 4.9. The logarithmic scale again serves to focus attention on counts due to thermal M decays in the vacuum, to the right of the target peak. The smaller of the two peaks on the left is from decays of μ^+ stopping in the beam scintillator. To obtain Fig. 4.9, the cuts applied are:

1. The χ^2 of the track must be less than 6.250 for tracks with 3 degrees of freedom, less than 4.605 for tracks with 2 degrees of freedom and less than 2.705 for tracks with 1 degree of freedom. These cuts represent the points which border 90% of the area under the χ^2 distribution from above, where

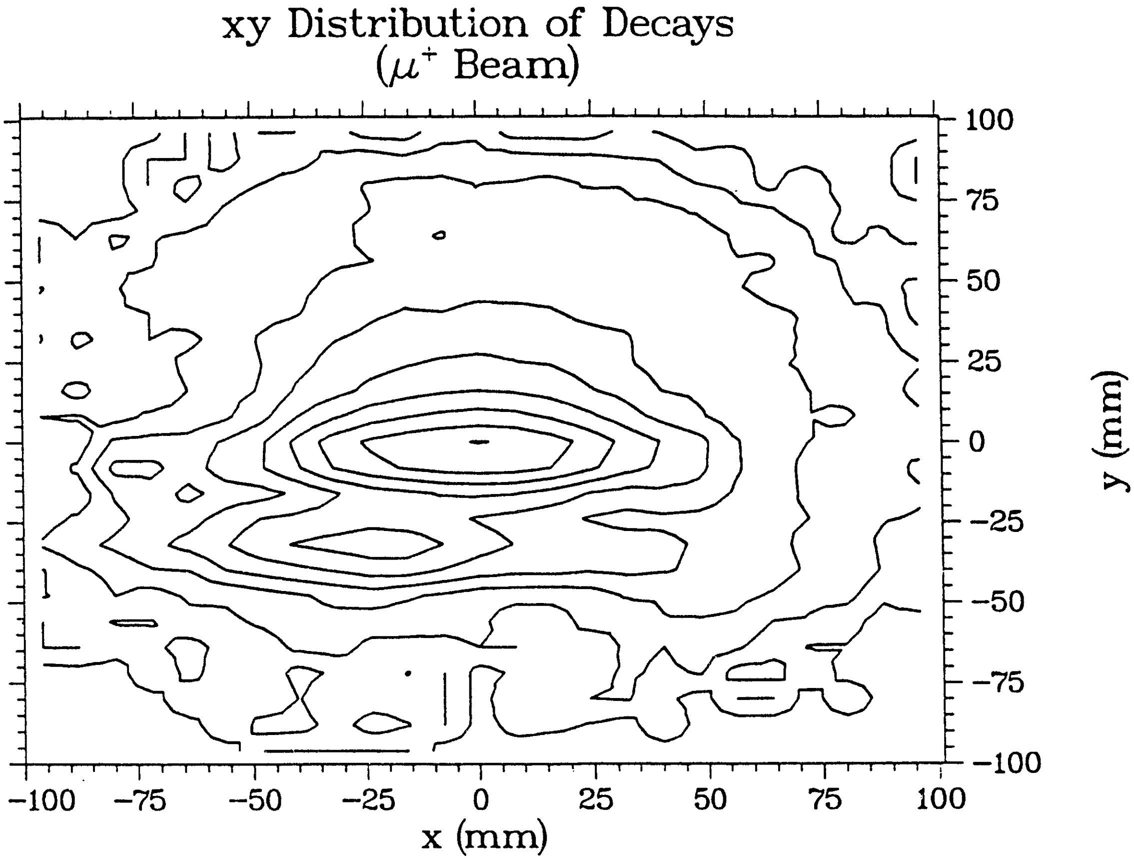


Figure 4.7: Distribution of decay points in the vertical plane $z = 0$. The contours are logarithmic to show the enhancement from the vacuum region due to thermal M decays there.

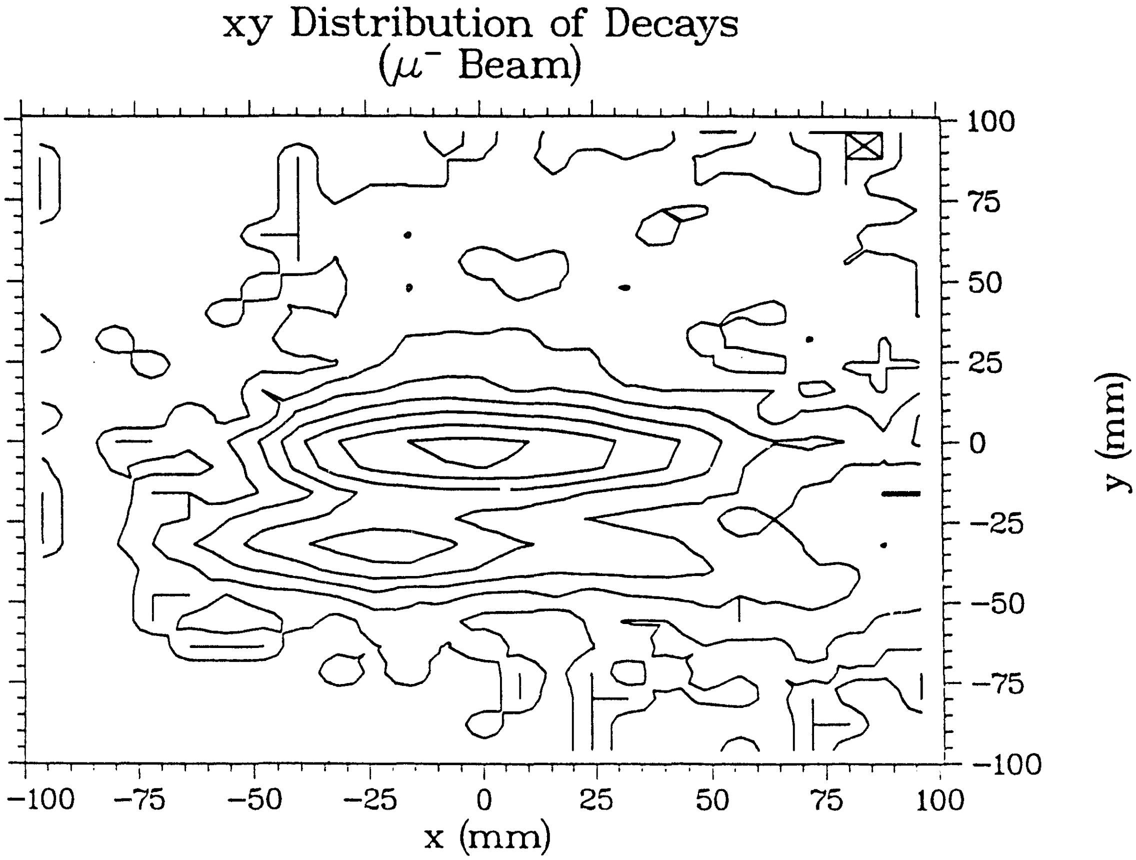


Figure 4.8: Logarithmic contours of distribution of μ^- decays. (Note the absence of any enhancement in the vacuum region as was observed for an incident μ^+ beam.)

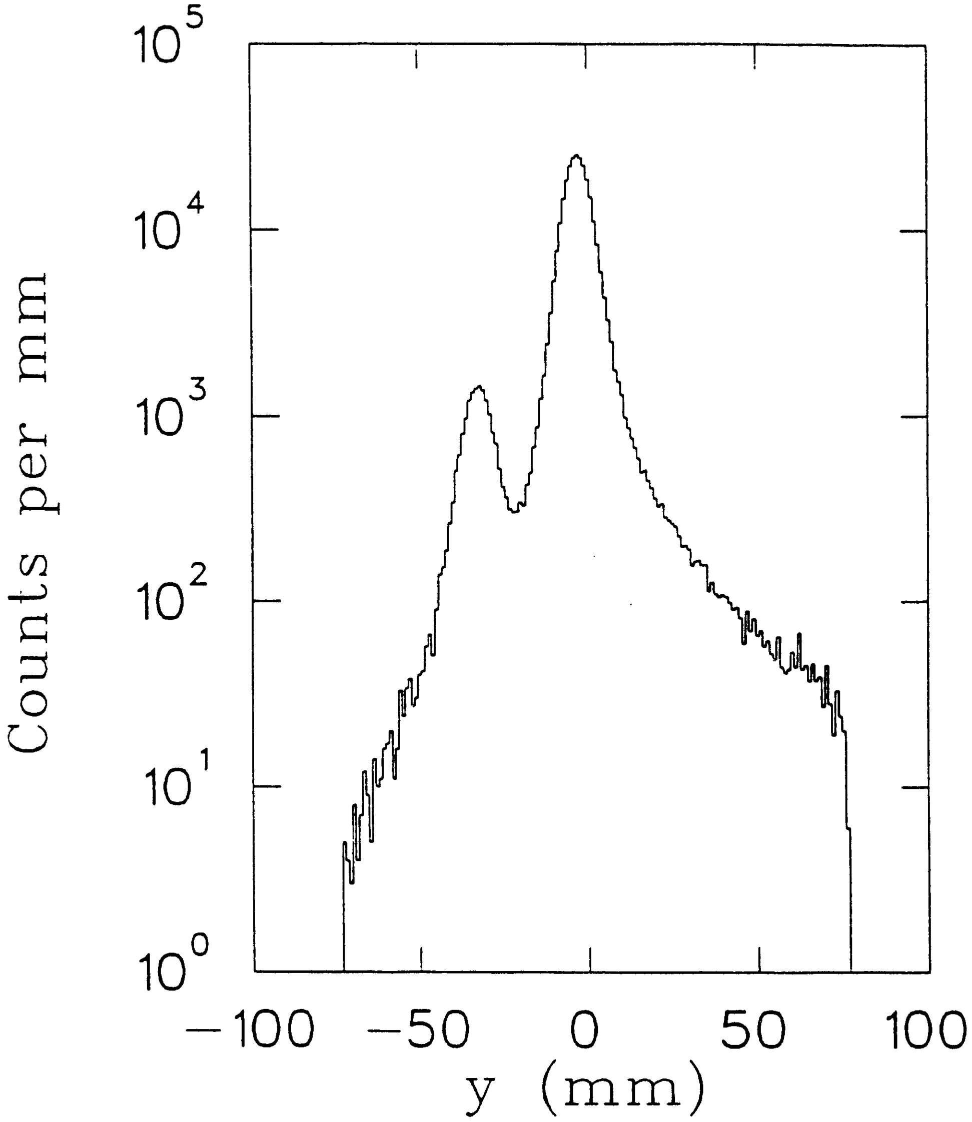


Figure 4.9: Projection of decay origins onto the y-axis. The vertical scale logarithmic to better present the enhancement in the vacuum region due to M.

Gaussian measurement errors have been assumed. The χ^2 distributions for 3 and for 2 degrees of freedom are shown in Fig. 4.10. Both of these distributions are a little more enhanced toward small values compared to χ^2 distributions for Gaussian measurement errors because the errors in the MWPC hits of the track are not quite Gaussian distributed. The assumption that the errors due to multiple scattering are Gaussian in the construction of the weight matrix (see Appendix E) causes this slight deviation. This cut removes events (about 5% of all) which are due to more than one particle leaving the spectrometer track or a single particle scattering very heavily in its passage through the chambers.

2. The track in the MWPCs must have passed through the vacuum window and the reconstructed decay origin must also lie in the same radial constraint. The alignment of the spectrometer onto the center of the vacuum window, as discussed above, is convenient at this point. This requirement removes most of the events that are due to decays from the first grid in the electrode structure of the slow $e^{\pm}$ collection system.

3. In the magnetic field, the tracks accepted must show the deflection of a positive particle. This serves to focus on the decay e^+ that are of interest when detecting M and rejects the small contribution from knock-on e^- .

The first two cuts are applied unchanged throughout on this position histogram in the further analysis. To search for $\overline{M}$, however, the third condition is reversed in polarity.

4.5 Time-of-Flight of Vacuum Muonium

When thermal muonium in vacuum was first observed [Bee86,Mil86], the signature relied on the observation of the time elapsed between the entry of a μ^+ into the apparatus and the detection of a decay e^+ from the vacuum region downstream of the target. It was found that the distribution of these decays in time and position is consistent with a Maxwell velocity distribution of M atoms (at the temperature

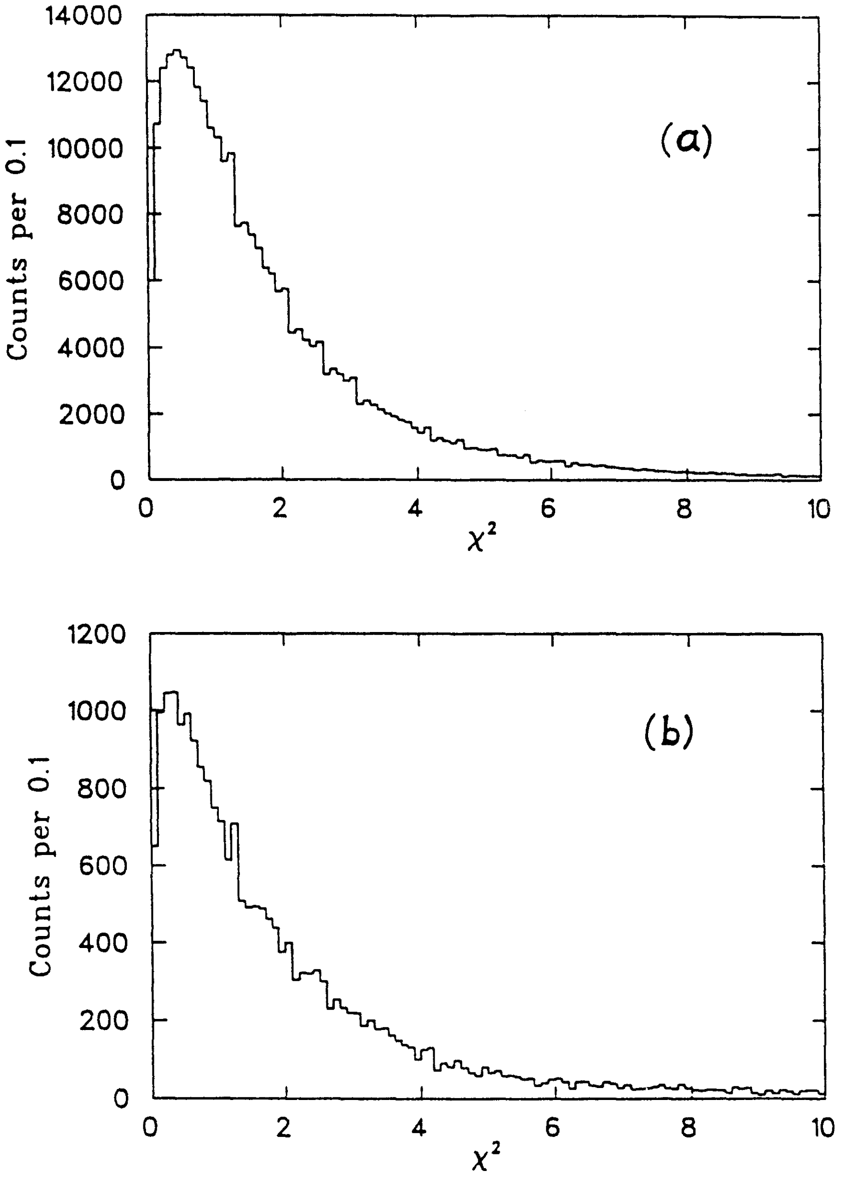
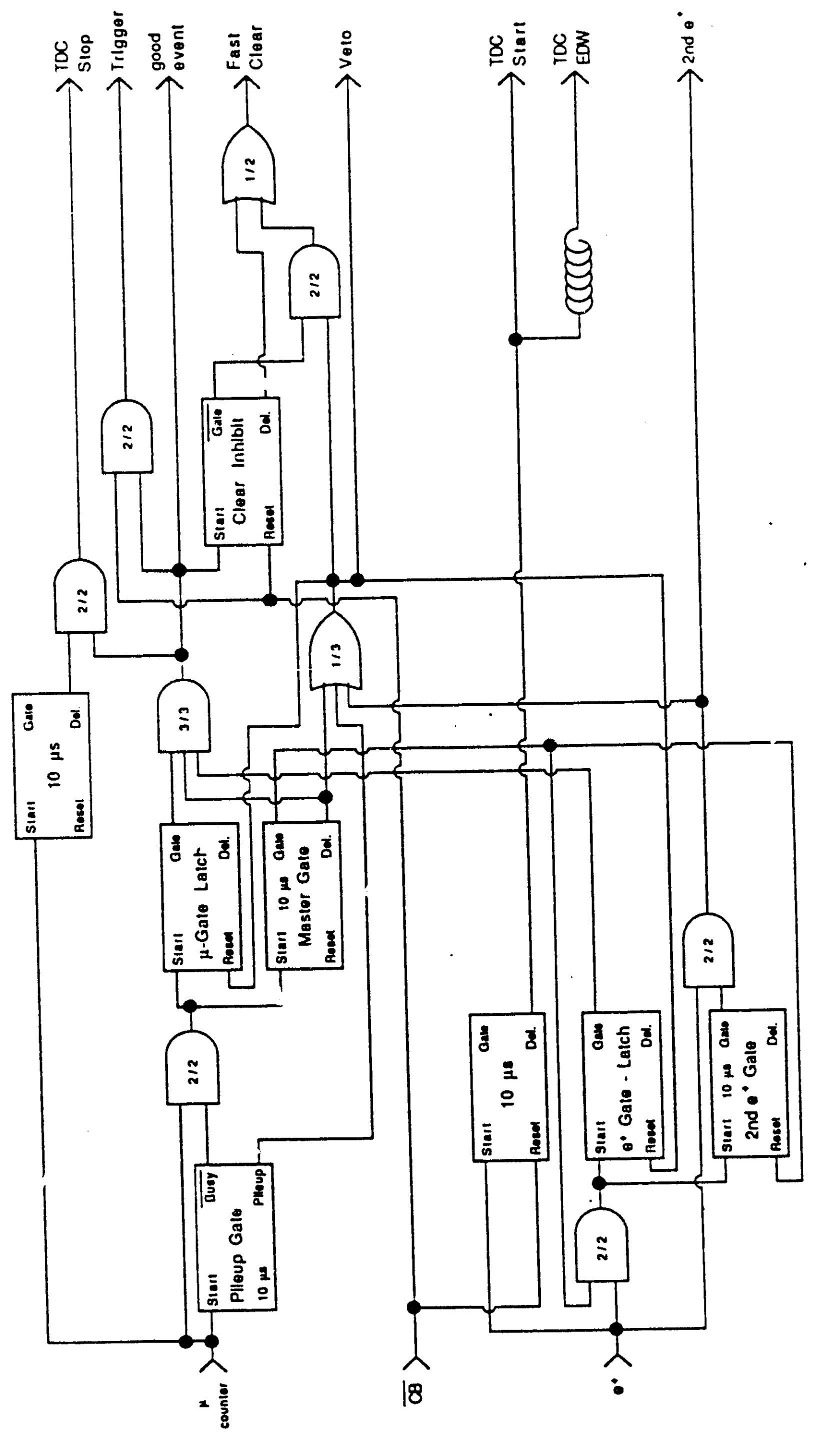


Figure 4.10: Distribution of χ^2 from track fits to decay e^+ with 3 degrees of freedom (a) and 2 degrees of freedom (b).

of the target) folded with the muon decay time distribution. One concluded that M atoms in vacuum with thermal energies had been produced. In order to verify the production of thermal M in vacuum in our apparatus, data were taken to observe the characteristic distributions that are the signature of thermal M in vacuum.

Because this signature requires correlating a decay e^+ track in the spectrometer to the incident μ^+ that gave rise to it, one can only tolerate a single μ^+ in the apparatus at a time. The gate length chosen during which to search for a decay e^+ after admitting a μ^+ was $10 \mu s$ long, corresponding to 4.5 μ^+ lifetimes. This meant that the rate of the incident beam needed to be small compared to $10^5\ s^{-1}$, instantaneous, so that pileup of μ^+ in the trigger did not diminish the detection efficiency appreciably. At a duty factor of 6.4 % in the accelerator, this restricted the allowed incident rate to less than about $6 \times 10^3\ s^{-1}$, average. The trigger used to ensure the detection of a μ^+ entering the apparatus and the observation of a correlated e^+ track in the spectrometer, without pileup from additional μ^+ or e^+, is shown in Fig. 4.11. When this trigger was not cleared by either pileup or by the absence of an observed e^+ in the spectrometer, it allowed the time-difference between the incident μ^+ and the decay e^+ to be recorded in a TDC (LeCroy 4208, 1 ns per channel).

This time-difference is expected, of course, to be distributed simply as a decaying exponential with the time-constant of muon decay. But when we place the condition on this histogram that the decay e^+ must have originated in particular regions along the y-axis, perpendicular to the target, then the e^+ must first reach these regions before being counted. More specifically, let us divide the y-axis into several regions: one that includes the reconstructed target image, and then 3 more adjacent intervals of 1 cm length that cover corresponding portions of the vacuum downstream of the powder target. In order for a decay e^+ to originate from these vacuum regions, a M atom must have reached this region and decayed there. Since M is produced in the powder target, the transit time from the formation in the target to the point of decay may be measured. The distribution of flight times to reach the different regions is characterized by the velocity distribution of the M atoms in the vacuum. These flight-time/decay-time distributions are shown in Fig. 4.12 for



CB = Computer Busy Status
e+ = MWPC pre-trigger
EDW = "end-of-window"

Figure 4.11: Trigger used to measure the time-of-flight of *M* atoms in the vacuum.

different spatial regions. The histogram in (a) contains counts that have been found to originate in the target y-distribution and therefore shows the expected muon decay lifetime distribution. In (b), (c), and (d), the y-position was required to lie in successive 1 cm wide zones of the vacuum, starting at the surface of the target. The departure from the exponential decay is apparent. Apparently, it has taken most of the counts entering here some time ($\sim$ 1.4 $\mu s/cm$ average) to reach these regions. This is due to the flight time of M at thermal energies from the powder to this point in the vacuum. A very coarse estimate of the temperature of the M from these histograms gives 300 K, in agreement with expectations.

The next section discusses the new signature that was developed for this experiment as an alternative to detect M and as a clean way to search for the $M \longrightarrow \overline{M}$ conversion.

4.6 Time-of-Flight of Atomic e^-

The most important development that led to the success of this experiment is the ability to detect the atomic e^- from M decay and therefore the use of a signature that includes the atomic e^+ in searching for $\overline{M}$ events. To associate counts on the MCP with particular decays observed in the spectrometer, the time-of-flight between a count in the spectrometer and the MCP was recorded. As the TOF peak of this correlation is expected at around 70 ns time difference between the spectrometer track and the arrival of a correlated atomic $e^\pm$ at the MCP, a 2 μs gate was opened by the spectrometer pulse. Within this gate, the time elapsed from the beginning of this gate to a delayed pulse on the MCP were recorded in a TDC (LeCroy 4208). The TDC logic is shown in Fig. 4.13. The zero-time channel for the TOF was not determined directly, but it was not needed for further analysis. Rather, in data replay, a gate of 75 ns length was located around the observed peak in the TOF histogram.

To properly identify atomic e^- and distinguish them from secondary e^- liberated from the SiO_2 powder target surface by decay e^+ leaving it, the TOF histogram was subjected to the same series of position cuts requiring the decay origin of the

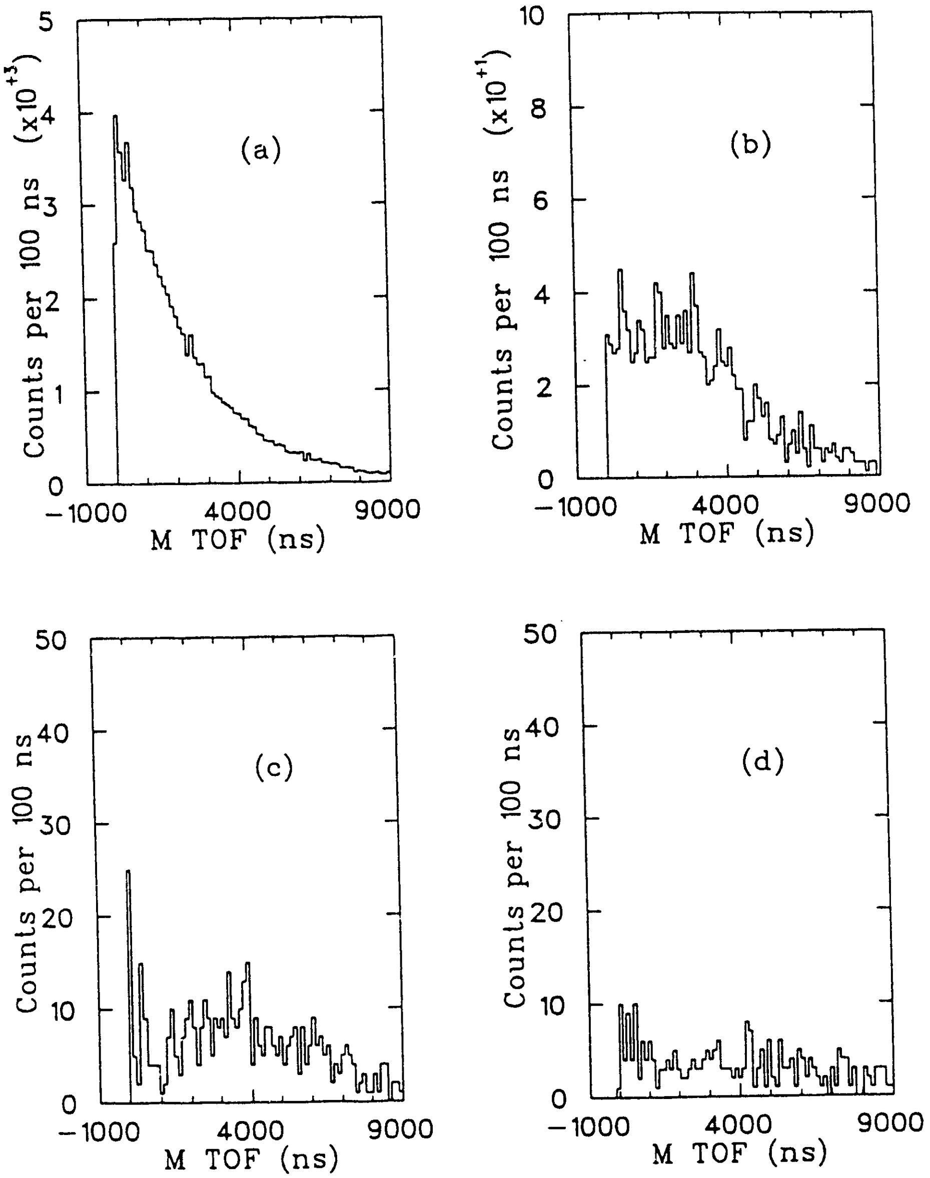


Figure 4.12: Time-distributions of M decays from different slices of the y-origin reconstruction. (See text for discussion.)

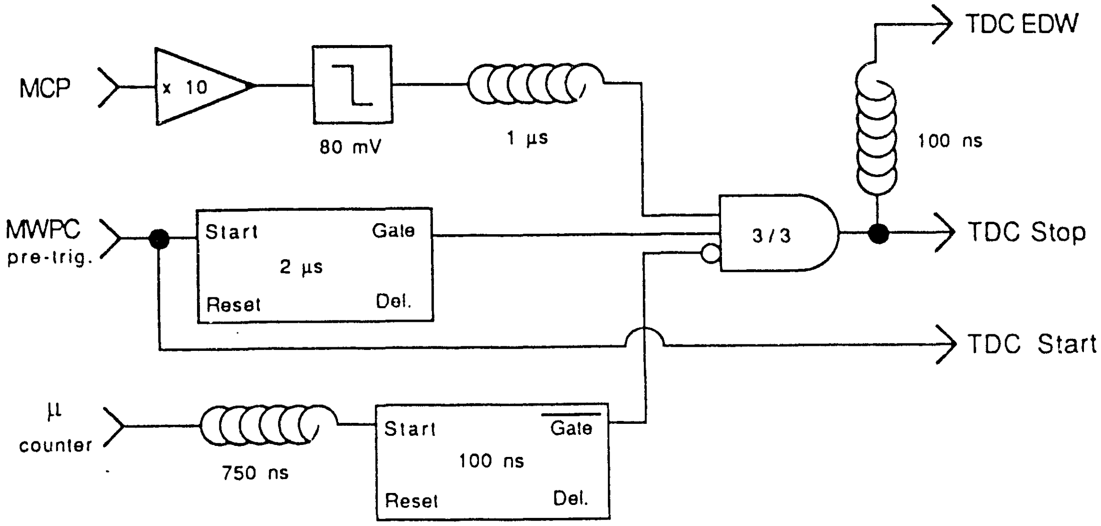


Figure 4.13: Logic for starting and stopping the TDC to acquire the TOF between spectrometer and MCP that gives the evidence for detection of atomic e^- .

event to lie in a specific region along the y-axis. The resulting histograms are shown in Fig. 4.14. The uncorrelated background (flat because at low rate) in these TOF histograms is mostly due to μ^+ passing through the target powder and emerging at its downstream surface, pulling with them a wake of secondary e^- . These are accepted by the low-energy transport system and directed onto the MCP. To suppress this background as far as possible, the TDC-stop logic included a veto generated from a MCP pulse in coincidence with a count in the beam scintillator. Due to finite efficiency of this veto, some of this background remains.

For data searching for the $M \longrightarrow \overline{M}$ conversion, this veto was removed, since the low-energy transport was then tuned for e^+ and no secondary e^- from the target are expected to be transported to the MCP. A second, decisive reason is that the full channel rate taken while searching for $\overline{M}$ events (10^6 μ^+ s^{-1} average) would cause nearly every MCP pulse to be vetoed by accidental coincidence.

To display in an alternative fashion that the atomic e^- have been observed, one cuts the y-position histogram on the presence of a count in the TOF gate of interest (channels 625 to 700), as in Fig.4.15. In this histogram, the contribution from thermal M in the vacuum is still patent, but the target appears narrower than it did without the TOF cut. This is because the requirement of a correlated TOF count will only accept decays from the target surface, not from within the powder. Secondary electrons from the powder can only be transported if they are liberated at its surface.

Again, comparing the TOF histograms of Fig. 4.14 to the corresponding histograms taken with an incident μ^- beam serves to confirm the presence of thermal M detected by the coincident observation of its decay e^+ and atomic e^- . For this check, the polarity of the collection system is also set to accept e^- , so that secondary e^- from the target surface still appear. Figure 4.16 shows these TOF histograms taken with a μ^- beam.

The striking feature is that there are very few counts with a decay origin in the vacuum that give a TOF in the gate of interest. As there is no M present, the only correlated events here are due to decays from the target that scatter strongly so as to give a reconstructed decay origin in the "vacuum," and to a few μ^- that scatter

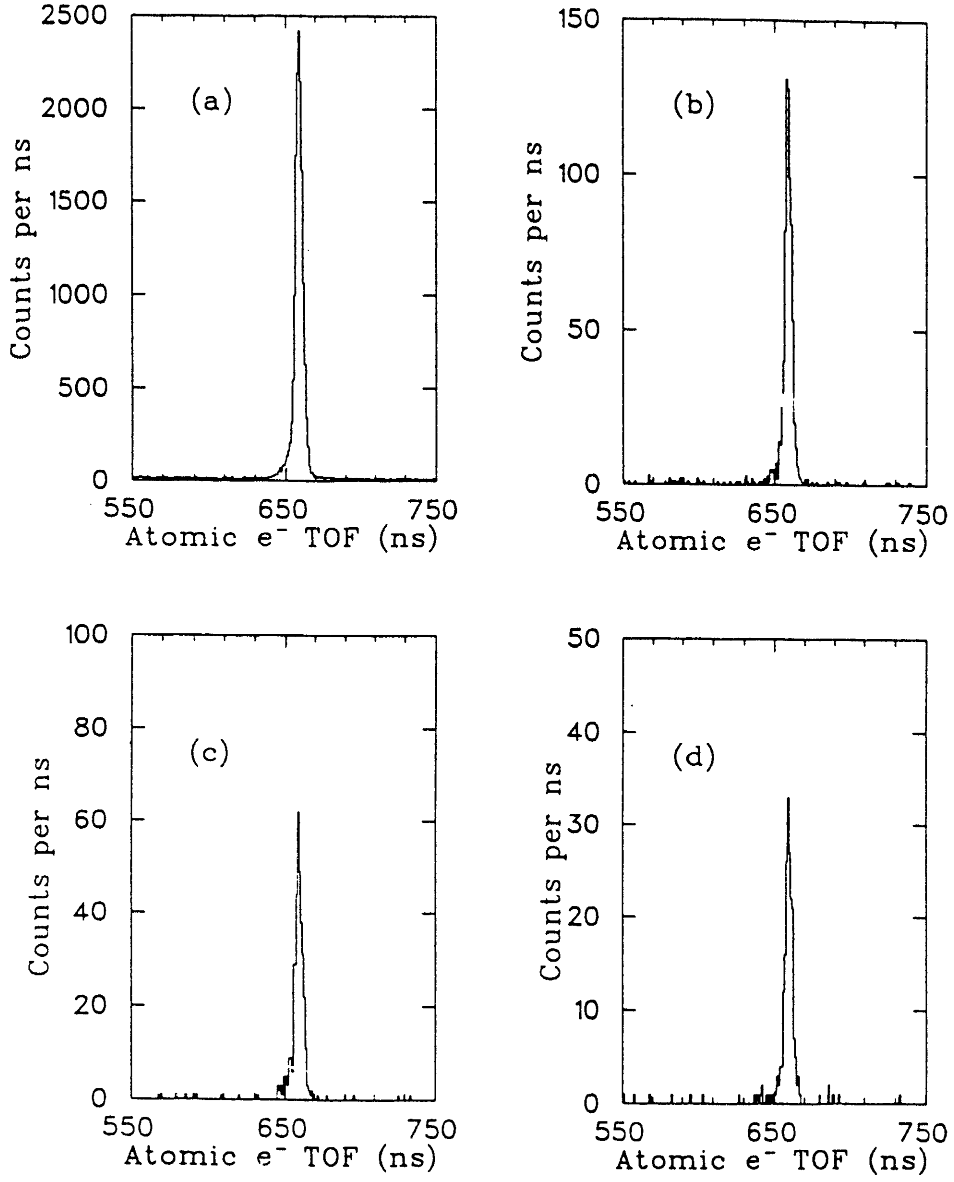


Figure 4.14: Time-of-flight distribution between spectrometer track and MCP. The TOF of secondary e^- from the target are shown in (a), whereas (b), (c), and (d) are the TOF from successive 1 *cm* regions starting at the powder surface.

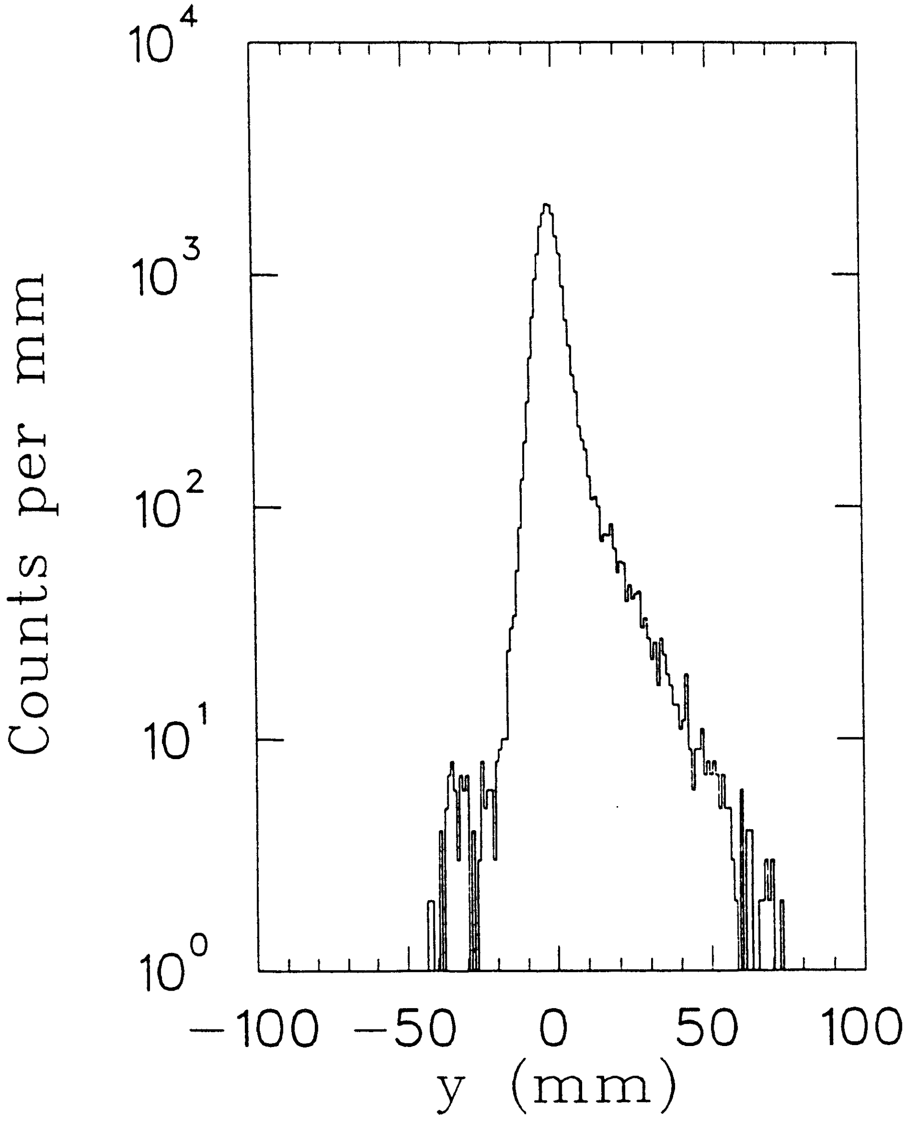


Figure 4.15: Distribution of decay positions along y-axis for events with a correlation between MWPC hits and MCP. Note that the thermal M counts in the vacuum region are still present.

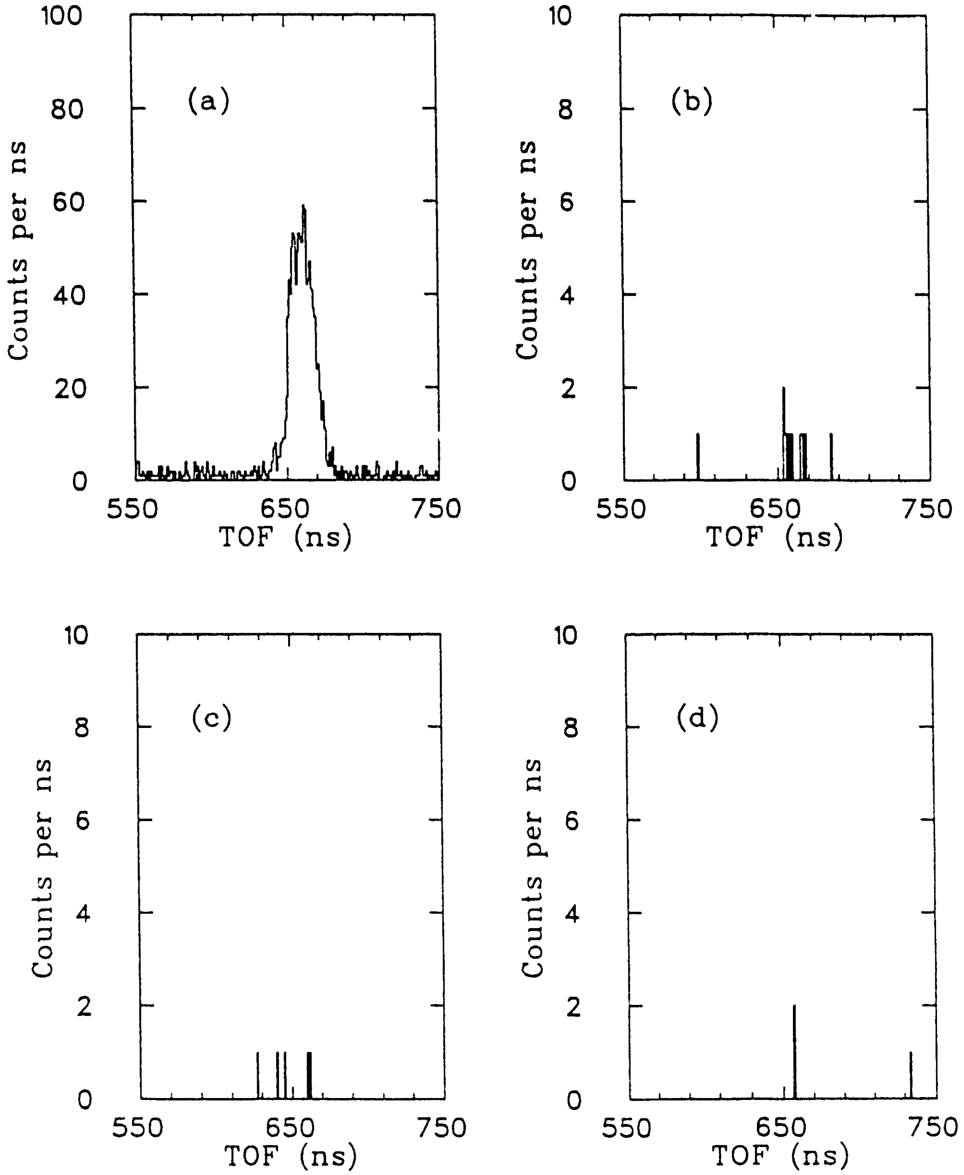


Figure 4.16: Spectrometer to MCP TOF taken with incident μ^- . Secondary e^- from the target give rise to (a). Note the scarcity of counts with origin in vacuum regions (b), (c), and (d).

in passing through the target, stop in the walls of the vacuum chamber in view of the spectrometer, and give a decay e^- that liberates secondary e^- from the walls.

Having thus established that one can indeed detect the atomic e^- liberated by the decay of M, it is clear that the analogous signature serves well to identify any $\overline{M}$ decays. As no direct source of $\overline{M}$ is available, the acceptance of the collection system for slow e^+ was tested, as described above, by a W-foil moderator slowing the β^+ from a ^{22}Na source.

4.7 The Maximum Likelihood Fit

The method of maximum likelihood is a very general prescription for extracting information from a measured data set. The necessary ingredients are a set of distributions, in the same parameter space as the data to be analyzed, that each describe one of the processes thought to contribute to the experimentally obtained distribution of events. These *reference distributions* may be obtained by analytic calculation when the model is confidently known, by a Monte Carlo of the process to be described when the analytic method is too complex for solution, or from other measurements in which the experimental conditions have been arranged in a controlled way to emphasize the particular process whose distribution is sought. It is then possible to state, separately for every event in the data to be analyzed, with what probabilities it derives from each of the reference distributions. In effect, for each physical process included by its characteristic distribution, one may give the probability that it caused a given event. These probabilities are simply the values of the normalized reference distributions at the point in the parameter space where the data count occurred. To determine the probability that a given event is at all described by a linear combination of the reference distributions, these values are added with the appropriate weights.

4.7.1 Definition of the Likelihood Function

One may extend this principle to an ensemble of data counts which are believed to be drawn from a combination of reference distributions characterizing the different processes contributing to the observed distribution of the data. In this case, the product of the probabilities for each event to come from the specified linear combination of reference distributions is a relative measure of the suitability of the combined probability distribution for describing the data set. This product is called the *likelihood function* and may contain parameters that are to be adjusted for the best description of the experimental data.

Denoting the reference distributions by R_j, the number of counts from each of these distributions present in the data by N_j, and the parameter dependence of the distributions (say, for example, on position of decay origin and on fitted track momentum for an event) by the vector $\vec{X}$, the likelihood function may be written as

$$\mathcal{L} = \prod_{i=1}^{N} \left[\sum_j \frac{N_j}{N} R_j \left(\vec{X}_i \right) \right] , \tag{4.8}$$

where N is the total number of events in the data set. In this expression the index i runs over all the events in the data distribution and the index j is taken over the set of reference distributions. Thus, the dependence on each event enters through its location in the parameter space, $\vec{X}_i$. Because there is a fixed number of counts in the data distribution, N, which are to be described by counts drawn in the appropriate admixture from each of the reference distributions, the constraint that

$$N = \sum_j N_j \tag{4.9}$$

must be imposed. The likelihood function may be thought of as a function of the number of counts from each of the reference distributions, but satisfying aforementioned condition on their sum. Thus, one may consider the adjustment to the best description of the data set as a fit, where the fit parameters are the N_j. Because of the constraint on their sum, the number of parameters in this fit is one less than the number of reference distributions.

4.7.2 Application of the Method

The histogram to which the maximum likelihood analysis is applied in this experiment is the y-position distribution of the decay points, as traced back from particle tracks in the spectrometer. This method is applied both to find the most probable number of $\overline{M}$ atoms observed over the course of data taking and to the determination of the number of M atoms formed during this time, from which these $\overline{M}$ counts would have resulted by conversion. The likelihood fit to these distributions must be handled separately.

To fit to the y-distribution of M decays, the likelihood function takes the form

$$\mathcal{L} = \prod_i \left[\left(\frac{N_T}{N}\right) T(y_i) + \left(\frac{N_M}{N}\right) M(y_i) + \left(\frac{N_B}{N}\right) B(y_i) \right] , \qquad (4.10)$$

where $T(y)$, $M(y)$, and $B(y)$ are the distributions of decays from the target, from thermal M in vacuum, and from other background sources, respectively. The product runs over all events in the histogram and the constraint $N = N_T + N_M + N_B$ applies. Therefore, this fit has two parameters.

The data distribution obtained from summing together all data taken to search for the $M \longrightarrow \overline{M}$ conversion is fit by

$$\mathcal{L} = \prod_i \left[\left(\frac{N_{\overline{M}}}{N}\right) \overline{M}(y_i) + \left(\frac{N_K}{N}\right) K(y_i) \right] , \qquad (4.11)$$

where $\overline{M}(y)$ is the distribution of vacuum antimuonium decays after conversion from an initial state of M at the target surface and $K(y)$ is the distribution of background events. The latter is dominated by knock-on e^- and includes structure from decays from the target, the beam scintillator and knock-ons from M decays in the vacuum.

So that the likelihood fit can be carried out, one needs the reference distributions. For the case of the target distribution in the M fit, it is simply determined by folding the resolution distribution with the finite thickness of the powder. The determination of the resolution function for the reconstruction of the decay origin is described in Appendix F. The background distribution in the M fit has been empirically found to be well-approximated by a uniform distribution ("flat" background).

The background in the $\overline{M}$ fit, however, is more complicated and must be obtained from the data; the relevant cuts will be described below. Finally, the distribution of vacuum M and $\overline{M}$ decays may be analytically calculated and numerically folded with the resolution distribution.

4.7.3 Reference Distributions for M and $\overline{M}$

The calculation of the characteristic distributions for vacuum decays of M and $\overline{M}$ assumes that the possibly coupled system is formed in an initial state of M. These M atoms may be taken to emerge from the target surface at time $t = 0$. Their further motion is determined by their Maxwellian velocity distribution, their $\cos\theta$ angular distribution, and their decay. To find the y-distribution of M decays, one calculates the folding integral

$$M(y) = \int_0^\infty dv_y \int_0^\infty \frac{dt}{\tau}\, \delta(y - y_0 - v_y t)\, F(v_y)\, e^{-t/\tau} \,, \tag{4.12}$$

where y_0 is the location of the target surface xz-plane and $F(v_y)$ is the distribution of the y-component of the M atoms velocity. The Dirac-delta function comes about by requiring the decay-position of the atom to be kinematically compatible with the y-velocity and time-of-flight until decay. The velocity integral may then trivially be carried out to give

$$M(y) = \int_0^\infty \frac{dt}{\tau t}\, e^{-t/\tau}\, F\left(\frac{y - y_0}{t}\right) \,. \tag{4.13}$$

The normalized distribution $F(v_y)$ is determined by integrating out the x- and z-dependence from the three-dimensional velocity distribution, with the $\cos\theta$ angular distribution accounted for. This integral is

$$\begin{aligned} F(v_y) &= 4\left(\frac{m}{2\pi kT}\right)^{3/2} v_y\, e^{-mv_y^2/(2kT)} \\ &\quad\times \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_z \frac{e^{-\left(mv_x^2+mv_z^2\right)/(2kT)}}{\sqrt{v_x^2 + v_y^2 + v_z^2}} \,, \end{aligned} \tag{4.14}$$

where m is the mass of the M atom, k is Boltzmann's constant, and T is the temperature of the M atoms. The angular distribution of $\cos\theta$ with respect to the target normal has been inserted by the factor

$$\frac{v_y}{\sqrt{v_x^2 + v_y^2 + v_z^2}} \tag{4.15}$$

and the normalization has been adjusted to satisfy

$$\int_0^\infty dv_y \, F(v_y) = 1 \ . \tag{4.16}$$

Negative values of v_y are not included, as they will not carry the M atoms from the target surface into the vacuum. The integral in Eq. 4.14 can be carried out to give

$$F(v_y) = \frac{2mv_y}{kT}\left[1 - \mathrm{erf}\left(\sqrt{\frac{m}{2kT}}\, v_y\right)\right] \ . \tag{4.17}$$

The M atom distribution, $M(y)$, is then numerically evaluated at the different y-values desired to obtain the M reference distribution. This distribution is shown before and after folding with the resolution function of the reconstruction in Fig. 4.17.

This distribution is actually that of the total M, $\overline{M}$ system, as we did not include in the calculation above the depletion of the M population by the conversion. As the $M \longrightarrow \overline{M}$ process is known to proceed at most very slowly, this distribution is an excellent approximation to the M distribution alone.

To obtain the distribution of thermal $\overline{M}$ atom decays in the vacuum, the time-dependence in the folding integral must be modified by a factor describing how the convergence populates the $\overline{M}$ state. It then reads

$$\begin{aligned} \overline{M} &= \int_0^\infty dv_y \int_0^\infty \frac{dt}{\tau}\, \delta(y - y_0 - v_y t)\, F(v_y) \\ &\times\ 2\left[1 + \left(\frac{\hbar}{\delta\tau}\right)^2\right] e^{-t/\tau} \sin^2\left(\frac{\delta t}{2\hbar}\right) \\ &= \int_0^\infty \frac{dt}{\tau t}\, F\left(\frac{y - y_0}{t}\right) \\ &\times\ 2\left[1 + \left(\frac{\hbar}{\delta\tau}\right)^2\right] e^{-t/\tau} \sin^2\left(\frac{\delta t}{2\hbar}\right) \ . \end{aligned} \tag{4.18}$$

The result of the numerical quadrature to give the reference distribution $\overline{M}$ is shown in Fig. 4.18, before and after smearing it with the resolution function.

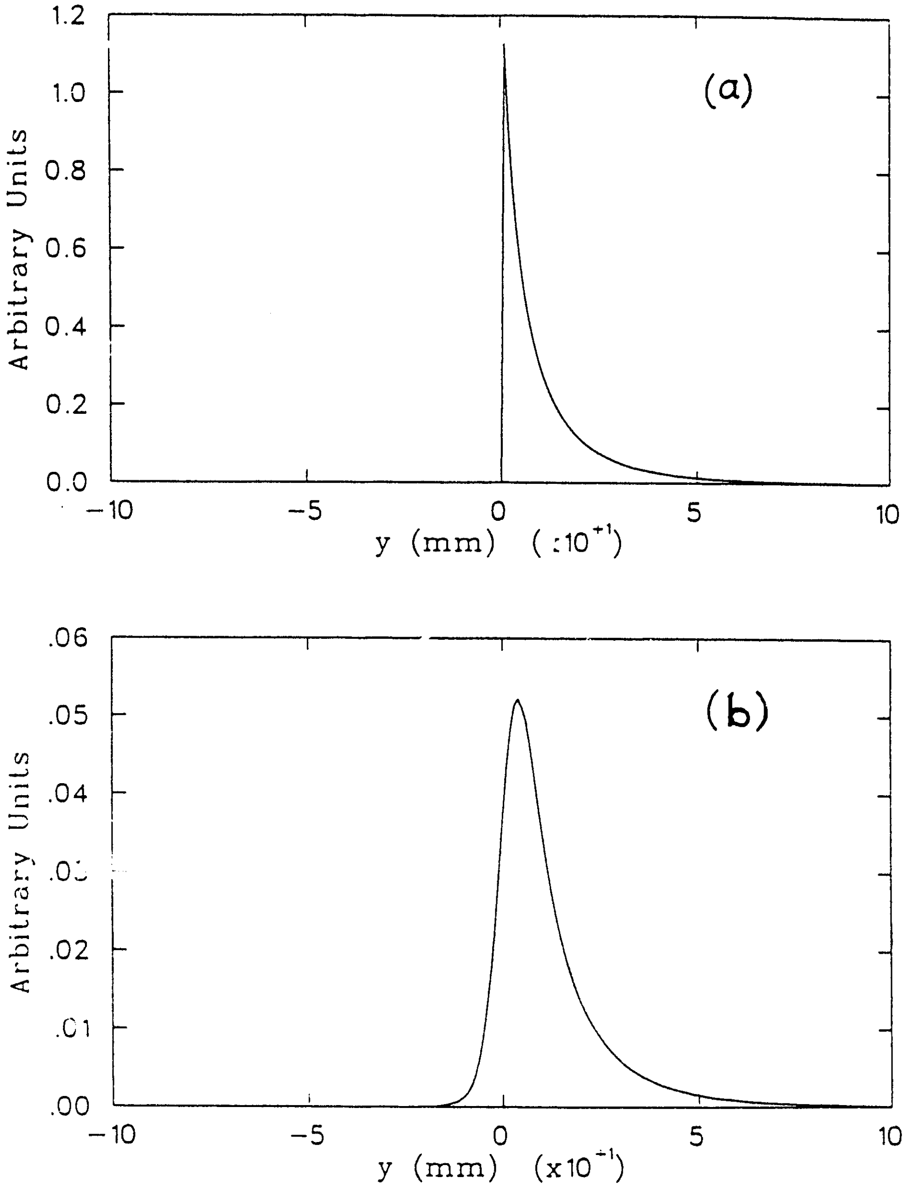


Figure 4.17: Reference distribution for decay positions of thermal M in vacuum, before (a) and after (b) folding with resolution function.

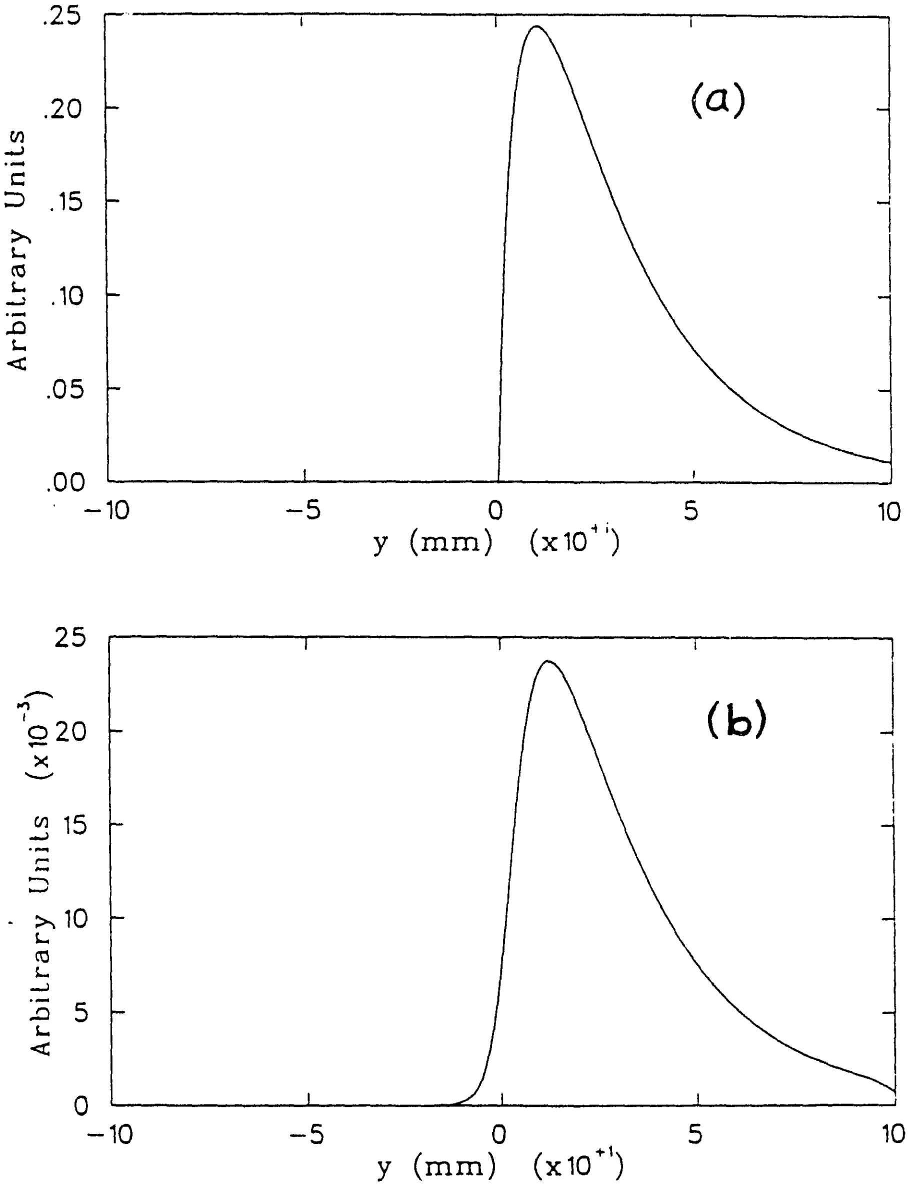


Figure 4.18: Reference distribution for vacuum decays of thermal $\overline{M}$ before (a) and after (b) folding with resolution function.

4.8 Number of $\overline{M}$ Atoms

The y-histogram in which evidence of $\overline{M}$ is sought requires several conditions for entry of an event. These are chosen to maximally emphasize any $\overline{M}$ decays that may be present. First, the tracks must pass the usual pattern recognition cuts and be successfully fitted with reasonable χ^2. The trajectories in the spectrometer must deflect in the direction of a negatively charged particle. Also, as established, the tracks must pass through the vacuum window and their origins must lie in the circular region defined by it when projected back to the $z = 0$ plane. Finally, there must be a count on the MCP within the time window established by the observation of the atomic e^- of M. The resulting y-histogram is shown in Fig. 4.19.

Slightly different cuts must be applied to determine the background distribution relevant to this histogram. To ensure that the spatial acceptance for the background tracks is not different from that for the $\overline{M}$ data histogram, the only change is to require that there be no count on the MCP within the time window where a conversion signal would appear. This ensures that no event enters into both the data and the background distributions. The resulting distribution is shown in Fig. 4.20.

With this background distribution and the $\overline{M}$ reference distribution, one fits the data distribution according to the method of maximum likelihood. The algorithm used is a grid search with adjustable step size in the single free parameter, for example $N_{\overline{M}}$. This ensured that the global maximum was found. In fact the likelihood function was found to be well-behaved, nearly Gaussian in shape and without subsidiary minima. The likelihood function is shown in Fig. 4.21 in dependence on the number of $\overline{M}$ counts. Its peak lies at zero counts and 90% of its area is contained in the region $N_{\overline{M}} < 7$ counts. The result is, then, a 90% confidence level upper limit of 7 $\overline{M}$ counts in a total of over 200 tapes of data searching for $M \longrightarrow \overline{M}$.

There is, however, more information that has not yet been used. The momentum distribution of the knock-on e^- that are presumed to be the cause of the background is quite different than the Michel distribution that is characteristic of $\overline{M}$ decays. The Michel distribution, determined from fits to e^+ tracks with a reversed C-magnet, and

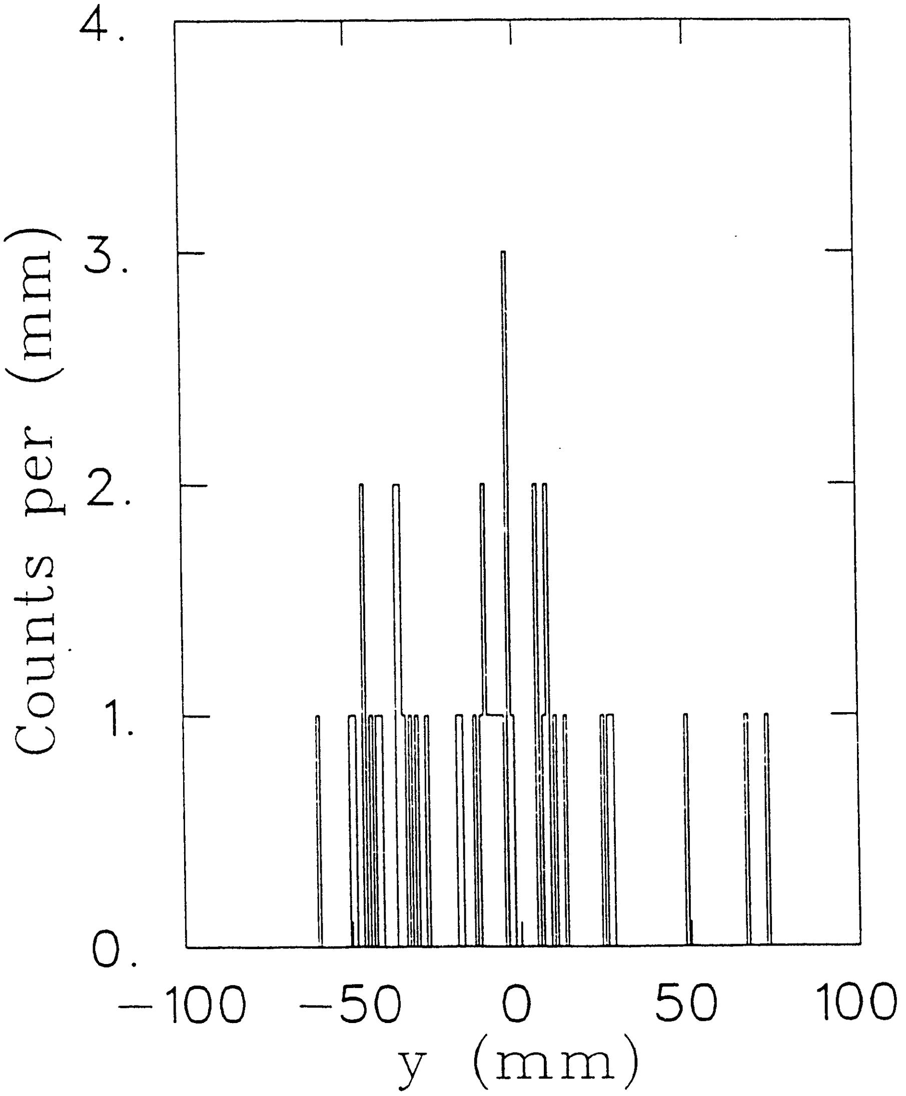


Figure 4.19: Position distribution of decay origin for $\overline{M}$ conditions.

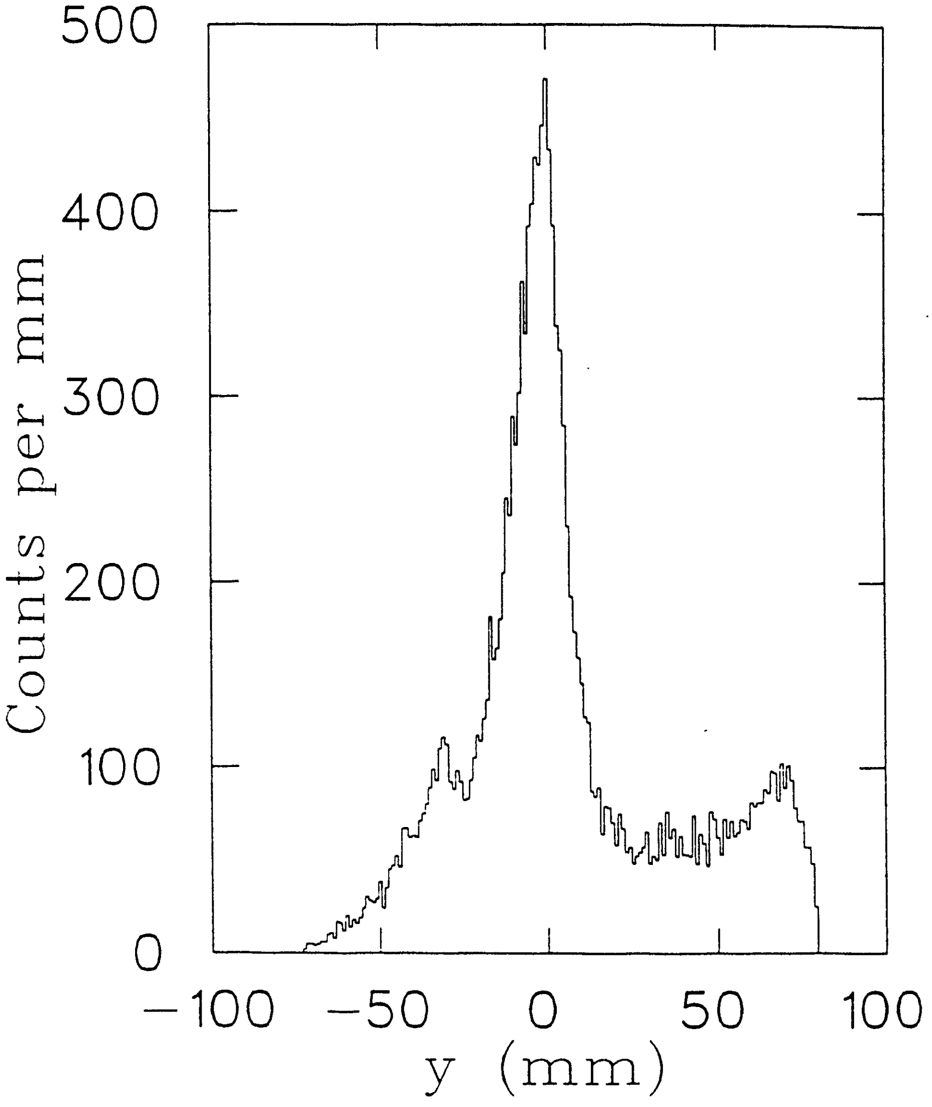


Figure 4.20: Background distribution for $\overline{M}$ detection in the y-position histogram.

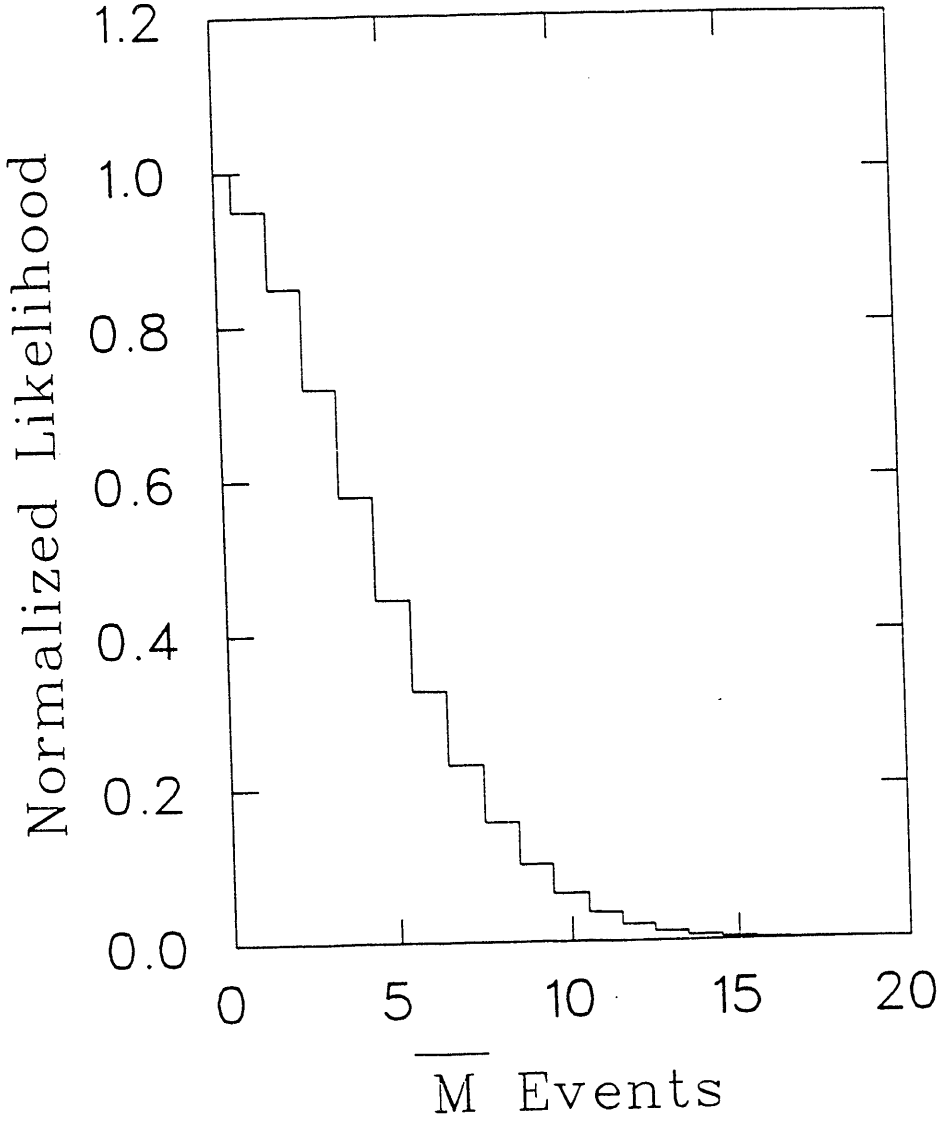


Figure 4.21: Likelihood function for fit to $\overline{M}$ data.

the measured Bhabha distribution in this momentum range, obtained by focusing on tracks with e^- curvature, are shown in Fig. 4.22. The low-momentum fall-off of the knock-on distribution is caused by the spectrometer momentum cut off that is a result of its finite transverse dimensions. It is clear that a cut on this momentum is very appropriate to remove much of the background from the $\overline{M}$ data while maintaining any conversion signal. The figure of merit used to judge the choice of lower momentum cut was chosen as the product

$$f.o.m. = \epsilon_{Michel}\ (1 - \epsilon_{knock-on}) \tag{4.19}$$

to optimize the rejection of knock-on electrons while maintaining acceptance for Michel e^- . The dependence of this figure on the location of the cut is given in Table 4.1. As the table indicates, the cut is optimally chosen as $p > 22.5\ MeV/c$. This condition removes over 89% of the knock-on e^- while preserving more than 87% of the Michel-distributed decay electrons.

The data distribution after application of this requirement is shown in Fig. 4.23. The distribution of background events passing the momentum cut is given in Fig. 4.24. The likelihood function that is obtained using these distributions, together with the $\overline{M}$ signal distribution, behaves as in Fig. 4.25. For this case, the likelihood function is maximal at $N_{\overline{M}} = 0$ with a 90% confidence level upper limit of $N_{\overline{M}} < 2$. This is the final result for the number of $\overline{M}$ atoms.

4.9 Number of M Atoms

To determine the total number of M atoms formed during the search for $M \longrightarrow \overline{M}$, several steps are necessary. The guideline to be followed in properly finding this normalization may be stated as the question: "If *all* the M atoms formed decayed as $\overline{M}$, how many $\overline{M}$ atoms would we detect?" In other words, the number of M atoms used as a normalization must be stated with those acceptances folded in that make the number compatible with the number of $\overline{M}$ atoms obtained above.

It was found that the matrix trigger accepts unequally over the field of view of the reconstruction. This may seem somewhat surprising, as the histogram being

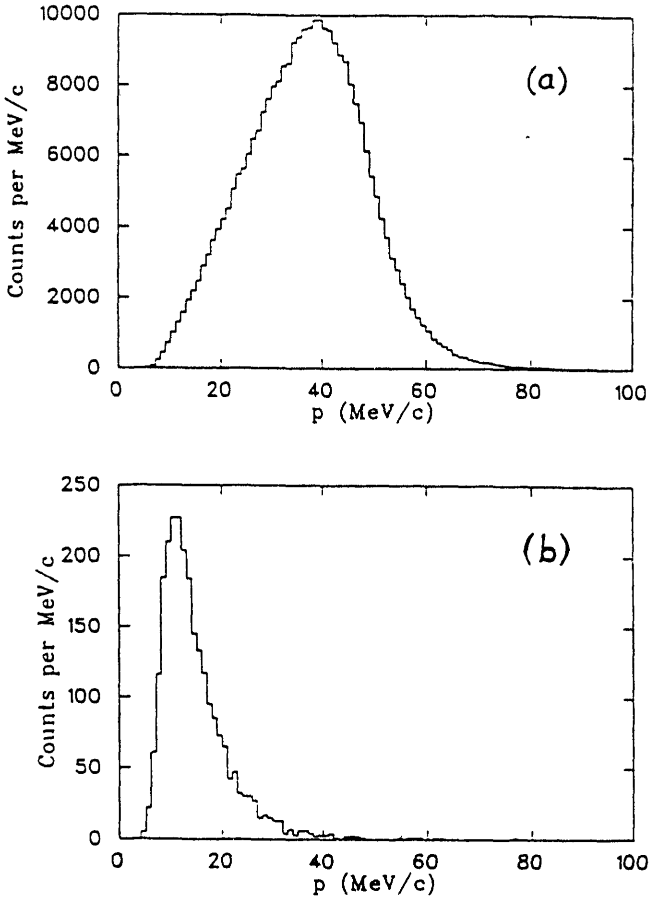


Figure 4.22: Momentum distributions (a) for decay electrons and (b) for knock-on electrons.

p_{min} (MeV/c)	ϵ_{Michel}	$\epsilon_{knock-on}$	$f.o.m.$
0.0	1.0000	1.0000	0.0000
10.0	0.9948	0.7596	0.2391
12.5	0.9843	0.5333	0.4594
15.0	0.9671	0.3636	0.6155
17.5	0.9432	0.2416	0.7152
20.0	0.9116	0.1617	0.7642
22.5	0.8730	0.1063	0.7802
25.0	0.8258	0.0750	0.7639

Table 4.1: Figure of merit of the momentum cut.

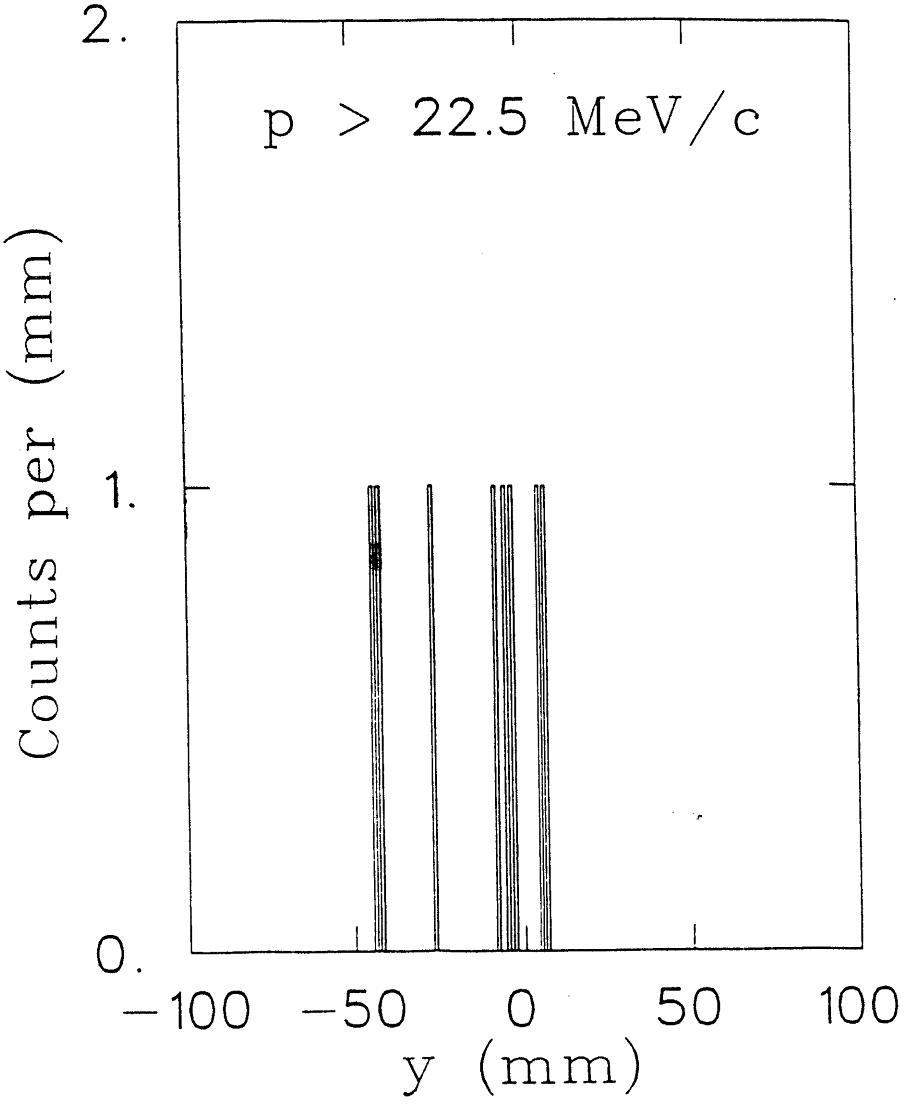


Figure 4.23: Position distribution of $\overline{M}$ data after momentum cut.

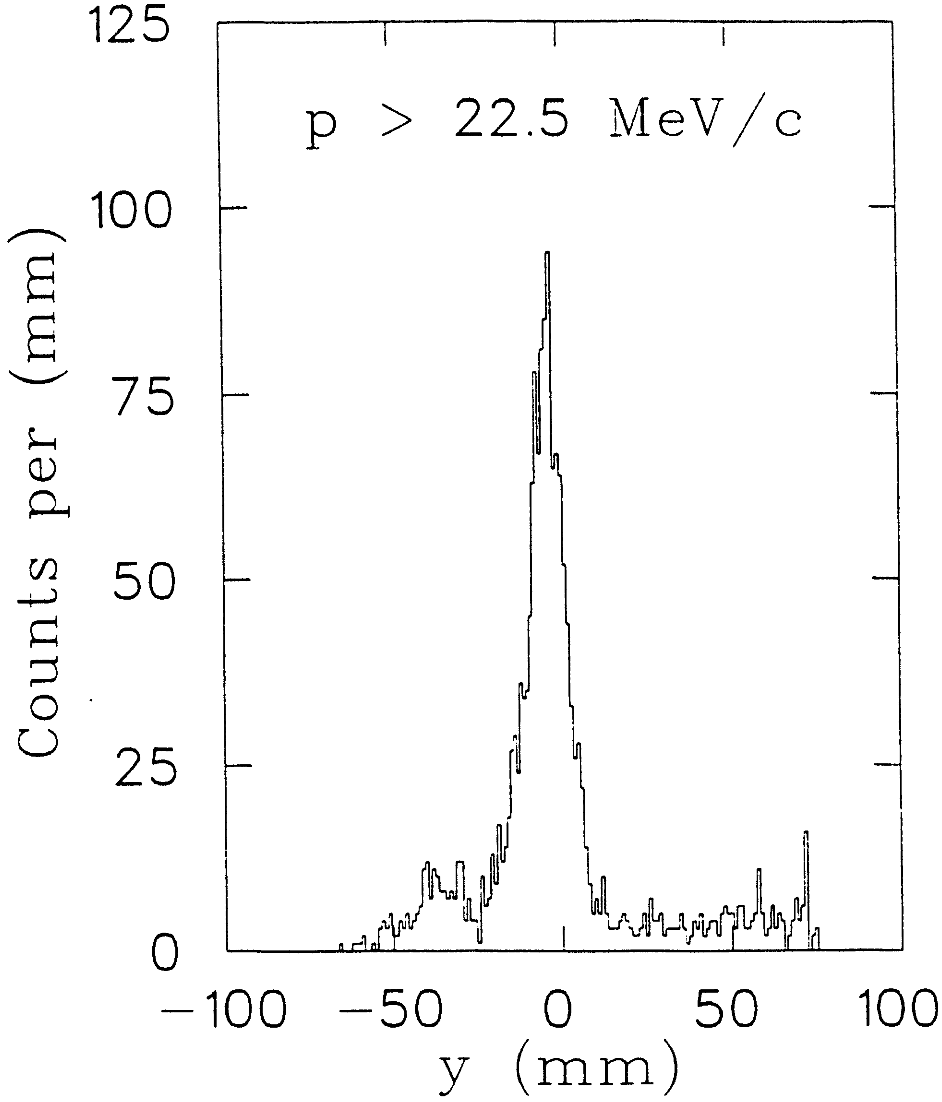


Figure 4.24: Background distribution passing momentum cut.

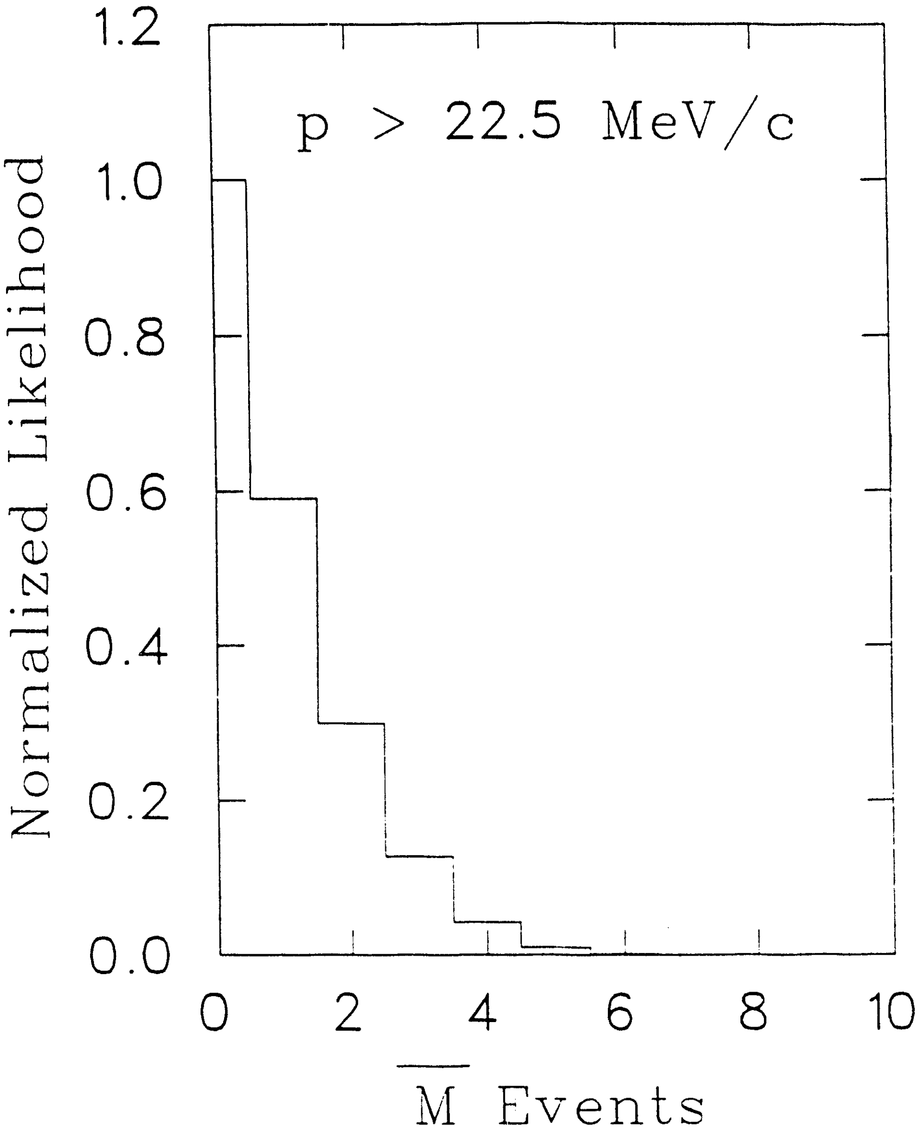


Figure 4.25: Likelihood function for fit to $\overline{M}$ data with momentum cut.

analyzed is the y-view of the reconstructed origin. However, the varying acceptance of the matrix trigger over the x-range of the reconstruction mixes in to the acceptance in y by the circular geometry of the vacuum window that constrains the valid decay origins. The problem is manifest in the failure of the M distribution to properly describe the distribution of counts in the vacuum region when the data were taken with the matrix trigger selecting against e^+ curvature.

To surmount this difficulty, instead of using the number of M atoms from the likelihood fits to $M \longrightarrow \overline{M}$ data directly in the normalization, the fitted target counts are taken as characteristic of the number of M atoms formed. Separately, with the MWPC trigger that does not suffer from the difficulty of uneven acceptance, the M formation fraction per target count is determined before and after a block of data taken searching for the conversion. The average M fraction is then used to translate the number of decays from the target to the number of M atoms that were formed.

Because the matrix trigger has unequal acceptance for e^+ from the target (used to find the number of M atoms) and e^- from the vacuum (used to find the number of $\overline{M}$ atoms), a correction factor accounting for this must be applied. This was determined separately for most of the SiO_2 targets used in data taking by taking M decays in the vacuum with the C-magnet reversed in polarity to simulate the vacuum decays of $\overline{M}$ atoms. The need for multiple determination arose because the location of the beam spot on the target along the x-axis was not necessarily constant for all targets. The ratio of the acceptance for e^- from the vacuum to that for e^+ from the target was averaged with equal weighting to obtain the value 60.08 ± 2.52 as the best estimate of this factor. Multiplying the number of $\overline{M}$ atoms by this correction then gives the number of M atoms that would have been observable as $\overline{M}$ events in the case of a conversion. The matrix trigger acceptance for e^- from the vacuum was about 66% and that for e^+ from the target was about 1.1%.

Since the full signature of the conversion also involves the detection of the atomic e^+ , the number of M atoms must be multiplied by the TOF acceptance for each target. This acceptance was measured by detecting thermal M with the coincidence signature and dividing the number of M counts that pass the TOF cut by their total.

target number	N_T (10^3)	$\frac{M}{T}$ (%)	ϵ_{TOF} (%)	N_M (10^3)
11	585 ± 2	3.71 ± 0.10	19.3 ± 1.2	252 ± 17
12	2207 ± 4	6.17 ± 0.08	15.3 ± 0.5	1249 ± 46
14	2212 ± 4	7.72 ± 0.33	13.0 ± 1.3	1328 ± 146
16	3476 ± 4	5.04 ± 0.14	16.1 ± 1.4	1694 ± 153
17	4835 ± 5	3.93 ± 0.09	18.1 ± 2.0	2063 ± 230
18	1408 ± 3	7.10 ± 0.12	8.1 ± 0.4	487 ± 27
average		5.39 ± 0.07	15.5 ± 0.8	
total				7073 ± 317

Table 4.2: Results for determination of number of M atoms formed during $\overline{M}$ data taking.

The conversions that are applied to the fitted number of target counts during the $M \longrightarrow \overline{M}$ data to get the number of M atoms formed that could produce observable conversions may be summarized by the equation

$$N_M = N_T \times \left(\frac{M}{T}\right) \times 60.08 \times \epsilon_{TOF} \;. \tag{4.20}$$

Table 4.2 summarizes the M formation per decay from the target and the TOF acceptances for each target together with the result for the number of "observable" M atoms. Averages have been formed by weighting with the number of fitted decays from the target.

The M formation per target count is not the true M formation fraction as usually quoted, since the M counts themselves are due to stopped μ^+ in the powder. For completeness, Table 4.3 gives the true formation fraction of M and the μ^+ stopping fraction for each target together with its projected thickness along the beam axis and the optimal momentum for M-formation.

Using these formation fractions and the spectrometer absolute acceptance of 2.5×10^{-3} the total number of incident and stopping μ^+ may be calculated. These are summarized in Table 4.4.

Now, the limit on $G_{M\overline{M}}$ may be calculated.

target number	thickness (mg/cm^2)	optimal momentum (MeV/c)	$\frac{\mu_{stop}}{\mu_{inc}}$ (%)	$\frac{M}{\mu_{stop}}$ (%)	$\frac{M}{\mu_{inc}}$ (%)
11	11.27	21.2	44.4 ± 2.7	3.58 ± 0.08	1.59 ± 0.09
12	5.64	20.5	45.0 ± 1.5	5.80 ± 0.07	2.61 ± 0.08
14	6.76	20.25	61.5 ± 3.5	7.16 ± 0.26	4.40 ± 0.19
16	10.71	21.0	63.7 ± 4.1	4.79 ± 0.12	3.05 ± 0.18
17	10.15	21.0	62.4 ± 5.0	3.56 ± 0.08	2.22 ± 0.17
18	9.58	20.75	40.4 ± 1.5	6.63 ± 0.10	2.68 ± 0.09
average			56.2 ± 1.7	5.02 ± 0.06	2.82 ± 0.08

Table 4.3: Summary of M formation characteristics for data taking targets.

target number	μ_{inc} (10^9)	μ_{stop} (10^9)
11	50 ± 5	22 ± 3
12	190 ± 11	86 ± 6
14	141 ± 22	87 ± 14
16	209 ± 29	133 ± 20
17	312 ± 54	194 ± 37
18	136 ± 11	55 ± 5
total	1038 ± 67	577 ± 45

Table 4.4: Total number of muons taken for each target.

4.10 Upper Limit on the Conversion

The total number of M atoms in the normalization has been determined to be $N_M = (7.07 \pm 0.32) \times 10^6$, corresponding to a total of 1.04×10^{12} incident μ^+. Of these M atoms, $(6.17 \pm 0.28) \times 10^6$ will have a decay e^+ with a momentum above 22.5 MeV/c. The most probable number of $\overline{M}$ atoms is zero and the upper limit on this number is $N_{\overline{M}} < 2$ at 90% confidence. Inserting this into Eq. 2.46 gives an upper limit on the probability per atom of a conversion at

$$S_{\overline{M}} < 6.5 \times 10^{-7} \ \ (90\% \ C.L.) \ , \tag{4.21}$$

where the suppression of the conversion by 50% due to the external field of about 10 G has been included. According to Eq. 2.48, this yields an upper limit of

$$G_{M\overline{M}} < 0.16 \ G_F \tag{4.22}$$

on the effective four-Fermion coupling constant of a $(V - A)$ interaction.

The vacuum region (11 to 50 mm) in the $\overline{M}$ data histogram which has been cut on $p > 22.5$ MeV/c contains no counts, so it is not possible to give a direct background estimate for this case. Without any momentum cut other than the spectrometer cutoff at around 10 MeV/c, there are 11 counts in this region. For the 7.07×10^6 M atoms in this data sample, this corresponds to a background of 1.6×10^{-6} events per M decay, which would limit the result at $G_{M\overline{M}} \sim 0.4 \ G_F$ if the maximum likelihood fit were not used to separate contributions of background and possible signal. It is clear that the method of maximum likelihood has the power to separate the contributions from processes described by appropriate reference distributions. Together with the leverage offered by the functioning track fitting in identifying useable spectrometer tracks, the maximum likelihood fit has proven itself a powerful tool in the analysis of this experiment.

Chapter 5

Results and Discussion

The experiment described in this dissertation has searched for the $M \longrightarrow \overline{M}$ conversion using a signature that required coincident detection of the decay e^- and the slow e^+ produced by $\overline{M}$ decay. To calibrate the apparatus, M atoms were observed by detecting both the decay e^+ and the atomic e^- from their breakup. The maximum likelihood analysis has shown that the most probable number of $M \longrightarrow \overline{M}$ conversion events seen is zero with less than 2 events at 90% confidence. Enough data were acquired to be successful in obtaining an improved upper limit on the conversion probability per atom of

$$S_{\overline{M}} < 6.5 \times 10^{-7} \quad (90\% \text{ C.L.}) \ . \tag{5.1}$$

This is an improvement of a factor of ~ 3 over the previous best limit [Hub90]. Assuming a conversion coupling of $(V - A)$ form, the corresponding upper limit on the effective four-Fermion coupling constant from the present experiment is

$$G_{M\overline{M}} < 0.16\ G_F \quad (90\% \text{ C.L.}) \ . \tag{5.2}$$

Thus, the $M \longrightarrow \overline{M}$ conversion is now constrained to be an appreciably rarer process than the most likely channel for muon decay, $\mu^+ \rightarrow e^+\nu_e\overline{\nu_\mu}$. The experimental progress of the upper limit on $G_{M\overline{M}}$ is depicted in Fig. 5.1. As the $M \longrightarrow \overline{M}$ conversion does not additively conserve lepton number for electron and muon families, this process is forbidden in the standard model. It is, however, permitted in a minimal left-right symmetric model [Moh81a,Her].

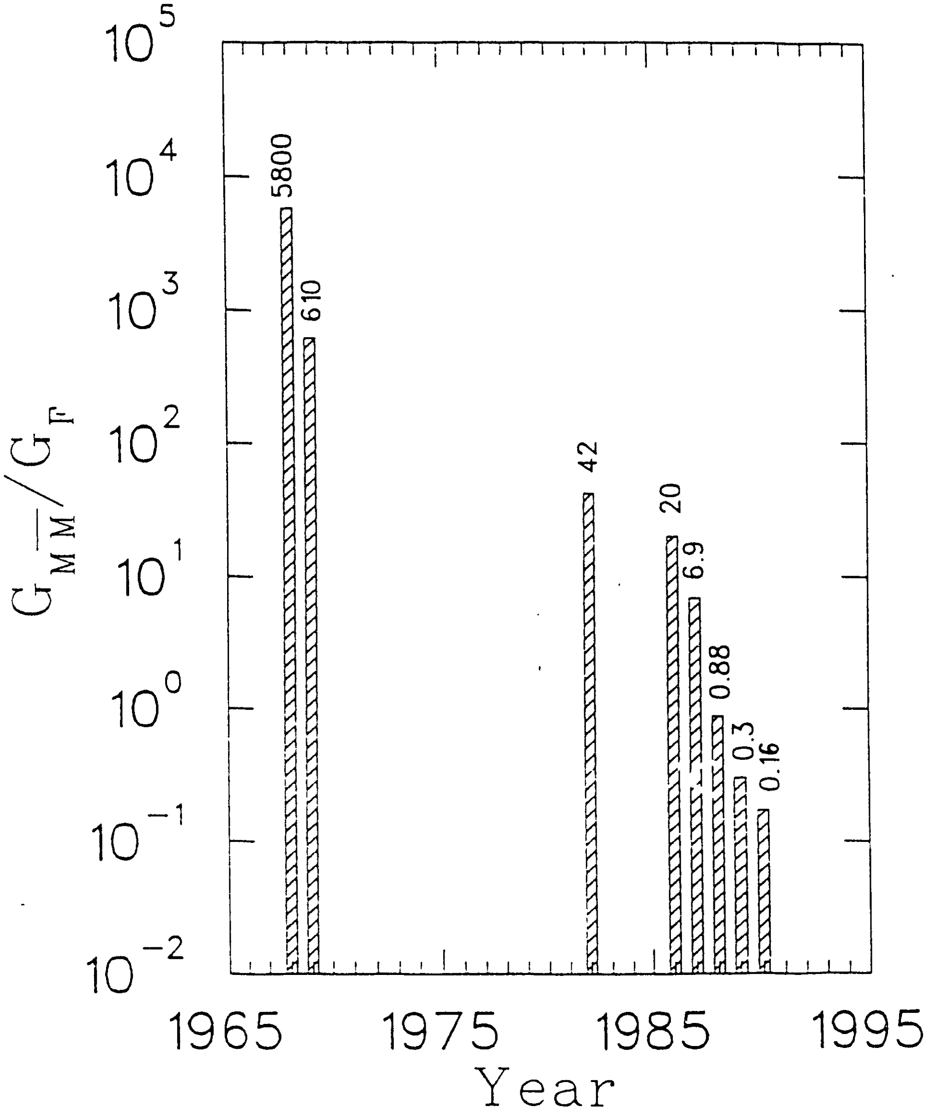


Figure 5.1: History of experimental upper limits on $G_{M\overline{M}}$.

In this model, there is a triplet of Higgs bosons which have lepton-number violating couplings. The doubly charged member of this triplet can mediate $M \longrightarrow \overline{M}$. The coupling constant for the conversion in this model is

$$G_{M\overline{M}} = \frac{\sqrt{2} f_{ee} f^*_{\mu\mu}}{8 m^2_{++}} \tag{5.3}$$

and the Fermi constant may be written as

$$G_F = \frac{\sqrt{2} g^2}{8 M^2_W} \; . \tag{5.4}$$

Therefore, the constraint this experiment places on the parameters of this model is

$$m_{++} > (202 \; GeV) \sqrt{\frac{f_{ee} f^*_{\mu\mu}}{g^2}} \; , \tag{5.5}$$

where $M_W = 80.9 \; GeV$ has been used for the mass of the W-boson. What makes matters exciting is that cosmological arguments and mass relations among the members of the Higgs triplet [Her91] in this model serve to place a *lower* bound on the $M \longrightarrow \overline{M}$ coupling. This lower bound is $G_{\overline{M}} > 4 \times 10^{-4} \; G_F$. Thus, the coupling constant for the conversion within this model is now constrained to a range of less than 3 decades.

The anomalous decay of the muon, $\mu^+ \rightarrow e^+ \overline{\nu_e} \nu_\mu$, may be considered as an analogous process to the $M \longrightarrow \overline{M}$ conversion, as it also violates additive muon and electron number conservation by two units each. Any conclusions about this process are, however, model-dependent [Her91]. Models exist which allow $M \longrightarrow \overline{M}$ while forbidding the anomalous muon decay in lowest order [Gel81,Cha89]. Under the assumption that the process $\mu^+ \rightarrow e^+ \overline{\nu_e} \nu_\mu$, proceeds with the same coupling constant as the $M \longrightarrow \overline{M}$ conversion, one may use the limit set by our experiment to speculate that the branching ratio for the anomalous muon decay is constrained by

$$\begin{aligned} R &= \frac{\Gamma(\mu^+ \rightarrow e^+ \overline{\nu_e} \nu_\mu)}{\Gamma(\mu^+ \rightarrow e^+ \nu_e \overline{\nu_\mu})} \\ &< \left(\frac{G_{M\overline{M}}}{G_F} \right)^2 \\ &< 3 \times 10^{-2} \; , \end{aligned} \tag{5.6}$$

where the estimate is straightforward since the phase space of the final state is the same for the ordinary and the anomalous muon decays. This limit is on the same order as the result of an experimental search [Wil80] for the decay $\mu^+ \rightarrow e^+\overline{\nu_e}\nu_\mu$.

Another interesting point to consider is the possibility of a contribution to the anomalous magnetic moment of the muon from couplings with a Δ^{++} Higgs boson. The lowest order contribution is a simple vertex correction to the muon-photon vertex that is proportional to

$$\frac{f_{\mu\mu}^2 m_\mu^2}{M_H^2} \; . \tag{5.7}$$

For a model with a singlet doubly charged Higgs boson, the contribution to the muon anomalous magnetic moment has been estimated [Cha89] as

$$\begin{aligned} \delta a_\mu &= \frac{1}{2}(g_\mu - 2) \\ &= -\frac{f_{\mu\mu}^2 m_\mu^2}{6\pi^2 m_{++}^2} \; . \end{aligned} \tag{5.8}$$

If one assumes that the couplings of the doubly charged Higgs boson to the muon and to the electron are equal, then our limit on the $M \longrightarrow \overline{M}$ conversion implies the upper limit

$$\delta a_\mu < 4 \times 10^{-9} \; . \tag{5.9}$$

The first use of the coincidence signature in searching for $M \longrightarrow \overline{M}$ has proven the approach to be feasible. An obvious improvement for a next-generation search for this process is to increase the acceptance of the detection of the decay e^- . This avenue is being pursued presently at the Paul-Scherrer-Institut (PSI, formerly SIN) [Jun89]. There, a new search for $M \longrightarrow \overline{M}$ is under construction that will use the same coincidence signature. The acceptance of the decay electron spectrometer (the SINDRUM I will be used) is approximately 300 times larger than that of this experiment. Furthermore, the signature of the atomic e^+ has been extended to include the observation of its annihilation γ's after its detection on a microchannel plate. The expected sensitivity of this experiment is in the range of $G_{M\overline{M}} \sim 10^{-3}\,G_F$ for the coupling constant, corresponding to $S_{\overline{M}} \sim 10^{-11}$. If it has not already, the search for $M \longrightarrow \overline{M}$ will come of age with this new experiment and join other rare muon decay searches as an effort of equal stature.

Appendix A

Neutral Scalar $M \longrightarrow \overline{M}$ Coupling

To date, the magnetic field dependence of the $M \longrightarrow \overline{M}$ conversion probability has been calculated only under the assumption of an interaction of $(V - A)$ form [Mor66,Ni88b,Sch88]:

$$\mathcal{H}_{M\overline{M}} = \frac{G_{M\overline{M}}}{\sqrt{2}}\, \overline{\mu}\, \gamma_\lambda\, (1 + \gamma_5)\, e\, \overline{\mu}\, \gamma^\lambda\, (1 + \gamma_5)\, e \; + \; H.c. \tag{A.1}$$

Since the form of the conversion matrix element resulting from this interaction, including the action on the spin projection of the particles involved, has been given as [Mor66]

$$\langle \overline{M}\ ; m'_{S_e}, m'_{S_\mu} | H_{M\overline{M}} | M; m_{S_e}, m_{S_\mu} \rangle = \left(\frac{\delta}{2}\right)\, \delta_{m'_{S_e}, m_{S_e}} \delta_{m'_{S_\mu}, m_{S_\mu}} \,, \tag{A.2}$$

one is led to ask for an alternative form that may result from the assumption of another type of interaction. In particular, another possibility seems to be

$$\langle \overline{M}\ ; m'_{S_e}, m'_{S_\mu} | H_{M\overline{M}} | M; m_{S_e}, m_{S_\mu} \rangle = \left(\frac{\eta}{2}\right)\, \delta_{m'_{S_e}, m_{S_\mu}} \delta_{m'_{S_\mu}, m_{S_e}} \,. \tag{A.3}$$

A completely general treatment of an arbitrary conversion coupling must include scalar, pseudoscalar, vector, axial vector, and tensor bilinear forms in arbitrary linear combination. One might expect the resulting matrix element to have the form of a linear combination of the two different forms given above, though it seems possible that conditions relating the incident and outgoing spins may occur.

The full calculation has not been done, but in order to demonstrate that there is an alternative to the usually assumed form of the conversion matrix element, the calculation of the matrix element for a neutral scalar conversion coupling is presented.

Let us call the neutral scalar boson being exchanged X^0. Then we can draw tree-level diagrams for s-, t-, and u-channels of $\mu^+ \, e^- \longrightarrow \mu^- \, e^+$ as shown in Fig. A.1. First, take the relevant interaction Lagrangian density to have the form

$$\mathcal{L}_{M\overline{M}} = f_{\mu e} \, \overline{\mu} \, e \, X^0 + H.c. \tag{A.4}$$

Then, consider the scalar propagator in momentum space,

$$\Delta_X(k) = \frac{1}{k^2 - m_X^2} \, , \tag{A.5}$$

which has the low energy limit ($k^2 \ll m_X^2$, that is, far below mass shell) of

$$\Delta_X(k) \longrightarrow \frac{-1}{m_X^2} \, . \tag{A.6}$$

The coordinate space propagator is

$$\Delta_X(x - x') = \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot(x-x')}}{k^2 - m_X^2 + i\epsilon} \, , \tag{A.7}$$

where the $i\epsilon$ term represents the prescription for the integral over k^0 to avoid the singularity on mass shell after doing the integral over $\vec{k}$. This propagator has the low energy limit

$$\Delta_X(x - x') \longrightarrow \frac{-1}{m_X^2} \, \delta^{(4)}(x - x') \, , \tag{A.8}$$

and so exhibits the same point-like nature as the $V - A$ coupling. Thus, the low energy effective interaction Hamiltonian density is

$$\mathcal{H}_{M\overline{M}} = \frac{f_{\mu e}^2}{m_X^2} \, \overline{\mu} \, e \, \overline{\mu} \, e \; + \; H.c. \tag{A.9}$$

The low energy effective interaction Hamiltonian is then given by

$$H_{M\overline{M}} = \int d^3\vec{x} \, \mathcal{H}_{M\overline{M}}(x) \, . \tag{A.10}$$

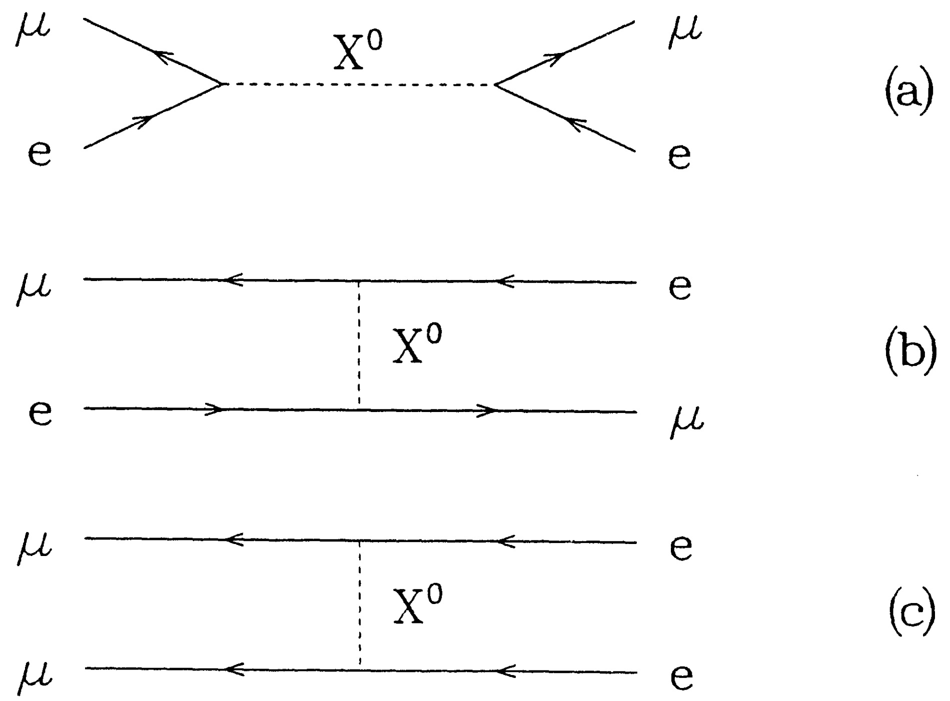


Figure A.1: $M \longrightarrow \overline{M}$ conversion mediated by neutral scalar boson in: (a) s-channel, (b) t-channel, and (c) u-channel (this is actually the process $\mu\,\mu \rightarrow e\,e$).

To be specific about conventions, we shall use the Dirac representation of the γ-matrices and the following form of the plane wave expansion to the Fermion fields:

$$e(x) = \sum_s \int \frac{d^3\vec{p}}{(2\pi)^3} \sqrt{\frac{m_e}{E_e(\vec{p})}} \left[b_e(p,s)\, u_e(p,s)\, e^{-ip\cdot x} + d_e^\dagger(p,s)\, v_e(p,s)\, e^{ip\cdot x} \right] \quad (A.11)$$

and

$$\bar{e}(x) = \sum_s \int \frac{d^3\vec{p}}{(2\pi)^3} \sqrt{\frac{m_e}{E_e(\vec{p})}} \left[b_e^\dagger(p,s)\, \bar{u}_e(p,s)\, e^{ip\cdot x} + d_e(p,s)\, \bar{v}_e(p,s)\, e^{-ip\cdot x} \right] , \quad (A.12)$$

and similarly for the μ and $\bar{\mu}$ fields. Here, $b_e^\dagger$ ($b_\mu^\dagger$) and b_e (b_μ) are the creation and annihilation operators, respectively, for e^- (μ^-) in Fock space. The operators $d_e^\dagger$ ($d_\mu^\dagger$) and d_e (d_μ) are the creation and annihilation operators for e^+ (μ^+). The anti-commutation relations among these operators are

$$\begin{aligned}
\left\{ b(p,s), b^\dagger(p',s') \right\} &= \delta_{s,s'}\, \delta^{(3)}(\vec{p}-\vec{p}\,') \\
\left\{ d(p,s), d^\dagger(p',s') \right\} &= \delta_{s,s'}\, \delta^{(3)}(\vec{p}-\vec{p}\,') \\
\{ b(p,s), b(p',s') \} &= 0 \\
\{ d(p,s), d(p',s') \} &= 0 \\
\left\{ b^\dagger(p,s), b^\dagger(p',s') \right\} &= 0 \\
\left\{ d^\dagger(p,s), d^\dagger(p',s') \right\} &= 0 \\
\{ b(p,s), d(p',s') \} &= 0 \\
\left\{ b(p,s), d^\dagger(p',s') \right\} &= 0 \\
\{ d(p,s), b(p',s') \} &= 0 \\
\left\{ d(p,s), b^\dagger(p',s') \right\} &= 0 ,
\end{aligned} \quad (A.13)$$

for each particle type separately. Muon and electron operators, however, commute with one another. The energy appearing in the plane wave expansions is given by

$$E(\vec{p}) = \sqrt{|\vec{p}|^2 + m^2} \,. \quad (A.14)$$

The Dirac spinors satisfy

$$(\not{p} - m)\, u(p,s) = 0 = \bar{u}(p,s)\,(\not{p} - m) \quad (A.15)$$

and

$$(\not{p} + m)\, v(p,s) = 0 = \bar{v}(p,s)\,(\not{p} + m)\,. \tag{A.16}$$

The relevant orthogonality relations are

$$\begin{aligned} \bar{u}(p,s)\, u(p,s') &= \delta_{s,s'} \\ \bar{v}(p,s)\, v(p,s') &= -\delta_{s,s'} \\ \bar{u}(p,s)\, v(p,s') &= 0 \\ \bar{v}(p,s)\, u(p,s') &= 0\,. \end{aligned} \tag{A.17}$$

Plane wave solutions to the Dirac equation have the form

$$u(p,s) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m} \end{pmatrix} \chi_s \tag{A.18}$$

and

$$v(p,s) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m} \\ 1 \end{pmatrix} \chi_s\,. \tag{A.19}$$

In these expressions, $\vec{\sigma}$ designates the Pauli matrices and χ_s, with $s \in \{1,2\}$, are the orthonormal 2-spinors that represent helicities of ±1. We now have all the required tools.

To begin the calculation, we specify the initial and final states. This will be most convenient if the center-of-mass coordinate frame (relative coordinates) is chosen. Let the expansion of the 1S spatial wave function in momentum space plane waves be given by

$$\phi_{1S}(\vec{x}\,) = \int \frac{d^3\vec{p}}{(2\pi)^3}\, \phi_{1S}(\vec{p}\,)\, e^{i\vec{p}\cdot\vec{x}} \tag{A.20}$$

for M and

$$\phi_{1S}(\vec{x}\,'\,) = \int \frac{d^3\vec{p}\,'}{(2\pi)^3}\, \phi_{1S}(\vec{p}\,'\,)\, e^{i\vec{p}\,'\cdot\vec{x}\,'} \tag{A.21}$$

for $\overline{M}$. This means that the initial state of M may be represented by

$$|M, 1S; s_e, s_\mu\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3}\, \phi_{1S}(\vec{p}\,) b_e^\dagger(p_e, s_e) d_\mu^\dagger(p_\mu, s_\mu)|0\rangle \tag{A.22}$$

and the final state may be specified by

$$|\overline{M}, 1S; s'_e, s'_\mu\rangle = \int \frac{d^3\vec{p}'}{(2\pi)^3} \phi_{1S}(\vec{p}')d^\dagger_e(p'_e, s'_e)b^\dagger_\mu(p'_\mu, s'_\mu)|0\rangle \ , \qquad (A.23)$$

where $|0\rangle$ is the Fock vacuum. In the center of mass system, the relative momenta are those in the integrations above; specifically

$$\vec{p} = \vec{p}_e - \vec{p}_\mu \qquad (A.24)$$

for M and

$$\vec{p}' = \vec{p}'_e - \vec{p}'_\mu \qquad (A.25)$$

for $\overline{M}$. The center-of-mass momenta are

$$\vec{P} = \vec{p}_e + \vec{p}_\mu = 0 \qquad (A.26)$$

for M and

$$\vec{P}' = \vec{p}'_e + \vec{p}'_\mu = 0 \qquad (A.27)$$

for $\overline{M}$.

Since the calculation is being carried out in the low energy limit, it is useful to note the low energy form of the Dirac spinors:

$$u(p, s) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \chi_s \qquad (A.28)$$

and

$$v(p, s) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \chi_s \qquad (A.29)$$

The resulting low energy form of the Fermion field operators in the plane wave expansion is

$$e(x) = \sum_s \int \frac{d^3\vec{p}}{(2\pi)^3} \left[b_e(p, s)u_e(p, s)e^{-ip\cdot x} + d^\dagger_e(p, s)v_e(p, s)e^{ip\cdot x}\right] \qquad (A.30)$$

and

$$\overline{e}(x) = \sum_s \int \frac{d^3\vec{p}}{(2\pi)^3} \left[b^\dagger_e(p, s)\overline{u}_e(p, s)e^{ip\cdot x} + d_e(p, s)\overline{v}_e(p, s)e^{-ip\cdot x}\right] \ , \qquad (A.31)$$

for the electron fields and similarly for the muon fields. The integrations are to be understood as restricted to a range of $|\vec{p}| \ll m_e$ to accord to the low energy limit.

The matrix element sought is

$$\mathcal{M}_{s'_e s'_\mu s_e s_\mu} = \langle \overline{M}, 1S; s'_e, s'_\mu | H_{M\overline{M}} | M, 1S; s_e, s_\mu \rangle \ . \tag{A.32}$$

Inserting explicitly the initial and final states and the interaction Hamiltonian gives

$$\begin{aligned} \mathcal{M}_{s'_e s'_\mu s_e s_\mu} &= \frac{f^2_{\mu e}}{m^2_X} \int d^3\vec{x} \int \frac{d^3\vec{p}\,'}{(2\pi)^3}\, \phi^*_{1S}(\vec{p}\,') \int \frac{d^3\vec{p}}{(2\pi)^3}\, \phi_{1S}(\vec{p}) \\ &\times \ \langle 0 \left| b_\mu(p'_\mu, s'_\mu) d_e(p'_e, s'_e) \, [\overline{\mu}\, e\, \overline{\mu}\, e \right. \\ &+ \ \left. \overline{e}\, \mu\, \overline{e}\, \mu] \, b^\dagger_e(p_e, s_e) d^\dagger_\mu(p_\mu, s_\mu) \right| 0 \rangle \ . \end{aligned} \tag{A.33}$$

The second term vanishes; it would induce $\overline{M} \longrightarrow M$ transitions. With the explicit low energy form of the Fermion field operators this becomes

$$\begin{aligned} \mathcal{M}_{s'_e s'_\mu s_e s_\mu} &= \frac{f^2_{\mu e}}{m^2_X} \int d^3\vec{x} \sum_{s_1} \sum_{s_2} \sum_{s_3} \sum_{s_4} \\ &\times \int \frac{d^3\vec{p}\,'}{(2\pi)^3}\, \phi^*_{1S}(\vec{p}\,') \int \frac{d^3\vec{p}}{(2\pi)^3}\, \phi_{1S}(\vec{p}) \\ &\times \int \frac{d^3\vec{p}_1}{(2\pi)^3} \int \frac{d^3\vec{p}_2}{(2\pi)^3} \int \frac{d^3\vec{p}_3}{(2\pi)^3} \int \frac{d^3\vec{p}_4}{(2\pi)^3} \\ &\times \langle 0 | b_\mu(p'_\mu, s'_\mu) d_e(p'_e, s'_e) \\ &\times \left[b^\dagger_{\mu_1}(p_1, s_1) \overline{u}_{\mu_1}(p_1, s_1) e^{ip_1 \cdot x} + d_{\mu_1}(p_1, s_1) \overline{v}_{\mu_1}(p_1, s_1) e^{-ip_1 \cdot x} \right] \\ &\times \left[b_{e_2}(p_2, s_2) u_{e_2}(p_2, s_2) e^{-ip_2 \cdot x} + d^\dagger_{e_2}(p_2, s_2) v_{e_2}(p_2, s_2) e^{ip_2 \cdot x} \right] \\ &\times \left[b^\dagger_{\mu_3}(p_3, s_3) \overline{u}_{\mu_3}(p_3, s_3) e^{ip_3 \cdot x} + d_{\mu_3}(p_3, s_3) \overline{v}_{\mu_3}(p_3, s_3) e^{-ip_3 \cdot x} \right] \\ &\times \left[b_{e_4}(p_4, s_4) u_{e_4}(p_4, s_4) e^{-ip_4 \cdot x} + d^\dagger_{e_4}(p_4, s_4) v_{e_4}(p_4, s_4) e^{ip_4 \cdot x} \right] \\ &\times b^\dagger_e(p_e, s_e) d^\dagger_\mu(p_\mu, s_\mu) | 0 \rangle \ . \end{aligned} \tag{A.34}$$

Of the 16 terms in this equation, only 4 have non-zero occupation number matrix elements. Without the integrations and streamlining the notation somewhat, these are

$$T_1 = \langle 0 | b'_\mu d'_e b^\dagger_{\mu_1} b_{e_2} d_{\mu_3} d^\dagger_{e_4} b^\dagger_e d^\dagger_\mu | 0 \rangle \, (\overline{u}_{\mu_1} u_{e_2} \overline{v}_{\mu_3} v_{e_4}) \, e^{i(p_1 - p_2 - p_3 + p_4) \cdot x} \tag{A.35}$$

$$T_2 \;=\; \langle 0|b'_\mu d'_e b^\dagger_{\mu_1} d^\dagger_{e_2} d_{\mu_3} b_{e_4} b^\dagger_e d^\dagger_\mu |0\rangle \,(\overline{u}_{\mu_1} v_{e_2} \overline{v}_{\mu_3} u_{e_4})\, e^{i(p_1+p_2-p_3-p_4)\cdot x}$$

$$T_3 \;=\; \langle 0|b'_\mu d'_e d_{\mu_1} b_{e_2} b^\dagger_{\mu_3} d^\dagger_{e_4} b^\dagger_e d^\dagger_\mu |0\rangle \,(\overline{v}_{\mu_1} u_{e_2} \overline{u}_{\mu_3} v_{e_4})\, e^{i(-p_1-p_2+p_3+p_4)\cdot x}$$

$$T_4 \;=\; \langle 0|b'_\mu d'_e d_{\mu_1} d^\dagger_{e_2} b^\dagger_{\mu_3} b_{e_4} b^\dagger_e d^\dagger_\mu |0\rangle \,(\overline{v}_{\mu_1} v_{e_2} \overline{u}_{\mu_3} u_{e_4})\, e^{i(-p_1+p_2+p_3-p_4)\cdot x} \;. \tag{A.36}$$

These terms reduce to

$$\begin{aligned} T_1 \;&=\; (-1)\,\delta_{s'_\mu,s_1}\,\delta^{(3)}(\vec{p}'_\mu-\vec{p}_1)\,\delta_{s'_e,s_4}\,\delta^{(3)}(\vec{p}'_e-\vec{p}_4) \\ &\times\; \delta_{s_2,s_e}\,\delta^{(3)}(\vec{p}_2-\vec{p}_e)\,\delta_{s_3,s_\mu}\,\delta^{(3)}(\vec{p}_3-\vec{p}_\mu) \\ &\times\; (-\delta_{s_1,s_2}\,\delta_{s_3,s_4})\,e^{i(p_1-p_2-p_3+p_4)\cdot x} \end{aligned} \tag{A.37}$$

$$\begin{aligned} T_2 \;&=\; 0 \\[2ex] T_3 \;&=\; 0 \\[2ex] T_4 \;&=\; (-1)\,\delta_{s'_\mu,s_3}\,\delta^{(3)}(\vec{p}'_\mu-\vec{p}_3)\,\delta_{s'_e,s_2}\,\delta^{(3)}(\vec{p}'_e-\vec{p}_2) \\ &\times\; \delta_{s_1,s_\mu}\,\delta^{(3)}(\vec{p}_1-\vec{p}_\mu)\,\delta_{s_4,s_e}\,\delta^{(3)}(\vec{p}_4-\vec{p}_e) \\ &\times\; (-\delta_{s_1,s_2}\,\delta_{s_3,s_4})\,e^{i(-p_1+p_2+p_3-p_4)\cdot x} \end{aligned} \tag{A.38}$$

when applying the anticommutation relations on the creation and annihilation operators and using the orthonormality of the two-spinors χ_s. Carrying out the spin sums over s_1, s_2, s_3, s_4, doing the integral over $d^3\vec{x}$, and assuming that the conversion takes place at time $t = 0$, we obtain

$$\begin{aligned} \mathcal{M}_{s'_e s'_\mu s_e s_\mu} \;&=\; \frac{f^2_{\mu e}}{m^2_X}\int\frac{d^3\vec{p}'}{(2\pi)^3}\,\phi^*_{1S}(\vec{p}'\,)\int\frac{d^3\vec{p}}{(2\pi)^3}\,\phi_{1S}(\vec{p}\,) \\ &\times\; \int\frac{d^3\vec{p}_1}{(2\pi)^3}\int\frac{d^3\vec{p}_2}{(2\pi)^3}\int\frac{d^3\vec{p}_3}{(2\pi)^3}\int\frac{d^3\vec{p}_4}{(2\pi)^3} \end{aligned} \tag{A.39}$$

$$\begin{aligned} &\times \quad \delta_{s'_\mu, s_e}\, \delta_{s'_e, s_\mu}\, (2\pi)^3 \delta^{(3)}(\vec{p}_1 - \vec{p}_2 - \vec{p}_3 + \vec{p}_4) \\ &\times \quad \Big[\delta^{(3)}(\vec{p}\,'_\mu - \vec{p}_1)\, \delta^{(3)}(\vec{p}\,'_e - \vec{p}_4)\, \delta^{(3)}(\vec{p}_2 - \vec{p}_e)\, \delta^{(3)}(\vec{p}_3 - \vec{p}_\mu) + \\ &\qquad \delta^{(3)}(\vec{p}\,'_\mu - \vec{p}_3)\, \delta^{(3)}(\vec{p}\,'_e - \vec{p}_2)\, \delta^{(3)}(\vec{p}_1 - \vec{p}_\mu)\, \delta^{(3)}(\vec{p}_4 - \vec{p}_e)\Big] \ . \end{aligned}$$

Doing the integrals over the momenta $\vec{p}_1$, $\vec{p}_2$, $\vec{p}_3$, and $\vec{p}_4$ and using the definition of the center of mass momentum gives

$$\begin{aligned} \mathcal{M}_{s'_e s'_\mu s_e s_\mu} &= \frac{f^2_{\mu e}}{m^2_X} \int \frac{d^3\vec{p}\,'}{(2\pi)^3}\, \phi^*_{1S}(\vec{p}\,') \int \frac{d^3\vec{p}}{(2\pi)^3}\, \phi_{1S}(\vec{p}) \\ &\times \quad 2\, \delta_{s'_\mu, s_e}\, \delta_{s'_e, s_\mu}\, (2\pi)^3\, \delta^{(3)}(\vec{P}\,' - \vec{P}) \ . \end{aligned} \qquad \text{(A.40)}$$

Now, the integrals over $\vec{p}$ and $\vec{p}\,'$ can be done. Inspecting the definition of the momentum components, $\phi_{1S}(\vec{p})$, of the spatial wavefunction in Eqs. A.20 and A.21 shows that

$$\phi_{1S}(\vec{0}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\, \phi_{1S}(\vec{p}) \qquad \text{(A.41)}$$

which results in

$$\mathcal{M}_{s'_e s'_\mu s_e s_\mu} = \frac{f^2_{\mu e}}{m^2_X}\, 2\, \delta_{s'_\mu, s_e}\, \delta_{s'_e, s_\mu}\, (2\pi)^3\, \delta^{(3)}(\vec{P}\,' - \vec{P})\, \left|\phi_{1S}(\vec{0})\right|^2 \ . \qquad \text{(A.42)}$$

The probability of finding the e^- overlapping the μ^+ enters as a result of the low energy limit which has rendered the coupling a contact interaction. One can integrate the matrix element we have found over the final state center of mass momentum to remove the remaining delta function. The result is

$$\mathcal{M}'_{s'_e s'_\mu s_e s_\mu} = 2\, \frac{f^2_{\mu e}}{m^2_X}\, \left|\phi_{1S}(\vec{0})\right|^2\, \delta_{s'_\mu, s_e}\, \delta_{s'_e, s_\mu} \ . \qquad \text{(A.43)}$$

The leading factor of 2 is a result of the nonvanishing terms T_1 and T_4 of above and may be interpreted as coming from equal contributions of s- and t-channels (see Fig. A.1).

Thus, we have shown that there is an alternative possibility for the action on the spins in a conversion coupling matrix element. In particular, for a neutral scalar coupling, we have shown that

$$\langle \overline{M}; m'_{S_e}, m'_{S_\mu} | H_{M\overline{M}} | M; m_{S_e}, m_{S_\mu} \rangle = \left(\frac{\eta}{2}\right)\, \delta_{m'_{S_e}, m_{S_\mu}}\, \delta_{m'_{S_\mu}, m_{S_e}} \ , \qquad \text{(A.44)}$$

where the helicities have been translated to the projection quantum numbers of the uncoupled angular momentum basis. The parameter η is given by

$$\frac{\eta}{2} = 2\,\frac{f_{\mu e}^2}{m_X^2}\,\left|\phi_{1S}(\vec{0})\right|^2 \;. \tag{A.45}$$

Inserting the form of the 1S spatial wave function, defining a new coupling constant, $G'_{M\overline{M}}$, and substituting for all other constants gives

$$\frac{\eta}{2} = \left(3.8 \times 10^{-13}\, eV\right) \times \left(\frac{G'_{M\overline{M}}}{G_F}\right) \;, \tag{A.46}$$

where

$$G'_{M\overline{M}} = \frac{f_{\mu e}^2}{m_X^2} \;. \tag{A.47}$$

If one were to work out the magnetic field dependence of the conversion probability under the assumption of a neutral scalar conversion coupling, I expect that the same low field behavior as for a $(V-A)$ coupling would be found (because of the symmetry of the M and $\overline{M}$ $m_F = \pm 1$ states under spin interchange of $m_{S_\mu} \leftrightarrow m_{S_e}$), but that the high field conversion would be unaffected (see Sec. 2.4). So, again, the important conclusion seems to be that the magnetic field dependence of the conversion probability is a model-dependent result, just as is the definition of a coupling constant for the conversion.

Appendix B

Effect of External Magnetic Field on $M \longrightarrow \overline{M}$

B.1 The M Atom in a Magnetic Field

The non-relativistic Hamiltonian for the M atom in an external magnetic field is

$$H = H_0 + H_{hf} + H_Z \ , \tag{B.1}$$

where

$$H_0 = \frac{p^2}{2m} - \frac{e^2}{r} \tag{B.2}$$

describes the unperturbed hydrogenic Bohr atom in relative coordinates and CGS units. The *reduced mass* of the system is defined as usual by

$$m = \left[\frac{1}{m_e} + \frac{1}{m_\mu}\right]^{-1} \tag{B.3}$$

and the eigenenergies of H_0 are

$$E_0 = -\frac{1}{n^2}\,\frac{\alpha^2 mc^2}{2} = -\frac{1}{n^2}\,\frac{me^4}{2\hbar^2} \ . \tag{B.4}$$

These energies depend only on the *principal quantum number*, n. The Fermi contact portion of the hyperfine interaction between the electron and the muon spin is given by

$$H_{hf} = a\,\vec{S}_\mu \cdot \vec{J} \ , \tag{B.5}$$

where a will be referred to as the *hyperfine structure interval*, $\vec{S}_\mu$ is the muon spin operator, and $\vec{J}$ is the total electron angular momentum, $\vec{J} = \vec{L} + \vec{S}_e$. In the following, we will restrict ourselves to the ground state, for which $n = 1$ and the orbital angular momentum of the electron, $\vec{L}$, has an expectation value of zero. Thus, the expression for H_{hf} becomes

$$H_{hf} = a\,\vec{S}_\mu \cdot \vec{S}_e \ . \tag{B.6}$$

Lastly, H_Z describes the Zeeman effect, the interaction of the magnetic moments of the electron and of the muon with the external magnetic field, $\vec{B}$. Designating the magnetic moments by $\vec{\mu}_e$ and by $\vec{\mu}_\mu$, we have

$$H_Z = -\vec{\mu}_e \cdot \vec{B} \ - \ \vec{\mu}_\mu \cdot \vec{B} \ . \tag{B.7}$$

The magnetic moments are related to the spin angular momenta by

$$\vec{\mu}_e = -g_e\, \mu_B\, \vec{S}_e \tag{B.8}$$

for the electron and

$$\vec{\mu}_\mu = g_\mu\, \mu_B \left(\frac{m_e}{m_\mu}\right) \vec{S}_\mu \tag{B.9}$$

for the muon. Here, $\mu_B = \frac{e\hbar}{2m_ec}$ is the Bohr magneton, m_e and m_μ are the electron and muon masses, respectively. The g-factors for the electron and the muon differ by $\mathcal{O}(\alpha^2)$, so for our purposes they may be taken to be approximately equal:

$$g_\mu \approx g_e \approx 2(1 + \frac{\alpha}{2\pi}) = 2.00232 \ . \tag{B.10}$$

To express the eigenenergies of H_{hf} and of H_Z, we introduce two possible bases for the state vectors. The *uncoupled basis* is labeled by six eigenvalues:

$$|n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle \ . \tag{B.11}$$

Here, n is the principal quantum number ($n = 1$ for our case), l is the orbital angular momentum quantum number, such that

$$\vec{L}^2\ |n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle = l(l+1)\ |n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle \ , \tag{B.12}$$

s_e and s_μ are the eigenvalues of the electron and muon spins, respectively, with

$$\vec{S}_e^2 \; |n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle = s_e\,(s_e + 1) \; |n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle \tag{B.13}$$

and

$$\vec{S}_\mu^2 \; |n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle = s_\mu(s_\mu + 1) \; |n, l, s_e, s_\mu, m_{s_e}, m_{s_\mu}\rangle \ , \tag{B.14}$$

and, finally, m_{s_e} and m_{s_μ} are the eigenvalues of the z-projection of the electron and muon spins, respectively. Since we are considering only states with $n = 1$, $l = 0$, $s_e = \frac{1}{2}$, and $s_\mu = \frac{1}{2}$, a shorthand is appropriate:

$$\begin{aligned} |\uparrow\uparrow\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},+\tfrac{1}{2},+\tfrac{1}{2}\rangle \\ |\uparrow\downarrow\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},+\tfrac{1}{2},-\tfrac{1}{2}\rangle \\ |\downarrow\uparrow\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},+\tfrac{1}{2}\rangle \\ |\downarrow\downarrow\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\rangle \ . \end{aligned} \tag{B.15}$$

In this basis,

$$S_{ez} \; |m_{s_e}, m_{s_\mu}\rangle = m_{s_e} \; |m_{s_e}, m_{s_\mu}\rangle \tag{B.16}$$

and

$$S_{\mu z} \; |m_{s_e}, m_{s_\mu}\rangle = m_{s_\mu} \; |m_{s_e}, m_{s_\mu}\rangle \ . \tag{B.17}$$

After introducing the total angular momentum operator, $\vec{F}$, as

$$\vec{F} = \vec{S}_e + \vec{S}_\mu \ , \tag{B.18}$$

we can define the *coupled basis* as

$$|n, l, s_e, s_\mu, f, m_f\rangle \ . \tag{B.19}$$

Again, there is an appropriate shorthand notation:

$$\begin{aligned} |1,+1\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},1,+1\rangle \\ |1,0\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},1,0\rangle \\ |1,-1\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},1,-1\rangle \\ |0,0\rangle &= |1,0,\tfrac{1}{2},\tfrac{1}{2},0,0\rangle \ . \end{aligned} \tag{B.20}$$

In this basis,

$$\vec{F}^2\,|f,m_f\rangle = f(f+1)\,|f,m_f\rangle \tag{B.21}$$

and

$$F_z|f,m_f\rangle = m_f\,|f,m_f\rangle\ . \tag{B.22}$$

Now, the eigenenergies of H_{hf} may be expressed in the coupled basis as

$$\langle H_{hf}\rangle = E_{hf} = \frac{a}{2}\left[f(f+1) - \frac{3}{2}\right] \tag{B.23}$$

after making use of the identity

$$\vec{S}_\mu \cdot \vec{S}_e = \frac{1}{2}\left(\vec{F}^2 - \vec{S}^2_\mu - \vec{S}^2_e\right)\ . \tag{B.24}$$

The representation of H_{hf} is, however, not diagonal in the uncoupled basis. Since the coupled and uncoupled bases are related by the orthogonal transformation

$$|m_{s_e}, m_{s_\mu}\rangle = \sum_{f,m_f} |f,m_f\rangle\langle f,m_f|m_{s_e},m_{s_\mu}\rangle\ , \tag{B.25}$$

where the factors $\langle f,m_f|m_{s_e},m_{s_\mu}\rangle$ are the Clebsch-Gordan coefficients:

$$\begin{aligned}
\langle 1,+1|+\tfrac{1}{2}\,,+\tfrac{1}{2}\,\rangle &= 1\\
\langle 1,0|+\tfrac{1}{2}\,,-\tfrac{1}{2}\,\rangle &= \frac{1}{\sqrt{2}}\\
\langle 1,0|-\tfrac{1}{2}\,,+\tfrac{1}{2}\,\rangle &= \frac{1}{\sqrt{2}}\\
\langle 1,-1|-\tfrac{1}{2}\,,-\tfrac{1}{2}\,\rangle &= 1\\
\langle 0,0|+\tfrac{1}{2}\,,-\tfrac{1}{2}\,\rangle &= -\frac{1}{\sqrt{2}}\\
\langle 0,0|-\tfrac{1}{2}\,,+\tfrac{1}{2}\,\rangle &= \frac{1}{\sqrt{2}}\ ,
\end{aligned} \tag{B.26}$$

the operator H_{hf} may be represented in the uncoupled basis by

$$\begin{aligned}
H_{hf} \;=\; & \left(|+\tfrac{1}{2}\,,+\tfrac{1}{2}\,\rangle\frac{a}{4}\langle+\tfrac{1}{2}\,,+\tfrac{1}{2}\,|\right) + \left(|-\tfrac{1}{2}\,,-\tfrac{1}{2}\,\rangle\frac{a}{4}\langle-\tfrac{1}{2}\,,-\tfrac{1}{2}\,|\right)\\
- & \left(|+\tfrac{1}{2}\,,-\tfrac{1}{2}\,\rangle\frac{a}{4}\langle+\tfrac{1}{2}\,,-\tfrac{1}{2}\,|\right) - \left(|-\tfrac{1}{2}\,,+\tfrac{1}{2}\,\rangle\frac{a}{4}\langle-\tfrac{1}{2}\,,+\tfrac{1}{2}\,|\right)\\
+ & \left(|+\tfrac{1}{2}\,,-\tfrac{1}{2}\,\rangle\frac{a}{2}\langle-\tfrac{1}{2}\,,+\tfrac{1}{2}\,|\right) + \left(|-\tfrac{1}{2}\,,+\tfrac{1}{2}\,\rangle\frac{a}{2}\langle+\tfrac{1}{2}\,,-\tfrac{1}{2}\,|\right)\ .
\end{aligned} \tag{B.27}$$

The eigenenergies of H_Z are

$$\langle H_Z \rangle = E_Z = g_e \, \mu_B \, B \, m_{s_e} - g_\mu \, \mu_B \left(\frac{m_e}{m_\mu} \right) B \, m_{s_\mu} \tag{B.28}$$

in the uncoupled basis. Here, one can apply a corrolary of the Wigner-Eckart theorem, the *Landé formula* [Wei78]:

$$\langle j,m | \vec{A} \, | j,m\prime \rangle = \frac{\langle j,m | \vec{A} \cdot \vec{J} \, | j,m \rangle}{j\,(j+1)} \langle j,m | \, \vec{J} \, | j,m\prime \rangle \ , \tag{B.29}$$

where $\vec{A}$ is any vector operator, $\vec{J}$ is an angular momentum operator, j is its eigenvalue, and m is its z-projection eigenvalue. Applying this to H_Z in the coupled basis gives

$$\langle f, m_f | H_Z | f, m_f \rangle = \frac{1}{2} \left(g_e - \frac{m_e}{m_\mu} \, g_\mu \right) \mu_B \, B \, m_f \ , \tag{B.30}$$

which may be rewritten be as

$$\langle H_Z \rangle = g_F \, \mu_B \, B m_f \ , \tag{B.31}$$

where we have defined the effective g-factor

$$g_F = \frac{1}{2} \left(g_e - \frac{m_e}{m_\mu} \, g_\mu \right) \ . \tag{B.32}$$

At small magnetic fields, $B \ll \frac{a}{\mu_B}$, the coupled basis states well approximate the eigenstates of the full Hamiltonian (Eq. B.1), but at larger fields, $\frac{a}{\mu_B} \ll B \ll \frac{\alpha^2 mc^2}{2\mu_B}$, it is the uncoupled basis states that closely approximate the eigenstates of H. The latter upper limit derives from the condition that the Zeeman effect be treated perturbatively, that it not influence the structure of the atom grossly. At field values intermediate to these limits, $B \sim \frac{a}{\mu_B}$, we must carry out the diagonalization of H in either of the state vector bases to obtain its generally valid eigenstates and eigenenergies.

Choosing the uncoupled basis to express the full Hamiltonian gives

$$H = \begin{pmatrix} E_1 & 0 & 0 & 0 \\ 0 & E_2 & \frac{a}{2} & 0 \\ 0 & \frac{a}{2} & E_3 & 0 \\ 0 & 0 & 0 & E_4 \end{pmatrix} \begin{array}{c} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{array} \ , \tag{B.33}$$

where the rows are labeled by the basis vectors and the columns are arranged similarly. The diagonal entries are given by

$$\begin{aligned} E_1 &= E_0 + \frac{a}{2}\left(\frac{1}{2} + Y\right) \\ E_2 &= E_0 + \frac{a}{2}\left(-\frac{1}{2} + X\right) \\ E_3 &= E_0 + \frac{a}{2}\left(-\frac{1}{2} - X\right) \\ E_4 &= E_0 + \frac{a}{2}\left(\frac{1}{2} - Y\right) , \end{aligned} \tag{B.34}$$

where

$$X = \frac{\mu_B\, B}{a}\left(g_e + \frac{m_e}{m_\mu}\, g_\mu\right) \tag{B.35}$$

and

$$Y = \frac{\mu_B\, B}{a}\left(g_e - \frac{m_e}{m_\mu}\, g_\mu\right) . \tag{B.36}$$

The solutions to the diagonalization of H are the well-known Breit-Rabi states and energies, here designated by $|\,\lambda_i^{(M)}\,\rangle$ and $\lambda_i^{(M)}$, respectively:

$$\begin{aligned} |\,\lambda_1^{(M)}\,\rangle &= |\uparrow\uparrow\,\rangle \\ |\,\lambda_2^{(M)}\,\rangle &= \frac{1}{N}\left[\left(X + \sqrt{1+X^2}\right)|\uparrow\downarrow\,\rangle + |\downarrow\uparrow\,\rangle\right] \\ |\,\lambda_3^{(M)}\,\rangle &= \frac{1}{N}\left[|\uparrow\downarrow\,\rangle - \left(X + \sqrt{1+X^2}\right)|\downarrow\uparrow\,\rangle\right] \\ |\,\lambda_4^{(M)}\,\rangle &= |\downarrow\downarrow\,\rangle \\ N &= \left[1 + \left(X + \sqrt{1+X^2}\right)^2\right]^{\frac{1}{2}} \end{aligned} \tag{B.37}$$

$$\begin{aligned} \lambda_1^{(M)} &= E_0 + \frac{a}{2}\left(\frac{1}{2} + Y\right) \\ \lambda_2^{(M)} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} + \sqrt{1+X^2}\right) \\ \lambda_3^{(M)} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} - \sqrt{1+X^2}\right) \\ \lambda_4^{(M)} &= E_0 + \frac{a}{2}\left(\frac{1}{2} - Y\right) . \end{aligned} \tag{B.38}$$

The Breit-Rabi energy levels are shown in Fig. B.1 as a function of the magnitude of the external magnetic field. It is straightforward to verify that, as $B \longrightarrow 0$, the

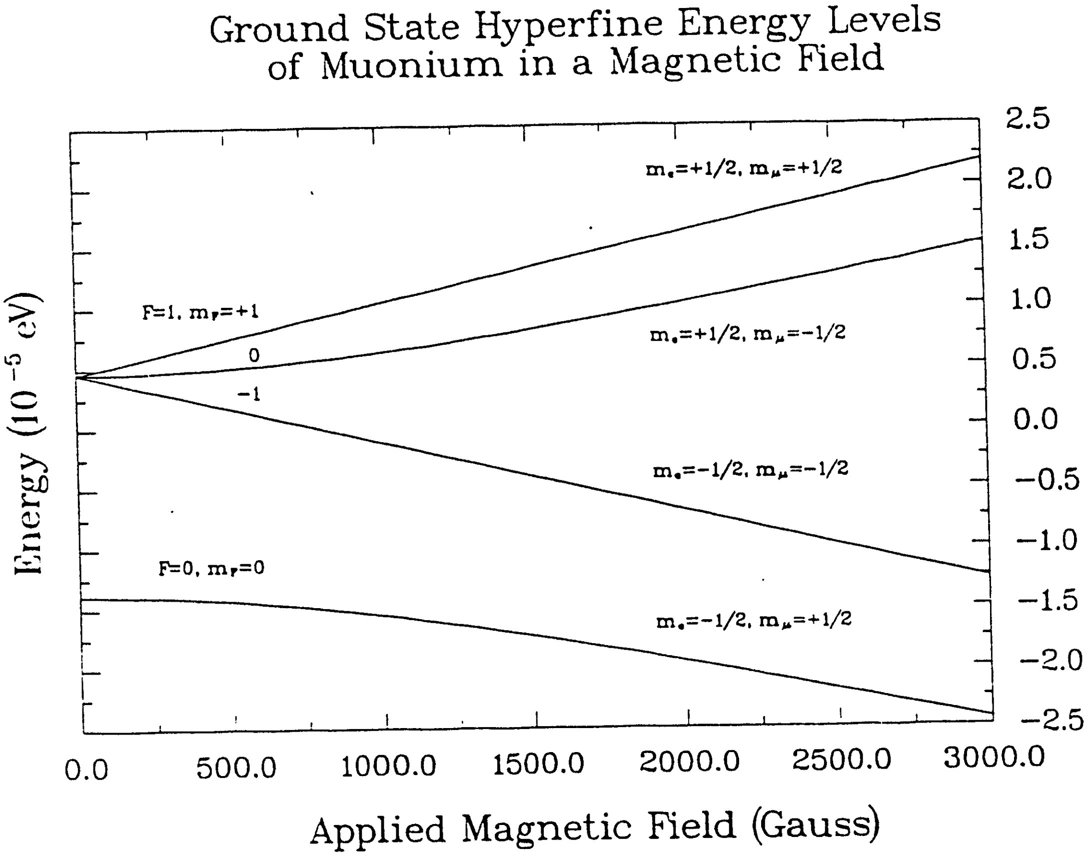


Figure B.1: Energy levels of ground state muonium in an external magnetic field.

Breit-Rabi eigenstates and -energies tend to the coupled states and energies and that, for $B \gg \frac{a}{\mu_B}$, they approach the uncoupled states and energies. One concludes that an increasing external field decreases the extent to which the coupling of the muon and electron spins affects the energy levels of the M atom in favor of an increasing interaction of the spins with the applied magnetic field. The external field *decouples* the spins.

B.2 The $\overline{M}$ Atom in a Magnetic Field

The $\overline{M}$ atom differs from the M atom only by the reversal of the charges of the muon and of the electron. Thus, the field-dependent description of the $\overline{M}$ atom may be obtained easily by noting that the only change needed in the above calculation for the M atom is a reversal of the signs of the magnetic moments of the muon and the electron from Eqs. B.8 and B.9:

$$\begin{aligned} \vec{\mu}_{e^+} &= g_e\, \mu_B\, \vec{S}_e && \text{(B.39)} \\ \vec{\mu}_{\mu^-} &= -g_\mu\, \mu_B \left(\frac{m_e}{m_\mu}\right) \vec{S}_\mu\ . && \text{(B.40)} \end{aligned}$$

Thus, the prescription for transforming the results for M to $\overline{M}$ is $X \rightarrow -X$, $Y \rightarrow -Y$, and $\sqrt{1+X^2} \rightarrow -\sqrt{1+X^2}$. The last sign change seems arbitrary, but it is allowed by making the opposite sign choice as for the case of M when diagonalizing H_{hf} and H_Z together for the $\overline{M}$ atom. This choice renders the index assignment of the levels more symmetric and eases the comparison of the limiting behavior of the M and $\overline{M}$ levels with the magnetic field. We retain the definitions of the coupled (Eq. B.19) and uncoupled (Eq. B.11) state vector bases, with the addition of a label to indicate a M or an $\overline{M}$ basis state. In other words, the label M or $\overline{M}$ designates the charge state of the component leptons. For $\overline{M}$, the Breit-Rabi states are

$$\begin{aligned} |\,\lambda_1^{(\overline{M})}\rangle &= |\overline{M}\ ,\uparrow\uparrow\rangle \\ |\,\lambda_2^{(\overline{M})}\rangle &= \frac{1}{N}\left[-\left(X+\sqrt{1+X^2}\right)|\overline{M}\ ,\uparrow\downarrow\rangle + |\overline{M}\ ,\downarrow\uparrow\rangle\right] \\ |\,\lambda_3^{(\overline{M})}\rangle &= \frac{1}{N}\left[|\overline{M}\ ,\uparrow\downarrow\rangle + \left(X+\sqrt{1+X^2}\right)|\overline{M}\ ,\downarrow\uparrow\rangle\right] \end{aligned}$$

$$\begin{aligned} |\lambda_4^{(\overline{M})}\rangle &= |\overline{M}\,,\downarrow\downarrow\rangle \\ N &= \left[1+\left(X+\sqrt{1+X^2}\right)^2\right]^{\frac{1}{2}} \end{aligned} \tag{B.41}$$

and the corresponding energy eigenvalues are

$$\begin{aligned} \lambda_1^{(\overline{M})} &= E_0 + \frac{a}{2}\left(\frac{1}{2} - Y\right) \\ \lambda_2^{(\overline{M})} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} - \sqrt{1+X^2}\right) \\ \lambda_3^{(\overline{M})} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} + \sqrt{1+X^2}\right) \\ \lambda_4^{(\overline{M})} &= E_0 + \frac{a}{2}\left(\frac{1}{2} + Y\right) . \end{aligned} \tag{B.42}$$

These energy levels are shown in Fig. B.2 as a function of the external magnetic field. It is between the M levels of Fig. B.1 and the $\overline{M}$ levels shown in Fig. B.2 that the possibility of conversion is to be considered. Since the conversion Hamiltonian is not diagonal in any basis considered so far – it acts in a space spanned by M *and* $\overline{M}$ ground state hyperfine levels – the rigorous procedure for calculating the conversion probabilities between given hyperfine levels of M and $\overline{M}$ is to diagonalize the full Hamiltonian describing the hydrogenic structure, the hyperfine interaction, the Zeeman effect, and the conversion coupling. Then one expresses the desired initial and final states in terms of the eigenstates of the full M, $\overline{M}$ system and calculates the overlap of these. The next section discusses the diagonalization of the full 8 dimensional Hamiltonian in the space of coupled M and $\overline{M}$ in the $n = 1$ state.

B.3 The Coupled M, $\overline{M}$ System in an External Magnetic Field

The full Hamiltonian that describes M and $\overline{M}$ in the $n = 1$ state with a possible conversion coupling is

$$H = H_0 + H_{hf} + H_Z + H_{M\overline{M}} , \tag{B.43}$$

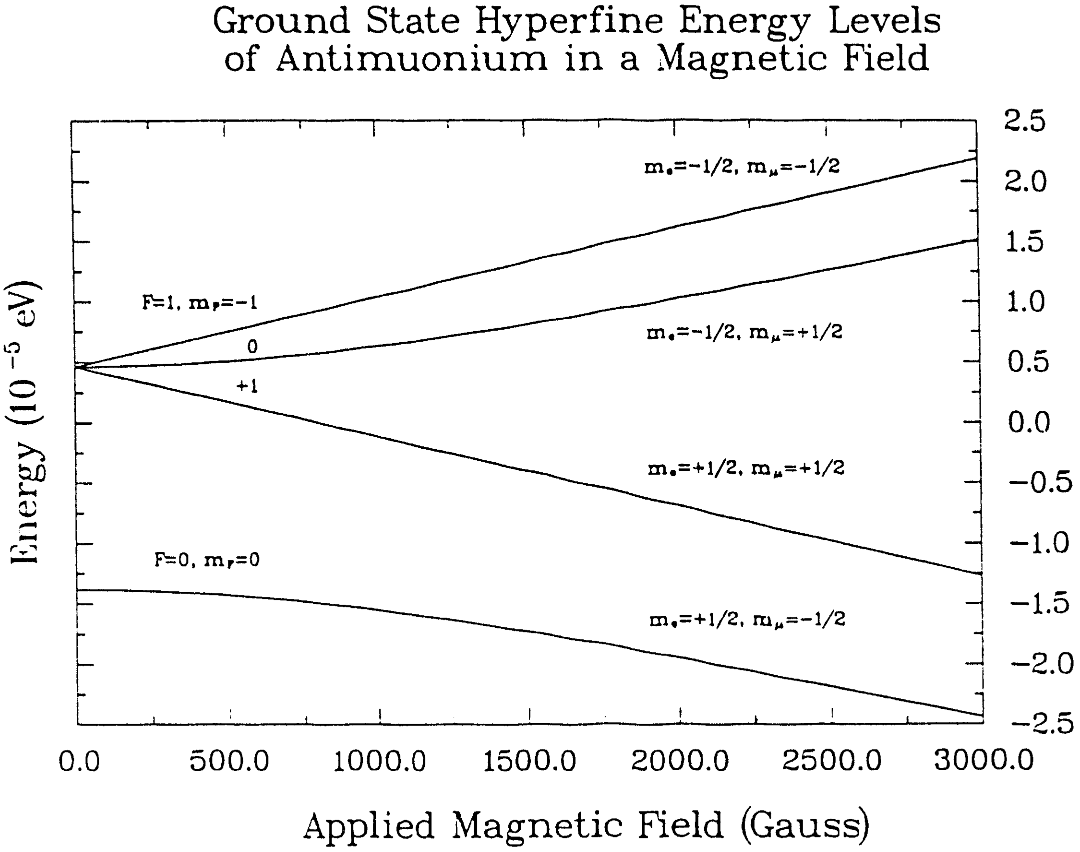


Figure B.2: Energy levels of ground state antimuonium in an external magnetic field.

where

$$\begin{aligned} H_0 &= \frac{\vec{p}^{\,2}}{2m} - \frac{e^2}{|\vec{r}\,|} \\ \vec{p} &= \vec{p}_e - \vec{p}_\mu \\ \vec{r} &= \vec{r}_e - \vec{r}_\mu \\ m &= \left[\frac{1}{m_e} + \frac{1}{m_\mu}\right]^{-1} , \end{aligned} \tag{B.44}$$

as before, for the spatial part of the Hamiltonian which acts the same for M and for $\overline{M}$,

$$\begin{aligned} H_{hf} &= a\,\vec{S}_\mu \cdot \vec{S}_e \\ &= \frac{a}{2}\left[\vec{F}^{\,2} - \vec{S}_\mu^{\,2} - \vec{S}_e^{\,2}\right] , \end{aligned} \tag{B.45}$$

also as before, for the hyperfine interaction which acts equally on M and $\overline{M}$,

$$H_Z = \pm\left(g_e\mu_B\vec{S}_e \cdot \vec{B} \;-\; g_\mu\mu_B\frac{m_e}{m_\mu}\vec{S}_\mu \cdot \vec{B}\right) , \tag{B.46}$$

where the sign choice is "+" for M and "−" for $\overline{M}$, and finally

$$\begin{aligned} H_{M\overline{M}} &= \sum_{m_{s_e},m_{s_\mu}} \sum_{m'_{s_e},m'_{s_\mu}} \left|\overline{M}\,;\, m'_{s_e}, m'_{s_\mu}\right\rangle \left(\frac{\delta}{2}\right) \delta_{m_{s_e},m'_{s_e}}\,\delta_{m_{s_\mu},m'_{s_\mu}}\,\left\langle M;\, m_{s_e}, m_{s_\mu}\right| \\ &+\ H.c. \\ &= \sum_{m_{s_e},m_{s_\mu}} \left|\overline{M}\,;\, m_{s_e}, m_{s_\mu}\right\rangle \left(\frac{\delta}{2}\right) \left\langle M;\, m_{s_e}, m_{s_\mu}\right| \;+\; H.c. \end{aligned} \tag{B.47}$$

for a $(V - A)$ conversion coupling, or

$$\begin{aligned} H_{M\overline{M}} &= \sum_{m_{s_e},m_{s_\mu}} \sum_{m'_{s_e},m'_{s_\mu}} \left|\overline{M}\,;\, m'_{s_e}, m'_{s_\mu}\right\rangle \left(\frac{\eta}{2}\right) \delta_{m_{s_e},m'_{s_\mu}}\,\delta_{m_{s_\mu},m'_{s_e}}\,\left\langle M;\, m_{s_e}, m_{s_\mu}\right| \\ &+\ H.c. \\ &= \sum_{m_{s_e},m_{s_\mu}} \left|\overline{M}\,;\, m_{s_e}, m_{s_\mu}\right\rangle \left(\frac{\eta}{2}\right) \left\langle M;\, m_{s_e}, m_{s_\mu}\right| \;+\; H.c. \ , \end{aligned} \tag{B.48}$$

for a neutral scalar conversion coupling. The convention of spin-projection labeling is still that the electron spin is given first and the muon spin second.

The uncoupled basis will be chosen for the diagonalization of H, since it seems easier to interpret the difference between the magnetic field dependence of the conversion probabilities for the two possible forms of the $M \longrightarrow \overline{M}$ conversion Hamiltonian in this representation. Here, the diagonalization will be carried out in detail for the $(V - A)$ form of the conversion Hamiltonian, Eq. B.47, as this form is believed at present to be more well-motivated physically than the alternative. For convenience, we recall here Eq. B.34

$$\begin{aligned}
E_1^{(M)} &= E_0 + \frac{a}{2}\left(\frac{1}{2} + Y\right) \\
&= E_1 \\
E_2^{(M)} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} + X\right) \\
&= E_2 \\
E_3^{(M)} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} - X\right) \\
&= E_3 \\
E_4^{(M)} &= E_0 + \frac{a}{2}\left(\frac{1}{2} - Y\right) \ , \\
&= E_4
\end{aligned} \qquad \text{(B.49)}$$

and its equivalent for $\overline{M}$

$$\begin{aligned}
E_1^{(\overline{M})} &= E_0 + \frac{a}{2}\left(\frac{1}{2} - Y\right) \\
&= E_4 \\
E_2^{(\overline{M})} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} - X\right) \\
&= E_3 \\
E_3^{(\overline{M})} &= E_0 + \frac{a}{2}\left(-\frac{1}{2} + X\right) \\
&= E_2 \\
E_4^{(\overline{M})} &= E_0 + \frac{a}{2}\left(\frac{1}{2} + Y\right) \\
&= E_1 \ .
\end{aligned} \qquad \text{(B.50)}$$

With these assignments and the choice of uncoupled basis, the matrix representation of the full Hamiltonian is

$$H = \begin{array}{cc} \left(\begin{array}{cccccccc} E_1 & 0 & 0 & 0 & \frac{\delta}{2} & 0 & 0 & 0 \\ 0 & E_2 & \frac{a}{2} & 0 & 0 & \frac{\delta}{2} & 0 & 0 \\ 0 & \frac{a}{2} & E_3 & 0 & 0 & 0 & \frac{\delta}{2} & 0 \\ 0 & 0 & 0 & E_4 & 0 & 0 & 0 & \frac{\delta}{2} \\ \frac{\delta}{2} & 0 & 0 & 0 & E_4 & 0 & 0 & 0 \\ 0 & \frac{\delta}{2} & 0 & 0 & 0 & E_3 & \frac{a}{2} & 0 \\ 0 & 0 & \frac{\delta}{2} & 0 & 0 & \frac{a}{2} & E_2 & 0 \\ 0 & 0 & 0 & \frac{\delta}{2} & 0 & 0 & 0 & E_1 \end{array}\right) & \begin{array}{l} |M, \uparrow\uparrow\rangle \\ |M, \uparrow\downarrow\rangle \\ |M, \downarrow\uparrow\rangle \\ |M, \downarrow\downarrow\rangle \\ |\overline{M}, \uparrow\uparrow\rangle \\ |\overline{M}, \uparrow\downarrow\rangle \\ |\overline{M}, \downarrow\uparrow\rangle \\ |\overline{M}, \downarrow\downarrow\rangle \end{array} \end{array} \qquad \text{(B.51)}$$

where the rows have been labeled with the uncoupled basis vectors and the columns are ordered correspondingly.

We seek a solution to the eigenvalue equation

$$H \, |\eta_i\rangle = \eta_i \, |\eta_i\rangle \quad , \qquad \text{(B.52)}$$

where the eigenenergies are represented by η_i, the eigenstates by $|\eta_i\rangle$, and the labeling index by $i \in \{1, \ldots, 8\}$. All quantum number dependence labeling charge and angular momentum state of the eigenvectors is contained in the label η_i. In fact, these states will turn out to be mixtures of the uncoupled basis states, so it is not possible to display any of the charge and angular momentum labels explicitly; there are components of several of these present. The necessary and sufficient condition for the existence of a solution for the eigenenergies η_i is the secular equation

$$\det(H - \eta I) = 0 \ , \qquad \text{(B.53)}$$

where I is the identity matrix and η is to have eight possible solutions. Straightforward, though tedious Gauss-Jordan manipulations on the 8×8 determinant transform Eq. B.53 into

$$\begin{aligned} 0 \ = \ & \left[\left(\frac{\delta}{2}\right)^2 - (E_1 - \eta)(E_4 - \eta)\right]^2 \\ \times \ & \left\{\left[\left(\frac{\delta}{2}\right)^2 - \left(\frac{a}{2}\right)^2 - (E_2 - \eta)(E_3 - \eta)\right]^2\right. \end{aligned}$$

$$-4\left(\frac{a}{2}\right)^2 (E_2-\eta)(E_3-\eta)\Big\} \ . \tag{B.54}$$

The solutions can be written as

$$\begin{aligned}
\eta_1 &= E_0+\frac{a}{2}\left[\frac{1}{2}+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]\\
\eta_2 &= E_0+\frac{a}{2}\left[-\frac{1}{2}+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]\\
\eta_3 &= E_0+\frac{a}{2}\left[-\frac{1}{2}-\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]\\
\eta_4 &= E_0+\frac{a}{2}\left[\frac{1}{2}-\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]\\
\eta_5 &= E_0+\frac{a}{2}\left[\frac{1}{2}-\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]\\
\eta_6 &= E_0+\frac{a}{2}\left[-\frac{1}{2}+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]\\
\eta_7 &= E_0+\frac{a}{2}\left[-\frac{1}{2}-\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]\\
\eta_8 &= E_0+\frac{a}{2}\left[\frac{1}{2}+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right] \ ,
\end{aligned} \tag{B.55}$$

where $\eta_5 = \eta_4$ and $\eta_8 = \eta_1$ have been written as separate solutions, though they are degenerate. They result from the first factor in Eq. B.54, in brackets, which gives two double roots. Since there must be eight solutions, due to the dimensionality of the Hamiltonian, we expect eight distinct eigenstate solutions, of which two pairs of states are energy degenerate.

The eigenstate solutions are found by substituting the solutions for η_i into Eq. B.52. For the degenerate pairs η_1, η_8 and η_4, η_5, there remains one free parameter for each eigenvector that must be chosen in determining the exact form of the eigenfunctions when forming them as linear combinations of the uncoupled basis states. The requirement to fulfill in these choices is the orthonormality of the eigenfunctions. The freedom for these choices is a direct result of the energy degeneracy of these eigenstate pairs. The choices are made to give the minimal representation in the uncoupled basis. The remaining eigenvectors are well-determined. After a lengthy calculation, the eigenvector solutions are

$$\begin{aligned}
|\eta_1\rangle &= \frac{1}{N_1}\left\{|M,\uparrow\uparrow\rangle + \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|\overline{M},\uparrow\uparrow\rangle\right\} \\
|\eta_2\rangle &= \frac{1}{N_2}\left\{|M,\uparrow\downarrow\rangle + \left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|M,\downarrow\uparrow\rangle\right. \\
&\qquad \left. + \left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|\overline{M},\uparrow\downarrow\rangle + |\overline{M},\downarrow\uparrow\rangle\right\} \\
|\eta_3\rangle &= \frac{1}{N_2}\left\{-\left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|M,\uparrow\downarrow\rangle + |M,\downarrow\uparrow\rangle\right. \\
&\qquad \left. + |\overline{M},\uparrow\downarrow\rangle - \left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|\overline{M},\downarrow\uparrow\rangle\right\} \\
|\eta_4\rangle &= \frac{1}{N_1}\left\{|M,\downarrow\downarrow\rangle - \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|\overline{M},\downarrow\downarrow\rangle\right\} \\
|\eta_5\rangle &= \frac{1}{N_1}\left\{-\left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|M,\uparrow\uparrow\rangle + |\overline{M},\uparrow\uparrow\rangle\right\} \\
|\eta_6\rangle &= \frac{1}{N_6}\left\{-\left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]|M,\uparrow\downarrow\rangle - |M,\downarrow\uparrow\rangle\right. \\
&\qquad \left. + |\overline{M},\uparrow\downarrow\rangle + \left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]|\overline{M},\downarrow\uparrow\rangle\right\}
\end{aligned}$$

$$
\begin{aligned}
|\eta_7\rangle &= \frac{1}{N_6}\Bigg\{ -|M,\uparrow\downarrow\rangle + \left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^{2}+X^{2}}\right]|M,\downarrow\uparrow\rangle \\
&\quad - \left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^{2}+X^{2}}\right]|\overline{M},\uparrow\downarrow\rangle + |\overline{M},\downarrow\uparrow\rangle \Bigg\} \\
|\eta_8\rangle &= \frac{1}{N_1}\left\{ \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^{2}+Y^{2}}\right]|M,\downarrow\downarrow\rangle + |\overline{M},\downarrow\downarrow\rangle \right\} ,
\end{aligned}
\tag{B.56}
$$

where

$$
\begin{aligned}
N_1 &= \sqrt{1+\left(\frac{a}{\delta}\right)^{2}\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^{2}+Y^{2}}\right]^{2}} \\
N_2 &= \sqrt{2\left\{1+\left(1+\frac{\delta}{a}\right)^{-2}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^{2}+X^{2}}\right]^{2}\right\}} \\
N_6 &= \sqrt{2\left\{1+\left(1-\frac{\delta}{a}\right)^{-2}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^{2}+X^{2}}\right]^{2}\right\}} .
\end{aligned}
\tag{B.57}
$$

An energy level diagram for the eigenstates of the coupled M, $\overline{M}$ system is presented in Fig. B.3.

Several checks have been carried out to verify the expected behavior of the eigenstates and -energies for the coupled M, $\overline{M}$ system. The orthonormality of the eigenvectors as given has been verified. Limits for $B \to 0$ and $\delta \to 0$ have been taken and found to be correct.

In previous work [Mor66,Ni88b,Sch88], perturbative approximations to the energy levels and states of the coupled M, $\overline{M}$ system were derived. It has been shown here that the solution may be obtained without resorting to a perturbation expansion of the eigenstates and -energies. The cited treatments give their results for the eigenstates in terms of the coupled basis. To facilitate the comparison of results, the eigenstates given above in the terms of the uncoupled basis may be cast into the coupled representation:

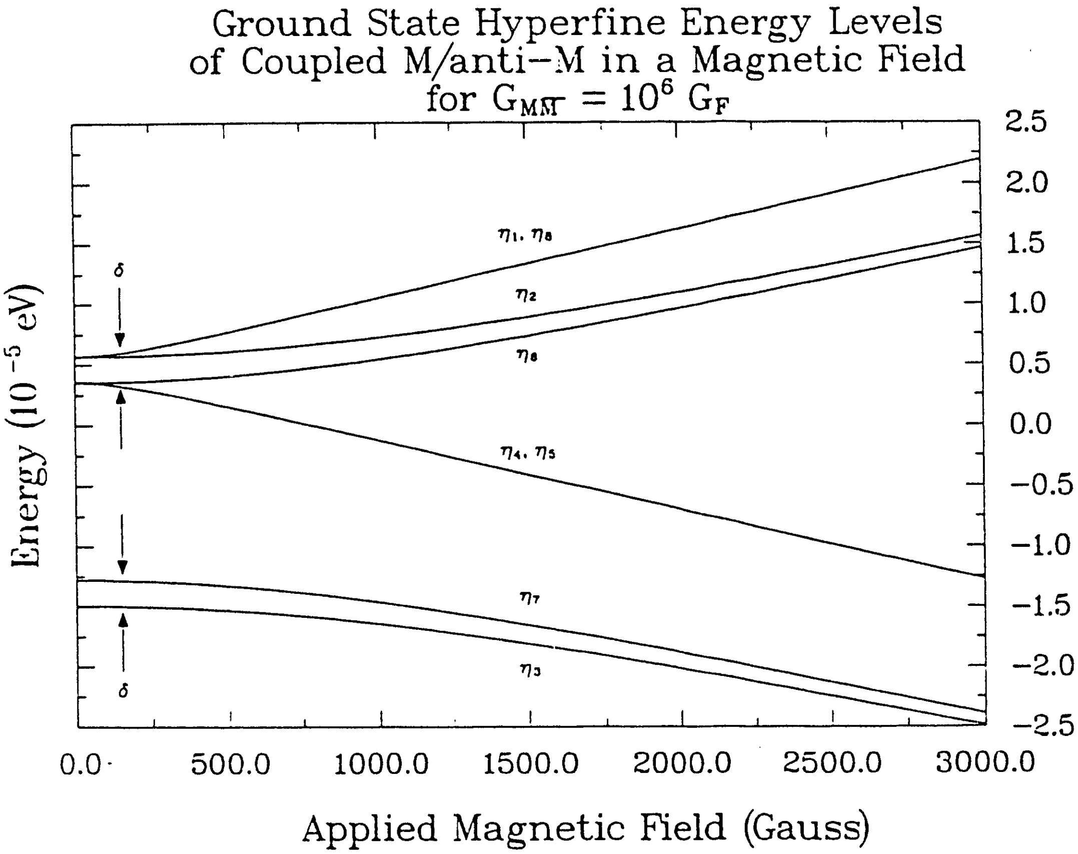


Figure B.3: Energy levels of coupled M, $\overline{M}$ system.

$$\begin{aligned}
|\eta_1\rangle &= \frac{1}{N_1}\left\{|M;\,1,\,+1\rangle + \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]\left|\overline{M};\,1,\,+1\right\rangle\right\} \\
|\eta_2\rangle &= \frac{1}{\sqrt{2}N_2}\left\{\left[1+\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,1,\,0\rangle\right. \\
&\quad+\left[1-\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,0,\,0\rangle \\
&\quad+\left[1+\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]\left|\overline{M};\,1,\,0\right\rangle \\
&\quad\left.+\left[-1+\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]\left|\overline{M};\,0,\,0\right\rangle\right\} \\
|\eta_3\rangle &= \frac{1}{\sqrt{2}N_2}\left\{\left[1-\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,1,\,0\rangle\right. \\
&\quad+\left[-1-\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,0,\,0\rangle \\
&\quad+\left[1-\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]\left|\overline{M};\,1,\,0\right\rangle \\
&\quad\left.+\left[1+\left(1+\frac{\delta}{a}\right)^{-1}\left(-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right)\right]\left|\overline{M};\,0,\,0\right\rangle\right\} \\
|\eta_4\rangle &= \frac{1}{N_1}\left\{|M;\,1,\,-1\rangle - \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]\left|\overline{M};\,1,\,-1\right\rangle\right\} \qquad \text{(B.58)}
\end{aligned}$$

and

$$|\eta_5\rangle = \frac{1}{N_1}\left\{-\left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|M;\,1,\,+1\rangle + \left|\overline{M};\,1,\,+1\right\rangle\right\}$$

$$\begin{aligned}
|\eta_6\rangle \;=\;& \frac{1}{\sqrt{2}N_6}\Bigg\{\left[-1-\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,1,\,0\rangle \\
&+\left[1-\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,0,\,0\rangle \\
&+\left[1+\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|\overline{M};\,1,\,0\rangle \\
&+\left[1-\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|\overline{M};\,0,\,0\rangle\Bigg\} \\
|\eta_7\rangle \;=\;& \frac{1}{\sqrt{2}N_6}\Bigg\{\left[-1+\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,1,\,0\rangle \\
&+\left[-1-\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|M;\,0,\,0\rangle \\
&+\left[1-\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|\overline{M};\,1,\,0\rangle \\
&+\left[-1-\left(1-\frac{\delta}{a}\right)^{-1}\left(X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right)\right]|\overline{M};\,0,\,0\rangle\Bigg\} \\
|\eta_8\rangle \;=\;& \frac{1}{N_1}\left\{\left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|M;\,1,\,-1\rangle+|\overline{M};\,1,\,-1\rangle\right\}\;. \qquad \text{(B.59)}
\end{aligned}$$

When these results and the eigenenergies are expanded to first order in the quantity (δ/a), they agree with results given in previous work [Ni88b]. For computational convenience, the results as given in terms of the uncoupled basis states will be used in the following.

To prepare for taking matrix elements of the $M \longrightarrow \overline{M}$ conversion operator when calculating the probabilities relevant to this experiment, the expressions for the $(M,\ \overline{M}\ )$ eigenstates have been inverted for the uncoupled basis states. The results are:

$$
\begin{aligned}
|M,\uparrow\uparrow\rangle &= \frac{1}{N_1}\left\{|\eta_1\rangle - \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|\eta_5\rangle\right\} \\
|M,\uparrow\downarrow\rangle &= \frac{1}{N_2}\left\{|\eta_2\rangle - \left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|\eta_3\rangle\right\} \\
&\quad +\frac{1}{N_6}\left\{-\left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]|\eta_6\rangle - |\eta_7\rangle\right\} \\
|M,\downarrow\uparrow\rangle &= \frac{1}{N_2}\left\{\left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|\eta_2\rangle + |\eta_3\rangle\right\} \\
&\quad +\frac{1}{N_6}\left\{-|\eta_6\rangle + \left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]|\eta_7\rangle\right\} \\
|M,\downarrow\downarrow\rangle &= \frac{1}{N_1}\left\{|\eta_1\rangle + \left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|\eta_8\rangle\right\}
\end{aligned}
\tag{B.60}
$$

and

$$
\begin{aligned}
|\overline{M},\uparrow\uparrow\rangle &= \frac{1}{N_1}\left\{\left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\right]|\eta_1\rangle + |\eta_5\rangle\right\} \\
|\overline{M},\uparrow\downarrow\rangle &= \frac{1}{N_2}\left\{\left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|\eta_2\rangle + |\eta_3\rangle\right\} \\
&\quad +\frac{1}{N_6}\left\{|\eta_6\rangle - \left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]|\eta_7\rangle\right\} \\
|\overline{M},\uparrow\downarrow\rangle &= \frac{1}{N_2}\left\{|\eta_2\rangle - \left(1+\frac{\delta}{a}\right)^{-1}\left[-X+\sqrt{\left(1+\frac{\delta}{a}\right)^2+X^2}\right]|\eta_3\rangle\right\} \\
&\quad +\frac{1}{N_6}\left\{\left(1-\frac{\delta}{a}\right)^{-1}\left[X+\sqrt{\left(1-\frac{\delta}{a}\right)^2+X^2}\right]|\eta_6\rangle + |\eta_7\rangle\right\}
\end{aligned}
$$

$$|\overline{M}, \downarrow\downarrow\rangle \;=\; \frac{1}{N_1}\left\{-\left(\frac{a}{\delta}\right)\left[-Y+\sqrt{\left(\frac{\delta}{a}\right)^2+Y^2}\,\right]|\eta_1\rangle + |\eta_8\rangle\right\} \;. \qquad \text{(B.61)}$$

In the next section, the time-dependent probabilities that the coupled M, $\overline{M}$ system is in the $\overline{M}$ state will be calculated assuming an initial state of pure M. These probabilities will be specific to particular angular momentum quantum numbers of the initial and final states. Since M in this experiment is formed in a mixed angular momentum state, this calculation will be followed by the density matrix treatment appropriate to the experimental conditions.

B.4 The $M \longrightarrow \overline{M}$ Conversion Probability

The action of the $M \longrightarrow \overline{M}$ conversion Hamiltonian on the spins of the particles involved figures centrally in the calculation of its matrix elements. In accord with the result given in previous work [Mor66], the form

$$\langle\overline{M};\, m'_{s_e},\, m'_{s_\mu}\big|\, H_{M\overline{M}}\, \big|M;\, m_{s_e},\, m_{s_\mu}\rangle \;=\; \left(\frac{\delta}{2}\right)\,\delta_{m_{s_e},m'_{s_e}}\,\delta_{m_{s_\mu},m'_{s_\mu}} \qquad \text{(B.62)}$$

will be assumed for the calculations in this section.

Since we have solved for the eigenstates and -energies of the coupled M, $\overline{M}$ system, the time-dependence of any superposition of these states may easily be written down. If we construct an initial state, for example of pure M in a well-defined angular momentum state, as

$$|\psi(t=0)\rangle \;=\sum_k c_k\,|\eta_k\rangle \;\;, \qquad \text{(B.63)}$$

then the time development of this state in the Schrödinger picture proceeds according to

$$|\psi(t)\rangle \;=\sum_k c_k\, e^{-i\eta_k t/\hbar}\,|\eta_k\rangle \;\;. \qquad \text{(B.64)}$$

Thus,

$$|\psi(t)\rangle \;= U(t,0)\,|\psi(t=0)\rangle \;\;, \qquad \text{(B.65)}$$

the charge state of the component leptons. For $\overline{M}$, the Breit-Rabi states are

$$\begin{aligned}
|\lambda_1^{(\overline{M})}\rangle &= |\overline{M}\,,\uparrow\uparrow\rangle \\
|\lambda_2^{(\overline{M})}\rangle &= \frac{1}{N}\left[-\left(X+\sqrt{1+X^2}\right)|\overline{M}\,,\uparrow\downarrow\rangle + |\overline{M}\,,\downarrow\uparrow\rangle\right] \\
|\lambda_3^{(\overline{M})}\rangle &= \frac{1}{N}\left[|\overline{M}\,,\uparrow\downarrow\rangle + \left(X+\sqrt{1+X^2}\right)|\overline{M}\,,\downarrow\uparrow\rangle\right]
\end{aligned}$$

where

$$U(t',t) = e^{-iH(t'-t)/\hbar} \tag{B.66}$$

and H is the full Hamiltonian. For the eigenstates of H we have, of course,

$$U(t,0)\,|\eta_k\rangle = e^{-i\eta_k t/\hbar}\,|\eta_k\rangle \;\; . \tag{B.67}$$

Now, we define the probability of the system being in the state $|\overline{M};\, m'_{s_e},\, m'_{s_\mu}\rangle$ at time t after starting in the state $|M;\, m_{s_e},\, m_{s_\mu}\rangle$ at time $t=0$ by

$$P_{\overline{M}}(m'_{s_e}, m'_{s_\mu}; m_{s_e}, m_{s_\mu}; t) = \left|\langle\overline{M};\, m'_{s_e},\, m'_{s_\mu}|U(t,0)|M;\, m_{s_e},\, m_{s_\mu}\rangle\right|^2 \;\; . \tag{B.68}$$

As we prepare the system initially in a state of pure M, we list now the probabilities of finding it in the $\overline{M}$ state at some later time. Displaying explicitly the magnetic field dependence in the coefficients of the time-dependence, they are

$$\begin{aligned}
P_{\overline{M}}(\uparrow\uparrow,\, \uparrow\uparrow;\, t) &= \frac{\left(\frac{\delta}{a}\right)^2}{\left(\frac{\delta}{a}\right)^2 + Y^2}\sin^2\omega_0 t \\
P_{\overline{M}}(\downarrow\downarrow,\, \downarrow\downarrow;\, t) &= P_{\overline{M}}(\uparrow\uparrow,\, \uparrow\uparrow;\, t) \\
P_{\overline{M}}(\uparrow\downarrow,\, \uparrow\downarrow;\, t) &= \frac{1}{4}\left\{\frac{\left(1+\frac{\delta}{a}\right)}{\sqrt{\left(1+\frac{\delta}{a}\right)^2 + X^2}}\sin\omega_+ t \right. \\
&\qquad \left. - \frac{\left(1-\frac{\delta}{a}\right)}{\sqrt{\left(1-\frac{\delta}{a}\right)^2 + X^2}}\sin\omega_- t\right\}^2 \\
P_{\overline{M}}(\downarrow\uparrow,\, \downarrow\uparrow;\, t) &= P_{\overline{M}}(\uparrow\downarrow,\, \uparrow\downarrow;\, t) \\
P_{\overline{M}}(\downarrow\uparrow,\, \uparrow\downarrow;\, t) &= \frac{1}{4}\left\{\left(\cos\omega_+ t \;-\; \cos\omega_- t\right)^2\right. \\
&\qquad \left. + \left[\frac{X}{\sqrt{\left(1+\frac{\delta}{a}\right)^2 + X^2}}\sin\omega_+ t \;-\; \frac{X}{\sqrt{\left(1-\frac{\delta}{a}\right)^2 + X^2}}\sin\omega_- t\right]^2\right\}
\end{aligned}$$

$$P_{\overline{M}}(\uparrow\downarrow,\downarrow\uparrow;t) = P_{\overline{M}}(\downarrow\uparrow,\uparrow\downarrow;t)\,. \tag{B.69}$$

In these expressions

$$\omega_{+} = \frac{a}{2\hbar}\sqrt{\left(1+\frac{\delta}{a}\right)^{2} + X^{2}}\,, \tag{B.70}$$

$$\omega_{-} = \frac{a}{2\hbar}\sqrt{\left(1-\frac{\delta}{a}\right)^{2} + X^{2}}\,, \tag{B.71}$$

and

$$\omega_{0} = \frac{a}{2\hbar}\sqrt{\left(\frac{\delta}{a}\right)^{2} + Y^{2}}\,. \tag{B.72}$$

All other possibilities give zero. The most important observation at this point is that all of the non-vanishing probabilities depend upon the magnitude of the external magnetic field through the quantities X and Y. More precisely, an increasing magnetic field will increasingly suppress the $M \longrightarrow \overline{M}$ conversion in all possible channels. This suppression is due to the breaking of the energy degeneracy of the M and $\overline{M}$ levels involved in the conversion, as is discussed in more detail in Sec 2.4. The suppression is strongest in the extreme levels of the triplet, $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$, for which a field of only 0.2 mG will halve the coefficient of the time-dependence. In addition, however, the time-development of any $\overline{M}$ component is somewhat speeded, due to the field-dependence in ω_0. For the unpolarized states, the dependence of the denominator in the coefficients on the field is not strong until $X \sim 1$, which corresponds to $B \sim 1.6$ kG.

The M initial state in this experiment is a mixed state of 50% each of $|M;\uparrow\downarrow\rangle$ and $|M;\downarrow\downarrow\rangle$, if the quantization axis is assumed to lie along the incident μ^{+} beam direction. Therefore, rather than calculating the integrated conversion probabilities for each of the separate angular momentum channels, we turn next to the density matrix description of the system.

Working in the uncoupled basis, the initial state of 1S M produced in this experiment may be represented by

$$\rho(t=0) = |M;\uparrow\downarrow\rangle\,\frac{1}{2}\,\langle M;\uparrow\downarrow| + |M;\downarrow\downarrow\rangle\,\frac{1}{2}\,\langle M;\downarrow\downarrow|\,, \tag{B.73}$$

where the quantization axis has been chosen to lie along the incident beam direction. This form is a direct result of the pickup of an unpolarized e^- from the SiO_2 powder by the incident μ^+ , which is fully polarized with helicity -1. The density matrix describing this mixed state at a later time is then given by

$$\rho(t) = U(t,0)\,\rho(0)\,U^\dagger(t,0) \ , \tag{B.74}$$

where $U(t,0)$ is the time-evolution operator of the full Hamiltonian, defined above in Eq. B.66. Next, an operator is needed that projects out any $\overline{M}$ components in the system with equal probabilities, that is, an operator that represents $\overline{M}$ detection that is not sensitive to the particular angular momentum state of the $\overline{M}$ atom. This operator, which will be called $\mathcal{P}_{\overline{M}}$, is then

$$\begin{aligned} \mathcal{P}_{\overline{M}} &= \left|\overline{M}\ ;\uparrow\uparrow\right\rangle \left\langle\overline{M}\ ;\uparrow\uparrow\right| + \left|\overline{M}\ ;\uparrow\downarrow\right\rangle \left\langle\overline{M}\ ;\uparrow\downarrow\right| \\ &+ \left|\overline{M}\ ;\downarrow\uparrow\right\rangle \left\langle\overline{M}\ ;\downarrow\uparrow\right| + \left|\overline{M}\ ;\downarrow\downarrow\right\rangle \left\langle\overline{M}\ ;\downarrow\downarrow\right| \ . \end{aligned} \tag{B.75}$$

Thus, the probability of the system being in an $\overline{M}$ state is given by

$$P_{\overline{M}}(t) = \mathrm{Tr}\,\left(\rho\,(t)\,\mathcal{P}_{\overline{M}}\right) \ , \tag{B.76}$$

which becomes

$$P_{\overline{M}}(t) = \mathrm{Tr}\,\left(e^{-iHt/\hbar}\,\rho\,(0)\;e^{iHt/\hbar}\,\mathcal{P}_{\overline{M}}\right) \ . \tag{B.77}$$

To evaluate $P_{\overline{M}}(t)$, we can either transform $\rho(0)$ and $\mathcal{P}_{\overline{M}}$ as given into the eigenstate basis $\{|\eta_i\rangle\}$ or express the time evolution operators into the uncoupled basis. In any case, to perform the trace, all matrices must be in the same basis. The former approach is chosen for simplicity, so the required transformation may be written as

$$|\eta_i\rangle = \sum_{ij} \Lambda_{ij}\;|u_j\rangle \ , \tag{B.78}$$

where Λ_{ij} is the basis transformation matrix and the uncoupled basis for the full M, $\overline{M}$ system is given by

$$|u_i\rangle \in \{|M;\, m_{s_e},\, m_{s_\mu}\rangle,\ |\overline{M}\ ;\, m_{s_e},\, m_{s_\mu}\rangle\} \ . \tag{B.79}$$

The matrix elements of the transformation may be read off the set of equations above (Eq. B.61) expressing the uncoupled state vectors in terms of the M, $\overline{M}$

eigenstates. The matrix multiplication is lengthy, but straight-forward. Taking the trace yields the result

$$\begin{aligned} P_{\overline{M}}(t) &= \frac{\left(\frac{\delta}{a}\right)^2}{2\left[\left(\frac{\delta}{a}\right)^2 + Y^2\right]} \sin^2 \omega_0 t \\ &+ \frac{1}{4}\left[1 - \frac{(1+X^2) - \left(\frac{\delta}{a}\right)^2}{\sqrt{\left(1+\frac{\delta}{a}\right)^2 + X^2}\sqrt{\left(1-\frac{\delta}{a}\right)^2 + X^2}}\right] \sin^2\left(\frac{\omega_+ + \omega_-}{2}\, t\right) \\ &+ \frac{1}{4}\left[1 + \frac{(1+X^2) - \left(\frac{\delta}{a}\right)^2}{\sqrt{\left(1+\frac{\delta}{a}\right)^2 + X^2}\sqrt{\left(1-\frac{\delta}{a}\right)^2 + X^2}}\right] \sin^2\left(\frac{\omega_+ - \omega_-}{2}\, t\right) \quad \text{(B.80)} \end{aligned}$$

where and ω_0, ω_+, and ω_- are given in Eqs. B.72, B.70, and B.71, respectively, and the magnetic field dependence of the coefficients has been explicitly displayed. The above is then the instantaneous probability of finding the system in any of the $\overline{M}$ sublevels after initially preparing it in the specified mixed state of M in our experiment. The first term may be identified as coming from conversions between the polarized states of M and $\overline{M}$ and is very sensitive to an external magnetic field, whereas the second and third terms are due to conversions between the unpolarized states, which exhibits a weaker degree of field dependence. The limit as $B \rightarrow 0$ of $P_{\overline{M}}(t)$ is simple and provides another check on the result:

$$P_{\overline{M}}(t;\, B = 0) = \sin^2\left(\frac{\delta t}{2\hbar}\right) . \qquad \text{(B.81)}$$

The quantity $P_{\overline{M}}(t;\, B = 0)$ is plotted against time in Fig. B.4 together with the normal muon decay time-dependence.

We note here that this result could also have been obtained by calculating separately the conversion probabilities between polarized and unpolarized states, since conversions between the two do not mix. The contribution to conversions between unpolarized states, however, has two components that do indeed interfere: the conversions from triplet-to-triplet (singlet-to-singlet) and triplet-to-singlet (singlet-to-triplet). Thus, using the density-matrix formalism will ensure that proper account is taken of this behavior. For convenience, the right-hand side of Eq. B.80 may be

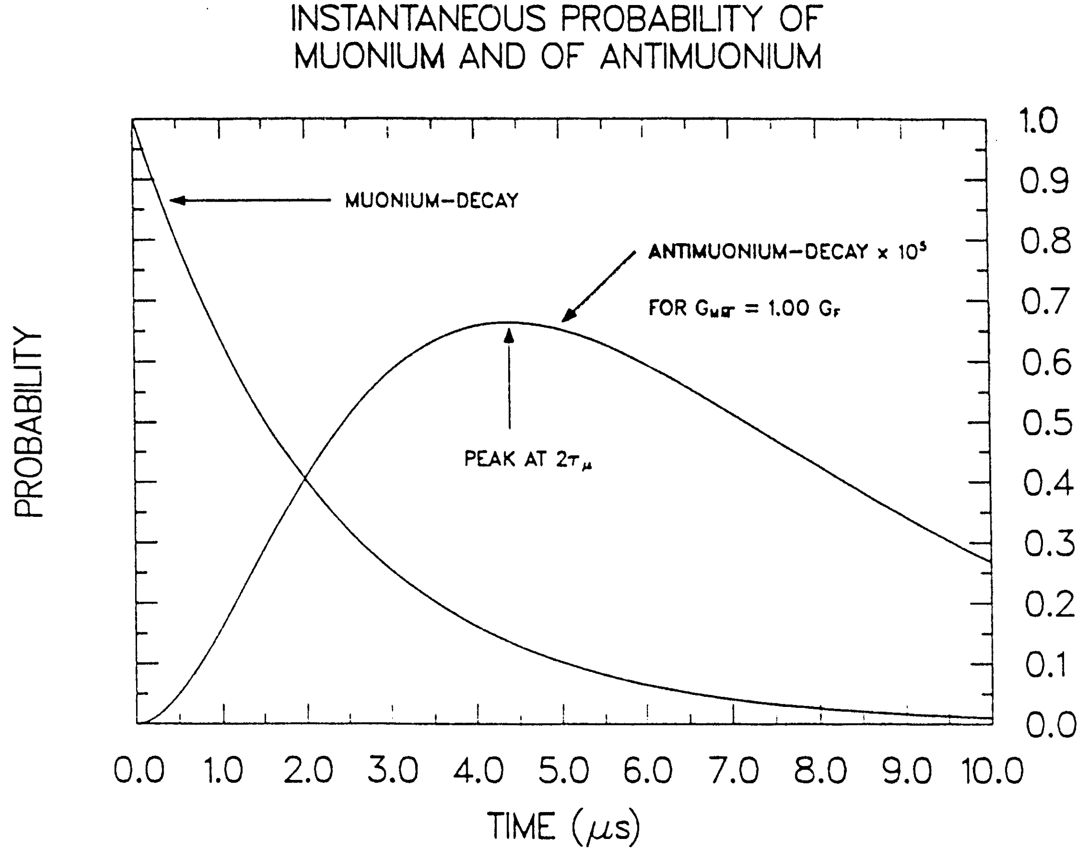


Figure B.4: Instantaneous probabilities of M and $\overline{M}$, after preparing an initial state of M, for $B = 0$.

expanded in powers of $\frac{\delta}{a}$, since $\frac{\delta}{a} \sim 10^{-7}$. The result is

$$\begin{aligned} P_{\overline{M}}(t) \approx{} & \frac{1}{2}\left(\frac{\delta}{aY}\right)^2 \sin^2\frac{aYt}{2\hbar} + \frac{1}{2}\sin^2\left(\frac{\delta t}{2\hbar}\frac{1}{\sqrt{1+X^2}}\right) \\ & + \frac{1}{4}\left(\frac{\delta}{a}\right)^2 \frac{X^2}{(1+X^2)^2}\left[\cos\left(\frac{\delta t}{\hbar}\frac{1}{\sqrt{1+X^2}}\right) - \cos\left(\frac{at}{\hbar}\sqrt{1+X^2}\right)\right] \\ & + \mathcal{O}(\left(\left(\frac{\delta}{a}\right)^3\right) . \end{aligned} \tag{B.82}$$

This section concludes by considering the integral of Eq. B.80 over the full observation time during which the apparatus may observe a decay of the M, $\overline{M}$ system. This integral represents the "signal" that is sought in this experiment and may be written as

$$S_{\overline{M}}(T) = \int_0^T \gamma e^{-\gamma t} P_{\overline{M}}(t)\, dt \ , \tag{B.83}$$

where γ is the muon decay rate

$$\gamma = \frac{1}{\tau_\mu} \tag{B.84}$$

and T is the duration of the observation. Since the mean speed of the thermal M in vacuum is about 0.7 $cm/\mu s$ and the size of the field of view of the spectrometer is on the order of 7 cm, the M, $\overline{M}$ system requires about 10 μs to leave this acceptance. As this time is long compared to the muon lifetime, the approximation of $T \to \infty$ is valid and will be used. Carrying out the integral on Eq. B.80 gives

$$\begin{aligned} S_{\overline{M}}(\infty) ={} & \frac{\left(\frac{\delta}{a}\right)^2}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + \left(\frac{\delta}{a}\right)^2 + Y^2\right]} \\ & + \frac{1}{8}\left[1 - \frac{(1+X^2) - \left(\frac{\delta}{a}\right)^2}{\sqrt{\left(1+\frac{\delta}{a}\right)^2 + X^2}\sqrt{\left(1-\frac{\delta}{a}\right)^2 + X^2}}\right]\frac{(\omega_+ + \omega_-)^2}{\gamma^2 + (\omega_+ + \omega_-)^2} \\ & + \frac{1}{8}\left[1 + \frac{(1+X^2) - \left(\frac{\delta}{a}\right)^2}{\sqrt{\left(1+\frac{\delta}{a}\right)^2 + X^2}\sqrt{\left(1-\frac{\delta}{a}\right)^2 + X^2}}\right]\frac{(\omega_+ - \omega_-)^2}{\gamma^2 + (\omega_+ - \omega_-)^2} \end{aligned} \tag{B.85}$$

Substituting numbers into the expression for $S_{\overline{M}}$ reveals that an external magnetic field inhibits 50% of the conversion for values above the order of 100 mG but below the order of 1 kG. It is on this "plateau" that this experiment was carried out, with a magnetic field in the conversion region of about 10 G. The dependence of $S_{\overline{M}}$ on the magnetic field is shown graphically in Fig. B.5.

Calculating the correct expansion in powers of $\frac{\delta}{a}$ can be a little tricky and has been treated in a cavalier fashion in previous work by keeping inconsistent orders. Expanding all quantities involved through second order in $\frac{\delta}{a}$ and collecting terms yields

$$
\begin{aligned}
S_{\overline{M}}(\infty) \quad = \quad & \frac{\left(\frac{\delta}{a}\right)^2}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + Y^2\right]} \\
& + \frac{\left(\frac{\delta}{a}\right)^2}{4\left(\frac{\hbar\gamma}{a}\right)^2 (1 + X^2)} \\
& + \frac{\left(\frac{\delta}{a}\right)^2 \frac{X^2}{1+X^2}}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + (1 + X^2)\right]} \; . \qquad \text{(B.86)}
\end{aligned}
$$

Here, the first term is due to conversions between the polarized states, the second term is due to conversions from triplet-to-triplet or singlet-to-singlet unpolarized states, and the third term describes conversions between unpolarized states going from singlet M to triplet $\overline{M}$ or from triplet M to singlet $\overline{M}$. As a check, the limit of $S_{\overline{M}}(\infty) \longrightarrow 0$ as $\delta \longrightarrow 0$ is correctly obtained. We notice that the last term vanishes as $X \longrightarrow 0$, as expected, because a non-zero magnetic field is required to mix the triplet and singlet unpolarized states in M (and in $\overline{M}$). In the absence of such mixing, the conversion cannot couple the unpolarized part of the M triplet with the $\overline{M}$ singlet or the M singlet with the unpolarized part of the $\overline{M}$ triplet state. The last two terms may be combined to get the result

$$
S_{\overline{M}}(\infty) = \frac{\left(\frac{\delta}{a}\right)^2}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + Y^2\right]} + \frac{\left(\frac{\delta}{\hbar\gamma}\right)^2 + \left(\frac{\delta}{a}\right)^2}{4\left[\left(\frac{\hbar\gamma}{a}\right)^2 + (1 + X^2)\right]} \; . \qquad \text{(B.87)}
$$

In this equation, the contributions from conversions between polarized states are collected in the first term and from those between unpolarized states in the second term.

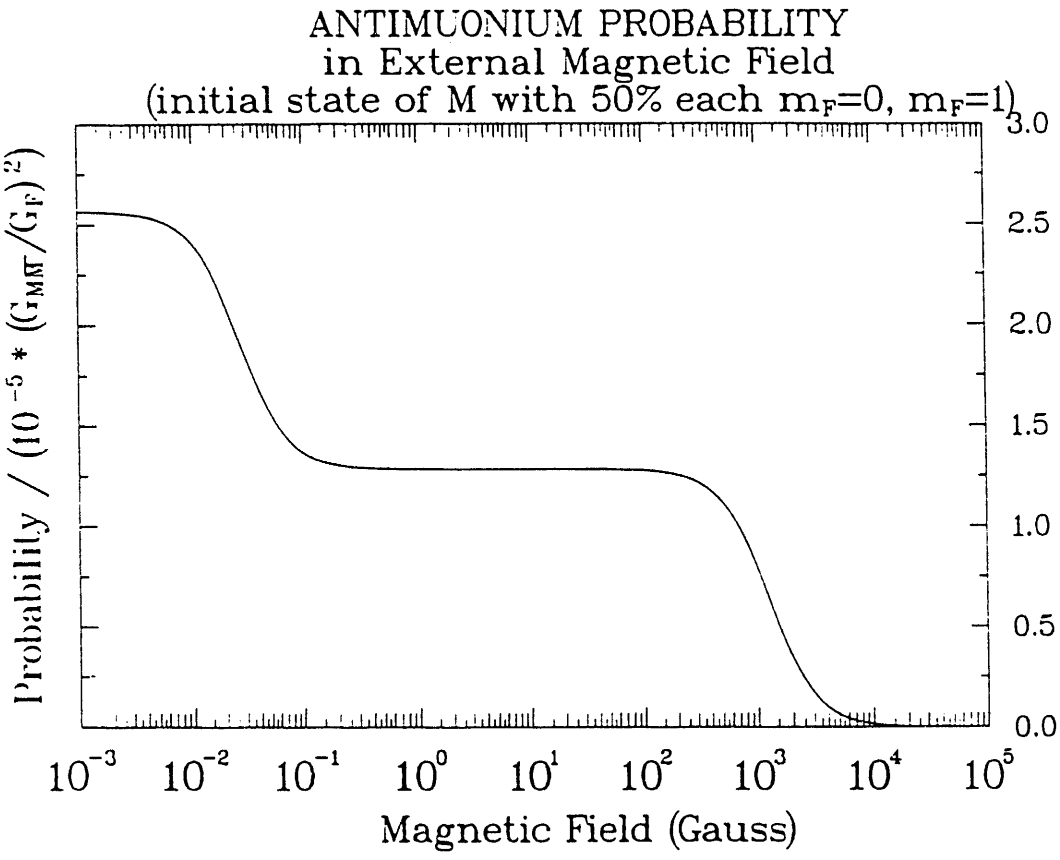


Figure B.5: Experimentally expected $M \longrightarrow \overline{M}$ conversion probability as a function of external magnetic field.

It is possible to take the $X, Y \rightarrow 0$ limit to get the form of $S_{\overline{M}}(\infty)$ valid at zero field without approximations. This limit is

$$S_{\overline{M}}(\infty) \xrightarrow{B \rightarrow 0} \frac{\frac{1}{2}\delta^2}{(\hbar\gamma)^2 + \delta^2}, \tag{B.88}$$

in agreement with results of earlier work. The value of the matrix element for the $M \longrightarrow \overline{M}$ conversion is

$$\begin{aligned} \frac{\delta}{2} &= \frac{8G_F}{\sqrt{2}\pi a_0^3} \left(\frac{G_{M\overline{M}}}{G_F}\right) \\ &= \left(1.07 \times 10^{-12}\right) \left(\frac{G_{M\overline{M}}}{G_F}\right) \end{aligned} \tag{B.89}$$

with $G_F = 8.96 \times 10^{-38}\, eV^{-2}$ and $a_0 = 0.532 \times 10^{-8}\, cm$. Also, $\hbar\gamma = 3.00 \times 10^{-10} eV$ is the muon decay constant in energy units, and we know [Hub90] that $G_{\overline{M}} < 0.29\, G_F$ (90% C.L.), so, assuming that $\delta \ll \hbar\gamma$, we may write

$$S_{\overline{M}}(\infty;\, B = 0) = \left(2.57 \times 10^{-5}\right) \left(\frac{G_{\overline{M}}}{G_F}\right)^2 . \tag{B.90}$$

This is the final result in the $B \rightarrow 0$ limit for the assumption of a $(V - A)$ conversion coupling.

Appendix C

Second-Order Taylor Expansion of the Magnetic Field

Since the track model that is fit to the measured particle hits in the spectrometer requires knowledge of all components, B_x, B_y, and Bz, of the field in the spectrometer magnet at arbitrary space points, it was necessary to develop a suitable method by which to determine these values from the measurements at the field map grid points.

In a vacuum region, the static magnetic field may be expressed as the gradient of a potential. Maxwell's equations require this potential to satisfy Laplace's equation:

$$B \equiv \vec{\nabla}\phi_B \tag{C.1}$$

$$\vec{\nabla} \cdot B = \vec{\nabla}^2 \phi_B = 0 \ . \tag{C.2}$$

One notes that Ampère's Law becomes

$$\vec{\nabla} \times \vec{\nabla}\phi_B = 0 \ , \tag{C.3}$$

and has been reduced to an identity by assuming Eq. C.1. Thus, one might, in principle, fit the magnetic field globally to the field map data, using a series of the eigenfunctions of the Laplace equation for the given boundary conditions. These boundary conditions are exceedingly complicated for the wide-gap field-clamped C-magnet used, so this approach was abandoned.

Instead, it is sufficient to have a locally valid interpolation procedure that uses the measured field map values at grid points surrounding the point at which the field is needed. The interpolation is accomplished by a Taylor expansion of each field component around every field map grid point up to second order in the spatial coordinates. The order was chosen to allow use of Maxwell's equations to constrain the expansion parameters and to handle regions of rapidly varying fields more accurately than a linear interpolation would.

The expansions of the field components around the point $\vec{x}_0 = (x_0, y_0, z_0)$ are

$$\begin{aligned}
B_x \;=\;& B_{x_0} + (x - x_0) \left.\frac{\partial B_x}{\partial x}\right|_{\vec{x}_0} + (y - y_0) \left.\frac{\partial B_x}{\partial y}\right|_{\vec{x}_0} + (z - z_0) \left.\frac{\partial B_x}{\partial z}\right|_{\vec{x}_0} \\
& + \tfrac{1}{2}\,(x - x_0)^2 \left.\frac{\partial^2 B_x}{\partial x^2}\right|_{\vec{x}_0} + \tfrac{1}{2}\,(y - y_0)^2 \left.\frac{\partial^2 B_x}{\partial y^2}\right|_{\vec{x}_0} + \tfrac{1}{2}\,(z - z_0)^2 \left.\frac{\partial^2 B_x}{\partial z^2}\right|_{\vec{x}_0} \\
& + (x - x_0)(y - y_0) \left.\frac{\partial^2 B_x}{\partial x \partial y}\right|_{\vec{x}_0} + (x - x_0)(z - z_0) \left.\frac{\partial^2 B_x}{\partial x \partial z}\right|_{\vec{x}_0} \\
& + (y - y_0)(z - z_0) \left.\frac{\partial^2 B_x}{\partial y \partial z}\right|_{\vec{x}_0} \\
& + \mathcal{O}(\Delta^3)\ , \qquad\qquad (C.4)
\end{aligned}$$

$$\begin{aligned}
B_y \;=\;& B_{y_0} + (x - x_0) \left.\frac{\partial B_y}{\partial x}\right|_{\vec{x}_0} + (y - y_0) \left.\frac{\partial B_y}{\partial y}\right|_{\vec{x}_0} + (z - z_0) \left.\frac{\partial B_y}{\partial z}\right|_{\vec{x}_0} \\
& + \tfrac{1}{2}\,(x - x_0)^2 \left.\frac{\partial^2 B_y}{\partial x^2}\right|_{\vec{x}_0} + \tfrac{1}{2}\,(y - y_0)^2 \left.\frac{\partial^2 B_y}{\partial y^2}\right|_{\vec{x}_0} + \tfrac{1}{2}\,(z - z_0)^2 \left.\frac{\partial^2 B_y}{\partial z^2}\right|_{\vec{x}_0} \\
& + (x - x_0)(y - y_0) \left.\frac{\partial^2 B_y}{\partial x \partial y}\right|_{\vec{x}_0} + (x - x_0)(z - z_0) \left.\frac{\partial^2 B_y}{\partial x \partial z}\right|_{\vec{x}_0} \\
& + (y - y_0)(z - z_0) \left.\frac{\partial^2 B_y}{\partial y \partial z}\right|_{\vec{x}_0} \\
& + \mathcal{O}(\Delta^3)\ , \qquad\qquad (C.5)
\end{aligned}$$

$$B_z \;=\; B_{z_0} + (x - x_0) \left.\frac{\partial B_z}{\partial x}\right|_{\vec{x}_0} + (y - y_0) \left.\frac{\partial B_z}{\partial y}\right|_{\vec{x}_0} + (z - z_0) \left.\frac{\partial B_z}{\partial z}\right|_{\vec{x}_0}$$

$$\begin{aligned}
&+\tfrac{1}{2}\,(x-x_0)^2\left.\frac{\partial^2 B_z}{\partial x^2}\right|_{\vec{x}_0} + \tfrac{1}{2}\,(y-y_0)^2\left.\frac{\partial^2 B_z}{\partial y^2}\right|_{\vec{x}_0} + \tfrac{1}{2}\,(z-z_0)^2\left.\frac{\partial^2 B_z}{\partial z^2}\right|_{\vec{x}_0}\\
&+(x-x_0)(y-y_0)\left.\frac{\partial^2 B_z}{\partial x\partial y}\right|_{\vec{x}_0} + (x-x_0)(z-z_0)\left.\frac{\partial^2 B_z}{\partial x\partial z}\right|_{\vec{x}_0}\\
&+(y-y_0)(z-z_0)\left.\frac{\partial^2 B_z}{\partial y\partial z}\right|_{\vec{x}_0}\\
&+\mathcal{O}(\Delta^3)\ ,
\end{aligned} \tag{C.6}$$

where $\vec{B}(\vec{x}_0) = (B_{x_0}, B_{y_0}, B_{z_0})$, $\Delta \sim (x-x_0),\ (y-y_0),\ (z-z_0)$, and all derivatives are evaluated at the central point of the expansion, $\vec{x}_0$. To render these formulae useful, one must determine all of the first and second partial derivatives of each field component at every one of the field map points. These may then be thought of as the coefficients in the spatial expansion of the magnetic field.

For easier reference to the derivatives in the expansion, one introduces the coefficients

$$D_{ab} = \left.\frac{\partial B_a}{\partial b}\right|_{\vec{x}_0}\ , \tag{C.7}$$

where $a, b \in \{x, y, z\}$, and

$$S_{abc} = \left.\frac{\partial^2 B_a}{\partial b\partial c}\right|_{\vec{x}_0}\ , \tag{C.8}$$

where $a, b, c \in \{x, y, z\}$. It is reasonable to assume that the field components are at least once continuously differentiable, so that $S_{abc} = S_{acb}$ holds for the mixed partial derivatives. This assumption has already been built in to Eqs. C.4, C.5, and C.6.

We thus have 10 expansion coefficients for each of the field components. To determine these coefficients, we use the field map data at the grid points given in Table C.1. In the expansion, (x_0, y_0, z_0) is chosen to be the point on the field map grid that lies closest to the point at which the field is needed. The remaining points are nearest neighbors and next-nearest neighbors to this central point of the expansion. At each of these points, the field map gives the value of 3 field components. This gives 57 field values with which to determine 30 expansion parameters. It turns out that, excluding the central point of the expansion, the remaining 54 values doubly determine each of the remaining 27 expansion parameters. Of two possible expressions for each of the expansion coefficients, at most one can be the

Field Map Grid Point	Field Components at this Point
(x_0, y_0, z_0)	$B_{x_0}, B_{y_0}, B_{z_0}$
(x_{-1}, y_0, z_0)	F_{1x}, F_{1y}, F_{1z}
(x_{+1}, y_0, z_0)	F_{2x}, F_{2y}, F_{2z}
(x_0, y_{-1}, z_0)	F_{3x}, F_{3y}, F_{3z}
(x_0, y_{+1}, z_0)	F_{4x}, F_{4y}, F_{4z}
(x_0, y_0, z_{-1})	F_{5x}, F_{5y}, F_{5z}
(x_0, y_0, z_{+1})	F_{6x}, F_{6y}, F_{6z}
(x_{-1}, y_{-1}, z_0)	F_{7x}, F_{7y}, F_{7z}
(x_{+1}, y_{+1}, z_0)	$F'_{7x}, F'_{7y}, F'_{7z}$
(x_{-1}, y_{+1}, z_0)	F_{8x}, F_{8y}, F_{8z}
(x_{+1}, y_{-1}, z_0)	$F'_{8x}, F'_{8y}, F'_{8z}$
(x_{-1}, y_0, z_{-1})	F_{9x}, F_{9y}, F_{9z}
(x_{+1}, y_0, z_{+1})	$F'_{9x}, F'_{9y}, F'_{9z}$
(x_{-1}, y_0, z_{+1})	$F_{10x}, F_{10y}, F_{10z}$
(x_{+1}, y_0, z_{-1})	$F'_{10x}, F'_{10y}, F'_{10z}$
(x_0, y_{-1}, z_{-1})	$F_{11x}, F_{11y}, F_{11z}$
(x_0, y_{+1}, z_{+1})	$F'_{11x}, F'_{11y}, F'_{11z}$
(x_0, y_{-1}, z_{+1})	$F_{12x}, F_{12y}, F_{12z}$
(x_0, y_{+1}, z_{-1})	$F'_{12x}, F'_{12y}, F'_{12z}$

Table C.1: Grid points and field values used in the Taylor expansion of the spectrometer magnetic field.

symmetric approximation to the first or second derivative at the central point. Since this is accurate to one higher order in the grid point spacing than any asymmetric approximation, the symmetric derivative approximations are used to find all of the coefficients. The resulting expressions for the D_{ab} are

$$\begin{aligned}
D_{xx} &= \frac{(x_{-1}-x_0)(F_{2x}-B_{x_0})}{(x_{+1}-x_0)(x_{-1}-x_{+1})} - \frac{(x_{+1}-x_0)(F_{1x}-B_{x_0})}{(x_{-1}-x_0)(x_{-1}-x_{+1})} \\
D_{yx} &= \frac{(x_{-1}-x_0)(F_{2y}-B_{y_0})}{(x_{+1}-x_0)(x_{-1}-x_{+1})} - \frac{(x_{+1}-x_0)(F_{1y}-B_{y_0})}{(x_{-1}-x_0)(x_{-1}-x_{+1})} \\
D_{zx} &= \frac{(x_{-1}-x_0)(F_{2z}-B_{z_0})}{(x_{+1}-x_0)(x_{-1}-x_{+1})} - \frac{(x_{+1}-x_0)(F_{1z}-B_{z_0})}{(x_{-1}-x_0)(x_{-1}-x_{+1})} \\
D_{xy} &= \frac{(y_{-1}-y_0)(F_{1x}-B_{x_0})}{(y_{+1}-y_0)(y_{-1}-y_{+1})} - \frac{(y_{+1}-y_0)(F_{3x}-B_{x_0})}{(y_{-1}-y_0)(y_{-1}-y_{+1})} \\
D_{yy} &= \frac{(y_{-1}-y_0)(F_{1y}-B_{y_0})}{(y_{+1}-y_0)(y_{-1}-y_{+1})} - \frac{(y_{+1}-y_0)(F_{3y}-B_{y_0})}{(y_{-1}-y_0)(y_{-1}-y_{+1})} \\
D_{zy} &= \frac{(y_{-1}-y_0)(F_{1z}-B_{z_0})}{(y_{+1}-y_0)(y_{-1}-y_{+1})} - \frac{(y_{+1}-y_0)(F_{3z}-B_{z_0})}{(y_{-1}-y_0)(y_{-1}-y_{+1})} \\
D_{xz} &= \frac{(z_{-1}-z_0)(F_{6x}-B_{x_0})}{(z_{+1}-z_0)(z_{-1}-z_{+1})} - \frac{(z_{+1}-z_0)(F_{5x}-B_{x_0})}{(z_{-1}-z_0)(z_{-1}-z_{+1})} \\
D_{yz} &= \frac{(z_{-1}-z_0)(F_{6y}-B_{y_0})}{(z_{+1}-z_0)(z_{-1}-z_{+1})} - \frac{(z_{+1}-z_0)(F_{5y}-B_{y_0})}{(z_{-1}-z_0)(z_{-1}-z_{+1})} \\
D_{zz} &= \frac{(z_{-1}-z_0)(F_{6z}-B_{z_0})}{(z_{+1}-z_0)(z_{-1}-z_{+1})} - \frac{(z_{+1}-z_0)(F_{5z}-B_{z_0})}{(z_{-1}-z_0)(z_{-1}-z_{+1})}
\end{aligned} \tag{C.9}$$

and the S_{abc} are given by

$$S_{xxx} = \frac{2}{(x_{+1}-x_{-1})}\left[\frac{(F_{2x}-B_{x_0})}{(x_{+1}-x_0)} - \frac{(F_{1x}-B_{x_0})}{(x_{-1}-x_0)}\right] \tag{C.10}$$

$$S_{yxx} = \frac{2}{(x_{+1}-x_{-1})}\left[\frac{(F_{2y}-B_{y_0})}{(x_{+1}-x_0)} - \frac{(F_{1y}-B_{y_0})}{(x_{-1}-x_0)}\right]$$

$$S_{zxx} = \frac{2}{(x_{+1} - x_{-1})}\left[\frac{(F_{2z} - B_{z_0})}{(x_{+1} - x_0)} - \frac{(F_{1z} - B_{z_0})}{(x_{-1} - x_0)}\right]$$

$$S_{xyy} = \frac{2}{(y_{+1} - y_{-1})}\left[\frac{(F_{1x} - B_{x_0})}{(y_{+1} - y_0)} - \frac{(F_{3x} - B_{x_0})}{(y_{-1} - y_0)}\right]$$

$$S_{yyy} = \frac{2}{(y_{+1} - y_{-1})}\left[\frac{(F_{1y} - B_{y_0})}{(y_{+1} - y_0)} - \frac{(F_{3y} - B_{y_0})}{(y_{-1} - y_0)}\right]$$

$$S_{zyy} = \frac{2}{(y_{+1} - y_{-1})}\left[\frac{(F_{1z} - B_{z_0})}{(y_{+1} - y_0)} - \frac{(F_{3z} - B_{z_0})}{(y_{-1} - y_0)}\right]$$

$$S_{xzz} = \frac{2}{(z_{+1} - z_{-1})}\left[\frac{(F_{6x} - B_{x_0})}{(z_{+1} - z_0)} - \frac{(F_{5x} - B_{x_0})}{(z_{-1} - z_0)}\right]$$

$$S_{yzz} = \frac{2}{(z_{+1} - z_{-1})}\left[\frac{(F_{6y} - B_{y_0})}{(z_{+1} - z_0)} - \frac{(F_{5y} - B_{y_0})}{(z_{-1} - z_0)}\right]$$

$$S_{zzz} = \frac{2}{(z_{+1} - z_{-1})}\left[\frac{(F_{6z} - B_{z_0})}{(z_{+1} - z_0)} - \frac{(F_{5z} - B_{z_0})}{(z_{-1} - z_0)}\right]$$

$$S_{xxy} = \frac{(F_{7x} + F_{8x} - F'_{7x} - F'_{8x})}{(x_{+1} - x_{-1})(y_{+1} - y_{-1})}$$

$$S_{yxy} = \frac{(F_{7y} + F_{8y} - F'_{7y} - F'_{8y})}{(x_{+1} - x_{-1})(y_{+1} - y_{-1})}$$

$$S_{zxy} = \frac{(F_{7z} + F_{8z} - F'_{7z} - F'_{8z})}{(x_{+1} - x_{-1})(y_{+1} - y_{-1})}$$

$$S_{xxz} = \frac{(F_{9x} + F_{10x} - F'_{9x} - F'_{10x})}{(x_{+1} - x_{-1})(z_{+1} - z_{-1})}$$

$$S_{yxz} = \frac{(F_{9y} + F_{10y} - F'_{9y} - F'_{10y})}{(x_{+1} - x_{-1})(z_{+1} - z_{-1})}$$

$$S_{zxz} = \frac{(F_{9z} + F_{10z} - F'_{9z} - F'_{10z})}{(x_{+1} - x_{-1})(z_{+1} - z_{-1})}$$

$$
\begin{aligned}
S_{xyz} &= \frac{(F_{11x} + F_{12x} - F'_{11x} - F'_{12x})}{(y_{+1} - y_{-1})(z_{+1} - z_{-1})} \\
S_{yyz} &= \frac{(F_{11y} + F_{12y} - F'_{11y} - F'_{12y})}{(y_{+1} - y_{-1})(z_{+1} - z_{-1})} \\
S_{zyz} &= \frac{(F_{11z} + F_{12z} - F'_{11z} - F'_{12z})}{(y_{+1} - y_{-1})(z_{+1} - z_{-1})} .
\end{aligned}
$$

These formulae are used to calculate the full set of expansion coefficients for each interior field map grid point. Those grid points on the boundary do not have all of the nearest neighbor points required for this procedure, so they represent the limits on where field values are available.

During the execution of the analysis code, all expansion coefficients are stored in memory, so the field at any point within the spatial limits of the field map is found by calculating it according to the expressions given in Equations C.4, C.5, and C.6. This minimizes the computational effort of this aspect of the replay program.

Finally, it must be noted that the constraints that may be placed on the expansion parameters by Maxwell's equations and their first spatial derivatives were not enforced. The absolute measurement error of 0.5 G together with the fact that not all field components are simultaneously large at most points in the field map causes some deviations from Maxwell's equations. These always occur at points where at least one field component was very small compared to the center gap field of 522 G and, hence, had a large relative error associated with it. Since this only happens for small values of a field component, it will not cause any difficulty or diminish the validity of the field values as obtained above. One may imagine a global "smoothing" algorithm that adjusts iteratively each field component at each grid point in a way to convergently bring the whole field map into precise agreement with Maxwell's equations, effectively using information about the whole field of the magnet to constrain the values at each point. As such a procedure seemed quite complicated to devise and since we did not expect any significant improvement in the reliability of the field information, it was not attempted.

Appendix D

Model of Spectrometer Tracks for Track Fit

The track model for this experiment is based directly on the equations of motion for a charged particle in a magnetic field, the Lorentz force law,

$$\frac{du^\alpha}{d\tau} = \frac{q}{m} F^{\alpha\beta} u_\beta \ , \tag{D.1}$$

where u^α is the 4-velocity, $d\tau = dt/\gamma$ is proper time, $q = \pm e$ is the charge of the e^+ or e^- , m is the electron mass, $F^{\alpha\beta}$ is the Faraday tensor, and SI units are used. Using

$$u^\alpha = (\gamma c, \gamma \vec{v}) = \frac{p^\alpha}{m} \tag{D.2}$$

and

$$F^{\alpha\beta} = \begin{pmatrix} 0 & \frac{-E_x}{c} & \frac{-E_y}{c} & \frac{-E_z}{c} \\ \frac{E_x}{c} & 0 & -B_z & B_y \\ \frac{E_y}{c} & B_z & 0 & -B_x \\ \frac{E_z}{c} & -B_y & B_x & 0 \end{pmatrix} ,$$

where $\vec{E}$ is the electric field and $\vec{B}$ is the magnetic field, the Lorentz force may be decomposed into equations in laboratory coordinates:

$$\frac{d\mathcal{E}}{dt} = \frac{q}{\gamma m} \vec{p} \cdot \vec{E} \tag{D.3}$$

$$\frac{d\vec{p}}{dt} = q\left(\vec{E} + \frac{1}{\gamma m}\,\vec{p} \times \vec{B}\right) ,$$

where $\mathcal{E}$ is the total energy of the $e^{\pm}$.

The specifics of the experimental conditions allow further simplifications. No electric fields are present in the spectrometer, so $\vec{E} = 0$. Since the energy loss of Michel-distributed $e^{\pm}$ has been estimated to be on the order of 1 MeV in the air and the MWPCs, a small value compared to their typical energy of several tens of MeV, energy loss in the spectrometer material will be neglected in the track model. This entails $\mathcal{E} = const.$ and $\gamma = const.$ The remaining equations are thus

$$\begin{aligned} \frac{dp_x}{dt} &= \frac{q}{\gamma m}\left(p_y B_z - p_z B_y\right) \\ \frac{dp_y}{dt} &= \frac{q}{\gamma m}\left(p_z B_x - p_x B_z\right) \\ \frac{dp_z}{dt} &= \frac{q}{\gamma m}\left(p_x B_y - p_y B_x\right) . \end{aligned} \qquad \text{(D.4)}$$

Since the magnetic field is a function of the spatial coordinates, these equations may only be solved locally, where the field can be assumed to be constant over a small region. To obtain this solution, one rewrites the equations as

$$\frac{d\vec{p}}{dt} = \mathcal{M} \cdot \vec{p} \qquad \text{(D.5)}$$

and postulates the *Ansatz*

$$\vec{p}(t) = e^{\mathcal{M}t}\vec{p}_0 , \qquad \text{(D.6)}$$

where

$$\mathcal{M} = \frac{q}{\gamma m}\begin{pmatrix} 0 & B_z & -B_y \\ -B_z & 0 & B_x \\ B_y & -B_x & 0 \end{pmatrix} \qquad \text{(D.7)}$$

and $\vec{p}_0$ is the momentum at the outset of the motion through the small region. Inserting the power series definition of the exponential of the matrix, $\mathcal{M}$, and using the recursive property found for powers of this matrix gives the result

$$\vec{p}(t) = \left[1 + \frac{\sin\omega t}{\omega}\mathcal{M} + \frac{(1 - \cos\omega t)}{\omega^2}\mathcal{M}^2\right]\vec{p}_0 , \qquad \text{(D.8)}$$

where

$$\omega = \frac{qB}{\gamma m} \tag{D.9}$$

is the cyclotron frequency of $e^{\pm}$ in a field of magnitude B.

Since the position of the particle in the spectrometer is related to its momentum by

$$\vec{p} = \gamma m \frac{d\vec{x}}{dt} , \tag{D.10}$$

the solution for the position of the particle follows immediately as

$$\vec{x} = \vec{x}_0 + \frac{\vec{p}_0 t}{\gamma m} + \frac{(1-\cos\omega t)}{\omega^2}\frac{\mathcal{M}\vec{p}_0}{\gamma m} + \frac{(\omega t - \sin\omega t)}{\omega^3}\frac{\mathcal{M}^2\vec{p}_0}{\gamma m} . \tag{D.11}$$

To use this equation as a step prescription in a simulation, a few approximations that follow from assuming small time steps are possible and indeed necessary, since the form of the equation is very susceptible to round-off errors. When these are carried out, one has an equation that is third order in the time step and so might promise to be computationally efficient. It is not, however, useful to us, since we are bound by the assumption that the magnetic field be locally constant, an assumption that holds for distances of at most on the order of 1 *cm*. Instead, the more robust 4th order Runge-Kutta integration method is chosen. It offers the advantage of being able to use information on how the magnetic field varies along each of the four intermediate points that are part of its step prescription.

For the differential equations (Eqs. D.5 and D.10) for transporting a charged particle through a magnetic field, the prescription for the nth 4th order Runge-Kutta step, transporting the particle from location $\vec{x}_n$ to $\vec{x}_{n+1}$ with a change in momentum from $\vec{p}_n$ to $\vec{p}_{n+1}$, is given by the equations

$$\begin{aligned} \vec{x}_{n+1} &= \vec{x}_n + \frac{1}{6}\left(\vec{k}_{1x} + 2\vec{k}_{2x} + 2\vec{k}_{3x} + \vec{k}_{4x}\right) \\ \vec{p}_{n+1} &= \vec{p}_n + \frac{1}{6}\left(\vec{k}_{1p} + 2\vec{k}_{2p} + 2\vec{k}_{3p} + \vec{k}_{4p}\right) , \end{aligned} \tag{D.12}$$

where the step parameters are

$$\begin{aligned} \vec{k}_{1x} &= \frac{\Delta t}{\gamma m}\,\vec{p}_n \\ \vec{k}_{1p} &= \Delta t\,\mathcal{M}(x_n)\,\vec{p}_n \end{aligned} \tag{D.13}$$

$$\begin{aligned}
\vec{k}_{2x} &= \frac{\Delta t}{\gamma m}\,(\vec{p}_n + \frac{1}{2}\vec{k}_{1p}) \\
\vec{k}_{2p} &= \Delta t\,\mathcal{M}(x_n + \frac{1}{2}\vec{k}_{1x})\,(\vec{p}_n + \frac{1}{2}\vec{k}_{1p}) \\
\vec{k}_{3x} &= \frac{\Delta t}{\gamma m}\,(\vec{p}_n + \frac{1}{2}\vec{k}_{2p}) \\
\vec{k}_{3p} &= \Delta t\,\mathcal{M}(x_n + \frac{1}{2}\vec{k}_{2x})\,(\vec{p}_n + \frac{1}{2}\vec{k}_{2p}) \\
\vec{k}_{4x} &= \frac{\Delta t}{\gamma m}\,(\vec{p}_n + \vec{k}_{3p}) \\
\vec{k}_{4p} &= \Delta t\,\mathcal{M}(x_n + \vec{k}_{3x})\,(\vec{p}_n + \vec{k}_{3p})\ .
\end{aligned}$$

The track model then begins with the particle at a given point in space with given momentum components, the above step rules are used to move this particle discretely through the whole spectrometer, and the hits in the MWPCs are recorded as this simulated track passes through each chamber. Two minor modifications, however, improve the performance of this method for the particular situation of this experiment.

Firstly, the magnitude of the momentum is apt to be perturbed by round-off errors. This can cause increasing errors in the position and momentum of a particle as it is transported through the spectrometer. A small error in momentum causes a deviation from the true track into a region of field different from where the particle was "supposed" to be. This, in turn, will cause a deflection in the particle's path that is different from what it would have been along the correct path. Thus, round-off errors compound with systematic deviations from the desired path for this particle. To cure this problem, the momentum components are renormalized at every integration step to maintain the correct magnitude. This, in effect, assumes that any round-off errors will occur with equal frequency in each of the momentum components.

The second modification addresses the optimal value for the step size. Adaptive step size control of the time step Δt was briefly attempted, but then removed. The spatial variation of the magnetic field in the spectrometer figured centrally in the determination of the optimal step size. However, the overhead of the step size adjustment to maximize speed and minimize the errors in the dependent variables cancelled any advantage in reducing the total number of steps through the spectrometer.

Rather, the most important modification to the general algorithm is that the particle is transported through the field-free regions before and after the spectrometer magnet in a single step each. The time step within the field region is, then, the only tunable parameter of the algorithm and its value was found optimal at $\Delta t = 0.15\ ns$. This is equivalent to a spatial step of about 4.5 *cm*. The objective in this determination was for maximum speed while maintaining accuracy in the MWPC hits. This accuracy was checked by replacing the measured C-magnet field map with a uniform field for which the particle trajectory was analytically calculable.

The time for passage of the $e^{\pm}$ through the entire spectrometer (210 *cm*) was about 7 *ns*, and about 80 *cm* of this distance was in the field region of the spectrometer magnet. Thus, a typical track requires on the order of 20 Runge-Kutta steps in tracing it through the full spectrometer. The fitting of an actual track from the data requires at least 14 – and usually not more than 14 – integration passes through the track model. For each of two iterations, one is for the initial value of the χ^2 of the track, five are for the calculation of the first derivatives of χ^2 with respect to the parameters, and one more is for the determination of χ^2 after the parameter correction. It is found that these typically 280 Runge-Kutta steps per event required an average of 130 *ms* of CPU time on the VAX6420/VAX8700/VAX8650 cluster computer system that was used for the data analysis.

Appendix E

Minimizing χ^2 for Track Fitting

E.1 Definition of the Weight Matrix

Since multiple scattering in the passage of decay $e^{\pm}$ through the spectrometer can cause appreciable deviations from an ideal, unperturbed track, this effect must be accounted for to obtain a meaningful χ^2 for the fit to each particle trajectory. In particular, if the particle leaving a track scatters at some point in the spectrometer, all further hits in the MWPCs will show correlated deviations from the ideal track. This correlation between hits and the relative importance of the hits in the different MWPCs must be incorporated into a carefully constructed weight matrix that appears in the definition of the appropriate χ^2 for the track fit.

To formalize this, define the measured hits in the MWPCs as

$$\vec{h} = \begin{pmatrix} \vec{x}_h \\ \vec{y}_h \end{pmatrix} , \tag{E.1}$$

where

$$\vec{x}_h = \begin{pmatrix} x_{1h} \\ x_{2h} \\ x_{3h} \\ x_{4h} \end{pmatrix} \tag{E.2}$$

and

$$\vec{y}_h = \begin{pmatrix} y_{1h} \\ y_{2h} \\ y_{3h} \\ y_{4h} \end{pmatrix} . \tag{E.3}$$

The fitted hits, for which the track parameters need to be determined, will be designated by

$$\vec{f} = \begin{pmatrix} \vec{x}_f \\ \vec{y}_f \end{pmatrix} , \tag{E.4}$$

where

$$\vec{x}_f = \begin{pmatrix} x_{1f} \\ x_{2f} \\ x_{3f} \\ x_{4f} \end{pmatrix} \tag{E.5}$$

and

$$\vec{y}_f = \begin{pmatrix} y_{1f} \\ y_{2f} \\ y_{3f} \\ y_{4f} \end{pmatrix} . \tag{E.6}$$

The χ^2 for the track fit is then defined as

$$\chi^2 = \left(\vec{h} - \vec{f}\right)^T \mathcal{W} \left(\vec{h} - \vec{f}\right) , \tag{E.7}$$

where $\mathcal{W}$ is the weight matrix that contains all the information about the relative importance of each hit and the correlations between them.

Short of including scattering angles as fit parameters to the track model, multiple scattering can not be "corrected" for any given measured track. Rather, one must treat scattering as a stochastic process which introduces uncertainty to the measured hits in addition to any intrinsic measurement error (*i.e.* wire spacing). The correct choice for the weight matrix in this case is the inverse of the covariance matrix of the errors in the MWPC hits:

$$\mathcal{W} = \Sigma^{-1} . \tag{E.8}$$

In terms of the measured and fitted MWPC hits, the covariance matrix is defined as

$$\Sigma_{ij} = \mathrm{cov}\left(\left(\vec{h} - \vec{f}\right)_i , \left(\vec{h} - \vec{f}\right)_j\right) . \tag{E.9}$$

Since the fitted hits derive from the track model presented in Appendix D, which describes "ideal" tracks that suffer neither energy loss nor scattering, the difference between such fitted and the measured hits will be characterized by the effects of multiple scattering and energy loss. The energy loss of a decay $e^{\pm}$ is negligible in the spectrometer, so the covariance matrix is constructed solely by considering multiple scattering and the intrinsic resolution of the chambers.

The random variables in which to express the covariances between the MWPC hits are chosen to be the projected planar scattering angles and uncorrelated measurement errors for each plane that are due to the intrinsic chamber plane resolution. With this choice of random variables, there are no correlations between measurements of x-coordinates and y-coordinates in the MWPCs and the covariance matrix becomes block-diagonal:

$$\Sigma = \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_y \end{pmatrix} . \tag{E.10}$$

This entails a block-diagonal weight matrix

$$\mathcal{W} = \begin{pmatrix} \mathcal{W}_x & 0 \\ 0 & \mathcal{W}_y \end{pmatrix} , \tag{E.11}$$

where

$$\begin{aligned} \mathcal{W}_x &= \Sigma_x^{-1} \\ \mathcal{W}_y &= \Sigma_y^{-1} . \end{aligned} \tag{E.12}$$

Thus, we need to evaluate

$$(\Sigma_x)_{ij} = \mathrm{cov}\left((\vec{\epsilon}_x)_i , (\vec{\epsilon}_x)_j\right) \tag{E.13}$$

and

$$(\Sigma_y)_{ij} = \mathrm{cov}\left((\vec{\epsilon}_y)_i , (\vec{\epsilon}_y)_j\right) , \tag{E.14}$$

where

$$\vec{\epsilon}_x = (\vec{x}_h - \vec{x}_f) \qquad \text{(E.15)}$$
$$\vec{\epsilon}_y = (\vec{y}_h - \vec{y}_f) \ .$$

The vectors $\vec{\epsilon}_x$ and $\vec{\epsilon}_y$ are the deviations of an actual track from the track model. Next, these deviations will be expressed in terms of the appropriate random variables describing the scattering and measurement process.

In the following, the intrinsic measurement errors in MWPC i are denoted by $\delta_{i,x}$ and $\delta_{i,y}$. The planar scattering angles in the air between MWPCs i and j are denoted by $\theta_{ij,x}$ and $\theta_{ij,y}$ and the planar scattering angles due to passage through MWPC i are $\theta_{MWPC\,i,x}$ and $\theta_{MWPC\,i,y}$. The length of the track path between MWPCs i and j is symbolized by l_{ij} and the planar angles of incidence on MWPC i are $\phi_{i,x}$ and $\phi_{i,y}$. With this notation, the deviations in the x-hits may be written as

$$\epsilon_{x1} = \delta_{1,x} \qquad \text{(E.16)}$$

$$\epsilon_{x2} = \delta_{2,x} + \frac{l_{12}\theta_{12,x}}{\sqrt{3}\cos\phi_{2,x}} \qquad \text{(E.17)}$$

$$\begin{aligned} \epsilon_{x3} = {} & \delta_{3,x} + \frac{l_{12}\theta_{12,x}}{\sqrt{3}\cos\phi_{2,x}} \\ & + \frac{l_{23}\theta_{12,x}}{\cos\phi_{3,x}} + \frac{l_{23}\theta_{MWPC\,2,x}}{\cos\phi_{3,x}} + \frac{l_{23}\theta_{23,x}}{\sqrt{3}\cos\phi_{3,x}} \end{aligned} \qquad \text{(E.18)}$$

$$\begin{aligned} \epsilon_{x4} = {} & \delta_{4,x} + \frac{l_{12}\theta_{12,x}}{\sqrt{3}\cos\phi_{2,x}} \\ & + \frac{l_{24}\theta_{12,x}}{\cos\phi_{4,x}} + \frac{l_{24}\theta_{MWPC\,2,x}}{\cos\phi_{4,x}} + \frac{l_{23}\theta_{23,x}}{\sqrt{3}\cos\phi_{3,x}} \\ & + \frac{l_{34}\theta_{23,x}}{\cos\phi_{4,x}} + \frac{l_{34}\theta_{MWPC\,3,x}}{\cos\phi_{4,x}} + \frac{l_{34}\theta_{34,x}}{\sqrt{3}\cos\phi_{4,x}} \ . \end{aligned} \qquad \text{(E.19)}$$

In complete analogy, the deviations from an ideal track in the y-hits are

$$\epsilon_{y1} = \delta_{1,y} \qquad \text{(E.20)}$$
$$\epsilon_{y2} = \delta_{2,y} + \frac{l_{12}\theta_{12,y}}{\sqrt{3}\cos\phi_{2,y}}$$

$$
\begin{aligned}
\epsilon_{y3} &= \delta_{3,y} + \frac{l_{12}\theta_{12,y}}{\sqrt{3}\cos\phi_{2,y}} \\
&\quad + \frac{l_{23}\theta_{12,y}}{\cos\phi_{3,y}} + \frac{l_{23}\theta_{MWPC\,2,y}}{\cos\phi_{3,y}} + \frac{l_{23}\theta_{23,y}}{\sqrt{3}\cos\phi_{3,y}} \\
\epsilon_{y4} &= \delta_{4,y} + \frac{l_{12}\theta_{12,y}}{\sqrt{3}\cos\phi_{2,y}} \\
&\quad + \frac{l_{24}\theta_{12,y}}{\cos\phi_{4,y}} + \frac{l_{24}\theta_{MWPC\,2,y}}{\cos\phi_{4,y}} + \frac{l_{23}\theta_{23,y}}{\sqrt{3}\cos\phi_{3,y}} \\
&\quad + \frac{l_{34}\theta_{23,y}}{\cos\phi_{4,y}} + \frac{l_{34}\theta_{MWPC\,3,y}}{\cos\phi_{4,y}} + \frac{l_{34}\theta_{34,y}}{\sqrt{3}\cos\phi_{4,y}} \ .
\end{aligned}
$$

The random variables $\delta_{i,x}$, $\delta_{i,y}$, $\theta_{ij,x}$, $\theta_{ij,y}$, $\theta_{MWPC\,i,x}$, and $\theta_{MWPC\,i,y}$ are independent by construction so that the covariance matrix can be easily evaluated. In the equations for the deviations, the values of these random variables for a particular track enter. Of course, we can never determine these values for an individual track, but, as the calculation of the covariance matrix involves finding the expectation values of the random variables for each entry, we will be left with error estimates valid for any track. These estimates then depend only on the models for the distributions of the random variables as this determines their expectation values.

The above expressions for the track perturbations at the MWPC planes have been designed assuming that one can estimate the r.m.s. planar multiple scattering angles in a slice of material according to an empirical fit to Molière theory [Yos88]:

$$\theta^{\rm rms}_{\rm plane} = \frac{14.1\ MeV/c}{p\beta} Z_{\rm inc}\sqrt{\frac{L}{L_R}}\left[1 + \frac{1}{9}\log_{10}\left(\frac{L}{L_R}\right)\right] \text{ radians} \ . \tag{E.21}$$

In this equation, p, β, and $Z_{\rm inc}$ are, respectively, the momentum in MeV/c, the velocity in units of c_0 (vacuum speed of light), and the charge in units of e (electron charge) of the incident particle. The ratio L/L_R is the thickness of the scattering medium in units of its radiation length. The range of validity of the expression is $10^{-3} < L/L_R < 10$.

The MWPCs are treated as scatterers of negligible thickness compared to the overall length of the track, whereas the air between them is treated as a bulk scattering medium. To find the deflection due to scattering in a region of air, it is necessary to estimate the length of the particle track between the two bounding

MWPCs. This may be found approximately from the coordinates of the measured MWPC plane hits. Also, since the simple estimate of transverse deflection to a particle due to scattering over a path of length L in a medium [Yos88],

$$\Delta^{\text{rms}}_{\text{plane}} = \frac{1}{\sqrt{3}} L \, \theta^{\text{rms}}_{\text{plane}} , \tag{E.22}$$

holds in a coordinate system affixed to its particular direction of motion, these deflections need to be transformed into the coordinate system of the MWPC planes. Both of these considerations require knowledge of the direction cosines of the particle at each point. These may again be estimated from the measured MWPC hits. The deflections of a particle in each part of the spectrometer are then added together over the portion of the track up to the MWPC plane for which the deviation is being calculated.

Using Eq. E.21 to estimate the r.m.s. values of the random variable scattering angles in MWPCs and air and assuming that the r.m.s. value of the intrinsic measurement error in all MWPC planes is

$$\sqrt{\langle \delta^2 \rangle} = \sigma , \tag{E.23}$$

the expectation values in the elements of the covariance matrix may be evaluated to give

$$\begin{aligned}
\Sigma_{x,1,1} &= \sigma^2 && \text{(E.24)} \\
\Sigma_{x,1,2} &= \Sigma_{x,2,1} = 0 \\
\Sigma_{x,1,3} &= \Sigma_{x,3,1} = 0 \\
\Sigma_{x,1,4} &= \Sigma_{x,4,1} = 0 \\
\Sigma_{x,2,2} &= \sigma^2 + \frac{l_{12}^2 \langle \theta_{12,x}^2 \rangle}{3 \cos^2 \phi_{2,x}}
\end{aligned}$$

$$\Sigma_{x,2,3} = \frac{l_{12}\langle\theta^2_{12,x}\rangle}{\sqrt{3}\cos\phi_{2,x}}\left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,x}} + \frac{l_{23}}{\cos\phi_{3,x}}\right]$$

$$\Sigma_{x,3,2} = \Sigma_{x,2,3}$$

$$\Sigma_{x,2,4} = \frac{l_{12}\langle\theta^2_{12,x}\rangle}{\sqrt{3}\cos\phi_{2,x}}\left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,x}} + \frac{l_{24}}{\cos\phi_{4,x}}\right]$$

$$\Sigma_{x,4,2} = \Sigma_{x,2,4}$$

$$\begin{aligned}\Sigma_{x,3,3} = {} & \sigma^2 + \left(\frac{l_{12}}{\sqrt{3}\cos\phi_{2,x}} + \frac{l_{23}}{\cos\phi_{3,x}}\right)^2 \langle\theta^2_{12,x}\rangle \\ & + \left(\frac{l_{23}}{\cos\phi_{3,x}}\right)^2 \langle\theta^2_{MWPC\,2,x}\rangle + \left(\frac{l_{23}}{\sqrt{3}\cos\phi_{3,x}}\right)^2 \langle\theta^2_{23,x}\rangle\end{aligned}$$

$$\begin{aligned}\Sigma_{x,3,4} = {} & \frac{l_{12}}{\sqrt{3}\cos\phi_{2,x}}\langle\theta^2_{12,x}\rangle\left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,x}} + \frac{l_{23}}{\cos\phi_{3,x}} + \frac{l_{24}}{\cos\phi_{4,x}}\right] \\ & + \frac{l_{23}\,l_{24}}{\cos\phi_{3,x}\cos\phi_{4,x}}\langle\theta^2_{MWPC\,2,x}\rangle + \frac{l_{23}}{\sqrt{3}\cos\phi_{3,x}}\langle\theta^2_{23,x}\rangle\left[\frac{l_{23}}{\sqrt{3}\cos\phi_{3,x}} + \frac{l_{34}}{\cos\phi_{4,x}}\right]\end{aligned}$$

$$\Sigma_{x,4,3} = \Sigma_{x,3,4}$$

$$\begin{aligned}\Sigma_{x,4,4} = {} & \sigma^2 + \left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,x}} + \frac{l_{24}}{\cos\phi_{4,x}}\right]^2 \langle\theta^2_{12,x}\rangle + \left(\frac{l_{24}}{\cos\phi_{4,x}}\right)^2 \langle\theta^2_{MWPC\,2,x}\rangle \\ & + \left[\frac{l_{23}}{\sqrt{3}\cos\phi_{3,x}} + \frac{l_{34}}{\cos\phi_{4,x}}\right]^2 \langle\theta^2_{23,x}\rangle + \left(\frac{l_{34}}{\cos\phi_{4,x}}\right)^2 \left[\langle\theta^2_{MWPC\,3,x}\rangle + \frac{1}{3}\langle\theta^2_{34,x}\rangle\right]\end{aligned}$$

for the matrix of x-covariances and

$$\Sigma_{y,1,1} = \sigma^2 \tag{E.25}$$

$$\Sigma_{y,1,2} = \Sigma_{y,2,1} = 0$$

$$\Sigma_{y,1,3} = \Sigma_{y,3,1} = 0$$

$$\Sigma_{y,1,4} = \Sigma_{y,4,1} = 0$$

$$\Sigma_{y,2,2} = \sigma^2 + \frac{l_{12}^2 \langle \theta_{12,y}^2 \rangle}{\sqrt{3}\cos^2\phi_{2,y}}$$

$$\Sigma_{y,2,3} = \frac{l_{12}\langle \theta_{12,y}^2 \rangle}{\sqrt{3}\cos\phi_{2,y}} \left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,y}} + \frac{l_{23}}{\cos\phi_{3,y}} \right]$$

$$\Sigma_{y,3,2} = \Sigma_{y,2,3}$$

$$\Sigma_{y,2,4} = \frac{l_{12}\langle \theta_{12,y}^2 \rangle}{\sqrt{3}\cos\phi_{2,y}} \left[-\frac{l_{12}}{\sqrt{3}\cos\phi_{2,y}} + \frac{l_{24}}{\cos\phi_{4,y}} \right]$$

$$\Sigma_{y,4,2} = \Sigma_{y,2,4}$$

$$\begin{aligned} \Sigma_{y,3,3} &= \sigma^2 + \left(\frac{l_{12}}{\sqrt{3}\cos\phi_{2,y}} + \frac{l_{23}}{\cos\phi_{3,y}} \right)^2 \langle \theta_{12,y}^2 \rangle \\ &+ \left(\frac{l_{23}}{\cos\phi_{3,y}} \right)^2 \langle \theta_{MWPC\,2,y}^2 \rangle + \left(\frac{l_{23}}{\sqrt{3}\cos\phi_{3,y}} \right)^2 \langle \theta_{23,y}^2 \rangle \end{aligned}$$

$$\begin{aligned} \Sigma_{y,3,4} &= \frac{l_{12}}{\sqrt{3}\cos\phi_{2,y}} \langle \theta_{12,y}^2 \rangle \left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,y}} + \frac{l_{23}}{\cos\phi_{3,y}} + \frac{l_{24}}{\cos\phi_{4,y}} \right] \\ &+ \frac{l_{23}\,l_{24}}{\cos\phi_{3,y}\cos\phi_{4,y}} \langle \theta_{MWPC\,2,y}^2 \rangle + \frac{l_{23}}{\sqrt{3}\cos\phi_{3,y}} \langle \theta_{23,y}^2 \rangle \left[\frac{l_{23}}{\sqrt{3}\cos\phi_{3,y}} + \frac{l_{34}}{\cos\phi_{4,y}} \right] \end{aligned}$$

$$\Sigma_{y,4,3} = \Sigma_{y,3,4}$$

$$\begin{aligned} \Sigma_{y,4,4} &= \sigma^2 + \left[\frac{l_{12}}{\sqrt{3}\cos\phi_{2,y}} + \frac{l_{24}}{\cos\phi_{4,y}} \right]^2 \langle \theta_{12,y}^2 \rangle + \left(\frac{l_{24}}{\cos\phi_{4,y}} \right)^2 \langle \theta_{MWPC\,2,y}^2 \rangle \\ &+ \left[\frac{l_{23}}{\sqrt{3}\cos\phi_{3,y}} + \frac{l_{34}}{\cos\phi_{4,y}} \right]^2 \langle \theta_{23,y}^2 \rangle + \left(\frac{l_{34}}{\cos\phi_{4,y}} \right)^2 \left[\langle \theta_{MWPC\,3,y}^2 \rangle + \frac{1}{3} \langle \theta_{34,y}^2 \rangle \right] \end{aligned}$$

for the matrix of y-covariances. By construction, there are no cross-correlations between x- and y-hits:

$$\Sigma_{x,i;y,j} = 0 \quad \forall\,(i,j)\ . \tag{E.26}$$

Some simplifying assumptions are introduced by specifying the estimates of l_{ij} and $\phi_{i,x}$, $\phi_{i,y}$. Denoting the measured MWPC plane hits for a given track by x_i and y_i for $i \in \{1, \ldots, 4\}$, the planar direction cosines are given approximately by

$$\begin{aligned}
\cos\phi_{1,x} &= \cos\phi_{2,x} = \frac{(x_2 - x_1)}{\sqrt{(x_2 - x_1)^2 + (z_2 - z_1)^2}} \\
\cos\phi_{1,y} &= \cos\phi_{2,y} = \frac{(y_2 - y_1)}{\sqrt{(y_2 - y_1)^2 + (z_2 - z_1)^2}} \\
\cos\phi_{3,x} &= \cos\phi_{4,x} = \frac{(x_4 - x_3)}{\sqrt{(x_4 - x_3)^2 + (z_4 - z_3)^2}} \\
\cos\phi_{3,y} &= \cos\phi_{4,y} = \frac{(y_4 - y_3)}{\sqrt{(y_4 - y_3)^2 + (z_4 - z_3)^2}}\ ,
\end{aligned} \tag{E.27}$$

where z_i is the location of the ith MWPC along the spectrometer axis. For both transverse dimensions, the planar direction cosines of incidence are taken to be equal for each pair of MWPCs, before and after the magnet, since the field of the spectrometer magnet extends only negligibly into the regions between each chamber pair. This assumption leads also to expressions approximating the length of the track between adjacent MWPCs:

$$\begin{aligned}
l_{12} &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \\
l_{23} &= \left(\frac{z_c - z_2}{z_2 - z_1}\right)\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \\
&\quad + \left(\frac{z_3 - z_c}{z_4 - z_3}\right)\sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2 + (z_4 - z_3)^2} \\
l_{34} &= \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2 + (z_4 - z_3)^2} \\
l_{24} &= l_{23} + l_{34}\ ,
\end{aligned} \tag{E.28}$$

where z_c is the center-gap point of the magnet along the spectrometer axis. One slight deficiency in these expressions is the consistent overestimate of the actual length of the track from MWPC2 to MWPC3 by neglecting the smooth curving of the particle track in the magnetic field region. This was not found to significantly impact the performance of the fit. It modestly overestimated the multiple scattering between the second and third chambers which caused a slight under-weighting of the track coordinates in MWPCs 3 and 4.

Referring to Eq. E.21, it is clear that the momentum of the particle leaving the track in the MWPCs enters into the estimate of the r.m.s. scattering angles in the covariance matrix. Thus, a value for this momentum is required to find the elements of the covariance matrix. For the case of data taken with the spectrometer magnet turned off with the purpose of providing data to be used for the final off-line alignment of the MWPCs, this estimate comes simply from the calibrated pulse-height measured in the $NaI(Tl)$ crystal calorimeter. Requiring a sensible pulse height of the $NaI(Tl)$ crystal reduced the acceptance of the spectrometer, which does not spoil the alignment data, but is clearly undesirable in the search for $M \longrightarrow \overline{M}$ events. For MWPC alignment data, the $NaI(Tl)$ pulse height is the *only* available momentum estimate for the track, but when the magnetic field is on, one may construct a momentum estimate from the observed MWPC coordinate data together with the field map data.

To this end, the field of the spectrometer magnet is modeled by a region of fixed length with sudden boundaries (see Fig. E.1). The value of the field in this region is estimated by taking the field integrals of the magnet along the spectrometer axis, for given transverse coordinates, and dividing by the length of this region along z. Assuming, then, that the particle experiences a constant field,

$$B_{eff} = \frac{1}{\Delta z} \int_{z_c - \frac{\Delta z}{2}}^{z_c + \frac{\Delta z}{2}} B_y(x, y, z)\, dz \ , \tag{E.29}$$

along its trajectory through the magnetic field region, one can relate the radius of curvature of the track, R, to the deflection angle, θ, and the secant across the portion of the track in the field region, s, by

$$s = 2R \sin\frac{\theta}{2} \ . \tag{E.30}$$

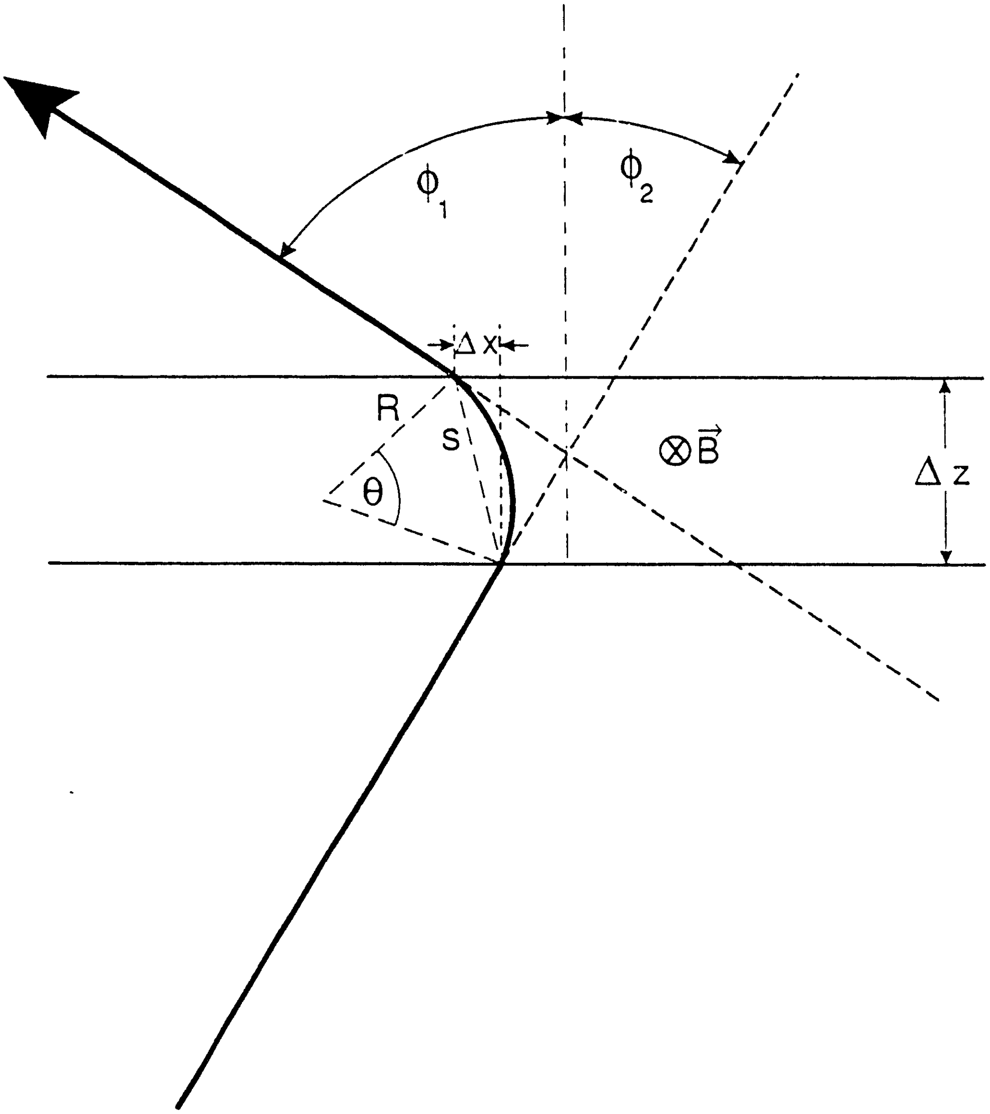


Figure E.1: Geometry for the momentum estimate to seed the track fit. The magnetic field is modeled by the values of the field integral over the effective length.

The secant is also related to the field length and the transverse deflection by

$$s^2 = \Delta x^2 + \Delta z^2 \ . \tag{E.31}$$

The angles of incidence entering and leaving the field region in the xz-plane add to give the total deflection angle

$$\theta = \phi_1 + \phi_2 \ . \tag{E.32}$$

Furthermore, the x-deflection and the field length may be related to these incidence and exit angles and the radius of curvature by

$$\begin{aligned} \Delta x &= R\,(\cos\phi_2 - \cos\phi_1) \qquad \text{(E.33)} \\ \Delta z &= R\,(\sin\phi_1 + \sin\phi_2) \ . \end{aligned}$$

Then, combining Eqs. E.30, E.31, E.32, and E.34, one can solve for the radius of curvature in terms of only the incidence angle, the transverse deflection, and the effective field length:

$$R = \frac{\Delta x^2 + \Delta z^2}{2\,(\Delta z \sin\phi_1 - \Delta x \cos\phi_1)} \ . \tag{E.34}$$

Finally, as the radius of curvature of a particle in a magnetic field is directly related to its momentum, the latter may be found from

$$p_{est} = e\,B_{eff}\,R \ . \tag{E.35}$$

The momentum thus estimated is used to seed the χ^2 minimization algorithm for fitting to the particle track.

E.2 Fit to Alignment Tracks

For the fit to straight-line tracks acquired to align the MWPCs, it is possible to specify the solution for the track parameters in closed form. This is a direct consequence of the linearity of the fit function (a straight line) in its parameters (slopes and intercepts in each of the xz- and yz-planes). The results of this fit are then used to estimate the necessary alignment corrections to the coordinates measured in each of the MWPC planes to ensure that they relate to a coordinate system common to all chambers.

The fit functions, one for each planar view of the track, are

$$x_{i,f} = (z_i - z_{ref})\, m_{xz} + x_{ref} \tag{E.36}$$

and

$$y_{i,f} = (z_i - z_{ref})\, m_{yz} + y_{ref} \ . \tag{E.37}$$

Here, the chamber index is $i \in \{2, \ldots, 4\}$, while the "reference chamber" is taken to be MWPC1. This means that no corrections will be applied to the coordinates of MWPC1; rather, the coordinate systems of all other MWPCs will be adjusted such that they agree with that of the reference chamber. Fixing the coordinate system of one of the MWPCs in this way prevents an overall "travelling" of the corrected spectrometer coordinate system, ensures minimal adjustments to the coordinates measured in the other chambers, and entails a simplification of the fit.

The x- and y-intercepts may be immediately identified as $x_{ref} = x_1$ and $y_{ref} = y_1$. This reduces the fit to determining the best slopes, m_{xz} and m_{yz}. Assuming that there are no xy-correlations in the measured MWPC hits, the fit decouples into separate linear fits in the xz- and in the yz-planes. Each of these have three remaining data points (discounting the hit in the reference chamber) with the slope as the only fit parameter, leaving 2 degrees of freedom. The χ^2 for the fit may be written as

$$\begin{aligned} \chi^2 &= \vec{\epsilon}^{\,T}\, \mathcal{W}\, \vec{\epsilon} \qquad\qquad \text{(E.38)} \\ &= \left(\vec{\epsilon}_x^{\,T}, \vec{\epsilon}_y^{\,T}\right) \begin{pmatrix} \mathcal{W}_x & 0 \\ 0 & \mathcal{W}_y \end{pmatrix} \begin{pmatrix} \vec{\epsilon}_x \\ \vec{\epsilon}_y \end{pmatrix} \\ &= \vec{\epsilon}_x^{\,T}\, \mathcal{W}_x\, \vec{\epsilon}_x + \vec{\epsilon}_y^{\,T}\, \mathcal{W}_y\, \vec{\epsilon}_y \\ &= \chi_x^2 + \chi_y^2 \ , \end{aligned}$$

where one may identify

$$\chi_x^2 = \vec{\epsilon}_x^{\,T}\, \mathcal{W}_x\, \vec{\epsilon}_x \tag{E.39}$$

and

$$\chi_y^2 = \vec{\epsilon}_y^{\,T}\, \mathcal{W}_y\, \vec{\epsilon}_y \ . \tag{E.40}$$

It is now possible to redefine the variables of the fit to be

$$\begin{aligned} \Delta x_i &= x_{i,h} - x_{ref} \ , \\ \Delta y_i &= y_{i,h} - y_{ref} \ , \end{aligned} \tag{E.41}$$

and

$$\Delta z_i = z_i - z_{ref} \ , \tag{E.42}$$

so the χ^2's become

$$\chi_x^2 = \sum_{ij} \left(\Delta x_i - m_{xz} \Delta z_i\right) \mathcal{W}_{x,ij} \left(\Delta x_j - m_{xz} \Delta z_j\right) \tag{E.43}$$

and

$$\chi_y^2 = \sum_{ij} \left(\Delta y_i - m_{yz} \Delta z_i\right) \mathcal{W}_{y,ij} \left(\Delta y_j - m_{yz} \Delta z_j\right) \ . \tag{E.44}$$

The minimization conditions are

$$\frac{\partial \chi_x^2}{\partial m_{xz}} = 0 \tag{E.45}$$

and

$$\frac{\partial \chi_y^2}{\partial m_{yz}} = 0 \ . \tag{E.46}$$

Though the weight matrix is dependent on the fit parameters through the length of the track between the MWPCs, the dependence is weak near the χ^2 minimum. One may, therefore, assume that the weight matrix is stationary for this problem, allowing the closed form solutions

$$m_{xz} = \frac{\Delta \vec{z}^{\,T} \mathcal{W}_x \Delta \vec{x}}{\Delta \vec{z}^{\,T} \mathcal{W}_x \Delta \vec{z}} \tag{E.47}$$

and

$$m_{yz} = \frac{\Delta \vec{z}^{\,T} \mathcal{W}_y \Delta \vec{y}}{\Delta \vec{z}^{\,T} \mathcal{W}_y \Delta \vec{z}} \ . \tag{E.48}$$

The 1-σ errors on the slopes can be determined from the condition that the χ^2 be increased by unity. The calculation is a little lengthy, but straightforward, and gives

$$\sigma_{m_{xz}} = \sqrt{\frac{1}{\Delta \vec{z}^{\,T} \mathcal{W}_x \Delta \vec{z}}} \tag{E.49}$$

and

$$\sigma_{m_{yz}} = \sqrt{\frac{1}{\Delta\vec{z}^{T}\,\mathcal{W}_y\,\Delta\vec{z}}}\,. \tag{E.50}$$

Since the weight matrix is symmetric, the denominator under the root is a proper quadratic form and thus positive definite.

E.3 Fit to Shift Corrections of MWPC Coordinates

With linear fits to a large sample of tracks taken with the spectrometer magnet turned off, one may determine the best shift corrections to the measured coordinates in MWPCs 2, 3, and 4 to align them with the transverse coordinate axes of MWPC1. This procedure is expressible as a fit to a constant for each x- and y-plane of chambers 2, 3, and 4. An important assumption at this point is that there are no rotational misalignments between the MWPCs, so that the x- and y-coordinates for a given chamber do not "mix." In practice there will be such misrotations, but they are known to be small. The procedure to correct for them is described in Section 4.3 and is performed iteratively in conjunction with the correction for shift misalignments described here.

The shift corrections for the MWPC planes according to a single fitted straight line track would simply be the deviations, $\vec{\epsilon}_x$ and $\vec{\epsilon}_y$, defined above. These may be rewritten, according to Eqs. E.16, E.36, E.37, E.41, and E.42, as

$$\begin{aligned} \vec{\epsilon}_x &= \Delta\vec{x} - m_{xz}\Delta\vec{z} \\ &= \Delta\vec{x} - \left(\frac{\Delta\vec{z}^{T}\,\mathcal{W}_x\,\Delta\vec{x}}{\Delta\vec{z}^{T}\,\mathcal{W}_x\,\Delta\vec{z}}\right)\Delta\vec{z} \\ \\ \vec{\epsilon}_y &= \Delta\vec{y} - m_{yz}\Delta\vec{z} \\ &= \Delta\vec{y} - \left(\frac{\Delta\vec{z}^{T}\,\mathcal{W}_y\,\Delta\vec{y}}{\Delta\vec{z}^{T}\,\mathcal{W}_y\,\Delta\vec{z}}\right)\Delta\vec{z} \end{aligned} \tag{E.51}$$

The hypothesis, now, is that these deviations are sampled from distributions whose means are the best corrections to the measured MWPC hits. Over the sample of all tracks recorded under alignment conditions, we may take the weighted average of the deviations for each track to obtain the best estimate for the required shift corrections to each MWPC plane coordinate.

This weighted average requires an error estimate for the deviations, $\vec{\epsilon}_x$ and $\vec{\epsilon}_y$. To calculate these errors, one notes that all random variable dependence in the deviations in Eq. E.51 is contained in the quantities $\Delta\vec{x}$ and $\Delta\vec{y}$. Estimating the uncertainties in the track deviations by the roots of their diagonal covariances gives

$$\begin{aligned} \sigma^2_{\epsilon_{x,i}} &= \Sigma_{x,ii} - \frac{(\Delta z_i)^2}{\Delta\vec{z}^T\,\mathcal{W}_x\,\Delta\vec{z}} \\ &= \Sigma_{x,ii} - \sigma^2_{m_{xz}}\,(\Delta z_i)^2 \end{aligned} \tag{E.52}$$

and

$$\begin{aligned} \sigma^2_{\epsilon_{y,i}} &= \Sigma_{y,ii} - \frac{(\Delta z_i)^2}{\Delta\vec{z}^T\,\mathcal{W}_y\,\Delta\vec{z}} \\ &= \Sigma_{y,ii} - \sigma^2_{m_{yz}}\,(\Delta z_i)^2 \ . \end{aligned} \tag{E.53}$$

These errors are used to weight the determinations of the best alignment shift corrections on the placements of MWPCs 2, 3, and 4 relative to the reference chamber, MWPC1, according to

$$\overline{\epsilon_{x,i}} = \frac{1}{D_x}\sum_n \frac{\epsilon^{(n)}_{x,i}}{\left(\sigma^{(n)}_{\epsilon_{x,i}}\right)^2} \tag{E.54}$$

and

$$\overline{\epsilon_{y,i}} = \frac{1}{D_y}\sum_n \frac{\epsilon^{(n)}_{y,i}}{\left(\sigma^{(n)}_{\epsilon_{y,i}}\right)^2} \tag{E.55}$$

where n is the index labeling the event number, which runs over all tracks recorded for alignment purposes at fixed conditions, and

$$D_x = \sum_n \frac{1}{\left(\sigma^{(n)}_{\epsilon_{x,i}}\right)^2} \tag{E.56}$$

and

$$D_y = \sum_n \frac{1}{\left(\sigma^{(n)}_{\epsilon_{y,i}}\right)^2} \ . \tag{E.57}$$

The errors on these shift corrections are

$$\sigma_{\overline{\epsilon_{x,i}}} = \frac{1}{\sqrt{D_x}} \tag{E.58}$$

and

$$\sigma_{\overline{\epsilon_{y,i}}} = \frac{1}{\sqrt{D_y}} \ . \tag{E.59}$$

Since the determination of the best alignment corrections is tantamount to fits to constants for each chamber plane, one may calculate a χ^2 for each of these to judge the overall quality of the alignment procedure:

$$\begin{aligned} \chi^2_{\epsilon_{x,i}} &= \sum_n \frac{\left(\epsilon^{(n)}_{x,i} - \overline{\epsilon_{x,i}}\right)^2}{\left(\sigma^{(n)}_{\epsilon_{x,i}}\right)^2} \\ \chi^2_{\epsilon_{y,i}} &= \sum_n \frac{\left(\epsilon^{(n)}_{y,i} - \overline{\epsilon_{y,i}}\right)^2}{\left(\sigma^{(n)}_{\epsilon_{y,i}}\right)^2} \ . \end{aligned} \tag{E.60}$$

For the alignment data, these χ^2's differed no more than a few % from the number of events fitted. This means that the χ^2 per degree of freedom is very nearly unity, indicating believable alignment shift results.

E.4 The χ^2 Minimization for Curved Tracks

Since there is no closed functional form for particle tracks in the spectrometer when the magnetic field is turned on, the fit to particle trajectories for most of the data must rely on an iterative numerical algorithm to converge upon the best parameters describing these tracks. To arrive at a prescription that accomplishes this, we can use the fact that the initial track parameter estimates rarely lie far from their best fit values. Futhermore, the behavior of the χ^2 as defined in Eq. E.7 along each of the parameter axes was found to be quite parabolic. Thus, one may fit a paraboloid to the χ^2 hypersurface and estimate the parameter values of the best fit from the minimum of this paraboloid surface. To ensure that the procedure has indeed converged, the paraboloid fit may be iterated until the resulting value of the χ^2 minimum changes by less than some given fraction from one step to the next.

We need, therefore, an expansion of the χ^2 through second order in the parameters of the track fit. It can be shown that this is completely equivalent to carrying out a second-order Taylor expansion of the fit function in the parameters [Bev69]. First, it is in order to restate the form of χ^2, displaying explicitly the parameter dependence in the fit function, which in this case is given by the track model discussed in Appendix D:

$$\begin{aligned} \chi^2 &= \sum_{ij} \left[(x_{i,h} - x_{i,f}(\vec{a}))\, \mathcal{W}_{x,ij}\, (x_{j,h} - x_{j,f}(\vec{a})) \right. \\ &+ \left. (y_{i,h} - y_{i,f}(\vec{a}))\, \mathcal{W}_{y,ij}\, (y_{j,h} - y_{j,f}(\vec{a})) \right] . \end{aligned} \tag{E.61}$$

The track parameters are symbolized by $\vec{a}$. They have been chosen as

$$\begin{aligned} a_1 &= \frac{1}{p} = \text{inverse magnitude of momentum} \\ a_2 &= x_{1,f} = x-\text{position of particle at MWPC1} \\ a_3 &= y_{1,f} = y-\text{position of particle at MWPC1} \\ a_4 &= c_x = x-\text{direction cosine at MWPC1} \\ a_5 &= c_y = y-\text{direction cosine at MWPC1} . \end{aligned} \tag{E.62}$$

Since the particle deflection in the spectrometer magnet depends inversely upon its momentum, the given choice for a_1 is superior to using the momentum directly as a fit parameter – the dependence of χ^2 of the track fit upon a_1 is parabolic, whereas it is rather pathological when the momentum itself is the fit parameter. The charge of the particle is determined by simply comparing the in-going and out-going x-direction cosines with respect to the magnetic field region, as determined from the track asymptotes. The latter are approximated by the lines passing through the chamber pair before and the chamber pair after the field region. As the particle charge is not a continuous parameter, it is determined once in this way and passed to the track fitting routine as a fixed condition.

With respect to the parameters thus defined, one seeks the global minimum of χ^2 determined by

$$\frac{\partial \chi^2}{\partial a_n} = 0 \quad \forall\, n \in \{1, \cdots, 5\} . \tag{E.63}$$

Since the initial guesses at the a_n lie close to the solution sought, there are no complications due to multiple minima. Also, there is no danger of converging upon a maximum, so the curvature of the χ^2 surface need not be checked to be concave upward. Inserting the definition of χ^2 from Eq. E.61 into Eq. E.63 gives

$$0 = \sum_{i,j} \left[\left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}} \mathcal{W}_{x,ij} \left(x_{j,h} - x_{j,f}(\vec{a})\right) + \left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}} \mathcal{W}_{y,ij} \left(y_{j,h} - y_{j,f}(\vec{a})\right) \right] , \qquad \text{(E.64)}$$

where the symmetry of the weight matrix has been used and it has been assumed that its dependence on the parameters may be neglected. This means that the initial parameter estimates are used throughout to determine the weight matrix, an approximation that holds because the starting parameter values are not far from the minimum sought. At this point, one needs the second-degree Taylor expansions of the fit function around the desired minimum:

$$\begin{aligned} x_{i,f}(\vec{a}) &= x_{i,f}(\vec{a}_0) + \sum_k \left.\frac{\partial x_{i,f}}{\partial a_k}\right|_{\vec{a}_0} \delta a_k + \sum_{k,m} \left.\frac{\partial^2 x_{i,f}}{\partial a_k \partial a_m}\right|_{\vec{a}_0} \delta a_k \delta a_m \\ y_{i,f}(\vec{a}) &= y_{i,f}(\vec{a}_0) + \sum_k \left.\frac{\partial y_{i,f}}{\partial a_k}\right|_{\vec{a}_0} \delta a_k + \sum_{k,m} \left.\frac{\partial^2 y_{i,f}}{\partial a_k \partial a_m}\right|_{\vec{a}_0} \delta a_k \delta a_m , \qquad \text{(E.65)} \end{aligned}$$

where $\delta\vec{a} = \vec{a} - \vec{a}_0$ and $k, m, n \in \{1, \ldots, 5\}$. The present parameter values are $\vec{a}$ and the parameter values at the minimum are $\vec{a}_0$. The first derivatives of the fit functions are thus approximately

$$\left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}} = \left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}_0} + \sum_m \left.\frac{\partial^2 x_{i,f}}{\partial a_n \partial a_m}\right|_{\vec{a}_0} \delta a_m \qquad \text{(E.66)}$$

and

$$\left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}} = \left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}_0} + \sum_m \left.\frac{\partial^2 y_{i,f}}{\partial a_n \partial a_m}\right|_{\vec{a}_0} \delta a_m . \qquad \text{(E.67)}$$

It is not necessary to retain terms beyond first order in the δa_n in Eq. E.64, since the intent is to keep χ^2 itself only up through second order in δa_n. Thus, the derivatives of $x_{i,f}$ and $y_{i,f}$ have been given only up to first order in δa_n. Substituting these expansions into Eq. E.64 and collecting orders gives

$$0 = \sum_{ij} \left[\left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left(x_{j,h} - x_{j,f}(\vec{a}_0)\right) + \left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left(y_{j,h} - y_{j,f}(\vec{a}_0)\right) \right]$$

$$
\begin{aligned}
-\sum_k \sum_{ij} \Bigg[& \left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left.\frac{\partial x_{j,f}}{\partial a_k}\right|_{\vec{a}_0} + \left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left.\frac{\partial y_{j,f}}{\partial a_k}\right|_{\vec{a}_0} \\
& - \left.\frac{\partial^2 x_{i,f}}{\partial a_n \partial a_k}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left(x_{j,h} - x_{j,f}(\vec{a}_0)\right) \\
& - \left.\frac{\partial^2 y_{i,f}}{\partial a_n \partial a_k}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left(y_{j,h} - y_{j,f}(\vec{a}_0)\right) \Bigg] \delta a_k \ ,
\end{aligned}
\tag{E.68}
$$

with $n, k \in \{1, \ldots, 5\}$ and $i, j \in \{1, \ldots, 4\}$. This may be written as

$$
\beta_n = \sum_k \alpha_{nk} \, \delta a_k \ , \tag{E.69}
$$

where

$$
\beta_k = \sum_{ij} \left[\left.\frac{\partial x_{i,f}}{\partial a_k}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left(x_{j,h} - x_{j,f}(\vec{a}_0)\right) + \left.\frac{\partial y_{i,f}}{\partial a_k}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left(y_{j,h} - y_{j,f}(\vec{a}_0)\right) \right] \tag{E.70}
$$

and

$$
\begin{aligned}
\alpha_{nk} \ = \ \sum_{ij} \Bigg[& \left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left.\frac{\partial x_{j,f}}{\partial a_k}\right|_{\vec{a}_0} + \left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left.\frac{\partial y_{j,f}}{\partial a_k}\right|_{\vec{a}_0} \\
& - \left.\frac{\partial^2 x_{i,f}}{\partial a_n \partial a_k}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left(x_{j,f} - x_{j,h}(\vec{a}_0)\right) \\
& - \left.\frac{\partial^2 y_{i,f}}{\partial a_n \partial a_k}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left(y_{j,f} - y_{j,h}(\vec{a}_0)\right) \Bigg] \ .
\end{aligned}
\tag{E.71}
$$

Since we wish to solve for the corrections δa_k that allow us to find the parameter values at the presumed χ^2 minimum from those at the current position on the χ^2 surface, we write

$$
\delta a_n = \sum_k \left(\alpha^{-1}\right)_{nk} \beta_k \tag{E.72}
$$

and the problem has been reduced to inverting the χ^2 curvature matrix, α. At this point one may, for calculational efficiency without loss of accuracy, drop the second derivative terms [Bev69] in α, since the factors $(x_{j,h} - x_{j,f}(\vec{a}_0))$ and $(y_{j,h} - y_{j,f}(\vec{a}_0))$ are very small for a good intial guess. This means that

$$
\alpha_{nk} \approx \sum_{ij} \left[\left.\frac{\partial x_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{x,ij} \left.\frac{\partial x_{j,f}}{\partial a_k}\right|_{\vec{a}_0} + \left.\frac{\partial y_{i,f}}{\partial a_n}\right|_{\vec{a}_0} \mathcal{W}_{y,ij} \left.\frac{\partial y_{j,f}}{\partial a_k}\right|_{\vec{a}_0} \right] \ . \tag{E.73}
$$

This ensures a positive-definite curvature matrix [Bev69] and significantly reduces computational effort, since it avoids the need to numerically find second derivatives. This approximation is equivalent to the assumption of relative independence of the fit parameters near the minimum.

A final practical concern is the fact that the entries in the curvature matrix, α, have differing dimensions according to the various dimensions of the parameters. As a result, the entries of α can have very disparate orders of magnitude. This situation is often fatal to the matrix inversion algorithm used [Bev69]. To avoid such trouble, the curvature matrix is rescaled to become dimensionless according to

$$\alpha'_{nk} = \frac{\alpha_{nk}}{\sqrt{\alpha_{nn}\alpha_{kk}}} , \tag{E.74}$$

which may be shown to entail

$$\left(\alpha^{-1}\right)_{nk} = \frac{(\alpha')^{-1}_{nk}}{\sqrt{\alpha_{nn}\alpha_{kk}}} . \tag{E.75}$$

Thus, the final form of the prescription to step from the present location in parameter space to the presumed minimum of χ^2 is

$$\vec{a}_0 = \vec{a} - \delta\vec{a} \tag{E.76}$$

with

$$\delta a_n = \sum_k \frac{(\alpha')^{-1}_{nk}\,\beta_k}{\sqrt{\alpha_{nn}\alpha_{kk}}} . \tag{E.77}$$

Appendix F

Determination of MWPC Resolution Function

It may be argued that the resolution of the reconstruction of the decay origin of e^+ tracks observed in the spectrometer may be modeled by a Gaussian distribution, as this is the approximate behavior of multiple scattering angles [Yos88]. But the multiple scattering distribution deviates from a Gaussian in such a way that we may not neglect it for this experiment. In particular, multiple scattering is more likely to give large angle deflections than the Gaussian approximation would imply [Bet53]. As it is of central importance to reliably determine the number of M atom decays from the vacuum, where these decays may originate from regions very close to the target, one must be able to distinguish the contribution from M decays from those of scattered target decays. Understanding the resolution of the reconstruction gives the tool to accomplish this.

Thus, to derive the correct y-position distributions that describe the decays of μ^+ and M in the target and the decays of M and $\overline{M}$ atoms from the vacuum, it is necessary to have a quantitative measure of the finite position resolution in the reconstruction of the decay origin by the MWPCs. The principal contributor to finite resolution is the multiple scattering in the material of the Al vacuum window, the air between the window and MWPC1, the material of MWPC1, and the air between MWPCs 1 and 2. The intrinsic resolution of about 1 mm in the chamber planes due to the 2 *mm* wire spacing is also a factor, though not dominant.

In order to measure directly the resolution function of the reconstruction, data were taken with an Al foil target of 0.026 *cm* thickness in place of the usual SiO_2 powder targets. The area density of this foil at 7.0 mg/cm^2 is comparable to that of the powder targets used while searching for the $M \longrightarrow \overline{M}$ conversion, but its physical thickness is negligible by comparison. Therefore, it presented in effect a point source along the direction perpendicular to its surface (y), so that its image from reconstruction along this axis is a direct measurement of the resolution function.

The data taken to measure this resolution are, however, of finite statistics and overlayed by a small amount of "background", counts that result mainly from μ^+ that scatter in the Al target, pass through it, but lodge themselves in the wall of the vacuum chamber in regions from where their decay e^+ may still be seen in the chambers of the spectrometer. Thus, one requires an algorithm that can extract the counts from the Al target while suppressing the effect of background. It also must be able to "smooth" out the statistical fluctuations that are present in the data in a way that is consistent with simple assumptions about the characteristics of the position-resolution function. A simplifying circumstance is that the background decays appear in the data to be distributed as a constant over the entire field of view of the reconstruction.

The first step is a fit to the data using Gaussian distributions to describe decays from the thin scintillator and the Al target foil plus a background constant. Since a least-squares fit is known to underestimate the area under a data distribution (by χ^2) in the case of Poisson-distributed measurements [Bev69], another strategy was adopted. The least-squares method derives from the more general maximum-likelihood method under the assumption of Gaussian distributed measurements for each data point. The likelihood function is a quantitative estimate of the probability that the observed data actually derive from some assumed probability distribution. More accurate for our situation is the assumption that the measurements originate from Poisson distributions around some mean for each channel in the y-position histogram. It is now possible to construct a fit algorithm by inserting the assumption of Poisson statistics into the likelihood function [Awa79]. For this case, the likelihood function is

$$\mathcal{L} = \prod_i \left(\frac{f(y_i)^{d_i} e^{-f(y_i)}}{d_i!} \right) , \tag{F.1}$$

where f is the fitting function and d_i are the data values. Since $\mathcal{L}$ is to be maximized by adjusting the parameters of the fit function, one may perform the equivalent but computationally more convenient task of minimizing the negative of the logarithm of $\mathcal{L}$:

$$-\ln \mathcal{L} = \sum_i \left[-d_i \ln f(y_i) + f(y_i) + \ln(d_i!) \right] . \tag{F.2}$$

In practice, one may omit the constant due to

$$\sum_i \ln(d_i!) \tag{F.3}$$

and minimize

$$G = \sum_i \left[f(y_i) - d_i \ln f(y_i) \right] . \tag{F.4}$$

It can be shown that, for Poisson distributed data, this method neither under- nor overestimates the area under a data distribution [Awa79]. This point is critical to this application, as the background must not be underestimated. If it were, one would get unreasonably long tails on the resolution distribution. Overestimating the background would suppress legitimate portions of the tails. The drawback of this procedure compared to the least-squares minimization is the loss of an absolute measure of the quality of the fit.

Next, an array of numbers is initialized to the fit values from the target Gaussian, as determined from the above method. This array is now to be treated in a fashion that will cause it to deviate from the Gaussian but approach the actual resolution distribution as judged by the background-subtracted *Al* target data. There are two basic assumptions made about the shape of the resolution distribution:

1. The resolution function must be symmetric about its peak.
2. The resolution function must fall off monotonically to each side of the peak, asymptotically approaching zero.

Under these constraints, one performs what might be termed a ***relaxation fit*** of the initial guess at the resolution distribution to the background-subtracted Al target data. This procedure does not consider a functional form for the fit. Rather, it tries to adjust an array of numbers to the best fit to a data set by adding or subtracting counts from each channel in the distribution; only the simple constraints listed prevent convergence to the "perfect fit." The quality of the "fit" is judged here by minimizing its square deviation from the data. This choice will be justified below.

This method proceeds as follows:

1. Choose a channel for which an alteration of the current estimate of the resolution distribution is being considered.
2. Evaluate the χ^2 between the current guess at the resolution distribution contained in the array reserved for it and the Al target distribution from the data, taking care to properly center the resolution function on the Al target peak.
3. Add a count on a trial basis to the channel being considered.
4. Reevaluate χ^2.
5. If χ^2 improves, accept the added count and return to step 1.
6. Subtract a count from this channel on a trial basis.
7. Again, evaluate χ^2.
8. If χ^2 has improved, accept the subtraction and return to step 1.
9. Select the next channel, looping over the entire relevant range of the data distribution.

When there is no change to the resolution distribution over one full loop of attempted additions and attempted subtractions, the current distribution is considered to be the final fit to the resolution. The constraints listed above are incorporated by actually adding or subtracting trial counts *symmetrically* from the resolution distribution and by not allowing a count addition or subtraction if it violates the monotonicity requirement, even if the χ^2 were to improve.

The sanity of the choice of χ^2 as the relevant estimator of fit quality must be proven. To this end, we define χ^2 by

$$\left(\chi^2\right)^{(n)} = \sum_i w_i \left(F_i^{(n)} - D_i\right)^2 , \tag{F.5}$$

where n labels the iteration number of the procedure outlined above, w_i is the weight with which channel i enters into χ^2, $F_i^{(n)}$ is the fit to the resolution distribution in channel i after n passes, and D_i is the background-subtracted contents of channel i in the data distribution. Next, the procedure for adding or subtracting a count in channel j is given by

$$\begin{aligned} F_j^{(n+1)} &= F_j^{(n)} \pm 1 \\ F_{-j}^{(n+1)} &= F_j^{(n)} \pm 1 \\ F_{i \neq \pm j}^{(n+1)} &= F_{i \neq \pm j}^{(n)} , \end{aligned} \tag{F.6}$$

where the symmetry constraint has been explicitly built in. Inserting the prescription of Eq. F.6 into Eq. F.5 gives χ^2 at iteration $n+1$ as

$$\begin{aligned} \left(\chi^2\right)^{(n+1)} &= \sum_i w_i \left(F_i^{(n+1)} - D_i\right)^2 \\ &= \left(\chi^2\right)^{(n)} + \Delta \left(\chi^2\right)^{(n)} , \end{aligned} \tag{F.7}$$

where

$$\Delta \left(\chi^2\right)^{(n)} = w_j \left[1 \pm 2 \left(F_j^{(n)} - D_j\right)\right] + w_{-j} \left[1 \pm 2 \left(F_{-j}^{(n)} - D_{-j}\right)\right] . \tag{F.8}$$

Using the symmetry requirements

$$F_{-j}^{(n)} = F_j^{(n)} \tag{F.9}$$

and

$$w_{-j} = w_j \tag{F.10}$$

gives

$$\Delta \left(\chi^2\right)^{(n)} = 2 w_j \left\{1 \pm \left[2 F_j^{(n)} - \left(D_j + D_{-j}\right)\right]\right\} . \tag{F.11}$$

Here, the sign choice is still "+" for adding a count and "−" for subtracting a count. In order for this iteration to give an improvement in the fit, we require

$$\Delta\left(\chi^2\right)^{(n)} < 0 \; . \tag{F.12}$$

With the Eq. F.11 this gives

$$F_j^{(n)} < \left(\frac{D_j + D_{-j}}{2}\right) - \frac{1}{2} \tag{F.13}$$

for the condition that must be fulfilled if a count is to be added to channel j and

$$F_j^{(n)} > \left(\frac{D_j + D_{-j}}{2}\right) + \frac{1}{2} \tag{F.14}$$

as the requirement to be satisfied before subtracting a count from channel j. One notes that the weights of the individual channels (as long as they are non-vanishing and symmetric) have cancelled out; they have no effect on the procedure to obtain the best resolution distribution. They only determine the absolute scale of the values of χ^2, which for this procedure has not been shown to have any deeper meaning.

The criteria obtained for the adding or subtracting of a count from a given channel in the fit to the resolution distribution are what one may have written down intuitively and the rationality of this result lends credibility to the form of χ^2 given. In practice, the fit was done not using the relational criteria for individual channels, but calculating the full χ^2 at every step. This made the algorithm slightly more flexible, for example allowing the range of the fit to be asymmetric about the peak. Since the data distribution contains counts not just from the *Al* target, but also from the beam scintillator, the range over which χ^2 was evaluated had to exclude the scintillator, as counts from it were not described by a fit with the single resolution distribution positioned at the target peak.

This procedure will work with any number of counts in the *Al* target data distribution, but the results certainly improve with statistics. Since these are finite, though, alterations to the resolution function on a scale finer than a single count would not be well-motivated by the level of confidence in the number of counts in any given channel. The data used for the relaxation fit are shown in Fig. F.1. The resulting resolution distribution, after normalization to unit area, is shown in Fig. F.2.

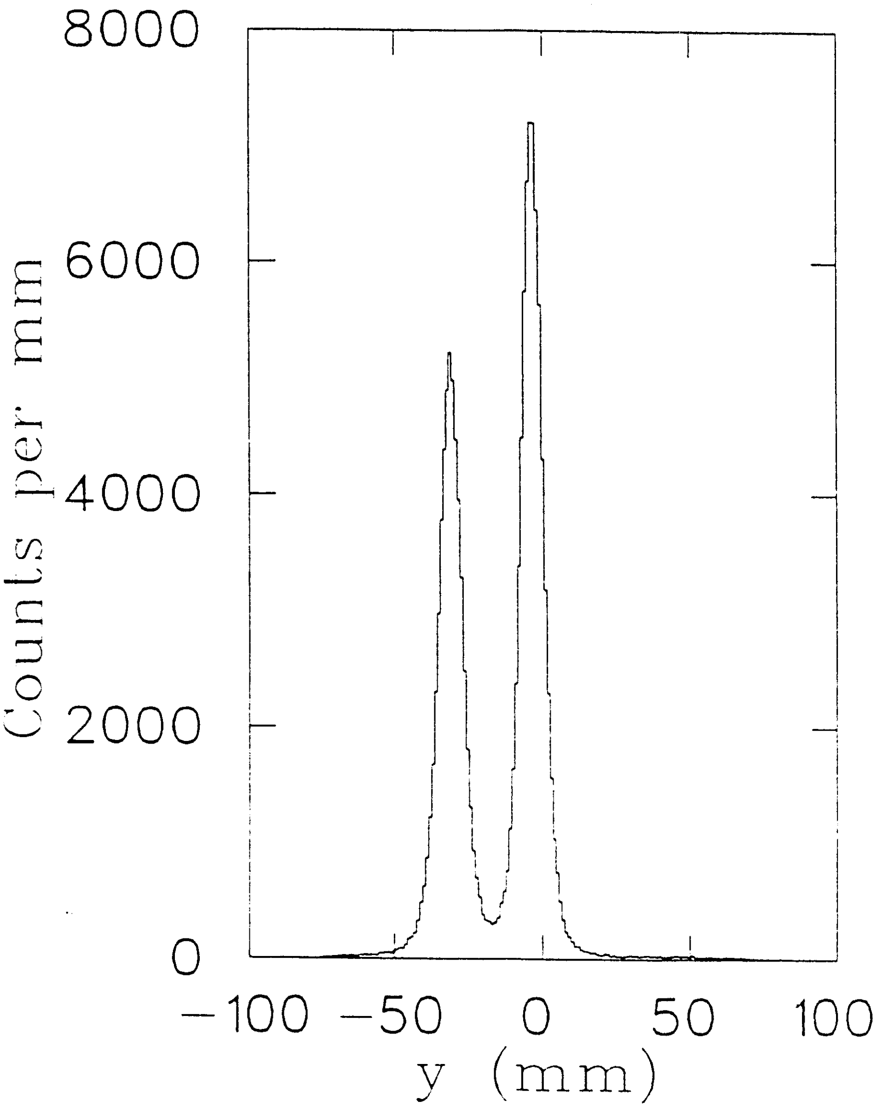


Figure F.1: Distribution of decay positions projected onto the y-axis for the Al target. Decays from the beam scintillator are also apparent to the left of the target.

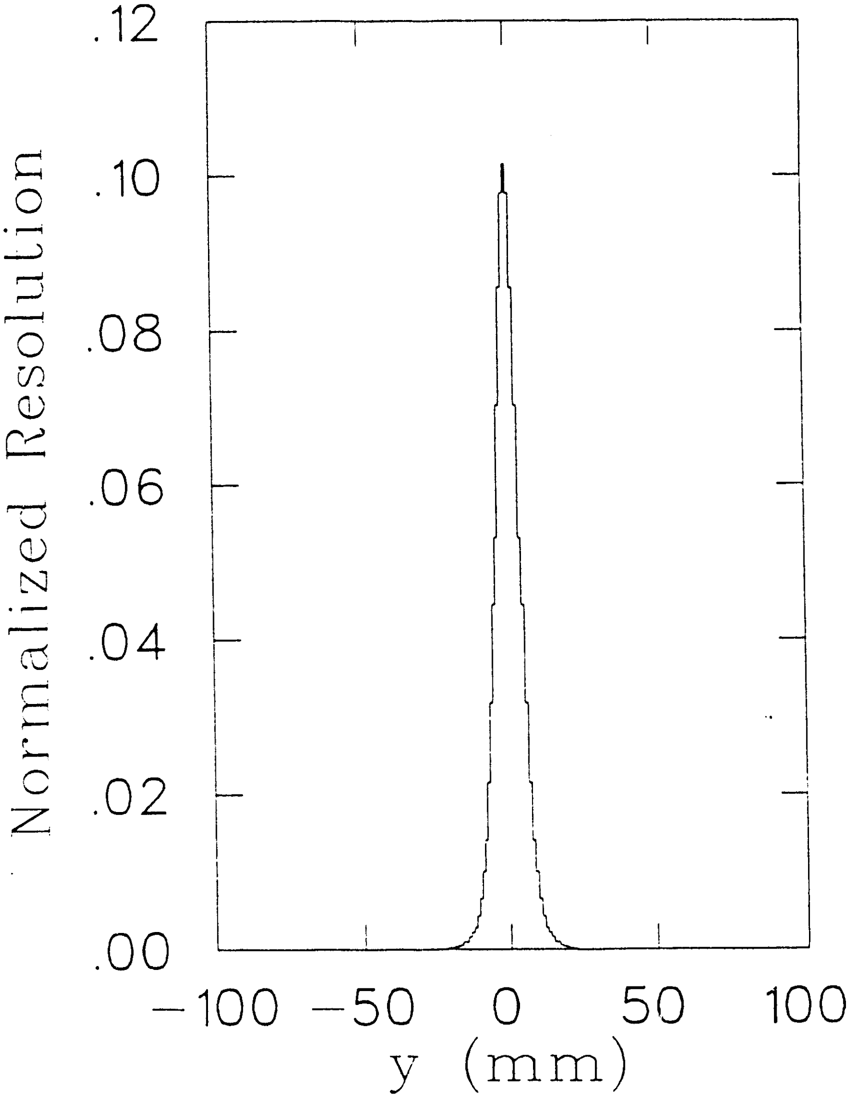


Figure F.2: Resolution function of the reconstruction of the decay origin.

Appendix G

Beam-Correlated Rates

Often it is desirable to know what fraction of a measured detector rate is correlated to the incident beam and what fraction is due to beam-uncorrelated background processes. This is possible when the accelerator has a pulsed beam structure, so that some scalers counting the rates of interest may be gated to count only during the beam pulses. Comparison of these counts with those from ungated scalers allow one to derive the information sought.

Specifically, we recorded counts from detectors and from a clock (10 kHz) once with the condition of "computer live in the beam gate" ($BG \bullet \overline{CB}$) and once with the condition of "computer live" ($\overline{CB}$). *Some* of these counts (for example, the μ-counter linear-Σ counts, the clock (CLK), and the beam-gated clock ($CLK \bullet BG$)) were also recorded without live-time gating.

What we would like to extract from what we have is the best possible estimate of the true raw detector rates averaged over the beam cycle as well as in the beam-gate. Also, often more intuitively useful rates are the beam-correlated and the beam-uncorrelated rates for a detector, again both undistorted for any data acquisition dead time.

The basic complication is that the CB signal was clearly correlated to the BG, because we triggered the MWPCs on hopefully beam-correlated events of interest occurring during the beam gate. This was the "Event 8" trigger and was responsible for most of the overall dead time and for all of the in-beam-gate dead time (Event

4 (collector HV supply) and Event 6 (CAMAC scalers) were forced between beam gates). Thus, the $\overline{CB}$ rates are not a good measure of truly ungated rates, because of the unequal dead times in- and outside of the BG.

To surmount this difficulty, start by defining:

$$\begin{aligned}
CLK &= \text{ungated clock (10 kHz)} \\
CLK \bullet BG &= \text{beam gated clock} \\
CLK \bullet \overline{CB} &= \text{computer live clock} \\
CLK \bullet BG \bullet \overline{CB} &= \text{computer live clock in the beam gate} \\
L &= \frac{(CLK \bullet \overline{CB})}{CLK} = \text{overall live time fraction} \\
L_{BG} &= \frac{(CLK \bullet BG \bullet \overline{CB})}{CLK \bullet BG} = \text{live time fraction in the beam gate} \\
N &= \text{raw detector counts during counting interval} \\
N \bullet BG &= \text{detector counts in the beam gate} \\
N \bullet \overline{CB} &= \text{detector counts while computer live} \\
N \bullet BG \bullet \overline{CB} &= \text{detector counts in beam gate while computer live} \\
N_{corr} &= \text{beam correlated detector counts} \\
N_{uncorr} &= \text{beam uncorrelated detector counts .}
\end{aligned} \tag{G.1}$$

The easiest way to argue for expressions relating the differently gated counts seems to be to construct the measured counts from the basic beam-correlated and beam-uncorrelated rates and the appropriate live times:

$$N \bullet \overline{CB} = N_{corr} \times L_{BG} + N_{uncorr} \times L \tag{G.2}$$

and

$$N \bullet BG \bullet \overline{CB} = N_{corr} \times L_{BG} + N_{uncorr} \times \delta \times L_{BG} \ , \tag{G.3}$$

where δ is the duty factor of the beam cycle. In words, the computer-live counts, $N \bullet \overline{CB}$, are the sum of the beam-correlated counts (which must occur in the beam-gate and thus have the beam-gated live time applied) and the beam-uncorrelated counts (which may occur anytime during the beam cycle and thus have the overall live time as a factor). The *beam-gated* computer-live counts are the sum of the same beam-correlated counts as above but only with that fraction of the beam-uncorrelated counts that happens to fall in the beam-gate (and is thus also subject to the beam-gated live time).

Solving the above equations for the beam-correlated and beam-uncorrelated counts gives

$$N_{corr} = \frac{N_{BG \bullet \overline{CB}}\, L - N_{\overline{CB}}\, \delta\, L_{BG}}{L_{BG}\,(L - \delta\, L_{BG})} \tag{G.4}$$

and

$$N_{uncorr} = \frac{N_{\overline{CB}} - N_{BG \bullet \overline{CB}}}{(L - \delta\, L_{BG})} \ . \tag{G.5}$$

From these we may also infer the raw detector counts and the raw detector counts in the beam gate:

$$N = N_{corr} + N_{uncorr} \tag{G.6}$$

and

$$N_{BG} = N_{corr} + \delta \times N_{uncorr} \ . \tag{G.7}$$

At this point one may choose to refer the counts to a clock of chosen gating. The online CAMAC scaler output typically used $CLK \bullet \overline{CB}$ as the normalization. Here let's define detector rates as:

$$R = \frac{N}{CLK} \ , \tag{G.8}$$

$$R_{BG} = \frac{N_{BG}}{CLK} \ , \tag{G.9}$$

$$R_{\overline{CB}} = \frac{N_{\overline{CB}}}{CLK_{\overline{CB}}} \ , \tag{G.10}$$

$$R_{BG \bullet \overline{CB}} = \frac{N_{BG \bullet \overline{CB}}}{CLK_{\overline{CB}}} \ , \tag{G.11}$$

$$R_{corr} = \frac{N_{corr}}{CLK} \ , \tag{G.12}$$

and

$$R_{uncorr} = \frac{N_{uncorr}}{CLK} \ . \tag{G.13}$$

Thus, solving for R_{corr} and R_{uncorr} in terms of the rates found in the online end-of-run-summaries, we have

$$R_{corr} = \frac{R_{BG\bullet\overline{CB}} - R_{\overline{CB}}\,\delta\,\frac{L_{BG}}{L}}{\frac{L_{BG}}{L}\left(1 - \delta\,\frac{L_{BG}}{L}\right)} \tag{G.14}$$

and

$$R_{uncorr} = \frac{R_{\overline{CB}} - R_{BG\bullet\overline{CB}}}{\left(1 - \delta\,\frac{L_{BG}}{L}\right)} \; . \tag{G.15}$$

These combine to give

$$R = \frac{L}{(L - \delta L_{BG})}\left[R_{\overline{CB}}(1 - \delta) + R_{BG\bullet\overline{CB}}\left(\frac{L}{L_{BG}} - 1\right)\right] \tag{G.16}$$

and

$$R_{BG} = R_{BG\bullet\overline{CB}}\,\frac{L}{L_{BG}} \; . \tag{G.17}$$

These expressions have been checked against a few cases where NIM scaler information and/or ungated CAMAC scaler information are available and were found to be in good agreement with both.

One can imagine taking approximate forms of these expressions that are valid for certain limiting conditions, such as live times of 100%. This provides a way to check the formulae obtained. If $L \rightarrow 1$ and $L_{BG} \rightarrow 1$ then $R \rightarrow R_{\overline{CB}}$ and $R_{BG} \rightarrow R_{BG\bullet\overline{CB}}$. This, of course, reads as an identity if $\overline{CB}$ is always true.

Other clock normalizations may be chosen to get, for example, instantaneous rates. Factors can be adjusted to refer back to the rates as given in the CAMAC scaler printouts or to the raw counts that are available on tape.

Bibliography

[Ahn90] H. Ahn *et al.* 1990. in preparation.

[Ahn91] H. E. Ahn. PhD thesis, Yale University, 1991. in preparation.

[Ama68a] J. J. Amato *et al.* Phys. Rev. Lett., **21**, 1709 (1968).

[Ama68b] James J. Amato, *A Search for Muonium-Antimuonium Conversion.* PhD thesis, Yale University, 1968.

[And36] C. D. Anderson and S. H. Neddermeyer. Phys. Rev., **50**, 263 (1936).

[Awa79] T. Awaya. Nucl. Instrum. Methods, **165**, 317 (1979).

[Bad84] A. Badertscher *et al.* Phys. Rev. Lett., **52**, 914 (1984).

[Bad85] A. Badertscher *et al.* Nucl. Instrum. Methods, **A238**, 200 (1985).

[Bar69] W. C. Barber *et al.* Phys. Rev. Lett., **22**, 902 (1969).

[Bee86] G. A. Beer *et al.* Phys. Rev. Lett., **57**, 671 (1986).

[Beg74] M. A. B. Bég and A. Sirlin. Ann. Rev. Nucl. Sci., **24**, 379 (1974).

[Ber77] H. G. Berry. Rep. Prog. Phys., **40**, 12 (1977).

[Bet53] H. A. Bethe. Phys. Rev., **89**, 1256 (1953).

[Bev69] Philip R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences.* (McGraw-Hill, New York, 1969).

[Bjo77] J. D. Bjorken and S. Weinberg. Phys. Rev. Lett., **38**, 622 (1977).

[Bol81] P. R. Bolton *et al.* Phys. Rev. Lett., **47**, 1441 (1981).

[Bol82] Paul R. Bolton, *Production of Muonium in Vacuum.* PhD thesis, Yale University, 1982.

[Bol88] R. D. Bolton *et al.* Phys. Rev., **D38**, 2077 (1988).

[Cab60] N. Cabbibo and R. Gatto. Phys. Rev. Lett., **5**, 114 (1960).

[Cab61] N. Cabbibo and R. Gatto. Nuovo Cim., **19**, 612 (1961).

[Cab88] . *Cab-O-Sil Properties and Functions.* Cabot Corporation, Cab-O-Sil Division, Tuscola, IL 61953, 1988. We are grateful for samples of fumed silica powder of grades M-5, EH5, and PTG.

[Can82] K. F. Canter and Jr. A. P. Mills. Can. J. Phys., **60**, 551 (1982).

[Cas77] D. E. Casperson *et al.* Phys. Rev. Lett., **38**, 956 (1977).

[Cha89] D. Chang and W.-Y. Keung. Phys. Rev. Lett., **62**, 2583 (1989).

[Che77a] T.-P. Cheng and L.-F. Li. Phys. Rev. Lett., **38**, 381 (1977).

[Che77b] T.-P. Cheng and L.-F. Li. Phys. Rev., **D16**, 1425 (1977).

[Che85] D. M. Chen *et al.* Phys. Rev., **B31**, 4123 (1985).

[Chu88] S. Chu *et al.* Phys. Rev. Lett., **60**, 101 (1988).

[Cle64] W. E. Cleland *et al.* Phys. Rev. Lett., **13**, 202 (1964).

[Dal80] J. M. Dale *et al.* Surf. Interface Anal., **2**(6), 199 (1980).

[Dan62] G. Danby *et al.* Phys. Rev. Lett., **9**, 36 (1962).

[Der79] E. Derman. Phys. Rev., **D19**, 317 (1979).

[Don83] J. B. Donahue, Radioactive gas. LAMPF internal memo, September 1983.

[dV70] R. de Voe *et al.* Phys. Rev. Lett., **25**, 1779 (1970).

[Fav71] D. Favart *et al.* Phys. Rev. Lett., **27**, 1336 (1971).

[Fei58] G. Feinberg. Phys. Rev., **110**, 1482 (1958).

[Fei59] G. Feinberg *et al.* Phys. Rev. Lett., **3**, 527 (1959).

[Fei61a] G. Feinberg and S. Weinberg. Phys. Rev. Lett., **6**, 381 (1961).

[Fei61b] G. Feinberg and S. Weinberg. Phys. Rev., **123**, 1439 (1961).

[Fey58] R. P. Feynman and M. Gell-Mann. Phys. Rev., **109**, 193 (1958).

[Gel81] G. B. Gelmini and M. Roncadelli. Phys. Lett., **99B**, 411 (1981).

[Gla61] S. L. Glashow. Nucl. Phys., **22**, 529 (1961).

[Gla70] S. L. Glashow *et al.* Phys. Rev., **D2**, 1285 (1970).

[Hal82] A. Halprin. Phys. Rev. Lett., **48**, 1313 (1982).

[Her] P. Herczeg and R. N. Mohapatra. to be published.

[Her89] P. Herczeg. in D. Bryman *et al.*, eds., *Rare Decay Symposium*, p. 24, Vancouver, 1988, (World Scientific, Singapore, 1989).

[Her91] P. Herczeg. 1991. private communication.

[Hub88] T. M. Huber *et al.* Phys. Rev. Lett., **61**, 2189 (1988).

[Hub90] T. M. Huber *et al.* Phys. Rev., **D41**, 2709 (1990).

[Hug60] V. W. Hughes *et al.* Phys. Rev. Lett., **5**, 63 (1960).

[Hug70] V. W. Hughes *et al.* Phys. Rev., **A1**, 595 (1970).

[Hug72] E.B. Hughes *et al.* IEEE Trans. Nucl. Sci., **NS-19**(3), 126 (1972).

[Jan90] A. C. Janissen *et al.* Phys. Rev., **A42**, 161 (1990).

[Jun89] Klaus Jungmann *et al.*, Search for Spontaneous Conversion of Muonium to Antimuonium. PSI Proposal No. R 89-06.1, 1989. K. Jungmann and W. Bertl, Spokesmen.

[Ket90a] S. H. Kettell *et al.* 1990. in preparation.

[Ket90b] Steven H. Kettell, *Measurement of the* $2^2S_{1/2}$-$2^2P_{3/2}$ *Fine Structure Interval in Muonium.* PhD thesis, Yale University, 1990.

[Kie79] R. F. Kiefl *et al.* Hyperfine Int., **6**, 185 (1979).

[Kob73] M. Kobayashi *et al.* Prog. Theor. Phys., **49**, 652 (1973).

[Kon53] E. J. Konopinski and A. H. Mahmoud. Phys. Rev., **92**, 1045 (1953).

[Kua87] Y. Kuang *et al.* Phys. Rev., **A35**, 3172 (1987).

[Kua89] Y. Kuang *et al.* Phys. Rev., **A39**, 6109 (1989).

[Lec69] C. Lechanoine *et al.* Nucl. Instrum. Methods, **69**, 122 (1969).

[Lee77] B. W. Lee and R. E. Shrock. Phys. Rev., **D16**, 1444 (1977).

[Lok55] S. Lokanathan and J. Steinberger. Phys. Rev., **98**, 240 (1955).

[Lon81] Anthony C. Longhitano, *Heavy Higgs Bosons and their Low-Energy Impact.* PhD thesis, Yale University, 1981.

[Lyn77] K. G. Lynn and B. T. A. McKee, *Some Investigations of Moderators for Slow Positron Beams.* Tech. Rep. BNL-24143, Brookhaven National Laboratory, 1977.

[Lyn80] K. G. Lynn and H. Lutz. Rev. Sci. Instrum., **51**(7), 977 (1980).

[Mar77] W. J. Marciano and A. I. Sanda. Phys. Rev. Lett., **38**, 1512 (1977).

[Mar78] G. M. Marshall *et al.* Phys. Lett., **65A**, 351 (1978).

[Mar81] G. M. Marshall. PhD thesis, University of British Columbia, 1981.

[Mar82a] F. G. Mariam *et al.* Phys. Rev. Lett., **49**, 993 (1982).

[Mar82b] G. M Marshall *et al.* Phys. Rev., **D25**, 1174 (1982).

[Mic50] L. Michel. Proc. Phys. Soc., **A63**, 514 (1950).

[Mil86] A. P. Mills *et al.* Phys. Rev. Lett., **56**, 1463 (1986).

[Moh75a] R. N. Mohapatra and J. C. Pati. Phys. Rev., **D11**, 2558 (1975).

[Moh75b] R. N. Mohapatra and J. C. Pati. Phys. Rev., **D11**, 556 (1975).

[Moh80] R. N. Mohapatra and G. Senjanovic. Phys. Rev. Lett., **44**, 912 (1980).

[Moh81a] R. N. Mohapatra and G. Senjanovic. Phys. Rev., **D23**, 165 (1981).

[Moh81b] R. N. Mohapatra and J. D. Vergados. Phys. Rev. Lett., **47**, 1713 (1981).

[Mor66] David L. Morgan, *Atomic Processes Involving Muonium and Antimuonium.* PhD thesis, Yale University, 1966.

[Mor70] D. L. Morgan and V. W. Hughes. Phys. Rev., **D2**, 1389 (1970).

[Mor73] D. L. Morgan and V. W. Hughes. Phys. Rev., **A7**, 1811 (1973).

[Ned37] S. H. Neddermeyer and C. D. Anderson. Phys. Rev., **51**, 884 (1937).

[Ni87] B. Ni *et al.* Phys. Rev. Lett., **59**, 2716 (1987).

[Ni88a] B. Ni *et al.* Nucl. Phys., **A478**, 757c (1988).

[Ni88b] Benwen Ni, *Search for Spontaneous Conversion of Muonium to Antimuonium.* PhD thesis, Yale University, 1988.

[Ora84] C. J. Oram *et al.* Phys. Rev. Lett., **52**, 910 (1984).

[Pat74] J. C. Pati and A. Salam. Phys. Rev., **D10**, 275 (1974).

[Pif76] A. E. Pifer *et al.* Nucl. Instrum. Methods, **135**, 39 (1976).

[Pon57] B. Pontecorvo. Zh. Eksp. Teor. Fiz., **33**, 549 (1957). [Sov. Phys. JETP **6**, 429 (1958)].

[Pon60] B. Pontecorvo. Sov. Phys. JETP, **10**, 1236 (1960).

[Pre61] R. Prepost *et al.* Phys. Rev. Lett., **6**, 19 (1961).

[Rei78] H.-W. Reist *et al.* Nucl. Instrum. Methods, **153**, 61 (1978).

[Sal68] A. Salam. in N. Svartholm, ed., *Elementary Particle Theory: Relativistic Groups and Analyticity*, p. 367, Nobel Symposium No. 8, (Almqvist & Wiksell, Stockholm, 1968).

[Sal72] A. Salam and J. Strathdee. Nuovo Cimento, **11**, 397 (1972).

[San85] V. D. Sandberg *et al.* Nucl. Instrum. Methods, **A234**, 512 (1985).

[Sch57] J. Schwinger. Ann. Phys., **2**, 407 (1957).

[Sch87] Reiner Schaefer *et al.*, Measurement of Muonium to Antimuonium Conversion with Improved Sensitivity. LAMPF Proposal No. 1073, 1987. Reiner Schaefer and Vernon Hughes, Spokesmen.

[Sch88] Willi Schäfer, *Myonium im Vakuum zum Studium der Myonium - Antimyonium - Konversion.* PhD thesis, Universität Heidelberg, 1988.

[Shl74] Sally Shlaer, *An MBD Primer.* Tech. Rep. LA-5511-MS, LASL, 1974.

[Sta74] R. D. Stambaugh *et al.* Phys. Rev. Lett., **33**, 568 (1974).

[Swa89] M. L. Swartz. Phys. Rev., **D40**, 1521 (1989).

[Tho69] P. A. Thompson *et al.* Phys. Rev. Lett., **22**, 163 (1969).

[Tho79] P. A. Thompson *et al.* Nucl. Instrum. Methods, **161**, 391 (1979).

[Tro66] W. Peter Trower, *High Energy Particle Data IV Range-Momentum and dP/dx Plots of Charged Particles in Matter.* Tech. Rep. UCRL-2426, Lawrence Radiation Laboratory, 1966.

[Ver83] A. Verhanen *et al.* Appl. Phys., **A32**, 163 (1983).

[Ver86] J. D. Vergados. Phys. Rep., **133**, 1 (1986). This reference reviews several standard model extensions.

[Wei67] S. Weinberg. Phys. Rev. Lett., **19**, 1264 (1967).

[Wei78] Mitchel Weissbluth, *Atoms and Molecules.* (Academic Press, Inc., New York, 1978).

[Wil77] F. Wilczek and A. Zee. Phys. Rev. Lett., **38, 531 (1977).**

[Wil80] S. E. Willis *et al.* Phys. Rev. Lett., **44**, 522 (1980).

[Win74] H. Wind. Nucl. Instrum. Methods, **115**, 431 (1974).

[Woo88] K. A. Woodle *et al.* Z. Phys., D9, 59 (1988).

[Woo90] K. A. Woodle *et al.* Phys. Rev., **A41**, 93 (1990).

[Yos88] G. P. Yost *et al.*, Review of Particle Properties. Phys. Lett., **B204**, 1 (1988).

[Zio62] K. Ziock *et al.* Phys. Rev. Lett., **8**, 103 (1962).

END

DATE FILMED

12/03/91