Dissertations, Theses, and Masters Projects   Theses, Dissertations, & Master Projects

1982

A precision measurement of the lifetime of the positive muon

Kevin Louis Giovanetti

College of William & Mary - Arts & Sciences

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

8229791

Giovanetti, Kevin Louis

A PRECISION MEASUREMENT OF THE LIFETIME OF THE POSITIVE MUON

The College of William and Mary in Virginia   .   PH.D.   1982

University
Microfilms
# International   300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __√__ .

1. Glossy photographs or pages _____

2. Colored illustrations, paper or print _____

3. Photographs with dark background _____

4. Illustrations are poor copy _____

5. Pages with black marks, not original copy _____

6. Print shows through as there is text on both sides of page _____

7. Indistinct, broken or small print on several pages __✔__

8. Print exceeds margin requirements _____

9. Tightly bound copy with print lost in spine _____

10. Computer printout pages with indistinct print _____

11. Page(s) _____ lacking when material received, and not available from school or author.

12. Page(s) _____ seem to be missing in numbering only as text follows.

13. Two pages numbered _____ . Text follows.

14. Curling and wrinkled pages _____

15. Other_____

University
Microfilms
International
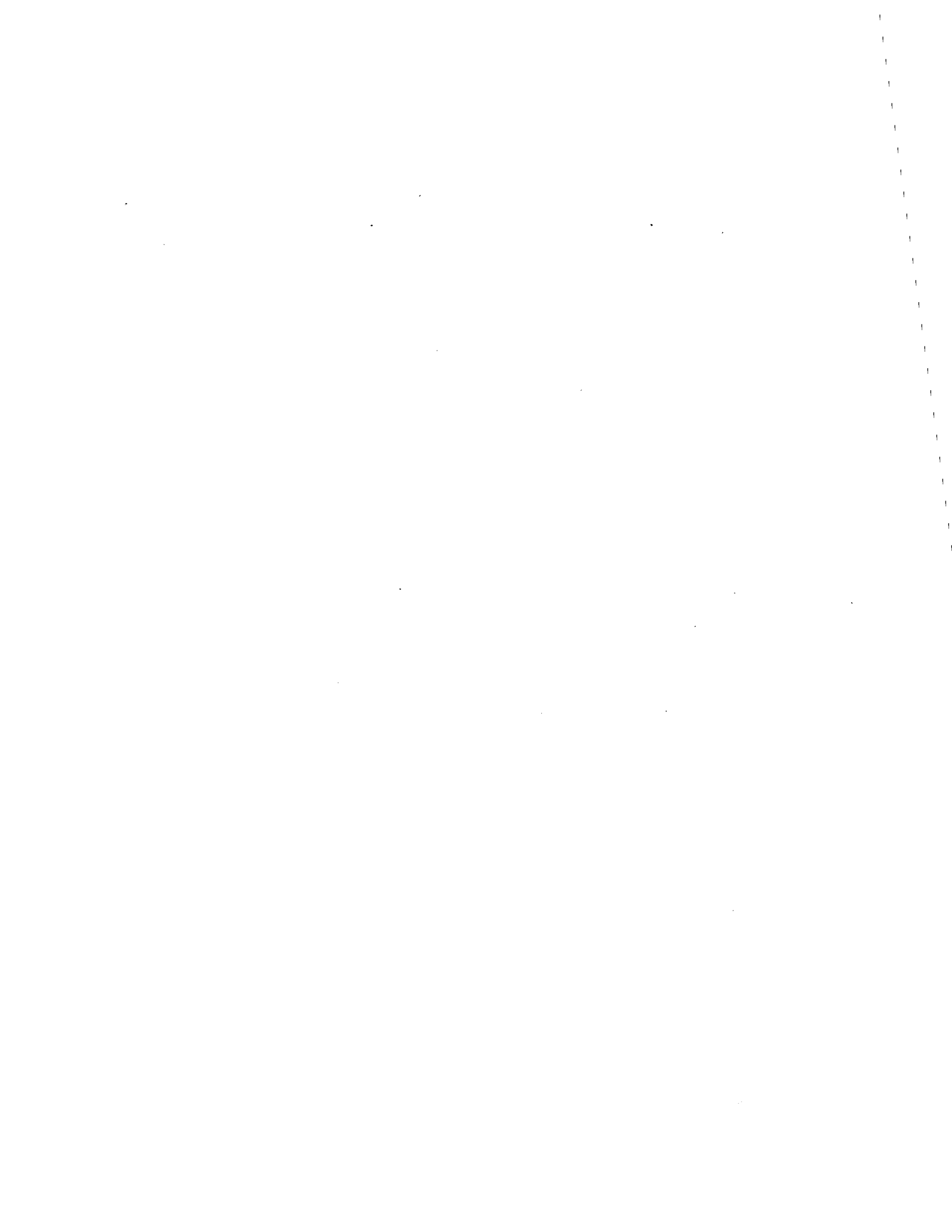
A PRECISION MEASUREMENT OF THE LIFETIME
OF THE POSITIVE MUON

———————————

A Dissertation

Presented to

The Faculty of the Department of Physics

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

———————————

by

Kevin Louis Giovanetti
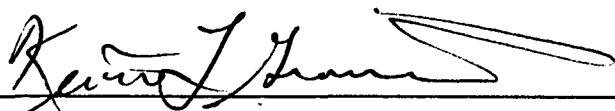
August 1982

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

（signature）

Approved, August 1982

（signature）
Robert E. Welsh

（signature）
Morton Eckhause

（signature）
John R. Kane

（signature）
Robert T. Siegel

（signature）
Rolf G. Winter
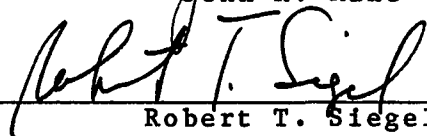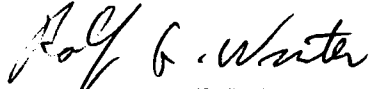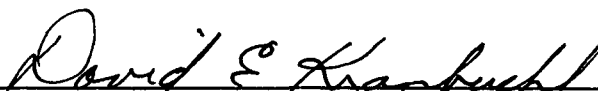
（signature）
David E. Kranbuehl
Department of Chemistry

DEDICATION

To Louis and Elizabeth Giovanetti
my parents

CONTENTS

ACKNOWLEDGMENTS

ABSTRACT

In an experiment performed at TRIUMF Laboratory, the time distribution of muon decay positrons arising from pions stopped in a water Cherenkov counter was measured with a high precision timing system. The pions served as a source of positive muons.

The data were used to determine a new value for the positive muon lifetime. This lifetime 2.19695 $\pm$ .00006 (μsec) is slighty lower than previous measurements but still in fair agreement. The data were tested for the various systematic effects that typically occur in these measurements.

KEVIN LOUIS GIOVANETTI

DEPARTMENT OF PHYSICS

THE COLLEGE OF WILLIAM AND MARY IN VIRGINIA

A PRECISION MEASUREMENT OF THE LIFETIME
OF THE POSITIVE MUON

Chapter I

INTRODUCTION

The first evidence of the weak interaction came from the study of nuclear β-decay. By now, a host of processes including muon decay and muon capture have been observed and are believed to be a manifestation of a universal weak interaction.[1,2] Universality implies that the coupling constants that govern weak processes are the same for all fundamental particles. The most widely accepted view is that weak processes are due to the exchange of massive (>40 GeV) vector bosons associated with the symmetries of nature deemed local gauge invariant. These bosons, $W^{\pm}$ and Z_o, in a manner analogous to the photon in quantum electrodynamics, are exchanged between the initial and final state particles. Since the $W^{\pm}$ and Z_o are so massive, the exchange at low momentum transfer appears to occur at one space-time point and so can be successfully approximated by a local coupling of two currents. This phenomenological model of current-current coupling is known as the local V-A theory.

Leptons, specifically the muon, the electron, the tau, and their associated neutrinos, are believed to be fundamental particles. The decay of the muon into an electron and two neutrinos

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$$

represents a unique opportunity to observe the muon-electron interaction as a dominantly weak process. (Electromagnetic contibutions to the total decay rate are less than 1 %.) Thus, there have been many experiments designed to measure the parameters of this decay.[1] One parameter, the lifetime τ_μ of the positive muon, can be used to determine G_F, the weak coupling constant. The history of the measurement of this parameter has been marked by advances to more precise values (see Table I). The Williams and Williams measurement is in disagreement with later work and has not been included in the latest Particle Data Group[3] value for τ_μ. In addition to G_F, the lifetime of μ^- in vacuum is inferred from τ_μ according to the conservation of CPT. A comparison of the μ^+ lifetime to the μ^- disappearance rate in hydrogen therefore yields the μ^- capture rate in hydrogen; that result can be used to determine g_p, the induced pseudoscalar coupling. Also, comparisons of rates for superallowed β-decay transitions to the free decay rate test the Cabbibo universality hypothesis.[4]

What follows in this report is a detailed account of a measurement of τ_μ obtained by recording the time interval between the arrival of π^+ or μ^+ and the subsequent decay positron. (Positive pions readily serve as a source of μ^+ <u>via</u>

TABLE I

Lifetime Determinations

Experiment	Reference	$\tau_{\mu}\pm$ (μsec)
Farley et al. (1962)	5	2.198 ± 0.001
Lundy (1962)	6	2.203 ± 0.002
Eckhause et al. (1963)	7	2.202 ± 0.003
Meyer et al. (1963)	8	2.197 ± 0.002
Williams & Williams (1972)	9	2.20026 ± 0.00081
Duclos et al. (1973)	10	2.1973 ± 0.0003
Balandin et al. (1974)	11	2.19711 ± 0.00008
Bardin et al. (1980)	12	2.19718 ± 0.0001

their primary decay $\pi^{+}\rightarrow\mu^{+}+\nu_{\mu}$.) These measured intervals are governed by a set of parameters. Those parameters characterize the physical processes that produce these intervals and the experimental strategy chosen to measure them. The limit in information on a parameter in an experiment such as this is due to the statistical nature of the physical process or to the systematic errors in the measurement technique. The statistical uncertainty can be reduced by an increase in the number of observations. In order to reach a fractional error of 10^{-5} at least 10^{+10} events need to be recorded. Therefore an efficient method which is free of systematic effects is essential.

Chapter II outlines the theory of the weak interaction while Chapter III provides a description of the experimental details of our measurement. Chapter IV covers the analysis involved in extracting τ_{μ} from the data and testing for

systematic errors. The results are presented and discussed in Chapter V. Two appendices are included. Appendix A describes the mathematical function that governed the distribution of the data. Appendix B covers some of the important aspects of the theory of statistical analysis and its application in this experiment.

Chapter II

THEORY

The weak interaction involving hadrons and leptons can be described at low energies by direct current-current coupling. The Lagrangian describing the interaction can be written

$$L_{weak} = \frac{G}{\sqrt{2}}(J^{\mu}J_{\mu} + h.c.), \text{ for which}$$

$$J = J_{lepton} + J_{hadron} \text{ and}$$

$$J^{\mu}_{lepton} = \bar{e}\gamma^{\mu}\frac{1}{2}(1-\gamma_5)\nu_e + muon + \ldots$$

where e is the electron-positron field and ν_e is the accompanying neutrino field. The lepton current can be enlarged to include all the known leptons by adding the appropriate particle-neutrino currents of similar form. The hadronic current in terms of quarks is similiar but is more complicated to calculate with because of the composite nature of the hadrons and the strong interaction effects. The weak current is composed of two parts with both vector and axial vector contributions resulting in the familiar V-A interaction. Although this interaction has been known for many years and describes most of the observed behavior of weak processes, it has only recently been understood on more fundamental grounds. The advent of successful (i.e. renor-

malizable) theories has provided a formalism with which to construct calculable particle theories.

A brief description of the highlights of the theories will follow. A more complete description can be found in the review articles by Abers and Lee,[13] Weinberg,[14] Taylor,[15] and Iliopoulos[16], among others. The premise upon which gauge theories are built is that global symmetries of nature which result in conserved currents and charges can be extended so that the theory is invariant not only globally but locally. This means that the group of transformations which previously left the equation of motion invariant is also expected to leave the equation of motion invariant even if applied in varying degree at different points of space-time. The parameters of the group are now a function of the Lorentz space-time four vector. The result of such an assumption is the need for gauge fields with definite group transformations to preserve the invariance of the Lagrangian. The quanta of these Yang-Mills vector fields associated with some exact local gauge invariance become mediators of particle interactions.

An initial difficulty in theories of this nature was the lack of mass terms in the Lagrangian. These terms could not appear because they were not invariants. The problem was surmounted by a technique known as spontaneous symmetry breaking. Here unphysical particles (Higgs particles) are

added to the theory in a special way so that the lowest energy state is degenerate. The Lagrangian can then be transformed by a particular choice of gauge so that the degrees of freedom for the Higgs particles are absorbed into the gauge fields, which then acquire the necessary masses. Other particle masses are generated by Yukawa couplings between the particles and the Higgs fields. The result is both a "physical" Lagrangian in which particle fields and interaction terms can be identified, and a Lagrangian which can be shown to be renormalizable.

Weinberg[17] and Salam[18] produced a successful unified weak and electromagnetic theory using the group SU(2) X U(1) as the invariant gauge group. Table II illustrates the results of such a prescription for electrons and electron neutrinos. The initial arrangement of fields is done to show the transformation properties under SU(2) X U(1). The derivatives transform as

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + 0^a T_a,$$

where 0^a are parameters and T_a are generators of the gauge group, in order to maintain local gauge invariance. This transformation or redefinition of the derivative operator produces many couplings between the fields. The complete interaction Lagrangian is discussed in detail in Weinberg[19]. The Higgs fields can be redefined so that the η_0, a scalar neutral spinless Higgs particle remains and the other three

Higgs degrees of freedom are absorbed into vector boson fields. After some reorganization the Lagrangian can be written in terms of "physical" particles and the five initial parameters combine in order to produce some of the more common parameters which characterize the weak and electromagnetic processes. A complete description can be found in the work of Taylor.[15]

Other the leptons can be included by adding the appropriate left-handed doublets of leptons and their neutrinos and right-handed singlets. The masses for these leptons can be set with a choice of Higgs coupling.

The correct addition of the hadrons into this theory is not obvious. One simple approach is to build left-handed doublets and right-handed singlets as follows

$$\begin{pmatrix} u \\ d' \end{pmatrix}_1 \quad \begin{pmatrix} c \\ s' \end{pmatrix}_1 \quad u_r \quad d'_r \quad c_r \quad s'_r$$

where u, d, c, and s are up, down, charmed and strange quarks and

$$d' = d \cos\theta_c + s \sin\theta_c,$$
$$s' = -d \sin\theta_c + s \cos\theta_c, \quad \text{and}$$
$$\theta_c = \text{Cabbibo angle}.$$

Semi-leptonic processes like β-decay, muon capture, pion decay, etc. have been examined, but the lack of a strong interaction theory that is easy to calculate with has hindered these computations. Sirlin,[4] using current algebra techniques along with some assumptions about the strong force, ex-

TABLE II

Summary of Weinberg-Salam Model

LAGRANGIAN

$$L=L_W+L_B+L_L+L_H+L_{HL}-V+ \quad (L_G)$$

TERM	PARTICLE	TYPE	FORM
L_W, L_B	Boson	Vector	$1/4\ F^{\mu\nu}F_{\mu\nu}$
L_L	Leptons	Dirac Spinor	$\bar{e}\gamma^\mu\partial_\mu e$
L_H	Higgs	Scalar	$\lvert\partial_\mu\phi\rvert^2$
L_{HL}	Higgs-Lepton Coupling		$-\sqrt{2}M_e/f\,(\bar{e}_R(\phi^+ l_e)+(\bar{l}_e\phi)e_R)$
V	Symmetry Breaking		$\lambda(\phi^+\phi-(1/2)f^2)^2$
L_G	Gauge Fixing		determines propagators

SYMMETRY

SU(2)xU(1)   SU(2)---Weak Isospin=I
U(1)--Weak Hypercharge=Y

FIELDS (INITIAL L) / **FIELDS** ("PHYSICAL" L)

	I	Y	# STATES		SPIN	# STATES
$l=\binom{\nu}{e}_1$	1/2	-1	2	ν_1	1/2	1
$r=e_r$	0	-2	1	e	1/2	2
$\phi=\binom{\phi^+}{\phi_0}$	1/2	1	4	A_μ	1	2
A_μ^i	1	0	6	$W_\mu^\pm$	1	6
B_μ	0	0	2	Z_μ^0	1	3
$(e_{r,1}=(1\pm\gamma_5)e)$				η_0	0	1

PARAMETERS

INITIAL L		"PHYSICAL" L	
g	SU(2) coupling	$M=2\lambda f$	$M_W=1/2fg$
g´	U(1) coupling	$M_Z=1/2(f(g^2+g'^2)^{1/2}$	
λ,f	Symmetry breaking	$TAN\theta_W=g'/g$	$G_F/\sqrt{2}=g/8M_W^2$
M_e	lepton-higgs term	$e(el.\ chg.)=gSIN\theta_W$	

amined β-decay and μ^+ decay for evidence of universality of the charged vector couplings (i.e. $W^{\pm}$ exchange).

In order to calculate with this theory beyond first order, the question of renormalization has to be addressed.[20] Although the theory has been shown to be renormalizable to all orders, there are many ways to absorb the infinities by redefinition of the fields, masses, and coupling constants. Sirlin[20] outlines a renormalization framework in which M_W and M_Z are the physical masses. The Weinberg angle θ_W as defined in Table II becomes $COS^2\theta_W = (M_W^2)/(M_Z^2)$, and establishing e as the physical charge, the relation $e = g \, SIN\theta_W$ with $TAN\theta_W = g'/g$ gives a set of gauge invariant parameters and a prescription for generating the renormalization counter terms. Sirlin then renormalizes diagrams that contribute to μ^+ decay to the order of (αG_F) where α is the fine structure constant and $(G_F/\sqrt{2}) = (g^2/8M_W^2)$. These diagrams can be found in Fig. 1.

Diagram a of Fig. 1 is the first order process that contributes to muon decay. It reduces to the ordinary local V-A theory when q/M_W is neglected, q being the momentum transfer. Since the decay of the muon involves only modest q values, the effect of the W propagation contributes to a term of order m_μ^2/M_W^2. Diagram a of Fig. 1 results in the following

$$\frac{1}{\tau_\mu} = \frac{G_F^2 m_\mu^5}{192\pi^3} \left[1 - 8\frac{m_e^2}{m_\mu^2} + 0\left(\frac{m_e^3}{m_\mu^3}\right)\right] \left(1 + \frac{3m_\mu^2}{5M_W^2}\right).$$

Higher order diagrams can be divided into two parts. The photon propagator can be written as

$$\frac{1}{k^2} = \frac{1}{k^2 - M_W^2} + \frac{M_W^2}{M_W^2 - k^2} \frac{1}{k^2} .$$

Then the photonic correction can be evaluated in two parts. Diagrams d, f, m with photons cut off by M_W along with bremsstrahlung diagrams (b,c of Fig. 1) reduce to the local V-A photonic corrections and contribute a term given by

$$\frac{\alpha}{2\pi} (\frac{25}{4} - \pi^2) .$$

The remainder of the terms can be absorbed by defining a coupling such that

$$\hat{g} = g(1 + \Delta\sqrt{})$$

where $\Delta\sqrt{}$ contains all the corrections of the remaining diagrams. Since all charged current interactions contain these terms, the definition of $G_\mu / \sqrt{2} = \hat{g}^2 / M_W^2$ provides a convenient parameter of the theory.

G_μ is related to the muon lifetime by

$$\frac{1}{\tau_\mu} = \frac{G_\mu^2 m_\mu^5}{192\pi^3} [1 - 8\frac{m_e^2}{m_\mu^2}] (1 + \frac{3m_\mu^2}{5M_W^2} + \frac{\alpha}{2\pi} [\frac{25}{4} - \pi^2]) .$$

Non-Abelian gauge theories, specifically the Weinberg and Salam model, provide a theory for a universal weak and electromagnetic interaction. The charged W exchange has the V-A form of the phenomenological local model and encompasses the conserved vector current CVC and partially conserved axial current PCAC hypotheses due to different renormalizations of vector and axial vector terms. The model predicts the neu-

tral Z_o exchange and is successful in its comparison of a wide range of experiments. It does not exhibit CP violation and additional modifications are needed to violate this symmetry.

In order to determine G_μ, a complete second order calculation has been outlined by Sirlin[20] with the Weinberg-Salam model so that no questions remain about corrections to this order, except for the effects of the strong interaction on charge renormalization. It should again be noted that most of these corrections are absorbed in the definition of G_μ so as to avoid unnecessary complication. G_μ as defined then serves as a convenient coupling for all charged weak interactions.

Along with the above detailed calculation of G_μ, a more general treatment of the whole process of muon decay is necessary since certain experimental details are based on several more salient aspects. A muon at rest will decay into any one of the allowable spin and momentum final states. The probability of observing a decay into a specific final state j at a time t later is

$$\mathcal{P}(t)dt = e^{-\lambda_{tot} t} \lambda_j dt \quad ,$$

where $e^{-\lambda_{tot} t}$ is the probability the muon has not decayed by t and λ_j is the probability per unit time of a decay into the final state j.

$$\lambda_{tot} = \sum_j \lambda_j \quad , \text{ with}$$

$$\lambda_j = \frac{2\delta^4 (P_\mu - P_e - P_\nu - P_\nu)}{(2\pi)^5 E_\mu E_e E_\nu E_\nu} \; |<j|S|i>|^2 d^3P_e d^3P_\nu d^3P_\nu .$$

λ_{tot} is the total decay rate determined by the sum over all possible channels and $<j|S|i>$ is a scattering matrix element between the intial state of the muon $|i>$ and the final state $|j>$. The initial state of the muon is characterized by its spin direction (polarization) and its four-momentum. The final state is characterized by the spins and momenta of the positron and the neutrinos. The $\int_j$ implies a sum over spins and an integral over momenta. Due to the elusive nature of neutrinos the positron in the final state is usually the only one detected. The probability that a muon at rest decays into a positron of momentum p and spin s is obtained by integrating over the neutrino final states. The differential decay probability is[1]

$$\frac{d^2\Gamma(x,\theta,\phi,\psi)}{dx \; d\cos\theta} = \frac{m_\mu^5 G_\mu^2}{384\pi^3} \; x^2 \{(3-2x)-(2x-1)\cos\theta\}(1-\cos\phi) ,$$

assuming V-A is the correct form for the process. The angles θ, ϕ and ψ are shown in Fig. 2. The standard reduced energy is $x=2E/(M_\mu)$, where x ranges from M_e/M_μ to 1.0. A more general description in terms of the muon decay parameters[1] allows a different form for the interaction but all measurements of these parameters are consistent with the V-A theory. The energy dependence of the decay probability (Michel spectrum) is shown in Fig. 3. Two important aspects for our experiment are the asymmetric decay with respect to the muon polarization and the decay positron energy. Since

the muon will precess in a magnetic field and the direction
of emission of the decay positron will depend on the muon
polarization, a time dependent modulation will be observed
for any detection efficiency that depends on direction.
This well known property of muon decay has been exploited to
measure magnetic fields in solids and forms the basis for
the field of μSR.

Chapter III

EXPERIMENTAL CONSIDERATIONS

3.1   INTRODUCTION

The key to a precise measurement of this nature is data
free of unknown systematic effects. First, the electronics
must record or reject events strictly according to the plan-
ned criteria and should behave identically under the various
conditions of data accumulation. If time intervals are to
be measured with a crystal-controlled oscillator, the sta-
bility and precision of the oscillator have to be maintained
with precision greater than that which is sought in τ_μ.
Secondly, the design of the experiment, specifically the
criteria chosen for sorting data, should not introduce un-
known time dependencies into the data distribution function.

The combination of the properties of the electronics
(e.g. resolution, deadtime, efficiencies), the conditions
(e.g. rates, beam composition), and logic criteria (e.g.
start and stop definitions, control bits, sorting schemes)
combine to effect the measured time distribution of data. A
complete knowledge of this distribution, in principle, es-
tablishes a basis from which the value τ_μ is determined and

its error obtained (see Appendix B). The distribution function, however, cannot be completely determined (see Appendix A). Thus, the approach to this measurement which is detailed in the following sections has been to vary the conditions under which data are accumulated and concurrently to record events of different multiplicity and of different detected positron energy in order to simultaneously estimate τ_μ while testing for systematic errors.

3.2 BEAM

The experiment was performed during two runs at the TRIUMF meson facility.[21,22] Secondary beam lines M-13 and M-11, as shown in Fig. 4, served as sources of π^+ and μ^+. The first data were recorded in July 1981 in M-13, followed by a second run in August 1981 in beam line M-11. Both beam lines transmitted particles produced when 500 MeV protons impinged upon the production target Tl. The time microstructure of the proton beam, a 2-5 nsec wide burst every 43 nsec, resulted in the microstructure of the secondary beam lines, as demonstrated by the time of flight (TOF) beam profile of M-11 shown in Fig. 5.

The arrival of protons was signalled by a capacitive probe at high proton intensities, and by the RF controlling modulation at low intensities. Both signals were available

to the experimenter and were used in beam-defining signatures. The time macrostructure of the beam was determined by a constant rate stochastic process, except for short periods when the proton beam was turned off. The duty cycle of the machine was a measure of this macrostructure. In normal operation, the duty cycle was approximately 99%, but could drop to a few percent.

The secondary beam lines were standard momentum-selective channels. M-13 focused particles from the back of the production target to the apparatus approximately 10 m away. This channel was tuned on various runs for beam momenta ranging from 150 MeV/c to 170 MeV/c. These momenta were beyond the original design specifications for the channel, and the saturation of the quadrupole magnets resulted in a relatively poor final focus.

M-11 is a forward beam designed for higher momenta than M-13, and provided a very sharp focus at the detectors situated 14 m from the production target T1. This beam length resulted in much better separation in time of arrival of π's, μ's, and e's, but in a larger percentage of muons and electrons.

The relevant running conditions are summarized in Table III. The variation in either channel in constituent particle makeup was due primarily to different production targets

placed at T1. During the course of the experiment, proton currents ranged from a few nA to 30 μA. The protons in secondary beams were removed by inserting degrader at the first focus of both M-13 and M-11. Both M-13 and M-11 data were accumulated over various incident beam intensities to provide a test for rate-dependent effects.

TABLE III

Characteristics of M-11 and M-13 beam lines

	M-13	M-11
Momenta	150-170(MeV/c)	150-170(Mev/c)
Composition (%) (/ /e)	93/4/3 - 96/4/.25	72/19/10 - 80/13/6
Intensities	5-30(kHz)	5-30(kHz)
Channel length	10 m	14 m
Takeoff angle	135°	2.5^{o}
Momentum Res.	2%	3%
T1 targets	28mm H_2O, 10mm C	2mm C, 10mm C, 28mm H_2O
Proton intensities	200nA-3 A	200nA-30 A

3.3 DETECTORS

The incident beam was defined by two scintillation counters. The detection of the decay positron was accomplished in a large-volume H_2O Cherenkov detector. The principal scheme, as outlined by Balandin et al.,[11] is to use a medium with sufficient stopping power to bring pions and muons to rest near the center of the target, thereby allowing the de-

cay positrons a long path length along which to radiate, and
providing 4π solid angle. The index of refraction should be
chosen so that positrons between 1-50 MeV will radiate but
charged particle backgrounds from the beam will be below the
threshold for Cherenkov radiation. The implementation of an
energy tag (high, medium, low), determined from the intensi-
ty of radiated light, allows events to be sorted not only by
the multiplicity of the detected beam or decay particles,
but also by positron energy.

The theory of Cherenkov radiation and its applications in
particle detection is outlined in many referenc-
es.[23,24,25,26] The following review will focus on those as-
pects relevant to this experiment.

A charged particle whose velocity exceeds c/n in a medium
with an index of refraction n will emit photons. The char-
acter of these photons, the angular emission (highly direc-
tional), polarization (along the momentum vector), and the
spectral intensity all characterize Cherenkov radiation.
The number of photons dN emitted per unit frequency interval
d$\checkmark$ is:

$$\frac{dN}{d\mathcal{V}} = \frac{2\pi l}{c} SIN^2 \mathcal{O}$$

for a trajectory of length 1 and for $\mathcal{O} = COS^{-1}(c/vn)$. For wa-
ter in which the index of refraction remains relatively con-
stant between the relevant wavelengths 1000$\overset{o}{A}$-7000$\overset{o}{A}$, most of

the photons are emitted in the ultraviolet (UV) portion of
the spectrum. At shorter wavelengths, the refractive index
of water is not large enough to satisfy the Cherenkov condi-
tion v>c/n, as seen in Fig. 6.

As a particle travels through a medium it loses energy by
several processes. The primary energy loss mechanism for
charged particles is ionization. As a particle slows to a
velocity v<c/n, the intensity of emitted Cherenkov radiation
drops to zero. To estimate the number of photons generated
by pions, muons and positrons traversing the water Cherenkov
counter, their trajectories were divided into small unit
cells. An estimate of the intensity of Cherenkov radiation
versus the intial particle momentum was accomplished by us-
ing the average velocity per cell to calculate the intensity
of Cherenkov radiation and the particle energy loss. Fig. 7
shows the results of this calculation by displaying the av-
erage number of Cherenkov photons versus incident momentum
as pion, muons, and positrons fall below the Cherenkov
threshold in H_2O. Since a pion (or muon) of a particular
momentum produces on the average the same number of photons
as a positron of some lower momentum, these signals from the
Cherenkov detector are indistinguishable. It is therefore
convenient to view the pion and muon signals from the Cher-
enkov counter in terms of the equivalent positron energy
which would produce the same signal. Fig. 8 shows this
equivalence.

Aside from normal energy loss mechanisms, positrons can lose a substantial fraction of their total energy by emission of one bremsstrahlung photon. The end point of the Michel decay positron spectrum lies below the critical energy for water (84 MeV) and these positrons travel less than one radiation length (37 g/cm^2), so the relative probability for this type of bremsstrahlung is small. Beam positrons, however, have much higher energies and therefore readily emit high energy bremsstrahlung photons. This energy loss mechanism results in a significant reduction in total Cherenkov radiation. These events will appear on the low energy side of the Cherenkov peak for our H_2O detector. Thus the pulse height spectrum of Fig. 9 has a low energy tail. Any process in which a positron loses a significant fraction of its energy (e.g. inelastic scattering or bremsstrahlung) decreases the amount of Cherenkov radiation and results in low energy tailing. These tails are evident in Fig. 10 for 56 MeV beam positrons, although reduced compared to the higher energy beam positron spectra. These tails represent a limit for the detection efficiency.

The consequences of using Cherenkov radiation for particle detection are: relatively low yields of photons (as compared to the scintillation process), directional dependence of the radiation on particle trajectory, and a spectrum concentrated in the UV.

There are two approaches to sensing the UV portion of the Cherenkov spectrum: either UV transmitting and reflecting materials can be used, or, as in the present experiment, a waveshifter can be used to absorb the UV photons, reemitting photons of lower energy. Di-potassium 2-amino-6, 8-naphthalene-disulfonic acid[27,28,29] (G-amino), diluted in the medium of distilled water, was chosen to act as the waveshifter. The isotropic reemission of the frequency shifted photons additionally removed the directional nature of the Cherenkov photons. The G-amino broadened the output pulse of the detector, slightly, since its lifetime for fluorescence is on the nsec level.

The Cherenkov counter, as shown in Fig. 11, consisted of the following components: a) a stainless steel containment vessel with two UVT plastic windows; b) two 5" RCA 8854 photomultiplier tubes, specially chosen for their low noise, good quantum efficiency, and high gain characteristics; c) a white diffuse reflective liner[30] (Kydex); and d) a water plus waveshifter medium.

As discussed in Appendix A care must be taken to insure that no time dependencies are introduced into the data distribution function via beam muon precession, muonium precession or depolarization. Therefore, special attention was paid to the ambient magnetic field at the detector center.

To avoid muon precession a three-axis bucking system[31] was used along with thin cylinders of shielding material.[32] The position of the field cancelling coils relative to the detection apparatus is shown in Fig. 12. The ambient magnetic field was monitored during the run by measuring B_x, B_y, and B_z at a reference point. The self correcting bucking coils then kept the ambient field below 50 mG. Aside from free muon precession, a muon can bind to an electron to form muonium. This system precesses with 100 times the frequency of free muons and thereby requires better field reduction to eliminate possible systematics. To observe the formation of muonium in water[33] using muon spin resonance techniques requires a carefully degassed target. The absence of precession-generated time dependence was insured by dissolving sufficient oxygen in the H_2O target in order to depolarize the muonium atoms rapidly.

Even with zero magnetic field (no precession) the depolarization of the free muon component produces unwanted effects in positron detection efficiency. Recent reports[34] for the mean depolarization T_{dp} for muons in water establish $T_{dp} > 100$ μsec. Although only small quantities of G-amino were dissolved in the water target, an estimate of its contribution to depolarization was deemed necessary. Tests using pulsed nuclear magnetic resonance techniques[35] determined the proton relaxation constants T_1 and T_2 for various

concentrations of G-amino in water. These results implied a
negligible contribution to muon depolarization by G-amino.

3.4 H_2O TARGET CHARACTERISTICS

The Cherenkov detector was subjected to a series of
tests[36] to determine its operational properties. Cosmic ray
and light emitting diode (LED) studies provided useful in-
formation for the final testing and run at TRIUMF.

Pulse height spectra were examined as a LED was moved
through the volume of the radiator. A variation of less
than 20% was typical. Studies with cosmic-ray and beam-po-
sitron trajectories also revealed a variation of less than
20% in light collection efficiency throughout the counter.

Concern with directional dependence resulted from both a
desire to maximize overall efficiency of positron detection
and to avoid light collecting efficiencies that were depen-
dent on the decay positron direction. There was less than
1% variation in efficiency detected over the central regions
of the detector.

Results from tests with cosmic rays and accelerator-pro-
duced positrons agreed with published reports[27,28,29] on the
performance of the organic G-amino waveshifter. An approxi-
mate fourfold increase in overall light collection resulted

from using an 8 mg/10^3cm^3 solution. No special purification technique was used in diluting the waveshifter. There was no evidence of any deterioration of the pulse height signal from the water counter over the 3-4 week periods that each solution was used.

Both the energy dependence of the intensity and the resolution were measured for the Cherenkov detector. The one, two, and three photoelectron peaks were observed from each photomultiplier tube using low-intensity LEDs. These LED tests satisfactorily set an approximate absolute calibration which then checked against the inverse square root dependence of the pulse widths on the number of photoelectrons.

The number of photoelectron events n_{pe} observed depends on the photon collection efficiency ϵ and on the quantum efficiency Q of the photomultiplier tubes:

$$n_{pe} = n_c \epsilon Q$$

A comparison of the total number of Cherenkov photons n_c versus the number of photoelectrons showed that about 1% of all the light is detected. This result is reasonable because the phototubes intercept a solid angle of 8%, typical quantum efficiencies Q are less than 30%, and a reflectivity of 80% is reported[30] for the Kydex reflective liner. Figure 13 shows the number of photoelectrons observed in the 5" photomultiplier tubes viewing the water medium.

Figure 10 indicates the response of the positron detection system to 20, 30, and 56 MeV beam positrons along with a decay positron spectrum. The observed shape of the Michel spectrum is due to a folding of the actual spectrum, the resolution function, and the cutoff effects. Figure 14 shows that the 120 MeV/c π^+ beam gives rise to a decay positron spectrum shifted lower in pulse height than the 170 MeV/c beam. This shift is a result of "energy cutoff" since the positrons can leave the upstream side of the detector without depositing all their energy.

The overall detection efficiency for decay positrons was estimated to be 95%, including that fraction of the spectrum which falls below the cutoff discriminator settings. This estimate was obtained by comparing the number of data events to the total number of event starts. Ten percent of all accepted event starts did not have accompanying decay positrons and five percent of these starts were assumed to be positive pions absorbed in nuclear processes.[37,38]

The discriminator levels which determined the low, medium, and high energy sorting of positron events were relatively easy to test and optimum conditions were established quickly. Most of the data were taken with the conditions shown in Table IV. The effects of discriminator settings are demonstrated by gating the Cherenkov spectra with the

high, medium, and low energy signals. These spectra are displayed in Fig. 15 along with a beam-muon gated spectrum and a spectrum gated by the coincidence of the photomultiplier tubes of the Cherenkov detector. The low-energy gated spectra then consist of low, medium and high Cherenkov signals while the high-energy gating results in accumulation of only high intensity Cherenkov events.

TABLE IV

Energy Sorting Conditions

Energy Window	Relative Percent of Decay Positrons	
	Predicted	Observed
Low 10-20 MeV	10	13
Med 20-50 MeV	80	83
High 50- MeV	10	4

Finally, Table V lists all the possible sources of Cherenkov light in the counter. For all testing done, there was no evidence that scintillation processes from incident charged particles contributed to the signals. All H_2O counter signals were apparently directly generated by Cherenkov radiation. However, positive pions and neutrons could constitute indirect sources of Cherenkov light by producing δ-rays which in turn can produce electron-positron pairs, Compton electrons, and photo-electrons which could subsequently radiate.

TABLE V

Candidate Cherenkov Sources

Particle	Source	Observations
Michel e^+	$95\% \pi^+ + 100\% \mu^+$	95% efficient
Beam e^+	100% beam e^+	100% efficient
Cosmic Rays		50Hz
Neutrals		50Hz
Beam Muons		25%
Beam Pions	5% nuclear absorbed	20% efficient
Pionic decay e^+	Branching Ratio $< 2 \times 10^{-4}$	

During the run, the ratio of Cherenkov EL signals to in-
cident beam particles (12 coincidences) ranged from 97% to
103%. The EL signal consisted of all signals above the
threshold of the low discriminator. In addition, the rela-
tive number of prompt (i.e. t=0) Cherenkov signals ranged
between 3%-6%. Over 80% of these prompt signals were in the
low-energy multiple spectra. They are therefore believed to
be caused primarily by Cherenkov radiation from beam muons.

3.5 LOGIC AND CLOCK

A complete list of logic signals used to implement the
data-taking scheme is shown in Table VI. The particle de-
tection logic is shown in Fig. 16 and 17. The identifica-
tion of the pions, muons and positrons in the beam was done
by TOF. Attempts were made to use the water target to sepa-
rate beam pions and muons, but this did not prove to be very

applicable. No lifetime data were obtained with the H_2O target incorporated into any of the beam-defining logic. Also, no beam logic vetos were imposed upon the positron detection circuit. The signals from the phototubes of the water counter were processed to provide the time of the event (EL) along with energy tags (EL,EM, EH). Care was taken to insure good time resolution and efficiency with minimal noise. The anode signals were clipped, summed, and filtered[39] before being applied to a constant fraction discriminator CFD for time pickoff. The inclusion of an AB (Cherenkov signal) coincidence was carefully tested and adjusted to eliminate the after-pulsing that has been reported for 8854 photomultiplier tubes.[40]

TABLE VI

Logic Signals

Signals	Description	Definition
S	Candidate event start	12 or 12RF
BS	Candidate post beam event	12 or 12RF
EL	Time of positron event	AB(A+B)
EM	Medium energy tag	(A+B)
EH	High energy tag	(A+B)
PG	Pile up "cleaning"	12 or 2
	Muon tag	12RF

The signals described above were then used to determine time intervals, event multiplicity and energy for appropriate routing of the data into histograms. Histogram defini-

tions may be found in Table VII. These histograms were stored temporarily in the memory of a PDP-11 computer.

TABLE VII

Histogram Definitions

Histogram	Character	Criteria
1	low clean	1 and only 1 positron in 20 μsec
2	medium clean	with pre and post pile up gate
3	high clean	cleaning for π^+ start.
4	low pre	up to a multiplicity of 3 positrons
5	medium pre	with 20 μsec post beam cleaning
6	high pre	
7	1 clean muon	1 and only 1 positron with 20 μsec
8	m clean muon	pre and post cleaning for μ^+ start.
9	h clean muon	
10	low post	up to 3 positrons and
11	med post	at least one post beam signal
12	high post	
13	beam pion	1 and only 1 post beam event in
14	beam muon	the 20 μsec period after a start

Two parallel timing channels were used so that data could be accumulating in one while the computer was processing the other. Each channel consisted of: a) a digitron (scaler bank), where up to four time intervals could be stored, b) a pattern unit, indicating the status of an event, and c) an event trigger. The time intervals, $(t_{e1}-t_{\pi 0})$, $(t_{e2}-t_{\pi 0})$, $(t_{e3}-t_{\pi 0})$, and $(t_1-t_{\pi 0})$, where $t_{\pi 0}$ is an event start, t_{ei}

is the time of arrival of the ith Cherenkov event, and $t_{\pi i}$ is a beam particle arriving after a start, were measured during the 20 μsec period after a start.

A logic unit was designed and built at William and Mary to accommodate this dual-digitron multiple-event routing. This unit, along with the digitrons, pattern units, oscillator, and event triggers, constitutes the clock logic displayed in Fig. 18. This system was interfaced to a PDP-11 computer via a CAMAC system controlled by a microprogrammable branch driver[41,42,43,44] (MBD). In all, 28 2048-channel histograms were accumulated and, upon overflow of any channel, the histograms were written to tape along with 30 scalers. The time between tape updates was typically 15 minutes, but could be modified by changing the histogram overflow limit.

A typical event sequence follows. A stopped pion signal S initiated a 20 μsec gate, provided a channel was available and an event was not already in progress. This pion signal then started four scalers (SC_A or SC_B) counting a 100 MHz crystal oscillator pulse train. Subsequent positron signals (t_{ei}) stopped scalers 2, 3, and 4 of the scaler bank (SC_A or SC_B) while the first beam signal after a start stopped scaler 1. These measured time intervals would eventually cause the incrementing of channels in the 28 histograms provided

the event did not overflow. More than three positron sig-
nals would cause all positron time intervals to be ignored.
More than one post beam signal would cause the beam time in-
terval to be ignored. If both a positron overflow and a
beam overflow occurred the event was immediately reset and
the system made ready to accept the next start. The 18th
bit associated with each measured time interval (SC 1,2,3,4)
was used to signify the multiplicity. If, for example, the
18th bit of scaler 3 was set then at least three positron
signals had occurred. If the 18th bit of scaler 1 was set
the multiplicity for pions was at least two and no beam his-
tograms would be incremented. The pile-up gate (PG) logic
generated 20 μsec pre and post cleaning by setting a bit in
the pattern unit (PU$_A$ or PU$_B$) if the start was preceded or
succeeded by a signal PG. The positron-energy information
was also recorded in the pattern unit. Energy bits were set
by a coincidence between EM and ei, and EH and ei, where ei
is the time of the ith positron event as generated by the
control unit. The deadtimes of both the positron logic and
the beam logic were fixed at 50 nsec. Upon completion of a
20 μsec period, the control logic either generated a comput-
er interrupt signal or reset the channel (indicating no
data).

The MBD was able to process an event completely in 50
μsec for the events with maximal multiplicity and in only 25

μsec if a single-positron event occurred. This processing time excluded computer deadtime losses for clean data.

The clock system was subjected to extensive "randoms" testing. In addition "fixed phase" (FP) testing was done in which the start of a time interval was varied with respect to the oscillator phase. These FP test were essential for understanding the possible effects of differential non-linearity in the system. The system showed no non-linearity or distortions to a level of 10^{-5}. The crystal controlled oscillator's frequency and stability were regularly checked against the network television station color carrier frequencies which are based on the NBS atomic clock standard[45]. The oscillator was adjusted to 100 MHz at the beginning of the run and remained stable to better than one part in 10^6 throughout the experiment.

Chapter IV

ANALYSIS OF DATA

4.1   **INTRODUCTION**

As stated in Chapter III, the data were stored in 28 histograms, two sets of 14, one set from each of the two parallel timing channels (A and B). These histograms are defined in Table VII. Two of each set of 14 histograms (13 and 14) served to monitor the time dependence of the arrival of beam particles following a start signal. Six of the histograms (4,5,6,10,11,12) provided information on the time dependence of the background. The remaining six histograms were termed the clean spectra because of their relatively low background levels. Three of these (1,2,3) were started primarily by the arrival of a pion and three (7,8,9) were started by an enhanced ratio of muons to pions. The clean muon spectra (7,8,9) were used to test for the presence of a time dependence generated by muon precession and depolarization.

The data were routed into the histograms according to the energy of the detected decay positrons. Thus spectra 1,4,7,10 contained low-energy positrons from muon decay along with the low-energy background. Spectra 2,5,8 and 11

contained the majority of the muon-decay positrons and the medium-energy background. The high-energy decay positrons along with high-velocity particles such as beam positrons, $\pi^+ \to e^+ + \sqrt{}$ and cosmic rays were routed to spectra 3,6,9 and 12. Plots of the histogram data can be found in Fig. 19 and Fig. 20.

These histograms were written to tape periodically along with the contents of 30 scalers that monitored the performance of the electronic circuitry and the running conditions during the course of that segment of data accumulation. These tape updates represent the smallest fraction of data available for separate analysis. In all approximately 800 updates were written to tape.

Since the analysis is very dependent on the theory of statistics a more general discussion of this topic is given in Appendix B. Appendix A provides a detailed description of the arguments used to arrive at the expected form for the data distribution function. The positive muon lifetime was obtained by performing a least squares fit of all the data consistent with the assumed form

$$a \ e^{-t/\tau_\mu} + b \quad .$$

The data from the two timing channels were always analyzed separately. All plots and references to histogram data or fits entail timing channel A results, timing channel B results and the summed results. The timing channels were examined separately as a test of the electronics. Although not completely independent, electronics for these timing channels did entail a fair number of different components.

The goals of the analysis were to choose the best subset of data for determining the positive muon lifetime, to verify the reliability of the data-taking system and to test the assumptions made about the data accumulation. This chapter consists of a description of the method used to scrutinize the data to determine this subset of good data and of the most important tests used to verify the results.

4.2 TESTING

Before combining the tape updates into larger portions of data two tests were completed. First the data from the clean spectra 2 were fitted to an exponential function plus a flat background. The results of these fits, the lifetime values and their errors, were plotted along with the normalized χ^2 values for timing channel A, timing channel B and the sum. The sequential arrangement of these lifetime values showed no evidence of any unexplained non-statistical behavior.

This complete set of lifetime values for spectra 2A and 2B was arranged in histograms by subtracting 2.19695 μsec from each value and dividing by the error obtained from the fit. In order to obtain a convenient range, these values were then multiplied by six, added to 15.5 and the result was truncated to give the histogram bin. These calculated values were then used to form a 30-channel histogram. If the lifetimes measured for the 800 update histograms fluctuates according to statistical principles the histograms should follow a Gaussian distribution with σ=6. The results of this analysis are shown in Fig. 21 and indicate the expected statistical behavior. The solid curve represents the Gaussian fit to this histogram. Also included in Fig. 21 are histograms that show the distribution of the normalized χ^2 values for every update fit. These were similarly formed. The central bin (15) contains the number of updates with χ^2 values close to one. The scale has been chosen so that a deviation of 1σ in χ^2 is three channels.

The updates were then examined to determine the performance of the electronic circuitry. A typical set of scaler rates is given in Table VIII . These scalers were monitored for each update along with various histogram areas and the length of time for the update by forming various ratios and tabulating the results. For example, the number of computer interrupts per timing channel divided by the number of 20

μsec gate starts provided information on the percentage of time that no stop signals were obtained during the 20 μsec gate. Also the number of 20 μsec gates per candidate start signal and the number of beam particles incident during the 20 μsec gate were examined. The ratios among low, medium, and high energy positron signals were similarly examined for each update. In all, approximately 15 types of ratios were compared for the data in each update. The two timing channels performed identically under these tests. All data rates and ratios were consistent with the expected operation of the system. Any changes could be attributed to the various changes in running conditions or to statistical fluctuation.

Not all the data were readily analyzable. Several criteria were used to judge updates. A small fraction of the updates was discarded because of tape errors. Further, some of the data could not be successfully retrieved after computer failures occurred during data accumulation. There were a few cases in which the data were corrupted by incorrect settings of the beam-defining logic and thus these updates were not included in the final data sample. The rest of the data was scanned to ascertain the relevant beam conditions. The duty cycle of the beam was monitored and all low-duty-cycle data were removed from the final sample. The reason for discarding these data were twofold: firstly,

TABLE VIII

Typical Scaler Rates

NAME	COUNTS/SEC
EL after deadtime setting disc. (dt)	13561.
EL (delayed)	13561.
EM (before dt)	12179.
EM (dt)	12154.
EM (delayed)	12154.
EH (before dt)	1801.
EH (dt)	1800.
EH (delayed)	1800.
Channel A Starts	4995.
Event Gates A	4995.
Computer Interrupts A	4588.
Positron Stops A	5732.
Beam Stops A	1252.
Channel B Starts	4992.
Event Gates B	4992.
Computer Interrupts B	4583.
Positron Stops B	5727.
Beams Stops B	1256.
RF (cap. probe)	22828.
1 2 (coin.)	13872.
Pi/Mu Start	12578.
Beam Muons	788.
Beam Positrons	2496.
2 (singles)	21554.
1 (singles)	16101.
Cherenkov A (low threshold)	20399.
Cherenkov B (low threshold)	19873.
Cherenkov A+B CFD	13744.
Cherenkov A B (coin.)	16452.
Cherenkov A B EL (coin.)	13608.

since the average rate of the incident beam was kept fairly
constant, the low duty cycle data had a very high instanta-
neous beam rate; secondly, there was concern that a low duty
cycle would introduce a non-flat background into the histo-
grams.

Finally, the data were summed to form the runs as they were recorded during the experiment. These were typically 8 hours in length and usually represented a constant environment for the accumulation of data. There was a series of runs in the early stages of the experiment which was found to contain an oscillation in the beam-monitoring histograms (13 and 14). This oscillation presumably occurred because of an unusual modulation of the incident beam. The cyclotron was operating in a non-standard low intensity mode during this period of the experiment. At no other time were any similar variations in the beam intensity observed in the beam-monitoring histograms.

With the exception of a ten hour period when timing channel B developed an odd-even fluctation, there was no evidence for any further difficulties with the data. An odd-even fluctuation results when the effective bin size is increased for odd (even) channels and is decreased in the even (odd) channels by the same amount. The variations induced on the ground plane of the discriminators following the coincidence of the oscillator and a gate were found to be the primary cause of odd-even type variations. Accompanying the odd-even fluctuations at a lower sensitivity level were similar variations for groups of n channels. On-line data inspection was used to locate and correct this effect. The odd-even fluctuation resulted only in an increase of the

value of χ^2 . Lifetime determinations themselves were not affected. This conclusion was based on computer simulations which duplicated the odd-even and modulus n effects. Subsequently the only result of the effect was an increase in the value of χ^2 for a few of the data sets used. Because these data constituted in any case a small fraction of the total data and because χ^2 is a very important measure of the validity of the functional form these data were not included in the final data sample. They were, however, used in the lifetime rate-dependent test.

The final sample of data represented about 70% of the total recorded. For most of the rejected data the lifetime and the χ^2 values deviated greatly from the average of the other sets. Some of the low-duty-cycle data however were consistent with the other data but were eliminated as a precaution.

In order to test the data for self-consistency, several data sets were generated by summing together various updates. As discussed in the Appendices there are possible distortions whose amplitudes depend on the instantaneous rate. An example of a rate-dependent effect is the deadtime distortion. The consistency of the lifetimes determined by updates subject to different incident beam rates provides a good test for possible systematics. Five data sets were

constructed by comparing the ratio of the area of the clean histograms to that of the beam histograms. This ratio was the most sensitve indicator of the average incident beam rate for any update. A plot of the lifetimes determined from these data sets is shown in Fig. 22 along with normalized χ^2 values. The estimated average rates during these data sets can be found in Table IX. The high χ^2 value for the spectrum of timing channel B labeled t7m101 is due to the large fraction of odd-even data in the data set. Also plotted in this figure are other groupings of the same data. Five data sets were constructed by summing the data that were collected during the same time period. The first two data sets were constructed from the 1981 July run (t1s04, t2s03) while the other three (t3s01, t4s03, t5so3) were sums of data from the 1981 August run. The values presented represent the results of fitting the medium clean spectra 2.

The final sample of data is data set t7s08 which represents the total data. Data set t7s08 has slightly larger values for χ^2 than is statistically expected. When the data are binned by two channels, the χ^2 values fall to within one standard deviation of their theoretical values and the lifetimes remain the same. The large χ^2 for the unbinned data is believed to be caused by a very small odd-even effect. For all binning factors 2,4,6,8 and 13 the lifetimes as determined from the data in t7s08 spectra 2 remained the same.

TABLE IX

Average Incident Beam Rates

Data set	Average Rate (kHz)
t7101	5 - 7
t7m101	7 - 11
t7m01	11 - 17
t7mh01	17 - 23
t7h01	23 - 30

The medium clean histograms 2 (π^+ start) and 8 (μ^+ start) were fitted over different ranges of channels. The results for spectra 2 are shown in Fig. 23 and the results for spectra 8 are shown in Fig. 24. Data from early times of the decay spectra were not expected to follow a simple exponential function with a flat background term because of the large numbers of pions that were used as sources of muons. These pions produce a growth function over the first few hundred nanoseconds succeeding a start.

Although 12 histograms for each timing channel were used to measure decay time intervals only spectra 2 and 8 were used to determine τ_μ. All 12 histograms were analyzed, however, as tests of the performance of the system and to search for background-induced effects. Two data sets, t7s10 and t7s12, were constructed from the updates for this analysis of τ_μ. The data accumulated in July were placed in T7s12 while the sum of all good data accumulated in August

was called T7s10. The reason for the distinction was two-fold. Firstly, the data were recorded under different conditions. The July data (t7s12) were taken in beam line M-13 while the August data (t7s10) were taken in M-11. The August run exhibited improved pile-up protection for the clean spectra with an improved signal-to-noise ratio. Secondly, the July run was not taken with a set of histograms started primarily by muons and cleaned with pile-up protection circuitry. During July, histograms 7,8,9 were multiple histograms and consequently had large backgrounds.

As mentioned in Appendices A and B the form assumed for the distribution of data is

$$ae^{-t/\tau_\mu} + b \quad .$$

The level at which deviations in the data from the assumed form is important depends on the sensitivity sought in the experiment. The inverse square root of the number of counts in a sample of data represents the fractional fluctuation of the data due to the statistical nature of the data sample. Fluctuations in the background below this level are then expected to be undetectable and incapable of causing deviations above statistical ones in the values of the parameters. The background levels in some of the histograms are considerably larger than those in the clean histograms. A small percentage of the background which does not conform to the assumed flat shape could in principle be detectable in

the high background spectra although negligible for the low
background spectra.

The data in all the histograms should follow similar dis-
tribution functions. The relative amount of background in
each spectrum could be drastically different but the overall
shape is expected to be the same. There could be differenc-
es in the background in histograms containing positrons of
different energy if the particles generating the background
had different time structure relative to a start. Cosmic
rays and beam positrons were responsible for a large frac-
tion of the data in the high-energy spectra. Decay positrons
from muons not associated with a start signal constituted a
large portion of the medium-energy background. Neutrons,
gamma rays and beam muons were perhaps the primary low-ener-
gy background constituents. If all the background had been
totally uncorrelated to a start, a flat distribution of
background would, in principle, be found in all the histo-
grams. The multiple histograms were intially included in
order to monitor any effects that would be sensitive to the
background.

The dominant effect observed in the multiple histograms
was a non-flat component in the background. The results of
the analysis showed that some of these spectra contained
time-dependent backgrounds sufficient to influence the de-
termination of the lifetime of these spectra. The results

from the fits to the 12 histograms for the July data (t7s12) and the August data (t7s10) are shown in Table X. This table gives the amplitude-to-background ratio a/b, along with the total number N of muon decays for each spectrum.

TABLE X

1981 July and August Fits

t7s10 (August)

Spec-trum	N (10^8)	a/b	τ_A (μsec)	τ_B (μsec)	$\Delta\tau$ (μsec)	χ^2 norm	Approx % non-flat bkgnd for 1τ
1	1.3	1600.	2.19665	2.19695	.0003	1.0	1.9
2	15.4	18000.	2.19695	2.19697	.00009	1.1	6.4
3	.8	3700.	2.19762	2.1968	.0004	.9	5.
4	.5	26.	2.1904	2.1913	.0007	3.5	0.04
5	6.6	100.	2.19699	2.1966	.0002	.9	0.06
6	.4	3.	2.204	2.206	.002	.8	0.07
7	.1	5600.	2.1964	2.1979	.001	1.1	27.
8	1.2	55000.	2.19698	2.19681	.0003	1.1	7.
9	.1	17000.	2.1973	2.1970	.001	.9	75.
10	.6	4.	2.145	2.146	.001		0.06
11	7.1	5.	2.1983	2.1980	.0003	4.8	0.03
12	.4	4.	2.208	2.205	.001	.9	0.08

t7s12 (July)

1	.9	610.	2.1947	2.1944	.0004	3.2	0.9
2	6.6	2400.	2.19680	2.19710	.00015	.9	1.27
3	.3	830.	2.19615	2.19698	.0007	1.0	2.
4	.2	12.	2.1998	2.203	.0015	1.3	0.04
5	1.0	40.	2.189	2.200	.0005	.9	0.05
6	.1	15.	2.1910	2.1961	.0025	1.1	0.1
7	.2	5.	2.24	2.194	.0015	1.8	0.02
8	1.2	7.	2.23	2.21	.0006	3.5	0.01
9	.1	7.	2.25	2.22	.003	1.2	0.1
10	.1	5.	2.22	2.185	.003	1.2	0.03
11	.7	7.	2.21	2.2	.0008	1.9	0.1
12		6.	2.26	2.205	.005	.9	0.05

The approach used to predict the level at which systematic deviation from flat background will affect a determination of τ_μ was to fit all the data with the assumed flat distribution. The deviation in the lifetimes obtained from these fits from that obtained using spectra 2A and 2B was a measure of the effect of the non-flat background. The percentage of the background that was non-flat should be the same for all histograms recording positrons of the same energy. The total number of counts associated with a non-flat background should be proportional to the amount of background itself. The sensitivity to this amount of "non-flat" data was assumed to scale with the square root of the number of true muon decay time intervals in the histogram. We expect that a spectrum requires a time dependent background whose area is comparable to the square root of the total decay area in order to cause a change in the lifetime on the same level as the statistical error. An estimate for the percentage of non-flat background can be made from the high-background data and its effect on the clean data can be considered. Included in the last column of Table X is an estimate of the percentage of non-flat background for each histogram which would be necessary to produce a systematic error on the lifetime comparable to the statistical error. The percentage of non-flat background in the high background histograms is below the sensitivity of the medium clean spectra. The validity of this approach does not depend on

the absolute value for the percentages given but only on the
relative value. The method is described merely to elucidate
the effects of the non-flat background. The more convincing
evidence that the fits of the medium-energy histograms re-
sult in an undistorted lifetime is the consistency of re-
sults obtained among histograms with various background lev-
els.

The various medium-energy spectra showed the smallest
non-flat component of background. This was the expected
situation since the primary source of the background in
these histograms was decay positrons from pre- and post-
muons. These positrons had minimal time correlation rela-
tive to a start signature. On the other hand, the low-ener-
gy and high-energy histograms displayed larger components of
non-flat background. The actual sources of the time-depen-
dent background could not be determined. The higher levels
of total background in the low- and high-energy spectra were
the result of the need for good efficiency for positron de-
tection. The low- and high-energy clean histograms were
therefore not used to determine τ_μ. These histograms for
August (t7s10), along with spectra 5,7,8 and 9 (A and B),
are consistent with the data from spectra 2A and 2B despite
the large variation in the relative amount of background.
When fitted, histograms 1,3 and 8 give approximately the
same error but with amplitude-to-background ratios that dif-

fer by as much as a factor of 30. The lifetimes obtained
are consistent with that of spectra 2. Spectra 5, with an
error on τ_μ that is only a factor of two larger than spectra
2 and an amplitude-to-background ratio that is 180 times
worse also results in a consistent determination of τ_μ.
Spectra 3 and 6 from the July data (t7s12) also agree with
the medium clean data. Spectra 1 from the July run (t7s12),
however, were slightly distorted due to the background lev-
els. The July run was notably more susceptible to background
in the clean histograms due to poorer pulse pile-up protec-
tion. The medium clean data for July (t7s12) were judged to
be free from any of the effects of the non-flat background.
The result of including the medium clean July data was a re-
duction in the error on the lifetime by approximately 20%.

Finally, the muon start data showed no evidence for any
oscillation due to the precession or depolarization of po-
larized muons. All the clean spectra were further fitted
with a term added to test for deadtime effects (d e^{-2t/τ_μ}).
The coefficient of this term (d) was consistent with zero,
and the value of χ^2 remained the same. These results agree
with the conclusion from the tests on rate dependence.

4.3 SUMMARY

It is impossible to be absolutely assured that all problems that could have caused this measurement of the μ^+ lifetime to be susceptible to systematic error have been addressed. Nevertheless, every questionable piece of data or result was scrutinized for evidence for some systematic effect. As indicated above, data with varying amounts of background collected under various beam conditions are consistent with the results from the medium clean data. Data with higher levels of background show the effects of a nonflat component, but considering the amount of background in these spectra they also support the conclusion that the medium clean data should be free from similar effects.

The value for the muon lifetime obtained from fitting the July and August data of spectra 2 and the August data of spectra 8 is

$$= (2.19695 \pm .00006) \,\mu\text{sec} .$$

This represents a precision of 27 ppm. Table XI contains a summary of the fits for individual timing channels.

TABLE XI

Positive Muon Lifetime Values

Data set	Timing Channel	Spectrum	Bin	τ_μ Lifetime (μsec)	$\Delta\tau$ (μsec)	χ^2 normalized
Medium Clean Muon Start Spectra (range 225-1900 channels)						
t7s10	8	A	1	2.19694	0.0003	0.97
t7s10	8	B	1	2.19677	0.0003	1.01
t7s10	8	A+B	1	2.19686	0.0002	1.02
Medium Clean Pion Start Total (range 225-1900 channels) data set t7s08 = t7s10 + t7s12						
t7s08	2	A	2	2.19691	0.00008	1.03
t7s08	2	B	2	2.19701	0.00008	1.04
t7s08	2	A+B	2	2.19697	0.00006	1.02
Medium Clean Pion and Muon Start Spectra (225-1900 channels) final = t7s08 (spec 2) + t7s10 (spec 8)						
final		A	2	2.19691	0.00008	1.04
final		B	2	2.19699	0.00008	1.03
final		A+B	2	2.19695	0.00006	1.02

Chapter V

CONCLUSION

The value for the positive muon lifetime measured in this experiment can be compared to the most recent measurements of Duclos et al.[10] (1973), Balandin et al.[11] (1974) and Bardin et al.[12] (1980). These lifetimes, along with the quoted errors, are shown in Fig. 25. The lifetime measured in this experiment (WM82) is also included.

The difference between the previous world average lifetime value[3] and the lifetime determined from this experiment is almost three standard deviations. The recently reported values for the μ^+ lifetime are shown in Table XII, with diferences given in terms of standard deviations. Because of the relatively large differences in lifetimes the error on the new world average is not substantially improved by this measurent, but the world average lifetime is lowered.

A new value for the weak coupling constant G_μ discussed in Chapter II can be determined using the lifetime reported in this experiment. The value of G_μ, following the method of Sirlin[20], is

$$G_\mu = (1.16637 \pm 0.00002) \times 10^{-5} \text{ GeV}^{-2}$$

TABLE XII

Recent Measurements of the Positive Muon Lifetime

Experiment	τ_μ (μsec)	$\Delta\tau_\mu$ (μsec)	difference with this experiment (std. dev.)
Duclos (1973)	2.1973	0.0003	1.2
Balandin (1974)	2.19711	0.00008	2.0
Bardin (1980)	2.19718	0.0001	2.3

The previous world average lifetime (WA)
(above values included)

WA (1982)	2.19714	0.00007	2.7

This measurement

2.19695	0.00006

The μ^+ lifetime measurement reported here can also be used to infer a capture rate for negative muons in liquid hydrogen. From the recently reported negative muon lifetime in liquid hydrogen[12] of (2.194903 ± 0.000066) μsec, we obtain 419 ± 19 $(sec)^{-1}$ for the capture rate, to be compared to the result of 460 ± 20 reported in ref. 12.

Appendix A

MATHEMATICAL DISTRIBUTION

A very important aspect of an experiment which is de-
signed to measure very precisely a physical quantity is a
knowledge of the distribution that governs the data. This
appendix is a review of the arguments that lead to the ex-
pected form for the data distribution function G.

In order to determine the positive muon lifetime, time
intervals between the arrival of a muon and its subsequent
decay product were measured. Along with these intervals a
large number of other time intervals was simultaneously re-
corded. All time intervals were measured by gating a crys-
tal-controlled oscillator and scaling the resulting pulse
train. The time intervals associated with a common start
signal compromised one event. The length of time T_g, as
measured from the start signal, during which gate-stops were
accepted, was chosen to be 20 μsec. The data from any one
event consisted of a maximum of four time intervals, three
or them associated with stop signals from the H_2O detector
and one resulting from a stop generated by the beam defining
logic. At the end of each event the data that had been re-
corded were stored in histograms. These histograms, there-

fore, consisted of n_j counts in the jth channel. The total number of channels available per histogram was 2048.

The probability for obtaining any histogram consisting of n_j data counts was

$$G(n_j; 0, \bar{N}) = e^{-\bar{N}} \prod_j \frac{P^{n_j}}{n_j!}$$

$$\text{with } P_j = \int_{channel} \mathcal{P}(t; 0) dt ,$$

where $\bar{N}$ and 0 are the parameters necessary to characterize the processes involved.

The probability density $\mathcal{P}(t; 0)$, corresponding to the single time interval distribution function, represents the folding together of the physical processes, criteria, and measurement distortions to give the probability that the time interval t will be recorded and included in the histogram described by G. If a function $\mathcal{P}(t; 0)$ can be found for each of the histograms then the data from each histogram can be interpreted as a measurement of N independent time intervals, where N is the total number of stops recorded in each histogram. At this point effects attributable to the random phase of the oscillator with respect to the time interval t have been ignored. A function $\mathcal{P}(t; 0)$ for each histogram recorded, along with the expected total number of counts $\bar{N}$ for each histogram, constitute a complete knowledge of the distribution of data.

There are two classes of physical processes that produce time intervals t. The first type of process generates a time interval as a result of some direct probability density $q_m(t;0)$. Here, we define the integral

$$\int_0^{T_g} q_m(t;0)dt = Q_m(T_g) \quad .$$

In the limit as T_g approaches infinity, Q_m approaches one. An example of a direct probability process is the arrival of a muon and its subsequent decay. One and only one time interval can be generated by this process. The differential probability for a time interval whose length is between t and t+dt is $q_m(t;0)dt$. The second class, processes of the non-homogeneous Poisson (NHP) type, are characterized by a rate $R(t)$ which can be a function of time. Those beam electrons that fall within an event gate are governed by such a process. For these processes, probability functions can be written only for specifically-stated conditions. For example, the probability of getting k and only k signals in an event gate T_g from a NHP process with a rate $R(t)$ is

$$P_k(T_g) = \frac{[\int R(t)dt]^k}{k!} e^{-\int R(t)dt} \quad .$$

Further, the probability density for obtaining the ith signal at time t if k total signals are found in T_g is

$$\mathcal{R}_j^k(t) = R(t)P_{j-1}(t)P_{k-j}(T_g-t) \quad .$$

The sum over j from 1 to k gives

$$\sum_j \mathcal{R}_j^k = R(t)P_{k-1}(T_g)$$

This final summation over the sequence of times demon-
strates two important principles used in determining G.
First, the sum retains the time dependence of the rate R(t)
due to the demand that only events with exactly k signals
are accepted and secondly, the summation demonstrates the
form of the probability function if all information as to
the ordering is ignored. This is the probability density
for obtaining a signal at time t out of the k signals per
event with no knowledge of whether this signal came first,
last, or at some intermediate time.

If all NHP processes are combined by summing over the
characteristic rates

$$R(t) = \sum_i R_i(t)$$

for these processes, and all direct probability processes
are determined by the q_m 's, then a similar summation can be
written for

$$\mathcal{P}(t) = \sum_m a_m q_m(t) + bR(t) .$$

The ordering of the time intervals in any event is lost
along with the knowledge of the process that generated each
time interval. The coefficients a_m and b are determined
primarily by the criteria imposed on each histogram and are
not dependent on time. For the "single cleaned" positron
histograms (1,2,or 3) the demand that there be one and only
one positron along with 20 μsec pre- and post-cleaning caus-

es the enhancement of the amplitude of all pion and muon decay sequences relative to other background processes.

Table XIII lists all the physical processes expected to contribute to the data distribution function. Both pions and muons are capable of producing a signal in the H_2O detector, thus providing both the start and stop signals for a time interval. Such intervals constitute less than 6% of the total number of intervals recorded. All of these intervals are short and therefore appear only in the channels near time zero. Approximately 95% of all incident pions decay into muons to serve as a source of muon decay time intervals. The other 5% are lost primarily through nuclear interactions in the target. Muons which constitute a fraction of the incident beam (<20%) also produce a substantial number of decay time intervals. A small sample of the time intervals in the beam histograms (13 and 14) obeyed an exponential decay probability of lifetime τ_μ. A positron of sufficient energy emitted so that it passes through counters 1 and 2 triggering a coincidence gives rise to such a distribution.

The rest of the time intervals recorded are due to processes that are independent; i.e. the start and stop that generate these intervals are not related in a direct way to each other. This does not rule out the possibility that a

non-constant time-dependent function might govern their distribution.

The beam is the primary source of these time intervals, either directly due to its composition or indirectly, by serving as a source of some secondary particles. A substantial number of the pions and muons that did not produce event starts nevertheless stopped in the H_2O target and decayed into positrons. If these positrons generated a stop signal during an event they could not be distinguished from the decay products of the actual particle that generated the start for the event. Cosmic rays were similiarly capable of generating a stop signal and consequently a recorded time interval if they occurred during an event. Other random signals compromise the so called "random component". Although it is important to understand the mechanisms for producing timing signals, the discovery of all sources giving rise to time dependent distortions was of fundamental concern. The "random component" constituted the remaining portion of the data and yielded a flat distribution in time. An example of a process believed to fall into this category is the accidental coincidence between phototube noise pulses. All of these processes then contributed to the time dependences that occurred in the experiment. Following a description of the behavior of undistorted processes, the relationship among the processes will be considered.

TABLE XIII

Primary Time Dependence

DIRECT PROCESS	FUNCTIONAL FORM
Muon Decay	$\frac{1}{\tau_\mu} e^{-t/\tau_\mu}$
Pion Decay	$\frac{1}{\tau_\mu - \tau_\pi} (e^{-t/\tau_\mu} - e^{-t/\tau_\pi})$
Muon(start+stop)	prompt
Pion(start+stop)	prompt
Upstream decay(beam histogram)	$A_o e^{-t/\tau_\mu}$

NHP Process	Rate
Beam	Constant Microstructure
Decay Positrons (Beam)	Constant
	$B_o e^{-t/\tau_\mu}$
	$C_o (1 - e^{-t/\tau_\mu})$
Cosmic Rays	Constant
Random Component	Constant

To discover the behavior of an NHP process the correlation between the start event and stop signal must be examined. If there is no correlation, only a flat distribution can be produced. One correlation that can relate two beam particles is the time microstructure of the beam. This arises from the pulsed nature of the proton beam. Every 43 nsec a pulse of pion protons arrives at the production tar-

get. If a secondary beam pulse can be characterized by a rate $f(t_a-t_o)$ with t_a the absolute time and t_o the center of the pulse, then the rate of the beam is

$$R_{beam}(t_a) = \sum_j f(t_a-[t_o+jT_B]) \ .$$

In this expression T_B is the seperation between proton pulses. The width of any pulse (<5 nsec) narrow, and therefore the function $f(z)$ is neglibibly small for z greater than 20 nsec. Because of the difference in flight times for the various constituents of the secondary beam a function f can be generated for the pions, muons and positrons:

$$f = f_{pion} + f_{muon} + f_{positron} \ .$$

There was of course some intensity variation between beam pulses. If this variation was significant for beam particles separated by less than 20 μsec and the variation was present for a significant sample of the events, then a time-dependent background could arise.

The beam then has two properties: one, the microstructure which causes any time interval that represents the separation between two beam particles to have a periodic structure, and the other, the macrostructure, which is a measure of the intensity variation among the pulses. The consequences of each will be considered separately.

Two histograms were recorded to check for any changes in the macrostructure that could cause the average intensity of the beam to vary during the course of an event. Although

the microstruture was always present, some of the histograms (e.g. "medium energy cleaned") had relatively few recorded beam time intervals and therefore the time dependencies due to the microstructure were suppressed. Other histograms (e.g. "high energy post" histogram) had a large fraction of the data contributed by time intervals started and stopped by beam particles.

As mentioned, the beam was a source of stopped pions and muons some of which did not generate event starts. The decay products of these particles could be associated in time with the event start only through the correlation of the parent particle with the start. Because the beam microstructure time scale was a factor of 50 smaller than the 2 μsec time scale of decay, the periodic correlation of the parent particles was averaged out. Computer simulations showed that the residual periodic fluctuations from this source were on the order of 0.2%. The macrostructure effects for these signals were evident in several histograms. Such effects occurred because of the selections imposed in the choice of an event start. As soon as the clock system became free the next available start signal initiated an event. The result was an average beam intensity preceeding a start that differed from the average beam intensity succeeding a start. Similarly, pre and post cleaning imposed a lower effective average beam intensity during the 40 μsec period surrounding an event start. It is important to note

that all of the histogram criteria changed the observable beam intensity only by increasing or decreasing it by a constant factor for the duration of an event or for the $20\,\mu sec$ period preceding an event. The criteria for routing the data into histograms were such that the relative probability for the beam particle to precede or succeed the start was different in the various histograms. The consequence for the background positron intervals was either an exponentially growing or exponentially decaying behavior. If, for example, the pion beam was abruptly turned off at t_1, the positrons from these pions would be characterized as an NHP process with a rate

$$R_e(t) = R\ e^{-(t-t_1)/\tau_\mu}\ .$$

For a beam abruptly turned on at t_2 the positron rate is

$$R_e(t) = R\ [1-e^{-(t-t_2)/\tau_\mu}]\ .$$

All histograms exhibit these growing or decaying components at some level. The "clean" histograms reflect this effect with an exponentially decaying backround from stopped particles arriving before the start of the $20\,\mu sec$ cleaning period. The amplitude at or below $t=0$ is already at the 0.01% level.

Along with the understanding of the distribution of time intervals an investigation of the possible distortions introduced by the measuring apparatus must be made. Two dis-

tortions were expected to contribute time-dependent varia-
tions in histograms at some level. Both rate dependent and
muon polarization distortions are crucial problems for meas-
urements of this nature. Rate dependence can result from
changes in the detection and logic systems due to a previ-
ously recorded signal. The deadtime of any component of the
system can induce a rate-dependent time variation. For
those particle-detection systems that cannot resolve two
distinct events within some small time separation, the de-
tection efficiency becomes dependent on the deadtime proba-
biltiy. This is just the probability that another event has
occurred within the resolving time of the system, causing an
inability to respond properly to the second signal. This
effect contributes a time dependence given by

$$\text{dead}(t) = \sum_{m} C_m q_m(t) R(t) + A R^2(t) \quad ,$$

the constants A and C_m being independent of time. For his-
tograms with very small time-dependent backgrounds, the
deadtime term is negligible. This was true for "clean" his-
tograms but not for the multiple histograms.

The second expected distortion stems from asymmetric de-
tection efficiency coupled to the polarization of the muon.
Since the positive muon has an enhanced probability to decay
into a positron whose momentum is parallel to the muon po-
larization, an effective time-dependent efficiency can arise
due to any time dependence in the direction of polarization.

If the muon either precesses or depolarizes as a function of time, this time dependence could manifest itself in the histograms. The amplitude for this distortion was determined by the number of muon starts and by the detection asymmetry. The precession frequency was determined by the ambient magnetic field. The depolarization contant depends on the stopping medium.[34] The inability to guarantee a negligible amplitude for these distortions imposed two constraints on the experiment. Firstly, the ambient field was held below 50 mG and secondly, the stopping medium was chosen for its long depolarization time constant. These choices insured that both muon precession and depolarization would not distort the measurement of the muon lifetime.

A further possible distortion that must be investigated occurs for those muons which capture an electron to form a muonium atom. For an H_2O target with dissolved oxygen the muonium depoloarizes too quickly to affect the histograms.[33]

Finally the digitron or gated-oscillator method used to measure time intervals in this experiment affects the distribution of data in the following manner. If a fixed time interval T_n is measured by the gated oscillator method, the number of pulses scaled will be n or n+1 depending on the phase of the oscillator with respect to the gate opening. The oscillator has a period ($\delta \sim 10$ nsec) with n= T_n/δ.

Given a general form $\mathcal{P}(t)$, the probability of incrementing the ith channel is

$$P_i = \int_{i\delta}^{(i+1)\delta} \frac{U(z)}{\delta} dz - \int_{(i-1)\delta}^{i\delta} \frac{U(z)}{\delta} dz$$

where
$$U(z) = \int_0^z \mathcal{P}(t) dt \quad .$$

Because of the small bin size used there is no significant effect for a slowly varying function of time.

A large number of difficulties is thus involved in understanding all contributions to a data distribution function, especially if the desired knowledge of the function is to be comparable to the precision of the attempted measurement. One consideration not yet mentioned is the possible deterioration of detector efficiency with time. Effects such as this should produce a random fluctuation or a time variation that is so slow that no appreciable change occurs over a 20 sec event. Perhaps the best test of the quality of the data taken is a demonstrated self consistency of results yielded in the analysis of the various histograms.

Appendix B

STATISTICAL ANALYSIS

In experimental measurements of this type there are prin-
ciples which justify estimating a physical quantity by ma-
nipulation of a recorded set of observations. Inherent in
this justification is the relationship between a set of pa-
rameters $\mathcal{O}_k$ that characterize the processes involved in the
experiment and a set of observations. These observations
could be a group of time intervals t_j or a histogram con-
sisting of n_j counts in the jth channel. The parameters may
or may not be known and the dependence of the data (<u>i.e.</u> ob-
servations) on a function $G(t_j;\mathcal{O}_k)$ may be derivable by theo-
retical considerations or could be the result of a phenome-
nological model. The distribution $G(t_j;\mathcal{O}_k)$ is the
differential probability of obtaining an experimental out-
come t_j, given the values of the parameters $\mathcal{O}_k$.

For the case of the direct measurement of the quantity $\mathcal{O}$,
the presence of the experimental uncertainties results in a
function $G(t_j;\mathcal{O})$ such that the variation of the measured
values t_j are normally distributed about the true value.
The lifetime measurement undertaken in this experiment, how-
ever, does not lend itself to direct measurement. Rather

another set of observations n_j are made and their dependence $G(n_j;\theta_k,\tau_\mu)$ on the lifetime τ_μ is theoretically derived. In principle one can invert the distribution $G(n_j;\theta_k)$ to give a function $H(\theta_k;n_j)$, which determines the probability that the parameters have values θ_j given an experimental outcome n_j[46]. A dilemma arises, however, for the case of a physical quantity believed to have only a single exact measurement (<u>i.e.</u> $H(\theta)\propto\delta(\theta-\theta^0)$). In this case the question of how much an experimental value is likely to differ from th true value is answered using $G(n_j;\theta_k)$. The quoted errors do not represent the width of a posterior probability function H, which assigns a probability to values of θ, but rather a "figure of merit" rating the estimation. The distinction is perhaps unimportant but in what follows the second approach is taken since τ_μ can in principle, have only one value.

In order to estimate a parameter of a known distribution G a set of statistical functions $\hat{\theta}=\hat{\theta}(n_j)$ is constructed from the data. These functions of the data will vary from one experiment to the next. The distribution of these statistical functions is governed by G where

$$\langle\hat{\theta}\rangle=\int\hat{\theta}(t_k)G(t_k;\theta_i)\ dt_1 dt_2 \ \cdots$$

is the mean value and $\langle\hat{\theta}^2\rangle-\langle\hat{\theta}\rangle^2$ is the variance.

There are many possible functions of the data that could be chosen for $\hat{\theta}$ in order to estimate θ. The important con-

cern is to choose a statistical function with good proper-
ties. It would be even better to find one with "best" be-
havior.

Peierls[47] suggests and examines several statistical func-
tions suitable for estimating τ_μ from a set of data obtained
in an experiment governed by the simple distribution

$$G(t_k; \tau_\mu) = \prod_k \frac{1}{\tau_\mu} e^{-t_k/\tau_\mu}$$

He obtains as the best statistical function $\hat{\tau}(t) = \frac{\Sigma t_i}{N}$, by
virtue of its minimal variance (mv)

$$\frac{\Delta\tau}{\tau} = \frac{1}{\sqrt{N}} \,,$$

where N is the total number of observations.

Along with minimal variance (mv), the properties of con-
sistency, bias , efficiency and sufficiency are defined in
treatises[48] on statistical analysis to aid in qualifying the
general properties of statistics. In principle a consistent
unbiased mv estimator can be found for all cases where G is
known and is well behaved. The values $\hat{O}(t)$ under the pre-
scription known as the maximum likelihood method (MLM) ap-
proach the true value O as the number of measurements in-
creases (i.e. satisfying the consistency property) while
maintaining the unbiased character $\langle\hat{O}\rangle = O$ independent of the
amount of data.

The MLM is detailed by Kendall[48] and Brandt[49], and also in review articles by Orear[50] and Solmitz[51]. Briefly the function $L(\sigma_k; t)$ is defined by setting $L = G$ but assuming the parameters are the independent variables. The likelihood function L is maximized with respect to the parameters σ by the condition

$$\frac{\partial L}{\partial \sigma_k} = 0,$$

with solutions

$$\sigma = \hat{\sigma}(t_j) .$$

The choice of the statistical functions $\hat{\sigma}$ is determined by maximizing the likelihood for obtaining the experimental result. These results can in principle be obtained analytically but often the calculations prove prohibitive and therefore numerical analysis techniques are used to obtain $\hat{\sigma}(t)$.

Another prescription often used to generate the statistical functions is the least squares method (LSM). Here the parameters are varied to minimize χ^2, a function which represents the weighted sum of the squares of the differences between the data and the predicted value for the data. With χ^2 defined as

$$\chi^2 = \sum_i \frac{(n_i - \bar{n}_i)^2}{\sigma_i^2} ,$$

the minimization equations are

$$\frac{\partial \chi^2}{\partial \sigma_k} = 0 .$$

Here $\bar{n}_i$, containing the dependence on the parameters, is defined as $\bar{n}_i = P_i N$. In this equation, P_i is the probability of getting a count in the ith channel of a histogram and N is the total number of observations. For data distributed normally about a set of mean values $\bar{n}_j$ the likelihood function L becomes

$$L = e^{-(\frac{1}{2}\chi^2)}$$

Thus the two techniques are equivalent for this special case.

As shown in Appendix A the data in this experiment are expected to obey a binomial probability function

$$G(n_k; 0, \bar{N}) = \bar{N}^N e^{-\bar{N}} \prod_k \frac{P^{n_k}}{n_k!}$$

with

$$P_k = \int_{k\delta - \delta/2}^{k\delta + \delta/2} \mathcal{P}(t; 0_i) dt \quad .$$

G determines L which can be expanded to give

$$L = e^{-\frac{1}{2}\sum_i \frac{(n_i - \bar{n}_i)^2}{\bar{n}_i}} (1 + \dots) \quad .$$

In the asymptotic limit of infinite data the two methods become equivalent if χ^2 is defined as

$$\chi^2 = \sum_i \frac{(n_i - \bar{n}_i)^2}{\bar{n}_i} \quad .$$

The minimization of χ^2 maximizes L. These equations for the binomial distribution of data become

$$\frac{\partial}{\partial \sigma_k} \{\frac{1}{2} \sum_i \frac{(n_i - \bar{n}_i)^2}{n_i} - \sum_i (n_i - \bar{n}_i) - \frac{1}{6} \sum_i \frac{(n_i - \bar{n}_i)^3}{\bar{n}_i^2} +$$

$$\frac{1}{12} \frac{(n_i - \bar{n}_i)^4}{\bar{n}_i^3} + \ldots \} = 0$$

with

$$\langle \frac{(n_i - \bar{n}_i)^k}{\bar{n}_i} \rangle = \bar{n}_i^{1 - k/2} .$$

Terms with odd powers in the numerator are small near the minimum $\hat{\sigma}$ and the first even term is of order $(\sum_i \frac{1}{\bar{n}_i})$. The statistical functions generated by both methods are equivalent when this term can be neglected. The LSM was used to analyze most of the data in this experiment. It is the standard approach used in most lifetime experiments.

Another useful property of the MLM and the LSM distributions is their asymptotic σ dependence. As the number of measurements increases

$$L = e^{-\frac{1}{2} H_{ij} \Delta \sigma_i \Delta \sigma_j}$$

or

$$\chi^2 = H_{ij} \Delta \sigma_i \Delta \sigma_j ,$$

where H_{ij} is the approximate Hessian matrix and $\Delta\hat{\sigma}=(\sigma-\hat{\sigma})$. In the asymptotic limit $\langle\Delta\sigma_i\Delta\sigma_j\rangle$ approaches H^{-1}_{ij} so H^{-1} can be used as a measure of the correlations between and the errors for the parameters.

In summary, the two methods mentioned are asymptotically equivalent. The MLM and the LSM give algorithms for generating statistics and determining the errors and correlations. The average value $\langle H_{ij}\rangle^{-1}$ can be used to examine the expected correlations and errors for any experiment with a given distribution. The LSM has been more widely used for the analysis of data in histograms because algorithms for the solution of the parameter-estimating equations are more readily available. For analysis of actual data the LSM was chosen. In attempting to uncover distorting or systematic effects the MLM is often used.

One troublesome aspect of these methods is the reliance of this framework on the functional form of G. The problem follows from the fact that a set of data serves very poorly as a means of determining which function out of the infinite set of alternative functions could have generated the data. The arguments that restrict the relationship between the parameters and the data are therefore very important. Even if the exact form of the function G is unknown, at least the alternative functions can perhaps be reduced to a small class of possible distributions. The class of alternative

functions for this experiment is obtained by the addition of terms f_α to the initial function to give the possible candidate distribution functions. These terms were known to affect the distribution at some level but were assumed to be negligble at the outset. In testing hypotheses on the form of the distributions, statistical functions like χ^2 are generated whose mean value and standard deviation are known. This function has a high probability of having a value close to its mean value if the correct form for G is used in determining the parameters and some probability of being significantly different for some choices of G that are incorrect. The ability to determine the incorrectness of a hypothesis is of course related to the type of alternatives and the level at which they enter. The primary concern is the effect of the inclusion or absence of these terms on the estimation of the parameters and the errors.

Another aid in establishing the form of the distribution of data is the examination of specific subgroups of data. For example, data may be sorted so as to have varying levels of a certain dependence. An example is the grouping of data according to the incident beam rate. Here the high-rate data will have a substantially increased deadtime term compared to the low-rate data. Previous measurements have relied on these groupings to determine rate-dependent distortions and to extrapolate to a zero rate or undistorted estimate of τ_μ.

To illustrate some of these principles two experiments can be considered. A decay process governed by

$$\rho(t; \tau_\mu) = \frac{1}{\tau_\mu} e^{-t/\tau_\mu}$$

is measured in both, but in experiment 1 each time interval is measured to a high degree of precision while in experiment 2 a digitron timer is used to build finite bin histograms. The data resulting from the two experiments are respectively t_k, with k ranging from 1 to N, N being the total number of observations and n_i with i ranging from 1 to m, m being the number of histogram channels. The MLM results for the estimates for τ are

$$\hat{\tau}_1 = \frac{\sum t_i}{N} \quad .. \quad \langle \Delta \tau_1^2 \rangle = \frac{\hat{\tau}^2}{N} \quad ; \qquad \langle \hat{\tau}_1 \rangle = \tau$$

$$\hat{\tau}_2 = \frac{\sum n_i i \delta}{N} \quad .. \quad \langle \Delta \tau_2^2 \rangle = \frac{\tau^2}{N} \left(1 + \frac{\delta^2}{\tau^2}\right) \quad ; \qquad \langle \hat{\tau}_2 \rangle = \tau$$

with δ as the bin size of the histograms. In the MLM solution for experiment 2 all terms of order $\frac{\delta}{\tau}$ are dropped in the expression for $\hat{\tau}_2$.

Another illustration of the MLM concerns data from experiment 3 measured to arbitrarily high precision and governed by

$$\rho(t; \tau_\mu, A, b) = \frac{A}{\tau_\mu} e^{-t/\tau_\mu} + \frac{2b}{\tau_\mu} e^{-2t/\tau_\mu} .$$

The second term is a standard deadtime term included to account for the finite response time of detectors and logic. Using the MLM, one obtains

$$0 = \sum_i t_i - N\tau_\mu + \frac{2b}{A} \sum_i t_i e^{t/\tau_\mu} + \dots$$

Applying the approximations

$$\sum_j f(t_j) \simeq N\langle f(t)\rangle \quad \text{and} \quad A \simeq 1,$$

we find

$$\hat{\tau}_\mu = \sum_i \frac{t_i}{N} + \frac{1}{2}b\,\tau_\mu.$$

Comparing this result with that of experiment 1 one can identify a systematic deviation if experiment 1 were truly to follow the distribution of experiment 3. The value $\hat{\tau}_\mu$ would then be systematically too low by a factor $\frac{1}{2}b\,\tau_\mu$.

Examination of $\langle H_{ij}\rangle^{-1}$ for experiment 3 gives two results: firstly, for the case when the amplitude of the deadtime term is completely unknown, and, secondly for the case when b has been previously determined to arbitrarily high precision. The expected variances are then

$$\langle \Delta\hat{\tau}^2\rangle_{b\ \text{known}} \simeq \frac{\tau^2(1+b/4)}{N}$$

$$\langle \Delta\hat{\tau}^2\rangle_{b\ \text{unknown}} \simeq 4\frac{\tau^2}{N}(1-2b)$$

Thus, if there is a deadtime term in the data and the magnitude of b is known, one gets a very small increase in the overall error. If b must be determined by estimating its value as some function of the data the correlation between $\mathcal{T}_\mu$ and b results in a fourfold increase in $\Delta\hat{T}^2$. The lack of information on the value of b permits a range of possible values for it and subsequently a larger range of acceptable values for $\hat{T}$.

The effects of additional terms in the distribution function appear in one of three categories: 1) expected and fully known, 2) expected but, of necessity, measured along with other parameters or 3) completely unsuspected and causing systematic deviation in the estimation of parameters. Experiments are not completely vulnerable to the third category. As stated earlier, test functions can be generated which can verify to some degree the original assumptions on the form for the data distribution. χ^2 as defined above is one such function.

The χ^2 distribution can be found in a book by Bevington.[52] Typically a value of χ^2 between $m \pm \sqrt{m}$ signifies that the data are distributed according to the assumed distribution. χ^2 tests do not strictly rule out other alternative dependences, however. Suppose distribution 1 was the true distribution

$$\rho_1 = \rho_0 + \rho_b$$

while alternate distribution 2 was

$$\rho_2 = \rho_0 + \sum_i a_i \zeta_i$$

the ζ_i being some set of completely defined functions with amplitudes a_i. Certainly the data could be fit with reasonable χ^2 for both distributions if enough terms were included in distribution 2.

If both of these distributions pass the χ^2 test one can ask about the predictions for the parameters. In general the predictions will be different. The parameters for case 2 may be over- or under-estimated with the differences being compensated for by the values for the a_i's. The errors involved in the parameter estimation should increase. The ability to absorb some of the fluctuation of data by a variation in parameters is reflected in the correlation between the parameters. As long as the errors are increased for case 2 such that the prejudices involved are smaller than the increase in error, the systematic effects are below the statistical error. The main disadvantage in this technique is then the increase in uncertainty in the estimated values. There is no guarantee, however, that a large enough increase in the error has occurred.

A more reliable approach to the problem is to esblish the additional terms $\sum_i a_i \zeta_i$ by fitting for the a's in a different region of data. Since any function can be expanded as a

set of functions of the form $\sum_i a_i \varphi_i$, the fitting of these a's in a different region may help to determine the form of the distribution.

In summary the best situation is a distribution of data that is completely understood and that has small correlations among the parameters of interest. If small deviations from the distribution hypothesized are evident one can hope that a parameterization of this deviation as an expansion of functions will yield the correct errors. The approach in this analysis is to establish by reasonable arguments the expected dependence of the data on the parameters, along with alternative dependences that have been suppressed by experimental safeguards (e.g. magnetic field reduction). The correctness of these assumptions is verified by χ^2-testing and comparisons of different groupings of data. Any evidence for functional dependences that may have generated significant amounts of data must be included when estimating τ and determining the errors. This is done either by hypothesising and justifying a functional form or by parameterizing this background with some set of coefficients and a basis of functions.

Chapter VI

LIST OF TABLES

REFERENCES

1. F. Sheck, Physics Reports 44C, 187 (1978).

2. N. C. Mukhopadhyay, Physics Reports 30C, (1977).

3. Particle Data Group, Phys. Letters B 11, 1 (1982).

4. A. Sirlin, Rev. Mod. Phys. 50, 573 (1978).

5. F. J. M. Farley et al.,Proc. 1962 Int. Conf. High Energy Phys. CERN (J. Prentki, ed.) p. 415 (CERN, Geneva).

6. R. A. Lundy, Phys. Rev. 125, 1686 (1962).

7. M. Eckhause et al., Phys. Rev. 132, 422 (1963).

8. S. L. Meyer et al., Phys. Rev. 132, 2693 (1963).

9. R. W. Williams and D. L. Williams, Phys. Rev. D6, 737 (1972).

10. J. Duclos, A. Magnon and J. Picard, Phys. Lett. 47B, 491 (1973).

11. M. P. Balandin et al., Zh. Eksp. Teor. Fiz. 67, 1631 (1974).

12. G. Bardin et al., Nucl. Phys. A352, 365 (1981).

13. E. S. Abers and B. W. Lee, Physics Reports 9C, 6 (1973).

14. S. Weinberg, Rev. Mod. Phys. 46, 255 (1974)

15. J. C. Taylor, Gauge Theories of Weak Interactions, (Cambridge University Press, Cambridge, 1976).

16. J. Iliopoulos, CERN Report 76-11, (1976)

17. S. Weinberg, Phys. Rev. Lett. 19, 1264 (1967)

18. A. Salam, in Elementary Particle Physics, edited by N. Svartholm (Almqvist and Wiksells, Stockholm, 1968), p. 367.

19. S. Weinberg, Phys. Rev. Lett. 27, 1688 (1971).

20. A. Sirlin, Phys. Rev. D 22, 971 (1980).

21. TRIUMF User's Handbook (1979).

22. C. J. Oram et al., Nucl. Instr. and Meth. 179, 95 (1981).

23. J. V. Jelley, Cherenkov Radiation and its Applications, (Pergamon , New York, 1958).

24. V. P. Zrelov, Cherenkov Radiation in High-Energy Physics, (Vol. I & II, Keter , Jerusalem, 1970).

25. P. J. Carlson, Nucl. Instr. and Meth. 158, 403 (1979).

26. C. Richard-Serre et al., Nucl. Instr. and Meth. 63, 173 (1968).

27. K. Saito and K. Suga, Nuovo Cimento 11, 600 (1959).

28. A. Bezaguet et al., Nucl. Instr. and Meth. 158, 303 (1979)

29. G. Cosme, S. Jullian, and J. LeFrancois, Nucl. Instr. and Meth. 70, 20 (1969).

30. R. Burman of LASL kindly supplied us with the Kydex which is described in the following reference.

    S. Willis, Ph. D. Thesis, Yale University (1979).

31. G. Marshall, Ph.D. Thesis, Unversity of British Columbia (1980).

32. G. J. Sellers, METGLAS Alloys: An answer to low frequency magnetic shielding, IEEE International Symposium on Electromagnnetic Compatibility, Seattle, WA (1977).

33. P. W. Percival et al., Chem. Phys. Lett. 39, 333 (1976).

34. Private Communication W. J. Kossler

35. Measurements courtesy of M. Conradi and P. Kuhns, College of William and Mary.

36. Preliminary testing was done at the National Bureau of Standards electron Linac . We are grateful to Dr. S. Penner and the NBS staff for their help in these tests.

37. I. Navon et al., Phys. Rev. Lett. 42, 1465 (1979).

38. D. Ashery et al., Phys. Rev. 23C, 2173 (1981).

39. G. Amsel, R. Bosshard, and C. Zajde, Nucl. Instr. and Meth. 71, 1 (1969).

40. G. P. Lamaze et al., Nucl. Instr. and Meth. 123, 403 (1975).

41. S.Schlaer, Los Alamos Informal Report 5511-MS (1974).

42. R. F. Thomas, Los Alamos Informal Report 5404-MS (1973).

43. L. R. Biswell and R. E. Rajala, Los Alamos 5144 (1973).

44. J. Chiba and E. H. Harvey, Jr., MBD Driver Manual, Version 0002, Lawrence Berkely Laboratory, UCB (1979).

45. to be obtained

46. M. Annis et al., Rev. Mod. Phys. 25, 818 (1953).

47. R. Peierls, Proc. Roy. Soc. A149, 467 (1953).

48. M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, VolII, (Hafner, New York, 1961).

49. S. Brandt, Statistical and Computational Methods in Data Analysis, (North-Holland, New York, 1961).

50. J. Orear, U. C. R. L. 8417 (1958).

51. F. Solmitz, Notes on the Least Squares and Maximum Likelihood Methods, Institute for Nuclear Studies Report, University of Chicago.

52. P. Bevington, Data Reduction and Error Analysis for the Physical Sciences, (McGraw-Hill, New York, 1966).

LIST OF FIGURE CAPTIONS

FEYNMAN DIAGRAMS FOR MUON DECAY

a

b

c

d

e

f

g

h

i

j

k

l

m

n
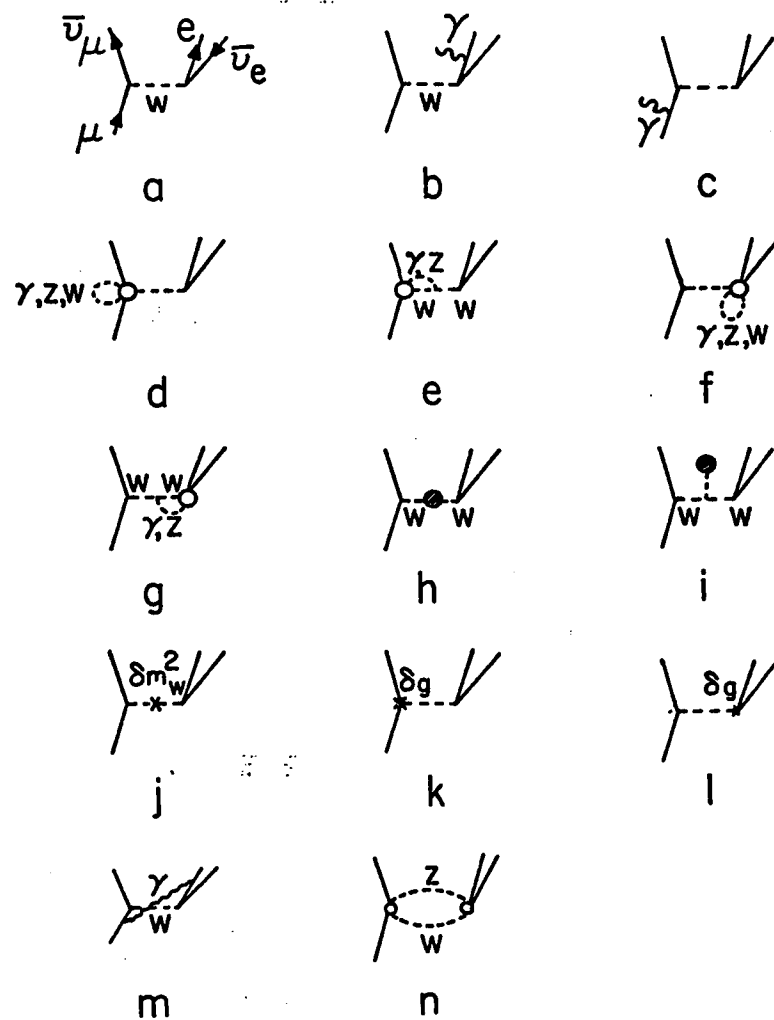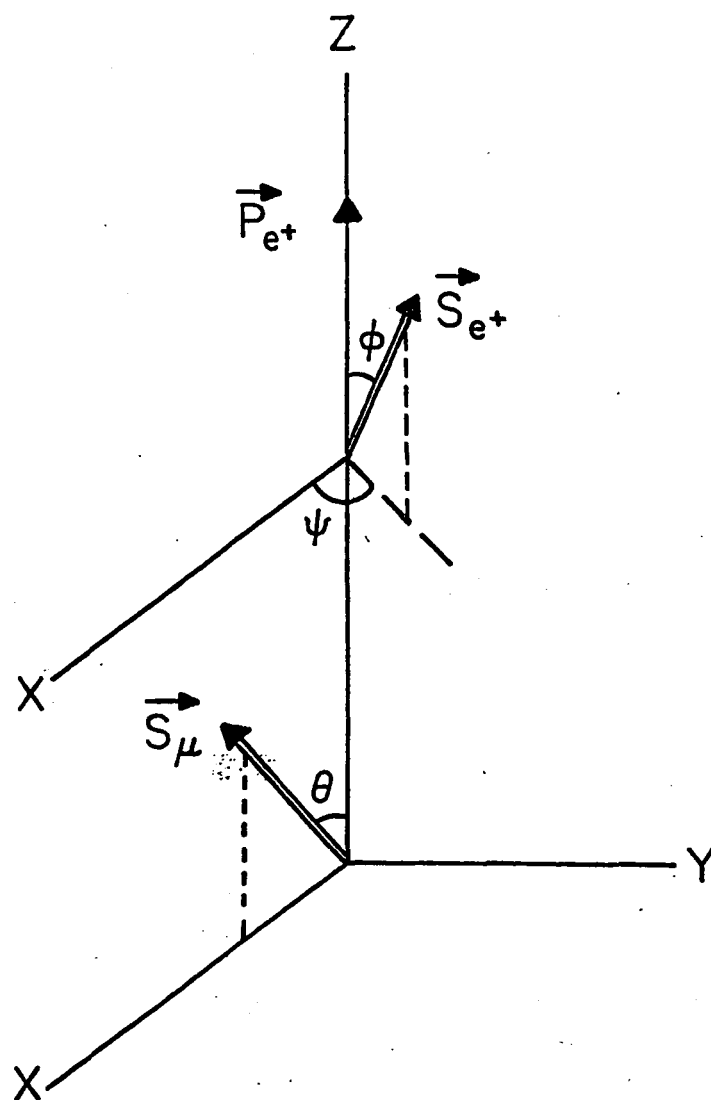
Figure 1

Figure 2

MICHEL SPECTRUM
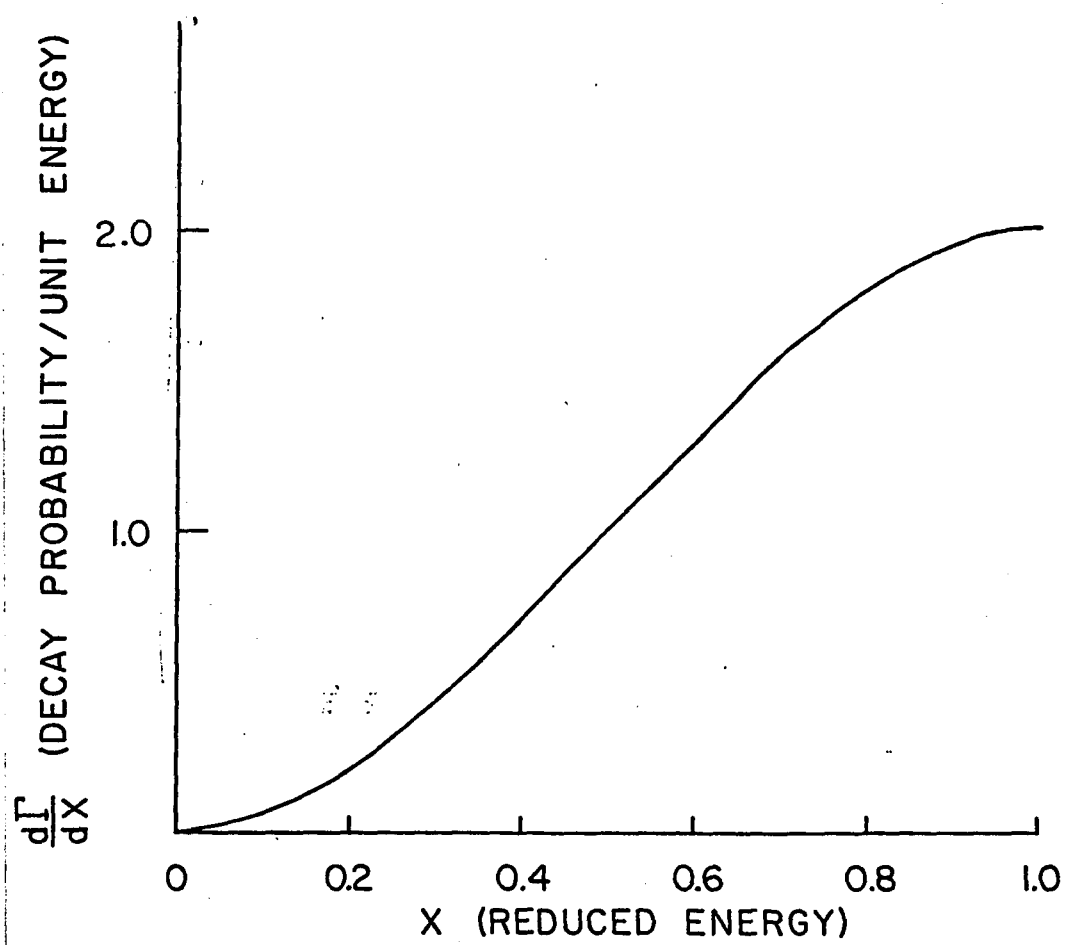
Figure 3
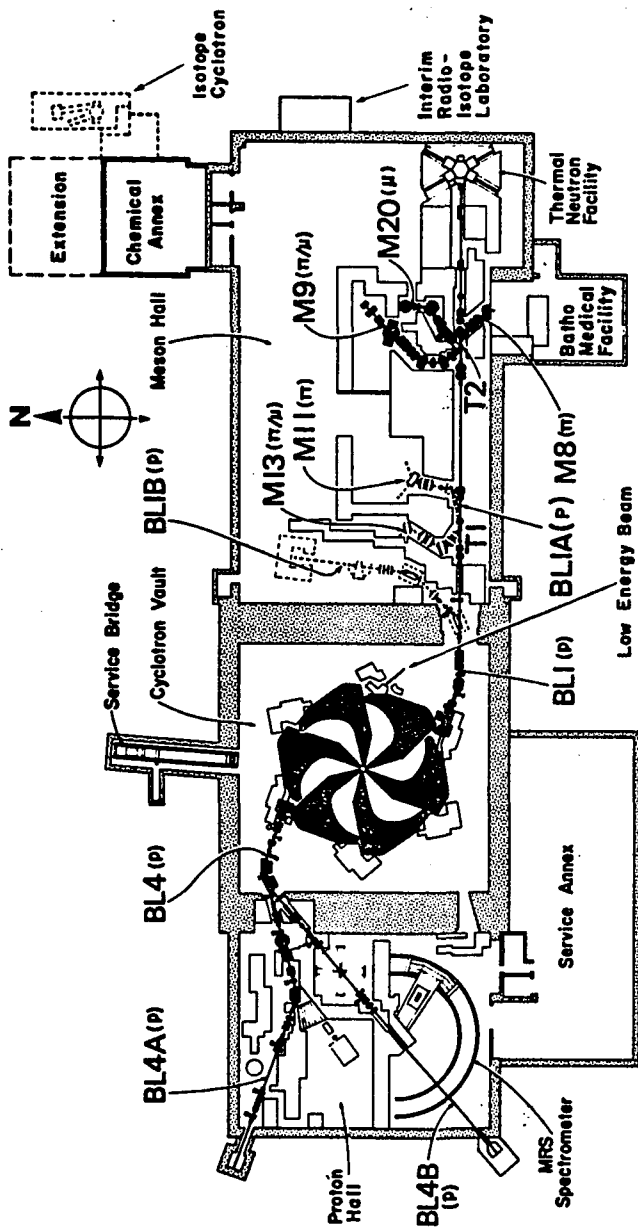
BEAM LINES AND EXPERIMENTAL FACILITIES

TRIUMF

Figure 4

TIME OF FLIGHT
(151 MeV/c)

Figure 5

Figure 6

93
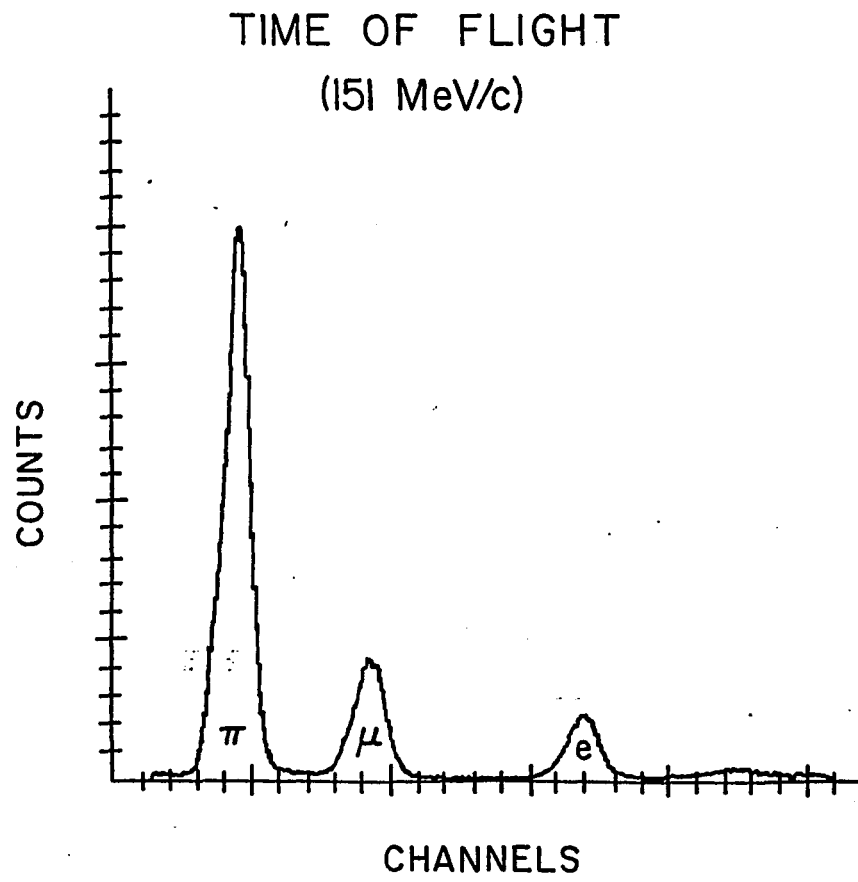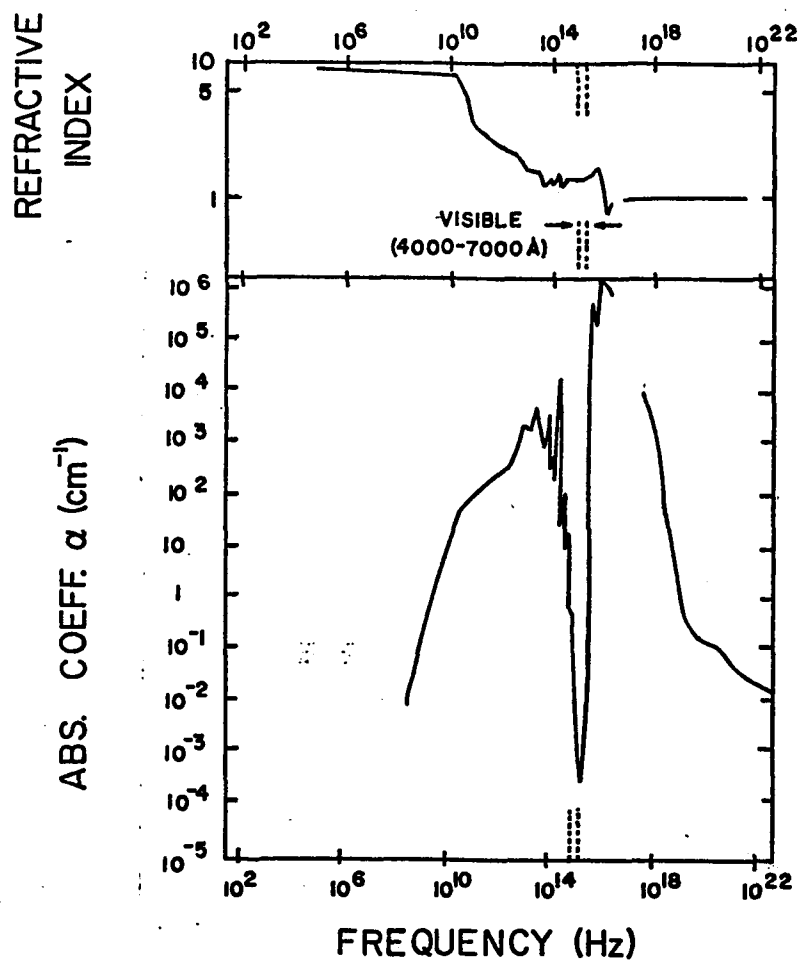
CHERENKOV PHOTONS PRODUCED IN H_2O (2000Å-6000Å)

POSITRON

MUON

muon thr. = 119.8 (MeV/c)

pion thr. = 158.3 (MeV/c)

PION

e thr. = 0.58 (MeV/c)

NUMBER OF PHOTONS (10^4)

MOMENTUM (MeV/c)
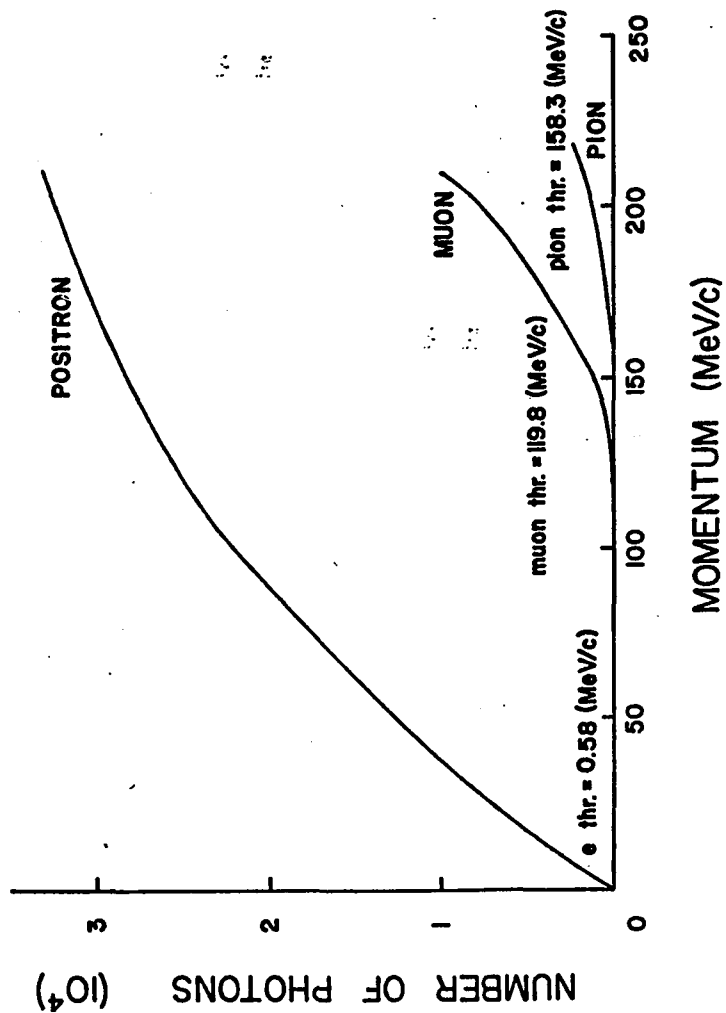
Figure 7

EQUIV. ČERENKOV PRODUCTION
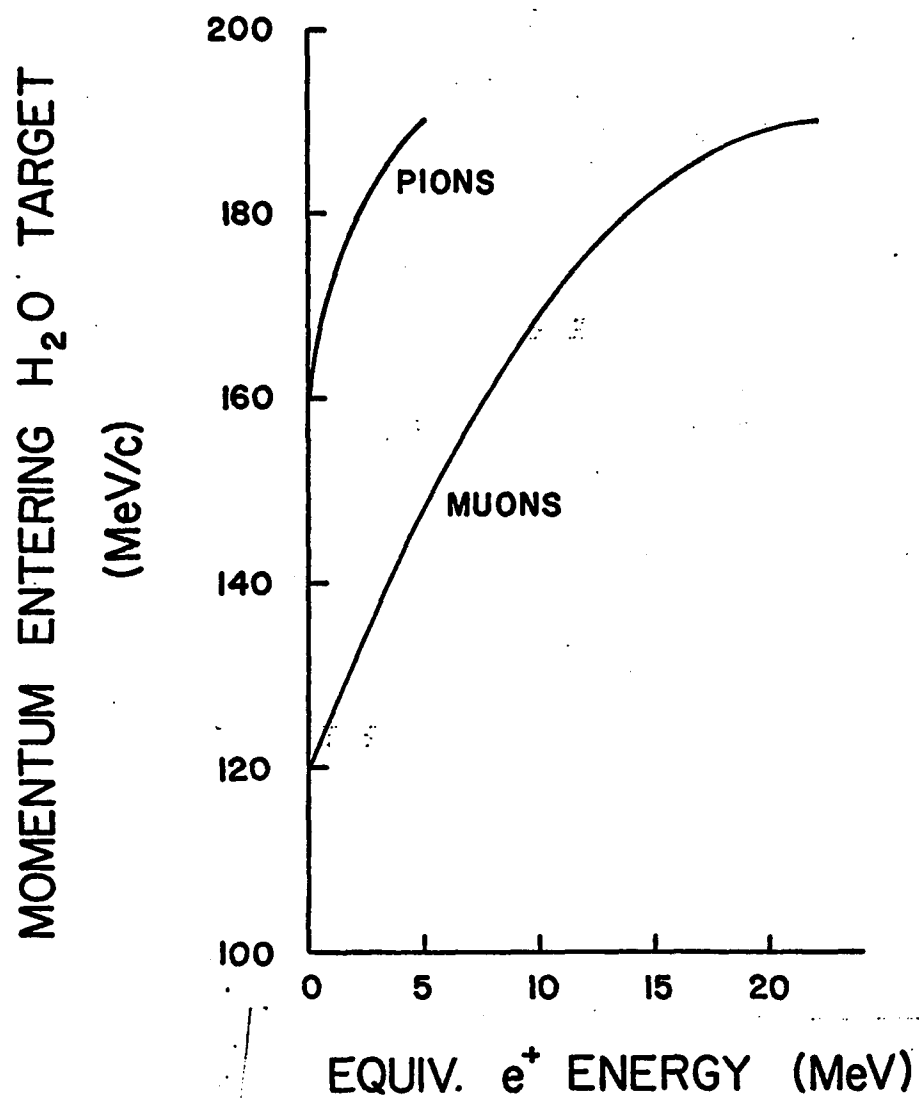
Figure 8

CHERENKOV RESPONSE TO 120 MeV BEAM POSITRONS
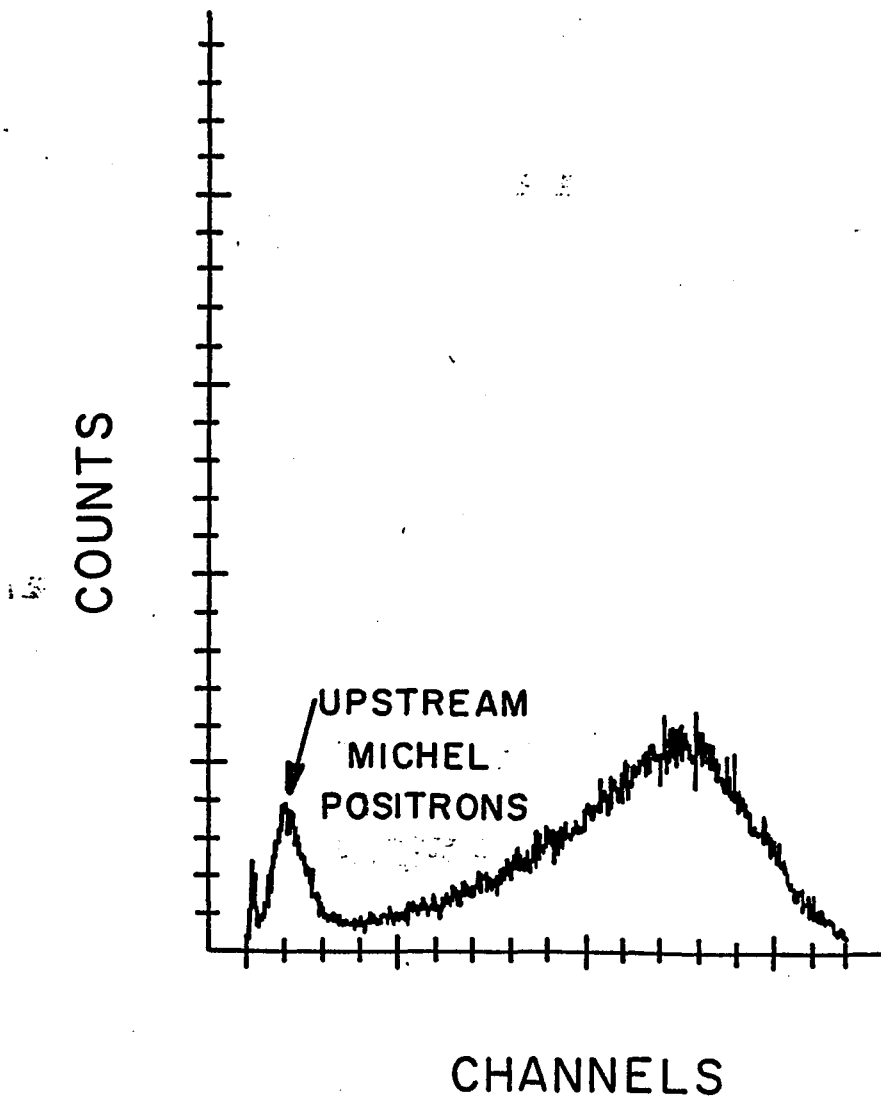
Figure 9

CHERENKOV COUNTER RESPONSE

A   20 MeV beam e$^+$
B   30 MeV beam e$^+$
C   56 MeV beam e$^+$
D   Michel e$^+$ from
     stopped μ^+ decay
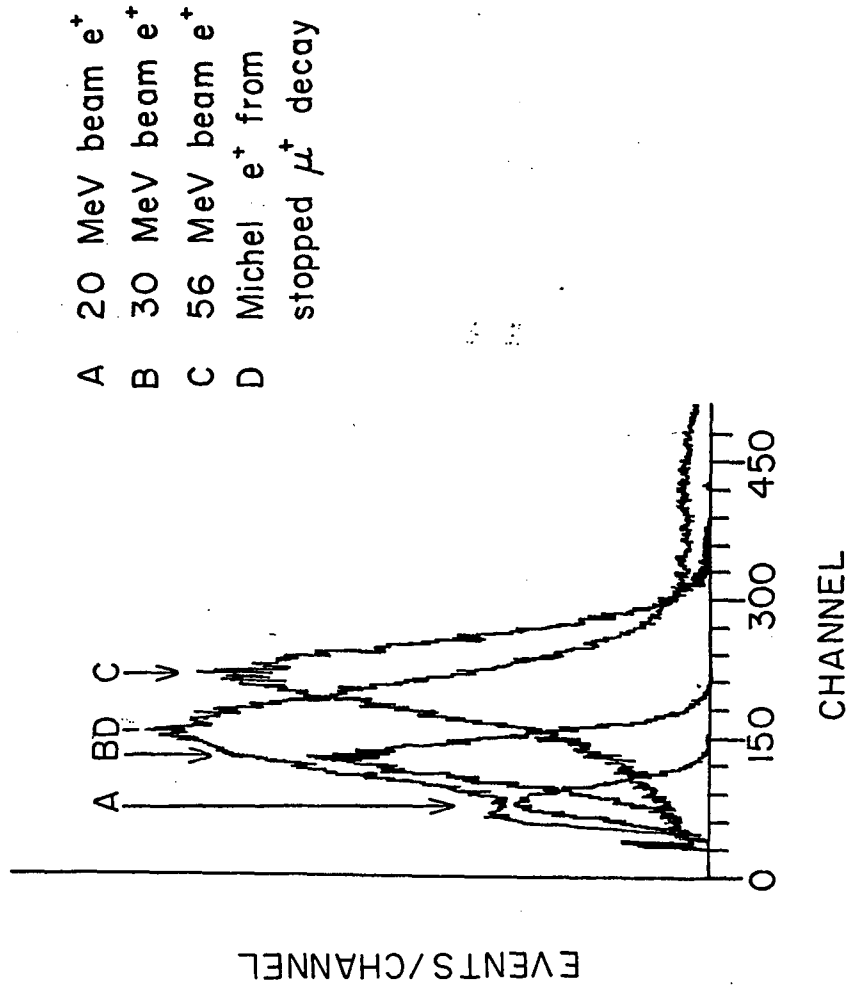
Figure 10

COUNTER ARRAY
(TOP VIEW)

B

RCA
8854
PM TUBE

.32 cm

.16 cm →

WATER AND

WAVE SHIFTER

48 cm

BEAM →

1

2

KYDEX
LINER

48 cm

FLEXIBLE
MAGNETIC
SHIELD

PLEXI-
GLASS
WINDOW

A
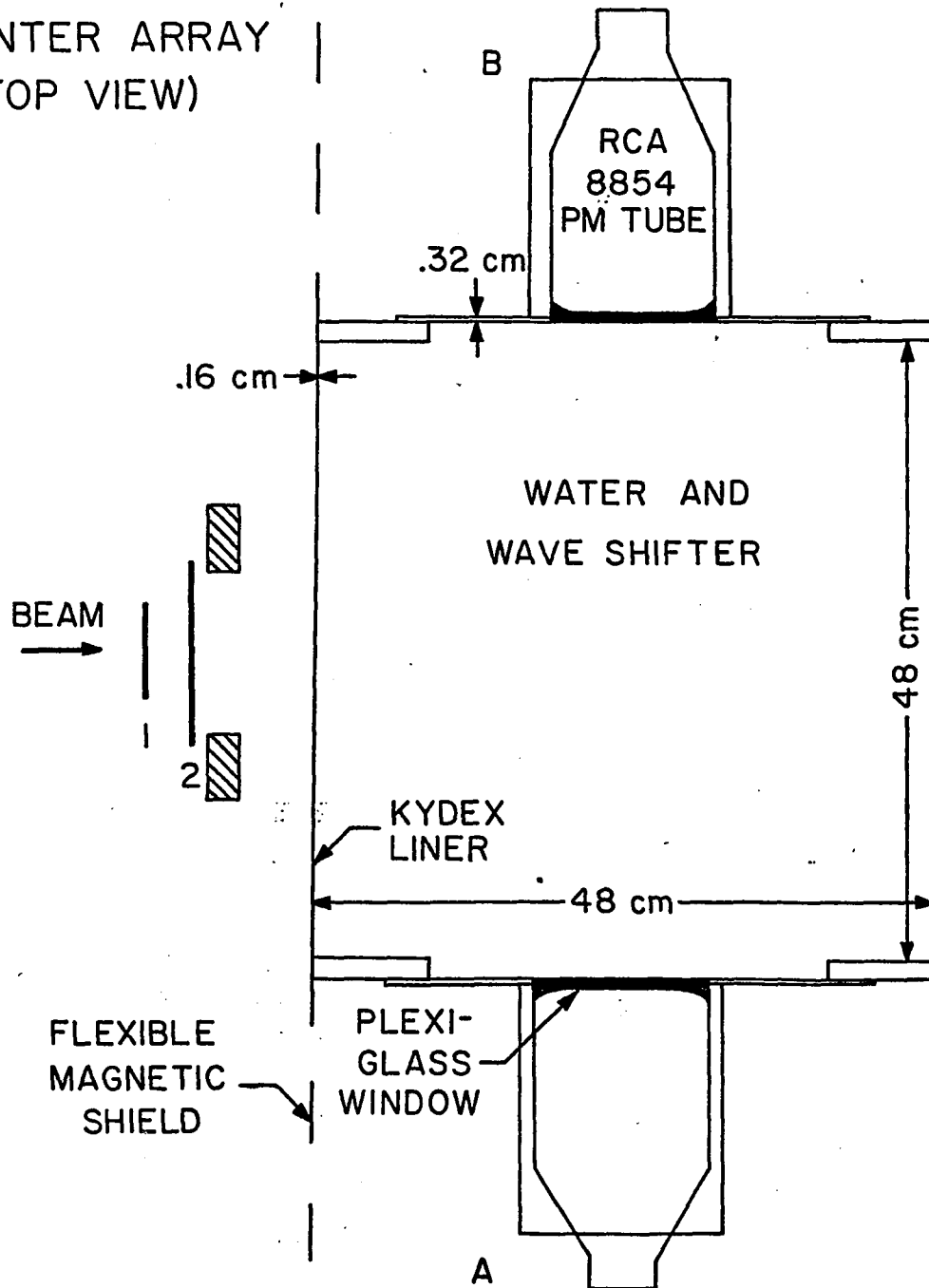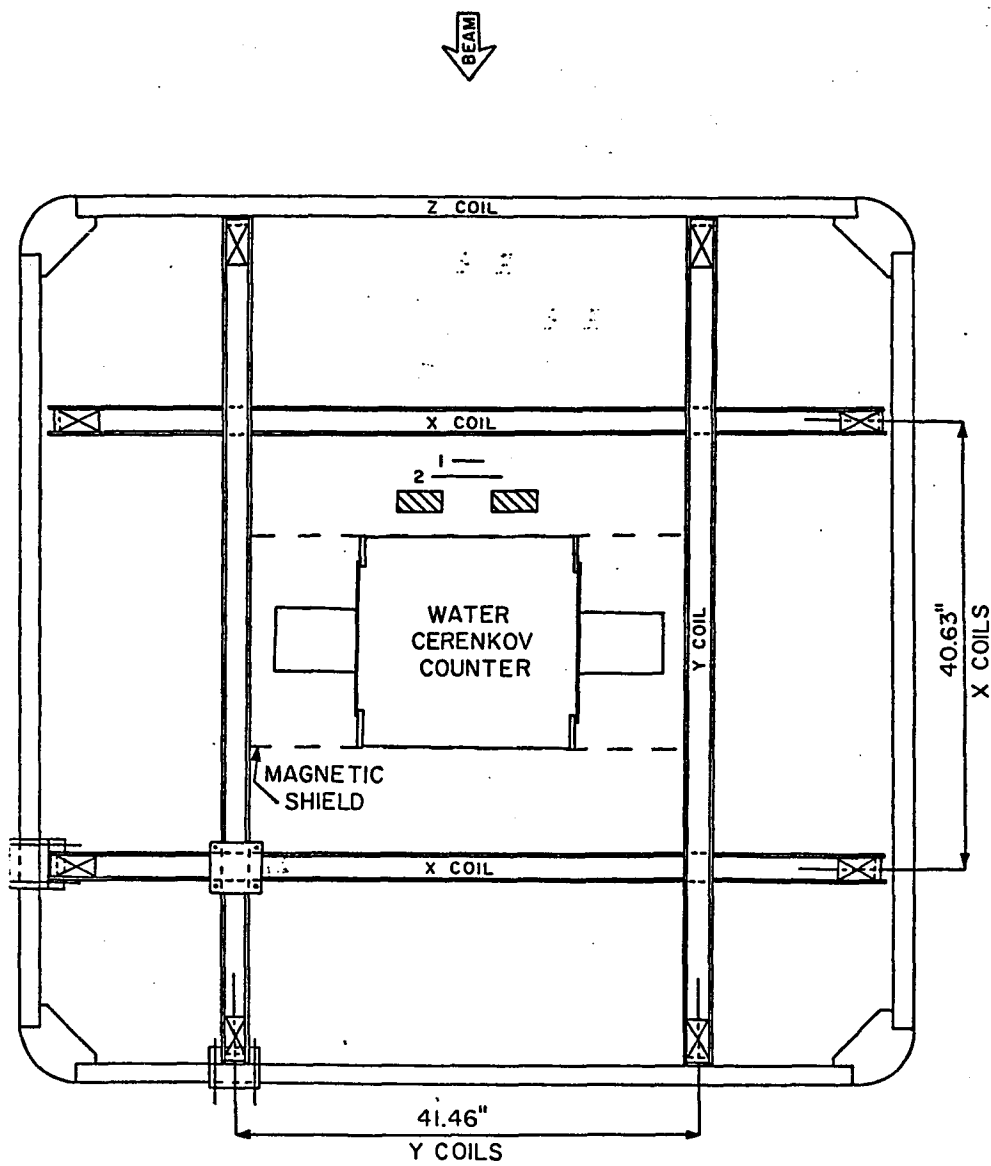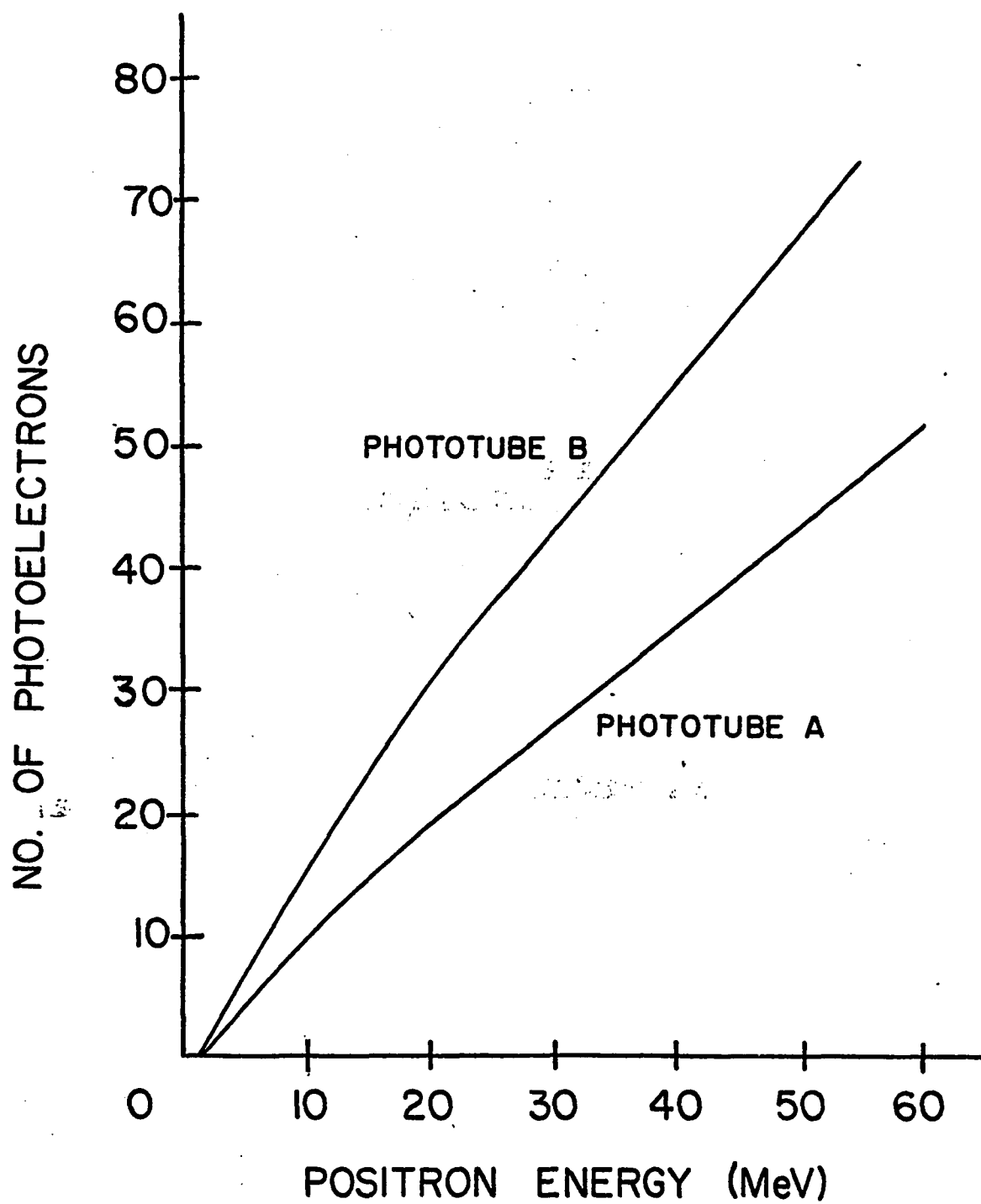
Figure 11

Figure 12

Figure 13

H₂O TARGET ENERGY SPECTRUM
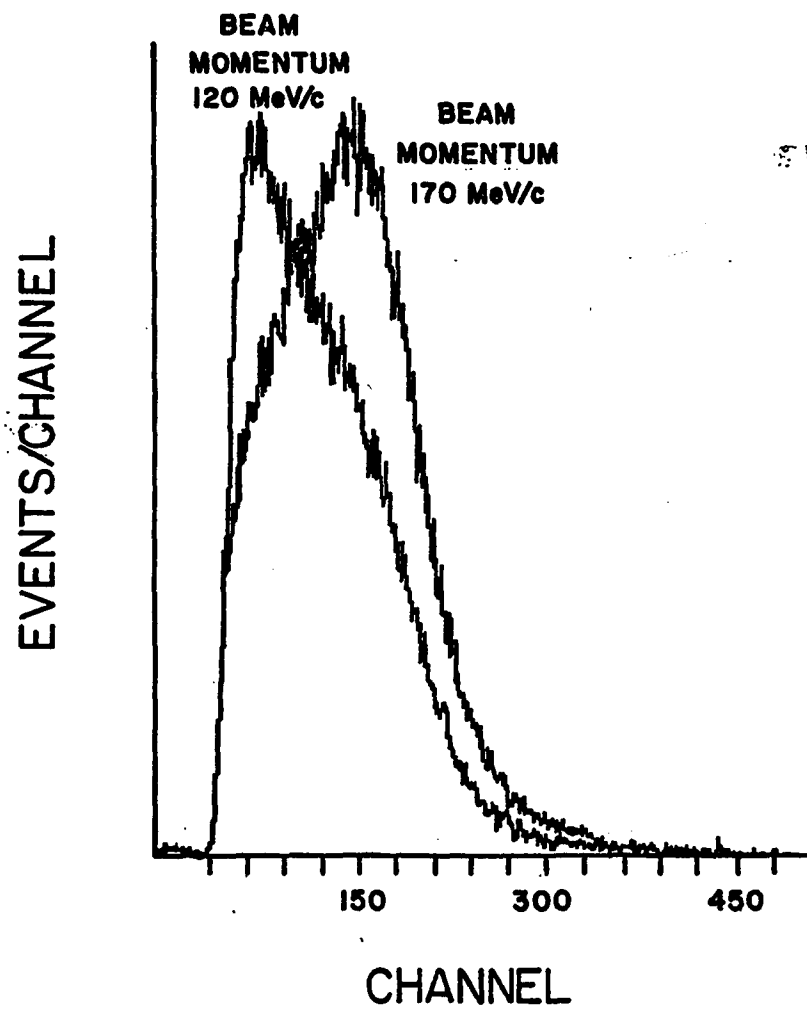
Figure 14
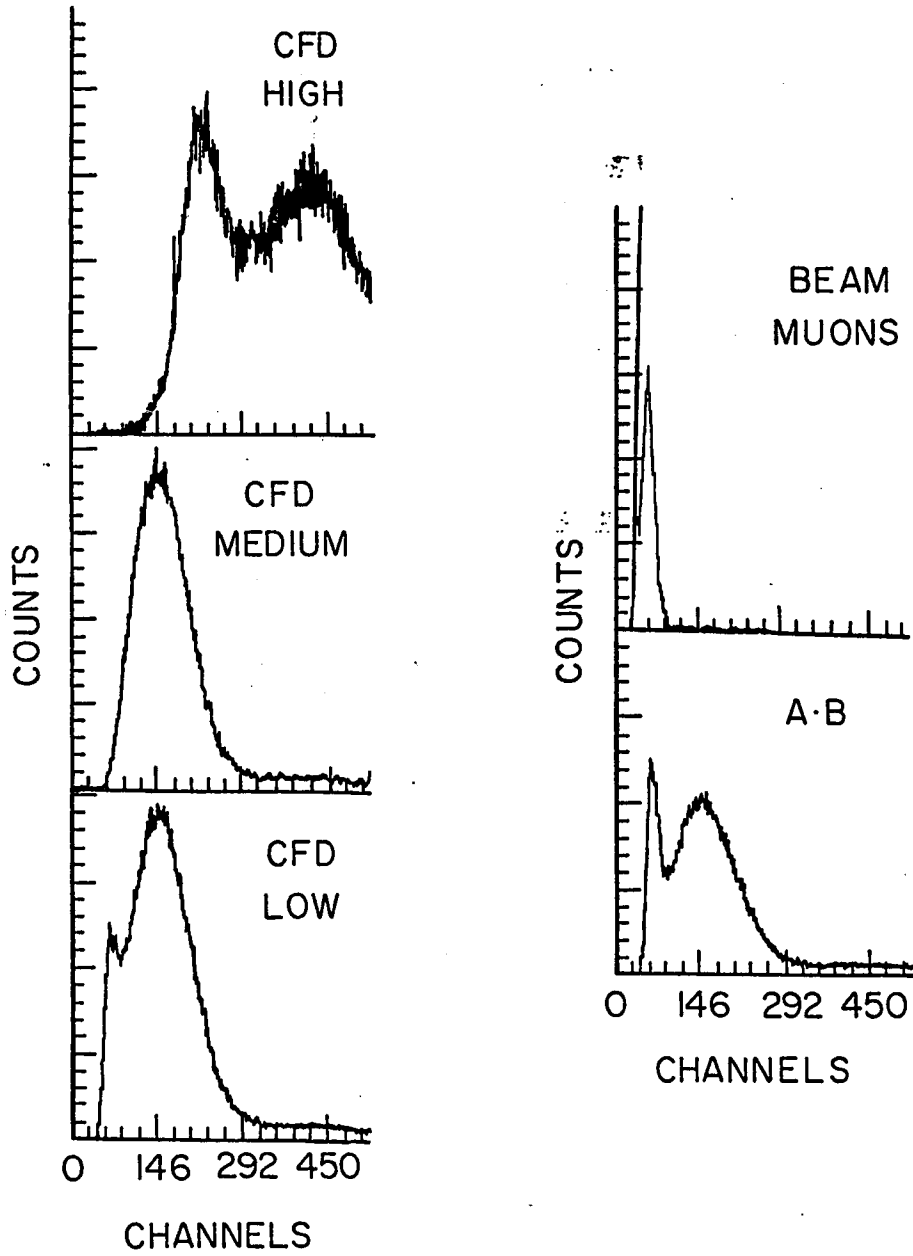
GATED H₂O COUNTER SPECTRA

Figure 15

Figure 16

Figure 17

Figure 18

FO fanout
D   discriminator
PG pile up gate
F   t-bar passive filter
DT deadtime discriminator
SC camac scalar bank
PU camac pattern unit
TM camac trigger module

Figure 19

Figure 20

CHANNEL A

CHANNEL A

CHANNEL B

CHANNEL B

DISTRIBUTION OF
WEIGHTED LIFETIMES

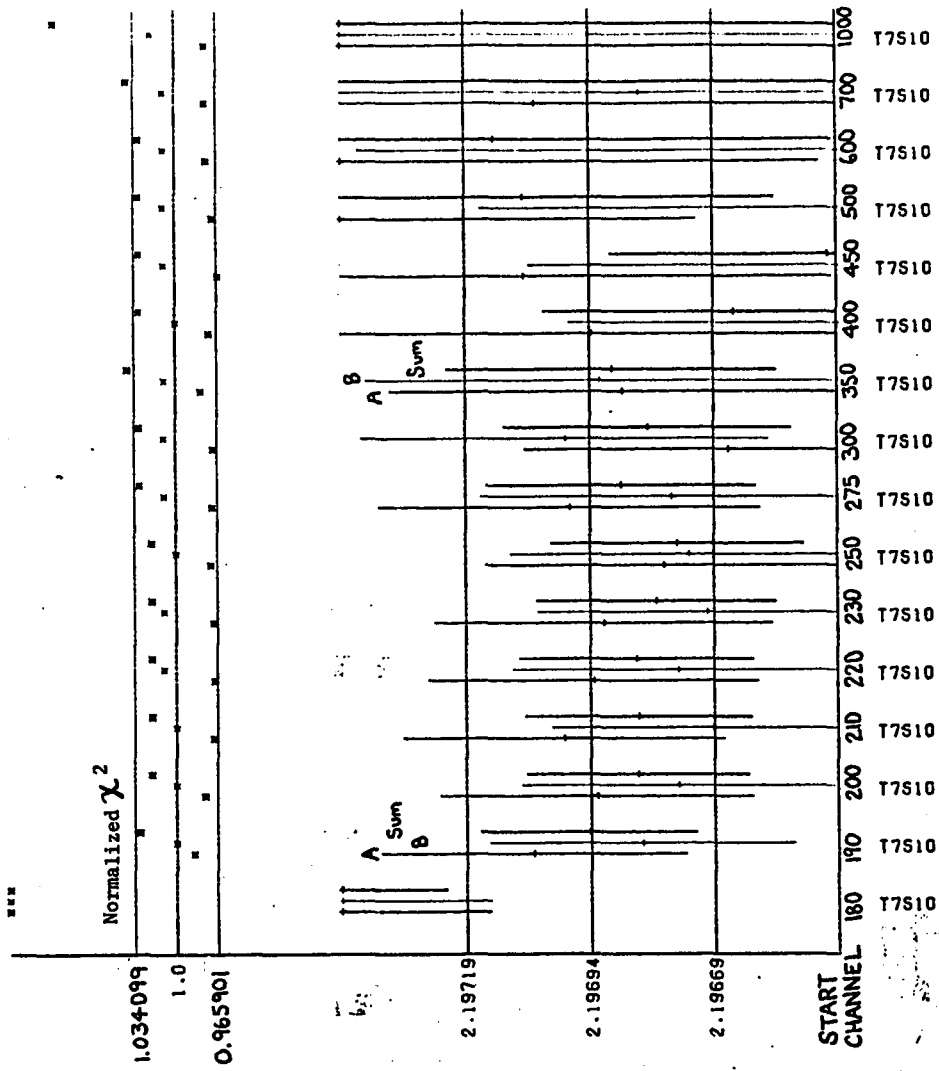$$\left(\frac{\tau - \tau_o}{\sigma} \right)$$

DISTRIBUTION OF
NORMALIZED χ^2
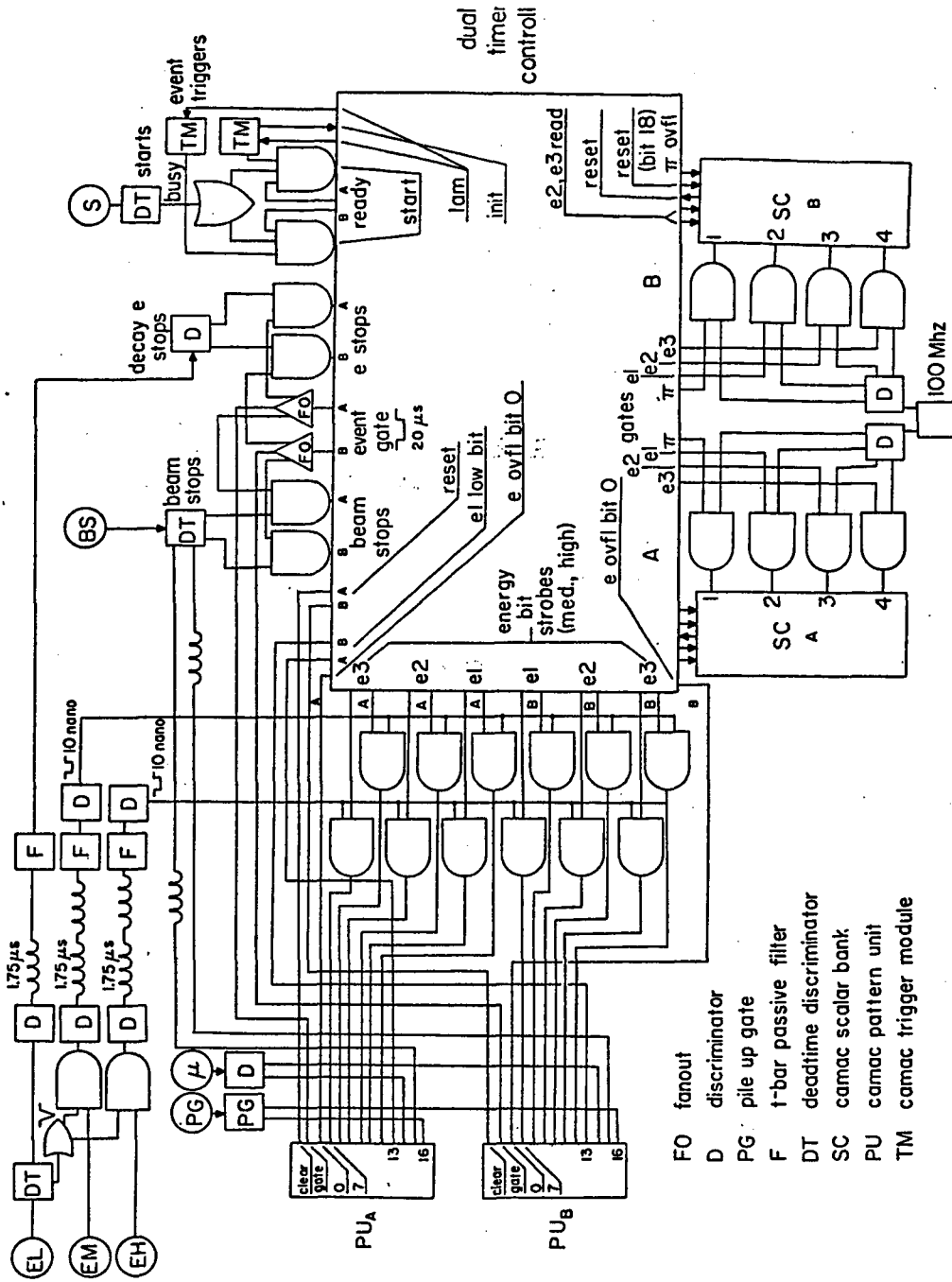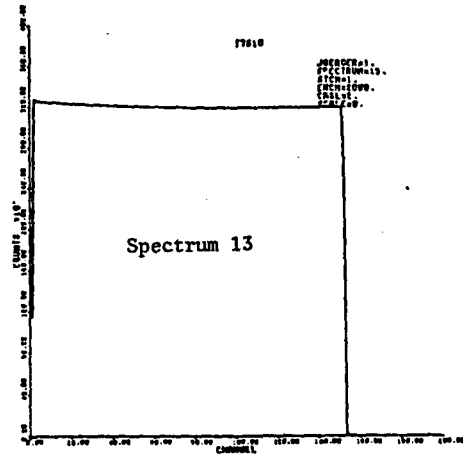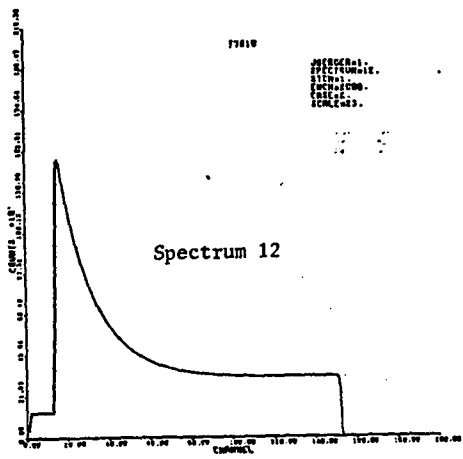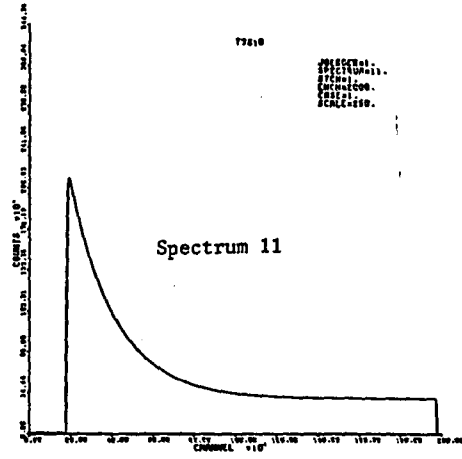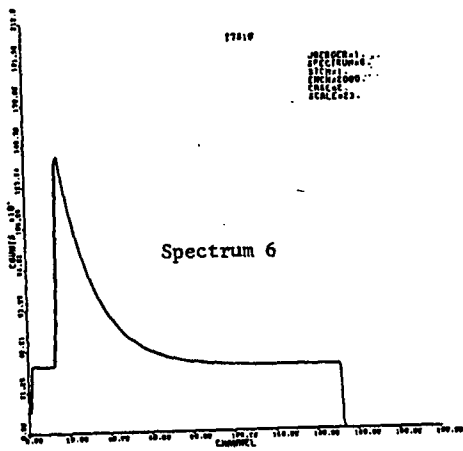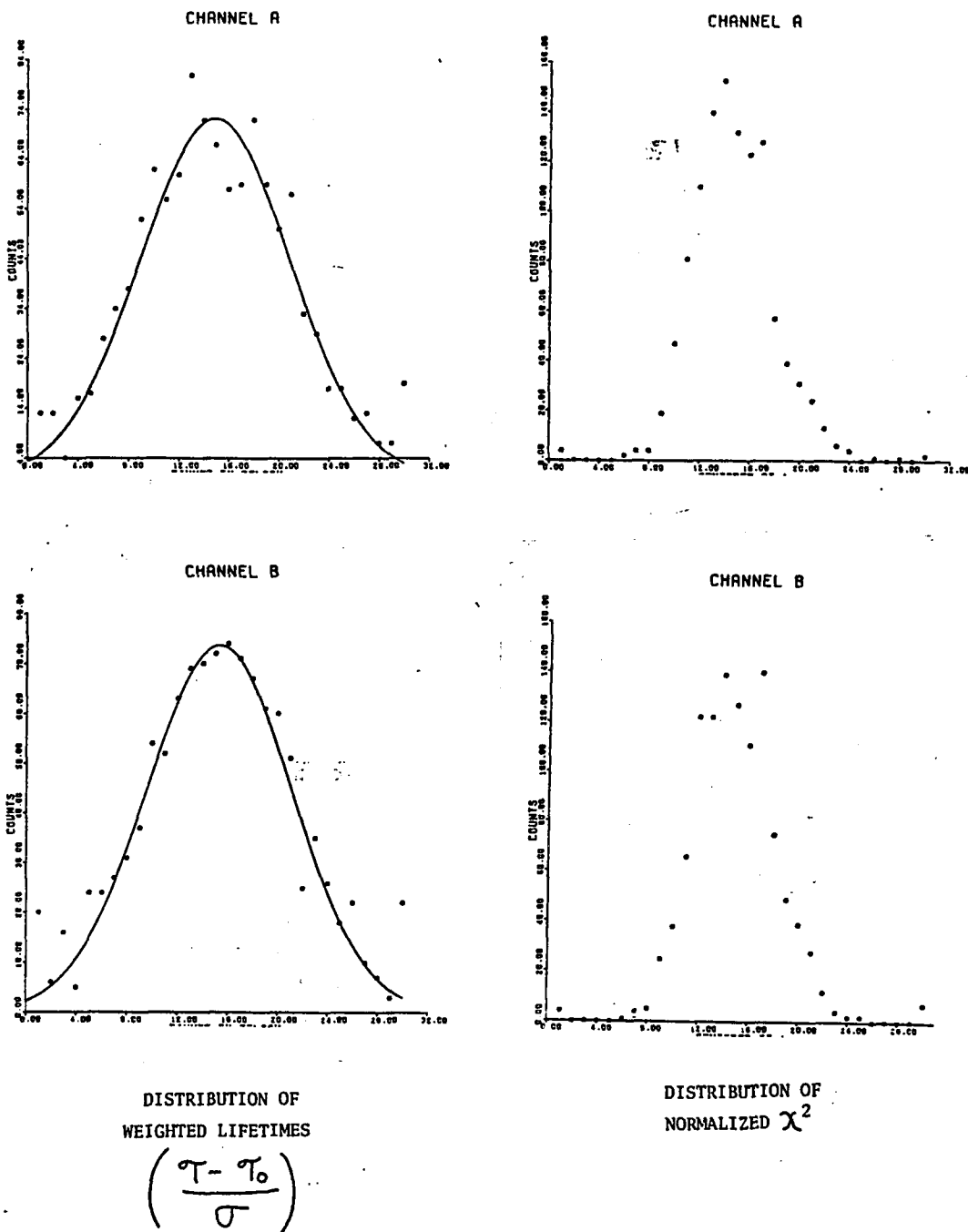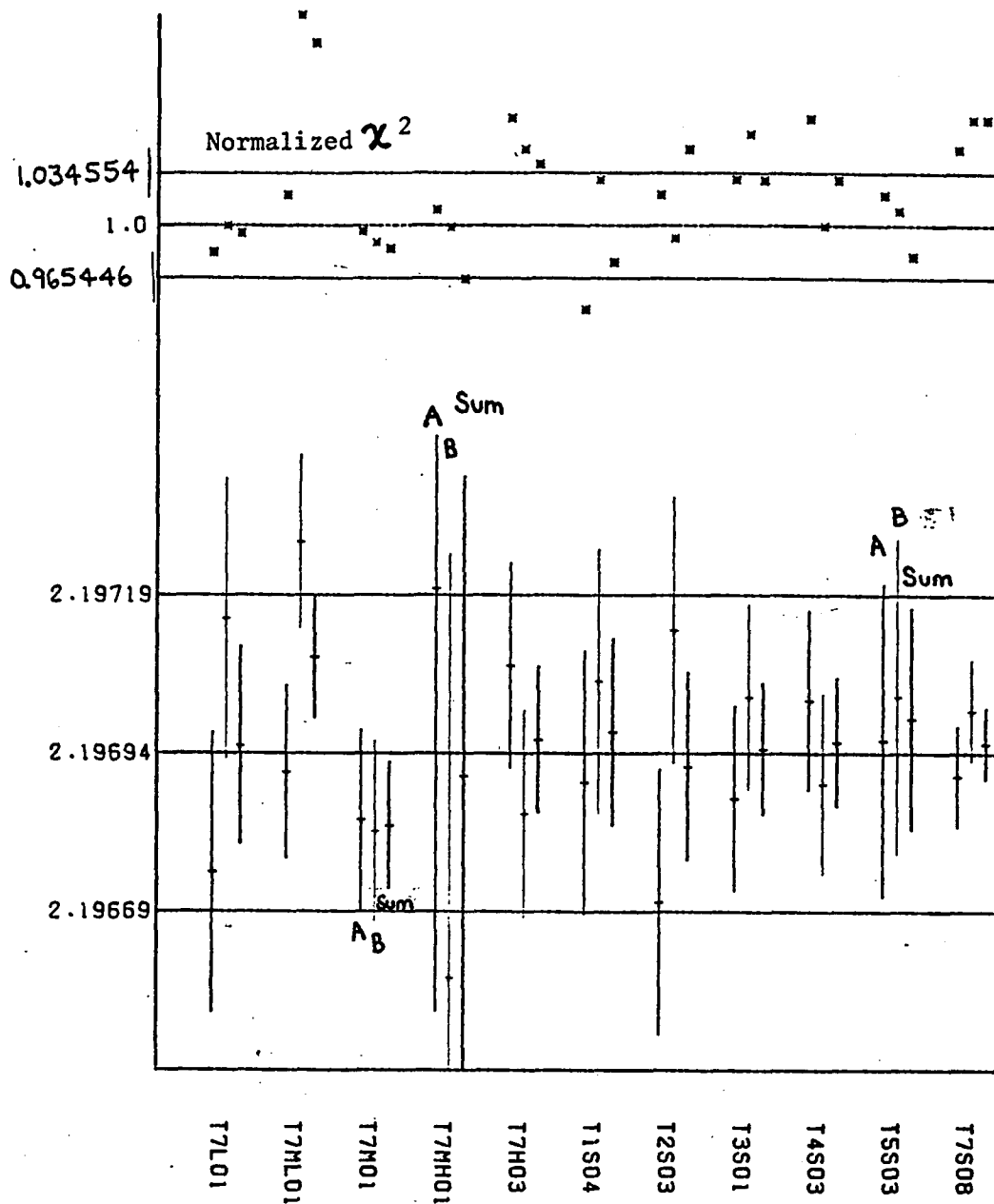
Figure 21

SPECTRA 2

RATE-DEPENDENT CHECK

Figure 22

SPECTRA 2
DIFFERENT RANGE OF CHANNELS
START CHANNEL – TO 1900

Figure 23

SPECTRA 8

DIFFERENT RANGE OF CHANNELS

START CHANNEL - TO 1900

Figure 24
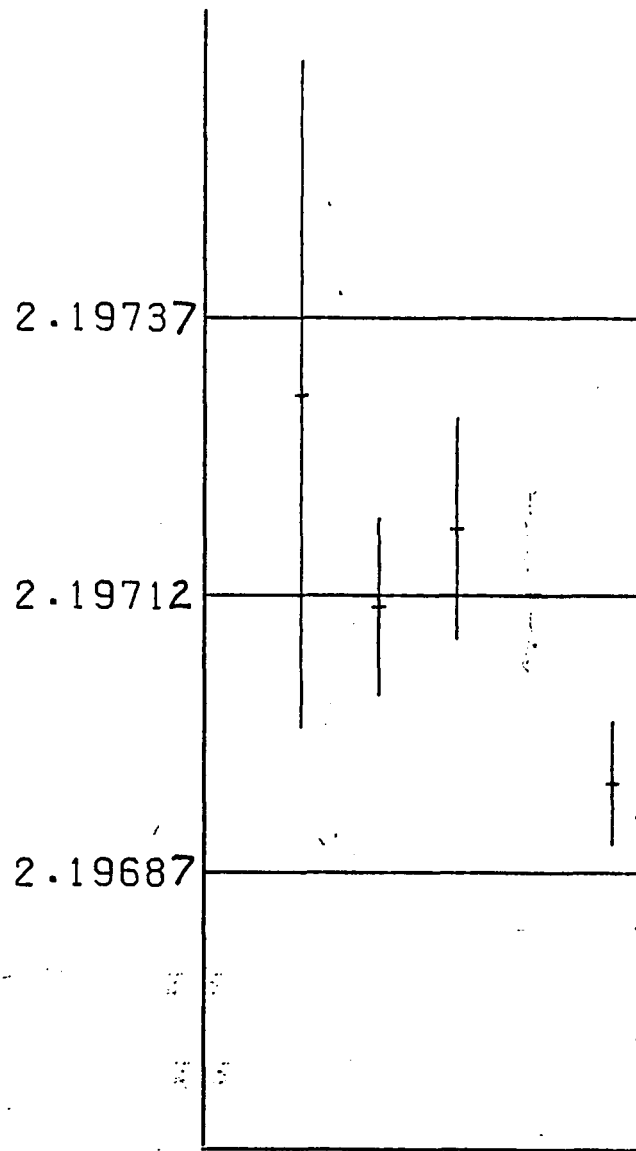
COMPARISON OF MEASURED

VALUES OF τ_μ^+

Figure 25