

**PARTICLE ASTROPHYSICS
ATOMIC PHYSICS
AND GRAVITATION**

Sponsored by:

- . CNRS (*Centre National de la Recherche Scientifique*)
- . IN2P3 (*Institut National de Physique Nucléaire et de Physique des Particules*)
- . CEA (*Commissariat à l'Energie Atomique*)
- . NSF (*National Science Foundation*)

XIVth Moriond Workshop

Villars sur Ollon, Switzerland - January 22-29, 1994

Particle Astrophysics, Atomic Physics and Gravitation

Series : Moriond Workshops

ISBN 2-86332-161-7

Copyright 1994 by Editions Frontières

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

EDITIONS FRONTIERES

B. P. 33

91192 Gif-sur-Yvette Cedex - France

Printed in Singapore

Proceedings of the XXIXth RENCONTRE DE MORIOND

Series : Moriond Workshops

Villars sur Ollon, Switzerland

January 22 - 29, 1994

**PARTICLE ASTROPHYSICS
ATOMIC PHYSICS
AND GRAVITATION**

edited by

**J. Trân Thanh Vân
G. Fontaine, E. Hinds**

EDITIONS
FRONTIERES

C
GERN
BIBLIOTHEQUE

M 81

1994

667
1994

XIVth Moriond Workshop on : Particle Astrophysics, Atomic Physics and Gravitation

was organized by

J. Trân Thanh Vân (*Orsay*)

with the active collaboration of :

E. Adelberger (Seattle)
R. Ansari (Orsay)
A. Blanchard (Meudon)
F. Boehm (Caltech)
G. Chardin (Saclay)
C. Cohen-Tannoudji (Paris)
T. Damour (IHES, Bures/Yvette)
O. Fackler (Livermore)
J. Faller (Boulder)
E. Fischbach (Purdue)
G. Fontaine (Paris)
G. Gerbier (Saclay)
E. Giacobino (Paris)
Y. Giraud-Héraud (Paris)
G. Greene (NBS)
I. Grenier (Saclay)
B. Guiderdoni (Paris)
E. Hinds (Yale)
J. Kaplan (Paris)
B. Kayser (Washington)
R. Pain (Paris)
S. Petcov (Sofia/Trieste)
J. Wilkerson (Los Alamos).

94' RENCONTRES DE MORIOND

The XXIXth Rencontres de Moriond were held in 1994 in Villars-sur-Ollon, Switzerland, and in Méribel les Allues, Savoie, France.

The first meeting took place at Moriond in the French Alps in 1966. There, experimental as well as theoretical physicists not only shared their scientific preoccupations but also the household chores. The participants in the first meeting were mainly French physicists interested in electromagnetic interactions. In subsequent years, a session on high energy strong interactions was also added.

The main purpose of these meetings is to discuss recent developments in contemporary physics and also to promote effective collaboration between experimentalists and theorists in the field of elementary particle physics. By bringing together a relatively small number of participants, the meeting helps to develop better human relations as well as a more thorough and detailed discussion of the contributions.

This concern of research and experimentation of new channels of communication and dialogue which from the start animated the Moriond meetings, inspired us to organize a simultaneous meeting of biologists on Cell Differentiation (1970) and to create the Moriond Astrophysics Meeting (1981). In the same spirit, we have started this year a new series on Condensed Matter Physics. Common meetings between biologists, astrophysicists, condensed matter physicists and high energy physicists are organized to study the implications of the advances in one field into the others. I hope that these conferences and lively discussions may give birth to new analytical methods or new mathematical languages.

At the XXIXth Rencontres de Moriond in 1994, four physics sessions, one astrophysics session and one biology session were held :

- * January 22-29 "Particle Astrophysics, Atomic Physics and Gravitation"
"Coulomb and Interference Effects in small electronic structures"
- * March 12-19 "Electroweak Interactions and Unified Theories"
"Clusters of Galaxies"
- * March 19-26 "QCD and High Energy Hadronic Interactions"
" Rencontre de Biologie - Méribel "

VI

I thank the organizers of the XXIXth Rencontres de Moriond :

- E. Adelberger, R. Ansari, A. Blanchard, F. Boehm, G. Chardin, C. Cohen-Tannoudji, T. Damour, O. Fackler, J. Faller, E. Fischbach, G. Fontaine, G. Gerbier, E. Giacobino, Y. Giraud-Héraud, G. Greene, I. Grenier, B. Guiderdoni, E. Hinds, J. Kaplan, B. Kayser, R. Pain, S. Petcov and J. Wilkerson for the session on Particle Astrophysics, Atomic Physics and Gravitation,

- H. Bouchiat, D. C. Glatli, D. Maily and M. Sanquer for the Condensed Matter Physics session,

- P. Binetruy, A. Blondel, R. Cahn, G. Coignet, L. Fayard, P. Fayet, J.-M. Frère, L. Krauss, L. Moscoso, C. Savoy and C. Verzegnassi for the Leptonic session,

- F. Durrel, A. Mazure and S. White for the Astrophysics session,

- P. Aurenche, E. Berger, A. Capella, D. Denegri, L. Montanet, B. Pietrzyck, D. Schiff, and C. Voltolini for the Hadronic session,

- A. Adoutte, M. Fellous, P. Goldstein, M. A. Gransbastien, J. C. Kader, M. Solignac, P. Sonigo, K. Trần Thanh Vân and A. Ullman for the Biology meeting,

and the conference secretaries : G. Ambonati, L. Besson, V. Chopin, S. Desbarbieux, C. Douillet, D. Josephson, A. Lecoeur, A. S. Lécuyer, F. Lefevre, D. Levêque, L. Norry, F. and N. Osswald, A. M. Perrin, J. Raguideau, R. Scattergood, V. Sobek and F. Warin.

I am also grateful to Mrs S. Müller, Ms B. Gautron, Mr C. Dallery, Ms M. Favet, Mr P. Bernard Granger, and Mr R. Theveniot who contributed through their hospitality and cooperation to the well-being of the participants enabling them to work in a relaxed atmosphere.

These Rencontres were sponsored by the Centre National de la Recherche Scientifique (SPM and FP), the Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) and by the Commissariat à l'Energie Atomique (DAPNIA). The workshop on Particle Astrophysics, Atomic Physics and Gravitation was also sponsored by the National Science Foundation. I would like to express my thanks to their encouraging support.

I sincerely wish that a fruitful exchange and an efficient collaboration between the physicists, the astrophysicists and the biologists will arise from these Rencontres as from the previous ones.

J. Trần Thanh Vân

Contents

<i>Foreword</i>		V
<i>I. HIGH ENERGY ASTROPHYSICAL PHENOMENA</i>		
J. Paul	Overview of the SIGMA Results.	3
M. G. Baring	Gamma-Ray Bursts.	13
I. Grenier	Egret Gamma-Ray Sources.	23
P. Duffy	Cosmic Ray Acceleration and Radio Emission from SNR 1987 A.	33
P.V. Ramanamurthy	Breaking Index of GEMINGA Pulsar.	41
Ch. D. Dermer	Particle Acceleration and Radiation in the Jets of Active Galactic Nuclei.	47
P. Eschstruth	THEMISTOCLE : Results and Prospects.	55
H. Meyer	Status of the HEGRA Experiment and some Early Results.	61
P. Sokolsky	Cosmic Rays above 0.1 EeV - Results from the Fly's Eye Experiment.	67
H. Dai	The HiRes Monte Carlo Simulation.	75
<i>II . DARK MATTER AND COSMOLOGY</i>		
S. Vauclair	On the Abundances of the Light Elements in the Universe.	83
S. Maurogordato	Dark Matter from Galaxy Surveys.	91
D. Pfenniger	Is Galactic Dark Matter Cold Molecular Gas ?	107
B. Guiderdoni	High-Redshift Measurements of Ω_0 .	115
J. W. Moffat	The Age of the Universe, the Hubble Constant and QSOs in a Locally Inhomogeneous Universe.	125

VIII

III. DARK MATTER AND PARTICLE PHYSICS

R. A. Flores	Particle Dark Matter and its Laboratory Detection.	141
L. Mosca	Review of Particle Dark Matter Experiments (Direct Detection).	147
P. Belli	New Limits on Axial Coupled WIMPs with Scintillators. Preliminary Measurements on Axial Coupled WIMPs with Xenon Target at Gran Sasso.	157 161
T. J. Sumner	Latest Limits on Cold Dark Matter Interaction Rates from the UK Underground NaI Detector.	163
M. Kamionkowski	Indirect Detection of WIMPs.	169
L. Bergström	Observing Gamma Lines from Higgsinos in Air Cherenkov Telescopes.	179
G. Forster	Cryogenic Detector Development for Direct Dark Matter Search.	185
TA. Girard	Prospects for Dark Matter Detection with Metastable Superconductors.	191
M. Yu Khlopov	Cosmoparticle Physics of Family Symmetry Breaking.	197

IV. GRAVITATIONAL LENSING

R. D. Blandford	Macro and Microlensing.	211
F. Queinnec	A Search for Low-Mass Objects in the Galactic Halo.	221
M. Szymanski	The Optical Gravitational Lensing Experiment. Discovery of the First Microlensing Events in the Direction of the Galactic Bulge.	229
J. Kaplan	Looking for Microlensing of Stars of the Andromeda Galaxy M31 by Monitoring Pixels.	237

V. NEUTRINOS

J. Kleinfeller	KARMEN : Neutrino Spectroscopy at Isis.	245
B. Maier	The Heidelberg-Moscow $\beta\beta$ Experiment with ^{76}Ge and the Present Status in Double Beta Decay Research.	251

VI. ATOMIC PHYSICS

S. M. Barr	Atomic Electric Dipole Moments and CP Violation.	261
S. K. Lamoreaux	A Test of Time Reversal Symmetry using ^{199}Hg .	271
S. B. Ross	The Berkeley Search for the Electron Electric Dipole Moment.	279
T. E. Chupp	The Search for a Permanent Electric Dipole Moment Using Dual Spin Exchange Pumped ^{129}Xe and ^3He Zeeman Masers.	287
D. Garreta	CP, T and CPT Violation in the Neutral Kaon System at the CPLEAR Experiment.	299
R. Le Gac	Recent Results of the CPLEAR Experiment.	305
M-A. Bouchiat	Parity Violation in Atoms : The Paris Experiment on Cesium.	311
D. Cho	Polarized Beam Measurement of Parity Nonconservation in Atomic Cesium.	325
S. K. Lamoreaux	High Accuracy Measurement of Weak Interaction Induced Optical Activity in Atomic Lead and Thallium.	331
D. DeMille	Nearly-Degenerate Opposite-Parity Levels in Atomic Dysprosium: A Novel System for the Study of Parity Non-Conservation.	339
I. B. Khriplovich	Atomic Physics: QED, Fundamental Symmetries.	347
A.M.Martensson-Pendrill	What do we know about the TL Nucleus from Atomic Spectroscopy ?	353
L. Julien	High Resolution Spectroscopy of Atomic Hydrogen ; A Tool for Metrology	361
C. Zimmermann	High Resolution Spectroscopy of Atomic Hydrogen.	369
K . Abdullah	Trapped Positrons for Ion Cooling and Antihydrogen.	377
J. Sucher	Potential Concepts, Long-Range Forces and Bound-State QED.	393
C. J. Joachain	Atoms in Intense Laser Fields.	403

X

VII. GRAVITATION

Th. Damour	The String Dilaton and the Equivalence Principle.	413
G. L. Smith	New Tests of the Equivalence Principle.	419
J-P. Blaser	Eötvös: A New Concept for Low Disturbance Measurements in Space.	427
J-P. Blaser and Th. Damour	Optimizing the Choice of Materials in Equivalence Principle Experiments.	433
R. C. Ritter	Preliminary Results for Anomalous Spin Coupling to the Dark Matter in our Galaxy.	441
H. Walesch	The Gravitational Force at Mass Separations from 0.6 m to 2.1 m and the Precise Measurement of G.	445
A. Cornaz	Determination of the Gravitational Constant with a Lake Experiment.	453
S. Buchman	Applications of Superconductivity to Gravitational Experiments in Space.	461
C. W. F. Everitt	Prospects for Measuring Gamma to a Part in 10^5 in the Gravity Probe B Experiment.	467
C. O. Escobar	Quantum Limits for Measurements on Macroscopic Bodies : a Decoherence Analysis.	473
Y. Fujii	New Proposed Experiment on Time-Variability of the Fine-Structure Constant.	479
A. Krolak	On the Detectability of the Post-Newtonian Effects in the Gravitational Wave Emission of a Coalescing Binary.	485

VIII. COMMON SESSION WITH MESOSCOPIC SYSTEMS

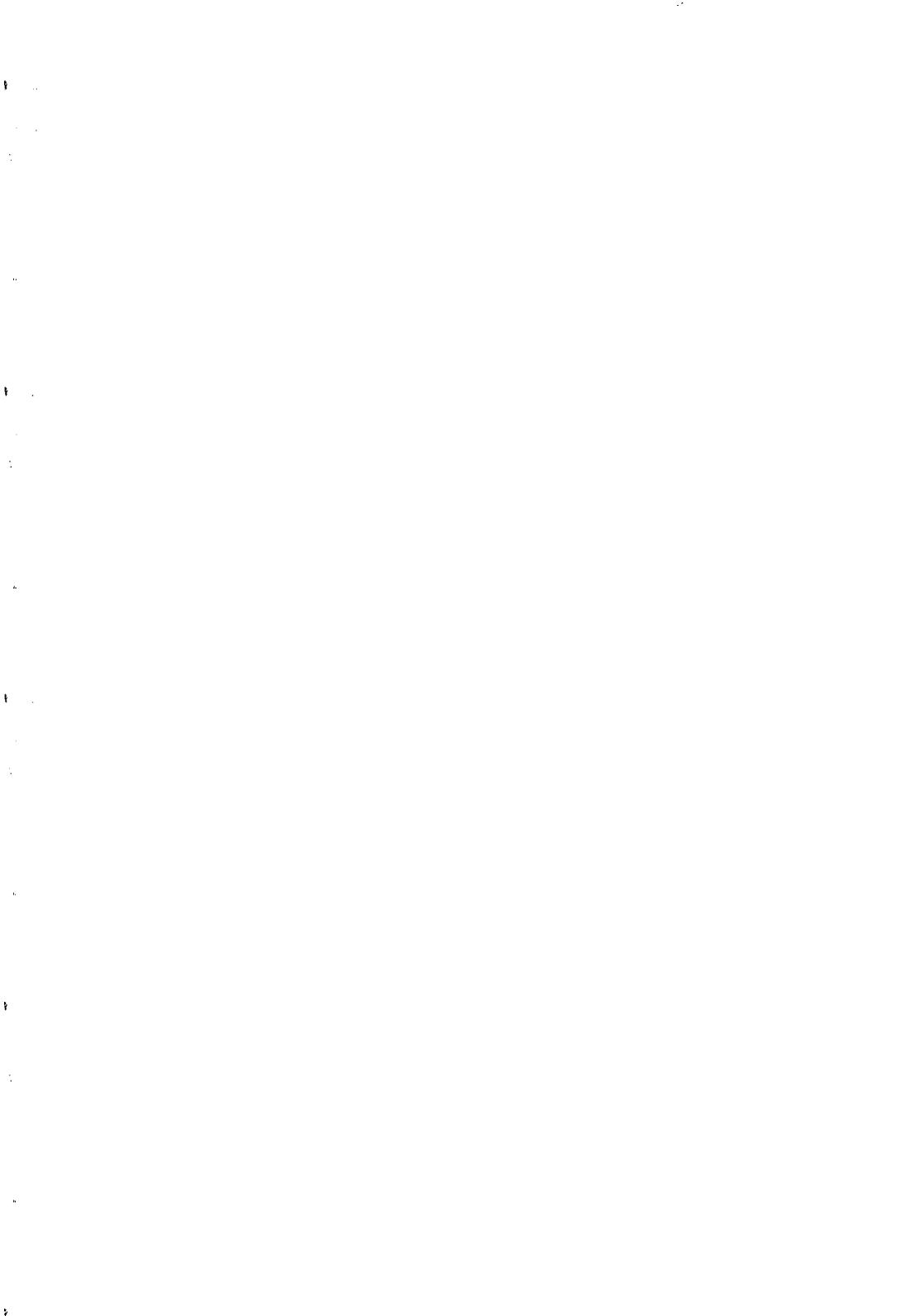
M. H. Holzscheiter	The Gravitational Acceleration of Antimatter.	491
G. Chardin	Testing the Discreteness of Spacetime at Low Energies.	503

IX. SUMMARY

H. Meyer	High, Very High and Ultrahigh Energy Astrophysical Phenomena: a Summary of the Sessions.	513
----------	--	-----

LIST OF PARTICIPANTS	517
-----------------------------	------------

HIGH ENERGY ASTROPHYSICAL PHENOMENA



OVERVIEW OF THE SIGMA RESULTS

J. PAUL

DAPNIA/Service d'Astrophysique, CE Saclay, 91191 Gif-sur-Yvette Cedex, France



ABSTRACT

The hard X-ray/soft γ -ray telescope SIGMA is one of the main instruments on board *GRANAT*, the Russian high energy astronomy space observatory, launched in 1989 December. Having examined hundreds of galactic and extragalactic fields during four years of successful in orbit operations, this coded aperture telescope, sensitive to radiation in the energy range 35 keV to 1.3 MeV, has succeeded both in disentangling severe source confusion problems and in providing firm identifications and reliable spectra of sources in such a spectral band. Tens of galactic sources, mainly X-ray binaries harbouring either a black hole (BH) or a neutron star (NS), as well as two isolated pulsars, have been detected up to 1993 December. Suspected BH systems seem to be particularly numerous among the source emitting above 100 keV, including the several X-ray novae so far detected. Bright active galactic nuclei (AGNs) have been detected and monitored during the SIGMA observations, indicating long term (one year time scale) flux variations. Even more interesting is the shorter time scale (few days) variability of these types of object which was observed for the first time at hard X-ray energies. A wealth of SIGMA results has been published so far, and we will summarize here some of the most prominent ones.

1. Introduction

Hard X-ray/soft γ -ray astronomy has long been considered as a particularly arduous field since focusing techniques are totally impracticable above ~ 10 keV, whereas estimating each photon direction by analyzing the pair tracks – as in EGRET¹⁾ – or the geometry of Compton diffusion – as in COMPTEL²⁾ – does not work below ~ 1 MeV. In between, astronomers have to rely on big collimated scintillator such as OSSE³⁾. At the beginning of the previous decade, it was recognized that a possible way to improve existing hard X-ray/soft γ -ray telescopes is the incorporation of the coded aperture technique to actually image celestial sources. The primary advantage of such a technique is to maintain the angular resolution of a single pinhole camera, while increasing the overall effective area of the instrument. Moreover, as the coded mask principle includes the simultaneous measurement of the sky and detector background, systematic effects due to temporal variations in the background are removed. The coded aperture concept had been successfully put to use in the X-rays domain by XRT on *Spacelab 2*⁴⁾, and at higher energy by the GRIP balloon borne experiment⁵⁾.

This paper reports on four years of successful high angular resolution observations performed by SIGMA, the first coded aperture telescope sensitive to radiation in the energy range from 35 keV to 1.3 MeV to be operated in space. We present here a selection of the salient scientific results obtained so far, with intent to emphasize the unique contribution of the imaging properties of the telescope, which were determinant not only to disentangle severe source confusion problems but also to enable firm identifications of sources. Even in the absence of source confusion problems, because the source is seen on the image and the background is measured at the same time and the same pointing as the source, imaging observations give source spectra which are free from any contamination from the background and/or from other nearby sources.

2. Instrument

The SIGMA telescope, one of the main devices on board the Russian *GRANAT* satellite, successfully launched on 1989 December 1st from Baikonour, Kazakhstan, was constructed by two French laboratories (Service d'Astrophysique at Saclay, and Centre d'Etude Spatiale des Rayonnements at Toulouse), both under contract to the Centre National d'Etudes Spatiales, the French Space Agency. This hard X-ray/soft γ -ray instrument of unprecedented size – weighing about one ton, it measures 3.50 m high and the diameter at the base is 1.20 m – features a coded mask, a position sensitive detector (PSD) whose design is based on the principle of the Anger cameras used in nuclear medicine, active and passive shielding devices, and the needful service modules (for a detailed description, see Paul *et al.*⁶⁾). The key imaging features of the telescope, whose total detection area is the 794 cm² central rectangular zone of the PSD matching the basic 29 \times 31 mask pattern, are the intrinsic angular resolution : 13', and the point source location accuracy, which is less than 2' for strong sources. Those properties are maintained over a 11.5° \times 10.6° field of view (FOV) at half sensitivity and over a 4.7° \times 4.3° FOV at full sensitivity.

Since the launch of the satellite into its nominal eccentric 4 day orbit (perigee 2000 km, apogee 200 000 km)¹, four years of in flight operations have been achieved, leading to a precise estimate of the in orbit performances of the instrument and of its long term evolution^{7,8)}. For broad band observations in a typical 10^5 s exposure, the 3σ sensitivity of SIGMA is $1.5 \cdot 10^{-5}$ photon $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$ (or 30 mCrab) at 100 keV. Identified spectral features observed from time to time during the mission have been used to confirm the instrument energy resolution as measured at the launch base before the flight : 16% at 60 keV, 12% at 122 keV and 8% at 511 keV. By the end of 1993 December, more than 500 celestial fields have been examined. In most cases, images with pixel size of 1.6' are accumulated in four contiguous wide energy bands, and images with pixel size of 3.2' are accumulated in 95 energy channels between 35 keV and 1.3 MeV.

In the usual data analysis procedure, the recorded images are first corrected in order to remove spatial non uniformities which are in part intrinsic to the detector and in part due to a structured background, and the sky images are reconstructed using standard deconvolution techniques. Point sources are located as maxima in the deconvolved images, with peak shape compatible with the SIGMA point spread function. Since the SIGMA telescope can observe in 95 consecutive energy bands, in each of which an image of the sky is derived, It is possible to obtain the count spectrum of a source independently of any other emission which originates from its surrounding, as e.g. from other sources within the FOV and from diffuse processes. As for other devices operating in the soft γ -ray regime, which is dominated by Compton scattering, a precise knowledge of the energy response function of the PSD is required to derive the actual source photon spectrum. Estimates of the SIGMA energy response are based on the ground calibrations performed in the launch base just before the flight, substantiated by the results of a precise Monte-Carlo simulation⁹⁾ to interpret the ground measurements.

3. General results

Table 1 presents the list of sources imaged by SIGMA up to 1993 December. The telescope operates in pointing mode only, and has mainly observed the Galactic Plane and particularly the Galactic Center (GC), where its imaging capabilities are best put to use. This list, which is just a compilation of already published positive source detections, is not a true statistical sample at a given sensitivity level. The source celestial coordinates (Equinox 1950.0), as presented in Table 1, have all been derived from the SIGMA data. Only two sources of this list, the Crab and 3C 273 are common to the list of sources so far detected by EGRET at 100 MeV¹⁾. Indeed, the hard X-ray spectra of the SIGMA sources are often cut off at a few hundred keV, whereas all the EGRET sources (except the Crab and 3C 273) present within the celestial field so far observed by SIGMA, have 100 keV flux below the SIGMA sensitivity limit. This may tell that thermal processes are still at work in the SIGMA range, pure non thermal processing taking over only at MeV energies.

¹ After a brief period of falling (from 2000 km down to 1000 km), the perigee is continuously rising, in such a manner that, from orbit to orbit, the time spent within the active fraction of the radiation belts becomes more and more reduced. Four years after launch, the perigee altitude was $> 30\,000$ km.

TABLE I
List of sources detected by SIGMA up to 1993 December

Name	$\alpha_{(1950)}$	$\delta_{(1950)}$	Activity	Nature
GRO J0422+32	4 18 29	32.79	Nova	BH system ?
M1/PSR 0531+21	5 31 30	21.98	Stable	Nebula/Fast pulsar
GS 0834-429	8 34 11	-42.94	Transient	Binary X-ray pulsar
GRS 1124-684	11 24 08	-68.38	Nova	BH system
NGC 4151	12 08 00	39.62	Variable	AGN (Seyfert 1)
NGC 4388	12 23 36	12.95	Weakly variable	AGN (Seyfert 2)
3C 273	12 26 40	2.36	Weakly variable	AGN (quasar)
GRS 1227+025	12 27 20	2.50	Transient	?
Cen A	13 22 18	-42.82	Strongly variable	AGN (radio galaxy)
PSR 1509-58	15 10 02	-59.02	Stable	Fast pulsar
TrA X-1	15 24 35	-61.66	Transient	BH system ?
OA0 1657-415	16 56 59	-41.57	Transient	Binary X-ray pulsar
GX 339-4	16 59 00	-48.73	Strongly variable	BH system ?
4U 1700-37	17 00 27	-37.78	Extremely variable	NS system
GRS 1716-249	17 16 34	-24.96	Nova	BH system ?
Terzan 2	17 24 14	-30.74	Variable	?
MXB 1728-34	17 28 45	-33.81	Flaring	X-ray burster
GX 1+4	17 29 02	-24.74	Rapidly variable	Binary X-ray pulsar
KS 1731-260	17 31 12	-26.09	Transient	X-ray burster
Terzan 1	17 32 38	-30.46	Variable	?
GRS 1734-292	17 34 00	-29.20	Transient	?
SLX 1735-269	17 34 58	-26.92	Persistent	?
IE 1740.7-2942	17 40 42	-29.72	Strongly variable	BH system ?
GRS 1741.9-2853	17 42 00	-28.89	Transient	?
A 1742-294	17 42 51	-29.49	Flaring	X-ray burster
GRS 1743-290	17 43 18	-29.05	Strongly variable	?
GRS 1747-341	17 47 26	-34.19	Persistent	?
GRS 1758-258	17 58 08	-25.74	Strongly variable	BH system ?
GRS 1915+105	19 12 51	10.86	Strongly variable	?
Cyg X-1	19 56 26	35.07	Weakly variable	BH system

4. Galactic hard X-ray sources viewed by SIGMA

A large fraction of the SIGMA sources listed in Table I are galactic stellar accreting systems, among which all are variable and many are transient sources, i.e. detectable only a small fraction of the time when in high state. The Crab (nebula and pulsar are not distinguished by SIGMA),

which is together with PSR 1509–58 the only source powered by rotation in the list presented in Table 1, is also the only mostly stable source. This extreme variability of the 100 keV sources is a direct consequence of their small size and the violent nature of the processes involved.

4.1. Black hole systems

Persistent sources. The prototype of BH binaries is Cyg X-1, whose spectrum in the normal state is characterized by a power law of index close to -2.0 , cut off above 200 keV. Such a spectrum is usually interpreted, following Sunyaev & Titarchuk¹⁰ as the result of Comptonization of soft (≤ 1 keV) photons by hot (~ 50 keV) electrons. Cyg X-1 was always caught by SIGMA in the same low state¹¹, while transient very low frequency (~ 0.04 Hz) quasi periodic oscillations (QPOs) have been discovered by SIGMA¹². GX 339–4, also suspected to harbour a BH, was repeatedly detected by SIGMA¹³. It exhibited large brightness variations, correlated with spectral hardening.

Other persistent sources are thought to be accreting BH binaries because of the similarity of their spectra to that of Cyg X-1, as the two hard sources 1E 1740.7–2942 and GRS 1758–258¹⁴, extensively monitored since the beginning of the SIGMA/GRANAT mission. Both sources exhibit roughly analogous long term behaviors^{15,16}, characterized by long periods (months) of bright state, followed by long periods of low state. Day to day variabilities are also apparent in both hard source light curves. Transient broad high energy features (beyond 200–300 keV) were detected in 1E 1740.7–2942^{17,18,19,20}, probably related to outbursts of pair formation.

The unprecedented statistics of the high resolution GC survey performed by SIGMA has made possible the determination of a more precise position of the hard source associated with 1E 1740.7–2942²¹. The improvement on the source position accuracy is such that the size of the error circle derived in the hard X-ray/soft γ -ray regime by a coded aperture instrument such as the SIGMA telescope, may now compare with that derived in the X-ray regime, where the source was first discovered. Radio observations with the VLA have revealed a compact variable radio source, the core of a double sided radio jet structure, close to the center of the improved SIGMA error circle²². The identification of 1E 1740.7–2942 with the VLA compact source was strengthened when comparing its time variations at 6 cm wavelength with the hard source light curve^{22,23}.

Soft X-ray transients. They are low mass X-ray binaries subject to episodic accretion, which become very bright in a few days, then decrease with a time scale around one month. In two members of this class, A 0620–00²⁴ and V404 Cygni²⁵, the mass of the compact object was shown to be larger than $3 M_{\odot}$, more than the NS limit. Their strong hard X-ray emission, although the spectra differ from one source to another, also put them on a par with other BH candidates such as Cyg X-1. SIGMA was conveniently on hand when a soft X-ray transient (GRS 1124–684) appeared in Musca in 1991 January, and could monitor the source from the earliest time of the flare^{26,27}. The optical spectroscopy allowing to measure the orbital elements of GRS 1124–684 had to wait for the optical emission of the nova to decrease below that of the companion ; its result put GRS 1124–684 in the list of BH candidate with a mass function of $3.1 \pm 0.5 M_{\odot}$ ²⁸.

The average spectral shape was close to a power law of photon index -2.4 . On 1991 January 20-21 SIGMA detected an emission line lasting more than 10 hours, centered around 480 keV, slightly below the energy of electron positron annihilation. It was the first such line observed in the spectrum of a soft X-ray transient. The fact that the source is clearly detected at the right position in the image derived in the energy interval where the line spectral feature is observed, substantially strengthens the overall confidence in the finding in such a case of a line detected in a spectral region where outstanding features are also detected in the background spectrum. A deeper analysis of the 1991 January 20-21 event reveals a complex picture of the source spectral evolution with the possible presence of two other high energy spectral features which could be interpreted as backscattered emission of the main line²⁹⁾.

Another such a source (GRO J0422+32), which appeared in Perseus in August 1992, was observed in detail by SIGMA. It was at the time the brightest hard X-ray source in the sky (~ 3 Crab). Its spectrum was comparable to that of Cyg X-1³⁰⁾ and its power spectrum revealed low frequency QPOs³¹⁾, most conspicuously around 0.3 Hz. By analogy with the other sources of this class, GRO J0422+32 is expected to be a new BH system. This is also the case of GRS 1716-249, the hard X-ray nova recently discovered by SIGMA in Ophiuchus³²⁾. Note also that SIGMA detected a sporadic hard emission from TrA X-1³³⁾, a soft X-ray transient whose largest outburst was observed on 1974 December 4³⁴⁾.

4.2. Neutron stars

NS systems. They form the bulk of X-ray binaries seen below 20 keV. Their spectra steepen earlier than those of BH systems, with cut off energies around 20 keV³⁵⁾. Indeed, none of the bright ($\sim 10^{38}$ erg s⁻¹, close to the Eddington limit) X-ray binaries has been detected by SIGMA, implying that their hard X-ray emission is less than 1% of the total. However, SIGMA has detected hard tails from two sub classes of X-ray binaries harbouring a NS. The X-ray bursters, which are relatively weak ($\sim 10^{37}$ erg s⁻¹) X-ray binaries, showed hard X-ray variabilities on time scales of a few days, as KS 1731-260³⁶⁾, MXB 1728-34 \equiv GX 354-0³⁷⁾ and A 1742-294³⁸⁾. Similarly, outbursts of hard X-rays were also observed in the emission of binary X-ray pulsars, as OAO 1657-415³⁹⁾, GX 1+4⁴⁰⁾, GS 0834-429⁴¹⁾ and Vela X-1⁴²⁾. The high mass X-ray binary 4U 1700-37, in which a NS orbits the wind of its massive companion, is a case in point. This source also features hard X-ray outbursts⁴³⁾, but the pulsar nature of the compact companion is not yet established.

Isolated fast pulsars. Although the time profile of the pulsed emission of PSR 0531+21 was measured soon after the launch of SIGMA/GRANAT⁴⁴⁾, no other pulsed emission from fast isolated pulsars was detected by SIGMA via time analysis. On the other hand, the total emission (pulsed plus unpulsed component) from PSR 1509-58 has been detected after 192 hours of observations preformed in 1990-1993⁴⁵⁾. By comparing the SIGMA flux to the pulsed flux measured by BATSE⁴⁶⁾, one can estimate the pulse fraction in the 40-300 keV to be $\sim 100\%$.

4.3. Black hole systems versus neutron star systems

On the basis of the results so far obtained by SIGMA, it has been proposed¹⁴⁾ that a direct relationship exists between the nature (NS or BH) of the compact object in a X-ray binary system and the overall spectral shape of its hard X-ray emission. Indeed, all sources known to be NS systems (X-ray bursters or binary X-ray pulsars) exhibit spectra which are clearly softer around 100 keV than the hard spectra of BH systems in their luminous states. As shown in Table 2⁴⁷⁾, the comparison of luminosities, taking distances into account, also clearly tells BH systems from NS systems. The 150-500 keV luminosity of NS systems was always less than a few 10^{35} erg s⁻¹, compared with positive detections at few 10^{36} erg s⁻¹ for BH systems. The brightest NS system so far detected by SIGMA (4U 1700-377) maintained this luminosity for a few hours only⁴³⁾, whereas the BH outbursts last a few weeks or more.

TABLE 2

Luminosity in the 150-500 keV band during luminous states in accreting binary systems⁴⁷⁾. An asterisk means that the distance is very uncertain.

Name	Nature	$L_{150-500 \text{ keV}}$ (10^{35} erg s ⁻¹)
GRO J0422+32	BH system ?	- 115 (★)
GRS 1124-684	BH system	- 50
GX 339-4	BH system ?	- 30
1E 1740.7-2942	BH system ?	- 40
GRS 1758-258	BH system ?	- 20
Cyg X-1	BH system	- 20
OA0 1657-415	Binary X-ray pulsar	≤ 1
4U 1700-37	NS system	- 5
MXB 1728-34	X-ray burster	≤ 3
GX 1+4	Binary X-ray pulsar	≤ 2
KS 1731-260	X-ray burster	≤ 2 (★)
A 1742-294	X-ray burster	≤ 5

5. Extragalactic hard X-ray sources viewed by SIGMA

Before the launch of SIGMA/GRANAT, it has been often predicted that hard X-ray/soft γ -rays should provide a promising spectral band for studying AGNs, since most of the few data available at this time on AGNs detected beyond 100 keV indicated that a substantial fraction of the total luminosity in these sources is emitted in the hard X-ray/soft γ -ray band. The results derived from the SIGMA pointing devoted to the study of AGNs of various types should temper these predictions, even if the sensitivity of the telescope is such that the extrapolation of the X-ray AGN spectra indicates that only the brightest ones can be detected.

Marginally significant (3σ) flux variations of the Seyfert 1 galaxy NGC 4151 were detected during the 1990-1992 SIGMA observations⁴⁸⁾ in both 35-60 keV and 60-170 keV energy bands, but no significant variation in the hardness ratio was found. The steep source spectrum in hard X-rays, first detected by SIGMA in 1990⁴⁹⁾, and confirmed in later SIGMA observations⁴⁸⁾, is much steeper than observed in lower energy X-rays and cannot be reconciled with previous detections in the hard X-ray/soft γ -ray regime. The 1990-1991 SIGMA observations of the quasar 3C 273⁵⁰⁾ indicate also a marginally significant (3σ) 40-120 keV flux variation by a factor of two in a time scale as short as 41 days. The combined data of the first two pointing are well fitted with a power law of index ~ -1.5 up to about 500 keV, with no evidence for a break or steepening in the spectrum. The radio galaxy Cen A has been observed by SIGMA on three occasions in 1990-1991⁵¹⁾. A comparison between the observations indicates a flux increase by a factor of three over a one year time scale. Even more interesting is a similar decrease which was observed in just four days. Despite such intensity variations, the spectral shape of the emission remains unchanged and is well fitted with a power law of index ~ -2 up to about 100 keV.

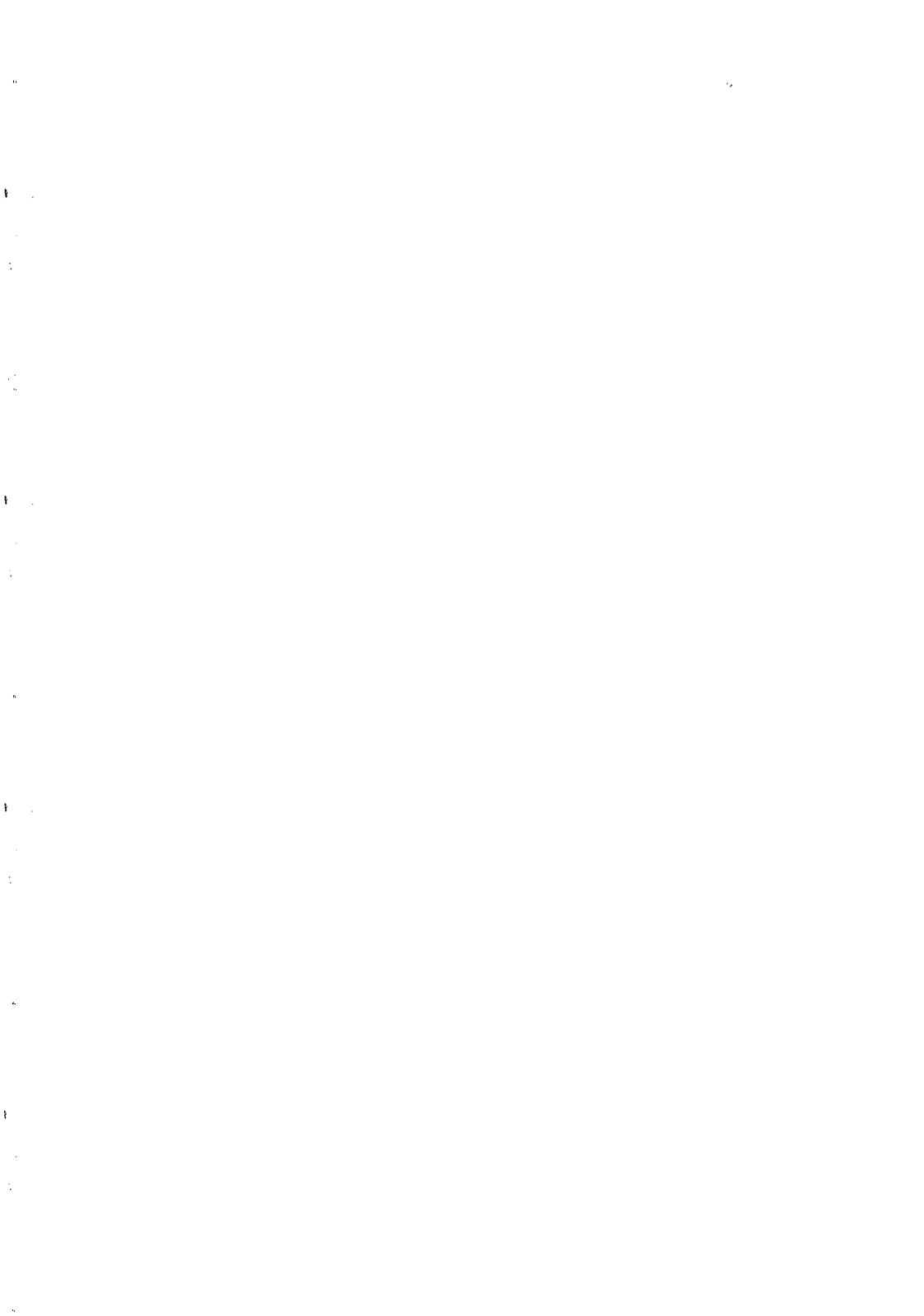
Source confusion is not an exclusive prerogative of galactic regions but may also affect extragalactic fields as the Virgo cluster, which has been observed 10 times by SIGMA in 1990-1991. Among the well known AGNs of the Virgo cluster, only the Seyfert 2 galaxy NGC 4388 was detected, near the limit of visibility⁵²⁾. Contrary to the soft X-ray band, where prevails the giant elliptical galaxy M 87, the emission of the Virgo cluster at higher energies is dominated by NGC 4388. Similarly, the sky region around 3C 273 not only shows emission from the quasar⁵⁰⁾, but also from a new hard source, GRS 1227+025, located 15' away from 3C 273⁵³⁾, and tentatively identified with an uncatalogued *Einstein* variable X-ray source, whose optical counterpart is a faint QSO at $z = 0.57$ ⁵⁴⁾.

7. Conclusions

It should be first emphasized that the SIGMA concept, based on a static coded mask aperture operating in the hard X-ray/soft γ -ray regime, is fully certified, at least in the rather stable background environment experienced on the *GRANAT* highly eccentric orbit. The unique capabilities of coded aperture have been clearly demonstrated, as e.g. the aptitude to disentangle confuse regions, to enable firm identifications at other wavelengths and to perform reliable spectral measurements. In addition to the fact that source confusion is not an exclusive prerogative of galactic regions as the GC vicinity, but may also affect extragalactic targets, one of the most fundamental results which comes out from four years of successful in orbit operations with SIGMA is the high variability of the hard X-ray/soft γ -ray sky. With the exception of the sources powered by rotation, and with the exception of few persistent sources which are too weak to allow time variability determinations, all the sources so far detected by SIGMA are all time variable, and most of them are even strongly variable. The scales of variability range from minutes to years. This is also true for the spectral features so far detected in the emission of the soft γ -ray sources: none of them were observed during more than a day.

References

1. Kniffen D.: 1993, in Proc. of 27th ESLAB Symp., in press
2. Bennett K.: 1993, in Proc. of 27th ESLAB Symp., in press
3. Johnson W.N. *et al.*: 1993, *ApJS*, in press
4. Skinner G.K. *et al.*: 1987, *Adv. Space Res.* 7 (5), 223
5. Althouse W.E. *et al.*: 1985, in Proc. of 19th Int. Cosmic Ray Conf. 3, 299
6. Paul J. *et al.*: 1991, *Adv. Space Res.* 11 (6), 289
7. Mandrou *et al.*: 1991, in *AIP Conf. Proc.* 232, 492
8. Leray J.P. *et al.*: 1991, in Proc. of 22nd Int. Cosmic Ray Conf. 2, 495
9. Barret D., Laurent P. 1991, *Nucl. Inst. and Meth.* A307, 512
10. Sunyaev R.A., Titarchuk L.G.: 1980, *A&A* 86, 121
11. Salotti L. *et al.*: 1992, *A&A* 253, 145
12. Vikhlinin *et al.*: 1992, *IAU Circ.* 5576
13. Bouchet L. *et al.*: 1993, *ApJ* 407, 739
14. Sunyaev R. *et al.*: 1991, *A&A* 247, L29
15. Cordier B. *et al.*: 1993, *A&AS* 97, 177
16. Gilfanov M. *et al.*: 1993, *ApJ* 418, 844
17. Bouchet L. *et al.*: 1991, *ApJ* 383, L45
18. Sunyaev R. *et al.*: 1991, *ApJ* 383, L49
19. Churazov E. *et al.*: 1993, *ApJ* 407, 752
20. Cordier B. *et al.*: 1993, *A&A* 275, L1
21. Cordier B. *et al.*: 1993, *A&A* 272, 277
22. Mirabel I.F. *et al.*: 1992, *Nature* 358, 215
23. Cordier B. *et al.*: 1993, *Adv. Space Res.* 13 (12), 611
24. McClintock J., Remillard R.A.: 1986, *ApJ* 308, 110
25. Casares J., Charles P.A., Naylor T.: 1992, *Nature* 355, 614
26. Sunyaev R. *et al.*: 1992, *ApJ* 389, L75
27. Goldwurm A. *et al.*: 1992, *ApJ* 389, L79
28. McClintock J., Bailyn C., Remillard R.: 1992, *IAU Circ.* 5499
29. Goldwurm A. *et al.*: 1993, *A&AS* 97, 293
30. Roques J.P. *et al.*: 1994, *ApJS*, in press
31. Vikhlinin *et al.*: 1992, *IAU Circ.* 5608
32. Ballet J. *et al.*: 1993, *IAU Circ.* 5874
33. Barret D. *et al.*: 1992, *ApJ* 392, L19
34. Kaluziensky L.J. *et al.*: 1975, *ApJ* 201, L121
35. White N.E., Stella L., Parmar A.N.: 1988, *ApJ* 324, 363
36. Barret D. *et al.*: 1992, *ApJ* 394, 615
37. Claret A. *et al.*: 1994, *ApJ*, in press
38. Churazov E. *et al.*: 1994, *ApJS*, in press
39. Mereghetti S. *et al.*: 1991, *ApJ* 366, L23
40. Laurent P. *et al.*: 1993, *A&A* 278, 444
41. Denis M. *et al.*: 1994, *ApJS*, in press
42. Ballet J. *et al.*: 1992, *IAU Circ.* 5536
43. Laurent P. *et al.*: 1992, *A&A* 260, 237
44. Natalucci L. *et al.*: 1991, *Adv. Space Res.* 11 (6), 289
45. Laurent P. *et al.*: 1994, *A&A*, in press
46. Wilson R.B. *et al.*: 1992, *IAU Circ.* 5429
47. Laurent P. *et al.*: 1993, *Adv. Space Res.* 13 (12), 139
48. Finogenov A. *et al.*: 1994, submitted to *A&A*
49. Jourdain E. *et al.*: 1992, *A&A* 256, L38
50. Bassani L. *et al.*: 1992, *ApJ* 396, 504
51. Jourdain E. *et al.*: 1993, *ApJ* 412, 586
52. Lebrun F. *et al.*: 1992, *A&A* 264, 22
53. Jourdain E. *et al.*: 1992, *ApJ* 395, L69
54. Grindlay J.E.: 1993, *A&AS* 97, 113



GAMMA-RAY BURSTS

Matthew G. Baring*

NASA Goddard Space Flight Center, Code 665, Greenbelt, MD 20771, U.S.A.



ABSTRACT

Gamma-ray Bursts are an exotic class of astronomical sources that are currently extremely topical within the astrophysical community and are also provoking considerable interest in a broader range of physics disciplines. This paper briefly reviews the generic properties of these transient sources, obtained from observations accumulated over the last two decades, and then discusses some of the latest observations, focusing on recent interpretations of the data, and theoretical approaches that may help unravel the mysteries of gamma-ray bursts.

AN INTRODUCTION TO GAMMA-RAY BURSTS

Gamma-ray Bursts (GRBs) have intrigued observers and confounded theorists ever since their discovery just over twenty years ago.¹ The results obtained by the Compton Gamma-Ray Observatory (CGRO) have perpetuated the confused picture we have of these enigmatic objects. Yet the GRB field is alive and more dynamic than ever, perhaps due to the exotic nature of these objects and the fascinating physics that may be needed to describe and understand them. The enigmatic nature of GRBs is largely because they are transient phenomena. They are bursts of gamma-rays, usually somewhere between about 10 keV and 10 MeV in energy (e.g. see ref. 2), and last for as little as milliseconds or as long as several minutes.²⁻⁶ Typical burst time histories are displayed in the First BATSE GRB Catalogue⁵ and an exceptionally good burst spectrum is presented in Fig. 1. While other transient astronomical sources exist, gamma-ray bursts are unique in that they are the only source population that emits only in the gamma-ray range of the spectrum; associated emission in

* NAS-NRC Research Associate

other wavebands has never been observed. In fact, searches (e.g. see papers in ref. 7) for convincing transient or quiescent steady-state counterparts to classical GRBs in radio, optical and X-ray bands have proved negative. Note that such counterparts have been found for the soft-gamma repeaters, a separate class of objects that will not be considered here. This null result has prolonged and enhanced the mystery of GRBs, since the discovery of associated sources would provide distances to GRBs and thereby tell astronomers their origin.

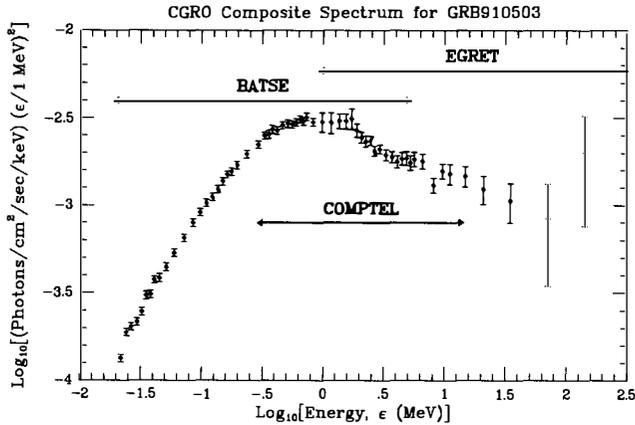


Fig. 1 The composite time-integrated spectrum⁸ for the burst GRB910503 (detected May 3, 1991) as obtained by the BATSE, COMPTEL and EGRET instruments on the Compton Gamma-Ray Observatory (CGRO), multiplied by the square of the photon energy ϵ to depict the emitted power per logarithmic energy bandwidth. The various detector responses have been accounted for in producing this source spectrum, and the energy ranges over which the three instruments are sensitive are indicated.

Gamma-ray bursts have rapid time-variability with intensity fluctuations on timescales t_{var} sometimes less than a millisecond,⁹ motivating the premise that they originate in compact environments of size $ct_{var} \sim 10^7$ cm. This naturally lead to the popularity of neutron stars as a site for GRBs, an idea that was bolstered by the realization that their extremely strong magnetic fields could contain the emitting plasma in the source for typical burst durations. The time-histories of GRBs, as illustrated in ref. 5, display a diversity of structure, with single or multiple peaks, subpeaks and lengthy gaps between peaks appearing in various sources. Typical burst spectra $f(\epsilon)$ are such that the power $\epsilon^2 f(\epsilon)$ (e.g. see Fig. 1) peaks in the range 50 keV – 5 MeV, a rare feat for astronomical sources; even more remarkable is the fact that less than 2% of the GRB flux appears in the range¹⁰ 2–10 keV. This paucity of X-rays renders bursts unique among classes of cosmic gamma-ray sources, and severely constrains theoretical GRB models. The observed GRB fluxes at earth imply source luminosities of $L \sim 10^{38}$ erg/sec if they are 1 kpc distant, suggesting a neutron star origin, or $L \sim 10^{50}$ erg/sec if 1 Gpc away, reminiscent of supernova events.

The early observations of gamma-ray bursts yielded spectra that were mostly quasi-thermal in nature (e.g. see the KONUS data in ref. 11), however the results¹² obtained by instruments aboard the Solar Maximum Mission (SMM) suggested that many if not most continuum spectra were non-thermal and extended well above 1 MeV, with a quasi-power-

law shape. This GRB property has been largely confirmed by the CGRO detectors, with COMPTEL and EGRET seeing several bursts with high energy emission,^{13,14} specifically at about 1 GeV for the case of GRB930131.¹⁴ This is a statistically significant detection rate that suggests that many (or most) GRBs emit at these energies. The earlier quasi-thermal observations may reflect intrinsic properties of those bursts, or that they may have been influenced by threshold effects at high energies in less sensitive detectors like KONUS.

Prior to the launch of CGRO, many observations of spectral lines from bursts were reported. About 20% of sources¹¹ showed single absorption dips in their spectra between 20 and 60 keV, which would yield magnetic field strengths of $2 - 5 \times 10^{12}$ Gauss in their emission regions if these broad features were interpreted as cyclotron absorption lines. Further, around 5% of bursts exhibited broad emission lines¹¹ at around 400 keV, which were naturally interpreted as electron-positron (pair) annihilation lines redshifted from 511 keV by the gravitational field of neutron stars. Therefore, around 1988, the weight of evidence in favour of a galactic neutron star association for bursts was substantial, and the stunning observation of double narrow absorption dips in two burst spectra by the GINGA X-ray satellite¹⁵ (interpreted as cyclotron harmonics) all but clinched the case for proponents of galactic burst sources. Since then, the CGRO instruments, in particular BATSE, have observed no lines of either type in over 800 burst detections, spawning a controversy over the existence of GRB line features. The consistency of the apparently conflicting BATSE and GINGA results concerning absorption lines is hotly debated, and presently it is uncertain whether or not the discrepancy can be explained by the different instrumental line-detection capabilities.

RECENT POPULATION STUDIES

The launch of the Compton Gamma-Ray Observatory created an even bigger upheaval to the "conventional wisdom" of the galactic neutron star hypothesis when the BATSE experiment observed that the burst population was isotropic but inhomogeneous with a comparative deficiency of fainter sources.¹⁶ Although this result was suggested by earlier experiments that observed significant numbers of GRBs, BATSE is a more sensitive detector and has been able to observe more bursts with a broader range of source fluxes. As such it was intended to observe to the edge of a galactic disk population and therefore detect an anisotropy in the spatial distribution of bursts; the contrary result has indicated that most (or all) bursts may in fact be cosmological in origin.¹⁷ This is currently the new conventional wisdom, and while the observed isotropy and inhomogeneity may be explained by GRBs originating in a large spherical galactic halo, the expectation that halo bursts from M31, the nearest large galaxy (Andromeda), should be detected and upset the observed isotropy is now severely constraining¹⁸ a possible halo population. A good review of the various population hypotheses is presented in ref. 19. In summation, it is now extremely unlikely that all bursts could be confined within the galaxy, though a significant minority disk population is still quite feasible.

Much discussion over the last year has concentrated on population studies with claims and counter-claims about evidence for two or more burst populations and repeating bursts. There is fairly strong evidence for a bimodality in the distribution of GRB durations,⁶ which may not necessarily be an indication of two classes of bursts. Recent work²⁰ on average burst temporal profiles has suggested a time dilation effect that is consistent with, but not

conclusively implying, a cosmological origin of bursts; the statistics on this result need to be improved in order to make more definitive deductions. There is even a hint of spectral reddening in fainter sources.

Suggestions this last year²¹ of evidence for repetition of bursts in the BATSE observations sparked a storm of controversy. It was argued that the first 260 bursts that BATSE detected had an improbably high clustering in the sky; specifically more bursts were correlated on four degree angular scales than expected, i.e. are near-neighbours in the sky. This lead to the provocative suggestion that BATSE was detecting repetitions of bursts, a naturally attractive idea to galactic GRB proponents since the enormous energies ($L \sim 10^{52}$ erg/sec) required to power GRBs at cosmological distances virtually precludes any chance of burst repetition. In fact an excess correlation on 176 degree scales²² is also observed in the same population, which cannot indicate repetition; these two properties taken together might suggest the presence of a minority anisotropic (perhaps a galactic disk) population. Preliminary analysis²³ of the next 482 bursts indicates that these correlations may disappear with better statistics, though complete correction for instrumental systematic effects has yet to be included. Therefore, presently the evidence for repeaters is unconvincing.

SPECTRAL CONSIDERATIONS

Information about individual bursts is still best deduced from their spectra, which can provide much insight concerning the physical processes occurring in GRB emission regions. The natural example is the interpretation of low energy (20-60 keV) absorption features seen in spectra obtained by the KONUS¹¹ and GINGA,¹⁵ referred to earlier, as being due to absorption of radiation at the cyclotron energy in the strong magnetic fields of neutron star magnetospheres. Since the BATSE instrument on CGRO has made no definitive observation of any spectral line features the GRB continuum must now provide the clues for theorists. Many such continuum spectra $f(\epsilon)$ have a general appearance similar²⁴ to Fig. 1 [which depicts $\epsilon^2 f(\epsilon)$], with a very flat spectrum at the lower energies (i.e. below 50 keV) "turning over" or "breaking" to a steep spectrum above 1 MeV. The energy ϵ_B of the turnover varies from burst to burst²⁵, and a significant fraction of bright bursts²⁶ detected by BATSE have $\epsilon_B > 100$ keV. The break energy ϵ_B and the spectral index $\alpha = -d[f(\epsilon)]/d[\log_e \epsilon]$ at energies above and below ϵ_B provide important constraints on theoretical models.

The most sophisticated model that generates turnovers in γ -ray continuum spectra is the $B = 0$ pair cascade model²⁷ that has been comprehensively developed for applications to active galactic nuclei (AGNs). This model used isotropic ultrarelativistic electrons, generated by some unspecified particle energization mechanism such as shock acceleration,²⁸ to produce a gamma-ray continuum by the inverse Compton scattering of UV photons, and then the two-photon pair production process $\gamma\gamma \rightarrow e^+e^-$ to absorb the γ -rays and initiate a cascade of several or many generations of pair production. The attenuation of the GRB continuum by $\gamma\gamma \rightarrow e^+e^-$ is an obvious candidate for a means to create the observed "MeV" turnovers. When this model is adapted²⁹ to try to fit the flatter hard X-ray continuum of GRBs, multiple inverse Compton scattering of the seed low energy photons occurs, spawning a peak at around 1 MeV where the Klein-Nishina decline of the Compton cross-section effectively shuts off the cascade. Refinements of the model³⁰ revealed that Coulomb scattering by the relativistic

electrons heats the ambient cool thermal electron gas to relativistic temperatures, so that a broad quasi-thermal bump due to pair annihilation and thermal bremsstrahlung appeared around 1 MeV, contrary to observations. It now seems unlikely³¹ that such field-free pair cascades can create the MeV breaks observed in bursts. Therefore it appears that the BATSE sources with breaks observed in the MeV energy range²⁸ require a mechanism other than $\gamma\gamma \rightarrow e^+e^-$ to generate their turnovers.³¹

For neutron star models of GRBs, MeV breaks can be generated using synchrotron pair cascades, where magnetic pair production $\gamma \rightarrow e^+e^-$ acts to attenuate the continuum,^{32,33} since in the strong fields of neutron stars ($B \gtrsim 10^{12}$ Gauss) this process dominates $\gamma\gamma \rightarrow e^+e^-$ as a means of absorbing gamma-rays. In these cascades, ultrarelativistic electrons are injected into the emission region with significant pitch angles θ so that a synchrotron continuum is radiated, with the radiation being beamed close to the direction of the electrons' momenta. The dramatic increase^{2,32,34} in the pair production optical depth above the effective threshold of $\epsilon = 2m_e c^2 / \sin \theta$ efficiently truncates the continuum above this energy,³² subsequently generating a pair cascade. Since this truncation energy is a function of pitch angle, spectra integrated over θ can generate broken power-laws given suitable angular distributions of electrons.³⁵ The spectral breaks in this scenario can be substantial or non-existent, an attractive versatility of this model, depending on how beamed the electrons are along the field lines. The drawback of the synchrotron cascades is that their spectrum below the break is never flatter than $\alpha = 3/2$, so that they cannot model many flat spectrum sources like GRB 910503 (see Fig. 1).

Another mechanism for generating MeV breaks in magnetized environments is resonant inverse Compton scattering. When $B \neq 0$, the Compton cross-section has a resonance at the cyclotron energy $\epsilon_c \propto B$ in the electron rest frame, which dominates the scattering process and dramatically alters the shape of the spectrum. For ultrarelativistic electrons with a power-law distribution γ^{-P} of Lorentz factors γ , colliding with seed soft photons (presumably X-rays from the surface of a neutron star) of energy ϵ_s , this process yields broken power-law spectra, flat at low energies (i.e. $\alpha < 1$) with a break³⁶ at $\epsilon_B = \epsilon_c^2 / \epsilon_s$. Resonant scattering has no problem dealing with those GRB spectra that are flat below 100 keV, and can easily predict breaks at MeV energies. The introduction of e^- cooling to this model^{37,38} steepens the spectrum somewhat and smooths the break; the observed spectra are strongly dependent on the angle of photon emission with respect to the field lines. The resonant scattering model has difficulty describing sources with small breaks such as GRB930131; possibly³¹ a synchrotron self-Compton hybrid cascade model might succeed in eliminating the deficiencies of the two neutron star models discussed here. The few cosmological models for GRB spectra that exist, including the fireball scenarios discussed below and AGN-like models,³⁹ are considerably less sophisticated than the neutron star variety and need much more development for serious comparison with burst continuum observations.

RELATIVISTIC MOTION IN BURSTS

The indication earlier that $\gamma\gamma \rightarrow e^+e^-$ is not responsible for MeV turnovers in gamma-ray bursts implies that it is probably not operating to attenuate the gamma-ray continuum within the range of observed burst energies. An important consequence of this is that rel-

ativistic bulk motion may be common in burst sources. Two-photon pair production was first considered as a potential mechanism for attenuating GRB spectra by Schmidt,⁴⁰ who proposed that it could be used to provide upper limits to the distance to burst sources that revealed no spectral breaks above 1 MeV. Schmidt's application focussed on the case of isotropic photons, and concluded that the bursts seen before then (at energies less than a few MeV) must be galactic, since they had no breaks, in direct conflict with the recent BATSE results' implications for burst populations. If the majority of bursts are at cosmological distances (~ 1 Gpc), then the enormous luminosities expected ($L \sim 10^{50}$ erg/sec), combined with the compact source size implied by the observed rapid time variability, give photon densities high enough to make GRBs optically thick to $\gamma\gamma \rightarrow e^+e^-$ by many orders of magnitude; i.e. the free escape of gamma-rays from burst sources would be impossible. Such a situation is incompatible with the observed spectra of GRBs detected by EGRET, which show no attenuation at energies in some cases as high as 1 GeV (see refs. 13, 14).

One way around this problem is to assume beaming of the radiation, which raises the pair production turnover to energies above those observed. Relativistic motion of the source is a natural way to achieve beaming of the emission and avoid attenuation in GRBs. The radiation from a source that is emitting isotropically in a reference frame that moves relativistically with respect to an observer on earth will be beamed roughly within an angle $1/\Gamma$ in the observer's frame,^{31,41} where Γ is the bulk Lorentz factor. Such *relativistic beaming* of radiation naturally suppresses the pair production rate, since the $\gamma\gamma \rightarrow e^+e^-$ mechanism has a threshold energy $2(m_e c^2)^2/[\epsilon(1 - \cos \Theta)]$ for photons of energy ϵ ; this is strongly dependent^{31,41} on the angle Θ between the photon directions. This implies that relativistic beaming of radiation in bursts with declining spectra ($\alpha = -d[f(\epsilon)]/d[\log_e \epsilon] > 0$: true for all EGRET sources) will "blueshift" the $\gamma\gamma \rightarrow e^+e^-$ turnover up in energy by a factor of Γ (since then $\Theta \sim 1/\Gamma$); that this moves above the EGRET energy range therefore provides a determination of a lower bound to Γ .

When the source opening angle is of order $1/\Gamma$, the pair production optical depth $\tau_{\gamma\gamma}$ is reduced by a factor $\Gamma^{-(1+2\alpha)}$ below the optical depth for isotropic radiation.^{31,41} The condition that no spectral attenuation occurs is that $\tau_{\gamma\gamma} < 1$ up to the highest energies observed (the optical depth is effectively the probability that $\gamma\gamma \rightarrow e^+e^-$ occurs in the source). The bulk Lorentz factors Γ consequently required for the bright "superbowl" burst GRB930131, detected by EGRET⁴⁴ up to an energy of 1 GeV, are $\Gamma \gtrsim 10^3$ if it is 1 Gpc distant and $\Gamma \gtrsim 10$ at 30 kpc (ref. 42). Similar estimates of bulk relativistic motion are obtained for the other EGRET sources,⁴³ indicating that this phenomenon is common in bursts. An advantage of relativistic beaming is that a smaller source luminosity is required because the observed flux is enhanced by a solid angle factor Γ^2 (ref. 41). However, then the number of sources must really be a factor Γ^2 higher than detected in order to account for the observed rate of gamma-ray bursts. In the case of cosmological GRBs, this factor could be as high as 10^6 for the typical Lorentz factors $\Gamma \sim 10^3$, which is unacceptably large for many models. Source geometry therefore has significant impact on the viability of models. Expanding the opening angle to become much larger than $1/\Gamma$ could ease this problem⁴⁴ if the high energy photons were able to escape; in fact estimates of the minimum Γ for completely spherical expanding sources are qualitatively similar to the highly beamed case.⁴⁴

FIREBALL MODELS

While relativistic motion in GRBs that have $\tau_{\gamma\gamma} < 1$ may be common, it is also probable in bursts of high optical depth to $\gamma\gamma \rightarrow e^+e^-$. Such situations lead to many generations of pair production and naturally arise out of several cosmological models for bursts. Among these is the so-called merger scenario, where two neutron stars or a neutron star and a solar mass black hole coalesce⁴⁵⁻⁴⁷ due to the loss of their orbital energy via the radiation of gravitational waves. Such mergers, which are envisaged as forming a cosmological population of GRBs, may occur at the rate of about one per galaxy per million years since there are four binary neutron star systems observed in our own galaxy,⁴⁸ and can release a large fraction of a solar mass of energy (about 10^{54} ergs) in an extremely small volume, typically about $R \sim 10^8$ cm in radius. The energy density is therefore enormous, as is $\gamma\gamma \rightarrow e^+e^-$, and thermal equilibrium is established rapidly, generating a pair plasma at relativistic temperatures⁴⁵ [$kT/(m_e c^2) \sim (k/m_e c^2)[M_\odot c^3/(4\sigma R^3)]^{1/4} \sim 20$]. This energy naturally must dissipate adiabatically,^{45,49} and the resulting expansion of pairs is called⁵⁰ a *fireball*. The early stages of such a dynamic fireball may involve rapid neutrino production (and $\nu\bar{\nu}$ annihilation to produce pairs with 0.1% efficiency) so that enormous numbers of neutrinos are emitted⁴⁶ ($E/c^2 \sim M_\odot \Rightarrow 10^{60}\nu$); unfortunately this leads to about 10^3 neutrinos per square cm on earth, a signal that is many orders of magnitude too small for detection.

Other models where fireballs arise include failed type 1b supernovae,⁵¹ in which a star with a massive iron core fails to expel its outer layers as a supernova, allowing the would-be ejecta to form a transient accretion disk and then collapse onto the condensed core (a newly-formed neutron star) and explosively depositing energy to form a fireball. Copious production of neutrinos would be expected in this model. It has also been suggested⁵² that rapidly spinning neutron stars in distant galaxies could form with extremely high magnetic fields, 10^{15} Gauss, much stronger than conventional pulsar fields. These so-called *magnetars* would result from the collapse of accreting white dwarfs with unusually high magnetic fields (about 10^9 Gauss), and would rapidly lose angular momentum through gravitational and magnetic-dipole radiation. A predominantly electromagnetic fireball would then be initiated, comprised mostly of electron-positron pairs and photons. Both models are envisaged as defining cosmological burst populations.

Non-relativistic temperatures (typically around 20 keV) are usually achieved^{45,49} in the adiabatic cooling of fireballs, which naturally imply that the bulk velocities attained by the expansions when they become optically thin to $\gamma\gamma \rightarrow e^+e^-$ are ultrarelativistic (typically $\Gamma \sim 10^3$); this is required in order to satisfy energy conservation. These properties are confirmed by hydrodynamic analyses of fireball evolution.⁵³ Rudimentary calculations⁴⁹ of the output spectrum of this fireball is a cool (narrow) Planck continuum blueshifted to MeV energies and distorted by bulk motion of the plasma. This is unlike the observed non-thermal spectra of bursts, leading immediately to a problem with pure fireball models. More detailed computations⁵⁴ of radiative transfer effects in late stages of fireball evolution indicate that high energy power-law tails may be possible, though refinement of these calculations is needed.

The fireball phenomenon is not as simple as an adiabatically expanding plasma of pairs and photons; realistic models must include some form of baryonic pollution swept up from the

environment of the progenitor. This contamination depends strongly on the geometry in the progenitor. It was realized⁵⁵ that the presence of baryonic matter in the fireball must rob it of its radiative efficiency since the baryons can dominate the kinetic energy of the expansion. Effectively the initial energy of the fireball is deposited as kinetic energy of the baryons rather than emitted as radiation. Simple computations⁵⁵ reveal that the final bulk Lorentz factor Γ attained by the fireball is less than the ratio of pair to baryon energy in its initial stages, and that the fraction of the available energy that appears as radiation is generally much less than unity, posing a severe problem for fireball models. A natural way of avoiding these problems is to allow the fireball to impact on the surrounding interstellar medium,⁵⁶ generate one or several shocks (much like the propagation of supernova ejecta), and accelerate particles in the shock environs i.e. convert the kinetic energy of the expansion to non-thermal energy. The resulting ultrarelativistic particles radiate efficiently via synchrotron radiation or inverse Compton scattering⁵⁷ guaranteeing that a large fraction of the initial fireball energy is extracted in the form of radiation. The clumpiness of the interstellar medium provides the GRB time variability as the fireball sweeps up matter on a timescale of order seconds.⁵⁶ The resulting spectra⁵⁷ can have spectral breaks in the GRB energy range, however there is no preferred energy for the breaks, so a broad range of ϵ_B would be anticipated in this scenario. In addition, it is presently unclear that the model can fit the variety of spectral slopes α present in GRB data.

CONCLUSION

This more or less summarizes one theorist's view of the gamma-ray burst dilemma and some recent observations and interesting theoretical work. Clearly the dilemma continues, and gamma-ray bursts remain as enigmatic as ever. It appears unlikely that any stunning new observations may appear on the scene from CGRO, which will largely be devoted to refinement of existing results by increasing the database. However, positive results from searches⁵⁸ for gravitationally-lensed GRBs (by distant galaxies) would provide the smoking gun for cosmological pundits, and any observation of 20–60 keV absorption features by BATSE would excite proponents of the galactic neutron star hypothesis. Neither discovery would conclusively prove that *all* bursts belonged to either population. In fact, searches for photoelectric absorption effects⁵⁹ in the very low energy X-ray spectra of bursts would have the potential to flip the cosmological/galactic coin either way. For theorists, who depend on the observations, the GRB question will almost certainly be resolved by the observations; however development of various cosmological (and galactic) models, particularly in regard to spectral issues, will aid in reducing the plethora of viable models.

REFERENCES

1. Klebesadel, R. W., Strong, I. B. and Olson, R. A.: *Ap. J.* **182**, L85 (1973).
2. Harding, A. K.: *Phys. Rep.* **206**, 327 (1991).
3. Fishman, G. J. et al.: *A&AS* **97**, 17 (1993).
4. Higdon, J. C. and Lingefelter, R. E.: *Ann. Rev. Astron. Astr.* **28**, 401 (1990).
5. Fishman, G. J. et al.: *Ap. J. Supp.* in press, (1994).
6. Kouveliotou, C. et al.: *Ap. J.* **413**, L101 (1993).
7. Paciesas, W. S. and Fishman, G. J., eds.: *Gamma-Ray Bursts*, (AIP 265, NY, 1991).
8. Schaefer, B. E. et al.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).

9. Laros, J.: *Nature* **318**, 448 (1985).
10. Laros, J. et al.: in *High Energy Transients in Astrophysics*, ed. S. E. Woosley (AIP 115, NY, 1984), p. 378.
11. Mazets, E. P. et al.: *Ap. Space Sci.* **60**, 3,85,119 (1981).
12. Matz, S. M. et al.: *Ap. J.* **288**, L37 (1985).
13. Schneid, E. J. et al.: *A&A* **255**, L13 (1992).
14. Sommer, M. et al.: *Ap. J.* **422**, L63 (1994).
15. Murakami, T. et al.: *Nature* **335**, 234 (1988).
16. Meegan, C. et al.: *Nature* **355**, 143 (1992).
17. Mao, S. and Paczyński, B.: *Ap. J.* **388**, L45 (1992).
18. Hakkila, J. et al.: *Ap. J.* **422**, 659 (1994).
19. Hartmann, D.: in *The Lives of Neutron Stars*, eds. A. Alpar and J. van Paradijs (Kluwer Academic, Dordrecht, 1994).
20. Norris, J. et al.: *Ap. J.* in press, (1994).
21. Quashnock, J. M. and Lamb, D. Q.: *MNRAS* **265**, L59 (1993).
22. Narayan, R. and Piran, T.: *MNRAS* **265**, L65 (1993).
23. Hartmann, D. et al.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).
24. Schaefer, B. E. et al.: *Ap. J. Supp.* **92**, in press (1994).
25. Band, D. et al.: *Ap. J.* **413**, 281 (1993).
26. Schaefer, B. E. et al.: *Ap. J.* **393**, L51 (1992).
27. Lightman, A. P. and Zdziarski, A. A.: *Ap. J.* **391**, 643 (1987).
28. Dermer, C. D.: in *Proc. Moriond Workshop 94* (these proceedings, 1994).
29. Zdziarski, A. A. and Lamb, D. Q.: *Ap. J.* **309**, L79 (1986).
30. Zdziarski, A. A., Coppi, P. S. and Lamb, D. Q.: *Ap. J.* **357**, 149 (1990).
31. Baring, M. G.: *Ap. J.* **418**, 391 (1993).
32. Baring, M. G.: *A&A* **225**, 260 (1989).
33. Preece, R. and Harding, A. K.: *Ap. J.* **347**, 1128 (1989).
34. Daugherty, J. K. and Harding, A. K.: *Ap. J.* **273**, 761 (1983).
35. Baring, M. G.: *MNRAS* **244**, 49 (1990).
36. Dermer, C. D.: *Ap. J.* **347**, L13 (1989).
37. Dermer, C. D.: *Ap. J.* **360**, 197 (1990).
38. Baring, M. G.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).
39. Dermer, C. D. and Schlickeiser, R.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).
40. Schmidt, W. K. H.: *Nature* **271**, 525 (1978).
41. Krolik, J. H. and Pier, E. A.: *Ap. J.* **373**, 277 (1991).
42. Harding, A. K.: in *Proc. 2nd Compton Symp.*, ed. N. Gehrels et al. (AIP, NY, 1994).
43. Baring, M. G. and Harding, A. K.: in *Proc. 23rd ICRC* **1**, 53 (1993).
44. Harding, A. K. and Baring, M. G.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).
45. Paczyński, B.: *Ap. J.* **308**, L43 (1986).
46. Eichler, D. et al.: *Nature* **340**, 126 (1989).
47. Mészáros, P. and Rees, M. J.: *Ap. J.* **397**, 570 (1992).
48. Narayan, R., Piran, T. and Shemi, A.: *Ap. J.* **379**, L17 (1991).
49. Goodman, J.: *Ap. J.* **308**, L47 (1986).
50. Cavallo, G. and Rees, M. J.: *MNRAS* **183**, 359 (1978).
51. Woosley, S. E.: *Ap. J.* **405**, 273 (1993).
52. Usov, V.: *Nature* **357**, 472 (1992).
53. Piran, T., Shemi, A. and Narayan, R.: *MNRAS* **263**, 861 (1994).
54. Carrigan, B. J. and Katz, J. I.: *Ap. J.* **399**, 100 (1992).
55. Shemi, A. and Piran, T.: *Ap. J.* **365**, L55 (1990).
56. Mészáros, P. and Rees, M. J.: *Ap. J.* **405**, 278 (1993).
57. Mészáros, P. and Rees, M. J.: *Ap. J.* **418**, L59 (1993).
58. Nemiroff, R. J. et al.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).
59. Schaefer, B. E.: in *Gamma-Ray Bursts*, ed. G. Fishman et al. (AIP 307, NY, 1994).

EGRET GAMMA-RAY SOURCES

Isabelle GRENIER, Equipe EUROPA, Université Paris VII et
Service d'Astrophysique, Centre d'Etudes de Saclay

Abstract

The EGRET telescope has observed so far 124 γ -ray sources in the sky among which 5 pulsars and 38 active galactic nuclei have been recognized and shed new light on these objects. This leaves 81 puzzles to solve, with interesting or promising counterparts in few cases. For instance, near the Galactic plane, several sources lie in the direction of known radio supernova remnants. At medium latitudes, half of the sources seem to belong to the local belt of giant molecular clouds and young OB associations.

THE SOURCE CATALOG

The EGRET instrument has been operating from 30 MeV to 30 GeV on board the Compton Observatory since 1991, together with three other detectors which offer a unique chance to record the often variable γ -ray sources simultaneously over 6 decades of energy. The first 19 months of the mission have been devoted to a whole sky survey to map the diffuse Galactic emission borne out of cosmic-ray interactions with the interstellar matter and radiation field, and to search for point-like sources. The negligible instrumental noise also allowed to detect the extragalactic γ radiation. The details of the telescope performances have been summarized by Thompson et al.¹⁾. The spatial distribution and intensity of the diffuse Galactic emission can be reliably modeled from our knowledge of the interstellar gas distribution (as mapped at 21 cm for H atoms, at 2.6 mm for CO thus H₂ molecules, and in H α for protons) and from various assumptions on the cosmic-ray distribution. The EGRET team favours a cosmic-ray density that is coupled to the gas density through the frozen-in magnetic field and is estimated from equipartition of energy between the

magnetic field and the cosmic-ray and gas pressures ². The old COS-B team preferred to measure the cosmic-ray density at various places and found a weak density gradient through the Galaxy ³. Both models must fit the H₂-to-CO conversion factor. As a first approximation, the two models rather well fit the observed diffuse emission which can therefore be subtracted to look for point-like excesses. This is why many (14) of the COS-B sources have been confirmed by EGRET. To a higher degree of precision, the models might lead to significant differences since they use different powers of the matter density. So, the two models should be carefully and quantitatively compared before looking for faint sources near the Galactic plane. The first EGRET catalog ⁴ safely concentrates on the brightest sources, most of them at large latitudes, but some weak excesses near the plane might slightly move or even disappear when improving the diffuse model.

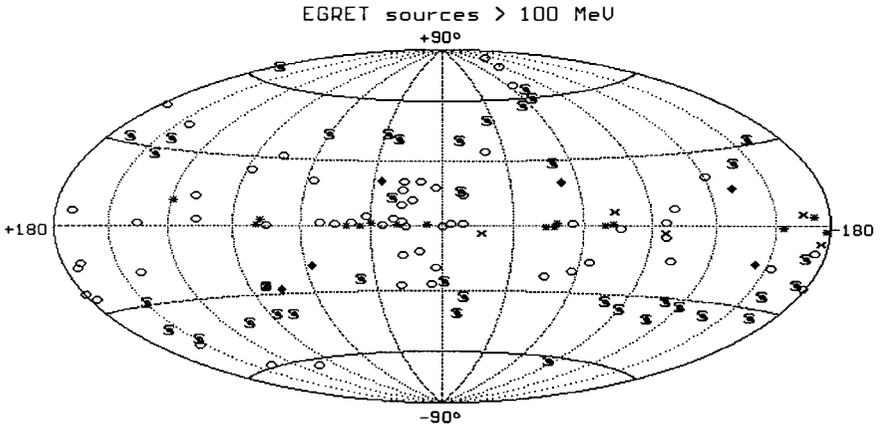


Figure 1: (l,b) positions of the 5 identified pulsars (x), sources toward a radio pulsar and/or a supernova remnant (*), identified AGN (§), possible AGN (◆) and unidentified sources (o)

Bearing this limitation in mind, a total of 124 sources have been recorded and are displayed in Fig. 1. 38 spectacular active galactic nuclei (AGN) have caused a lot of upheaval at the time of their discovery. EGRET really opened our eyes to the extragalactic world at high-energy. Six other sources shine near the direction of an AGN, but not close enough or too near other counterparts to be firmly identified. The three brightest sources in the GeV sky are the familiar Vela, Geminga, and Crab pulsars. Two fainter ones, PSR1706-44 and 1055-52, clearly emit pulsed γ rays, but five other sources which are seen in the direction of a known radio pulsar, are not pulsed. In the same spirit, sources have been marked on Fig. 1 when located near the direction of a radio supernova remnant and Fig. 2 gives close-up views to judge the spatial coincidences. A few cases look rather promising, like CTA1 away from the confused regions of the Galactic plane. Supernova shells are likely to be soft γ -ray sources because the electrons seen in radio synchrotron emission should also emit γ rays by bremsstrahlung with the compressed gas ⁵. Above 100 MeV, supernova remnants being the only respectable sources of cosmic rays below 10^{15} eV should indirectly produce γ rays

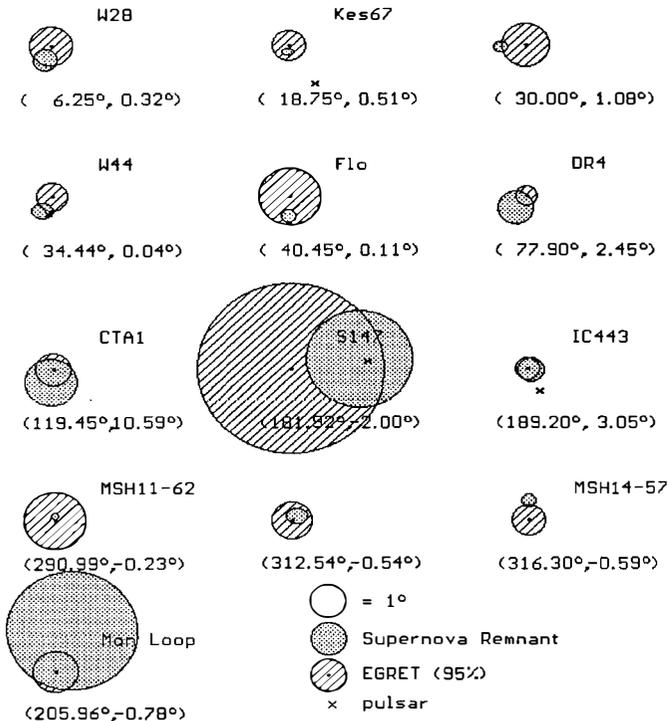


Figure 2: Examples of spatial coincidences between EGRET sources, supernova remnants, and pulsars

hundred parsecs if the remnant expands in a low density (1 cm^{-3}) medium. A more promising situation arises when a nearby dense cloud serves as a target for the nucleons ⁷⁾. DR4 and W28 give nice examples of such combinations ⁸⁾. The flux predicted for DR4 is consistent with the COS-B and EGRET observations ⁷⁾. The proximity of a molecular cloud would also be needed to explain the sources seen toward W44 and IC443. On the contrary, the source toward CTA1 requires an ambient density of $\sim 10 \text{ cm}^{-3}$ that is quite easily achieved, without a dense cloud, in the HI component of the Local Arm where CTA1 is located. The same low-density medium suffices around S147 and the Monoceros Loop, even though the latter interacts with the Rosette Nebula ⁹⁾. The more distant remnants MSH11-62, MSH14-57, and Kes67 should be expanding, very unlikely, inside very dense clumps (10^3 cm^{-3}) to be detected by EGRET. Moreover, current expansion and acceleration theories do not apply to a shell in a dense environment.

Of particular interest is the improving visibility of the γ -ray supernova remnants with increasing energy because their spectrum is harder than that of the underlying Galactic emission ⁶⁾. The predicted fluxes are accessible to the present TeV telescopes in a reasonable time. So, the search for TeV supernova remnants becomes a new prospect as well as an extremely important

via the interaction of accelerated nucleons and the ambient gas. Drury et al. ⁶⁾ have used their acceleration model and the tight constraint of the total Galactic cosmic-ray power to predict the γ -ray luminosity of a typical remnant. Interestingly, the luminosity peaks about a thousand years after the initial explosion and remains high for 10^5 years or more. But, the predicted flux falls below EGRET's reach beyond several

challenge since it would provide the first direct measurement of the nucleon acceleration rate by these objects.

The number of remaining unidentified sources is large enough to try to gather information on their origin from their spatial distribution. Fig.3 shows their latitude distribution. The central peak contains the sources that were just discussed because of their possible pulsar or supernova remnant counterpart, plus a larger number of totally unidentified objects. Their concentration along the Galactic plane clearly indicates that they lie beyond 1 or 2 kpc, a distance scale consistent with the remnants presented in Fig.2. Indeed, closer remnants would be resolved by EGRET.

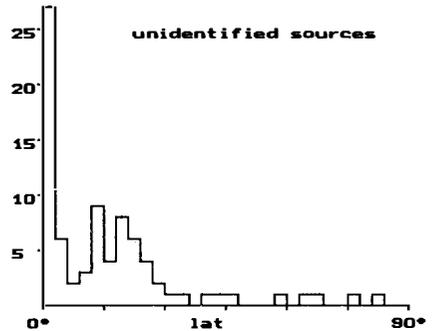


Figure 3: latitude number distribution of all sources but the identified pulsars and AGN.

The wide secondary peak around $b = 15^\circ - 20^\circ$ in the latitude profile is real, i.e. not consistent with a random fluctuation of an isotropic, AGN-like, population. The (l,b) source positions above $b = 3^\circ$ have been tested against various parent distributions using a likelihood analysis taking into account the effects of the non-uniform exposure and non-uniform detection threshold due to the structure of the underlying Galactic and extragalactic emissions¹⁰. The combination of an isotropic population plus a component distributed as Galactic pulsars, or as the HI column densities, or else as the total HI+H₂ column densities have been considered. The analysis significantly favours the last solution which implies that half of the sources above $b = 3^\circ$ belong to the local interstellar medium and, more specifically, to the giant cloud complexes of the Gould Belt, as can be seen on Fig.1. These very interesting sources have therefore a low γ -ray luminosity of $10^{33} - 10^{34}$ erg/s and their location with respect to young objects like OB associations, HII regions, and X-ray sources embedded in the clouds should be carefully studied to circumvent their nature¹¹.

ACTIVE GALACTIC NUCLEI

EGRET will be remembered for discovering the γ -ray emission from AGN. The catalog⁴⁾ presents a list of 38 AGN that shine up to several GeV, either as steady sources over months or blinking in and out of EGRET's reach. Ten of them have occasionally exhibited a rapid change in their intensity, by a factor of 2 to 10, over a few days or weeks¹²⁾. The proposed identifications are mostly based on the presence of a bright, flat spectrum, radiogalaxy within the EGRET error box and it is strengthened in half of the cases by peculiar AGN features such as radio and/or optical flares and jets, and superluminal motions. Because, even at high latitudes, a 1° error box may contain other counterparts or several AGN, some identifications need confirmation, the most reliable way being to look for correlated time variations at different wavelengths.

A thrilling surprise came from the huge distance scale over which γ -ray AGN have been

detected, namely from $z = 0.03$ to 2.3 . The source sample is small, but the distance scale large enough to trigger cosmological studies. For instance, a V/V_{\max} test shows evidence for an evolution of the γ -ray luminosity with z . An AGN luminosity function "de-evolved" to the present epoch could be computed¹³⁾. Interest was further stimulated when realizing that the νF_{ν} spectrum of these AGN peaks in the MeV-GeV band. Against previous ideas, the γ -ray energy release dominates over that of the "blue bump" and X-ray emission. Power-law spectra with indices ranging from 1.5 to 2.6 nicely fit to the γ -ray data¹⁴⁾.

According to present wisdom, the ultimate source of all AGN activity is a supermassive black hole of 10^6 - $10^8 M_{\odot}$, accreting 10 to $100 M_{\odot}$ per year to account for the overall luminosities. While UV and X-ray photons are emitted somewhat isotropically from the heated accretion disc, γ rays must be born in collimated jets in most cases. Indeed, cosmological distances and rapid time variability imply that very large luminosities of typically 10^{48} erg/s would arise isotropically from a compact source region less than 0.01 pc in size, flooded moreover with softer disc radiation. The source region would then be quite optically thick to photon-photon pair production ($\gamma\gamma \rightarrow e^+ e^-$ and $\gamma X \rightarrow e^+ e^-$) unless the γ rays are highly beamed to reduce their cross-section. Bulk relativistic motion of the emitting particles in a jet provides a natural way to collimate the radiation, avoid absorption, and enhance the observed flux. The required Lorentz factor is consistent with the radio observations of jets and superluminal motions¹⁵⁾.

Collimation of the γ radiation gives support to the recent "grand unification of the AGN zoo". Only blazars (named after BL Lac galaxies and quasars) and no Seyfert galaxies have been detected by EGRET because blazar jets are directed close to our line of sight whereas Seyfert jets are seen at large angles. γ rays from the latter, if there, would miss the Earth. Only upper limits on the beaming factors can be derived from opacity considerations. Typical values are expected in the 10^{-3} - 10^{-2} range. With this reducing factor f in mind, the closest detected AGN, Mrk421, releases $2 f 10^{44}$ erg/s above 100 MeV and its steady radiation needs not be beamed. On the contrary, the luminosity of PKS0528+134, at $z = 2.06$, reaches $8 f 10^{48}$ erg/s and is variable. Such luminosities can be achieved by current AGN models.

A large γ -ray flare from 3C279 lasted about 10 days in June 1991 while the source was in a low X-ray state¹⁶⁾. Yet, one class of AGN jet models uses relativistic electrons, accelerated by MHD shocks inside the jet, to boost soft photons from the disc thermal emission (or from a halo of scattered radiation surrounding the disc) up to γ -ray energies. A flare would thus result from a sudden increase in the primary particle density or energy. Compton scattering of synchrotron emission born inside the jet provides another means to produce γ rays. Here, the whole jet emission from UV to γ rays should vary simultaneously. In both cases, the radio emission is born further out in the jet and the radio flare would lag the high-energy one. Variability may be caused by sporadic injection of particles in the jet or by their sporadic acceleration by shocks propagating in a turbulent jet. Instead of shocks, another class of models uses huge ($> 10^{20}$ eV) potential drops at the base of the jet to accelerate primary electrons. A third class considers hadron-driven jets. The accelerated protons interact with the ambient matter or with photons to produce pions. The neutral

ones decay in two γ rays, the charged ones produce electrons that up-scatter disc or jet soft photons up to GeV energies. This list just gives a flavour of the models that are now hotly debated. In this book, C. Dermer gives a nice review of all the scenarii and their weaknesses, with many references. An important idea is that correlated multi-wavelength observations of the variability will undoubtedly help to discriminate between the models. The γ rays are crucial to probe the mechanisms at work in the jet formation near the central engine since they carry away most of the jet energy, at least for e^\pm powered jets. Understanding them will also help to clarify why quite a few blazars have not been seen by EGRET while they closely resemble those detected.

The very close Mrk421 galaxy has been observed up to TeV energies by the Whipple telescope ¹⁷. Its νF_ν spectrum is impressively nearly constant from the optical to the TeV band ¹⁸. The lack of detection of brighter GeV AGN, such as 3C279, corroborates the idea that TeV photons are effectively absorbed by the diffuse IR radiation over intergalactic distances. The optical depth to pair production ranges from ~ 1 at $z = 0.1$ to ~ 40 at $z = 1$ ¹⁹. The trip should be safer for a 100 GeV and a 10 GeV photon against the fainter optical and UV background radiation. But a rapid estimate shows that the optical depth exceeds 1 beyond $z \sim 0.2$ for the 200 GeV photons to be observed by the future CAT telescope in France. In the best of worlds, this absorption would be a useful tool to measure the poorly known diffuse background radiation in the infra-red, visible and ultra-violet, but it will be difficult to distinguish the spectral breaks induced by intergalactic absorption from the breaks inherent to the physics of these highly variable sources. Already at GeV energies, each AGN has its own personality.

Table 1: γ -ray pulsar characteristics

<i>Name</i>	<i>P ms</i>	<i>age kyr</i>	<i>B_{pc} TG</i>	<i>c/R_{km}</i>	<i>γ sp. index</i>
Crab	33.3	0.94	7.6	1592	2.15 ± 0.04
Vela	89.3	11	6.7	4263	1.70 ± 0.02
1706-44	102.4	17	6.3	4891	1.72 ± 0.08
1055-52	197.1	535	2.2	9411	1.18 ± 0.16
Geminga	237.1	342	3.3	11320	1.50 ± 0.08

GAMMA-RAY PULSARS

This club holds five members, listed in Table 1, which radiate a large fraction of their spin-down power above 1 MeV. Their νF_ν spectrum peaks in the GeV range (see Geminga in Fig.4). Only the Crab pulsar

presents a flat spectrum from 0.1 MeV to 5 GeV. The five pulsars differ in many ways: in their age, the size of their magnetosphere thus the strength of the magnetic field threading it, their spin-down power, the beaming of their radiation at various wavelengths (Fig.5), and in the energy spectrum of their beams and their variability over weeks and years. Finally, the Crab and PSR1706-44 pulsars power an active γ -ray nebula around them up to TeV energies while the others don't. To face this diversity, our tools are the knowledge of the rotational period P and its derivative, a typical moment of inertia of 10^{45} g/cm², and the properties of synchrotron and curvature radiation, inverse Compton scattering, and pair production processes ($\gamma B \rightarrow e^\pm$ and $\gamma\gamma \rightarrow e^\pm$). We are allowed ... two free parameters, i.e. the inclination angle χ between the magnetic and rotation axes and the aspect angle from Earth, ... and a lot of imagination!

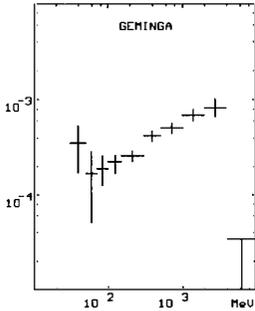


Figure 4: $E^2 S(E)$ spectrum of Geminga in $\text{MeV cm}^{-2} \text{s}^{-1}$

Fig.5 displays the impressive diversity of light curve shapes among the five pulsars and at different wavelengths. Amazingly, the Crab double peaks are synchronized at all energies and their separation in phase is exactly the same as for Vela. It should be noted that nonthermal emission dominates at all wavelengths but in X rays. Except for Crab, the X rays mostly come from the heated (10^6 K) surface of the neutron star and thus are weakly modulated.

One should be extremely careful with so-called γ -ray efficiencies, i.e. the ratio of the γ -ray luminosity to the spin-down power, since the emission geometry is unknown. Although an apparent increase in efficiency with age exists when assigning 1 or 4π sr to all pulsars ²⁰, reality plays with fan and pencil beams to probably scatter the points around. The increase in emission hardness with age looks like a more reliable trend (Table 1). The COMPTEL results in the MeV range show that power-law spectra are valid in a limited domain; they break down around 0.1 MeV for Crab, 10 MeV for Vela, Geminga, and PSR1706-44, and 100 MeV for PSR1055-52 and a sharp cut-off occurs at several GeV for all (Fig.4). Moreover, the spectra change with phase. For Crab and Vela, the interpeak emission is harder than from the main peaks ^{21, 22, 23}. The situation is more confused for Geminga ^{24, 25}. Interestingly, core and wings of the Vela peaks could be separated to show that the emission gets softer when moving out of a peak beam ²³. These spectra generally vary with time.

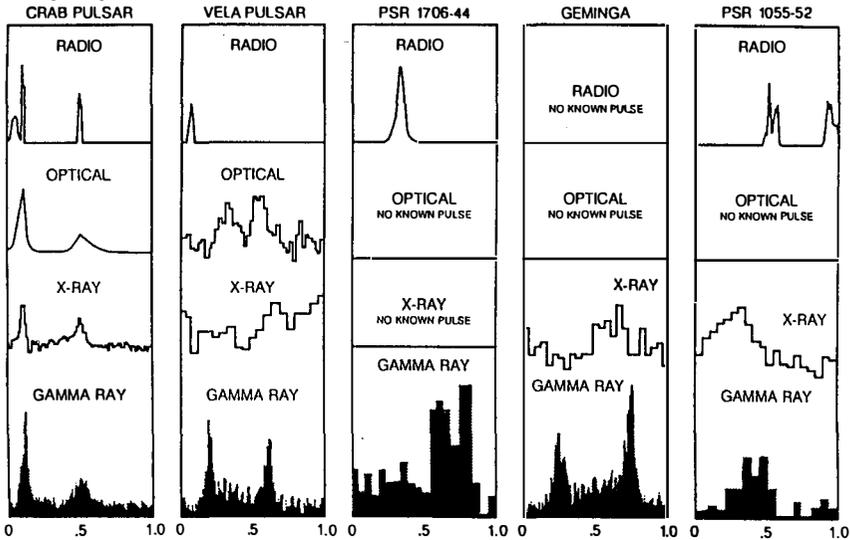


Figure 5: Pulsar lightcurves at different wavelengths (references in ²⁰)

Thankfully for the calibration of γ -ray satellites, the Crab does not vary much. Its emission, however, slightly softened between COS-B and EGRET, and the relative heights of the two peaks changed ^{21, 26}. The flux from Vela doubled within 11 days in 1975 and again in 1981, due to a clear softening of the emission at all phases but the first peak. These events triggered the idea that stable and narrow peaks may appear on top of a third variable component that covers most of the light curve and is weakly modulated, perhaps in phase with the optical pulses ²². Since then, EGRET has seen Vela in a low state of hard spectral index. The giant glitches have notably no effect on the output flux ²³. Geminga mimics Vela: the emission significantly softened once ^{24, 27} (except in the first peak), and EGRET saw it in 1991 in a harder state than in the 1970s ²⁵. The COS-B source 2CG342-02, that has been later identified by EGRET as the pulsar PSR1706-44, showed signs of variability in the 70s ²⁸, while it remained stable in 1991-92.

Despite the gravitational pull of the neutron star and because of the electric field \mathbf{E} induced by the rotating magnetic field \mathbf{B} , a pulsar magnetosphere must be filled with a force-free plasma of charge density $\rho_{\text{GJ}} \sim -\Omega \cdot \mathbf{B} / 2\pi c$ for $\chi=0$ ²⁹. Because charges leave the magnetosphere along the open field lines, regions called gaps are created in which $\rho \ll \rho_{\text{GJ}}$, so acceleration along \mathbf{B} takes place ($\mathbf{E} \cdot \mathbf{B} \neq 0$). One such gap may be maintained near the star polar cap because the star crust mostly gives away electrons. The ions screening effect causes the supply of charge to fall short of ρ_{GJ} above the cap. Electrons are accelerated to typically 10 TeV. They radiate curvature radiation γ rays that are subject to magnetic pair production thus initiating a synchrotron-pair cascade ³⁰. The gap closes where the pair density reaches ρ_{GJ} ($\mathbf{E} \cdot \mathbf{B}=0$). A variant polar cap model uses 100 GeV primary electrons to boost beams of optical or X-ray photons from the star up to γ -ray energies by resonant Compton scattering ³¹. The development of a cascade follows as before. In both cases, emission is produced within a narrow cone of aperture $\sim 1.5 \theta_{\text{pc}}$ (with $\theta_{\text{pc}} \approx$ polar cap radius $R_{\text{pc}}/\text{star radius}$). For nearly aligned rotators ($\chi - \theta_{\text{pc}}$) and aspect angles of the same order, all the γ -ray light curves of Fig.5 can be reproduced.

A naive estimate of the maximum power that can be radiated is $L_{\text{max}} \sim E_0 \rho_{\text{GJ}} c \pi R_{\text{pc}}^2$ where all possible charges, $\rho = \rho_{\text{GJ}}$, above the polar cap are accelerated and radiate $E_0 \sim 10$ TeV away as γ rays. With a vacuum B field at the pole estimated as usual from the slowing down of the pulsar and dipole radiation for $\chi = \pi/2$, one finds $L_{\text{max}} \propto B P^2 = 8.8 \cdot 10^{37} E_{0,\text{reg}} P^{-1.5} (dP/dt)^{0.5}$ erg/s. It is only 50 times larger than the 0.1-10 GeV luminosities L_γ observed by EGRET in the narrow cone just described (Fig.6a). Decreasing E_0 would yield too soft spectra, but decreasing $\rho \ll \rho_{\text{GJ}}$ is necessary to get an accelerator! Besides, only a fraction of E_0 may be radiated by the inefficient curvature radiation process. The primary γ spectrum is reprocessed to lower energies by the cascades, but it has little effect on L_γ as shown by Monte Carlo simulations ³²: $L_\gamma \propto B^{0.95} P^{-1.7} \propto P^{-1.225} (dP/dt)^{0.475}$. The five detected pulsars nicely follow this trend (Fig.6a). Many others should have been seen by EGRET unless their beam sweeps past us. The second polar cap model ³¹ uses a more sophisticated $E_0 \propto P^{-0.75} (dP/dt)^{0.25}$ and predicts a luminosity $L_\gamma = 5.8 \cdot 10^{39} P^{-2.25} (dP/dt)^{0.75}$ erg/s which does not fit the observed trend as well and falls short of detected luminosities even though $\rho = \rho_{\text{GJ}}$ (Fig.6b). By using a higher than usual B field, implied by $\chi \sim 1.5 \theta_{\text{pc}}$, in the

prediction of $L_{\max}(P, dP/dt)$, both fits get significantly worse.

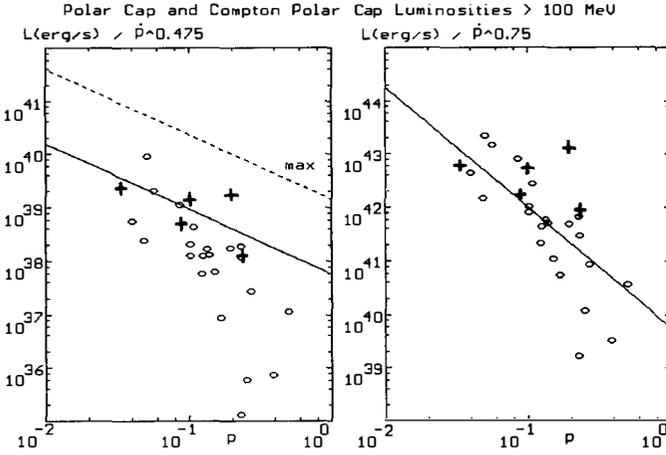


Figure 6: 0.1-10 GeV luminosities L , observed (x) or 3σ upper limits (o), have been computed for known pulsars assuming emission within a cone of aperture $1.5 \theta_{pc}$, see text; (a) $L/(dP/dt)^{0.475}$ are plotted as function of period P (seconds) and compared to the maximum luminosity expected from the standard polar cap model (dotted line) and an eye fit of the same power law through the detected pulsars (solid line); (b) $L/(dP/dt)^{0.75}$ are plotted as a function of P and compared to the prediction of the comptonized polar cap model (solid line).

on soft photons. These crossing beams are used to produce pairs and close the gap since magnetic pair production is ineffective out there. A symmetrical gap exists from the other pole, so four fan beams of radiation are available with much light aberration to construct almost any light curve shape³⁴. The radiation pattern is not predictable at present to compare predicted and observed luminosities. With increasing periods, a larger fraction of the spin-down power must be used to sustain the current through the light cylinder via pair production above the gap. The latter grows in size and the expanded power is radiated as harder γ rays because curvature radiation dominates over cascade radiation. The accelerator will eventually be quenched at a death line³⁵: $5 \log B - 12 \log P = 69.5$.

Ironically, the discovery of new γ -ray pulsars has not favoured one type of gap against the other, but has threatened and reinforced both. Polar cap models would be safer without Crab: optical synchrotron emission requires $B \ll 10^8$ G, therefore an emission site of all radio to γ radiation too high above the cap. The spectra of Geminga and PSR1055-52 seem harder than direct primary curvature radiation. Comptonization as in the second polar cap model would help here, but the comptonization rate may be actually too low because of the lack of soft, beamed photons. The limited visibility of the narrow beam originating from a polar cap will become a severe problem whenever new pulsars are discovered. On the contrary, outer-gap emission can be seen from many

Another gap may be found near the light cylinder, just above the closed magnetosphere³³. Because ρ_{CI} changes sign along these last open field lines, the charges that escape out cannot be replenished from inner regions and $\rho \ll \rho_{CI}$. In this long gap, electrons and positrons can be accelerated up to 100 TeV in opposite directions. Both emit beamed γ rays from curvature radiation or inverse Compton

orientations. In fact, the radiation pattern has so many facets that a single γ -ray pulse from PSR1706-44 and PSR1055-52 may be difficult to explain. Both Geminga and PSR1055-52 are dangerously close and in fact a little beyond the outer-gap death line. On the other hand, the slow Geminga exhibits wider peaks than Crab or Vela, as expected from a thicker gap.

In this friendly competition, both gaps face a common problem: the unknown origin of the variability and of its phase dependence. Interpeak emission can be understood as the inner or outer edges of a hollow polar cap beam as well as one or two of the four outer gap beams. But how can it behave independently of the rest of the light curve? Are both gaps active at the same time? A careful comparison of the presumably distinct spectra originating from these different gaps should help to answer this question. Before EGRET, we had two γ -ray pulsars and two gaps. The same two gaps have survived against five pulsars. So, either we lack imagination or there is hope that ideas may soon converge to find out the exact location of the acceleration sites in a pulsar magnetosphere.

References

1. Thompson D.J. et al., 1993, *ApJ Supp.* 86, 629
2. Bertsch D.L. et al., 1993, *ApJ* 416, 587
3. Strong A.W. et al., 1988, *A&A* 207, 1
4. Fichtel C. et al., 1994, *ApJ Supp.* in press
5. Blandford R.D. and Cowie L.L., 1982, *ApJ* 260, 625
6. Drury L. O'C., Aharonian F.A. and Völk H.J., 1994, submitted to *A&A*
7. Aharonian F.A., Drury L. O'C. and Völk H.J., 1994, *A&A* 285, 645
8. Pollock A.M.T., 1985, *A&A* 150, 339
9. Huang Y.L., 1985, PhD Thesis, Columbia University
10. Grenier I.A., 1994, submitted to *A&A*
11. Montmerle T., 1979, *ApJ* 231, 95
12. Michelson P.F. et al., 1993, "the 2nd Compton Symposium", eds. Fichtel C. et al., AIP
13. Chiang J. et al., 1993, "the 2nd Compton Symposium", eds. Fichtel C. et al., AIP
14. Hartman R.C. et al., 1993, "the 2nd Compton Symposium", eds. Fichtel C. et al., AIP
15. Mattox J.R. et al., 1993, *ApJ* 410, 609
16. Kniffen D.A. et al. 1993, *ApJ*, 411, 133
17. Punch M. et al., 1992, *Nature*, 358, 477
18. Lin Y.C. et al., 1993, "the 2nd Compton Symposium", eds. Fichtel C. et al., AIP
19. Stecker F.W., De Jager O.C., and Salamon M.H., 1992, *ApJ*, 390, L49
20. Thompson D.J., 1993, "The 2nd Compton Symposium", eds. Fichtel C. et al., AIP
21. Nolan P.L. et al., 1993, *ApJ* 409, 697
22. Grenier I.A., Hermsen W., and Clear J., 1988, *A&A* 204, 117
23. Kanbach G. et al., 1994, *A&A*, in press
24. Grenier I.A. et al., 1994, *ApJ Supp.*, 90, 813
25. Mayer-Haefelwander H. et al., 1994, *ApJ*, in press
26. Clear J. et al., 1987, *A&A* 174, 85
27. Grenier I.A., Hermsen W., and Henriksen R.N., 1993, *A&A* 269, 209
28. Grenier I.A. et al., 1991, AIP Proc. 220 "High-energy γ -ray astronomy", p 303
29. Goldreich P. and Julian W.H., 1969, *ApJ* 157, 869
30. Daugherty J.K. and Harding A.K., 1994, *ApJ*, in press
31. Dermer C.D. and Sturmer S.J., 1994, *ApJ* 420, L75
32. Harding A.K., 1981, *ApJ* 245, 267
33. Cheng K.S., Ho C., and Ruderman M., 1996, *ApJ* 300, 500
34. Halpern J.P. and Ruderman M., 1993, *ApJ* 415, 286
35. Chen K. and Ruderman M., 1993, *ApJ* 402, 264

Cosmic ray acceleration and radio emission from SNR 1987A

P. Duffy¹, Lewis Ball², J. Kirk¹

¹ Max Planck Institut für Kernphysik, Heidelberg, Germany

² Research Centre for Theoretical Astrophysics, University of Sydney, N.S.W. 2006, Australia

Observations of radio emission from SNR 1987A can be explained by the diffusive acceleration of electrons at the expanding blast wave only if its compression ratio is about 2.7 rather than the value of 4 expected for strong shocks. We therefore put forward a model where both protons and electrons are accelerated. The accelerated protons contribute significantly to the fluid pressure and exert a back reaction on the fluid and modify the shock. The electrons then see a weakening gas subshock where they are accelerated before emitting synchrotron radiation in the swept up magnetic field. The resulting spectrum is compared with observations.

1. Introduction

In February of 1987 the explosion of a star in a neighbouring galaxy, the Large Magellanic Cloud, was detected as the brightest supernova seen at earth in almost four hundred years. Although this supernova (SNR 1987A) has appeared in all bands of the electromagnetic spectrum since ^[1] we are primarily interested here in the emission at radio frequencies. After an initial radio flare which lasted for a few days, SNR 1987A remained radio quiet until July 1990 when it reappeared ^[2] at frequencies between 843MHz and 4.8GHz. The intensity of this emission has continued to rise up to the present day. The frequency spectrum which was initially quite steep has flattened off so that the flux density, $F(\nu)$, now has a power law dependence on frequency $F(\nu) \propto \nu^{-\alpha}$ with $\alpha \sim 0.9$. This emission is thought to be the synchrotron radiation given off by energetic electrons in the magnetic field around the expanding blast wave of the explosion. It can be estimated ^[3] that electrons with energies up to 10^9 eV are needed to explain this emission, if the magnetic field at the blast wave was roughly 1mG when the emission switched on roughly 1200 days after the explosion. The first question we must ask is what mechanism could produce electrons with these energies in the time available between the explosion and the onset of radio emission.

2. Electron Acceleration

One possibility put forward by Ball and Kirk ^[3] is that electrons are accelerated at the expanding blast wave of the explosion by the diffusive shock acceleration mechanism. Some preliminary remarks on particle transport in astrophysical magnetic fields are necessary before this mechanism can be properly described. Typically the magnetic fields in the interstellar medium have a random as well as a mean component so that energetic particles do not simply follow a helical trajectory in the average field but are also scattered by the magnetic turbulence. Consequently after travelling for some mean free path λ , which can be related to the magnetic field and particle energy, particles forget their original direction of motion. The result is that energetic particles moving at speed v will diffuse with a diffusion coefficient given by $\kappa = \lambda v/3$. Of course a field that is, in some frame of reference, purely magnetic everywhere is not useful from an acceleration point of view, for this we need to turn to shock waves (although there are other astrophysical acceleration sites and mechanisms which shall not concern us here).

For simplicity a shock wave is defined as a sharp transition across which a fluid's density, pressure and velocity change discontinuously. Here we are interested in a star which explodes and ejects mass at some speed into the surrounding medium. The blast wave then sweeps up the undisturbed (upstream) medium which is compressed and heated to produce the post-shock (downstream) region. If we move outwards at the speed of the expanding shock front (i.e. we are in the shock rest frame) the upstream region moves towards the shock at speed v_1 where it is decelerated to a downstream speed v_2 . Consider now energetic particles ($v \gg v_1, v_2$) that initially reside in the upstream region. Without diffusion these particles would be convected through the shock downstream where they could not return upstream. However, since diffusion is, in general, an important process for energetic particles they can be scattered back and forward across the shock many times, by the magnetic turbulence frozen into the fluid flow, before escaping downstream. With an energetic particle distribution that is close to isotropy both upstream and downstream of the shock the average cycle time to cross and re-cross the shock can be calculated to be $t_c = 4\kappa(1/v_1 + 1/v_2)/v$ while the probability per cycle of being convected away from the shock downstream is $P_{\text{esc}} = 4v_2/v$. Since there is a relative speed of $(v_1 - v_2)$ between the upstream and downstream media, or more precisely between the scattering centers that are frozen into the respective media, it can be shown ^[4] that particles will gain a small amount of momentum every cycle. The fractional increase in momentum is given by $\Delta = 4(v_1 - v_2)/3v$. Consequently we have an acceleration process competing with a possibility of escape downstream away from the shock. We assume that electrons are injected at some rate Q with an initial momentum p_0 . Defining $N(p)dp$ to be the number of electrons per unit volume in the vicinity of the shock with momenta between p and $p + dp$ we can write

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial p} \left(\frac{p\Delta}{t_c} N \right) + \frac{P_{\text{esc}}}{t_c} N = Q\delta(p - p_0) \quad (1)$$

Ball and Kirk solved this equation with a momentum independent diffusion coefficient, a steady state shock structure and a constant level of electron injection after some time t_a . The spectrum of electrons is then a power law $N(p) \propto p^{-q}$ between p_0 and $p_0 \exp((t - t_a)\Delta/t_c)$ where $q = (\rho + 2)/(\rho - 1)$ with $\rho = v_1/v_2$ the compression ratio of the shock. It is one of the appealing features of diffusive shock acceleration that the power law spectral index depends only on the shock compression ratio.

Once these electrons escape downstream they suffer adiabatic losses in the expanding, constant speed downstream medium. It is also in this region that they will emit synchrotron radiation in the magnetic field. A power law spectrum of electrons $N(p) \propto p^{-q}$ will produce power law synchrotron radiation with $F_\nu \propto \nu^{(1-q)/2}$ [5]. Since the energetic electrons are accelerated diffusively at the shock the spectral index of the radio emission depends only on the shock compression ratio by $\alpha = 3/2(\rho - 1)$. For the case of a supernova explosion we should have a strong shock wave expanding into an undisturbed medium with a non-relativistic equation of state. Thus the compression ratio ought to take its maximum value of 4 implying a radio spectral index of 0.5. This is in conflict with the radio observations of SNR 1987A which has a steeper index of 0.9 (implying a compression ratio of 2.7). Therefore, either the radio emission from SNR 1987A is not due to electrons accelerated by the diffusive shock mechanism or alternatively our picture of a strong blast wave expanding into an undisturbed medium is incorrect. For the remainder of this paper we address the second possibility, in particular how the acceleration of ions could alter the hydrodynamics and hence the compression ratio to give a steeper radio spectral index. In the next section we consider this effect for the simplest case of a single species proton population.

3. Proton Acceleration

The acceleration model for electrons presented above also applies to energetic protons although with one important difference. While accelerated electrons may be treated as test particles (they exert only a negligible force on the fluid), protons can have an important effect on the hydrodynamics. In what follows we refer to the proton population as Cosmic Rays (CRs) by way of analogy with our own galaxy where the nucleonic component of galactic cosmic radiation is believed to originate in supernova remnants. The pressure exerted on the fluid then has two components; the internal pressure of the thermal gas itself (P_G) and also that of the non-thermal CRs (P_C). To see the effect of P_C on the hydrodynamics consider an expanding blast wave which accelerates protons diffusively in the manner described above. The upstream medium will then contain a diffusive tail of protons falling off exponentially away from the shock with scale length $L \sim \kappa/v_1$ (where κ is now the proton diffusion coefficient). Therefore the blast wave does not quite encounter an undisturbed medium but rather one that has been accelerated by the upstream spatial gradient of P_C . Our picture of a strong shock expanding into a constant

speed upstream medium is therefore modified to one of a weaker subshock sweeping up material that has already been accelerated by the CRs. The compression ratio of the evolving subshock will be less than 4 which would then accelerate test particle electrons and produce synchrotron emission with a spectral index steeper than 0.5, which is what we require to explain the radio observations of SNR 1987A. Of course all of this depends on P_C being a dynamically important component of the total pressure; when $P_C \ll P_G$ our modification effects are unimportant. We need to examine the possibility that P_C can evolve quickly during the first few years after the explosion to a level that is comparable to P_G . In what follows test particle theory, i.e. before P_C alters the hydrodynamics, is used to estimate the initial timescale for the evolution of P_C .

We assume that prior to explosion we assume that the star gave off a freely expanding, constant speed stellar wind as a result of mass loss from its surface. The upstream density at the time of explosion is then proportional to r^{-2} , (where r is the radius) with an azimuthal magnetic field (B) dropping off as $1/r$. We assume that once the star explodes then the shock expands ballistically $r_s = u_s t$ where r_s is the shock radius and u_s is the constant shock speed. Just after the explosion at some time t_0 we start injecting protons at a constant rate Q with some initial momentum p_0 at the expanding shock front. Since the compression ratio of the shock is 4, as long as P_C does not alter the shock structure, diffusive shock acceleration will produce a spectrum of CR protons at the shock with $N(p) \propto p^{-2}$ (see section 2) from p_0 to some time dependent cut-off $p_{\max}(t)$. If the area of the shock front is $A(t)$ then, as long as they behave as test particles, the time dependence of the CR pressure is given by

$$P_C \propto \frac{1}{A(t)} \int_{p_0}^{p_{\max}(t)} \frac{pc}{3} \left(\frac{p}{p_0} \right)^{-2} dp \quad (2)$$

where we have assumed all of the protons to be ultrarelativistic, i.e. they all travel at essentially the speed of light. For a constant speed shock wave $A \sim t^2$ so that $P_C \propto t^{-2} \ln(p_{\max}(t)/p_0)$. The increase of the upper cut-off with time is, from section 2, determined by $\dot{p}_{\max} = p_{\max} \Delta/t_c$, with adiabatic expansion losses ignored. The crossing time is proportional to the diffusion coefficient which, in turn, scales with the mean free path for scattering so that $t_c \propto \lambda$. The smallest mean free path we can use, which gives the most rapid acceleration, is just the gyroradius (λ_g) of a proton in the mean magnetic field which falls off as $1/r$ away from the star. With $\lambda_g \propto B^{-1}$ and $\lambda \sim \lambda_g$

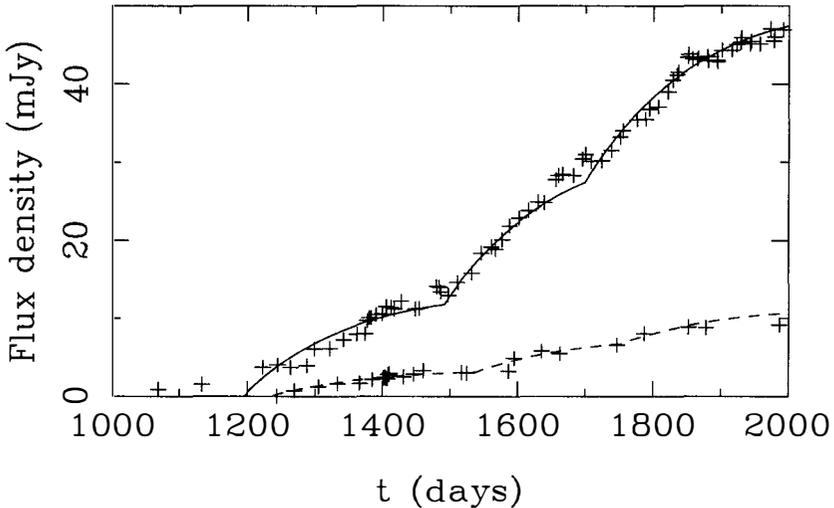


Figure 1. Observed and predicted light curves at 843MHz (solid line) and 4.8GHz (dashed line)

the crossing time at the freely expanding shock then increases linearly with time $t_c \propto t$. Therefore $\dot{p}_{\max} \propto p_{\max}/t$ so that $P_C \propto t^{-2} \ln(t/t_0)$. This is to be compared with a $1/r^2$ (or $1/t^2$) decrease for the post-shock thermal pressure. Therefore, initially ($t < 1.7t_0$) P_C increases rapidly on a timescale much shorter than the expansion timescale of the shock (r_s/u_s). It then falls off but only when $t \gg t_0$ does P_C decline as rapidly as P_G . Depending on the absolute values of the parameters (Q , κ etc.) it is possible that the non-thermal component of the pressure could dominate over the thermal part. Our test particle estimates would then no longer apply and the evolution of the coupled thermal and CR system must be solved numerically.

The details of our model can be found in Ball and Kirk [3] Kirk et al. [6] and Duffy et al. [7]. Here we restrict ourselves to some general remarks. We first treat the hydrodynamic part of the problem as a two fluid system where accelerated protons provide an important part of the pressure but none of the inertia. A time dependent solution then gives us a picture of the fluid profile as it evolves over the first few years after the explosion. With the hydrodynamics solved we then turn to the electron acceleration described in section 2. Now the quantities Δ , P_{esc} and t_c are all time dependent as the subshock compression and flow profile change. As before the accelerated electrons at

the shock ultimately escape downstream where they lose some energy as a result of the adiabatic expansion of the flow. The electron spectrum calculated at and behind the shock, and the magnetic field profile, is then combined to calculate the synchrotron emission which can be compared with observations.

4. Discussion

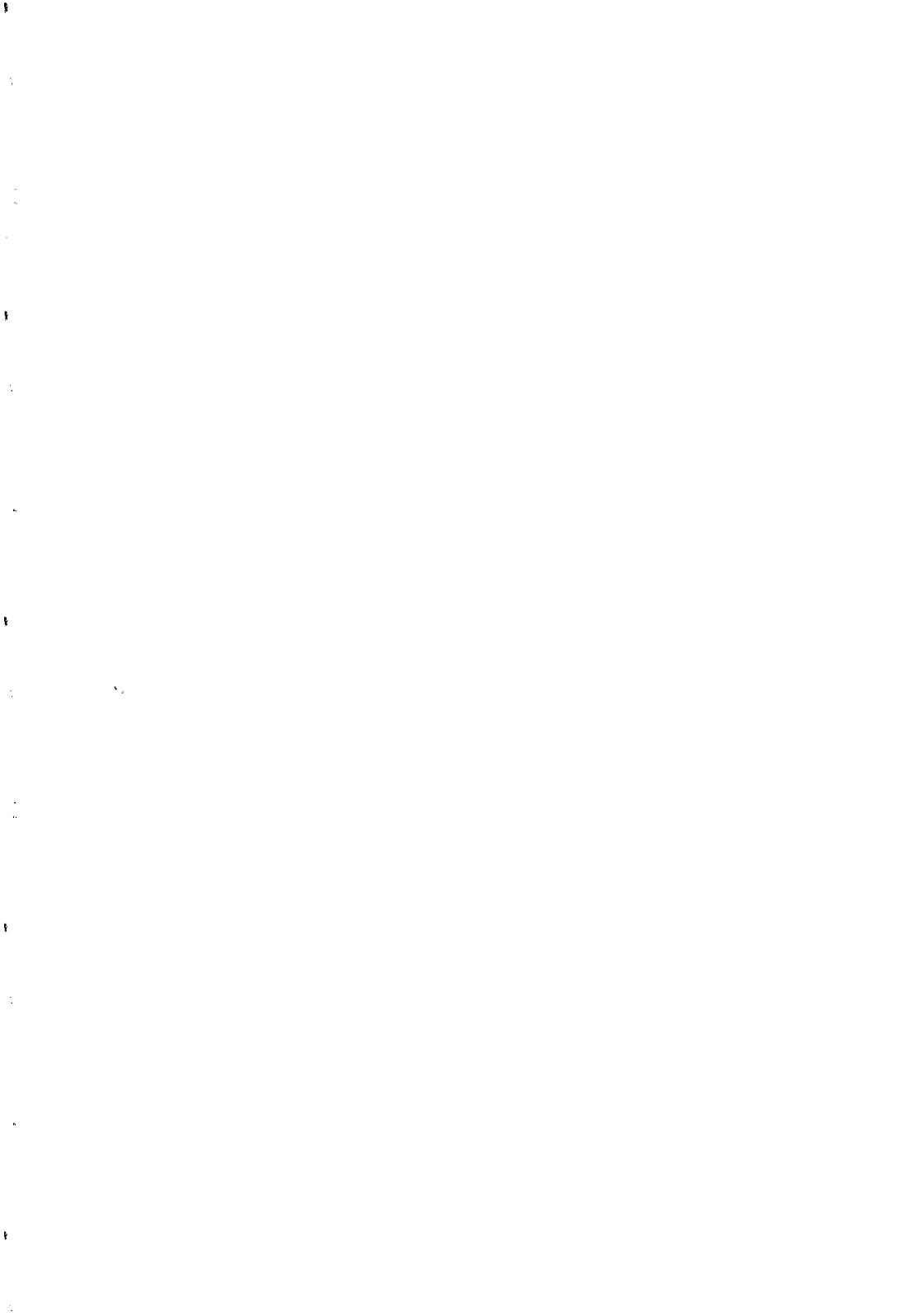
Using parameter values that are in agreement with other studies of CR acceleration [8] we can produce a subshock compression ratio that is close to the value of 2.7 after a few years. Electrons which are accelerated at this subshock will have a spectrum that is steeper than p^{-2} and will therefore emit synchrotron radiation that is steeper than $\nu^{-0.5}$. In figure 1 we make a comparison between the observations at two frequencies and our predicted light curves. As suggested by the observation of clumpiness in the supernova [9] we inject electrons at three stages (after 2.67, 3.3 and 3.6 years) at the respective rates of $1.6 \cdot 10^{44} \text{s}^{-1}$, $6.8 \cdot 10^{44} \text{s}^{-1}$ and $1 \cdot 10^{45} \text{s}^{-1}$. While we need to assume precise levels of electron injection to obtain the correct *absolute* radio flux, the *relative* fluxes emitted at the two frequencies is fully determined by the history of the subshock. Many other questions regarding acceleration at the shock of 87A, such as the complex circumstellar environment and the possibility of X- and γ -ray emission when the shock hits the ring, have not been addressed here but are subjects of considerable interest.

Acknowledgements

We are grateful to D. Campbell-Wilson and L. Staveley-Smith for permitting us access to data before publication.

References

- [1] McCray, R. 1994, *Ann. Rev. Astron. Astrophys.*, 31, 175
- [2] Staveley-Smith, L. et al. 1992, *Nature*, 355, 147
- [3] Ball, L. & Kirk, J.G. 1992, *Ap. J. Lett.*, 396, L39
- [4] Bell, A.R. 1978, *MNRAS*, 182, 147
- [5] Longair, M.S. 1992, *High Energy Astrophysics Vol. 1*, CUP
- [6] Kirk, J.G., Duffy, P. & Ball, L. 1994, *Ap. J. Suppl.*, 90, 807
- [7] Duffy, P., Kirk, J.G. & Ball, L. 1993, *Proc. of the 23rd ICRC*, 2, 358
- [8] Dorfi, E. 1990, *A&A*, 234, 419
- [9] Staveley-Smith, L. et al. 1993, *Nature*, 366, 136



BRAKING INDEX OF GEMINGA PULSAR

Ramanamurthy, P. V. and Mattox, J. R.
 Laboratory for High Energy Astrophysics
 NASA/ Goddard Space Flight Centre
 Greenbelt, Md. 20771, U.S.A.
 on behalf of the *EGRET* Science Team



ABSTRACT

The pulsar Geminga, also known as 2CG195+04, IE0630+178 and the faint star G'' , is a remarkable object in the sense that its energy output is almost all in high energy gamma rays. The braking index,

$$n = f \times \ddot{f} / (\dot{f})^2$$

as obtained from the elements given by *any one* group of authors (Hermsen *et al.*; Bertsch *et al.*) based on their own data set appears to be too high or has a very large upper limit compared with 3, the value expected for magnetic dipole radiation.

Rather than fitting a polynomial in elapsed time to the event phases over different data sets, we have taken a different approach to determine \ddot{f} . Hermsen *et al.* (1992) and Mattox *et al.* (1994) have determined f and \dot{f} from *COS-B* and *EGRET* data respectively at two widely separated epochs. Assuming that there were no glitches, we obtained \ddot{f} by dividing the difference in f values at the two epochs by the time difference between the two epochs; the resulting value of \ddot{f} is $(3.85 \pm 1.19) \times 10^{-26} \text{ s}^{-3}$. Combining this with the f and \dot{f} values we obtained a value of (4.3 ± 1.3) for the braking index. This value agrees well with the expected.

1. Introduction

High Energy Gamma Rays (HEGR) with energies ≥ 50 MeV from celestial sources were observed by a series of satellite-borne detectors — *SAS - 2*¹⁾, *COS - B*²⁾ and *EGRET*³⁾ — with progressively increasing sensitivity. Based on the results [Fichtel *et al.*⁴⁾] from the latest and most sensitive detector (*EGRET*), one can characterise the HEGR sky as consisting of a strong diffuse radiation by the Galactic center and disk, emission from 5 pulsars, atleast 21 AGNs and 16 unidentified sources. If marginal detections with $\geq 4\sigma$ are included, these numbers increase further. Geminga is one of the 5 HEGR pulsars detected so far, the others being the Crab and the Vela pulsars and PSR1055-52 and PSR1706-44.

Geminga was discovered by Fichtel *et al.*⁵⁾ with the *SAS - 2* detector. It is now known that the objects variously referred to in the literature as Geminga, 2CG195+04, the X-ray object 1E0630+178 and, most probably, the faint blue star *G''* all refer to the same object with the *J2000* co-ordinates of *R.A* = 6h 33m 54.02s and $\delta = +17^\circ 46' 11''.5$ and located at a distance of ~ 100 to 1000 pc from Solar system. During the seventies and eighties there were several erroneous claims for a periodicity ~ 59 s in X-ray, HEGR and TeV gamma ray emissions. Recently Halpern and Holt⁶⁾, using a deep *ROSAT* X-ray observation have convincingly demonstrated that the true period of Geminga pulsar is ~ 237 msec. Soon Bertsch *et al.*⁷⁾ found the same periodicity in HEGR from the *EGRET* data and further concluded that Geminga is a rotating magnetised ($B \sim 1.6 \times 10^{12}$ G) neutron star, with a slow-down energy loss rate of $\sim 3.5 \times 10^{34}$ erg.s⁻¹ and a characteristic age of $\sim 3.2 \times 10^5$ yr. Searches of the *COS - B* and *SAS - 2* archival data have also confirmed⁸⁻¹⁰⁾ the 237 msec periodicity. Bignami *et al.*¹¹⁾ measured the proper motion of *G''* to be $\sim 0''.17$ yr⁻¹. A search for radio pulsations, however, did not find¹²⁾ any 237 msec pulsations at ≥ 1 mJy. In the TeV energy range, there were two claims^{13,14)} of detection and one¹⁵⁾ of non-detection of pulsations at 237 msec periodicity.

It is clear that Geminga is a unique object shining brightest in HEGR, with luminosity ratios of $L_{HEGR}/L_{X-rays} \sim 10^3$, $L_{HEGR}/L_{opt} \sim 2 \times 10^6$ and $L_{HEGR}/L_{TeV\gamma} \geq 150$. Geminga can truly be called a high energy gamma ray pulsar.

2. Braking Index

A rotating non-aligned magnetic neutron star loses energy by emission of electromagnetic radiation and it is expected theoretically

$$\dot{\Omega} = -k \times \Omega^n \quad (1)$$

where Ω is the angular velocity of rotation of the neutron star, $\dot{\Omega}$ its first time derivative and n , the braking index. From (1) it follows

$$n = \Omega \ddot{\Omega} / \dot{\Omega}^2 = f \ddot{f} / \dot{f}^2 \quad (2)$$

Here f is the rotation frequency (cycles/sec) of the pulsar. For an ideal rotating magnetic dipole, the theoretical value for the braking index is 3. In reality the braking index could differ from this value; for example, the Crab pulsar, PSR0540-69 and PSR1509-58 exhibit braking indices of 2.50 ± 0.01 , 2.04 ± 0.02 and 2.80 ± 0.01 respectively. Glitches, timing noise and moving crust of the neutron star might be responsible for the discrepancy. To estimate the braking index to within an error of, say, 1%, one has to obtain f , \dot{f} and \ddot{f} to accuracies better than 1%.

3. Determination of f , \dot{f} and \ddot{f}

First one reduces the observed event times of gamma rays to solar system barycenter to correct for the earth's orbital motion and rotation. One, then, expresses the number of rotations of the pulsar upto i^{th} event time by

$$m_{rot} = f \times T_i + \frac{1}{2} \times \dot{f} \times T_i^2 + \frac{1}{6} \times \ddot{f} \times T_i^3 \quad (3)$$

where $T_i = t_i - t_0$. Here t_0 is an arbitrary epoch at which the phase of the pulsar equals 0. The difference between m_{rot} and its nearest integer is called residual. One minimizes the residuals to obtain the best fit for f , \dot{f} and \ddot{f} . While f can be obtained reasonably accurately in a relatively short data span, the other two terms \dot{f} and \ddot{f} require rather long duration data spans. One requires data time streams of durations of $\sqrt{2 \times \delta m / \delta \dot{f}}$ and $\sqrt[3]{6 \times \delta m / \delta \ddot{f}}$ to obtain \dot{f} and \ddot{f} to within errors of $\delta \dot{f}$ and $\delta \ddot{f}$ respectively. As an example, for $\delta m = 0.03$, $\delta \dot{f} / \dot{f} \approx 0.01$ and $\delta \ddot{f} / \ddot{f} \approx 0.01$, one needs, in the case of Geminga, data streams of 63 d and 27 yr respectively. The data streams of *COS - B* and of *EGRET* (upto now) are only of approximately 7 and 2.8 years long. This is the major reason why Hermsen *et al.*⁸ could determine \ddot{f} only to an accuracy of 60%.

As examples we reproduce below two of the published ephemerides of the Geminga pulsar: the first set by Hermsen *et al.*⁸

$$\begin{aligned} t_0 &= JD\,2443946.5 \\ f &= 4.21775012277(24) s^{-1} \\ \dot{f} &= -0.1952379(24) \times 10^{-12} s^{-2} \\ \ddot{f} &= (0.28 \pm 0.16) \times 10^{-24} s^{-3} \end{aligned} \quad (4)$$

and the second set by Mattox *et al.*¹⁶

$$\begin{aligned}
 t_0 &= JD\ 2448750.5 & (5) \\
 f &= 4.21766909413(5)\ s^{-1} \\
 \dot{f} &= -0.195221(8) \times 10^{-12}\ s^{-2} \\
 \ddot{f} &\leq 3 \times 10^{-24}\ s^{-3}
 \end{aligned}$$

Using these ephemerides one obtains, using formula (2), values of (31.1 ± 17.8) and < 332 respectively for the braking index, n , of the Geminga pulsar. It may be pointed out that the rather large errors in n are due to the correspondingly large errors in \ddot{f} . Though the derived values of n , with their large errors, are not seriously in conflict with the expected value 3, there have been a few suggestions in the literature to explain these 'high values' of braking index in terms of glitches¹⁷ and of pulsar proper motion¹⁸.

4. Our method to determine \ddot{f}

The errors in \dot{f} in the ephemerides (4) and (5) preclude joining the *COS - B* and *EGRET* data sets into a *single* data set of 18yr long lever-arm allowing one to determine \ddot{f} more accurately. Instead, we adapted a simple method to obtain \ddot{f} . Since Geminga did not show any evidence for any glitch in its observed history, we have divided the difference in the \dot{f} values by the difference in the epochs given by the two ephemerides (4) and (5); see Figure 1. The resulting \ddot{f} value is given by $\ddot{f} = (4.1 \pm 2.0) \times 10^{-26}\ s^{-3}$, leading to the value for braking index $n = (4.5 \pm 2.3)$. There has been one more observation on Geminga by *EGRET* in December, 1993. If we include these data and reanalyse the *EGRET* data, we obtain a more accurate value $\dot{f} = -0.1952219(43) \times 10^{-12}\ s^{-2}$. When we combine this value with the ephemerides (4), we obtain $\ddot{f} = (3.85 \pm 1.19) \times 10^{-26}\ s^{-3}$. This leads to a value $n = (4.3 \pm 1.3)$.

5. Conclusions

We conclude that, while neither the *COS - B* data nor the *EGRET* data stream is long enough to determine \ddot{f} accurately, one can obtain a reasonably accurate value for \ddot{f} from the \dot{f} values given by the two groups. The newly determined value of \ddot{f} leads to a value for the braking index of Geminga pulsar of $n = (4.3 \pm 1.3)$. This value is in agreement, within errors, with the theoretical value, 3, for a rotating non-aligned magnetic dipole, obviating any need to look for explanations for a 'large' braking index.

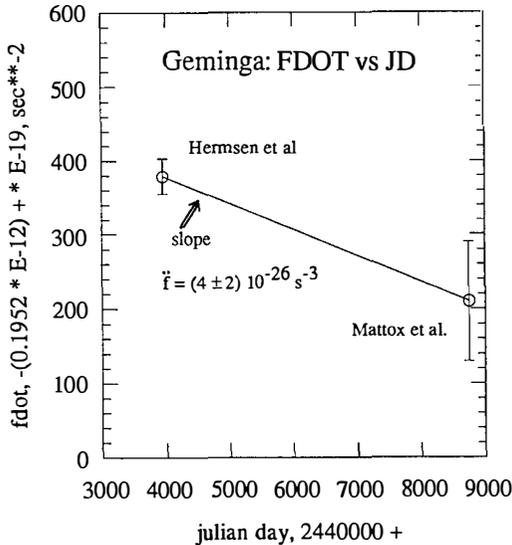


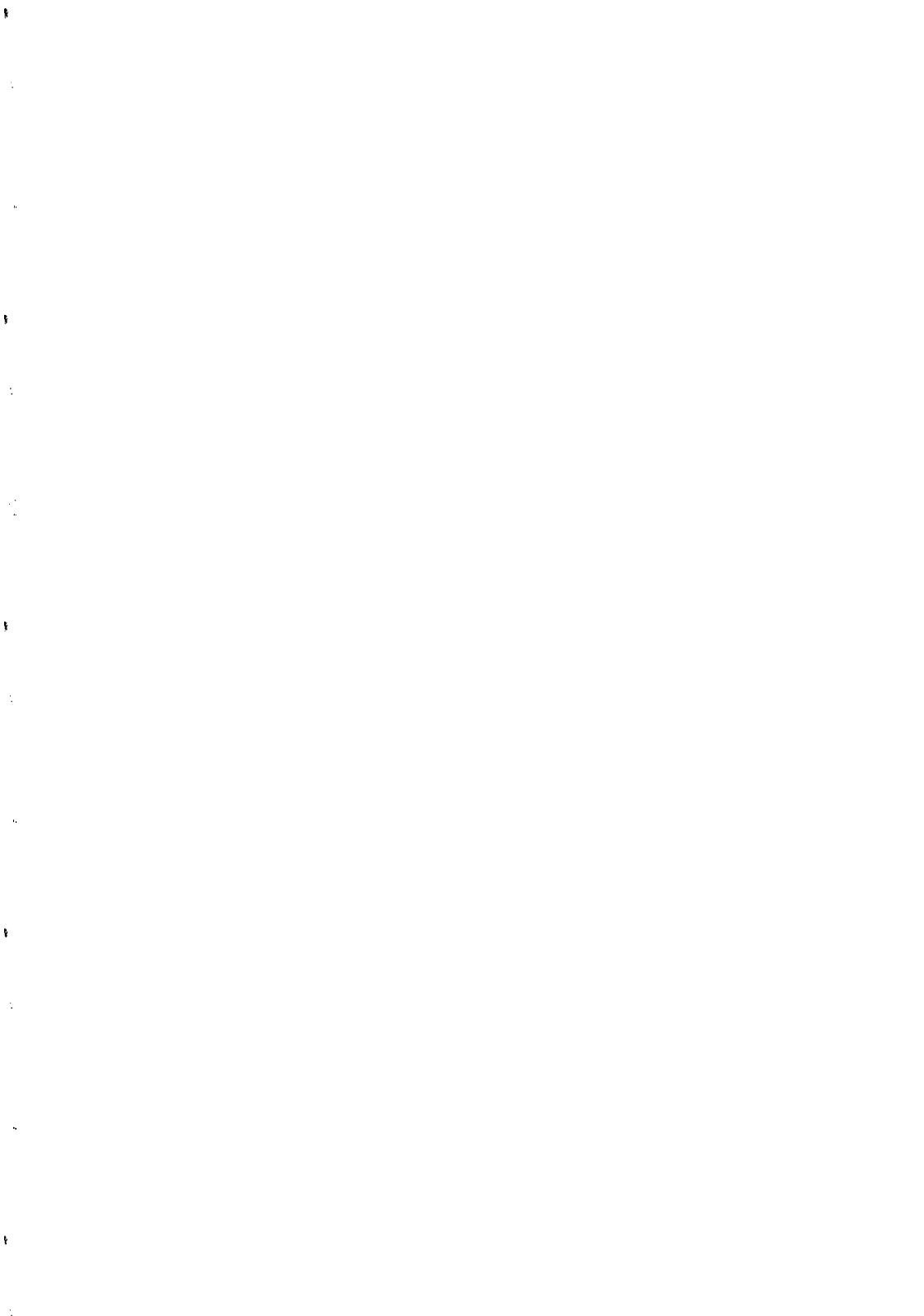
Fig.1: Time derivative of Geminga pulsar frequency vs time in Julian Days. The data points are from the ephemerides given by Hermesen *et al.* and Mattox *et al.*. The slope of the line joining the two points is a measure of the second time derivative of frequency of the pulsar.

Acknowledgements

One of the authors, P. V. Ramanamurthy, expresses his thanks to Dr. Carl Fichtel for his hospitality at Goddard Space Flight Center.

References

- (1) Kniffen, D.A. *et al.*, *Nature* **251**, 397 (1974)
- (2) Bennett, K. *et al.*, *Astron. Astrophys.* **61**, 279 (1977)
- (3) Kanbach, G. *et al.*, *Space Science Reviews* **46**, 69 (1988);
Thompson, D.J. *et al.*, *Ap.J.Suppl.* **86**, 629 (1993)
- (4) Fichtel, C.E. *et al.*, submitted to *Ap.J. Suppl.* (1993)
- (5) Fichtel, C.E. *et al.*, *Ap.J.* **198**, 163 (1975)
- (6) Halpern, J.P. and Holt, S.S., *Nature* **357**, 222 (1992)
- (7) Bertsch, D.L. *et al.*, *Nature* **357**, 306 (1992)
- (8) Hermesen, W. *et al.*, *IAU circular* 5541 (1992)
- (9) Bignami, G.F. and Caraveo, P.A., *Nature* **357**, 287 (1992)
- (10) Mattox, J.R. *et al.*, *Ap.J.* **401**, L23 (1992)
- (11) Bignami, G.F. *et al.*, *Nature* **361**, 704 (1993)
- (12) Seiradakis, J.H., *IAU circular* 5532 (1992)
- (13) Vishwanath, P.R. *et al.*, *Astron. Astrophys.* **267**, L5 (1993)
- (14) Bowden, C.C.G. *et al.*, *J.Phys.G* **19**, L29 (1993)
- (15) Fegan, D.J. *et al.*, *Proc. 1st Compton Symp.* (St. Louis, U.S.A.), p223 (1993)
- (16) Mattox, J.R. *et al.*, *Proc. 2nd Compton Symp.* (College Park), in press (1994)
- (17) Alpar, M.A. *et al.*, *Astron. Astrophys.* **273**, L35 (1993)
- (18) Bisnovatyi-Kogan, G.S., *Proc 2nd Compton Symp.* (College Park), in press (1994)



PARTICLE ACCELERATION AND RADIATION IN THE JETS OF ACTIVE GALACTIC NUCLEI

CHARLES D. DERMER

*E. O. Hulburt Center for Space Research
Code 7653, Naval Research Laboratory
Washington, DC 20375-5352 USA*

Abstract

The discovery of a class of active galactic nuclei that emit a large fraction of their energy in the medium energy (~ 30 MeV - 30 GeV) gamma-ray regime has important implications on particle acceleration and radiation processes near supermassive black holes, and the intensity of lower energy radiation fields in the sources. The essential facts connected with these high energy sources are reviewed. The necessity of beaming is argued on the basis of source energetics and variability data, assuming Eddington-limited accretion luminosities. Constraints on bulk outflow velocities are derived, which depend critically on whether the radiating agents are assumed to be leptons or hadrons.

1. Gamma-Ray Observations of Blazars

The Energetic Gamma Ray Experiment Telescope (EGRET) on the *Compton Gamma Ray Observatory* (CGRO) has strongly ($> 5\sigma$) detected 25 active galactic nuclei (AGNs) and marginally ($4 - 5\sigma$) detected an additional 13 AGNs during its Phase I all sky survey.¹⁾ All of these sources are identified with radio-loud extragalactic sources, yet none of the famous radio galaxies, such as Centaurus A²⁾ or Cygnus A,³⁾ are detected at the threshold sensitivity levels of EGRET. The EGRET threshold flux, which depends on direction due to obscuration by galactic diffuse emission, is in the range $0.5 - 3 \times 10^{-7}$ photons (> 100) MeV s⁻¹ cm⁻². The extragalactic sources discovered with EGRET are the flat radio-spectrum, compact radio core sources, variously categorized as radio quasars (in particular, highly polarized quasars), BL Lac objects, and optically violently variable quasars.^{4,5)} These objects are commonly called “blazars,” which are thought to be radio galaxies favorably oriented so that the axes of the radio jets are nearly aligned in our observing direction. None of the closer and more numerous radio-quiet Seyfert AGNs have been detected with EGRET,⁶⁾ implying that acceleration processes in the jets of the radio-loud objects are responsible for energizing particles to high energies.

The jet sources detected with EGRET have also been monitored at lower gamma-ray energies with the Oriented Scintillation Spectrometer Experiment (OSSE) and the Compton Telescope (Comptel) on CGRO. OSSE detects photons with energies between ~ 50 keV and 10 MeV, and Comptel is sensitive between ~ 0.7 MeV and 30 MeV. OSSE is a relatively small field-of-view ($3.8^\circ \times 11.4^\circ$) instrument and therefore requires specific source pointing. Of the 5 blazar targets observed with OSSE during Phase I, three (3C 279, 3C 273, PKS 2155-304) were detected and upper limits were reported for the other two (Mrk 421, PKS 0528+134).^{7,8)} The Comptel team reported detections of 4 AGNs during its Phase I all sky survey, three of which are blazars (3C 279, 3C 273, PKS 0528+134) and the fourth of which is Centaurus A.⁹⁾

One of the most exciting recent discoveries in gamma-ray astronomy is the detection of Mrk 421 at TeV energies with the ground-based Whipple telescope.¹⁰⁾ Mrk 421 is a lineless BL Lac object at the center of an elliptical galaxy, and is the closest and one of the weakest EGRET blazar sources. Several other EGRET blazars (e.g., 3C 273, 3C 279, PKS 0528+134, 3C 454.3, 0716+714, 4C 38.41) have also been observed with the Whipple observatory,¹¹⁾ but so far with negative results.

By constructing multiwavelength νF_ν diagrams, it is becoming certain that the bulk of the power output in many of the EGRET blazar sources (e.g., 3C 279, PKS 0528+134, 3C 454.3, PKS 0208-512, PKS 0235+164, PKS 0208-

512, CTA 102) is radiated in the medium energy gamma-ray regime. Although generally based on noncontemporaneous data, the fact that so many sources show similar rounded featureless continua peaking between $\sim 10^{12} - 10^{14}$ Hz and another peak between ~ 30 MeV and 1 GeV some 1-2 orders of magnitude higher in the power spectrum strongly suggests that the energetics of the blazar sources are dominated by emission in the gamma ray regime. Moreover, the high-energy radiation displays variability on time scales as short as \approx four days for 3C 454.3,¹²⁾ and \approx two days for 3C 279¹³⁾, 1633+382,¹⁴⁾ and PKS 0528+134.¹⁵⁾ The spectra between 100 MeV and 1 GeV are well-described by power laws, although more complicated spectral shapes are difficult to rule out because of the paucity of photons. Photon spectral indices α range from $\alpha = 1.5 \pm 0.2$ for NRAO 629 to $\alpha = 2.6 \pm 0.1$ for PKS 0528+134.¹⁾ Direct spectral evidence,⁸⁾ combined with the statistics of the relative number of source detections with Comptel and EGRET, indicate that the spectra of blazars soften between the soft gamma-ray and medium energy gamma-ray regimes. Source models are forced to contend with these observational constraints.

2. Evidence for Beaming

The 100 MeV-5 GeV luminosity in the source rest frame has been calculated for blazars with known redshift z .¹⁾ For a Hubble constant $H_0 = 75$ km s^{-1} Mpc $^{-1}$, the gamma-ray blazar luminosities range from $L = 2 \times 10^{44} f$ ergs s^{-1} for Mrk 421 with $z = 0.031$, to $L = 8 \times 10^{48} f$ ergs s^{-1} for PKS 0528+134 with $z = 2.06$. Here f denotes an unknown beaming factor. If we assume Eddington-limited accretion, then the implied black-hole masses necessary to power these objects range from $M_8 \approx 0.01f$ to $M_8 \approx 650f$, where M_8 is the black-hole mass in units of $10^8 M_\odot$.

Consider now the gamma-ray time-variability data mentioned above. From light-travel time arguments, variations of luminosity by a factor of 2 on the time scale δt_{obs} imply that the characteristic size R of the emitting region is given by $R \lesssim c\delta t_{obs}/(1+z)$. This argument holds for stationary unbeamed sources. We argue that beaming or bulk relativistic motion is required because the assumption of Eddington-limited luminosities and size scales implied by stationary emitters are incompatible.

For simplicity, we use the Schwarzschild metric. The luminosity radiated interior to radius R for a zero-stress boundary condition is given by $L_{Edd}(1 - R_I/R)$, where $L_{Edd} \cong 1.26 \times 10^{46} M_8$ ergs s^{-1} , the gravitational radius $R_g \equiv GM/c^2 \cong 1.5 \times 10^{13} M_8$ cm, and $R_I = 6R_g$ is the radius of the innermost stable orbit. This luminosity must be greater than the observed luminosity modulo the beaming factor f . This sets a lower limit on the black hole mass. An upper limit is set by the variability argument given above. We find that

$$\frac{80fL_{48}}{1-6R_g/R} \lesssim M_8 \lesssim \frac{170 \delta t_{\text{obs}}[\text{days}]}{(1+z)(R/R_g)}. \quad (1)$$

In Figure 1 we plot equation (1) as a function of $\rho = R/R_g$. Depending on the observed relative luminosity ($= L_{48}$) given by Fichtel *et al.*¹⁾ and the assumed beaming factor f , we obtain a lower limit for M_8 given by the solid curves. The observed variability time scale gives an upper limit for M_8 . The gamma-ray data shows that the radiation from some blazar sources must be beamed, because there is no solution to equation (1) unless $f \ll 1$. We specifically consider the sources 3C 279, 1633+382, and PKS 0528+134 which have, respectively, $z = 0.538$, 1.81, and 2.06, and measured values of $f = 1.2$, 6.2, and 8.0. Each varies on time scales $\delta t_{\text{obs}} \lesssim 2$ days (see above).

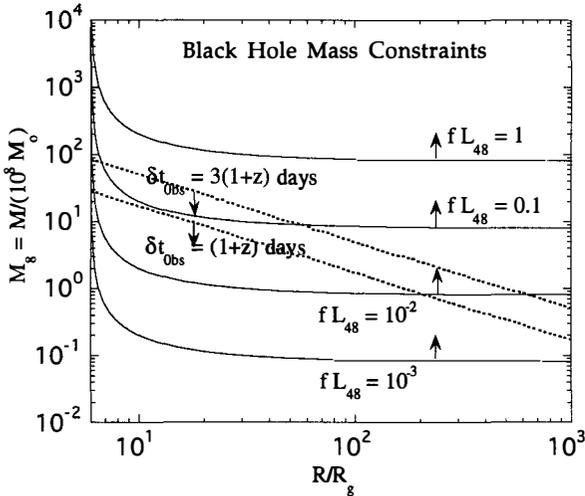


Fig. 1—Constraints on black hole masses M and beaming factors f from gamma-ray observations of variability time scale and luminosity, assuming radiation at the Eddington limit and a Schwarzschild metric. Use of a Kerr metric will relax the constraints.

In order to obtain solutions with the minimum amount of beaming, we find that $M_8 \cong 19$, 10, and 9 for 3C 279, 1633+382, and PKS 0528+134, respectively. The respective beaming factors are $f \cong 9.5\%$, 1% and 0.7%, and the emission size scales correspond to $\approx 12R_g$ in all three cases. Smaller mass black holes would be implied by smaller beaming factors. If the beaming is due to the Doppler boosting of the radiation by bulk relativistic motion, then $\mu \approx \beta_\Gamma \approx 1 - 2f$, where $\cos^{-1} \mu$ is the observing angle with respect to the direction of the bulk motion, $\beta_\Gamma = (1 - \Gamma^{-2})^{1/2}$, and Γ is the bulk Lorentz factor of the outflow. This implies that $\Gamma \gtrsim 1.7$, 5, and 6 for 3C 279, 1633+382, and PKS 0528+134, respectively. We propose the arguments leading to equation (1) as new evidence for relativistic motion in blazar sources. Intrinsic anisotropy of the radiation field must be required to jointly satisfy the two relations in

equation (1), because geometrical obscuration of some fraction of the source does not reduce the luminosity requirements. A use of the Kerr metric will, however, relax the constraints, but then it is also important to consider the gravitational redshifting and time dilation of the emergent radiation.

3. Constraints on Radiation Processes in Relativistic Jets

Beaming of the radiation reduces luminosity requirements by a factor $\sim f^{-1}$, thus considerably easing source energetics. Relativistic bulk outflow in a jet produces beamed emission through Doppler boosting and, moreover, avoids the catastrophic Compton cooling of electrons implied by the inferred radio luminosities and variability size scales for systems assumed to be at rest with respect to the observer.¹⁶⁾ Bulk relativistic motion also provides a natural explanation for apparent superluminal motion in the transverse motions of radio emitting blobs in compact radio sources. Lower limits to the bulk Lorentz factors Γ are derived from the measured apparent transverse velocities: most inferred values of Γ are between ≈ 3 and 10 .¹⁷⁾

We consider three models for EGRET blazar gamma-ray emission, and derive constraints on parameters in order that the models are in accord with gamma-ray variability data. In the model of Dermer & Schlickeiser¹⁸⁾ (hereafter DS), nonthermal relativistic electrons, which are assumed to be quasi-isotropically distributed in the comoving frame of a relativistically outflowing plasma blob, Compton upscatter accretion disk radiation to gamma ray energies. In the model of Sikora *et al.*¹⁹⁾ (hereafter SBR), the electrons in the plasma blob, instead of scattering disk radiation, scatter central source emission that has been rescattered by surrounding diffuse gas. In the model of Mannheim & Biermann²⁰⁾ (hereafter MB), high-energy protons initiate a cascade by photomeson production with photons produced through synchrotron emission by nonthermal electrons in the outflowing plasma jet.

For the first two lepton-dominated models, we require that the time scale for electron energy loss, as measured in the observer's frame, be less than $\delta t_{obs} \approx 1$ day. Because $-\dot{\gamma} = 4c\sigma_T u'_{ph} \gamma^2/3$ for Thomson scattering, where $U'_{ph} = m_e c^2 u'_{ph}$ is the comoving soft photon radiation-field energy density, and the observed time interval $t_{obs} = \Gamma t'(1 - \beta_\Gamma \mu)$, where the electron energy loss time scale $t' = |-\dot{\gamma}/\gamma|^{-1}$, we have

$$t_{obs} = \frac{3\Gamma(1 - \beta_\Gamma \mu)}{4c\sigma_T u'_{ph} \Gamma^2(1 + \beta_\Gamma^2/3)\gamma} \lesssim \delta t_{obs}. \quad (2)$$

In equation (2), we replace u'_{ph} with $u_{ph}\Gamma^2(1 + \beta_\Gamma^2/3)$, where u_{ph} is the quasi-isotropic soft photon field in the stationary accretion disk frame.²¹⁾

In model DS, $u_{ph} \approx L_{ad}R_g/4\pi z^3 m_e c^3$, where $L_{ad} = \zeta L_{Edd}$ is the accretion disk luminosity scaled by the Eddington luminosity, and $z \equiv \bar{z}R_g$ is the

height of the scattering blob above the central black hole. In the SBR model, $u_{ph} \approx L_{ad}\tau_{sc}/4\pi R_{sc}^2 m_e c^3$, where $\tau_{sc} \equiv 0.01\tau_{-2}$ is the Thomson scattering depth of the clouds and $R_{sc} \equiv 0.1R_{-1}$ pc is the radial extent of the scattering cloud. For simplicity in what follows, we assume that the observer is oriented in the superluminal direction $\mu = \beta_\Gamma$. We obtain the following constraint on Γ for model DS:

$$\Gamma(1 + \beta_\Gamma^2/3)^{1/3} \gtrsim \frac{1.7\tilde{z}_3 M_8^{1/3}}{(\zeta\gamma_3\delta t[\text{day}])^{1/3}}, \quad (3)$$

where $\tilde{z}_3 = \tilde{z}/1000$ and $\gamma_3 = \gamma/1000$. We choose a characteristic value of $\gamma \sim 10^3$ because electrons with Lorentz factors of this order will Compton scatter “blue” bump photons ($\sim 0.1 - 1$ keV) to energies near 100 MeV. For the SBR model, we obtain

$$\Gamma(1 + \beta_\Gamma^2/3)^{1/3} \gtrsim \frac{3.7R_{-1}^{2/3}}{(\zeta M_8 \tau_{-2} \gamma_3 \delta t[\text{day}])^{1/3}}. \quad (4)$$

We constrain the MB model on the basis of the observed radio-through-optical continua of blazars, which is the soft photon field with which protons produce pions through photomeson production. Using an approximation²²⁾ for the energy-loss rate of protons through photomeson production, we obtain the energy-loss time scale $t_{p\gamma}(\text{sec}) = 2.5 \times 10^{15}/U'_\gamma[\text{ergs s}^{-1}]\gamma_p$, where the radiation field in the comoving frame is assumed to be described by a power law with photon spectral index = 2, and γ_p is the proton Lorentz factor. Relating the energy-loss time scale to the observer’s time interval, we find that $2.5 \times 10^{15}/U'_\gamma[\text{ergs s}^{-1}]\gamma_p\Gamma \lesssim \delta t_{obs}$.

The observed frequency-integrated energy flux Φ_E , can be shown to be related to the soft photon energy density through the expression $\Phi_E = D^4 c R'^2 U'_{ph}/3d_L^2$, where $D \equiv [\Gamma(1 - \beta_\Gamma\mu)]^{-1}$ is the Doppler factor, R' is the radius of the blob (assumed spherical) in the comoving frame, and d_L is the luminosity distance. Substituting the implied expression for U'_{ph} into the energy-loss time scale, using the definitions $d_L \equiv 10^{28}d_{28}$ cm, $\gamma_p \equiv 10^8\gamma_8$, and $\Phi_E \equiv \Phi_{-11}10^{-11}$ ergs $\text{cm}^{-2} \text{s}^{-1}$, we obtain

$$R'[\text{cm}] \lesssim 2 \times 10^{16} \left(\frac{\Phi_{-11}\delta t[\text{day}]\gamma_8}{\Gamma^3} \right)^{1/2} d_{28}, \quad (5)$$

where we let $D \approx \Gamma$.

Equations (3), (4), and (5) are constraints imposed by the gamma-ray data on models DS, SBR, and MB, respectively, which must be satisfied if these models are valid. We find that the bulk Lorentz factors required in

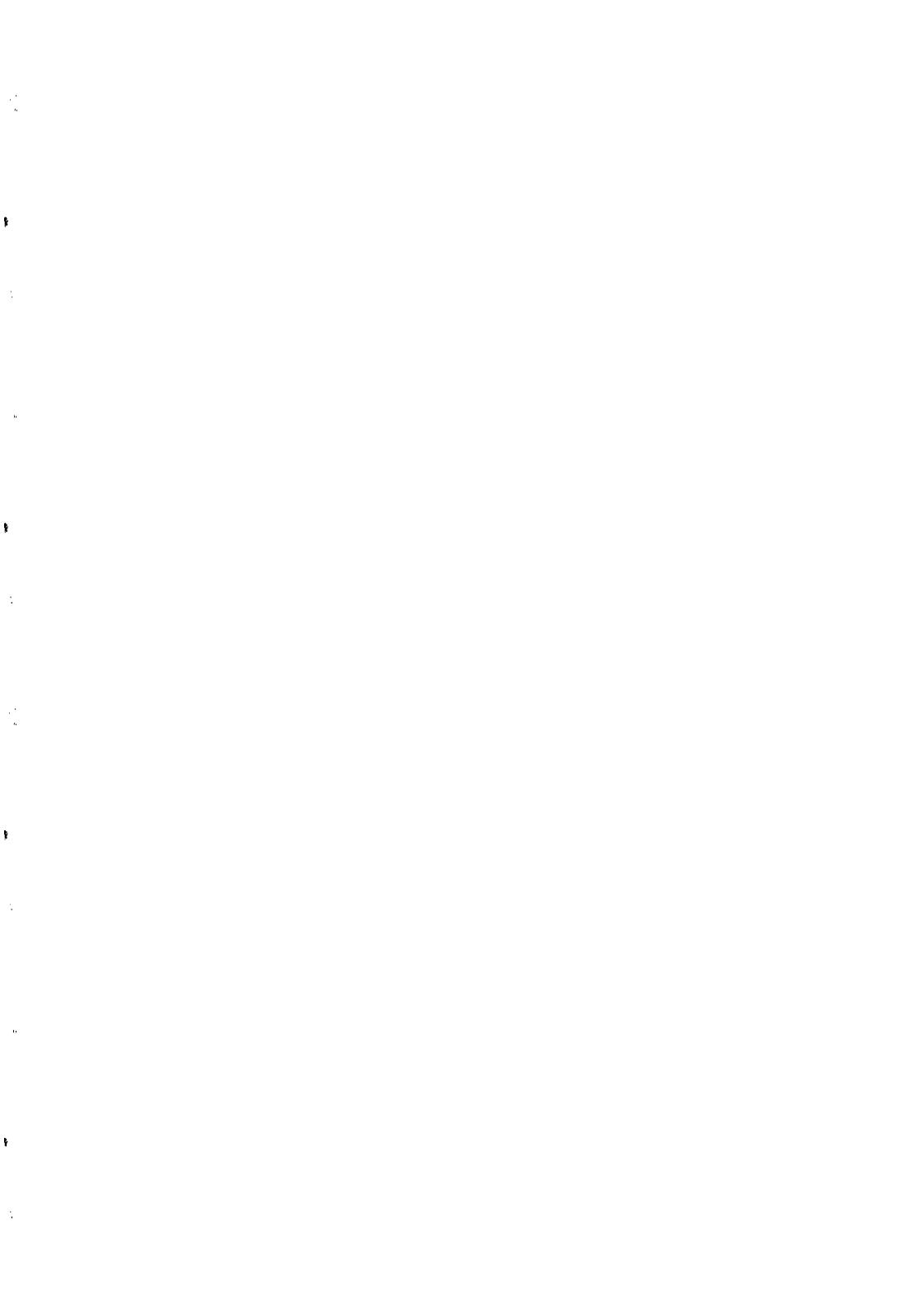
equations (3) and (4) are not at all unreasonable; only if the gamma-ray variability time scales were measured to be much shorter than one day would the implied values of Γ conflict with deductions from superluminal observations. The size of the soft photon source implied by equation (5) may be uncomfortably small if $\Gamma \sim 5 - 10$; it is additionally important to determine if medium-energy gamma rays can escape from a source this size because of $\gamma - \gamma$ pair production opacity.

In summary, we have briefly reviewed gamma-ray observations of blazars, and argued that the emission must be beamed. We have also constrained the bulk outflow velocities and emission size scales for three models of gamma-ray blazars.

Acknowledgments: I thank Professor Gérard Fontaine for his kind invitation to attend the Moriond Conference and Dr. Jeff Skibo for discussions.

REFERENCES

- 1) Fichtel, C. E., *et al.* 1994, ApJS, in press
- 2) Hartman, R. C. 1993, private communication
- 3) Mattox, J. R. *et al.* 1993, in Compton Gamma-Ray Observatory, ed. M. Friedlander, N. Gehrels & D. J. Macomb (New York: AIP), 365
- 4) Dermer, C. D., & Schlickeiser, R. 1992, Science, 257, 1642
- 5) Gehrels, N., & Dermer, C. D. 1994, in preparation
- 6) Lin, Y. C., *et al.* 1993, ApJ, 416, L53
- 7) Johnson, W. N., *et al.* 1994, in Second Compton Symposium, ed. C. E. Fichtel, N. Gehrels & J. P. Norris (New York: AIP), in press
- 8) McNaron-Brown, K., *et al.* 1994, in Second Compton Symposium, ed. C. E. Fichtel, N. Gehrels & J. P. Norris (New York: AIP), in press
- 9) Schönfelder, V. 1994, ApJS, in press
- 10) Punch, M. *et al.* 1992, Nature, 358, 477
- 11) Kerrick, A. D., *et al.* 1993, Proc. 23rd Intl. Cosmic Ray Conf., Calgary, Canada, 1, 408
- 12) Hartman, R. C., *et al.* 1993, ApJ, 407, L41
- 13) Kniffen, D. A., *et al.* 1993, ApJ, 411, 133
- 14) Mattox, J. R. *et al.* 1993, ApJ, 410, 609
- 15) Hunter, S. D., *et al.* 1993, ApJ, 409, 134
- 16) Begelman, M. C., Blandford, R. D., and Rees, M. J. 1984, RMP, 56, 255
- 17) Impey, C. D. 1987, in Superluminal Radio Sources, ed. J. A. Zensus & T. J. Pearson (New York: Cambridge University Press), p. 233
- 18) Dermer, C. D., & Schlickeiser, R. 1993, ApJ, 416, 458
- 19) Sikora, M., Begelman, M. C., & Rees, M. J. 1994, ApJ, 421, 153
- 20) Mannheim, K., & Biermann, P. L. 1992, A&A, 253, L21
- 21) Dermer, C. D., & Schlickeiser, R. 1994, ApJS, 90, 945
- 22) Mannheim, K., & Schlickeiser, R. 1994, A&A, in press



THEMISTOCLE : RESULTS AND PROSPECTS

P. Eschstruth*

Laboratoire de l'Accélérateur Linéaire

IN2P3-CNRS et Université de Paris-Sud, 91405 Orsay Cedex, France

The recently published results on the gamma-ray flux from the Crab Nebula in the 3 to 15 TeV range are presented. Refinements now being undertaken in the analysis methods, in particular relating to Monte-Carlo calculations, are described, with emphasis on the technique used to calculate the energy. Present activity with the Thémistocle detector, which includes data-taking on MKN-421 as well as certain technical studies, is mentioned.

*For the Thémistocle Collaboration: A. Djannati-Ataï, P. Espigat, C. Ghesquière and Ph. Schune¹, LPC, Collège de France, 11 Pl. M. Berthelot, F-75231 Paris; P. Baillon, CERN, CH-1211 Geneva 23; L. Behr, S. Danagoulian², B. Dudelzak, P. Eschstruth, R. Riskalla and Ph. Roy, LAL, Bât. 200, Université de Paris-Sud, F-91405 Orsay; R. George, F. Kovacs, Y. Pons, M. Rivoal, T. Socroun and A.-M. Touchard, LPNHE, Tour 33, Université de Paris 6-7, 4 Pl. Jussieu, F-75252 Paris; G. Fontaine and J. Vrana, LPNHE Ecole Polytechnique, F-91128 Palaiseau; B. Fabre and C. Meynadier, Université de Perpignan, Ave. de Villeneuve F-66025 Perpignan.

¹Now at DAPNIA, CEN Saclay, F-91191 Gif-sur-Yvette

²Permanent address: Yerevan Physics Institute, Alikhanian Br. Street 2, Yerevan 375036 Armenia

Description of the Experiment

The Thémistocle experiment, located on the THEMIS solar power plant site near Font-Romeu in the eastern French Pyrenees, is an air Cerenkov device run in a way similar to an air shower particle array; i. e. shower direction is determined by timing measurements. Eighteen telescopes, each having an 80 cm parabolic mirror and a single photomultiplier tube, are directed using the THEMIS alt-azimuth solar mirror mounts. These telescopes are distributed over the roughly elliptical 200 by 300 meter field of the solar installation. Unlike an air shower array, the Thémistocle detector is very directional, the optical aperture being limited by a diaphragm in the focal plane of the mirror to a 20 mrad half-angle. Due to the small mirror size, the device has a relatively high threshold of approximately 3 TeV.

At these energies, the shape of the swarm of Cerenkov photons falling to earth is distinctly conical, departing from a plane by about 1° . This feature makes it possible to make a cone fit to the timing data thus obtaining both the shower direction and its projected impact point on the ground, given by the vertex of the cone. The angular error, based on the Crab Nebula data, is 2.3 mrad, and Monte-Carlo calculations indicate that the impact point error for gamma-rays is better than 15 m. (These values characterize the respective 2-D Gaussian distributions).

Crab Nebula Results

Data were taken on the Crab Nebula during the 1990-1991 and 1991-1992 winter seasons. A statistically significant indication of a signal was obtained during the first period, and with the combined data sets, it was possible to provide a flux measurement and an energy spectrum^[1] as shown in Fig. 1. The 15 TeV gamma-rays at the high end of this spectrum are believed to be those of the highest energy yet observed with clear statistical significance from any astronomical source. The results obtained are on the whole compatible with the 1991 Whipple result^[2] and in good agreement with the preliminary points from HEGRA^[3]. The disagreement is considerable, however, with the reanalyzed Whipple data^[4]. (We have represented both the old and new Whipple results on this plot by integrating the authors' best-fit power law formula, given for their differential spectrum).

Analysis Improvements

Work in progress indicates that it is possible to improve the angular resolution of the experiment by correcting for systematic errors present in the cone

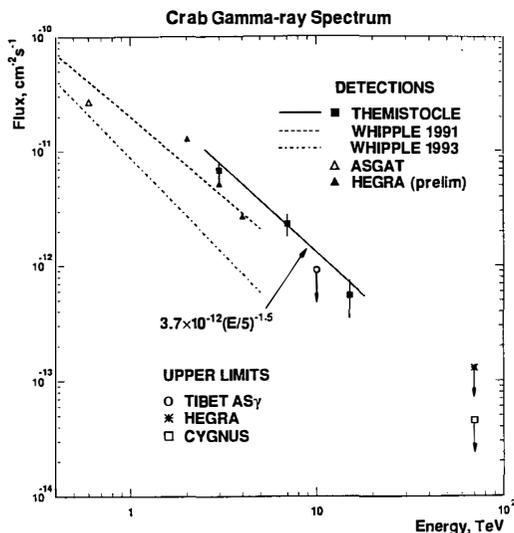


Figure 1 : Measurements and upper limits for the Crab gamma-ray flux.
Data points from refs. 1-4 and ref. 6, vol. 1.

fitting procedure. Fig. 2 shows results from the cone fit applied to Monte-Carlo generated gamma-ray events in the 3-7 TeV range. The field is divided into 13 sectors and the arrow in each one represents the average directional error for the events whose impact point fell within the sector. A preliminary application of this simple sector-dependent correction has been shown to reduce the width of the angular distribution for the Crab data, and it is expected that by further developing this technique it will be possible to reduce the average angular error from the present 2.3 mr value to one somewhat below 2 mr.

More important, however, is work being done on the hadronic Monte-Carlo simulations which should improve our estimate of the energy for photon showers. In the absence of an absolute photometric calibration method to determine the response of the detector for Cerenkov light signals, the energy scale is established indirectly by reference to the well-known flux of hadronic showers. The method can be outlined as follows. In a detailed Monte-Carlo calculation, the Cerenkov light yield is obtained for as large a cosmic-ray shower sample as possible covering various parameters such as primary energy and composition, arrival direction, and impact position with respect to the detector. A model must be supplied for the high-energy physics of the initial interactions. The secondary particles generated, and in particular the electromagnetic cascades that occur and the Cerenkov photons they produce, are followed in the atmosphere. A model of the experimental setup simulates its response to the Cerenkov light pulses and produces the equivalent of

pulse times and amplitudes. The overall trigger condition of the experiment can be applied to this Monte-Carlo output and a trigger rate calculated. This rate, as well as various distributions, can be compared with their experimental counterparts, and a compensation factor can then be introduced to take account of the unknown effects which caused simulation and experiment to disagree.

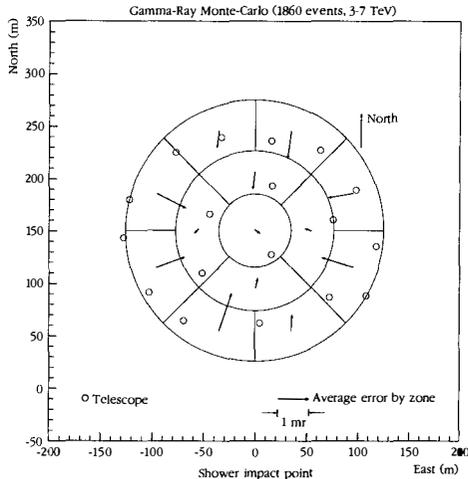


Figure 2 : Average error in direction for 13 impact-point zones.

The simulations just described for cosmic-ray showers are now repeated for gamma-ray events. The purpose is to obtain a set of curves giving average Cerenkov light intensities as a function of telescope distance from the shower center for a number of primary energies. Then the set of measured telescope light yield values for each experimental event is compared with these curves. (The events are treated here as though they were all gamma-rays). The energy of the event is taken to be that which provides the best fit. Finally, the compensation factor obtained from the hadronic comparison can be used to produce a corrected gamma-ray energy.

Until recently, our hadronic Monte-Carlo had included only proton showers. This commonly-used simplification can perhaps be justified in the high-energy limit where the Cerenkov light from n nucleons of energy E/n is equivalent to that of a proton of energy E . But at low energy, the light produced by a nucleon per TeV of energy falls off, and our Monte-Carlo studies indicate, for example, that a 10 TeV iron nucleus produces only 10% of the light of a 10 TeV proton on the shower axis. (Far from the axis this effect is less pronounced). In addition, the

chemical composition is far from being proton-dominated. When one classifies the species by total energy per nucleus, as is appropriate in our case, the tendency is a decrease in the proton component over the range 10 GeV to 10 TeV from 45% to 35%, approximately, while the alpha-particle content increases from 18% to 30%. The sum of all heavier elements thus remains stable, representing a third of the total. The heaviest (iron group) elements account for about one tenth of the total. (See refs. 5 and 6). If all these particles are simulated by protons, the number of events which trigger will be overestimated.

The original Thémistocle Monte-Carlo (based on DELSIM which was written for the CERN-DELPHI experiment), used extrapolations of the Grant model to describe the high energy processes. Recently we have made comparisons with two others developed respectively by Capdevielle^[7] and by Kertzman and Sembroski^[8] Although their approach is different, these two models give average Cerenkov light yields between 5 and 25 TeV which are identical to within 10%, while the Grant model gives 30% more light. Their more complete treatment of high-energy processes leads us to favor the newer models, and finally we have chosen that of Capdevielle.

These effects lead to an over-estimation of the trigger rate by the Monte-Carlo. In consequence, the compensation factor mentioned above was significantly different from 1. Preliminary results after incorporating these changes give us a new factor much closer to 1. In addition, the quality of the fit obtained using the cone fitting procedure has become worse on the average, bringing the Monte-Carlo χ^2 distributions closer to the experimental ones. The net result on the energy estimation for photon events is a significant reduction in the energy values. Our data points in Fig. 1 will thus be shifted to the left. Nonetheless we expect the overall agreement with the preliminary points from the HEGRA imaging telescope to remain satisfactory. The disagreement with the reanalyzed Whipple data will be reduced.

Technical Studies

As previously reported^[9], two technical studies are under way, both having the goal of improved hadronic shower rejection. The first one involves pulse-shape analysis with a fast photomultiplier on a 19th telescope. In the second study, crude imaging using a 64-pixel multi-anode tube on a 20th telescope is being explored. The information obtained from these new devices is correlated with that from the 18-telescope array operated in the usual manner.

Data-taking on MKN-421

The Whipple group has reported a 6 s.d. signal in the TeV range^[10] from the AGN MKN-421, previously shown by EGRET^[11] to be a high-energy gamma-ray emitter. Part of the Thémistocle collaboration has begun data-taking on this object in order to confirm the existence of these emissions and, if possible, measure their spectrum in energy. The signal seen by Whipple is relatively weak (1/3 of the Crab strength) and this source may show some variability. However, the spectrum seems to be flatter than that of the Crab, showing a spectral index of 2.1 from GeV to TeV energies. If one continues the Thémistocle Crab spectrum down to 1 TeV, multiplies the flux by a third and then extends this flux to higher energies using the above index, one concludes that the signal level should be comparable to that of the Crab in the 10 TeV range. A cutoff in the spectrum due to absorption by pair production on the cosmological flux of infrared photons, whose level has never been measured, might be observable somewhere in this energy region. A detection or even an upper limit could help to constrain the cosmological IR flux level under certain assumptions. At the date of the conference, 50 hours of data, as yet unanalyzed, have been taken in the spring of 1993 and in December and January 1993-1994.

Acknowledgments

This work was funded by the Institut National de Physique Nucléaire et Physique de Particules of the CNRS, France. We wish to thank Electricité de France for granting us the use of the facilities of the THEMIS site without charge.

References

- [1] P. Baillon et al., *Astroparticle Physics*, **1** (1993) 341.
- [2] G. Vacanti et al., *Astrophys. J.*, **377** (1991) 467.
- [3] H. Meyer, this conference.
- [4] D. A. Lewis, et al., *Proc. 23rd Inter. Cosmic Ray Conf. (Calgary)*, **1** (1993) 279.
- [5] T. K. Gaisser, *Cosmic Rays and Particle Physics*, Cambridge Univ. Press (1990).
- [6] Recent work on composition is reported in *Proc. 23rd Inter. Cosmic Ray Conf. (Calgary)* **2** (1993).
- [7] J.N. Capdevielle, *J. Phys. G : Nucl. Part. Phys.* **15** (1989) 909
- [8] M. P. Kertzman and G. H. Sembroski, submitted to *Nucl. Instrum. Methods*. See also reference 5, p. 250.
- [9] C. Ghesquière, in *Towards a Major Atmospheric Cerenkov Detector II* (Proc. Calgary Workshop 1993, R. C. Lamb, ed.), in press.
- [10] M. Punch et al., *Nature* **358** (1992) 477.
- [11] D. J. Thompson et al., *Astrophys. J.* **415**, (1993) L13.

STATUS OF THE HEGRA EXPERIMENT
AND SOME EARLY RESULTS

Hinrich Meyer
University of Wuppertal
D-42097 Wuppertal
Germany

Abstract: A brief account of the status of the construction of HEGRA will be given. The array combines 4 different experimental components into one and as such unique array. Data have been taken since some time with parts of the final setup. Some data on the performance of the experiment and also some early results will be presented.

Introduction: The HEGRA experiment is located on the canarian island La Palma, 2.200 m a.s.l. inside the area of the observatorio Roque de los Muchachos (ORM) of the Instituto de Astrofísica de Canarias (IAC), at a longitude of $17^{\circ} 53' 27.4''$ west and a latitude of $28^{\circ} 45' 42.8''$ north. It is above the tree line which follows the upper level of the passat clouds at about 1.800m. Viewing conditions are usually excellent which invites for extensive use of the Tscherenkov light generated in airshowers. The array was started in 1988 as a conventional airshower array of $37 \times 1 \text{ m}^2$ counters by the Kiel group. By now it combines $253 \times 1 \text{ m}^2$ scintillation detectors, $17 \times 16 \text{ m}^2$ Muontowers, $49 \times 0.125 \text{ m}^2$ Tscherenkov stations (AIROBICC) and two telescopes for Tscherenkov light detection. The HEGRA collaboration consists now of two Institutes of the Max-Planck-Gesellschaft in München and Heidelberg, the Yerevan Institute of Physics, Armenia, and Physics Institutes at the Universities of Madrid, Hamburg, Kiel and Wuppertal. In what follows I will give a brief description of the status of the different components of HEGRA, discuss some of the results obtained so far and give an outlook on future developments.

Scintillator Array: A regular array of $169 \times 1 \text{ m}^2$ counters in a 13×13 square with 15 meter spacing defines the boundary of the array. In the center an octagon of $84 \times 1 \text{ m}^2$ counters again with 15 m spacing but shifted by $15/2 \text{ m}$ in both coordinates defines an inner area of 20.000 m^2 with twice the detector density that gives 0,8% coverage.

In a scintillator hut of the 13×13 array two PM's deliver amplitude signals in two dynamic ranges and one time signal from the PM in the lower part of the amplitude range. In the inner part only one PM is used for readout of time and amplitude the high dynamic range required is provided by a second amplitude signal from one of the dynodes. Light pulsers in each counter box are used for frequent calibration data sets (every 10 min) a key procedure for the performance of the array. The master trigger requires at least any 14 stations to fire inside a gate of 150 nsec. The trigger rate is rather stable at about 12 Hz.

AIROBICC: The AIROBICC detectors form a regular 7×7 array with 30 m spacing thus matching the 13×13 scintillator array. Each AIROBICC hut contains a $20\text{cm}\varnothing$ PM, hemispherical in shape with a 40 cm diameter conical lightcatcher, a

filter plate to reduce the night sky background and a LED for calibration purposes. The lid of the box can be opened by remote control. The PM has only 6 stages and a fast amplifier to filter out the very short Tscherenkov pulses from the (almost) constant night sky background. The AIROBICC array gives a master trigger on (more \geq) 6 stations. Due to excellent night sky conditions at ORM the ontime of AIROBICC is about 14% of total time, averaged over 1 year. At good viewing conditions the total trigger rate of AIROBICC and the scintillator array (logic OR of both) is about 18 Hz.

Muontower: In the inner part of the array 17 muontowers of size $3.5 \times 7.0 \times 1.5 \text{ m}^3$ are placed with a spacing of about 30 m. In each tower there are 6 planes of Geiger-tubes 6 m long and 2.7 m wide. Each plane has 160 individual Geigertubes 1.5 cm^2 in cross section and 6 m long. A Geigerpulse is about 0.5 Volt high and 60 μsec long. Below the first and the second plane a layer of lead, 4.5 r.l. thick is placed, this absorbs enough of the very low energy photons and electrons in an airshower to allow a safe identification of muons as a straight track through all planes. For electrons and photons the muon tower also serves as a calorimeter with about 40% energy-resolution from about 100 MeV to 10 GeV. The muontower readout is started on the basis of a master trigger from either component of the HEGRA setup.

Air Tscherenkov Telescopes: Inside the array two Tscherenkov telescopes have been constructed. The first telescope near the center of the array has polar mount, 18 mirrors of 60 cm diameter and a 37 pixel PM camera at 5 meter distance. The telescope tracks sources with better than 0.05° precision and is triggered (usually) by any two pixels. The second telescope has an alt-azimuth mount, 30 mirrors and a camera with 61 pixels both of the some types as in the first telescope. The trigger is again two out of 61 and each trigger starts the readout of the other telescope and all compounts of the array.

Performance of HEGRA and some early results: The arrays have somewhat more than 1 sterad acceptance. The center of the shower is reconstructed with about 10 meter uncertainty and the direction of the particles that started the shower with about 0.6° for the scintillator array and about 0.15° for AIROBICC. The absolut poin-

ting of the scintillator array has been checked to about 0.1° by observing the shadows of moon and sun on the otherwise isotropic TeV cosmic particle flux. The shadow of the sun is slightly less significant than for the moon, this may indicate the influence of the solar magnetic field and deserves careful study at somewhat higher angular resolution.

The scintillator array was used to search for sources of cosmic rays using a method based on the rate in an annular (2° - 5°) off source bin to be subtracted from an on-source bin 1.12° wide. Based on a total data sample of 249×10^6 events, reduced after cuts to 179×10^6 events no source was detected for shower energies of order 50 TeV at a flux level of $(0.7 - 2.8) \times 10^{-13} \text{ cm}^{-2} \text{ sec}^{-1}$. In particular no significant excess has been observed from sources previously claimed like Cyg X3, Her X1, Mk 421, Crab and Geminga at either higher or lower energies. The limits will be further improved with more statistics collected and also some refinements in analysis. More recently (after this meeting) similar limits have been obtained using AIROBICC, where the reduction in statistics due to the ontime is balanced off by the much better angular resolution.

Very significant improvements will be expected based on separation methods for gamma induced showers from showers produced by hadrons. This is also one of the main goals of the muontower analysis both on the basis of muon identification and the radial flow of electromagnetic particle energy. Also the radial shape of the Tscherenkov light pool contains very useful information on the nature of the primary particle. We expect to have developed the analysis methods for separation and collected enough data with all three components of the array by autumn this year, to have significant impact of source searches below 100 TeV.

The first telescope has been used to search for gamma emission from Crab, Her X1, 4UO 115 + 63 and 2CG 078 + 01. While there is now significant signal from the latter three sources, the Crab has been positively identified. There are still rather large uncertainties in acceptance and energy calibration of the telescope, such that only preliminary flux values (or limits) can be quoted. We have for the CRAB signal at three energies 2, 3 and 4 TeV integral fluxes of 1.1×10^{-11} , 4.7×10^{-12} and 2.3×10^{-12} in units $\text{cm}^{-2} \text{ sec}^{-1}$ respectively. This is quite consistent with the observation of Themistocle at higher energies as well as Asgat at lower energy however our flux

is higher than observed by Whipple.

Future developments: 4 more of the second type telescope are under construction at MPI in München and Heidelberg. We plan to have significantly more pixels in new cameras, 217 or more. We expect great improvements in angular resolution and hadron rejection. The telescopes will have 60 - 80 meter distance to each other, a compromise between increase in acceptance of all telescopes added and the coincidence rate between telescopes. This part of the array should be ready by about summer 1995. We further hope to solve the question of energy calibration and acceptance mentioned earlier from the analysis of events that trigger a telescope and will be observed in the arrays as well. Studies along that line have already been started, thus getting closer to the resolution of one of the major experimental problems in TeV airshower detection.

Conclusion: When completed this and next year the HEGRA experiment will be the most powerful airshower array to operate in the TeV energy range. We thus expect significant results on searches for the origin of cosmic rays to come up in the next few years.

COSMIC RAYS ABOVE 0.1 EeV -RESULTS FROM THE FLY'S EYE EXPERIMENT

Presented by P. Sokolsky

D.J. Bird,¹ S.C. Corbató,³ H.Y. Dai,³ B.R. Dawson,²
J.W. Eibert,³ T.K. Gaisser,⁴ M.H.A. Huang,³ D.B. Kieda,³ S. Ko,³
C.G. Larsen,³ E.C. Loh,³ M.H. Salamon,³

J.D. Smith,³ P. Sokolsky,³ P. Sommers,³ T. Stanev,⁴ S. Tilav,⁴ J.K.K. Tang³ and S.B. Thomas³

¹*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA*

²*Department of Physics and Mathematical Physics, University of Adelaide, Adelaide, South Australia 5001
Australia*

³*High Energy Astrophysics Institute, Department of Physics, University of Utah, Salt Lake City UT 84112 USA*

⁴*Bartol Research Institute, University of Delaware, Newark DE 19716*

ABSTRACT

We present recent results from the Fly's Eye experiment. The results from the F.E. I and F.E. II stereo detectors indicate structure in the cosmic ray spectrum above 10^{18} eV. The composition of cosmic rays is found to change from a predominantly heavy composition near 10^{17} eV to a predominantly light one near 10^{19} eV. The spectral structure and changing composition can be accounted for by a simple two-component model. Higher statistics but lower resolution data from the F.E. I detector alone shows the presence of an outstanding 320 EeV energy event.

1. INTRODUCTION

The Fly's Eye experiment has been taking data since 1981 [Baltrusaitis et al. 1985]. A second fly's eye, F.E. II was built in 1986. This second eye views a significant fraction of events that trigger F.E. I. Such events have two independent views and are thus easily reconstructed. In addition to much improved geometrical reconstruction of the EAS trajectory, these events have two independent measurements of their energy and the position of shower maximum in the atmosphere (X_{max}). Such redundant measurements allow us to study the detector resolution in energy and X_{max} . Understanding detector resolution is crucial if structure in spectra are to be believable. Combining the measurements also results in a more accurate value for these variables. We report on results on the cosmic ray spectrum and composition for stereo data taken from 11/86 to 7/92 and monocular data taken since 1/82.

2. DATA ANALYSIS AND RESOLUTION

The air shower reconstruction process falls into two major divisions: geometric reconstruction and shower profile reconstruction. When a shower is seen simultaneously by Fly's Eye I and II, a shower-detector plane for each detector can be determined and the intersection of these planes defines the shower trajectory. If a shower is seen by only one detector, a fit to the relative arrival times of the signals is done and the shower trajectory determined. Given the track geometry, the EAS longitudinal size $N_e(x)$ is calculated via an iterative process to remove the contributions due to direct and scattered Cherenkov radiation. The residual photoelectrons (after the Cherenkov contribution is removed and various attenuation effects between source and detector are taken into account) result from the atmospheric scintillation process and therefore are directly proportional to the charged particle size within a tube's field of view. Each resultant longitudinal shower profile is approximated by a Gaussian function and the total primary energy and X_{max} are found from the fitted parameters.

We compare the energy calculated independently by F.E. I and F.E. II for events registered by both eyes. If E_1 is the energy measured by F.E. I and E_2 is the energy measured by F.E. II, the systematic shift between F.E. I and F.E. II is 2.5%. The distribution in the variable $(E_1 - E_2)/E_{average}$ has a standard deviation of 0.47 for events below $2 \cdot 10^{18}$ eV and 0.40 for events above. The resulting stereo energy resolutions are 24% and 20% respectively. The systematic error in energy is dominated by the scintillation efficiency uncertainty of 20%.

The resolution in the variable X_{max} is found in the same way. The result is consistent with a resolution of 45 gm/cm^2 . Uncertainties in the makeup of the atmosphere and residual Cherenkov light produce a systematic error in X_{max} of 20 gm/cm^2 .

3. SPECTRUM

Fig. 1 shows the stereo energy spectrum derived from the number of events observed and the calculated stereo exposure [Bird et al. 1993a]. It is clear that the spectrum becomes steeper right after $10^{17.6}$ eV and flattens after $3 \cdot 10^{18}$ eV. The change in the spectral slope forms a dip where a minimum lies between $2 \cdot 10^{18}$ and $5 \cdot 10^{18}$ eV. We divide the spectrum into three energy regions determined by eye and fit them to a power law in each region. Table 1 gives the normalization and slope in each region. Also listed in the table is the overall fit to a single power law. The latter is clearly a poor fit. The expected number of events from the overall fit can be compared to the actual number of events. The comparison of 5936 expected and 5477 observed between $10^{17.6}$ eV and $10^{19.6}$ eV gives a deficit of 5.96 sigma. An alternative way of showing the significance is to show the excess of observed events above $10^{18.5}$. Here we use the normalization and slope from the overall fit up to $10^{18.5}$ to calculate the expectation. The total observed events are 283 while the expectation is 230, leading to a 3.49 sigma excess.

Table 1. Spectral slopes and normalizations of J(E) ($\text{m}^{-2}\text{sr}^{-1}\text{s}^{-1}$)

Energy range (eV)	Power index	$\log(\text{normalization})$	Normalized at
$10^{17.3} - 10^{19.6}$	-3.18 ± 0.01	-29.593	10^{18} eV
$10^{17.3} - 10^{17.6}$	-3.01 ± 0.06	-29.495	10^{18} eV
$10^{17.6} - 10^{18.5}$	-3.27 ± 0.02	-29.605	10^{18} eV
$10^{18.5} - 10^{19.6}$	-2.71 ± 0.10	-32.623	10^{19} eV

The stereo spectrum shows clear evidence of structure. The structure is consistent with appearance of a new, harder spectrum which dominates the softer component at energies above $5 \cdot 10^{18}$ eV. It is of great interest to ask if the composition of cosmic rays undergoes a change in this energy interval as well.

4.COMPOSITION MEASUREMENT

The position of shower maximum in the atmosphere (X_{max}) in gm/cm^2 is sensitive to the composition of the parent particle. Protons, for instance, will interact more deeply in the atmosphere than heavy nuclei. Air showers produced by protons are also expected to have larger fluctuations than those produced by heavy nuclei. As a result, measuring the distribution of shower X_{max} can be used to infer the primary cosmic ray composition.

5. THE DATA AND MONTE CARLO SIMULATION

We impose the following cuts on the data: $\delta X_{max}/X_{max} \leq 0.12$, minimum viewing angle of pmt tube to track ≥ 20 degrees, and energy ≥ 0.1 EeV. The first cut removes poorly reconstructed events, the second removes events with strong Cherenkov light contamination and the third insures that the data is above the detector trigger threshold. After cuts, we have 8790 events in the energy range from .1 to 30 EeV and 5129 events from .3 to 30 EeV.

Monte Carlo data is generated by using two different kinds of simulations [Gaisser et al. 1993]. The first simulates the extensive air shower development. Because the nature of the hadronic interactions at these energies is not well known, we use three different hadronic models. The first two are high inelasticity models (QCD Pomeron and QCD minijet) while the third is a low inelasticity statistical model. The parameters of the models are described in [Gaisser et al 1993] in more detail. All three models use the same inelastic total cross-section energy dependence (near $\log(s)$). The models are all good fits to accelerator data for p-p interactions but represent different extrapolations of the behaviour of the fragmentation region to the energy range of interest. Nucleus-air interactions are modelled by following the reinteractions of the beam in the nucleus in a way that is appropriate for each hadronic model. The statistical model leads to a very poor fit to our data for any composition assumption and is hence ruled out. The QCD Pomeron and QCD minijet models result in adequate fits (they differ in their predictions by order of $10 gm/cm^2$ whereas, as will be seen below, the position of the X_{max} peak in any given model changes by $80 gm/cm^2$ from purely protonic to a purely iron composition). The QDC-Pomeron model gives the best overall fit to the data, hence we will use it alone in the following discussion.

The second simulation takes the generated extensive air shower profile and generates the appropriate scintillation and cherenkov light, propagates the light through the atmosphere and determines which tubes in F.E. I and F.E. II would trigger. For triggered tubes, the program determines the resulting pulse integrals and relative firing times. The program then generates a set of fake events, whose format is identical to the real data. This fake data set is then reconstructed using the same programs that we use for real data.

6. COMPARISON OF DATA WITH MONTE CARLO

We compare the measured X_{max} distributions to expectations for pure Fe and pure proton monte carlo data in four energy bins. Fig. 2 shows data and Fe and pure proton X_{max} simulated data normalized to equal areas. Note that the falling part of the X_{max} distribution (the decrement) depends on the inelastic cross section with air nuclei, hence the proton distribution falls more slowly than iron. The falling slope of the data is consistent with the proton slope but the rise and the peak of the data distribution is the same as predicted for iron, at least in the lower energy bins. The data thus requires a significant admixture of iron or similar heavy nuclei in the lower energy bins. Data in the higher energy bins clearly require a larger admixture of protons while the 3 to 10 EeV bin data can be largely accounted for by protons alone. Detailed fits show that the proportion of protons rises from $.21 \pm 0.07$ near .3 EeV to $.43 \pm 0.04$ for events with > 1.0 EeV energy.

7. ELONGATION RATE

The elongation rate is the rate of change of the average depth of shower maximum per decade of shower energy. Fig. 3 shows this dependence for data, pure Fe and pure proton showers using the QCD Pomeron hadronic model. The elongation rate for any pure elemental composition is $50 gm/cm^2$ per decade of energy. It follows that a mixed but unchanging composition is also expected to have the same elongation rate. The elongation rate for the data, from 0.1 to 10 EeV is $69 \pm 1.9 gm/cm^2$. Examination of the data shows that a single straight line is not a good fit. The elongation rate from 0.3 to 10 EeV is $78.9 \pm 3 gm/cm^2$ while the elongation rate below 0.3 EeV is consistent with the expected $50 gm/cm^2$. This picture also implies a composition that is growing lighter with energy, going from a heavy composition near 0.3 EeV to a proton dominated composition near 10 EeV.

8. CONSISTENCY OF SPECTRUM AND COMPOSITION DATA

The fact that both the spectrum and the composition of cosmic rays undergo a change near 3 EeV may imply that we are observing a new cosmic ray component emerging from the lower energy flux. To test this

hypothesis, we fit the stereo spectrum to a two component power law. Fig. 5 shows the result of such a fit. The best fit results from $10^{17.6}$ to $10^{19.6}$ are: $\log(j(E)(m^{-2}s^{-1}sr^{-1}eV^{-1})) = 33.185 - 3.496 \cdot \log(E)$ for the softer energy component and $\log(j(E)(m^{-2}s^{-1}sr^{-1}eV^{-1})) = 16.782 - 2.610 \cdot \log(E)$ for the harder energy component. If we associate the softer energy component with a predominantly heavy (Fe) composition and the harder energy component with a predominantly light (p) composition, we can predict the ratio of Fe to protons as a function of energy. The ratio can be expressed as: $ironflux/protonflux = (E/10^{18.5}eV)^{-0.887}$.

Given this ratio and the expected elongation rates for Fe and proton showers shown in Fig. 4, we can predict the position of the average X_{max} as a function of energy. The results are shown as the diamonds in Fig. 4. The excellent agreement between this simple two component model which comes from a fit to the spectrum and the composition results which come from the X_{max} distribution supports the conclusion that we have indeed observed a transition between two different spectra.

9. MONOCULAR SPECTRUM

Compared to the stereo data, the monocular data set is much bigger but the time fitting leads to larger geometrical errors and thus a larger energy uncertainty. Fig. 5 shows the total monocular spectrum. Due to limited energy resolution this spectrum does not show the degree of structure found in the stereo data. The best fit to the overall spectrum yields $J(E) = 10^{-29.95}(E/10^{18}eV)^{-3.07 \pm 0.01} m^{-2} s^{-1} eV^{-1}$. The figure shows a flattening of the spectrum above 10^{19} eV. The significance of the flattening based on the spectrum slope and normalization below $10^{19}eV$ is 3.5σ . The flattening extends for only a decade and appears to steepen after $10^{19.7}eV$. We expect 20.6 events above this energy if the flattened spectrum continues but observe 10. More statistics is needed to resolve this.

An event with an energy of 320 EeV is observed. This well measured event is consistent with expectations for either a proton or an Fe nucleus. Because of its great energy, this event is unlikely to come from a distance of greater than .30 Mpc or to be bent by more than 10 degrees by magnetic fields. No obvious astrophysical sources exist in the resulting error box.

9. DISCUSSION

Taken together, the Fly's Eye results on the cosmic ray energy spectrum and chemical composition strongly suggest a transition near $10^{18.5}$ eV to a new source, most likely of extragalactic origin. The present stereo detector runs out of sensitivity above 10 EeV because of limited aperture. Monocular data suffers from limited energy and X_{max} resolution. It is clearly of great interest to extend the X_{max} and spectrum measurements to this region with high statistics. Confirmation of this result requires a detector with at least an order of magnitude increase in aperture at energies above 10 EeV and similar or better resolution in X_{max} . The High Resolution Fly's Eye (HiRes) detector [Bird et al. 1993d] is presently in the prototype stage. Data from the prototype indicates that a full HiRes detector will have an aperture that is more than adequate to answer these questions. The construction and operation of this detector in the next few years should shed further light on this extremely interesting energy region.

REFERENCES

- Baltrusaitis, R.M. et al.: 1985, Nucl. Instr. Meth., A240, 410
 Bird, D. et al.: 1993a, submitted to Ap. J. Bird, D. et al.: 1993b, Proc. 23rd ICRC, Vol. 2, 55
 Bird, D. et al.: 1993c, Proc. 23rd ICRC, Vol. 2, 51
 Bird, D. et al.: 1993d, Proc. 23rd ICRC, Vol. 2, 458
 Chi, X. et al.: 1992, J. Phys. G., 18, 539
 Gaisser, T.K. et al.: 1993, Phys. Rev. D. 47, 1991

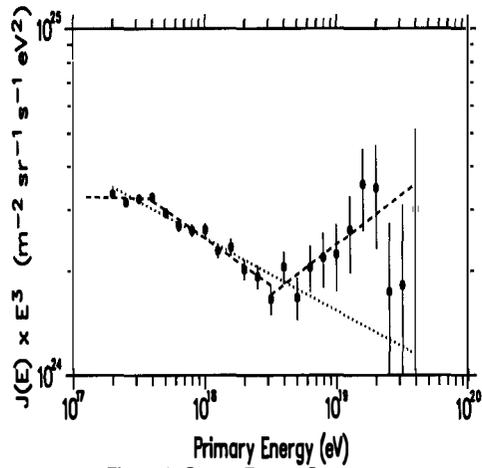
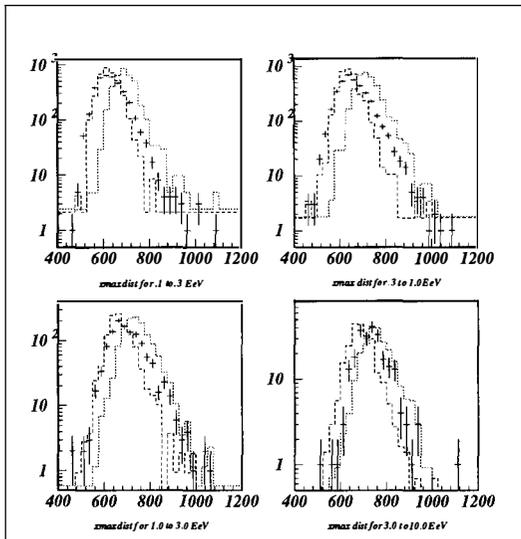


Figure 1: Stereo Energy Spectrum

Figure 2: X_{max} distribution for data (crosses), Fe (short dashes) and protons (long dashes) for four energy bins: 0.1 to 0.3 EeV; 0.3 to 1.0 EeV; 1.0 to 3.0 EeV; and 3.0 to 10.0 EeV.

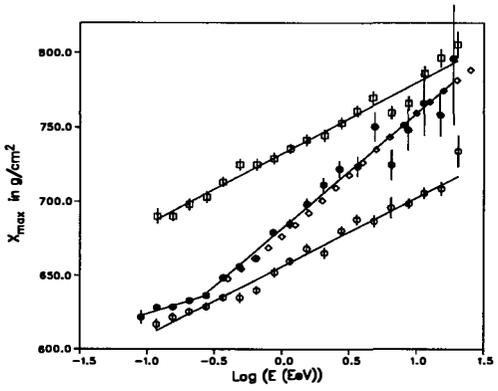


Figure 3: Elongation rate data (crosses), Fe (stars) and protons(circles). X axis is $\log_{10}(E)$. Y axis is X_{max} in gm/cm^2 . The error bars indicate the statistical error on the mean.

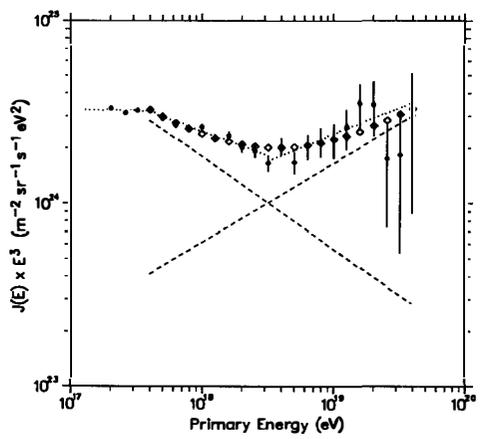


Figure 4: Two Component Fit to Stereo Spectrum

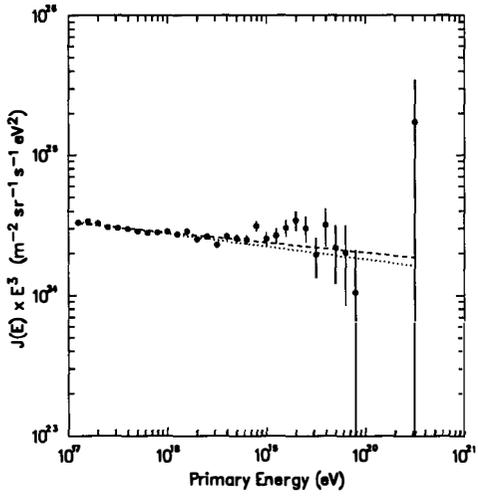
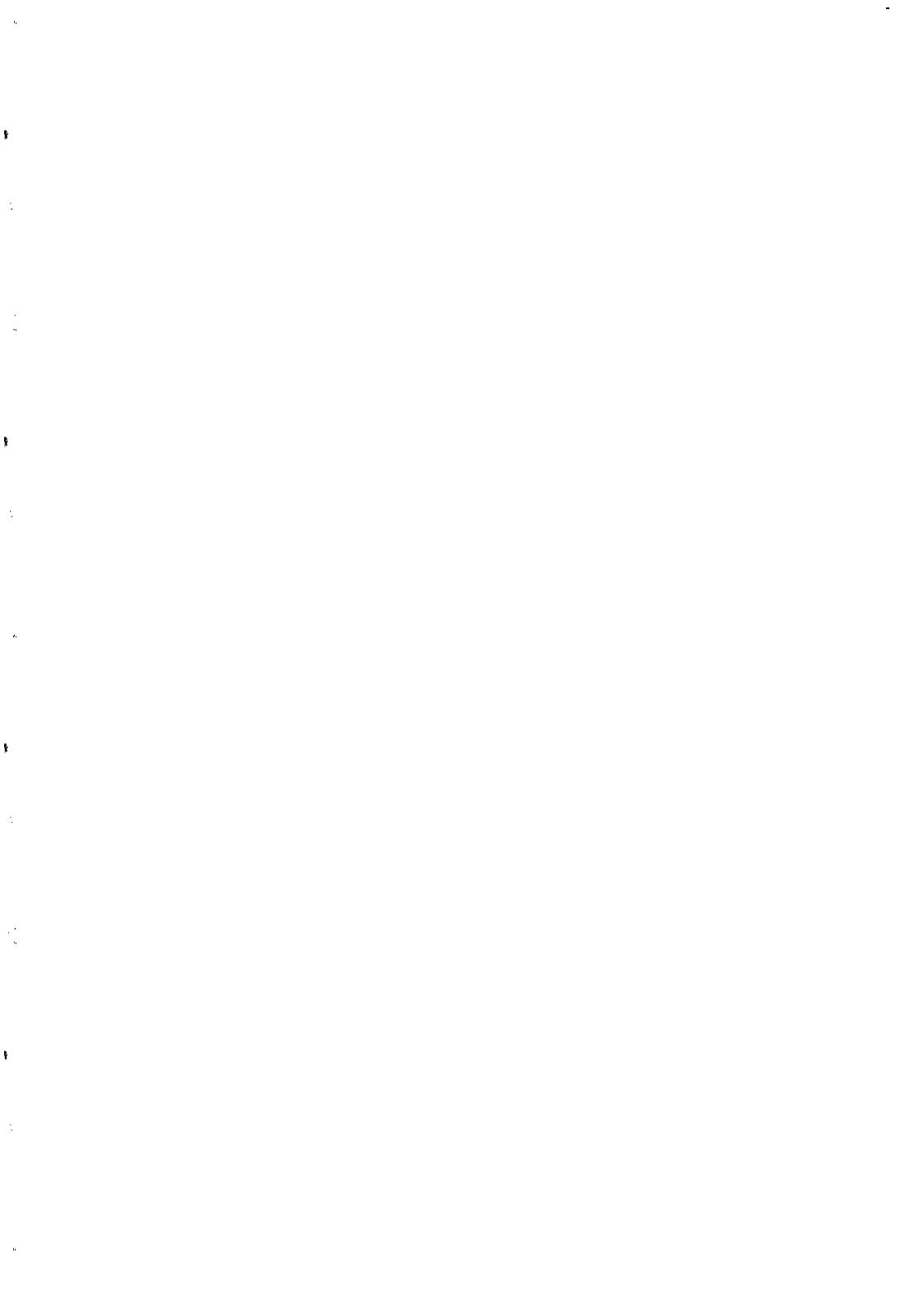


Figure 5: Monocular Energy Spectrum



The HiRes Monte Carlo Simulation

Hongyue Dai

High Energy Astrophysics Institute,

Department of Physics, University of Utah, Salt Lake City UT 84112 USA

Abstract

A Monte Carlo study of the HiRes energy resolution, X_{max} resolution, and aperture is reported in this paper.

1. Introduction

The University of Utah is collaborating with three other institutions to construct a high resolution Fly's Eye (HiRes) at Dugway, Utah. So far, a 14 mirror prototype has been taking data for more than one year. The first stage construction (half of the total number of mirrors) will soon start simultaneously at two sites separated by 12.6 kilometers. This paper will concentrate on the aperture, the energy and the composition resolution of the HiRes detector modeled by the Monte Carlo simulation.

The detector Monte Carlo simulation is a very useful tool for understanding and predicting the system performance. The HiRes Monte Carlo program was written based on our experience from the Fly's Eye experiment and the HiRes prototype. The reliability of the Fly's Eye detector Monte Carlo has been carefully checked by the Fly's Eye data. Compared to the Fly's Eye, the HiRes performance is more predictable because of its greatly improved optics and electronics, as well as its carefully redesigned calibration system.

2. Brief description of the HiRes detector simulation

a) Detector configuration

The HiRes detector in this simulation includes 2 eyes facing each other separated by 13 kilometers with each eye consisting of 54 mirrors. In the focal plane of each mirror, 256 hexagonal phototubes (16 by 16) view a $16 \times 13.5^\circ$ section of the sky. The 54 mirrors are arranged into 4 elevation rings (10.24° , 23.70° , 37.00° and 50.09° are the elevation angles for the ring centers), with each ring covering approximately 240° in azimuth angle.

An optical filter is added in front of the PMT (Photo-Multiplier Tube) cathode to reduce the sky background noise. In addition, for each PMT there is a choice of electronic

filter (either an RC or a Bessel filter) to smooth its pulse. The spot size due to the mirror aberrations is modeled by ray tracing. The PMT gain, parameters for the electronics, and the threshold used in this simulation are the actual numbers used in the HiRes prototype.

b) Shower track generation

For a given energy, the longitudinal shower profile can be generated either by an analytic formula (Gaisser-Hillas function) or by a full shower Monte Carlo (for composition study, we use Gaisser and Stanev's shower Monte Carlo program). The track geometry is randomly generated in space. Given the track geometry and shower profile, the Cerenkov beam and the air scintillation light associated with the shower are calculated according to the actual air density. Finally, the PMT signal is calculated by taking into account various scattering and attenuation mechanisms (see Baltrusaitis et al 1985). If the fired tubes satisfy the trigger pattern in both eyes, the tube signals are recorded.

c) Track and shower profile reconstruction

The shower track and the detector uniquely determine a plane, and we call this plane the shower-detector plane. The PMT hits can be used to determine such a plane. If a track is viewed by both eyes, the intersection of two shower-detector planes defines the track geometry.

The shower profile is fit to a 3-parameter Gaisser-Hillas function by the minimum χ^2 method. The attenuation and scattering between the light source and the detector is taken into account also in this reconstruction procedure.

3. Energy resolution of the HiRes detector

We get the HiRes energy resolution by comparing the reconstructed energy with the energy input to the detector Monte Carlo. Fig.1 shows the histogram of $(E_{reconstructed} - E_{input})/E_{input}$ at primary energy of 10 EeV. For each primary energy, 2000 events are generated and analyzed. The only cut applied to these histograms is the requirement of the shower maximum bracketing, that is, at least one eye should see the shower before it reaches its maximum. The sky background noise (0.02 photo-electrons per ns per tube) has been included into the simulation. At present, atmospheric fluctuation is not incorporated in the Monte Carlo, but according to our experiences with the Fly's Eye simulation, for the selected nights when we run the detector, the contribution to the energy resolution from atmospheric fluctuation is less than 10%. For the HiRes experiment, a Polaris monitor plus some computer controlled lasers will be mounted around each site to monitor the

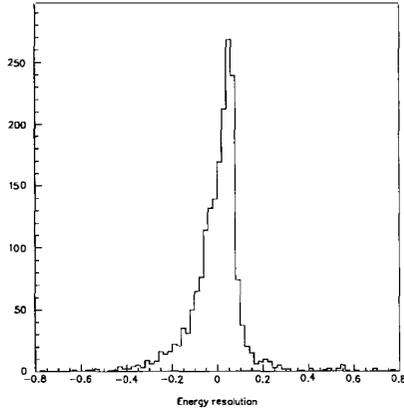


Figure 1: HiRes energy resolution at 10 EeV. X axis: $(E_{reconstructed} - E_{input})/E_{input}$; Y axis: number of events.

atmospheric conditions, which will improve our atmospheric correction significantly. The energy resolution for different energies are summarized in Table 1.

Table 1. Energy Resolution (FWHM) for the HiRes Detector

Energy (EeV)	1.0	3.0	10.0	30.0	100.0
Resolution	16.6%	14.0%	12.1%	11.5%	7.6%

4. X_{max} resolution

The X_{max} (depth at shower maximum) is chosen as a parameter for the composition study by the Fly's Eye group because of its intrinsic separation between proton and iron primaries (see Table 2). The energy of the events discussed in this section is from 10 to 30 EeV and is sampled from the Fly's Eye stereo energy spectrum. To test the sensitivity of the HiRes detector to X_{max} , we used Gaisser and Stanev's shower Monte Carlo to generate shower profiles for proton, carbon and iron. The generated showers are then input to the HiRes detector Monte Carlo to account for the detector bias and resolution. Fig.2 shows the histograms of the difference between the reconstructed X_{max} and the input X_{max} . The mean X_{max} and resolution for three primaries are summarized in Table 3. The X_{max} discussed in this paper is the X_{max} after the elongation rate deconvolution, that is, $X_{max} = X_{max}(E) - 55 \times \log_{10}(E)$, where E is the primary energy in EeV and

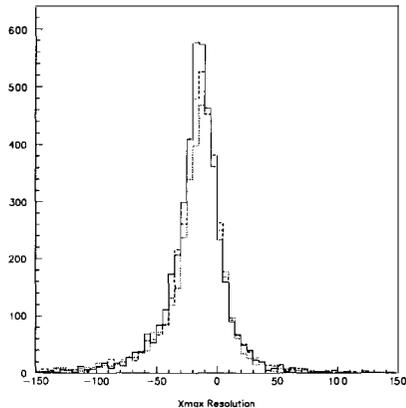


Figure 2: HiRes X_{max} resolution for proton (dotted line), carbon (dashed line), and iron (solid line) initiated showers. X axis: difference between the reconstructed X_{max} and the input X_{max} in g/cm^2 ; Y axis: number of events.

55 is the mean elongation rate. From Table 3 we see that the X_{max} resolution is about $30 g/cm^2$ in FWHM regardless of the primary composition. The cut applied here is also the X_{max} bracketing cut. The systematic shift due to triggering and reconstruction bias is approximately constant for different compositions, therefore the difference in X_{max} for proton and iron is kept after the reconstruction.

Table 2. P, C, and Fe separation and intrinsic width
(g/cm^2), 10 to 30 EeV

(from Gaisser and Todor's shower Monte Carlo)

	P	C	Fe
\bar{X}_{max}^*	739.0	688.6	652.3
$FWHM^*$	101.9	79.0	63.7

* With elongation rate deconvoluted, i.e., $X_{max} - 55 \times \log_{10}(E(EeV))$

Table 3. P, C, and Fe resolution
(g/cm^2), 10 to 30 EeV
(after detector Monte Carlo)

	P	C	Fe
FWHM	32	32	30
\bar{X}_{max}^*	726.1	680.7	641.9

* Reconstructed mean X_{max} , with elongation rate deconvoluted

5. HiRes aperture and expected results for 5 years

The estimated HiRes apertures at different energies are listed in Table 4. The cut applied in the aperture calculation requires ΔX_{max} to be less than $30 g/cm^2$. The expected number of events based on the Fly's Eye stereo spectrum for 5 years running period is also listed in Table 4 (the continuation of the spectrum to above 10^{21} eV and 10% duty cycle are assumed).

Table 4. HiRes aperture ($km^2 sr$) and
number of events expected (5 years)

Energy (EeV)	1.0	3.0	10.0	30.0	100.0
Aperture	1001	2255	4042	6009	7240
Events	31349	5834	1316	278	60

We have also performed Monte Carlo simulations of the HiRes energy spectrum for 5 years running time for two scenarios: with or without a cutoff at 60 EeV. The HiRes will observe a sharp turn-off around $10^{19.7}$ eV if the spectrum does cut off. The observed number of events above $10^{19.7}$ eV is about 15 including the spill-over by the energy resolution. Whereas the observed number of events above $10^{19.7}$ eV is about 100 if the spectrum continues to 10^{21} eV. Clearly the HiRes will be able to tell one story from the other in 5 years of operation. There are many calculations about the cutoff and pile up of the cosmic ray spectrum due to the interaction between the extremely high energy cosmic rays and the 2.7° micro-wave background (see for example, Yoshida & Teshima 1990). The exact shape of the energy spectrum near the cut off depends on the source distance. Dr. Sommers and Dr. Elbert in the Fly's Eye group also did such a calculation and their results show that the HiRes should be able to tell whether the extremely high energy cosmic rays mainly come from local clusters or cosmological sources through careful work.

The X_{max} distribution for events between 10 EeV and 30 EeV in 5 year's operation

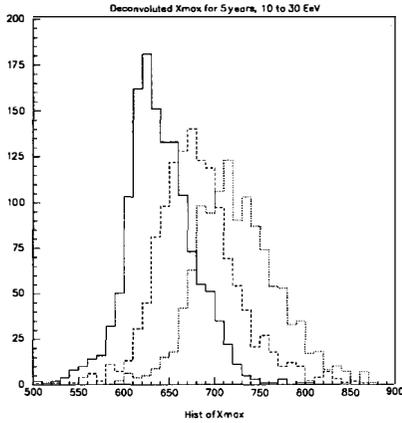


Figure 3: Reconstructed X_{max} distribution for proton (dotted line), carbon (dashed line), and iron (solid line) initiated showers with energy between 10 and 30 EeV. X axis: X_{max} in g/cm^2 , with the elongation rate deconvoluted (see text); Y axis: number of events for 5 years of statistics.

is shown in Fig.3. The HiRes should have enough resolution and statistics to tell the composition in this energy range.

Summary

The HiRes has good energy resolution, good X_{max} resolution and reasonably big aperture to measure the cosmic ray spectrum and composition before and around 10^{20} eV.

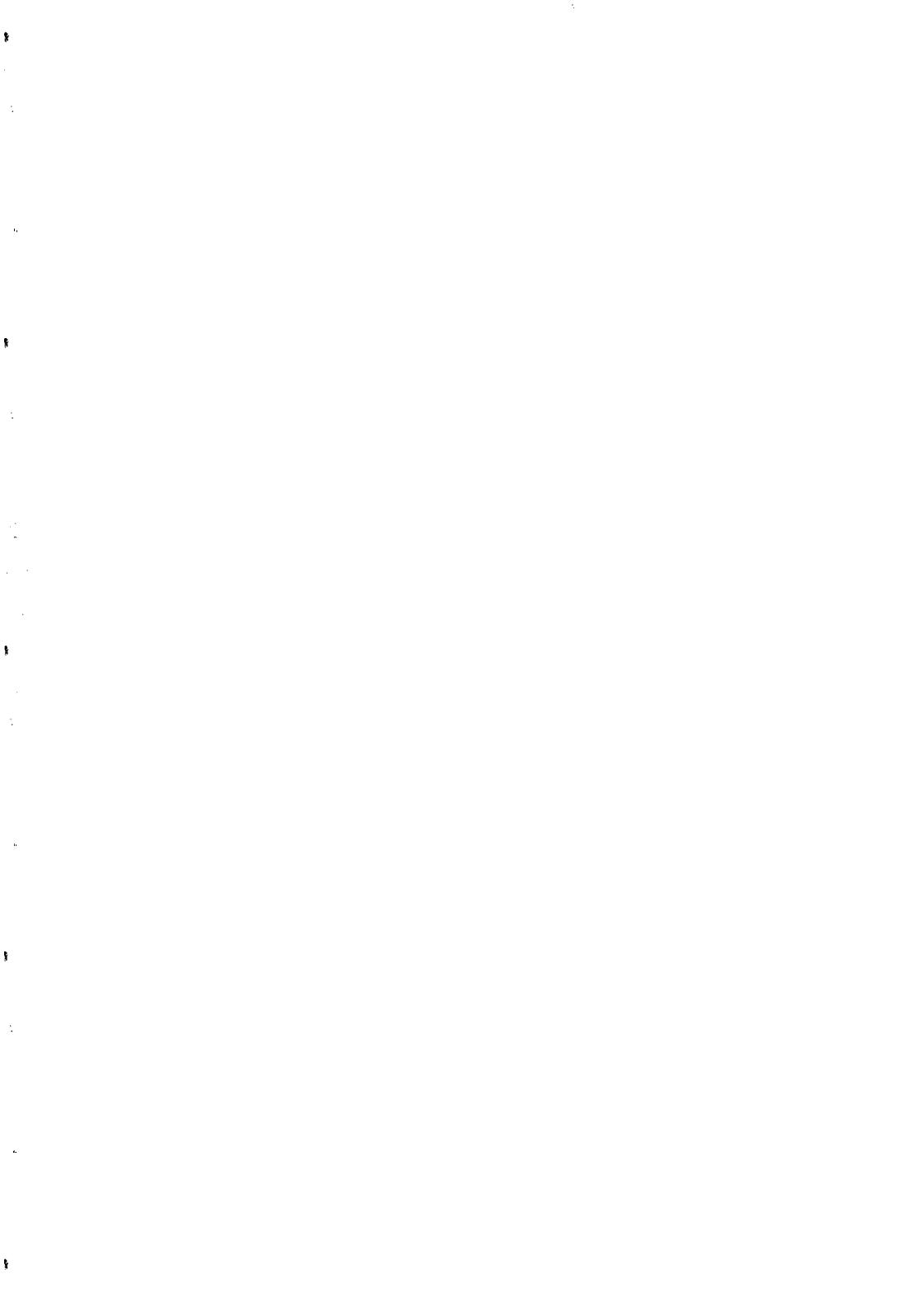
Acknowledgements

I thank Dr. Jerry Elbert and Dr. Paul Sommers for reading this manuscript. Stan Thomas contributed the ray tracing program for the HiRes simulation. The work has been supported in part by NSF grant PHY-91-00221 and NSF travel grant through Prof. E. Hinds of Yale University.

References

- Baltrusaitis R.M. et al, 1985, Nuclear Instruments and Methods, A240, 410
- Yoshida S. & Teshima M., 1990, Proc. 21st Int. Cosmic Ray Conf., 3, 165

DARK MATTER AND COSMOLOGY



ON THE ABUNDANCES OF THE LIGHT ELEMENTS
IN THE UNIVERSE

Sylvie Vauclair

Observatoire Midi-Pyrénées
and Université Paul Sabatier
Toulouse

ABSTRACT

The light elements D, ^3He , ^4He , ^7Li are known to be formed in substantial quantities in Standard Big-Bang Nucleosynthesis. For this reason, they can be used as constraints on the baryonic density of the universe (average value and inhomogeneities). It is thus important to be able, from the available observations, to obtain precise values of their primordial abundances. This paper presents a short review of the observed abundances of these elements in the Universe and their evolution with time. The uncertainties on their primordial abundances are discussed.

Most of the chemical elements which compose the matter presently observed in the Universe have been formed in stars, either during quiet evolutionary phases or during explosive events (novae or supernovae). The “light elements” deuterium, helium, lithium, beryllium and boron are special in this respect, as they cannot have been formed in stars at least during standard evolutionary sequences up to their observed abundances. Part of them (${}^6\text{Li}$, ${}^9\text{Be}$, ${}^{10}\text{B}$ and ${}^{11}\text{B}$) may have been formed in the interstellar matter, due to spallation reactions induced by cosmic rays. The remaining elements: D, ${}^3\text{He}$, ${}^4\text{He}$ and ${}^7\text{Li}$ are precisely those which are naturally formed in standard models of the primordial Universe. This so-called “Standard Big-Bang Nucleosynthesis” (SBBN) thus represents a real success of the theory.

It is important to gather informations on the abundances of these elements so as to obtain precise constraints on the physical conditions in the primordial universe, and particularly on the baryonic density. We will first study the deuterium and helium isotopes, and then we will focus on the ${}^7\text{Li}$ problem.

1. Deuterium and Helium

The hydrogen and helium isotopes D, ${}^3\text{He}$, ${}^4\text{He}$ are all formed in significant amounts in Standard Big-Bang Nucleosynthesis. The subsequent evolution of their relative abundances is however quite different from one another. Deuterium is destroyed by proton capture at about 600 000 K, leading to ${}^3\text{He}$. The ${}^3\text{He}$ abundance evolution is basically due to this effect, while the total D+ ${}^3\text{He}$ abundance does not vary by more than a factor 2. As for ${}^4\text{He}$, its variation with time depends basically on the amount of stellar matter ejected into the interstellar gas. It can be inferred from photometric observations of galaxies. We will first study the case of deuterium, then ${}^3\text{He}$ coupled with D, and finally ${}^4\text{He}$.

1.1. Deuterium

The most recent and best observations of deuterium in the local interstellar medium have been obtained with the Hubble Space Telescope by Linsky et al. 1992 and Lemoine et al. 1994.

Linsky et al. 1992 presented the first detection of the deuterium Ly α line in the light-of-sight of Capella. They gave the value:

$$\left(\frac{D}{H}\right) = 1.65(+0.07, -0.18) \times 10^{-5}$$

Lemoine et al. 1994 observed the same deuterium line in the line-of-sight of the white dwarf G191-B2B. A fit of 16 interstellar lines of other elements observed in the same direction allowed them to derive a good analysis of the structure of the interstellar medium in this line-of-sight. They found that it could be divided into three different clouds, one of them (their “cloud B”) being the same as Linsky’s

one. They derive for their cloud C a deuterium abundance value compatible with Linsky's one (inside 1σ), namely:

$$\left(\frac{D}{H}\right) = 1.5 \times 10^{-5}$$

However the lines corresponding to clouds A and B are blended so that assuming for cloud B a value compatible with Linsky's one leads to assume for cloud A a smaller value. With $\left(\frac{D}{H}\right) = 1.4 \times 10^{-5}$ for cloud B, they obtain $\left(\frac{D}{H}\right) = 1.0 \times 10^{-5}$ for cloud A.

Two questions have then to be solved before reaching the primordial deuterium abundance:

- 1) Is the deuterium abundance scatter in the interstellar medium real?
- 2) What is the deuterium abundance variation with age in the Galaxy?

The galactic evolution of deuterium has been computed theoretically, taking into account that all the deuterium which passes through stars is burned out. The result is that the deuterium abundance may have varied by a factor about 2 to 5 since the epoch of Big-Bang Nucleosynthesis (Vidal-Madjar 1991, 1993).

1.2. Helium 3 and deuterium

^3He is basically observed in the solar wind and in meteorites. Space observations of $^3\text{He}/^4\text{He}$ in the solar wind obtained with Apollo (Geiss and Reeves 1972) and ISEE-3 (Coplan et al 1984) experiments give an average value of $^3\text{He}/^4\text{He} = (4.5 \pm 0.4)10^{-4}$. Considering that D has been transformed into ^3He and taking $^4\text{He}/\text{H}=0.10$ as the original solar value, this gives a protosolar value (Geiss 1992):

$$\left(\frac{D + ^3\text{He}}{H}\right) = (4.1 \pm 1.0)10^{-5}$$

The uncertainty takes into account the fact that ^3He may have been slightly enriched in the solar convection zone since deuterium burning due to mixing inside the sun.

The ^3He abundance observed in meteorites varies from $(1.52 \pm 0.04) 10^{-5}$ for carbonaceous chondrites to $(4.03 \pm 0.19) 10^{-5}$ for gas-rich meteorites. According to Geiss and Reeves 1972, gas-rich meteorites include solar deuterium transformed into ^3He , while carbonaceous chondrites are unaffected by deuterium burning. These results are consistent with a protosolar (D/H) value of about $2.5 \cdot 10^{-5}$, so that:

$$\left(\frac{D + ^3\text{He}}{H}\right)_p = (6 \pm 2)10^{-5}$$

1.3. Helium 4

As ^4He is the second most abundant element in the universe, many observations of absorption and emission lines are available, and it becomes possible to draw regression curves with time and extrapolate them down to the primordial abundance. Helium abundances are generally given in terms of Y , the helium mass fraction. Reviews of the recent observations of helium in extragalactic HII regions may be found in Pagel et al 1992 and Skillman et al. 1993 and 1994. The basic problems in these regression curves are the dispersion of the observed values, and the fact that they are drawn against oxygen and nitrogen, which may suffer some bias. Pagel et al 1992 discuss the bias which are due to Wolf-Rayet emission features. They claim that if HII galaxies known to have broad WR emission features are excluded from the sample, the regression curves are linear. The Y value that they deduce for the primordial abundances is:

$$Y_p = 0.228 \pm 0.005$$

Skillmann et al. 1993 and 1994 have added new observations to the Pagel's sample, and discuss possible systematic errors due to He I emissivity calculations. They are lead to increase Y up to values between $Y = 0.231$ and 0.239 .

2. The lithium problem

Since the pioneering work of Reeves, Fowler and Hoyle 1970, it is generally believed that ^6Li is formed by spallation reactions in the interstellar matter as well as Be and B, while most of ^7Li has to be formed elsewhere. An obvious formation site for ^7Li is Big-Bang Nucleosynthesis. After Spite and Spite 1982 observations of ^7Li in halo (old) stars, which showed that it is about ten times less abundant than in galactic (young) stars, the ^7Li primordial abundance is generally supposed to be the halo one. This assumes that no lithium destruction occurred in halo stars, during more than ten billion years. This assumption is difficult to believe however, as lithium destruction is indeed observed in younger stars.

The observations of lithium abundances in galactic clusters show evidences of lithium variations with stellar effective temperature and age. Several interpretations of these variations have been proposed in the litterature (Michaud 1986, Vauclair 1988, Charbonneau and Michaud 1988, Pinsonneault et al 1990 and 1992, Garcia-Lopez and Spruit 1991, Dearborn et al 1992, Charbonnel et al. 1992).

The observations of lithium in halo stars are compared to those in the Hyades in figure 1. The observational points follow average lines which cross at low temperature. This is probably due to the depth of the convection zone which is smaller in low metallicity stars (halo stars) than in high metallicity (Hyades) stars due to opacity effects.

Note that a similar dip as observed in the Hyades could exist for halo stars, at a slightly larger effective temperature, but most of these stars must have evolved away from the main sequence.

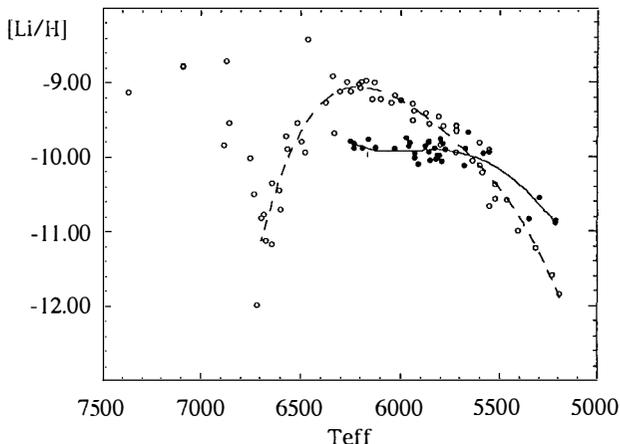


Figure 1: Observational data on the lithium abundances in the Hyades (open circles) and halo stars (filled circles). The curves represent average fits of the observed values. The lithium decrease at low temperature is due to nuclear destruction below the convection zone. The curves crossing is due to the fact that the convection zones are shallower in halo stars than in galactic disk stars. The yet unsolved question is about the plateau in halo stars: does it represent the primordial value or has lithium been destroyed in those stars?

Figure 2 gives the observations of lithium, beryllium and boron abundances in stars as a function of metallicity. The upper envelope of the lithium abundances becomes flat at low metallicities, with a value of $\frac{Li}{H}$ about 2.10^{-10} , while no evidence of such "plateau" is observed for beryllium and boron. The dispersion of the abundances is much larger for lithium than for beryllium and boron: this is due to lithium nuclear destruction in stellar envelopes.

The effect of gravitational settling on lithium in halo stars has been studied by Michaud, Fontaine and Beaudet 1984, Deliyannis et al 1990, Proffitt and Michaud 1991. It leads to a non uniform depletion with an average value of about a factor 2. Macroscopic motions can prevent this depletion, but they induce lithium mixing and destruction. Rotation induced mixing could even deplete lithium by a factor 10 and lead to a plateau-shape as observed (Vauclair 1988, Pinsonneault et al 1992). The basic problem for these computations is that we do not know the rotation history of these stars, which has to be postulated.

Is the maximum lithium abundance observed in low metallicity stars the primordial one, or has lithium been depleted in all these stars? For the lithium abundance in halo stars to represent indeed the primordial one, a fine tuning of turbulence is needed so that gravitational settling may be prevented or at least slowed down

without lithium being burned. This problem is not yet solved. More observations of lithium abundances at the two extremes of the plateau, and detailed spectroscopic observations of metals will help giving constraints on the process involved.

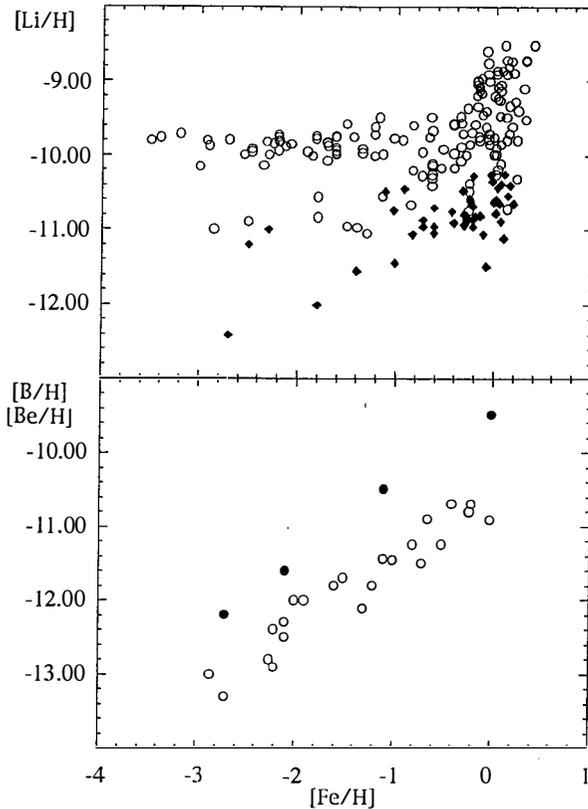


Figure 2: Lithium (top) and beryllium and boron (bottom) observations versus metallicity in stars. Most data are those gathered by Walker et al. 1991 and Prantzos et al. 1993. In the top graph, the open circles represent real values while the diamonds represents upper limits. In the bottom graph, the open circles represent the beryllium abundances and the filled ones the boron abundances. The dispersion is larger for lithium, due to nuclear destruction. The upper envelope of the lithium abundances becomes flat at low metallicities while no evidence of such "plateau" can be found for beryllium and boron.

3. Conclusion

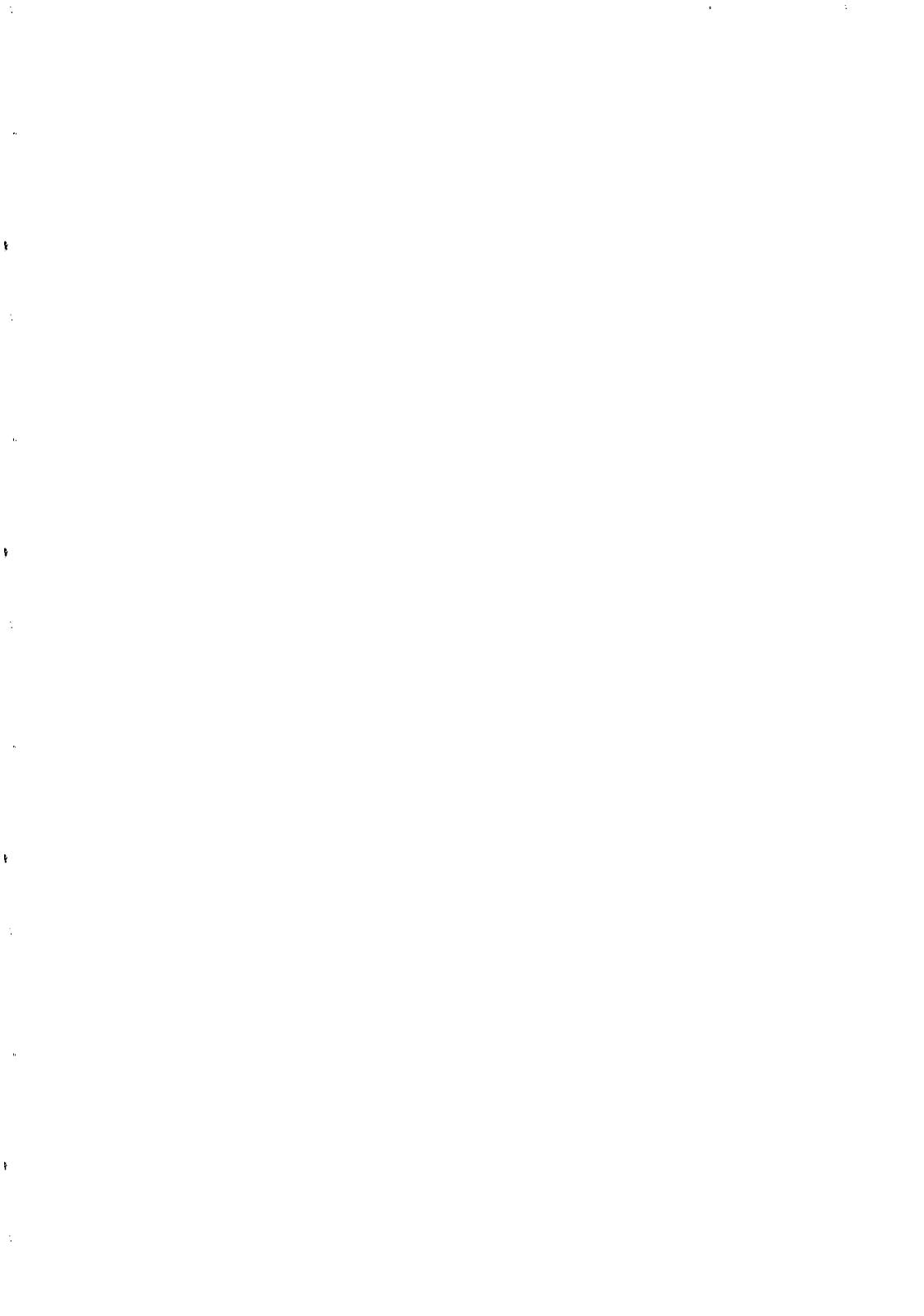
Within the present uncertainties on the observed values, the abundances of the primordial elements D, ^3He , ^4He and ^7Li are all compatible with the same baryonic number $\eta \simeq 3$ to 4×10^{-10} , in the frame of the SBBN (see Reeves, 1994), which corresponds to $\Omega_B = 0.01$ to 0.03 . More precisions are to be expected from new observations with the Hubble Space Telescope and with large ground-based instruments.

Note added in proof

After this conference we learned that a deuterium line has been discovered in a metal poor intergalactic cloud with the KECK telescope at Hawai by A. Songaila, L.L. Cowie, C. Hogan and M. Rugers (preprint submitted to Nature). If the detection is confirmed, it gives for the primordial deuterium abundance a value $\frac{D}{H} \simeq 2.10^{-4}$ larger than previously expected. This would suggest a larger deuterium burning in the Galaxy and would tend to a smaller Ω value.

References

- Charbonnel, C., Vauclair, S., Zahn, J.-P., 1992, A&A 255, 191
 Charbonneau, P., Michaud, G. 1988, ApJ 334, 746
 Coplan, M.A., Ogilvie, K.W., Bochsner, P., Geiss, J., 1984, Solar Phys. 93, 415
 Dearborn, D.S.P., Schramm, D.N., Hobbs, L.M., 1992
 Deliyannis, C.P., Demarque, P., Kawaler, S.D., 1990, ApJS, 73, 21
 Garcia Lopez, R.J., Spruit, H.C., 1991, ApJ, 377, 268
 Geiss, J., Reeves, H., 1972, A&A 18, 126
 Geiss, J. 1992, in "Origin and Evolution of the Elements", Cambridge University Press, p. 89
 Lemoine, M., Vidal-Madjar, A., Ferlet, R., Bertin, P., Gry, C., Lallement, R., 1994, preprint
 Linsky, J.L., Brown, A., Gayley, K., Diplas, A., Savage, B.D., Ayres, T.R., Landsman, W. Shore, S.N., Heap, S.R., 1992, preprint
 Michaud, G. 1986, Ap. J. 302, 650.
 Michaud, G., Fontaine, G., Beaudet, G. 1984, ApJ, 282, 206
 Pagel, B.E.J., Simonson, E.A., Terlevich, R.J., Edmunds, M.G., 1992, MNRAS 255, 325
 Pinsonneault, M.H., Kawaler, S.D., Demarque, P., 1990 ApJS, 74, 501
 Pinsonneault, M.H., Deliyannis, C.P., Demarque, P., 1992, ApJS, 78, 179
 Prantzos, N., Cassé, M., Viangoni-Flam, E., 1993, ApJ 403, 630
 Proffitt, C.R., Michaud, G., 1991, ApJ, 371, 584
 Reeves, H., Fowler, W., Hoyle, F., 1970, Nature 226, 727
 Reeves, H., 1994, this conference
 Skillman, E.D., Terlevich, R.J., Terlevich, E., Kennicutt, R.C., Garnett, D.R. 1993, in "Texas/Pascos 92", Annals of the New York Academy of Sciences, vol. 688 p. 739
 Skillman, E.D., Terlevich, R.J., Kennicutt, R.C., Garnett, D.R., Terlevich, E., 1994, preprint
 Spite, F., Spite, M., 1982, A&A, 115, 357
 Vauclair, S., 1988, ApJ, 335, 971
 Vidal-Madjar, A., 1991, Adv. Space Res. 11, 97
 Vidal-Madjar, A., 1993, private communication
 Walker, T.P., Steigman, G., Schramm, D.N., Olive, K.A., Kang, H.S., 1991, ApJ, 376, 51



DARK MATTER FROM GALAXY SURVEYS

SOPHIE MAUROGORDATO

Centre National de la Recherche Scientifique UA 173
Département d'Astrophysique Extragalactique et de Cosmologie
Observatoire de Paris - Meudon, 92195 Meudon, France

Abstract

We briefly review the implications on dark matter from galaxy surveys. The amount and nature of dark matter play a fundamental role in the formation of large-scale structures. Once the primordial density fluctuations and the cosmological model are specified, the distribution of mass can be predicted. It can then be confronted to the galaxy distribution from 2D and 3D galaxy surveys, assuming a simple relation between the mass and the galaxy distribution. We will present a general view of the 'archetypes' of recent galaxy surveys and of the statistical formalism developed to quantify the analysis of the galaxy distribution and to compare to theoretical predictions at the light of COBE results. A second very promising approach will be addressed which allows to recover the mass density field from the measurement of peculiar velocities without any hypothesis on the relation between mass and galaxy distribution. We will show that presently available results are not still conclusive about the (baryonic or non-baryonic) dark matter components of models and about the value of Ω . However the size and depth of surveys is evolving so rapidly that one should hope to obtain an answer in the next future.

1 Introduction

The history of large-scale clustering in the Universe has undergone a rapid evolution in the last decades. Several groups have focused on the more salient high density structures evidenced from the angular distribution on the sky introducing the concept of supercluster as for instance the Coma supercluster (Gregory and Thomson 1978), Hydra-Centaurus (Chincarini and Rood 1979), Hercules (Tarengi et al. 1979), Pisces-Perseus (Gregory et al. 1981), or the Local supercluster (Yahil et al.1980), and verifying their reality by intensive redshift measurements. Underdense regions have been discovered too, the most spectacular one being the 'Bootes Void' (Kirschner et al. 1981, Kirschner et al. 1987) which spreads over a diameter of $6000 km s^{-1}$. On the other side, the statistical analysis of the projected galaxy distribution under the care of Peebles began to give substantial results (Hauser and Peebles 1973, Davis and Peebles 1977, Fry and Peebles 1978). The need of representative tridimensional samples became urgent and led to the completion of flux-limited redshift catalogs which should allow a real quantitative analysis of the frequency and size of structures in the Universe. This was realized thanks to the Center for Astrophysics redshift survey. While the wide-angle surveys give a detailed vision of the structure of the nearby Universe, the 'pencil-beam' surveys, complete to much fainter magnitudes on a very small region on the sky, allow to probe very deep regions of the Universe. The recent development of distance indicators independent of redshift gives access to the velocity and to the dynamical density field. The statistical indicators computed from the different surveys are then confronted to models of galaxy formation, allowing to set constraints on the nature of dark matter ingredients, on the value of Ω and on the 'biasing' mechanism.

2 Wide-angle 3D catalogs

The wide-angle, flux-limited redshift surveys have been determinant to analyse the statistical properties of the galaxy distribution. The CfA redshift survey (Huchra et al. 1983) allowed to make the first jump forward in the measurement of spatial statistical indicators. The galaxies of magnitude $m_{pg} \leq 14.5$ have been selected from a mixing of the Nilson and of the Zwicky catalogs. The sample, mostly in the Northern Hemisphere, is divided in two parts of 1845 and 556 galaxies centered respectively on the Northern (1.83 sr) and Southern galactic caps (0.83 sr), and satisfying the following criteria: $b_{II} \geq 40^\circ$ and $\delta \geq 0^\circ$, $b_{II} \leq -30^\circ$ and $\delta \geq -2.5^\circ$.

The Southern Sky Redshift Survey (da Costa et al. 1989) has been completed in the Southern Hemisphere, providing 2028 galaxies centered on the Southern galactic cap: $b_{II} \leq -30^\circ$, $\delta \leq -17.5^\circ$. The galaxies have been selected from the ESO catalog with a limit on the face-on diameter $\Phi(0) \geq 1'$. From CfA and later from SSRS, the first real statistical analysis of the spatial galaxy distribution of galaxies in our Universe up to distances of $100h^{-1}Mpc$ has been performed.

In a second phase, the CfA team has been carrying on the extension of the redshift catalog up to magnitude 15.5. For optimization of both observational time and scientific analysis, the extension is completed by strips 6° wide in declination. The first strip (de Lapparent et al. 1986) brings a new vision of our Universe until $150h^{-1}Mpc$: bubbles of diameter typically $25h^{-1}Mpc$ are surrounded by overdense thin sheets $5h^{-1}Mpc$ thick. The completion of 8 slices (5800 galaxies, Geller and Huchra, 1989) reveals the existence of a very extended sheet-like structure spreading over: $60h^{-1}Mpc \times 170h^{-1}Mpc$: The Great Wall, and of empty regions reaching such diameter as $50h^{-1}Mpc$. Once finished, the full survey will include the redshifts of more than 15000 galaxies with $b_{II} \geq 30^\circ$ and $m_{B(0)} \leq 15.5$.

On the other side, the extension of the SSRS (4 slices of 10° wide in declination) consists in 3600 galaxies brighter than 15.5 covering over 1 steradian of the southern galactic cap in the opposite direction of the CfA2 north survey and adjacent to the CfA2 south survey.

Together, CfA2 and SSRS2 cover one third of the sky and provide a panoramic view of the nearby Universe sampling different structures (Fig. 1). The similarity of the two samples is striking. They both present voids reaching sizes as large as $50h^{-1}Mpc$, and a high-density "wall" (da Costa, Geller et al. 1994) which generates large-scale fluctuations of the density field up to shells of $100h^{-1}Mpc$. Large-scale inhomogeneities do exist up to sizes nearly comparable to the size of the surveys, and the fair sample may not be yet reached.

The general philosophy of the first generation of wide-angle catalogs has been to complete in redshift already existing angular catalogs as the Zwicky or ESO catalogs. To fainter magnitudes, the 2D mapping has to be done. The first alternative is to proceed to scans of Schmidt plates combined with an automatic software of star/galaxy separation. This technique has been used for the APM galaxy survey (Maddox et al. 1990) and for the Edinburgh/Durham Southern Galaxy Catalogue (Collins et al. 1988).

The APM galaxy survey covers 4300 square degrees of the sky and gives the angular positions

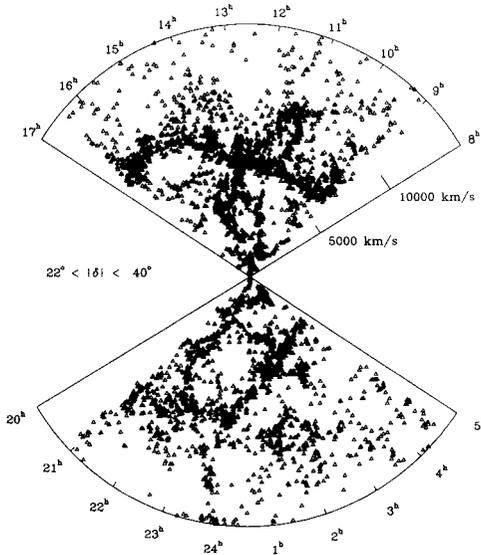


Figure 1: By courtesy of L.da Costa, M.J. Geller, J.P. Huchra, 1994. This cone diagram shows galaxies of CfA2 and SSRS2 with $22^\circ < |\delta| < 40^\circ$ and $cz \leq 15000 \text{ km/s}$. Note the Northern Great Wall, the Southern Wall and the presence of large voids.

and magnitudes of about 2 millions of galaxies with a limiting magnitude $b_j = 20.5$ in the area delimited by $\delta \leq -20^\circ$ and $b_{II} \leq -40^\circ$. The Stromlo-APM redshift survey (Loveday et al. 1992) consists in the spectroscopic follow-up of a bright sub-sample of galaxies of magnitudes $b_j \leq 17.5$ selected randomly at a rate of 1 in 20. The EDSGC provides about 10^6 galaxies up to a limiting magnitude of $b_j = 20$.

Several redshift surveys have been completed as the Durham/AAT and the Durham/SAAO (Metcalf et al. 1988). The ESO Slice Project, an ESO Key program (Vettolani et al. 1992), is based too on the EDSGC. It aims to measure the redshifts of about 3000 galaxies in a slice of $1 \times 22^\circ$ up to magnitude $b_j = 19.4$ in the South Galactic Pole region. The first slice is now complete, and shows a very dense structure at $\sim 300 h^{-1} \text{ Mpc}$ spreading over nearly all the range in right ascension surrounded by large voids.

The second strategy to get a 2D catalog is to map the survey area with CCD frames, leading to very high astrometric and photometric precision, but it requires a considerable effort in view of the amount of data and the complexity of data reduction. This method has been chosen for the Las Campanas Deep Redshift Survey (Oemler et al. 1992), a digital and spectroscopic

survey of ~ 20000 galaxies up to R magnitude 18 ($\sim B = 20$) undertaken on 600 squares degrees on the sky split in "bricks" of $1.5 \times 3^\circ$ which once all completed should form contiguous strips of 1.5° wide in declination. The visual impression resulting from the cone diagrams with 12000 galaxies is the recurrence of CfA-like structures, with maximum void sizes of the order of $50h^{-1}$ Mpc.

3 Infrared and radio surveys

3.1 The IRAS catalog

Complementary to the optical approach, surveys at the other wavelengths have undergone a rapid rise. In particular, the catalog of galaxies detected by the InfraRed Astronomical Satellite has been used to complete flux-limited redshift surveys which have the advantages of a nearly all-sky coverage with an uniform calibration, avoiding (mostly) the problems of the galactic plane extinction. Below $|b_{II}| = 30^\circ$, optical surveys are affected by confusion and obscuration due to the galactic plane, and statistical analysis becomes difficult. However, surveys of galaxies focused on the galactic plane (Dressler et al. 1991, Kraan-Korteweg et al. 1992) show the importance of clustering through the 'zone of avoidance'.

From IRAS was then completed the '2 Jy' redshift survey (Strauss et al. 1990), which consists of 2600 galaxies with a flux limit at wavelength $60\mu m$, $S_{60} \geq 2Jy$. Going fainter in flux led to the '1.2 Jy' survey (Fisher et al. 1992) of 5313 galaxies and to the QDOT sparse sample (rate 1 in 6) of 2163 galaxies up to 0.6 Jy, which is currently being fully completed (Saunders et al. 1994) These surveys, which provides an unique combination of sky coverage and depth (74% of the sky, down to galactic latitude $|b_{II}| = 10^\circ$) are particularly adapted to derive the density field of the IRAS galaxy distribution. This allows to derive the galaxy motions. Moreover, the confrontation of the properties of the optical/IRAS distribution is particularly fruitful for determining if and how galaxies trace the mass and the values of the combination $b\Omega^{-0.6}$.

3.2 The Pisces Perseus Supercluster Survey

This redshift sample has been assembled by Giovanelli and Haynes during the last decade, using the neutral hydrogen 21cm line to measure the redshifts of galaxies (PPSS, Giovanelli and Haynes 1993). The boundaries of the sample are: $b \leq -10^\circ$ and $\delta \geq 0^\circ$ leading to ~ 6000 galaxies with $m_{pg} \leq 15.7$. The most salient feature is the main ridge of the Perseus-

Pisces Supercluster, spreading across the sample along more than $80h^{-1}Mpc$, which is a severe challenge for models in terms of formation times.

4 Deep surveys

A complementary strategy to wide-angle 3D surveys which map the 'nearby' Universe is the 'pencil beam' survey approach. Broadhurst et al. (1990) combined two fields with $20'$ of diameter centered on the Northern and Southern galactic poles, complete to a limiting magnitude $b_J = 21.5$, in order to test the distribution on scales of $\sim 2000h^{-1}Mpc$. From their analysis, clustering is still present at separations larger than $100h^{-1}Mpc$. They detected regular peaks in the redshift distribution with an unexpected periodicity of $128h^{-1}Mpc$. The two nearby peaks have recently been shown to coincide respectively with the Northern Great Wall and with the Southern Wall (da Costa et al. 1994). Several deep surveys are on-going (Koo et al. in progress) and will test the reproducibility of this periodicity on other areas of the sky.

The ESO Key Programme: 'A Redshift Survey of galaxies with $z \leq 0.6$ ' (de Lapparent et al. 1989) aims to cover $0.4deg^2$ up to magnitude $R \leq 20.5$ at high southern galactic latitude, reaching an effective depth of about $1400h^{-1}Mpc$. This survey should go about 15 times deeper than CfA and should provide the diameter and frequency of distribution of the 'shells' evidenced in wide-angle nearby surveys.

The Canada-France Deep Redshift Survey (Crampton, Hammer, Le Fèvre, Lilly 1994, Tresse et al. 1993) covers several $10'$ by $10'$ fields up to magnitude $I \leq 22.1$ ($\sim B \leq 23.5$) providing a sample complete up to $z \sim 1$. Another important implication of the deep 'pencil beam' surveys is the spectrophotometric evolution of field galaxies (Coless et al. 1990, LDSS).

5 Statistical indicators

On flux-limited 2D and 3D catalogs, various statistical tools are used to characterize the density distribution. As a first approximation, the mass and galaxy fluctuations are often assumed to be related by a linear relation: $\delta\rho_g/\rho_g = b\delta\rho_M/\rho_M$, (linear biasing). The amplitude of large-scale fluctuations is now fixed by COBE, as well as a range for the index n of the spectrum of initial fluctuations (Smoot et al. 1992). Confrontation of the statistics from the data to the predictions of the models can be done at the light of COBE's results, setting severe limits on the value of b and of Ω .

5.1 The two-point correlation function

The two-point correlation function has been the most popular statistical indicator since it was introduced by Peebles in the 70's to analyze the galaxy distribution. The spatial two-point correlation function $\xi(r)$ is directly connected to the power-spectrum by a Fourier transform, and so is very useful to confront the theoretical predictions from models of galaxy formation to data. From the 2D catalogs, the projected angular two-point correlation function $w(\theta)$ is measured, which scales with the depth of the survey as $w(\theta) = D^{-1}W(\theta D)$ where W is a function only of the shape of the catalog, so a comparison of $w(\theta)$ for 2D catalogs with different flux limit is possible. The availability of flux-limited redshift catalogs has allowed to estimate the spatial correlation function. When the scales involved are sufficient to ignore the effect of random peculiar velocities (above $\sim 2h^{-1}Mpc$), the real space correlation function $\xi(r)$ is approximated by the direction-averaged redshift correlation function $\xi(s)$ with: $s = \frac{(V_1^2 + V_2^2 - 2V_1 V_2 \cos\theta_{12})^{1/2}}{H_0}$ which can be directly derived from the 3D redshift catalogs.

The two-point correlation function is commonly estimated by weighted counts of the galaxy-galaxy pairs $DD(s)$ and of the galaxy-random pairs $DR(s)$ in a random catalog with the same geometry as the galaxy one. $DD(s) = \sum_i \sum_j w_i w_j N_{DD}$ and $DR(s) = \sum_i \sum_j w_i w_j N_{DR}$.

Then, $1 + \xi(s) = \frac{DD(s)}{DR(s)} \frac{n_R}{n_D}$. This method has the great advantage of accounting directly for edge effects. Several weighting schemes have been advocated: the uniform weighting, the selection function weighting, and the minimum variance weighting. The uniform weighting is generally used when working on complete sub-samples limited both in distance and in absolute magnitude, so of roughly constant density. When analysing magnitude-limited samples, one has to take into account the fall-off of the density with radial distance. In this case, the uniform weighting gives too much weight to foreground galaxies. A classical weighting is to multiply the contribution of each galaxy by the inverse of the selection function (Davis and Peebles, 1980) $w_i = 1/\Phi(r)$. This however gives quite a high weight to distant pairs and increases the white noise for small separations. In order to determine the optimal weighting function, one can try to minimize the variance of the mean density (Efstathiou, 1990). This gives a weight $w_i = 1/4\pi\rho\Phi_i J_3(s)$ depending on $J_3(s)$ which is the integral of the correlation function we try to measure. One has then to assume a form for the correlation function. It seems however that this method improves the estimation of $\xi(s)$ at small scales giving less weight to the few pairs at large distance (Vogele et al.1993).

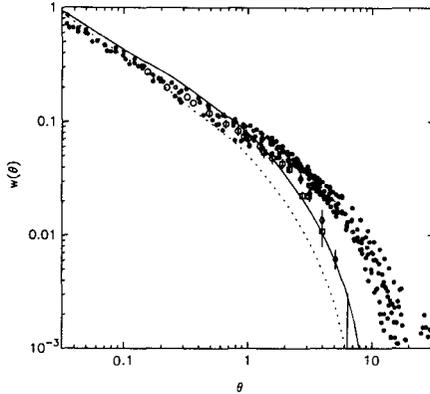


Figure 2: From Maddox et al. 1990. The angular correlation function $w(\theta)$ for magnitude limited sub-samples of APM scaled to Lick (filled symbols) compared to Groth and Peebles (1977) estimates on the Lick catalogue (open symbols), and to CDM models with $h=0.5$ (dotted line) and $h=0.4$ (solid line).

The correlation function is however very difficult to measure correctly at large scales where fluctuations are small compared to the mean, and where it is limited by the uncertainty on the mean density.

The first estimation of $w(\theta)$ was performed on the Lick catalog of counts by Groth and Peebles in 1977, showing a clear power law behaviour $w(\theta) \propto \theta^{1-\gamma}$ with $\gamma = 1.77$. On the CfA data, Davis and Peebles (1983), fixing the slope to $\gamma = 1.77$ found $\xi(r) = (r/r_0)^{-\gamma}$ with $r_0 = 5.4 \pm 0.3 h^{-1} Mpc$. The values of the slope and of the correlation length have not changed dramatically during the last ten years: $s_0 = 7.5 h^{-1} Mpc$ and $\gamma = -1.6$ (de Lapparent et al. 1988 on the first two slices of the CfA2), from volume-limited sub-samples of the CfA1 and SSRS1 to 40 and $80 h^{-1} Mpc$, for $\gamma = -1.8$, r_0 is equal to 6.0 and 8.0 (CfA1) and 6.4 and 7.5 (SSRS1) $h^{-1} Mpc$.

The APM catalog, providing the angular positions of about 2 million galaxies, has allowed to improve the precision on $w(\theta)$ and thanks to its depth to measure it up to large angular separations, up to $\theta = 20^\circ$ (Fig. 2). The power law is confirmed up to 3° with a slope $\gamma = 1.668$. Above 3° , a break is apparent in $w(\theta)$. At large scales, the data show an excess of power which cannot be reconciled with the standard CDM model. A recent analysis of the 3D Stromlo-APM redshift survey gives a slightly shallower slope $\gamma = 1.34 \pm .08$ and $s_0 = 5.7 \pm 0.8 h^{-1} Mpc$.

A derived statistic is the variance, the r.m.s. fluctuation in cells. The variance of galaxy

counts reaches unity at $\sim 8h^{-1}Mpc$ (Loveday et al. 1992), so in the context of linear bias, the variance of mass fluctuations in $8h^{-1}Mpc$ boxes is directly connected to the bias factor b_8 by $\sigma_8 = 1/b_8$.

5.2 Clustering in redshift space and in real space

The existence of peculiar velocities induces some distortion between the redshift and the real space distribution.

At small scales ($s \leq 1h^{-1}Mpc$) $\xi(s)$ and $\xi(r)$ can be very different and one has to compute the components of the correlation function parallel and perpendicular to the line of sight. The distribution of peculiar velocities can be approximated successfully by an exponential model: $f(V) \propto \exp(-2^{1/2})\frac{V}{\sigma}$ (Davis and Peebles 1983). The value of pairwise r.m.s. relative peculiar velocity is particularly interesting as it can be related to Ω through the cosmic virial theorem (Peebles 1980). It was estimated to be $\sigma = 340 \pm 40km/s$ at $1h^{-1}Mpc$ (Davis and Peebles 1983 on the CfA). The small value of σ at small scales is in contradiction with the predictions from the COBE normalized CDM model ($\sigma = 970 \pm 160km/s$, Bertschinger et al. 1990, Ostriker 1993). At large scales, where the coherent velocity field may be important, the redshift direction averaged correlation function is enhanced towards the real space one. In the frame of the linear perturbation theory, the multiplicative factor is a constant which does only depend on Ω (Kaiser 1987): $\xi_s(s) = (1 + \frac{2}{3}\Omega^{0.6} + \frac{1}{5}\Omega^{1.2})\xi_r(s)$. The anisotropy of the clustering pattern at large scales should then yield an independent way of determining Ω .

5.3 The Power Spectrum

The clustering at large scales is particularly interesting to quantify as this regime should be linked directly to initial conditions through the perturbation theory, without need to call for non-linear evolution. We noted previously that at these scales precisely where $\xi(r)$ is small, its measurement is severely affected by local fluctuations of the mean density. In order to extract the statistical information on large scales, 3D statistics can be calculated in the Fourier space as the Fourier components belonging to different wave vectors are statistically independent. The measurement of the power spectrum at one scale should then be independent of the measurement at another scale, while the correlations on different scales are highly correlated. The power spectrum of the mass fluctuations today can be directly related to the angular power spectrum of the temperature fluctuations. Assuming the Harrison-Zeldovich form for the power spectrum

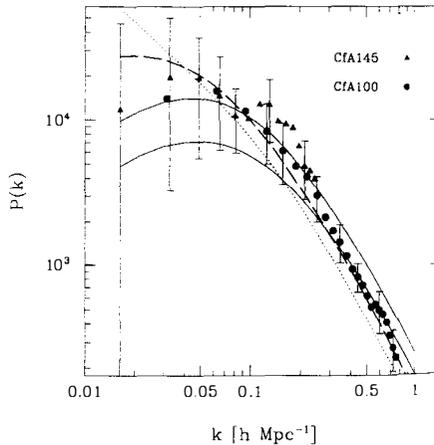


Figure 3: From Vogeley et al. 1992. The power spectrum for galaxy samples extracted from CfA2 (limited respectively to 100 and $145h^{-1}$ Mpc) compared to the predictions from models. From top to bottom: open $\Omega h = 0.2$ unbiased CDM (dashed line), standard $\Omega h = 0.5$ CDM with $\sigma_8 = 1.4$ and $\sigma_8 = 1$.

of initial fluctuations $P(k) = Ak^n$ with $n=1$ (in agreement with the first estimate from COBE $n = 1 \pm 0.6$, Smoot et al. 1992), the amplitude A is constrained by the RMS sky variation (Efstathiou et al. 1992): $Q_{rms}/T_0 \propto H_0^2 \Omega_0^{0.77} A^{1/2}$. So COBE fixes the amplitude of the power spectrum at large scales (corresponding to wavenumbers $k \leq 0.003 \Omega_0^{0.01} h Mpc^{-1}$). From galaxy redshift surveys, the power spectrum is measured significantly for $k \geq 0.02 h Mpc^{-1}$, so there is nearly a factor 10 in scale between the higher scale tested and the scales involved by COBE. In order to link both measurements, one needs to extrapolate the power spectrum assuming a form dependent on the model. In the standard CDM frame, the normalization required to fit COBE data is $\sigma_8 = 1.1 \pm 0.2$ (Efstathiou et al. 1992), implying a low bias $b_8 \sim 0.9$.

The power spectrum has been widely measured on galaxy catalogs. Baumgart and Fry (1991) on the CfA1 and on PPSS, Kaiser (1991) on the IRAS QDOT, Peacock and Nicholson (1991) on a sample of radio galaxies, Park et al. (1992) on the SSRS and CfA1. Vogeley et al. 1992 summarise the successive estimates and derives its value on the CfA2 (Fig. 3).

This last analysis shows that the $\Omega h = 0.5$ CDM model fails in reproducing the shape of the power spectrum (Fig.3). This is independent of the normalization which can be set to fit or the small-scales, or the large-scales, but never both together, as the standard CDM power

spectrum shows too high a ratio of small scale over large scale power in confront to the data. Up to now, the open model with $\Omega h = 0.2$ and non-zero cosmological constant seems to fit the data, with all the critics that can be done to the 'fine-tuning' of Λ .

5.4 High order indicators

In order to describe the density distribution, the N -point correlation functions should be known up to high values of N . The first orders (up to $N=4$) can be calculated directly and show the existence of a 'hierarchical' relation, where the N -point correlation function is a sum of $N - 1$ products of 2-point correlation functions. Testing the existence of such a relation is very interesting as it reflects directly the type of initial fluctuations. For instance, in the case of non-Gaussian initial fluctuations, terms of higher order do appear. The information on high-order correlation functions can be reached too calculating the moments (Bouchet et al. 1993 on IRAS), or through other indicators, like the Void Probability Function which involves indirectly correlations of all orders (Maurogordato and Lachièze-Rey 1987 on CfA1, Vogeley et al. 1991 on CfA2).

6 The velocity field

In parallel to the acquisition of redshift samples, the independent measurement of distances on still larger areas on the sky has strongly increased in the last five years. Empirical tight relations have been determined as the Tully-Fisher (for spiral galaxies) or the $D_n - \sigma$ for ellipticals and S0's (Dressler et al. 1988) and allowing to determine relative distance of the objects from the ratio between apparent and absolute luminosities. The main difficulty comes from the Malmquist bias which can affect seriously the extragalactic distance scale (Bottinelli et al. 1986). The radial peculiar velocities can then be deduced as : $v_{pec}^r = cz - d$. From the pioneering samples of Burstein et al. (1987), a considerable effort has been endeavoured on the measurement of peculiar velocities, leading today to the compilation of more than 3000 galaxies (Faber et al. 1993). Large scale velocity gradients have been shown to exist, converging to very high density concentrations as the so-called 'Great Attractor' (Lyndell-Bell et al. 1988). From these data, through the POTENT method (Bertschinger and Dekel 1989), the dynamical field of velocity and mass-density fluctuations has been recovered (for a review see Dekel 1994). The main stages are to smooth the radial velocity field, to reconstruct the velocity field assuming a

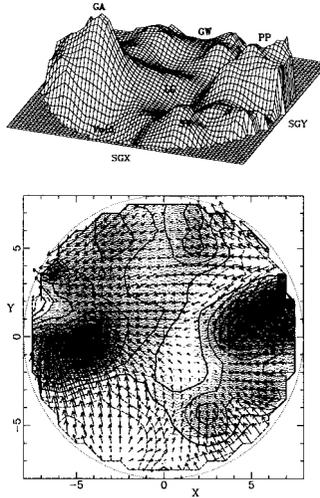


Figure 4: By courtesy of A.Dekel, 1994, Annual Review of Astronomy and Astrophysics. The Potent reconstructed mass density and velocity fluctuation fields. The Local Group is at the center. Large-scale flows (and overdensities) are apparent as Great Attractor (left), Pisces-Perseus (right), Cetus Southern Wall (bottom), and the Great Wall (top).

potential flow so the velocity potential can be calculated by integrating the radial component of the velocity along the line of sight, and the two other components obtained by differentiation. One can then derive the density field under the Zeldovich approximation. The mass density field recovered through the peculiar velocities can then be compared to the galaxy density field. The optical density maps (Hudson and Dekel 1993) and the IRAS 1.2 Jy density ones (Strauss et al. 1993) are strikingly similar to the potent mass density (Dekel et al.1992), showing the main high-density features as Pisces-Perseus, the Great Attractor, Cetus Southern wall.... This similarity makes sense to assume a linear bias between optical galaxies, IRAS galaxies and mass, $\delta_{opt} = b_{opt}\delta$ and $\delta_{IRAS} = b_{IRAS}\delta$. In the linear approximation where $\delta_{Potent} \propto \Omega^{0.6}\delta$, the confrontation of these density fields sets constraints on both optical and IRAS bias and on Ω_0 . From Strauss et al. 1993, $\Omega^{0.6}/b_{IRAS} = 1.28 \pm_{0.59}^{0.75}$ and from Hudson and Dekel 1993, $\Omega^{0.6}/b_{IRAS} = 0.5-0.7$. These results are fully consistent with an $\Omega = 1$, $b_{opt} = 1.5$, $b_{IRAS} = 1$ cosmology.

7 Discussion and conclusion

The statistical analysis of galaxy surveys has provided quantified information which combined to COBE results allows to test models of formation of structures, varying the dark matter

components (Cold Dark Matter, Hot Dark Matter, Mixed Dark Matter), the type of initial fluctuations (for instance through the index of the primordial spectrum), and the parameter Ω . From wide-angle redshift surveys, the shape and amplitude of the correlation functions, the low value of pair-wise dispersions argues against an $\Omega = 1$ CDM standard Universe and rather to an open $\Omega \sim 0.2, \Lambda \neq 0$ Universe, while the large scale dynamics from peculiar velocities strongly rules out $\Omega \leq 0.3$ and strongly favours high values of Ω . So, a consensus is still not reached. On the other side, very strong constraints come from galaxy clusters: both from the abundance of rich clusters and from the cluster-cluster correlation function (Bahcall and Cen 1992). Some variations to the standard CDM are developed to match the whole set of constraints, as lowering the index $n \sim 0.8$ ('tilted' spectrum), or adding hot neutrinos keeping $\Omega = 1$ (Mixed Dark Matter, Davis et al. 1992), or introducing non-gaussianness in the initial fluctuations (as 'texture' models).

In future, the measurement of the power-spectrum at very large scales, on samples big enough to minimize the density fluctuations (as the Century Survey, Geller et al. in progress, or the Digital Survey Gunn and Knapp in progress), will be determinant.

The progress of full sky catalogs in other wavelengths, like 2-mass and Denis at $2\mu m$, combined to optical redshift surveys and to the measurement of peculiar velocities will have great insights on the determination of the density field. Recent results (Lauer and Postman 1994) from a full-sky velocity survey of Abell clusters of galaxies within 15000 km/s, show that the Abell cluster frame is moving at $689 \pm 178 km/s$ in respect to the CMB dipole, implying the existence of high-density mass concentrations beyond $100h^{-1} Mpc$. On the side of CMB measurements, the ground-based experiments at smaller angular scales will, combined to COBE, constrain more the index of the primordial spectrum.

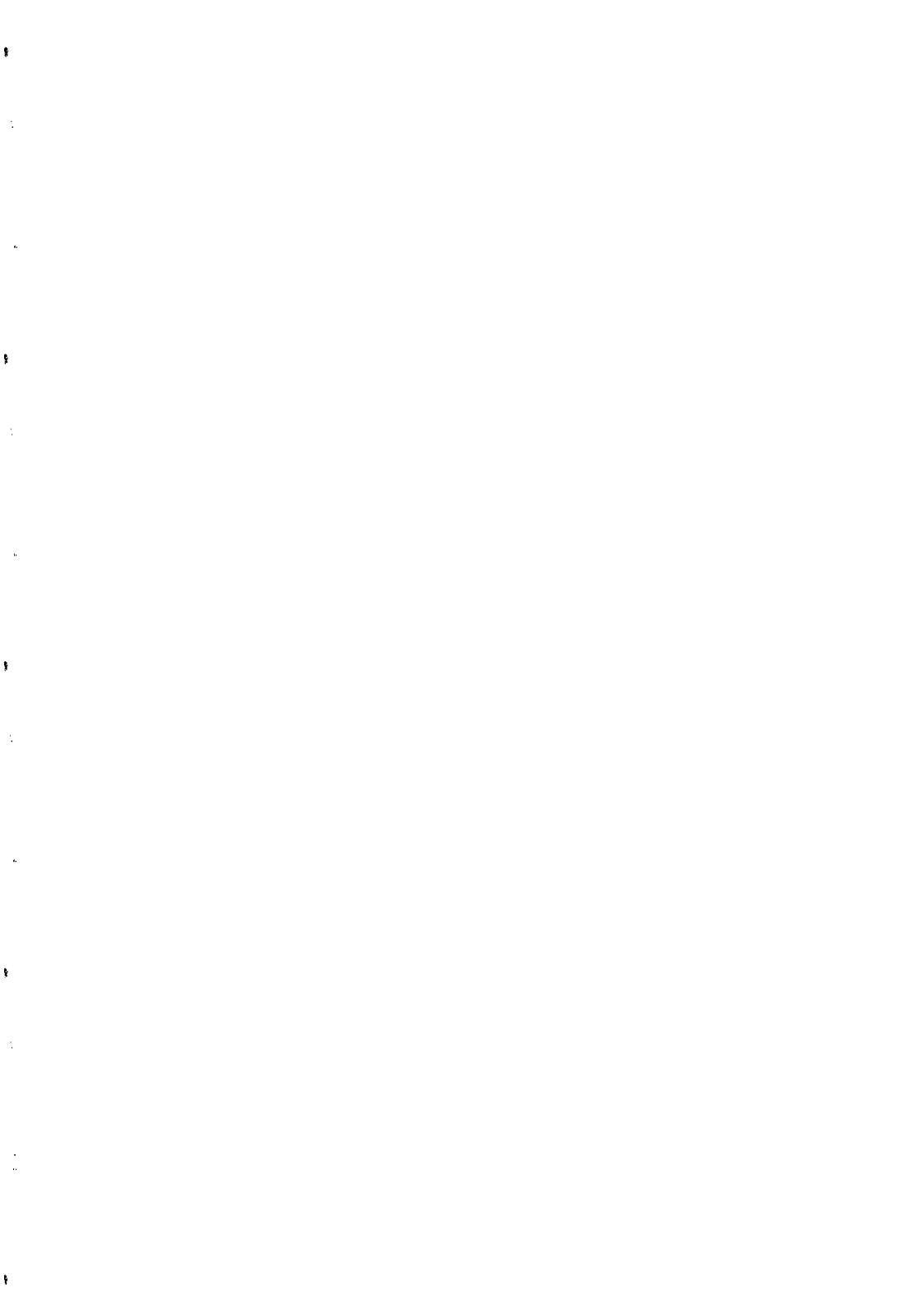
8 Acknowledgments

I would like to thank all the authors who helped me a lot for the graphic task of this review, providing me postscript files or allowing me to use their graphs: Luiz da Costa, Margaret Geller and John Huchra for providing me the cone diagram of CFA2 and SSRS2, Avishai Dekel for the updated figure of the density and velocity maps, Stefan Maddox and Michael Vogeley. I am grateful to Chantal Balkowski and to Alberto Cappi for reading this manuscript.

References

- [1] Bahcall, N., and Cen, R. 1992, Ap.J., 398, L81
- [2] Baumgart, D.J. and Fry, J.N. 1991, Ap.J., 375, 25
- [3] Bertschinger, E. and Dekel, A., 1989, Ap.J.(Letters), 336, L5
- [4] Bottinelli, L., Gouguenheim, L., Paturel, G., and Teerikorpi, P., 1987, Astron.Astrophys, 181,1
- [5] Bouchet, F.R., Strauss, M., Davis, M., Fisher, K., Yahil, A., Huchra, J.P., 1993, Ap.J., 417, 36
- [6] Broadhurst, T.J., Ellis, R.S., Koo, D.C., and Szalay, A.S., 1990, Nature, 343, 726
- [7] Burstein, D., Davies, R.L., Dressler, A., Faber, S.M., Lyndell-Bell, D.,
- [8] Wegner, G., and Terlevich, R. 1987, Ap.J.Suppl., 64, 601 da Costa, L.N., Pellegrini, P.S., Davis, M., Meiksin, A., Sargent, W.L.W., Tonry, J.L., 1991, Ap.J.(Suppts), 75, 935
- [9] da Costa, L.N., Geller, M.J., Pellegrini, P.S., Latham, D.W., Fairall, A.P., Marzke, R.O., Willmer, C.N.A., Huchra, J.P., Calderon, J.H., Ramella, M., and Kurtz, M.J., 1994, Ap.J., 424, L1
- [10] Chincarini, G. and Rood, H.J., 1979, Ap.J., 230, 648
- [11] Colless, M., Ellis, R.S., Taylor, K., Hook, R.N., 1990, MNRAS, 244, 408
- [12] Collins, A.A., Heydon-Dumbleton, N.H., MacGillivray, H.T., MNRAS, 236, 7P
- [13] Crampton, Hammer, Le Fevre, Lilly 1994
- [14] Davis, M. and Peebles, P.J.E., Ap.J., 1977, Ap.J.Suppl., 34, 425
- [15] Davis, M. and Peebles, P.J.E., Ap.J., 1983, 267
- [16] Davis, M., Meiksin, A., Strauss, M.A., daCosta, L.N. and Yahil, A., 1988, Ap.J., 333, L9
- [17] Dekel, A., 1993, in Observational Cosmology, International Symposium held in Milano, Italy, 21-25 sept 1992, Astronomical Society of the Pacific Conference Series, volume 51
- [18] Dekel, A., 1994, Annual Review of Astronomy and Astrophysics (to be published)
- [19] Dressler, A., Faber, S., Burstein, D., Davies, R., Lundel-Bell, D., Terlevich, R.J. and Wegner, G. 1987, Ap.J., 313, 42
- [20] Efstathiou, G., Bond, J.R., and White, S.D.M., 1992, M.N.R.A.S, 258, 1P
- [21] Faber, S.M. et al. 1993, in preparation
- [22] Fisher, K.B., Davis, M., Strauss, M., Yahil, A., and Huchra, J.P., 1993, 402, 42
- [23] Fry, J.N. and Peebles, P.J.E., 1978, Ap.J.221, 19
- [24] Geller M.J. and Huchra J.P., 1989, Science, 246, 897
- [25] Geller, M.J., Kurtz, M.J., Huchra, J.P., Schild, R., Fabricant, D.G., Thorstensen, J.R., and Wegner, G., 1994 in preparation
- [26] Gregory, S.A. and Thomson, L.A., 1978, Ap.J., 222, 784
- [27] Gregory, S.A., Thompson, L.A., and Tift, W.G., 1981, Ap.J., 243, 411
- [28] Groth, E.J., and Peebles, P.J.E., 1977, Ap.J., 217, 385
- [29] Hauser, M.G. and Peebles, P.J.E., 1973, Ap.J., 185, 757
- [30] Huchra, J., D., Davis, M., Latham, D., Tonry, J. 1983 Ap. J. Suppl., 52, 89
- [31] Hudson, M. and Dekel, A., 1993 in preparation
- [32] Kaiser, N. 1987, M.N.R.A.S., 227, 1
- [33] Kaiser, N., 1991, in Texas/ESO-CERN Symposium on Relativistic Astrophysics, EDS Barrow et al., 295, New York Academic Science Publishers

- [34] Kirshner, R.P., Oemler, A. Jr., Schechter, P.L., and Schectman, S.A. 1987, Ap.J., 314, , 493
- [35] Kirshner, R.P., Oemler, A. Jr., Schechter, P.L., and Schectman, S.A. 1981, Ap.J. (Letters), 248, L57
- [36] Kraan-Korteweg, R., Fairall, T., Balkowski, C, Cayatte, V., 1992 in Observational Cosmology, International Symposium held in Milano, Italy, 21-25 sept 1992, Astronomical Society of the Pacific Conference Series, volume 51
- [37] de Lapparent, V., Geller, M.J., and Huchra, J.P., 1986, Ap.J.(Letters), 202, L1
- [38] de Lapparent, V., Geller, M.J., and Huchra, J.P., 1988, Ap.J., 332, 44
- [39] de Lapparent, V., Mazure, A., Mathez, G., and Mellier, Y., 1989, The Messenger
- [40] Lauer, T. and Postman, M., 1994, Ap.J., 425, 418
- [41] Le Fevre, O., Crampton, D., Hammer, F., Lilly, S.J. and Tresse, L., 1994, Ap.J., 423, L89
- [42] Lyndell-Bell, D., Faber, S., Burstein, D., Davies, R.L., Dressler, A., Terlevich, R.J., and Wegner, G., 1988, Ap.J., 326, 19
- [43] Loveday, J., Efstathiou, G., Peterson, B.A. and Maddox, S.J., 1992, ApJ, 400, L43
- [44] Maddox, S.J., Efstathiou, G., Sutherland, W.J. and Loveday, J., 1990, MNRAS, 242, 43P
- [45] Maurogordato, S., and Lachièze-Rey, M., 1987, Ap.J., 320, 13
- [46] Metcalfe, N., Fong, R., Shanks, T., Kilkenny, D., M.N.R.A.S., 236, 207
- [47] Oemler, A., Tucker, D., Kirshner, R., Lin, H., Schechter, S., Schechter, P., 1993 in Observational Cosmology, International Symposium held in Milano, Italy, 21-25 sept 1992, Astronomical Society of the Pacific Conference Series, volume 51
- [48] Ostriker, J. 1993, Annual Review of Astron. Astrophys.
- [49] Park, C., Gott, J.R., and da Costa, L.N., 1992, Ap.J., 392, L51
- [50] Peacock J.A. and Nicholson D., 1991, MNRAS, 253, 307
- [51] Peebles, P.J.E, 1980, The Large-Scale Structure of the Universe, Princeton University Press
- [52] Saunders et al. 1994 in preparation
- [53] Smoot, G.F et al. 1992, Ap.J., 396, L1
- [54] Strauss, M.A., Davis, M., Yahil, A., Huchra, J.P. 1990, Ap.J.361, 49
- [55] Tarengi, M., Tift, W.G., Chincarini, G., Rood, H.J., and Thompson, L.A., 1979, Ap.J., 234, 793
- [56] Tresse, L., Hammer, F., Le Fevre, O., Proust, D., A.A., 277, 53
- [57] Vogeley, M., Geller, M., Huchra, J., P., 1991, Ap.J., 383, 44
- [58] Vogeley, M., Park, C., Geller, M.J., Huchra, J.P., 1992, Ap.J., 391, L5
- [59] Vettolani et al., 1993, in Observational Cosmology, International Symposium held in Milano, Italy, 21-25 sept 1992, Astronomical Society of the Pacific Conference Series, volume 51
- [60] Yahil, A., Sandage, A., and Tammann, G.A., 1980, Ap.J., 242, 448



IS GALACTIC DARK MATTER COLD MOLECULAR GAS?

D. Pfenniger¹, F. Combes²

¹Geneva Observatory, University of Geneva, CH-1290 Sauverny, Switzerland

²DEMIRM, 61 Av. de l'Observatoire, F-75014 Paris, France

Abstract

We propose that dark matter around spiral galaxies is in the form of a cold fractal gas, essentially molecular, supported by rotation, and lying mostly in the outer galactic disks^(1),2),3). This hypothesis is suggested by several observational clues, such as 1) the rapid evolution of galaxies along the Hubble sequence from Sd to Sa shown by dynamical studies, 2) the gas consumption problem in stellar population models, 3) the increasing dark/luminous matter ratio along the Hubble sequence towards late-type galaxies, 4) the constant ratio (≈ 20) between gas and dark matter surface densities in the outer parts of galaxies together with the flat rotation curve conspiracy, 5) the excess of gas in interacting galaxies, and 6) the high frequency of Ly α absorption systems in front of remote quasars.

To understand how measurements of neutral and molecular gas might only detect a small fraction of the total gas mass, we take into account the observations of the nearby cold gas showing its fractal structure over widely different physical conditions. We build a model of fractal gas, in near isothermal equilibrium with the 3 K cosmological background. We need to extrapolate only over 2 orders of magnitude (down to 30 AU) the already known fractal interstellar medium, observed over 4 orders of magnitude from 100 pc to 0.01 pc. By numerical simulations of fractal clouds we show that the gas mass can be easily underestimated by a factor 20, due to the highly inhomogeneous structure of the fractal, and to the low temperature and metallicity of the gas. Such a fractal structure with a fractal dimension derived from Larson's relations typically displays both optically thin and thick parts, the optically thick part covering only a small fraction of the sky, yet containing still a large fraction of the mass.

The smallest scale of the fractal is determined by the transition from the isothermal to the adiabatic regime. The smallest clumps, the clumpuscles, have a mass of about $10^{-3} M_{\odot}$, a radius of about 30 AU, and a low temperature near 3 K. They are in a dynamical state, constantly colliding, merging and disrupting, which prevents the formation of stellar like objects. The clumpuscule collision time is similar to the dynamical time, about 2000 yr. Beside traces of CNO elements, the coupling with the thermal background radiation may be ensured by some fraction of the mass being in the form of H₂ ice.

1. The gas – dark matter connection

It has been known for a long time that the gas surface density in the outer parts of galaxies decreases as $1/R$, from HI 21cm measurements⁴⁾. The surface density of dark matter around spiral galaxies also decreases as $1/R$, since the rotation curves are flat. It is remarkable that the dark matter and gas surface densities are always observed in a near constant ratio in the outer parts of galaxies, this ratio being between 10 and 30^{4),5),6)}. The coexistence of gas and dark matter strongly suggests an intimate link between them.

The flat rotation curves around spiral galaxies are a puzzle for two reasons: not only they do point towards the existence of unseen mass at the outskirts of visible disks, but this external mass conspires with the internal mass to maintain the same rotational velocity. This conspiracy suggests that the external dark mass is of the same nature and shares the same evolution as the visible mass during its condensation towards the center.

The recently acquired large data base of CO emission in nearby galaxies reveals that apparently 4 to 5 times more gas in interacting and perturbed galaxies is measured than in isolated normal ones⁷⁾. This result can only be interpreted by the existence of large reservoirs of hydrogen gas in the outer parts of galaxies. During a tidal interaction between two galaxies gravitational torques drive this unseen gas towards the center where it can form stars. After the starburst triggered by the interaction has consumed a part of the fresh gas, the gas amount then returns to normal.

From UBV colours and H α equivalent widths, Kennicutt⁸⁾ has shown that the current rates of star formation in spiral galaxies is comparable to the past rates averaged over the age of the disk, i.e. galaxies have formed stars at a nearly constant rate during their lives. Since the time scale for gas consumption by star formation is of the order of 2 Gyr now, a small fraction of the galaxy life-time, a constant star formation rate is only possible with continuous infall of fresh gas in the disk, as already proposed by Larson et al.⁹⁾ The gas consumption time-scale can be extended by a factor 2 – 4 when stellar winds and gas recycling is taken properly into account¹⁰⁾, but it can also be shortened by an even larger factor if only the gas within the optical disk is considered.

Galaxy disks are dynamically fragile and react to perturbations via spiral waves and bars in a few rotational periods. The resulting dynamical evolution coupled to gas dynamics may lead to the bar destruction and bulge formation^{11),12),13)}, as demonstrated by numerical simulations. The evolution along the Hubble sequence appears then to be from late-types to earlier types with possible time-scales much shorter than the galaxy ages. HI observations^{14),6)} have shown that the dark to luminous mass ratio in early-type galaxies is less than in late-type ones. This suggests that at least 90% of the galactic dark matter is transformed into stars during evolution, and therefore most of dark matter in late-type spirals has to be hydrogen and helium in a sufficiently diluted form such as to permit star formation at later times.

Finally, the large frequency of Ly α absorption features in front of remote quasars remains unexplained. If the absorption systems are associated to galaxy disks, their effective radius should be of the order of 400 – 500 kpc¹⁵⁾. Absorbing systems with large column density are rarer, as expected from a fractal medium, but molecular hydrogen has already been observed in absorption in one of these systems¹⁶⁾.

All the above observational clues support the hypothesis that dark matter around galaxies is in the form of cold gas; then one should understand how the gas density in the outskirts of galaxies may be underestimated by a large factor (≈ 20). This is possible if the gas is

structured in a highly hierarchical way, in which the densest clumps are in cold state consisting mainly of molecular hydrogen, so like the nearby molecular clouds. H_2 is a symmetric molecule that does not radiate through rotation at radio frequencies and therefore has to be traced by the CO molecule. But the abundance of CO constantly decreases with galactic radius, conformably to the view that the CNO elements are formed by stars, and we expect that the gas is nearly primordial beyond about 30 kpc. There are however some evidences of CO emission up to 28 kpc from the Galactic center¹⁷⁾, with a much higher H_2/CO conversion ratio which support the existence of large amounts of molecular hydrogen in the outer galaxy.

Any form of gaseous hydrogen in the Galaxy is also traced by γ -rays, which result from the interaction of cosmic rays (CR) and matter. But the potential source of CR are supernovae, and more generally stars, which have a radial scale of 3 – 5 kpc only. Yet the γ -rays radial scale is much larger¹⁸⁾ (15 kpc), and may result from an extended gas component more massive than expected.

2. Fractal state of cold gas

It has been known since several years that the nearby cold interstellar medium has a fractal structure over 4 orders of magnitude in scale^{19),20)}, i.e. from giant molecular clouds (100 pc) to Bok globules or tiny clumps (0.01 – 0.1 pc). The latter is the smallest scale reachable by current mapping techniques in nearby clouds. But VLBI in HI absorption²¹⁾, or extreme scattering events due to compact objects moving in front of remote radio-loud quasars²²⁾ have revealed structures at scales of the order of only 20 AU. Fiedler et al.²²⁾ deduce from the scattering amplitude of the occulting objects an electron density of the order of 4000 cm^{-3} , corresponding to a total ionised mass per object of $\approx 10^{-10} M_\odot$. However, in dense clouds the typical electron abundance is 10^{-7} , which leads to total masses of $10^{-3} M_\odot$, just enough to virialise the objects at low temperature. From their statistics of scattering events, Fiedler et al. estimate that the number density of their compact structures in the solar neighbourhood is 10^3 times the star one. These compact objects can then account for most of or all the unseen mass in the solar neighbourhood²³⁾.

A natural hypothesis is therefore to extrapolate the fractal structure of the interstellar medium by two orders of magnitude down to 20 AU in scale, i.e. the self-similarity relation between mass and size $M \sim r^D$, where D is the fractal dimension. With the hypothesis of virial equilibrium at each scale, the velocity dispersion is $v \sim r^{(D-1)/2}$, already verified for the atomic and molecular clouds^{24),25)}. The best match to the broad variety of observations is $D = 1.5 - 2$.

What is the cause of the fractal structure? Any amount of gravitating gas in quasi-isothermal equilibrium in a thermal bath, provided it has no excess of kinetic energy to expand to infinity, tends ideally to follow an isothermal distribution, i.e. a sphere where the density decreases as r^{-2} . However, the gravothermal catastrophe occurs as soon as the density contrast between the center and the edge reaches 32.1^{26),27)}. At this stage, the sphere is unstable and fragments in $N \approx 10$ clumps, while the outer parts evaporate, gaining energy from the background bath. This instability recurs at the next lower scale, as long as the gas remains isothermal, and a hierarchical self-similar fractal structure is automatically built in. The recursion ends down at the scale where the isothermal character breaks down.

This happens²⁸⁾ when the dynamical time (or free-fall time) $\tau_{\text{F}} = 1/\sqrt{G\rho}$ becomes shorter than the Kelvin-Helmholtz time τ_{KH} ,

$$\tau_{\text{KH}} = \frac{3 M k T}{2 \mu m_{\text{p}}} / L, \quad \text{with} \quad L = 4 \pi f R^2 \sigma T^4, \quad (1)$$

where μ is the mean molecular weight, m_p the proton mass, f is a factor of the order of one, or smaller, depending on the departure from spherical geometry and from black-body radiation, and σ is the Stefan-Boltzmann constant. Along the hierarchy of scales of the isothermal fractal, the ratio τ_{KH}/τ_{ff} varies as $r^{(3D-7)/2}$, i.e. decreases with r since D is inferred from observations to be lower than 2.33. The time-scale for cooling is short at large-scale, which ensures the isothermal character. And there exists a bottom scale where the adiabatic regime is reached when $\tau_{KH} = \tau_{ff}$. Together with the virial relation $M = 3kTR/G\mu m_p$, this gives the mass of the smallest scale fragments, that we name “clumpuscles”, i.e.,

$$M_* \approx 10^{-3} \left(\frac{T}{3\text{K}} \right)^{1/4} \left(\frac{\mu}{2.3} \right)^{-9/4} \left(\frac{f}{0.5} \right)^{-1/2} [M_\odot]. \quad (2)$$

The corresponding average size, volumic and surface densities are 30 AU, $3 \cdot 10^9 \text{ cm}^{-3}$, and $2 \cdot 10^{24} \text{ cm}^{-2}$ respectively, and $\tau_{KH} = \tau_{ff} \approx 2000 \text{ yr}$. This remarkably fits with the smallest observed scales^{21),22)}. It is also remarkable that these clumpuscles extend to smaller scales the Larson relations for the known interstellar medium, for a fractal dimension $D = 1.7$, and a velocity-size relation of $v \sim r^{0.35}$, perfectly compatible with the observations. The line of sight velocity dispersion of a single clumpuscle is thermal, of the order of 0.1 km s^{-1} .

At low temperature the condensed character of the fractal makes the collision time-scale between clumpuscles of the order of their dynamical time with $D \approx 1.7$ and $N \approx 10$ suggested by observations. This prevents the clumpuscles from collapsing slowly in the adiabatic regime. The collisions are supersonic for $D > 1$ with a Mach number $\lesssim 2$; temperature may increase during $\tau_{ff} = \tau_{KH}$ by a factor less than 3, and decreases after a few τ_{ff} since gravitating systems receiving energy cool. Also finite nearly isothermal clumps evaporate in about $100 \tau_{KH} = 100 \tau_{ff}$. The isothermal equilibrium of the fractal in the thermal bath is then a dynamic one: a quasi-steady state is reached by the balance between coalescence and fragmentation-evaporation of the clumps. In the first process the gas loses energy to the benefit of the background radiation, and the reverse occurs in the second process. Between collisions, the clumps tend to reach a r^{-2} density law, to be colder than the bath, and their envelopes evaporate. Perhaps not coincidentally the average kinetic energy density in turbulent gas motions in the outskirts of galaxies is of the same order of magnitude as the energy density of the 3 K radiation ($4\sigma T^4/c \approx 0.26 \text{ eV cm}^{-3}$).

Since forces in a hierarchical structure are mainly gravitational at every scale except at the smallest one, energy dissipation is strongly reduced. The very low dissipation rate resulting from this dynamical equilibrium explains the long life-time of the gas in outer gaseous disks, and the absence of star formation. But the gas in the optical disk gets in contact with heating sources, the gas temperature departs significantly from 3 K, the ratio of collision time to dynamical time increases considerably, and the clumpuscles may collapse to form stars¹⁾.

3. Simulations of fractal structures

We have carried out a large number of simulations of the fractal medium, in order to more quantitatively estimate the effective surface filling factor of the gas, and thus the amount of mass that can be “hidden” by such structures^{1),2)}. In Fig. 1 we show the projection of two simulated clouds with $D = 2.5$ and $D = 1.7$. Such clouds consist of $N = 10$ smaller subclouds, which are made of N smaller subsubclouds, etc. over 9 levels (10^9 elementary

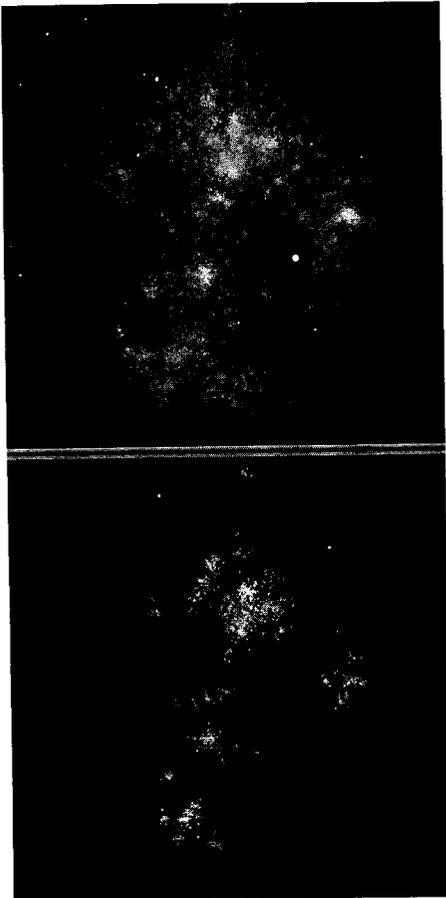


Fig. 1: Projected density of computer generated fractal clouds.
Top: $D=2.5$, **bottom:** $D=1.7$.

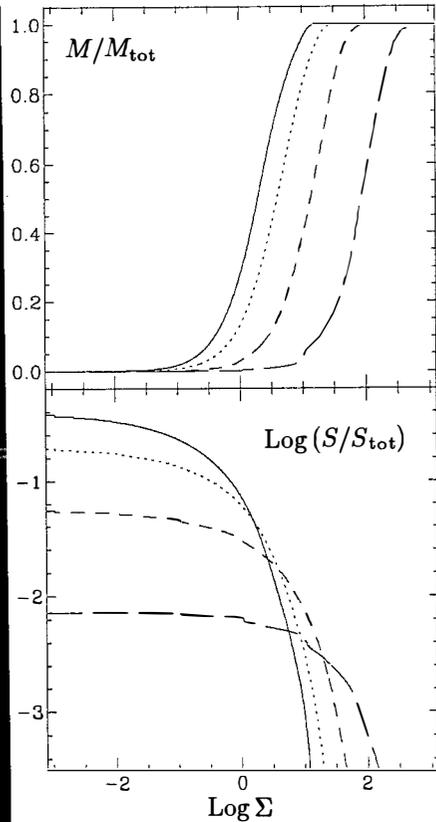


Fig. 2: Cumulated mass M at lower surface density Σ (**top**), and surface filling factor at higher Σ (**bottom**) as a function of $\log(\Sigma)$, for fractal clouds as in Fig. 1 of fractal dimensions $D=3$ (solid), 2.5 (dot), 2 (short dash) and 1.5 (long dash).

clouds). It is apparent that the sky coverage of the mass decreases considerably below $D = 2$. Then the observing properties are directly dependent on the smallest scale at which the fractal behaviour stops. The log grey scale used in Fig. 1 covers 6 decades in column density, showing the intrinsic difficulty to observe fractal clouds, since *most observing devices are limited to only 2 – 3 decades in column density*. The fundamental implications of this instrumental limitation on our perception of the ISM influencing its theoretical description has been critically discussed by Scalo²⁹⁾.

Although a range of possible parameters characterising the fractal are compatible with observations, we show in Fig. 2 that the surface filling factor of the dense clumpuscles can be only of the order of 1% or less. More quantitatively, the surface encompassing half of the mass is 12.7, 1.9 and 0.3% of the whole surface for $D = 3, 2$ and 1.5, in the more realistic model¹⁾. This means that optically thick gas in HI 21 cm radiation could have been missed

by factors between 10 to 100. Anyway, if most of the gas is close to 3 K emission would be almost impossible to detect. Moreover, the gas turns molecular as soon as the density is over 10^7 cm^{-3} , and metallicity problems then make its tracing very difficult. The low surface filling factor that we derive is useful however to estimate the probability of absorption: we expect roughly one optically thick H_2 absorption feature in front of a remote quasar for about $10^3 - 10^4 \text{ Ly}\alpha$ absorbing systems.

4. Cold gas – radiation coupling

A very important aspect of the physics of clumpuscles is the nature of their radiative coupling with the 3 K background radiation. In a primordial mix some relevant rotational lines from the gas are those of LiH and HD, at temperatures equal to a few tens of K, since they are optically thick for the column densities of the clumpuscles³⁰. Traces of CNO elements offer much more possibilities and the cooling is then efficient³¹.

But as soon as the temperature falls below $\approx 5 \text{ K}$, the coupling may be also ensured by H_2 ice, since a solid can radiate and absorb as a black-body. At these low temperatures, the H_2 gas is in equilibrium with a kind of H_2 snow³², because the average pressure P_* in the clumpuscles is then equal to the saturated vapour pressure P_s . The latter can be expressed as^{1),33} $P_s = 5.7 \cdot 10^{20} T^{5/2} \exp(-91.75/T) [\text{K cm}^{-3}]$. On the other hand the clumpuscle gas pressure is given by $P_* = 1.1 \cdot 10^8 T^{7/2} \mu^{1/2} f [\text{K cm}^{-3}]$. Below about 3.3 K, P_* is higher than P_s . At $T = 3 \text{ K}$, $P_s/P_* \approx 2 - 20$ and increases to 19-fold at $T = 2.736 \text{ K}$. The saturated pressure is reached in most of the clumpuscle. The question of the actual amount of solid H_2 depends critically on the presence of condensation sites such as dust, that we are presently unable to quantify for the outer part of galaxies. Since the heat liberated by freezing³³ ($\approx 100 \text{ K}/\text{H}_2$) is larger than the thermal energy by a factor 30, the clumpuscle contraction is correspondingly slowed down, as stars by nuclear reactions. Even if all the H_2 freezes out in some form of snow, the remaining He can sustain the H_2 flakes by viscosity. We estimate that for flake diameters smaller than 0.1 cm the dynamical time is much longer than the clumpuscle collision time, and the clumpuscle is optically thick to mm band radiation.

We note that clumpuscles radiating at the effective temperature $T \lesssim 3.3 \text{ K}$ would be exceedingly difficult to observe against the microwave background at $T_{\text{MWB}} = 2.736 \text{ K}$, essentially because the fraction of the sky covered by the clumpuscles in galaxies at a redshift $z < 0.2$ is much less than 1% of the galaxy surface if $D \approx 1.7$. The luminosity per unit mass ($\sim T^4 - T_{\text{MWB}}^4$) would be $\lesssim 0.002 \text{ erg g}^{-1} \text{ s}^{-1}$ in the mm band. The time-scale to dissipate this energy for reaching the galaxy rotation speed $v_c = 220 \text{ km s}^{-1}$ would be $\gtrsim v_c^2/0.002 \approx 8 \text{ Gyr}$. This lower bound would also be increased if we take into account the hotter T_{MWB} in the past. So a luminosity of $0.002 \text{ erg g}^{-1} \text{ s}^{-1}$ can be sustained for several Gyr by galactic rotation energy cascading down to the smallest level of the fractal gas.

5. Cold gas at the galaxy cluster scale

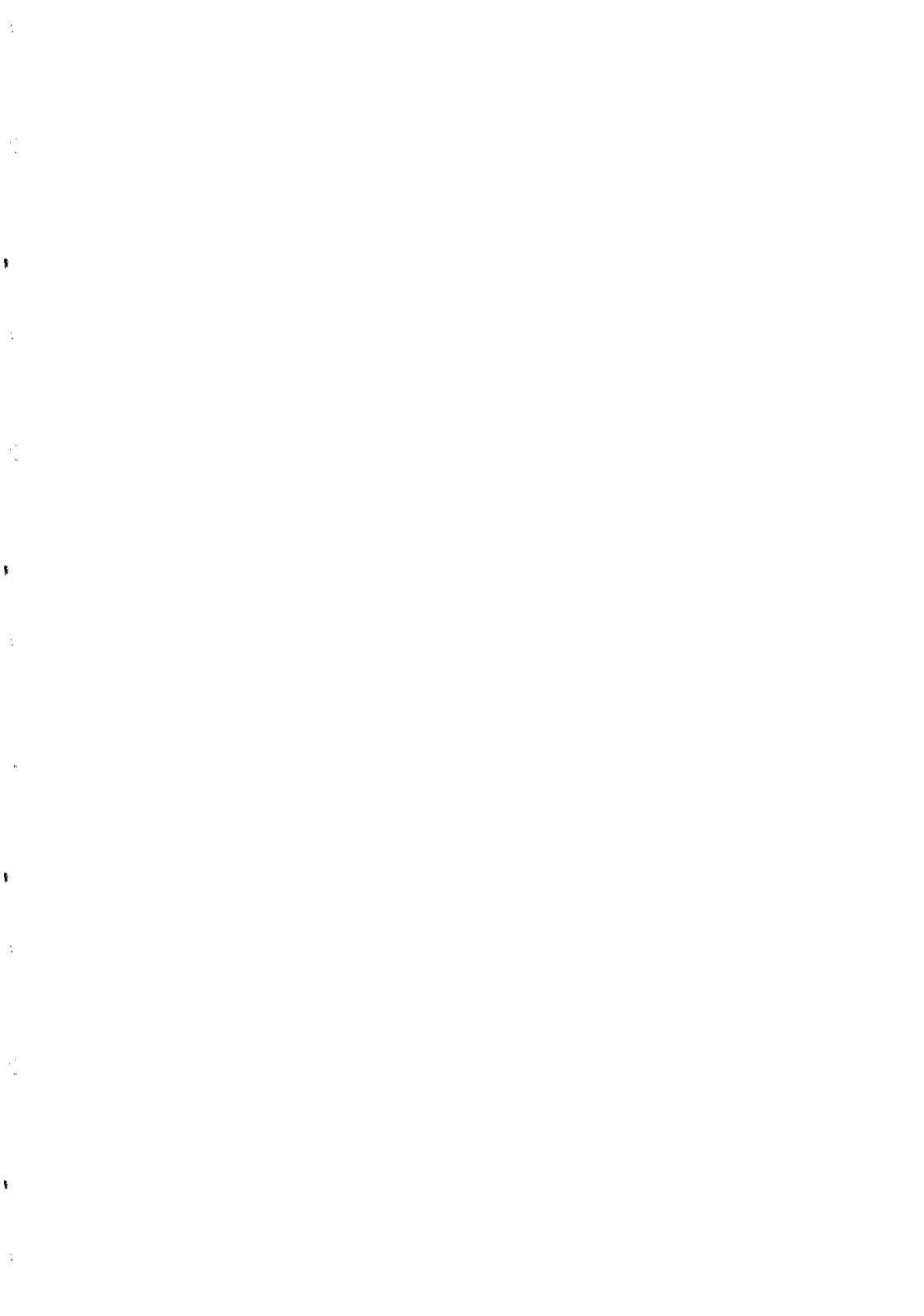
The observations of hot gas in galaxy clusters are perfectly compatible with our hypothesis. During galaxy interactions, one can expect that part of the fractal cold gas is destroyed and heated, and this explains why the mass of the intergalactic gas is so high, between 2 and 8 times the visible (i.e. stellar) mass³⁴. Clusters are therefore special regions where large amounts of baryonic dark matter becomes visible in X-rays. But this phase is only transient, and cooling flows bring this matter towards the very center of the cluster, where it cools and disappears from view into cold tiny clumps³¹. Recent observations in favor of a very small core radius for dark matter in clusters are then understandable^{35),36}.

Until now, we have only proposed that most of the dark matter around spiral galaxies is under the form of molecular cold gas^{1),3)}. However, we cannot rule out that most of the dark matter at the scale of clusters of galaxies is also in this form, and only baryonic. Indeed, the missing mass in clusters is compatible with the constraints brought by the standard big-bang nucleosynthesis³⁷⁾: $\Omega_b \lesssim 0.1$.

We are very grateful to Louis Martinet and John Scalo for interesting discussions.

References

- 1) Pfenniger D., Combes F. *Astr. Astrophys.* in press (1994)
- 2) Pfenniger D. in *Back to the Galaxy* S. Holt, F. Verter (eds.) AIP, New York, p. 584 (1993)
- 3) Pfenniger D., Combes F., Martinet L. *Astr. Astrophys.* in press (1994)
- 4) Bosma A. *Astron. J.* **86** 1971 (1981)
- 5) Freeman K.C. in *Physics of Nearby Galaxies: Nature or Nurture?* T.X. Thuan, C. Balkowski, J.T.T. Van (eds) Editions Frontières, p. 201 (1992)
- 6) Broeils A. *PhD thesis, University of Groningen* (1992)
- 7) Braine J., Combes F. *Astr. Astrophys.* **269** 7 (1993)
- 8) Kennicutt R.C. *Astrophys. J.* **272** 54 (1983)
- 9) Larson R.B., Tinsley B.M., Caldwell C.N. *Astrophys. J.* **237** 692 (1980)
- 10) Kennicutt R.C. in: *Star Formation in Stellar Systems*, G. Tenorio-Tagle, M. Prieto, F. Sánchez (eds.) Cambridge Univ. Press, Cambridge, p. 193 (1992)
- 11) Combes F., Debbasch F., Friedli D., Pfenniger D. *Astr. Astrophys.* **233** 82 (1990)
- 12) Pfenniger D., Norman C.A. *Astrophys. J.* **363** 391 (1990)
- 13) Friedli D., Benz W. *Astr. Astrophys.* **268** 65 (1993)
- 14) Casertano S., van Gorkom J.H. *Astron. J.* **101** 1231 (1991)
- 15) Sargent W.L.W. in *QSO Absorption Lines* J. Blades, D. Turnshek, C. Norman (eds.) Cambridge Univ. Press. p. 1 (1988)
- 16) Foltz C.B., Chaffee F.H., Black J.H. *Astrophys. J.* **324** 267 (1988)
- 17) Digel S., de Geus E., Thaddeus P. *Astrophys. J.* **422** 92 (1994)
- 18) Bloemen H. *Ann. Rev. Astr. Astrophys.* **27** 469 (1989)
- 19) Scalo J.M. in *Protostars and Planets II* D.C. Black, M.S. Matthews (eds.) Univ. of Arizona Press, Tucson, p. 201 (1985)
- 20) Falgarone E., Puget J-L, Perault M. *Astr. Astrophys.* **257** 715 (1992)
- 21) Diamond P.J., Goss W.M., Romney J.D. et al *Astrophys. J.* **347** 302 (1989)
- 22) Fiedler R.L., Dennison B., Johnston K.J., Hewish A. *Nature* **326** 675 (1987)
- 23) Bahcall J.N., Flynn C., Gould A. *Astrophys. J.* **389** 234 (1992)
- 24) Larson R.B. *Mon. Not. Roy. Astr.* **194** 809 (1981)
- 25) Larson R.B. *Mon. Not. Roy. Astr.* **256** 641 (1992)
- 26) Lynden-Bell D., Wood R. *Mon. Not. Roy. Astr.* **138** 495 (1968)
- 27) Katz J. *Mon. Not. Roy. Astr.* **183** 765 (1978)
- 28) Rees M.J. *Mon. Not. Roy. Astr.* **176** 483 (1976)
- 29) Scalo J. in *Physical Processes in Fragmentation and Star Formation* R. Capuzzo-Dolcetta et al. (eds.) Kluwer, Dordrecht, p. 151 (1990)
- 30) Lepp S., Shull J.M. *Astrophys. J.* **280** 465 (1984)
- 31) Ferlan G.J., Fabian A.C., Johnstone R.M. *Mon. Not. Roy. Astr.* **266** 399 (1994)
- 32) Sandford S.A., Allamandola L.J. *Astrophys. J.* **409** L65 (1993)
- 33) Encyclopédie des gaz—Gas Encyclopaedia, L'Air Liquide, Elsevier, (1976)
- 34) Edge A.C., Stewart G.C. *Mon. Not. Roy. Astr.* **252** 428 (1991)
- 35) Gerbal D., Durret F., Lima-Neto G., Lachièze-Rey M. *Astr. Astrophys.* **253** 77 (1992)
- 36) Wu X.P., Hammer F. *Mon. Not. Roy. Astr.* **262** 187 (1993)
- 37) Smith M.S., Kawano L.H., Malaney R.A. *Astrophys. J. Supp.* **85** 219 (1993)



HIGH-REDSHIFT MEASUREMENTS OF Ω_0

Bruno Guiderdoni

*Institut d'Astrophysique de Paris
98bis Boulevard Arago, F-75014 Paris*

ABSTRACT. This paper reviews the present status of the classical cosmological tests which use high-redshift objects: the angular-size test, the Hubble diagram, the Loh-Spillar test, the faint galaxy counts and the ages of distant galaxies. The review focusses on the crucial problem of galaxy evolution, and mentions new attempts based on supernovae, compact radio-sources and quasars. As a matter of fact, the simplest analysis of most of the data leads to a low-density open universe. Nevertheless, because of large error bars, the rejection of $\Omega_0 = 1$ is only marginal in most of the tests. Moreover, it is possible to accommodate the tests using galaxies with the flat universe and "non-standard" scenarios of evolution (involving starbursts, merging, bimodal IMF, etc.).

1 Introduction

The curvature of the universe is measurable by means of geometrical tests involving high-redshift objects. Some of these tests have been proposed in the early epoch of modern cosmology, but the past status of observations has hampered any significant conclusion. The recent development of sensitive detectors now allows one to revisit the question. The interest of these classical tests is their sensitivity to *all* matter, baryonic and non-baryonic, on very large scales comparable to the curvature radius. Here high- z objects are simply used as landmarks tracing the space-time structure. In principle, all the tests are straightforward and should give unambiguous results. Unfortunately, the high- z objects are not inert beacons. For instance, galaxies strongly evolve with time scales of a few Gyr. The analysis of the tests must take this evolution into account, by means of a model of spectrophotometric evolution which allows one to compute evolving synthetic spectra of galaxies from a minimum set of assumptions about the star formation history. By coupling these spectra with classical cosmological models, apparent magnitudes and colours can be derived, which take into account the entangling of the cosmological and evolutionary effects in a consistent way. The purpose of this paper is to describe the state-of-the-art in this rapidly-evolving field, by focussing on the problems arising from the modelling of galaxy evolution. The review also describes recent attempts based on other objects (supernovae, radio-sources and quasars) the evolution of which might have less influence on the tests. The reader wishing to know the general background of the tests is referred to the good review proposed by Sandage (1988).

2 Principles and fundamental formulae

In Friedman-Lemaître cosmological models, the curvature radius of the universe is:

$$R_0 = \frac{c}{H_0|\alpha_0|^{1/2}} \quad \text{with} \quad \alpha_0 = \Omega_0 + \lambda_0 - 1 \quad (1).$$

Here c is the speed of light, H_0 is the Hubble constant, Ω_0 is the density parameter, and λ_0 is the reduced cosmological constant. If $k = \text{sign}(\alpha_0) = -1$, the geometry is hyperbolic. If $k = +1$ it is spherical. The case $k = 0$ corresponds to a flat universe, with "infinite" curvature radius. This is the case predicted by inflation. We shall hereafter assume $\lambda_0 = 0$ (so $\Omega_0 = 2q_0$ the deceleration parameter), and come back to a non-zero λ_0 in section 8.

If one defines $\xi \equiv 1 + z$, the space-time properties of the universe can be computed by means of the time element and the comoving radial coordinate (see e.g. Weinberg, 1972):

$$dt(z) = -\frac{1}{H_0} \frac{d\xi}{\xi(\Omega_0\xi^3 - \alpha_0\xi^2 + \lambda_0)^{1/2}} \quad (2)$$

$$r_s(z) = S_k \left(\int_1^\xi \frac{|\alpha_0|^{1/2} d\xi}{(\Omega_0\xi^3 - \alpha_0\xi^2 + \lambda_0)^{1/2}} \right) \quad (3)$$

$$S_k(r) = \begin{cases} r & \text{if } k = 0; \\ \sinh r & \text{if } k = -1; \\ \sin r & \text{if } k = +1. \end{cases}$$

Then one can show that:

(i) The apparent bolometric flux f from an object of rest-frame bolometric luminosity L is given by

$$f = L/4\pi D_L^2(z) \quad \text{with} \quad D_L(z) \equiv R_0 r_s(z) \xi \quad (4).$$

$D_L(z)$ is the luminosity distance

(ii) The apparent angular size θ of an object of rest-frame linear size l is given by

$$\theta = l/D_A(z) \quad \text{with} \quad D_A(z) \equiv R_0 r_s(z)/\xi \quad (5).$$

$D_A(z)$ is the angular distance.

(iii) The volume element of a spherical shell at redshift z is

$$dV(z) = -4\pi D_A^2(z) c dt(z) \quad (6).$$

Thus any measurement of the variation of one of these quantities with z allows one to constrain the values of Ω_0 and λ_0 . For that purpose, several tests have been introduced:

(1) *Apparent magnitude m versus z* . This test is a measure of $D_L(z)$ and assumes that objects of constant luminosity L , the so-called "standard candles", are available.

(2) *Angular size θ versus z* . This test is a measure of $D_A(z)$ and assumes that objects of constant linear size l , the so-called "rigid rods", are available.

(3) *Surface density of objects N versus z* . This test, recently called the "Loh-Spillar" test, is a measure of the volume $V(z)$ of spherical shells by counting objects as far as the comoving number of the latter is conserved. It uses a large sample of objects while the previous tests use a class of special objects which can have anomalous behaviours.

(4) *Surface density of galaxies N versus apparent magnitude m .* This test, the classical faint galaxy counts, is a measure of $V(D_L)$. As the “Loh–Spillar” test, it also needs conservation of the comoving number of galaxies. Since V and D_L have related behaviours, it is a priori less sensitive to cosmology. Nevertheless its great advantage is to use only magnitudes and to probe the universe very deeply.

(5) *Ages of galaxies t_{gal} versus z .* It gives at each z a lower value of the age of the universe $t(z)$. The test is mainly sensitive to the Hubble constant H_0 . Its implicit assumption that the reddest galaxies are the oldest ones is not necessarily correct.

Among the various kinds of objects which have been used in the tests, galaxies are known to evolve strongly. Nevertheless the modelling of their evolution is a priori more accessible than the evolution of active nuclei, for instance. The computation of the k - and e - corrections (for red shift of the spectra and intrinsic evolution) needs specific models with the following principles (Bruzual 1981, 1983; Guiderdoni & Rocca-Volmerange 1987, 1988; Rocca-Volmerange & Guiderdoni 1988; Yoshii & Takahara 1988; Charlot & Bruzual 1991; Bruzual & Charlot 1993): at each time step, stars form from the gas content according to the classical parameters (Star Formation Rate and Initial Mass Function). These stars are placed on the Zero–Age Main Sequence of the $\log L - \log T_{eff}$ HR diagram. The models use compilations of stellar evolutionary tracks taking into account the main stages of stellar evolution in order to compute at each time step the distribution of the stellar populations in the HR diagram. This distribution is combined with a library of stellar spectra (or with stellar colours) in order to estimate the synthetic spectrum of the stellar population. The apparent magnitudes and colours at redshift z are derived from the spectra after coupling with the standard cosmological models and convolving with the response curves of filters.

Finally, it can be shown that all tests are independent of H_0 except through the evolutionary time scales in the e -corrections. Hereafter we shall take $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

3 The Hubble diagram

A “standard candle” with current absolute magnitude M_λ (in a filter centered at λ) has an apparent magnitude at redshift z :

$$m_\lambda = M_\lambda + (m - M)_{bol}(z) + k_\lambda(z) + e_\lambda(z) \quad (7)$$

with the bolometric distance modulus $(m - M)_{bol}(z) = 5 \log(D_L(z)/\text{Mpc}) + 25$. The status of the test was renewed by the observations of optical counterparts of radiogalaxies in the visible and in the near-IR at redshifts $z > 1$. The objects are selected on the basis of their radio properties in catalogues of powerful sources (3C, 1Jy, Parkes, 4C). The spectroscopic redshifts of the optical counterparts can be obtained from their strong emission lines.

The observer–frame B or V bands probe the UV in the galaxy rest frame, which is very sensitive to bursts of star formation and some scatter is observed. Nevertheless, the near-IR bands could be less sensitive to the history of star formation. So less scatter is expected and there was hope to derive a value of Ω_0 from K observations. Spinrad & Djorgovski (1987) proposed $\Omega_0 = 0.4 \pm 0.6$. More recently Weir et al. (1990) showed that the uncertainty in the history of star formation erases any cosmological information. Figure 1 shows predictions in the K -band Hubble diagram for various SFR histories with $\Omega_0 = 1$

and 0.1. It is difficult to discriminate between the models. There is some debate about the origin of the scatter which could be due to the mass of the objects and to starbursting evolution.

Tammann (1979) proposed to use supernovae as standard candles. Type Ia SNe are less affected by evolutionary effects than galaxies. After the first discovery of a SN at $z = 0.31$ (Norgaard-Nielsen *et al.* 1989), Perlmutter *et al.* (1994) recently measured the light curve of a SN at $z = 0.458$. The fit of the time-dilated light curve with light curves of "standard" SNe observed in nearby galaxies gives $\Omega_0 = 2q_0 = 0.2 \pm 0.6 (\pm 1.1)$. The computed value of Ω_0 should be increased if there is extinction in the host galaxy. The first type of errors is due to uncertainties in measurement and redshift corrections while the second type comes from the uncertainty in the true width of the absolute magnitude distribution which is known with a spread of ~ 0.3 mag. After the observation of more SNe at the Coma distance and beyond (i.e. without distance problems due to proper velocities), the absolute magnitude distribution might be narrower. With an increase of the sample of high- z SNe, the test would lead to a smaller computed range for Ω_0 . For the moment, the low-density universe seems to be favoured but the flat universe is rejected by the observations at only 1.3σ .

4 The angular-size test

The angular-size test used the diameters of cluster cores (Bruzual & Spinrad 1978; Hickson & Adams 1979) and the separation between the hot spots of double-lobe radio-sources (Wardle & Miley 1974) as "rigid rods". While there is too much scatter with cluster cores (see Sandage 1988), the intrinsic evolution of double-lobe sources makes the result of the test mimic the behaviour of the tired-light model ($D_A(z) \propto z^{-1}$). More recently, Kellermann (1993) applied the test to a sample of 82 compact radio-sources associated with active galaxies and quasars. He claims that, since compact sources are short-lifetime systems and are considerably smaller than the host galaxies, they are probably less affected by evolutionary effects than double-lobe sources. Moreover, the compact sources are seen at higher redshifts than the double-lobe sources. Although the sizes of compact sources are not unambiguously defined, he finds that the test is consistent with $\Omega_0 = 2q_0 = 1$.

5 The Loh-Spillar test

If $\phi(L, z)$ is the luminosity function (LF) at z , the surface density of galaxies in solid angle $d\omega$ with flux (per Hz) brighter than f_l is:

$$N(> f_l) = z^2 dz d\omega \frac{c^3}{H_0^3} A(z) \int_{L_l}^{\infty} \phi(L, z) dL \quad (8)$$

with the luminosity L_l per Hz:

$$L_l = \frac{4\pi z^2 c^2}{B(z) H_0^2} f_l \quad (9)$$

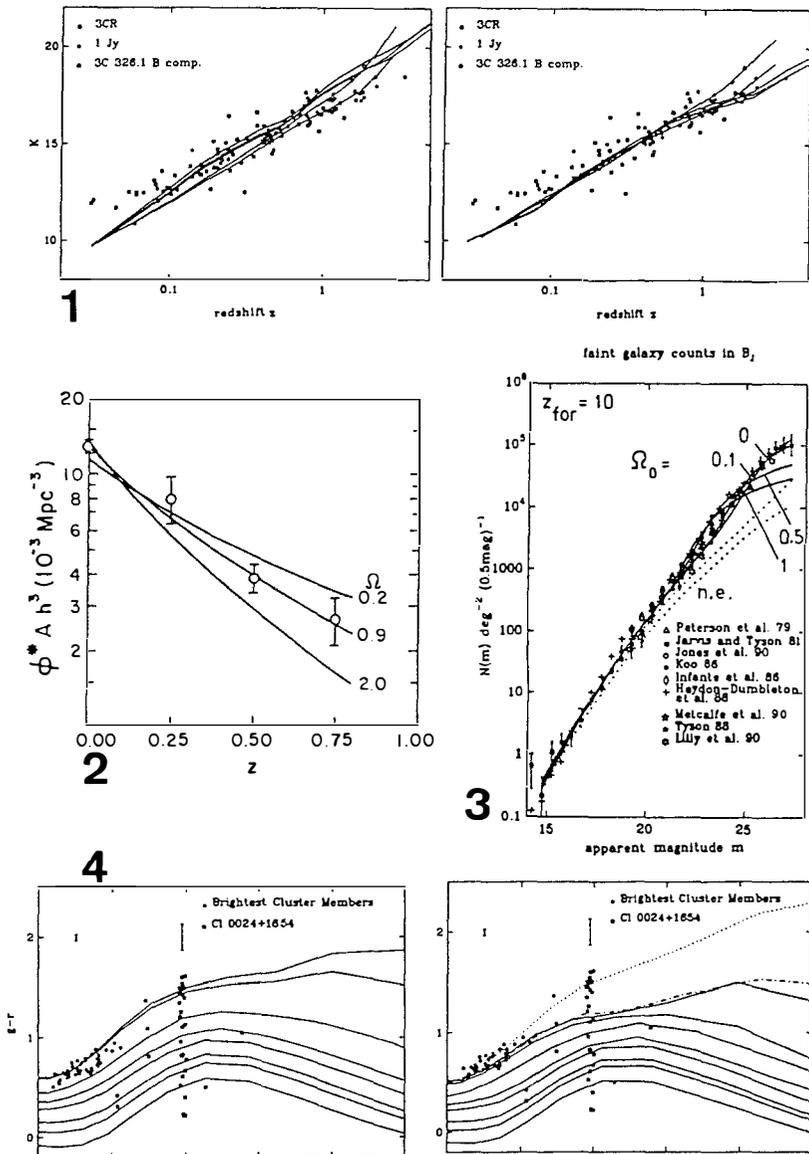


Figure 1: The K -band Hubble diagrams for $\Omega_0 = 0.1$ (left-hand panel) and 1 (right-hand panel). Various SFR time scales and redshifts of galaxy formation can accommodate the data within the observed scatter. **Figure 2:** The result of the Loh-Spallar test in its early version. From Loh & Spillar (1986). **Figure 3:** Sensitivity of the predicted counts to Ω_0 . Dotted lines: no evolution. **Figure 4:** Optical colour against redshift for various SFR histories and $\Omega_0 = 0.1$ (left-hand panel) and 1 (right-hand panel). Dashes-and-dots: "reddest" model, with SFR time scale 0.1 Gyr and redshift of galaxy formation $z_{\text{for}} = 40$. Dotted line: no evolution.

The quantities $A(z)$ and $B(z)$ only depend on z and on the values of Ω_0 and λ_0 and can be computed from eq. (6) and (4). The total flux (per Hz) from galaxies at redshift z is:

$$F(> f_l) = \frac{dz d\omega}{4\pi} \frac{c}{H_0} A(z) B(z) \int_{L_l}^{\infty} \phi(L, z) L dL \quad (10).$$

Providing (i) the LF has a constant shape (e.g. the Schechter form with characteristic luminosity L^* and constant slope α), and (ii) all galaxies evolve at the same rate (i.e. single-type evolution):

$$C \equiv \frac{F(> f_l)}{f_l N(> f_l)} = \frac{\int_{x_l}^{\infty} \phi(x, z) x dx}{x_l \int_{x_l}^{\infty} \phi(x, z) dx} \quad \text{with } x_l \equiv \frac{L_l}{L^*} \quad (11).$$

The method consists in measuring $N(> f_l)$ and $F(> f_l)$ at fixed z , computing C , and deriving x_l from eq. (11). Then, by putting this value into the integral of eq. (8) it is possible to determine $\phi^* A(z) c^3 / H_0^3$ where ϕ^* is the absolute normalisation of the LF (in Mpc^{-1}). This quantity does not depend on H_0 . Providing (iii) there is no number-density evolution (constant ϕ^*), it is easy to derive Ω_0 from the variation of $A(z)$ versus z .

This method has been used in Loh & Spillar (1986) and Loh (1986, 1988). They obtained narrow-band multicolor photometry of ~ 1000 galaxies in five $7' \times 10'$ fields at $m_{800nm} < 22$. The redshifts (up to $z \sim 0.75$) were determined from the spectrophotometry. Finally the authors' last analysis yielded $\Omega_0 = 1.1^{+0.9}_{-0.5}$ at the 95 % confidence level (see figure 2). Formally, the method has two advantages. It avoids the k -corrections by using shifted passbands for the various redshifts, as well as the e -corrections since luminosity evolution is directly measured (by the depth x_l of the survey at redshift z) under the assumptions (i) and (ii). Nevertheless, there has been much debate about the effect of departures from assumptions (i), (ii) and (iii).

First some-type dependent evolution would make the apparent $A(z)$ lower than the true $A(z)$, and consequently $\Omega_{0app} > \Omega_0$. Various estimates have been proposed: $\Delta\Omega_0 \simeq -0.16$ (crude estimate in Loh & Spillar early paper); $\Delta\Omega_0 \simeq -0.4$ (Loh 1988; Yoshii & Takahara 1990); up to $\Delta\Omega_0 \simeq -1$ in Bahcall & Tremaine (1988). These estimates (except the last one) are based on models of spectrophotometric evolution. The basic problem of the present status of the Loh-Spillar test is that the *measured* evolution is $\Delta M_B \equiv M_B(z=0) - M_B(z=0.75) = 0.3 \pm 0.2$ for red galaxies and 0.2 ± 0.2 for blue galaxies while all models predict $\Delta M_B = 0.6$ to 1 mag for red galaxies (and much weaker evolution for later galaxies). We must admit that we *do not* understand the evolution in the Loh-Spillar sample.

Secondly number-density evolution would make the apparent $A(z)$ higher than the true $A(z)$, and consequently $\Omega_{0app} < \Omega_0$. The estimates span from $\Delta\Omega_0 \simeq 0.1$ in Loh & Spillar (1986) and Loh (1988), to $\Delta\Omega_0 \simeq 2$ in Bahcall & Tremaine (1988). The reason of this discrepancy is that Loh & Spillar use a "low-merging scheme" based on observational estimates of pairs of bright galaxies while Bahcall & Tremaine give a theoretical estimate of the accretion of companions due to dynamical friction. Moreover Caditz & Petrosian (1989) showed that assumption (i) is not consistent with the data, independently of any assumed value of Ω_0 . In particular, the slope α of the LF evolves and can be as flat as $\alpha = -0.2$ at redshift $z > 0.4$. Such an evolution is opposite to that expected in any simple merging scenario. The authors suggest that the most likely value would be $\Omega_0 \sim 0.2$. Finally Omote

& Yoshida (1990) claimed that gravitational lensing could severely affect the method. All these works make the early result of the test very doubtful.

The test was recently re-examined by Schade & Hartwick (1994) with a compiled sample of AGNs and quasars. It is known that the quasar luminosity function evolves according to pure luminosity evolution (Boyle et al. 1987). So this evolution satisfies conditions (i) with a LF characterized by two power laws, (ii), and (iii). As a matter of fact, there might be a complicated evolutionary scheme with generations of short-lived quasars, which reproduces the “pure luminosity evolution”. Nevertheless, the Loh-Spillar method is independent of the way objects evolve as far as conditions (i) to (iii) are satisfied. By gathering data of high- z quasars with a sample of “nearby” AGNs, the authors find $\Omega_0 = 0.2^{+0.16}_{-0.1}$ (1σ errors) with a zero cosmological constant and $\Omega_0 = 0.4^{+0.1}_{-0.06}$ with $\Omega_0 + \lambda_0 + 1 = 0$. The strongest assumption of the work is the continuity of “pure luminosity evolution” between high- z quasars and nearby AGNs. With no reference to the nearby LF, the sample of QSOs gives $\Omega_0 = 0.45^{+0.3}_{-0.2}$. The flat universe is rejected by this conservative value at only 1.8σ .

6 The faint galaxy counts

The number of galaxies with type j at redshift $[z, z + dz]$ contributing to the counts per magnitude bin around apparent magnitude m_λ (through the λ filter) is:

$$d^2 N_j(m_\lambda, z) = \phi_j(M_\lambda)(1+z)^3 \frac{dV}{dz} dm_\lambda dz \quad (12).$$

dV is given in eq. (6). $\phi_j(M_\lambda)$ is the comoving LF for type j determined from the analysis of *nearby* galaxies. The absolute magnitude M_λ is computed from m_λ by means of eq. (7). Predictions for magnitude, redshift and colour distributions are computed by integrating the elementary contributions.

Guiderdoni & Rocca-Volmerange (1990) (and references therein) gave an analysis of the deepest surveys. The straightforward interpretation is based on the nearby luminosity functions for each spectral type, and on the extrapolation in the past of the standard scenarios reproducing the properties of the nearby galaxies, under the assumption of conservation of the comoving number density. Figure 3 shows the sensitivity of the standard predictions in the $N(m)$ diagram to Ω_0 . It turns out that the fits of the number counts (at $B \leq 28$ and in all bands from UV to red) and redshift distributions (at $B \leq 22$) require a low Ω_0 (~ 0.1) and a high value of z_{for} (~ 10). After some debate, the analysis of the K -band counts seems to lead to the same result (Cowie & Songaila 1993). It is not easy to get rid of this conclusion. As a matter of fact, it seems to be insensitive to uncertainties in the inputs of the model, since it is simply due to the smaller volume elements at high redshifts for $\Omega_0 = 1$.

The deep redshift surveys revealed the existence of a population of sub- L^* and dwarf galaxies at redshifts $z \sim 0.2$ to 0.5 (Broadhurst et al. 1988; Colless et al. 1990, 1993; Cowie et al. 1991). Most of the excess in the counts at $B = 24$ is due to this “new” population, with which it is possible to accommodate the deep counts down to $B = 28$ in an $\Omega_0 = 1$ universe. The corresponding slope of the LF is $\alpha \sim -1.6$ to -1.8 . It is not easy to understand why this population is not observed in the recent nearby surveys which give a LF

slope $\alpha \sim -1$ (Binggeli *et al.* 1990; Loveday *et al.* 1992). There might be a rapid evolution of the number of galaxies after merging of the small objects into bigger ones (Rocca-Volmerange & Guiderdoni 1990; Guiderdoni & Rocca-Volmerange 1991; Broadhurst *et al.* 1992). Alternatively, subtle selection effects might hide the low-surface brightness galaxies in nearby surveys which are not as sensitive as deep surveys (McGaugh 1994). For a recent review, see Guiderdoni (1993).

7 The ages of distant galaxies

The discovery of optical counterparts of radio-galaxies at redshifts $z > 3$ lead some authors to use the age of these distant galaxies as constraints on the values of the cosmological parameters. The first work by Lilly (1988) based on 0902+34 at $z = 3.395$ gave an age ≥ 1 Gyr. As a matter of fact, the age of the Universe at redshift 3.395 is 1.42 Gyr with $H_0 = 50$ and $\Omega_0 = 1$. Then several authors gave very small ages for the distant galaxies (Bithell & Rees 1990; Chambers & Charlot 1990). Thus the situation appears confused. In fact, Rocca-Volmerange & Guiderdoni (1990) showed that any age between ~ 0.5 Gyr and ~ 2.5 Gyr can fit the spectrophotometry within the error bars. Consequently it is impossible to put constraints on cosmology from these objects in the present state of the data.

May be the galaxies at redshifts ~ 0.5 are more interesting. The age of the universe was about two thirds of its current age. It turns out that the reddest galaxies at these redshifts are well reproduced by models *without* evolution (Colless *et al.* 1990). Guiderdoni & Rocca-Volmerange (1988) showed that models in a low-density universe, which have weak evolutionary corrections, can fit the data. On the contrary, it is difficult to accommodate the data with an $\Omega_0 = 1$ universe as shown in figure 4. The situation still degrades if H_0 is larger than 50. Several solutions could be imagined to this problem: (1) A low-value of H_0 . (2) Intrinsic extinction. It is not plausible since cluster galaxies which are devoid of gas show exactly the same behaviour as field galaxies. (3) Metallicity effect. The reddest galaxies might be metal rich. Nevertheless Yoshii & Takahara (1988) find the same difficulty with their model taking into account chemical evolution. (4) The IMF is not standard. Charlot *et al.* (1993) showed that, if strong bursts of star formation occur with an IMF which is deficient in low-mass stars (e.g. no stars below $2M_\odot$), the spectra of the galaxies can become very red when the stars with the lower masses become red giants. (5) The stellar clock is not accurate. This is related to the question of the ages of globular clusters which are only marginally consistent with the age of the universe 13 Gyr if $H_0 = 50$ and $\Omega_0 = 1$.

8 Conclusions

If we summarize the most simple results of the tests, it appears that they all seem to suggest an *open* universe with a low value of Ω_0 . Only the angular-size test based on the questionable assumption of the absence of cosmological evolution for the radii of compact radio-sources clearly points at the flat universe. We could be satisfied for that if the inflationary theory did not exist. But the inflationary paradigm is so attractive that it is worthwhile to revisit the tests.

Most problems for reconciling the data with inflation vanish if we introduce a non-zero cosmological constant. For instance, to save the faint counts, Fukugita et al. (1990) introduced a non-zero cosmological constant. The effect of $\lambda_0 \neq 0$ is to increase the observable volume of the universe and consequently the observable number of galaxies. The authors reproduce the counts with $\Omega_0 = 0.1$ and the corresponding $\lambda_0 = 0.9$ for the flat universe. On the other hand, $\lambda_0 = 0.9$ increases the age of the universe up to 25 Gyr and eliminates the difficulty with the colours of mild-redshift galaxies. Nevertheless, despite these successes (Efstathiou et al. 1990), cosmologists are very reluctant to introduce a non-vanishing λ_0 since it would require the fine-tuning that the inflationary paradigm precisely tries to avoid.

In fact, most of the current tests give only a *marginal* rejection of $\Omega_0 = 1$. Moreover, the evolution of galaxies is probably more complicated than the standard scenarios which are used in the classical approaches. On the one hand, we must admit that we *do not* understand the fate of the numerous sub- L^* and dwarfs seen at mild redshifts, but they seem to be what is needed to get an $\Omega_0 = 1$ universe from the counts and to destroy the original version of the Loh-Spillar test which uses galaxies (assumption (iii) is not valid any more). On the other hand, changes in the IMF of starbursts could alter the colour evolution of galaxies and make age determinations based on colours very hazardous.

Finally, it is interesting to recall that Hubble (1936) attempted to measure the curvature of the universe from faint counts. With the value $H_0 \simeq 530$, he found $\Omega_0 \simeq 14$. Some way has been done since this value and we can be reasonably optimistic for the future. In particular, the SNe Hubble diagram should give interesting results.

Acknowledgements. I am grateful to David Schade and Alfonso Aragón-Salamanca for providing informations. I would like to thank Yannick Giraud-Héraud for his help and the pleasant and stimulating discussions on this subject. Part of this work is drawn from a review published in *Observational Tests of Cosmological Inflation*, 1991, T. Shanks et al., Kluwer.

References

- Bahcall, S.R., & Tremaine, S., 1988. *Astrophys. J.*, **326**, L4.
 Binggeli, B., Tarenghi, M., & Sandage, A., 1990. *Astr. Astrophys.*, **228**, 42.
 Bithell, M., & Rees, M.J., 1990. *Mon. Not. R. astr. Soc.*, **242**, 570.
 Boyle, B.J., Fong, R., Shanks, T., & Peterson, B.A., 1987. *Mon. Not. R. astr. Soc.*, **227**, 717.
 Broadhurst, T.J., Ellis, R.S., & Shanks, T., 1988. *Mon. Not. R. astr. Soc.*, **235**, 827.
 Broadhurst, T.J., Ellis, R.S., & Glazebrook, K., 1992. *Nature*, **355**, 55.
 Bruzual, G., 1981. Ph.D. dissertation, University of California Berkeley.
 Bruzual, G., & Spinrad, H., 1978. *Astrophys. J.*, **220**, 1.
 Bruzual, G., 1983. *Astrophys. J.*, **273**, 105.
 Bruzual, G., & Charlot, S., 1993. *Astrophys. J.*, **405**, 5538.
 Caditz, D., & Petrosian, V., 1989. *Astrophys. J.*, **337**, L65.
 Chambers, K.C., & Charlot, S., 1990. *Astrophys. J.*, **348**, L1.

- Charlot, S., Ferrari, F., Mathews, G.J., & Silk, J., 1993. *Astrophys. J.*, **419**, L57.
- Charlot, S., & Bruzual, G., 1991. *Astrophys. J.*, **367**, 126.
- Colless, M., Ellis, R.S., Taylor, K., & Hook, R.N., 1990. *Mon. Not. R. astr. Soc.*, **244**, 408.
- Colless, M., Ellis, R.S., Broadhurst, T.J., Taylor, K., & Peterson, B.A., 1993. *Mon. Not. R. astr. Soc.*, **261**, 19.
- Cowie, L.L., Songaila, A. & Hu, E., 1991. *Nature*, **354**, 39.
- Cowie, L.L., & Songaila, A., 1993. in *First Light in the Universe: Stars or QSO's ?*, p. 147, ed. B. Rocca-Volmerange *et al.*, Kluwer.
- Efstathiou, G., Sutherland, W.J., & Maddox, S.J., 1990. *Nature*, **348**, 705.
- Fukugita, M., Takahara, F., Yamashita, K., & Yoshii, Y., 1990. *Astrophys. J.*, **361**, L1.
- Guiderdoni, B., & Rocca-Volmerange, B., 1987. *Astr. Astrophys.*, **186**, 1.
- Guiderdoni, B., & Rocca-Volmerange, B., 1988. *Astr. Astrophys. Suppl. Ser.*, **74**, 185.
- Guiderdoni, B., & Rocca-Volmerange, B., 1990. *Astr. Astrophys.*, **227**, 362.
- Guiderdoni, B., & Rocca-Volmerange, B., 1991. *Astr. Astrophys.*, **252**, 435.
- Guiderdoni, B., 1993. in *First Light in the Universe: Stars or QSO's ?*, p. 299, ed. B. Rocca-Volmerange *et al.*, Kluwer.
- Hickson, P., & Adams, P.J., 1979. *Astrophys. J.*, **234**, L91.
- Hubble, E., 1936. *Astrophys. J.*, **84**, 517.
- Kellermann, K.I., 1993. *Nature*, **361**, 134.
- Koo, D.C., 1986. *Astrophys. J.*, **311**, 651.
- Lilly, S., 1988. *Astrophys. J.*, **333**, 161.
- Loh, E.D., & Spillar, E.J., 1986. *Astrophys. J.*, **307**, L1.
- Loh, E.D., 1986. *Phys. R. Letters*, **57**, 2865.
- Loh, E.D., 1988. *Astrophys. J.*, **329**, 24.
- Loveday, J., Peterson, B.A., Efstathiou, G., & Maddox, S.J., 1992. *Mon. Not. R. astr. Soc.*, **390**, 79.
- McGaugh, S.S., 1994. *preprint*.
- Norgaard-Nielsen, H.U., *et al.*, 1989. *Nature*, **339**, 523.
- Omote, M., & Yoshida, H., 1990. *Astrophys. J.*, **361**, 27.
- Perlmutter, S., *et al.*, 1994. *preprint*.
- Rocca-Volmerange, B., & Guiderdoni, B., 1988. *Astr. Astrophys. Suppl. Ser.*, **75**, 93.
- Rocca-Volmerange, B., & Guiderdoni, B., 1990. *Mon. Not. R. astr. Soc.*, **247**, 166.
- Sandage, A., 1988. *Ann. Rev. Astron. Astrophys.*, **26**, 561.
- Schade, D., & Hartwick, F.D.A., 1994. *Astrophys. J.*, **423**, L85.
- Spinrad, H., & Djorgovski, S., 1987. in *Observational Cosmology, IAU Symp. 124*, p. 129, ed. A. Hewitt *et al.*, D. Reidel.
- Tammann, G.A., 1979. in *Astronomical Uses of the Space Telescope*, p. 329, ed. F.D. Macchetto *et al.*, ESO.
- Wardle, J.F.C., & Miley, G.K., 1974. *Astr. Astrophys.*, **30**, 305.
- Weinberg, S., 1972. *Gravitation and Cosmology*, Wiley, New York.
- Weir, W.N., Djorgovski, S., & Bruzual, G., 1990. in *Evolution of the Universe of Galaxies*, p. 25, ed. R. Kron, Astr. Soc. Pac..
- Yoshii, Y., & Takahara, F., 1988. *Astrophys. J.*, **326**, 1.
- Yoshii, Y., & Takahara, F., 1990. *Astrophys. J.*, **346**, 28.

THE AGE OF THE UNIVERSE, THE HUBBLE CONSTANT AND
QSOs IN A LOCALLY INHOMOGENEOUS UNIVERSE

J. W. Moffat [†] & D. C. Tatarski [‡]

Department of Physics

University of Toronto

Toronto, Ontario M5S 1A7, Canada

Abstract

A local void in the globally Friedmann-Robertson-Walker cosmological model with the critical density ($\Omega_0 = 1$) is studied. The inhomogeneity is described using a Lemaître-Tolman-Bondi solution for a spherically symmetric distribution of matter. The scale of the central underdense region is ~ 150 Mpc. We investigate the effects this has on the cosmological time scale, the measurement of the Hubble constant and the redshift–luminosity distance for moderately and very distant objects ($z \sim 0.1$ and more). The results indicate that if we happened to live in such a void, but insisted on interpreting cosmological observations through the FRW model, we could go wrong in a few instances. For example, the Hubble constant measurement could give results depending on the separation of the source and the observer, the quasars could be younger than we think and also less distant (less energetic).

[†]: moffat@medb.physics.utoronto.ca

[‡]: tatarski@medb.physics.utoronto.ca

1 Introduction

It seems, particularly after the introduction of the inflationary paradigm [1], that the isotropic and homogeneous Friedmann-Robertson-Walker (FRW) cosmological models are best suited for the description of the global structure and the evolution of the universe. However, a similar statement is not necessarily true when cosmologically moderate scales are thought of.

There exists direct observational evidence in favour of the isotropy of the observed universe, namely, the COBE data confirming a high degree of isotropy of the cosmic microwave background radiation (CMBR) [2, 3]. However, there is no observationally based reason supporting the assumption of homogeneity. On the contrary, there seems to exist observational evidence in favour of larger and larger structures [4].

Recent work on modelling voids in the expanding universe [5] has shown that it is possible to construct an asymptotically FRW universe containing an expanding spherical void of the Lemaître-Tolman-Bondi type ¹. Moreover, the void can be a region of under-density rather than vacuum.

We think that there exists sufficient observational evidence (briefly discussed later in this paper) to support a conjecture that we may live in a relatively large underdense region embedded in a globally FRW universe. Exploring physical properties of such a model is the aim of the present paper.

In the following section, we briefly discuss the LTB model. Section 3 consists of a brief discussion of the observational background and the simple toy model of a local void presented here. The closing section contains a description of our results of numerical calculations as well as conclusions.

Throughout this paper we use units in which $G = c = 1$, unless stated otherwise.

2 The Model

First, for the sake of notational clarity, let us recall the FRW line element:

$$ds^2 = dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right], \quad (1)$$

with $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$.

¹A cosmological solution spherically symmetric about one point was first proposed by Lemaître [6]. However, it is usually called the Tolman-Bondi solution [7, 8].

Now, let us consider a Lemaitre-Tolman-Bondi [7, 8] model for a spherically symmetric inhomogeneous universe filled with dust. The line element in comoving coordinates can be written as:

$$ds^2 = dt^2 - R'^2(t, r)f^{-2}dr^2 - R^2(t, r)d\Omega^2, \quad (2)$$

where f is an arbitrary function of r only, and the field equations demand that $R(t, r)$ satisfies:

$$2R\dot{R}^2 + 2R(1 - f^2) = F(r), \quad (3)$$

with F being an arbitrary function of class C^2 , $\dot{R} = \partial R/\partial t$ and $R' = \partial R/\partial r$. We have three distinct solutions depending on whether $f^2 < 1, = 1, > 1$ and they correspond to elliptic (closed), parabolic (flat) and hyperbolic (open) cases, respectively.

The proper density can be expressed as:

$$\rho = \frac{F'}{16\pi R'R^2}. \quad (4)$$

Whatever the curvature, the total mass within comoving radius r is:

$$M(r) = \frac{1}{4} \int_0^r dr f^{-1} F' = 4\pi \int_0^r dr \rho f^{-1} R' R^2, \quad (5)$$

so that

$$M'(r) = \frac{dM}{dr} = 4\pi \rho f^{-1} R' R^2.$$

Also for $\rho > 0$ everywhere we have $F' > 0$ and $R' > 0$ so that in the non-singular part of the model $R > 0$ except for $r = 0$ and $F(r)$ is non-negative and monotonically increasing for $r \geq 0$. This could be used to define the new radial coordinate $\bar{r}^3 = M(r)$ and find the parametric solutions for the rate of expansion.

In the flat (parabolic) case $f^2 = 1$, we have

$$R = \frac{1}{2}(9F)^{1/3}(t + \beta)^{2/3}, \quad (6)$$

with $\beta(r)$ being an arbitrary function of class C^2 for all r . After the change of coordinates $R(t, \bar{r}) = \bar{r}(t + \beta(\bar{r}))^{2/3}$, the metric becomes:

$$ds^2 = dt^2 - (t + \beta)^{4/3} (Y^2 dr^2 + r^2 d\Omega^2), \quad (7)$$

where

$$Y = 1 + \frac{2r\beta'}{3(t + \beta)}, \quad (8)$$

and from (4) the density is given by

$$\rho = \frac{1}{6\pi(t + \beta)^2 Y}. \quad (9)$$

Clearly, we have that ($t \rightarrow \infty$) the model tends to the flat Einstein–de Sitter case.

For the closed and open cases the parametric solutions for the rate of expansion can be written as [9]:

$$R = \frac{1}{4}F(1 - f^2)^{-1} [1 - \cos(v)], \quad f^2 < 1, \quad (10a)$$

$$t + \beta = \frac{1}{4}F(1 - f^2)^{-3/2} [v - \sin(v)], \quad f^2 < 1, \quad (10b)$$

and

$$R = \frac{1}{4}F(f^2 - 1)^{-1} [\cosh(v) - 1], \quad f^2 > 1, \quad (11a)$$

$$t + \beta = \frac{1}{4}F(f^2 - 1)^{-3/2} [\sinh(v) - v], \quad f^2 > 1, \quad (11b)$$

with $\beta(r)$ being again a function of integration of class C^2 and v the parameter.

The flat case ($f^2 = 1$) has been rather extensively studied elsewhere [10]. The model depends on one arbitrary function $\beta(r)$ and could be specified by assuming the density on some space-like hypersurface, say $t = t_0$.

The cases of interest to us, (10) and (11), correspond to closed and open models, respectively.

Before we proceed (in the next section) to discuss the observational grounds for modelling a local void, we need to amplify the discussion of the LTB model by introducing basic features of the propagation of light in our model. The high degree of isotropy of the microwave background forces us to the conclusion that we must be located very close to the spatial centre of the void. In our discussion, for the sake of simplicity, we place an observer at the centre ($t_{Ob} = t_0, r_{Ob} = 0$).

The luminosity distance between an observer at the origin of our coordinate system ($t_0, 0$) and the source at $(t_e, r_e, \theta_e, \phi_e)$ is [8]:

$$d_L = \left(\frac{\mathcal{L}}{4\pi\mathcal{F}} \right)^{1/2} = R(t_e, r_e) [1 + z(t_e, r_e)]^2, \quad (12)$$

where \mathcal{L} is the absolute luminosity of the source (the energy emitted per unit time in the source's rest frame), \mathcal{F} is the measured flux (the energy per unit time per unit area as measured by the observer) and $z(t_e, r_e)$ is the redshift (blueshift) for a light ray emitted at (t_e, r_e) and observed at $(t_0, 0)$.

The light ray travelling inwards to the centre satisfies:

$$ds^2 = dt^2 - R^2(t, r)f^{-2}dr^2 = 0, \quad d\theta = d\phi = 0,$$

and thus

$$\frac{dt}{dr} = -R'(t, r)/f(r). \quad (13)$$

Without getting into a detailed discussion, which can be found in [8, 10], let us state that if the equation of the light ray travelling along the light cone is:

$$t = T(r), \quad (14)$$

using (13) we get the equation of a ray along the path:

$$\frac{dT(r)}{dr} = -\frac{R'}{f}[T(r), r], \quad (15)$$

where

$$\dot{R}'[T(r), r] = \left. \frac{\partial^2 R}{\partial t \partial r} \right|_{r, T(r)} = \left. \frac{\partial R'}{\partial t} \right|_{r, T(r)}.$$

The equation for the redshift considered as a function of r along the light cone is:

$$\frac{dz}{dr} = (1+z)\dot{R}'[T(r), r], \quad (16)$$

and the shift z_1 for a light ray travelling from (t_1, r_1) to $(t_0, 0)$ is:

$$\log(1+z_1) = -\log(1-a_1) - \int_0^{r_1} dr \frac{M'(r)}{r(1-a_1)}, \quad (17)$$

where

$$a_1(r) = \dot{R}[T(r), r],$$

and, in obtaining the second equation, we used (4) and (5). Thus we have two contributions to the redshift. The cosmological redshift due to expansion, described by the first term with $a_1 = \dot{R}$, and the gravitational shift due to the difference between the potential energy per unit mass at the source and at the observer. Obviously, in the homogeneous case ($M'(r) = 0$) there is no gravitational shift.

3 The modelling of the local void

If we restrict ourselves to spatial scales that have been well probed observationally, i.e. up to a few hundred Mpc, the most striking feature of the luminous matter distribution is the existence of large

voids surrounded by sheet-like structures containing galaxies (e.g. [11]). The surveys [11],[12] give a typical size of the voids of the order 50–60 h^{-1} Mpc. There has also been some evidence [13] –with less certainty– for the existence of larger underdense regions with characteristic sizes of about 130 h^{-1} Mpc. Also, dynamical estimates of the FRW density parameter Ω_0 give very different results on different scales. The observations of galactic halos on scales less than about 10 to 30 Mpc typically give (see e.g. [14]) $\Omega_{10-30} \simeq 0.2 \pm 0.1$. On the other hand, smoothing the observations over larger scales (> 20 Mpc, say ~ 100 Mpc) indicates (e.g. [12]) the existence of a less clustered component with a contribution exceeding 0.2, and perhaps as high as $\Omega_{\sim 100} \simeq 0.8 \pm 0.2$.

At the same time, the large scale galaxy surveys (some of the recent literature is given in [15]) firmly indicate a considerable excess in the number–magnitude counts for faint galaxies relative to predictions of homogeneous, “no-evolution” models. This excess could be the result of a non-standard galactic evolution or could be caused by rather exotic FRW cosmology (i.e. the deceleration parameter $q_0 \ll 0.5$ or a non-zero cosmological constant Λ). However, it can also be treated as an observational indication of a very large (on the scale of the redshift $z \sim 0.5$) void. A model of a local void with a density distribution based on the faint galaxies number counts is presently being studied [16].

In the model presented here, we confine ourselves to the simple density distributions. We study two cases: a void with the central density equal to that of an FRW model with the density parameter $\Omega_0 = 0.2$, asymptotically approaching the FRW model with $\Omega_0 = 1$, and a very similar void “distorted” on an intermediate scale by a peak ($\Omega \lesssim 1$) in the density distribution. The two distributions are:

$$\Omega_v(r) = \Omega_{min} + (\Omega_{max} - \Omega_{min}) \left[1 - \left(\frac{r}{L} \right)^2 \frac{\exp(r/L)}{[\exp(r/L) - 1]^2} \right], \quad (18a)$$

and

$$\Omega_{vp}(r) = \Omega_{min} + (\Omega_{max} - \Omega_{min}) \left[1 - \left(\frac{r}{L} \right)^2 \frac{\exp(r/L)}{[\exp(r/L) - 1]^2} \right] + \left(\frac{r}{l} \right)^2 e^{-(r/l)^2}. \quad (18b)$$

In the numerical calculations presented in the next section we used the values $\Omega_{min} = 0.2$, $\Omega_{max} = 1$ and $L = l = 30$ Mpc. This assures that the void converges satisfactorily fast to the outside critical FRW universe ($\Omega \simeq 0.86$ for $r \simeq 150$ Mpc and $\Omega \simeq 0.95$ for $r \simeq 200$ Mpc) and that the intermediate peak is observationally acceptable ($\Omega \simeq 0.64$ for $r \simeq 33$ Mpc).

4 The results and discussion

In general, an LTB model depends on three arbitrary functions, see section 2, $F(r)$, $\beta(r)$ and $f(r)$. Since $F(r)$ can be interpreted as twice the effective gravitational mass within comoving radius r ([8]), then, in accordance with the discussion following (5), assuming its form is equivalent to a coordinate choice. In our calculations we used $F(r) = 4r^3$. The second function, $\beta(r)$, sets the initial singularity (“big bang”) hypersurface of the model. Since we want the outside region in our toy model to be fully equivalent to the critical FRW universe, we set $\beta(r) = 0$, thereby assuming a universally simultaneous big bang. We also set the time coordinate of constant time hypersurface “now” so that it is equal to the age of the universe t_0 in the FRW model with $\Omega_0 = 1$. In doing so, we give up a very important feature of an LTB model: an extra (with respect to FRW) degree of freedom that would allow the age of the universe to be different from FRW or even position dependent. The third (“curvature”) function, $f(r)$, is an unknown to be solved for in our calculations. From the work done in [5], we conclude that the LTB case to be used in modelling an underdense comoving void in an FRW universe is the hyperbolic ($f^2 > 1$) one.

In a manner similar to that employed in [10], we assume that since all cosmological observations are necessarily done by detecting some form of electromagnetic radiation, the solution should progress along the light cone. The final set of equations we solve consists of the equations (15), (16) and the equation describing the density distribution (4) through either of the relations (18) taken along the light cone, e.g.:

$$\rho[T(r), r] = \frac{F'(r)}{16\pi R'[T(r), r]R^2[T(r), r]} = \Omega_v(r). \quad (19)$$

Since $T(r)$ (the time of emission t_e of a light ray observed at $r = 0$ at t_0) is now given by (11b), the functions to be solved for are $f(r)$, $z(r)$ and $v(r)$, where the parameter v becomes the function of position. The initial conditions for the integration have to be set at $r \neq 0$, since the analytic expressions (11) are singular at $r = 0$, where $f^2 = 1$ (we have a flat $\Omega_0 = 0.2$ FRW universe there). We assume that for the initial radius $r_i \ll 1$ (we use dimensionless radius and time in the calculations) corresponding z_i and t_i are given by their standard FRW values. Then $z(r_i)$, $v(r_i)$ and $f(r_i)$ can be obtained from (11b), (15) and (16).

Once the equations have been numerically integrated we use (11b) to obtain t_e for a given r_e . The luminosity distance d_L corresponding to this event is obtained with the use of (12) and $z(t)$ (useful in studying the cosmological time scale) is given by the parametric relation $[T(r), z(r)]$.

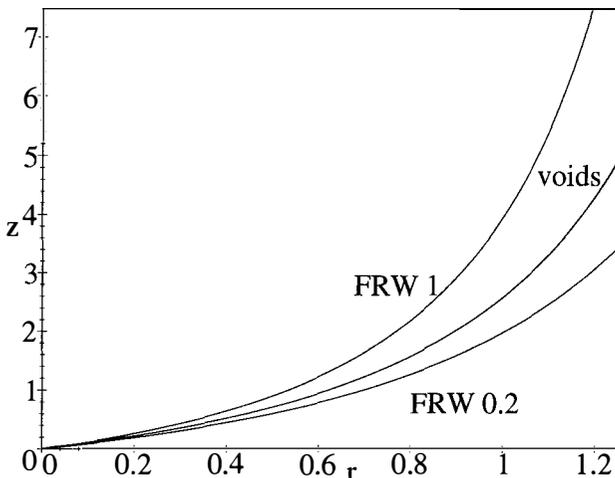


Figure 1: The redshift z as a function of comoving radius r . The FRW results for $\Omega = 1$ and $\Omega = 0.2$ are denoted by “FRW 1” and “FRW 0.2”, respectively. The results for the LTB modelled voids are denoted “voids”.

The results of our numerical calculations are as follows. Figure 1 depicts $z(r)$, where r is the dimensionless comoving radius used in the calculations. The coordinate distance has no direct physical relevance, but our units here are such that $r = 1$ corresponds to $2997.95h^{-1}$ Mpc, where h is the usual coefficient in the observationally determined value of the Hubble constant: $H_0 = 100h \text{ km s}^{-1}\text{Mpc}^{-1}$.

The departure of $z(r)$ from its FRW behaviour does not seem to be dramatic. In fact, on the distance scale used in Figure 1 it is hardly noticeable. More details, on a smaller spatial scale, can be seen in Figure 2. The redshift z , after being influenced noticeably by rapid changes in the density ρ on smaller scales, asymptotically tends to a limit that could, in accordance with FRW interpretation, correspond to the universe with the density parameter in the range $\Omega \in (0.2; 1)$.² The increase in z on intermediate scales is clearly induced by the additional gravitational shift caused by the mass distribution of the void. The large scale behaviour is controlled by our assumption of the equality $t_0 = t_{0FRW}$.

²This situation does not change considerably if we model the central void in a similar manner but with respect to the observable z as a variable [16].

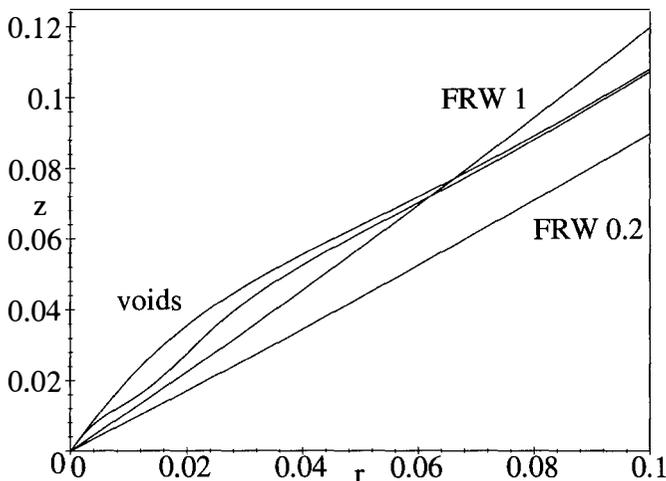


Figure 2: The redshift z as a function of comoving radius r . Smaller scales. Notation as in Fig. 1.

However, one should not forget that the comoving distance is *not* an observable, whereas the luminosity distance d_L is. In principle, provided we know its absolute luminosity \mathcal{L} , we can establish the luminosity distance, defined by $4\pi d_L^2 = \mathcal{L}/\mathcal{F}$, by measuring the energy flux \mathcal{F} of an observed object (for a discussion of usual caveats associated with so-called “standard candles” see e.g. [17]).

Due to the lack of space we do not present the $z(t)$ relation here. There is, again, some small scale ($t_0 - t \ll 1$) divergence from the FRW behaviour, but for early cosmological times ($t \ll 1$) the relation tends to the critical FRW ($\Omega = 1$) one. This is in accordance with our assumption of a simultaneous big bang, $\beta(r) = 0$, and with our setting the age of the universe to be equal to that of the critical FRW case. Objects with redshifts of order a few are younger than their FRW counterparts, but not significantly.

Figures 3 and 4 present the redshift–luminosity distance relation on cosmologically large (Figure 3) and intermediate (Figure 4) scales. Now the departure from the FRW behaviour becomes apparent. This comes as no surprise if one recalls the formulae for d_L in both LTB (12) and FRW:

$$d_L^2 = a^2(t_0)r_e^2(1+z_e)^2. \quad (20)$$

The important question is whether our results contradict the linearity of the Hubble relation $z = Hd$, well established on small scales.

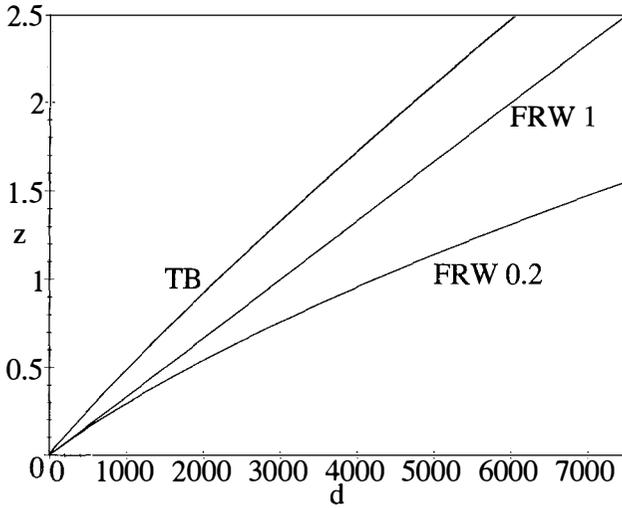


Figure 3: The redshift z as a function of the luminosity distance d_L (in Mpc).

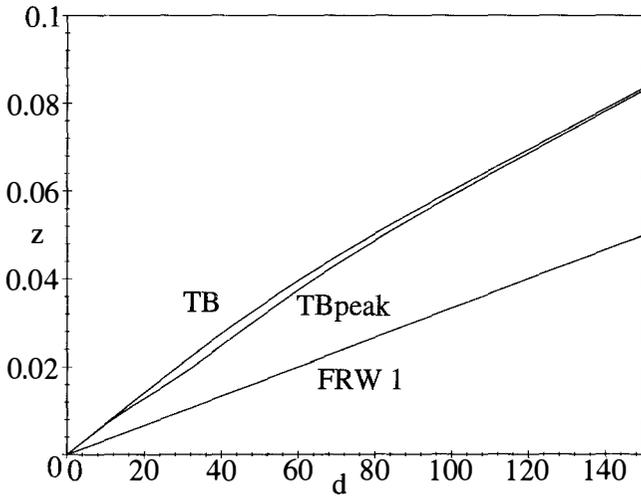


Figure 4: The redshift z as a function of the luminosity distance d_L (in Mpc). Smaller scales.

Due to our choice of the cosmological time scale and FRW embedding, the asymptotic ($z \rightarrow \infty$) behaviour of $z(d_L)$ is that of the critical ($\Omega = 1$) FRW case. On intermediate scales, however, objects of comparable redshifts are located at smaller luminosity distances. The ratio d_{LTB}/d_{FRW} is ≈ 0.7 (≈ 0.5) for $z \approx 1$, ≈ 0.8 (≈ 0.4) for $z \approx 2.5$ and ≈ 0.9 (≈ 0.36) for $z \approx 4.5$, where the most distant quasars are observed. (Values in parentheses correspond to the FRW $\Omega = 0.2$ case.) Since the absolute luminosity \mathcal{L} of the source scales as the square of d_L this reduces the energy output of QSOs (up to an order of magnitude). Also, the angular diameter distance $d_A = D/\delta = (1+z)^{-2}d_L$ now gives a smaller proper distance D across the source for the same observed angular diameter δ . This helps resolve problems with seemingly acausal signals (correlations in luminosity bursts) observed across some quasars.

At the same time, inspection of Figure 4 shows that on small scales a very nearly linear (in fact, observationally indistinguishable from linear) ‘‘Hubble diagram’’ is obtained. However, a different value for the Hubble parameter (constant) is inferred (position, or rather d_L , dependent on larger scales), if we insist on interpreting the results of cosmological observations through an FRW model.

To explore this possibility let us recall that in FRW cosmology the exact result for the Hubble relation (z versus d_L) in the matter dominated universe is [17]:

$$H_0 d_L = q_0^{-2} \left[z q_0 + (q_0 - 1) \left(\sqrt{2z q_0 + 1} - 1 \right) \right], \quad (21)$$

where $q_0 \equiv -\ddot{a}(t_0)/a(t_0)H_0^2$ is the deceleration parameter.

Let us assume that we live in a local LTB void and the z vs. d_L relation differs from the FRW one as described in this paper, but we are biased by our theoretical prejudice and interpret cosmological observations through the FRW model.

On cosmologically small distances we measure the same value of H_0 independently of the model (we call this value ‘‘the local measurement’’). This stems from the fact that, due to our assumptions, very close to the centre ($r \ll 1$) the model is well approximated by the FRW universe with $\Omega = 0.2$. Obviously, if the universe were LTB rather than FRW, then the Hubble parameter based on the observed (LTB) values of z and d_L , but inferred through an FRW relation (21), would be position (redshift) dependent as shown in Figure 5.

The values of H_0 reported to date span the range 40 to 100 km s⁻¹Mpc⁻¹ (with standard errors quoted frequently as 10 km s⁻¹Mpc⁻¹ or less!). Inhomogeneities similar to the LTB void presented here might provide an explanation for this.

The LTB void, such as the one presented here, decreases its density contrast (the depth of the

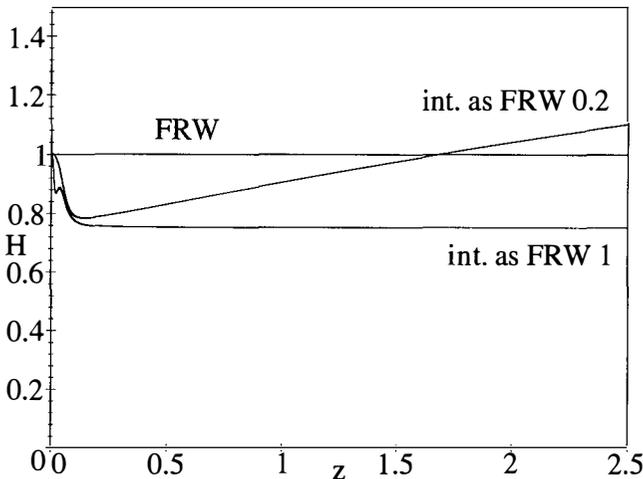


Figure 5: The “observed” Hubble constant H_0 (in units of the local measurement) as a function of the redshift z . The results interpreted through FRW $\Omega = 0.2$ and $\Omega = 1$ models, respectively, are denoted “int. as FRW 0.2” and “int. as FRW 1”.

void with respect to the FRW background) when evolved back in time [9, 10]. At early times it is almost homogenized (at $t/t_0 \simeq 10^{-5}$ we have $|\rho_{LTB}(r)/\rho_{FRW} - 1| < 10^{-6}$). This corresponds to a universe which at the beginning is very similar to the FRW one, but different at late stages.

In this manner, while retaining all accomplishments of the FRW cosmology in dealing with epochs preceding the matter dominated era, we can gain new freedom in modelling the more recent universe. We can solve the age of the universe problem (by assuming $\beta(r) = const \neq 0$), provide the excess power observed on scales of 5—10,000 km s $^{-1}$ in modelling structure formation (see [10]), alleviate a few old problems associated with quasars (their age, luminosity and size) and provide an explanation for the wide range of reported values of the Hubble constant.

Acknowledgement

The authors thank Ken-Ichi Nakao for helpful discussions and suggestions.

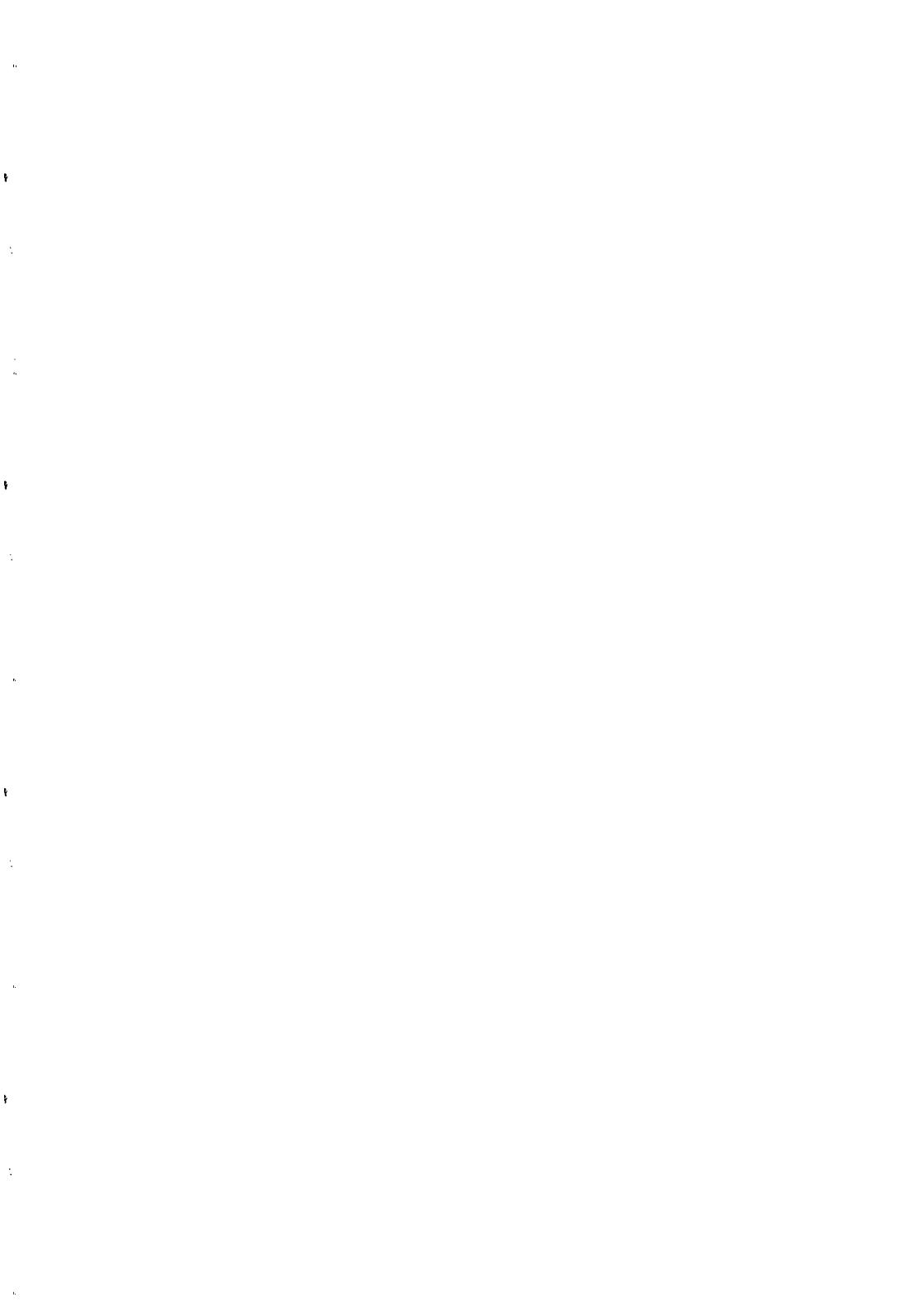
References

- [1] Guth A. H., *Phys. Rev. D*, **23**, 347, 1981.

- [2] Mather J. C. et al., *Astrophys. J.*, **354**, L37, 1990; Hogan C. J., *Nature*, **344**, 107, 1990.
- [3] Smoot G. F. et al., *Astrophys. J.*, **396**, L1, 1993; Wright E. L. et al., *Astrophys. J.*, **396**, L13, 1993.
- [4] Lynden-Bell D. et al., *Astrophys. J.*, **326**, 19, 1988; Geller M. J. & Huchra J. P., *Science*, **246**, 897, 1989; Clowes R. G. & Campusano L. E., *M.N.R.A.S.*, **249**, 218, 1991.
- [5] Bonnor W. B. & Chamorro A., *Astrophys. J.*, **361**, 21, 1990; *Astrophys. J.*, **378**, 461, 1991.
- [6] Lemaître G., *M.N.R.A.S.*, **91**, 490, 1931; *Ann. Soc. Sci. Bruxelles* **A53**, 51, 1933.
- [7] Tolman R. C., *Proc.Nat.Acad.Sci.*, **20**, 169, 1934.
- [8] Bondi H., *M.N.R.A.S.*, **107**, 410, 1947.
- [9] Bonnor W. B., *M.N.R.A.S.*, **159**, 261, 1972; *M.N.R.A.S.*, **167**, 55, 1974.
- [10] Moffat J. W. & Tatarski D. C., *Phys. Rev. D* **45**, 3512, 1992.
- [11] Geller M. J. & Huchra J. P., *Science*, **246**, 897, 1989.
- [12] Efstathiou G. et al., *M.N.R.A.S.*, **247**, 10, 1990; Saunders W. et al., *Nature*, **349**, 32, 1991.
- [13] Broadhurst T. J. et al., *Nature*, **343**, 726, 1990.
- [14] Sancisi R. & van Albada T. S., in *'Dark Matter in the Universe'*, eds. J. Kormendy & G. Knapp, Reidel, Dordrecht, 1987.
- [15] Maddox S. J. et al., *M.N.R.A.S.*, **240**, 43, 1990; Tyson J. A., *Astronom. J.*, **96**, 1, 1988; Heydon-Dumbleton N. H. et al., *M.N.R.A.S.*, **238**, 379, 1989; Lilly S. J., *Astrophys. J.*, **411**, 501, 1993.
- [16] Moffat J. W. & Tatarski D. C., work in progress.
- [17] Weinberg S., *'Gravitation and Cosmology'*, John Wiley & Sons Inc., New York, 1972; Kolb E. W. & Turner M. S., *'The Early Universe'*, Addison-Wesley Publ. Co., Redwood City, 1990; Peebles P. J. E., *Principles of Physical Cosmology*, Princeton University Press, Princeton, 1993.



DARK MATTER AND PARTICLE PHYSICS



PARTICLE DARK MATTER AND ITS LABORATORY DETECTION

Ricardo A. Flores, Department of Physics and Astronomy
University of Missouri–St. Louis, St. Louis, MO 63121, USA.

ABSTRACT

Weakly interacting particle dark matter, if it exists, must have been clumped into galactic halos by gravity. Thus these particles' density and velocities in the solar neighborhood must be several orders of magnitude larger than those of a smooth sea of such particles. This enhancement, which makes their direct detection in the laboratory feasible, would have occurred *even if some of the dark matter in our galaxy consists of sub-solar-mass objects*, as may be indicated by recent detections of microlensing. Here we briefly summarize the main reasons we have to believe that weakly interacting particle dark matter is the most abundant form of matter in the universe, and the methods and prospects for the direct detection of the most commonly discussed particle dark matter candidates: axions and neutralinos.

There is very good evidence that the universe is filled with an invisible form of matter, which we detect by its gravitational pull, but are otherwise unable to identify. We know this dark matter clumps around galaxies¹); it is present in clusters of galaxies²), spread throughout the cluster, not just clumped with the galaxies³); and it seems to be present in the intercluster medium as well, since the distribution and motion of galaxies on very large scales now provides fairly good evidence that the mean density of *matter* in the universe is close to the critical density⁴), which is about an order of magnitude larger than the mean density contributed by clusters. A critical density of matter is about an order of magnitude larger than the maximum density in ordinary (baryonic) matter that is consistent with primordial nucleosynthesis constraints. Thus the need for particle dark matter.

It is widely believed that the very-small-amplitude temperature inhomogeneities detected by COBE⁵) are due to density perturbations present in the early universe. In an expanding universe, regions of slightly enhanced density (above the mean) get gravitationally slowed down in their expansion, eventually ceasing to expand and collapsing to turn into the stable structures we see in the universe at the present. In this gravitational instability scenario⁶), a weakly interacting matter component can naturally allow inhomogeneities to grow denser before recombination without excessively distorting the microwave background (because of its feeble interaction with ordinary matter and radiation). Weakly interacting dark matter would also naturally explain the fact that galactic halos are so much more extended than galaxies (whose gas is able to radiate, and thus cool and contract inside the halo). Thus, it is very likely that particle dark matter is only weakly interacting.

Since the dark matter is gravitationally dominant on all scales larger than galactic cores, it is an essential ingredient⁶) of cosmological models⁶). In the most successful models this particle dark matter is "cold"⁷) * (*i.e.* have negligible random velocities during the expanding phase of even subgalactic structures). Galactic halos form by the collapse of overdense regions early on in such models. The collapse enhances the halo density by about an order

* It is enough that it be dominantly of the cold type on galactic scales only for the argument below to apply. See the next paragraph.

of magnitude relative to the mean density, and remains constant thereafter. Thus, the density of weakly interacting matter around our own galaxy must be well above the mean density by the present. The velocities of the dark matter particles remain high after the collapse in order to keep the halo in dynamical equilibrium. Thus, their characteristic velocity in our own galaxy must be well above those of a smooth sea of relic particles. It is these two features, enhanced velocities and density, that make the task of direct particle dark matter detection feasible. Note, however, that these features do not depend on these specific cosmological models. Even if galaxies had formed by some other mechanism, or even if some of the halo dark matter were baryonic, as suggested by the observations of microlensing reported at this conference⁸⁾, particle dark matter would have been accreted by galaxies as long as it was really there, thus giving it its enhanced density and velocities⁹⁾. Observational parameters allow one to estimate the density and characteristic velocity of dark matter particles to be $\approx 10^{-24}$ g/cm³ and ≈ 300 km/s respectively.

Two particles that have been speculated to exist for reasons entirely outside of cosmology remain as leading candidates to be the cold dark matter: a light axion (mass $\approx 10^{-5}$ eV) and a neutralino in the mass range $O(10 - 1000)$ GeV¹⁰⁾. Each of them requires an entirely different technique in order to detect them directly in laboratory experiments, which we briefly discuss below. A third well motivated candidate is a light neutrino. Not only do we know they exist, but solar neutrino experiments provide a strong hint that neutrinos could have a mass¹¹⁾. Also, cosmological models with a subdominant neutrino dark matter component can successfully account for large scale structure observations with fluctuation amplitudes normalized to the COBE observations¹²⁾. However, because their presence would not significantly alter the abundance of cold dark matter in galactic halos¹³⁾ on the one hand, and there are no planned experiments that could detect them as dark matter on the other, we will not discuss them any further here.

The invisible axion could be all the cold dark matter we ever wanted if it were light enough, and fortunately it would not be entirely invisible. Many years ago, Sikivie¹⁴⁾ pointed out that dark matter axions could be detected directly in the laboratory, despite their tiny mass and very weak coupling to matter, by looking for their conversion into photons in a tuned microwave

cavity of very high Q and in the presence of strong magnetic fields. The current level of sensitivity (a power $O(10^{-21}$ Watt) has only made possible to constrain the axion-photon-photon coupling so far¹⁵⁾. A much more sensitive search is planned for the future by the University of Florida group, and will perhaps achieve the required sensitivity.

Dark matter neutralinos are being searched for in an entirely different fashion. Goodman and Witten and Wasserman¹⁶⁾ first suggested that heavy cold dark matter relics could be searched for by detecting the recoil energy deposited in matter by their elastic collisions with nuclei. With not even a handful of dark matter particles per cc, very small cross sections, and very small recoil energies, the task is very hard indeed. Many techniques have been proposed to detect this recoil energy, with many experiments planned for R&D, and even some that might reach the level of sensitivity required for neutralinos. Mosca nicely summarized all these developments at this conference, therefore I refer the reader to his article in these Proceedings. (An entirely different method, albeit indirect, is to look for byproducts of neutralino annihilations. This was reviewed by Kamionkowski at this conference)

The elastic scattering of neutralinos involves small momentum transfers (unless the particle's mass is significantly larger than 100 GeV) because of their small velocities, thus they scatter coherently off of an entire nucleus in the quantum mechanical sense. However, neutralinos have spin-dependent couplings to the fundamental constituents (quarks), and amplitudes added over an entire nucleus do not add up to much: the total nuclear spin. This is typically referred to as incoherent scattering. Neutralinos also have spin-independent couplings that give amplitudes that add up to a lot more: the total nuclear mass. This is typically referred to as coherent scattering. Which coupling dominates depends on the composition of the neutralino, which even in the simplest of supersymmetric models is a superposition of four states. And adding spins is not as simple as it sounds due to polarization effects. Elaborate nuclear physics calculations are possible for individual target nuclei to reliably estimate nuclear spin amplitudes¹⁷⁾, but more uncertain is the spin contribution within nucleons¹⁸⁾. The complexities add up and I refer the reader to recent literature¹⁹⁾ for the many intricacies and adequate references to earlier literature. We have recently surveyed expected interaction rates for

a large variety of target materials (see the first article of ref. 19), and find rates in the range $O(.1 - .01)$ collisions per kg-day for neutralinos lighter than the W boson, with rates more likely to be at the lower end of the range if future searches for supersymmetry at LEP II fail. An exciting development in the near future will be the experiment planned with the CfPA kg-size Ge detector²⁰⁾, which might reach a level of sensitivity of $.1/\text{kg-day}^{21)}$ and thus be sensitive enough for neutralinos if the pseudoscalar Higgs is light²²⁾.

During the decade past since the methods for detecting axions and massive dark matter particles were proposed, efforts to detect dark matter directly in the laboratory have gathered a lot of momentum. A lot of progress was made in R&D, and even some candidates were eliminated (*e.g.* cosmions, and sneutrinos and heavy neutrinos with help from accelerator experiments). The next decade is likely to bring significantly more progress as detectors of much better sensitivity are built; and with luck, maybe the gold medal itself.

This work was supported by NSF grant PHY-9207026.

REFERENCES

1. For a recent review see *e.g.* K. Ashman, *PASP* **104** (1992) 1109
2. See *e.g.* the article by Mellier in these Proceedings, especially for the evidence from gravitational lensing
3. A. Tyson, F. Valdez and R. Wenk, *ApJ* **349** (1990) L1
4. For a recent review see *e.g.* A. Dekel, to appear in *ARA&A* (astro-ph/9401022)
5. G. F. Smoot *et al.*, *ApJ* **396** (1992) L1
6. See *e.g.* J. Peebles, "Principles of Physical Cosmology", Princeton University Press, 1993
7. See *e.g.* J. Ostriker, *ARA&A* **31** (1993) 689
8. See the articles by Bennett, Cavalier, and Queinnec in these Proceedings
9. See *e.g.* E. Gates and M. Turner, Fermilab preprint (astro-ph/9312008)
10. For a recent discussion on dark matter axions see *e.g.* P. Sikivie, University of Florida preprint UFIFT-HEP-93-26; for a recent discussion on dark matter neutralinos see *e.g.* J. Ellis, CERN preprint CERN-TH.7083/93

11. For a recent review see *e.g.* N. Hata, University of Pennsylvania preprint UPR-0612T (hep-ph/9404329)
12. See R. Nolthenius, A. Klypin and J. Primack, *ApJ* **422** (1994) L45; and references therein
13. J. Ellis and P. Sikivie, *Phys. Lett.* **B321** (1994) 390
14. P. Sikivie, *Phys. Rev. Lett.* **51** (1983) 1415
15. See C. Hagmann *et al.*, *Phys. Rev.* **D42** (1990) 1297; and references therein
16. M. Goodman and E. Witten, *Phys. Rev.* **D31** (1985) 3059; I. Wasserman, *Phys. Rev.* **D33** (1986) 2071
17. See M. Ressel *et al.*, *Phys. Rev.* **D48** (1993) 5519; and references therein
18. See J. Ellis and M. Karliner, CERN preprint CERN-TH.7022/93 (hep-ph/9310272)
19. See J. Ellis and R. Flores, *Nucl. Phys.* **B400** (1993) 25. See also M. Nojiri and M. Drees, *Phys. Rev.* **D48** (1993) 3483, who have pointed out a new spin-independent contribution to neutralino-nucleon scattering. It is important, however, only if the neutralino mass is rather finely tuned to a sfermion mass.
20. See the article by Lu in these Proceedings
21. T. Shutt, private communication
22. J. Ellis and R. Flores, *Phys. Lett.* **B300** (1993) 175

REVIEW of PARTICLE DARK MATTER EXPERIMENTS (DIRECT DETECTION)

presented by Luigi MOSCA

C.E. - SACLAY (DAPNIA/SPP) Laboratory



ABSTRACT

After a brief reminder of the physical motivations for this field of research and of the main experimental challenges that have to be faced, the different types of detectors are reviewed both from the point of view of the development under way and of the physical results.

I. INTRODUCTION

If at least a significant fraction of the Dark Matter in our Galaxy is made of a halo of elementary particles, the problem of their detection becomes of primary importance.

Two restrictions are adopted in this review : only direct detection type experiments and only particles able to transfer at least a few KeVs to some target nucleus are considered. The second condition essentially excludes light neutrinos and axions, while **WIMPs (SIMPs)**, that is : Weakly (Strongly) Interacting Massive Particles, are the main candidates.

The basic process in all these experiments is the elastic scattering of WIMPs (SIMPs) off nuclei of a given absorber.⁽¹⁾

The challenge is twofold because of the **small rate** of these events (at most 0.1 per KeV, per kg of absorber and per day) and the **small recoil energies** to be detected (< 100 KeV).

As a consequence the relevant and common properties of these detectors must be : a sufficiently **low threshold of recoil detection** (a few KeV), a **high sensitivity** and a **very efficient protection against cosmic rays** (laboratories must be deep underground!) and **radioactivity** both external and internal to the detectors.

Three complementary methods allow to fight against radioactivity : i) an **appropriate shielding** (mainly old lead and copper), ii) **selection (purification)** of very low activity **materials** for the set-ups, including the absorbers themselves and possibly iii) **increase of the intrinsic background rejection power of the detectors** by taking advantage of their **different response** to the background (**electrons**) and to the signal (**recoil nuclei**) : namely by a Pulse Shape Discrimination (**PSD**) or/and by the **simultaneous detection of 2 different physical quantities** (heat and ionization or heat and scintillation).

The WIMPS' velocity distribution ($\langle V_{\text{WIMPS}} \rangle \approx 300 \text{ Km/sec}$) is the result of a convolution of the solar system velocity ($\approx 220 \text{ km/sec}$) with a Maxwell-Boltzman gas distribution in a non dissipative halo. The main WIMPs signature would be given by the **annual flux** (and nuclear recoil energy) **modulation** (5 to 10% effect) requiring **large mass** detectors and **very stable** running conditions.

The presently **best WIMP candidate** is the Lightest Supersymmetric Particle (**LSP**), the **neutralino**, which could be a dominant cold Dark Matter relic.

The specific interaction of neutralino with ordinary matter depends not only on nuclear properties, but also on a set of internal parameters of the SUSY theory and includes, a priori, both the **coherent** and the **spin-dependent** couplings.

Let us now review the different types of detectors already used in this field of research or under development.

II. SEMICONDUCTORS : Ge and Si

Since several years Germanium crystals are used, at liquid N₂ temperature, as powerful γ and β detectors, mainly due to their impressive resolution (≈ 1 KeV at 1 MeV). One remarkable application is in the 2β experiments (MeV energy range), because in this case Ge is at once the source and the detector.

Then these detectors have been optimized for Dark Matter (WIMPs) search (energy range of a few KeV). Table 1 summarize the experimental situation from this point of view : here the recent performances of the Heidelberg-Moscow experiment have to be stressed, also if threshold is much higher then in the other quoted experiments.

Limitations of these detectors are the microphony problem at low energies, the electronic noise and, more generally, the technical and financial problems rapidly increasing with the mass of the crystals.

Table 1. Semiconductor Experiments

Detector crystal	Collaboration	Underground Laboratory	Mass of the crystal	Threshold (KeV) (+)	Rate (*) (evts/KeV /kg/day)	Comments (year)
Ge	UCSB/UCB/ /LBL (2)	Oroville (≈ 500 mwe)	900 g	3	≈ 3	(1988)
Ge	Avignone/ /Drukier (3)	Homestake (≈ 4400 mwe)	0.25 to 1 kg	4	≈ 2	(1991)
Ge	Caltech/Psi/ /Neuchatel (4)	St. Gotthard (≈ 3000 mwe)	800 g	2	≈ 3	low threshold (1991)
Ge	Zaragoza/PNL/ /USC (5)	Canfranc tunnel (675 mwe)	234 g	(1.6)	≈ 12	Cosme - 2 detector (1993)
Ge	Heidelberg/ /Moscow (6)	Gran Sasso (≈ 3500 mwe)	2.9 kg	12	0.2	low rate (1994) high threshold
Si	UCSB/UCB/ /LBL/SACLAY (7)	Oroville (≈ 500 mwe)	4 x 17 g	1.1	≈ 60	for Cosmions (high σ and low masses) (1990)

(+) Electron Equivalent Energy

(*) after "microphonic events" subtraction

III. SCINTILLATORS : NaI(Tl), CaF₂(Eu), liquid Xe,

In spite of their poorer resolution compared to Ge detectors, scintillators play at present an important role in this field for several reasons. Scintillation detection is a well known and simple technique (usually no cooling needed), the ratio cost/mass is reasonably low, and there is a wide choice of target nuclei.

In particular some of the scintillators (e.g. NaI(Tl)) have a good light output and a possible (also if difficult at low energy) Pulse Shape Discrimination (PSD) , as well as non zero spin target nuclei (²³Na, ¹²⁷I, ¹⁹F, ¹²⁹Xe,.....) . Other specific reasons of interest are the very high predicted axial coupling for the ¹⁹F nucleus and the high estimated Quenching Factor of liquid Xenon (see below).

The Quenching Factor (QF)⁽⁺⁾ measurements for nuclear recoils is in fact a crucial piece of information in all these studies, allowing to compute the real recoil energy from the measured one (Electron Equivalent Energy). The method used to determine it consists in exposing a reference scintillator to a neutron beam of a few MeV of kinetic energy, then measure the scintillation energy produced by the nuclear recoils (using a known γ source for calibration) and finally compare it with the real recoil energy calculated from the neutron beam energy and its scattering angle. Table 2 summarizes the present results.

Table 2. "Quenching Factor"⁽⁺⁾ of nuclear recoils, measured with Neutron beams

Scintillator	Target Nucleus	Saclay and IN2P3 ⁽⁸⁾ at Bruyères Le Chatel	Imp. College/ ⁽⁹⁾ Oxford/Rutherford	Osaka ⁽¹⁰⁾ (Japan)
NaI(Tl)	²³ Na	0.25 ± 0.02	≈ 0.30	0.4 ± 0.2
	¹²⁷ I	0.08 ± 0.01	≈ 0.08	0.05 ± 0.02
CaF ₂ (Eu)	⁴⁰ Ca	0.049 ± 0.005	0.08 ± 0.01	-----
	¹⁹ F	0.069 ± 0.005	0.12 ± 0.01	-----

- for Liquid Xenon : QF ≈ 0.8 (as estimated from Lindhard theory + recombination effects)
(while for α particles the measured QF ≈ 1.2) (11)

Moreover let us remind that :

- for Semiconductors : QF(Ge) ≈ 0.25 (12); QF(Si) ≈ 0.30 (13)

- for Bolometers (Heat channel) : QF ≈ 1.0 (to be checked at low energy)

(+) The QF is the scintillation efficiency relative to that of electrons of the same kinetic energy.

Concerning the present situation of the scintillator experiments, summarized in Table 3, the low rate near threshold in a 7kg NaI(Tl) crystal and the first results

obtained with a CaF₂(Eu) scintillator have to be noticed (see for more details the talk of P. Belli - BPRS Collaboration - given at this Conference). Also the huge mass of NaI(Tl) crystals available in the Osaka group⁽¹⁰⁾ and the Quenching Factor measurement for several nuclear recoils by the Boulby mine group (see the T.J. Sumner's talk, also given at this Conference) and by the BPRS collaboration⁽⁸⁾ have to be mentioned.

Table 3. Scintillator Experiments

Collaboration	Underground Laboratories	Type of scintillators	Mass	Threshold (KeV) (+)	Rate near threshold (evts/KeV /kg/day)	Resolution σ/E at 60 KeV	Comments
BPRS (Beijing, Paris, Roma, Saclay) (8)	Gran Sasso, Frejus, Mentogou	NaI(Tl)	7 kg (++)	≈ 4	≈ 2	$\approx 7\%$	dedicated expt.(R&D up to 10 kg)
		CaF ₂ (Eu)	370 g	≈ 4	≈ 15	$\approx 12\%$	
Saragoza/PNL/USC (5)	Canfranc tunnel	NaI(Tl)	3x10.7 kg	≈ 6	≈ 14	$\approx 12\%$	-----
Osaka (Japan) (10)	Kamioka mine (≈ 2400 mwe)	NaI(Tl)	36.5 kg *)	≈ 5	≈ 6	$\approx 10\%$	-----
Imp.College/ /Oxford/ (9) /Rutherford	Boulby mine (≈ 3000 mwe)	NaI(Tl) (**)	1 kg	≈ 4	≈ 7	$\approx 10\%$	dedicated expt.
DAMA (Roma) (11)	Gran Sasso	Liq. Xe	3.5 kg	≈ 15 (***)	≈ 70 (***)	$\approx 26\%$	project : enrich ¹²⁹ Xe
Imp.College/ /Oxford/ (9) /Rutherford	Boulby mine	Solid Xe	1 kg	-----	-----	-----	detector just achieved

(+) Electron Equivalent Energy

(++) New (preliminary) result obtained just after this Conference. The previously published result (1992) ⁽¹⁴⁾ was obtained with a 760 g NaI(Tl) crystal, with a Rate near threshold ≈ 7 events/KeV/kg/day.

(*) 17 such modules (≈ 600 kg) have been used to investigate the inelastic channel with $^{127}\text{I}^* \rightarrow \gamma$ (58 KeV). ⁽¹⁵⁾

(**) This group started to investigate a nearly pure NaI at a Temp of ≈ 100 K. ⁽⁹⁾

(***) Note that the Xe quenching factor is especially favourable (≈ 0.8). This result has been obtained with a no-low background prototype and big improvements in this respect are under way. ⁽¹¹⁾

Up to now only the WIMPs search has been considered. Concerning the SIMPs investigation, a preliminary result has been obtained by the BPRS⁽¹⁶⁾ Collaboration on the basis of delayed coincidences between two NaI(Tl) crystals. One candidate event in about 5 months running time was found, which is compatible with the estimated background rate : this allowed to somewhat enlarge the previously excluded regions in the σ vs mass plot both for coherent and spin dependent couplings.

IV. BOLOMETERS : Ge, Al₂O₃, LiF,

From the ancient Greek : $\beta\omicron\lambda\eta'$ (= ray, lit.: something thrown), this term is presently used to designate very sensitive thermometers (including, by extension, the associated heat absorber) operating at very low temperature (≈ 10 mK).

Table 4. Bolometers : present status of the R&D

Colla- -boration	Undergr. Laborat.	Crystal	Mass (g)	Threshold (KeV) (*)	Temp. (mK)	Rate at ≈ 10 KeV evts/KeV /kg/day	Resolution σ/E	Com- -ments
CFPA(Berkeley Stanford/UCSB /INR) Sadoulet (17)	shallow site at Stanford 17 mwe	Ge	≈ 60	≈ 3	24	—	1.3% at 60KeV (0.5 in Ioniz.)	both : - HEAT - IONIZ.
MUNICH Max Plank/Gar- -ching (18)	Gran Sasso (1995 - 96)	Al ₂ O ₃ Sapphire)	31	≈ 0.3	15 44	—	3% at 1.5 KeV 1.5% at 6 KeV ≈ 1 % at 1MeV	only HEAT (WIMPS < 15 GeV)
CEA/IN2P3/ INSU (France) (19)(**)	Fréjus 4800 mwe	Al ₂ O ₃ Ge	24 7 (from Berkeley)	≈ 1.5 ≈ 1.7	55 34	≈ 800 (1st- -run)	2.5% at 60 KeV 2.0 % (1.9% Ionizat.)	only HEAT both:HEAT and IONIZ.
TOKYO (20)	Kamioka	LiF	2.8	≈ 4	12		6% at 60 KeV	only HEAT
Milano Univ. Fiorini et al. (21)	Gran Sasso 3400 mwe	TeO ₂ " LiF NaF CdWO ₄ CaF ₂	334 73 100 30 60 2	—	≈ 25	—	0.4 % at 1 MeV 1 % at 60 KeV 5 % at 1 MeV 2 % at " 0.2 % at " 2 % at "	for $\beta\beta$ " for DarkMat for DarkMat HEAT+SC? HEAT+SC.

(*) Real Energy (Heat channel) (provided that the QF ≈ 1.0 down to low energies).

(**) The Lyon group is starting a study for HEAT + SCINTILLATION with CaF₂ crystals.

Why using bolometers to detect WIMPs ? As discussed in the introduction the main challenge is the suppression and/or rejection of the radioactive background. To do that , the maximum information is highly desirable, while in usual detectors the heat, which represents a large fraction ($\approx 2/3$ in Silicium) of the deposited energy, is generally lost. In addition the quanta of thermal energy (phonons) are very tiny (10^{-4} - 10^{-5} eV) so a potentially very low threshold and very good resolution are expected. Moreover the possibility of measuring simultaneously another form of deposited energy (ionization or scintillation) made these detectors especially adapted to background rejection.

Since ≈ 1984 (Fiorini, Coron, ...) the problem of detecting the heat deposited by individual elementary particles is basically solved (remind that 1 KeV of deposited energy $\approx 1.6 \times 10^{-16}$ Joule!). The problem now is to minimize microphony, electronic noise, radioactive background at low energy and reach long time (years) stable operating conditions with large masses of crystals (at least several kgs).

The present status of the R&D is summarized in Table 4. Concerning heat detection the most impressive results (threshold and resolution) have been obtained by the Munich group (see G.Forster talk at this Conference). Concerning the simultaneous measurement of heat and ionization, the first partially successful result has been obtained by the Berkeley group (see A.Lu talk at this Conference), while the French group has recently well improved the stability of the ionization channel (a crucial point!). On the other hand both the Milano and Lyon groups are developing scintillation detection at mK temperatures.

V. TRACK PROJECTION CHAMBERS (TPC)

A TPC at low pressure (10 to 20 Torr) has been tested by the Saclay⁽²²⁾ group ⁽²²⁾with Ethane gas. This type of detector has several good properties (measurable proton range : about 1 cm at 2 KeV in C_2H_6 at 10 Torr, directionality and sensitivity to axial coupling) making it well adapted to investigate low mass WIMPs (< 20 GeV), but unfortunately also one big drawback : to get , e.g., 2 kg of H_2 at 20 Torr it would need a TPC of 1000 m^3 !

The Grenoble, Münster, Neuchâtel, Padova, Zürich Collaboration⁽²³⁾ are studying an high pressure TPC (≈ 5 bars) with ^{19}F (≈ 18 kg of CF_4) and Xe target nuclei. But this detector has no directionality : in fact it was mainly proposed for ν magnetic moment and 2β neutrinoless investigations , so that the neutralino search is a byproduct. In addition Ypsilantis et al. proposed a WIMPs search as a byproduct for Hellaz, the high pressure Helium + Xenon TPC dedicated to Solar Neutrino detection. Finally the San Diego group⁽²⁴⁾ is still working on another low pressure TPC

VI. OUTLOOK

Some examples of present experimental projects for a better WIMPs investigation are :

a) the Heidelberg-Moscow Collaboration is preparing a new detector with a ^{73}Ge enriched crystal.(6)

b) a NaI(Tl) detector , of about 100 kg , which is being installed in the Gran Sasso Laboratory , by the BPRS Collaboration , mainly to study the annual modulation effect.(8)

c) the replacement of the vessel (low activity) and of natural liquid Xe by another one enriched in ^{129}Xe to increase the spin dependent sensitivity (DAMA group at Gran Sasso Laboratory).(11)

d) the CfPA group is preparing an "ice box" low activity cryostat to be installed in a shallow site at Stanford as a test bench for future R&D.(17)

e) the Munich group is preparing a set of 4 x 250 g Sapphire (Al_2O_3) bolometers to be installed in Gran Sasso Laboratory by the end of 1996. They are focussing on phonons detection only and low mass WIMPs ($< 20 \text{ GeV}$).⁽¹⁸⁾

f) the French group is preparing a new bolometer installation (EDELWEISS) in the Fréjus Laboratory (LSM) with a low radioactivity dedicated cryostat and a new type of sensors (thin Niobium/Silicium films).⁽¹⁹⁾

VII. CONCLUSIONS

i) As discussed throughout this paper , **there is no "universal" detector for WIMPs!**

ii) The **"reference"** detectors are the **classical Ge crystals** operating at **LN2 temperature**.

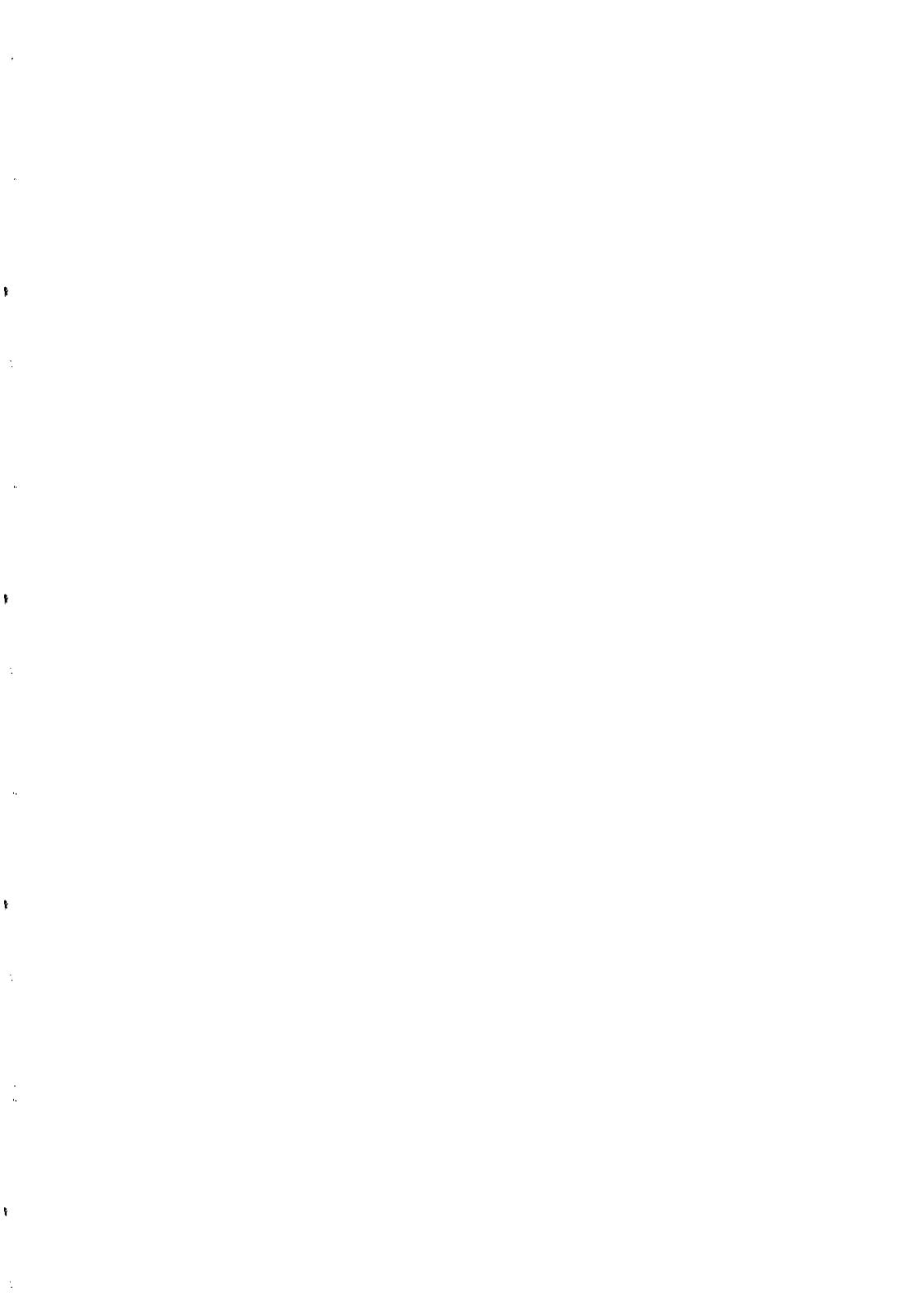
iii) **Scintillators** have proved to be **very useful and realistic** detectors for WIMPs search , mainly in view of an investigation of the **annual modulation effect**. Further efforts to still reduce and/or reject residual background are in progress.

iv) **Bolometers** represent the **most ambitious** developments, in particular by measuring **2 physical quantities** : Heat + Ionization or Heat + Scintillation. The remaining problems are a big challenge for the concerned groups.

REFERENCES

1. M.W. GOODMAN and E. WITTEN, Phys.Rev.D, 31(1985) 3059
2. D.O. CALDWELL et al, Phys.Rev.Lett., 61(1988) 510
3. A.K. DRUKIER et al, Nucl. Phys. B(Proc.Suppl.) 28A (1992) TAUP'91 (September 1991)
4. D. REUSSER et al., Phys. Lett.B, 255 (1991) 143
5. M.L. SARSA (Zaragoza/PNL/USC Collab.) TAUP'93 (Gran Sasso)

6. M. BECK (Heidelberg-Moscow Collab.) Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ. (Roma II) June 1993
7. D.O. CALDWELL et al. Phys Rev. Lett. 65 (1990) 1305
8. P. BELLI (BPRS Collab.) talk given at this Conference ; C.BACCI et al."Dark Matter Search with Calcium Floride Crystals" accepted by Astroparticle Phys. and E. GAILLARD-LECANU Doctor Thesis "DéTECTEURS à cristaux scintillants pour la recherche de Matière Noire non-baryonique." June 1st, 1994
9. G.J. DAVIES et al., Phys. Lett. B322 (1994)159 ; N.J.C. SPOONER and P.F. SMITH, Phys. Lett. B314 (1993) 430 and T.J. SUMNER (Imp. College/Oxford/Rutherford Collab.) talk given at this Conference.
10. K. FUSHIMI et al. , Phys. Rev.C, 47 (1993) R425
11. P. BELLI, "A first result on Axial Coupling WIMPs with Xenon Target at Gran Sasso" TAUP'93 (to appear on Nucl. Phys.B)
12. EDELWEISS Collab.(IPN Lyon, IN2P3, SACLAY, CdF, IAP, CSNSM) "Calibration of a Ge crystal with Ge nuclear recoils for the development of a Dark Matter detector" (to be published)
13. G. GERBIER et al., Phys. Rev.D, 42 (1990) 3211
14. C. BACCI et al. Phys Lett.B293(1992)460
15. H. EJIRI et al. "Search for Spin Coupled Dark Matters by studying Axial Vector Excitation of Nuclei (Draft 1993).
16. C. BACCI et al. "A search for Strongly Interacting Massive Particles with NaI scintillators" paper accepted by Astroparticle Physics.
17. A. LU (CfPA Collab.) talk given at this Conference.
18. G. FORSTER (Munich/Max Plank, Garching/ Collab.) talk given at this Conference.
19. P. de MARCILLAC : "Underground measurements at the Fréjus Tunnel in the 3 KeV- 5MeV energy range with a massive bolometer." TAUP'93 (September 1993) Gran Sasso; and D. YVON : "Bolometer development with simultaneous measurement of Heat and Ionization signals at Saclay." 5th Int. Workshop on Low Temperature Detectors for Neutrino and Dark Matter (LTD5) Berkeley, July 29th -August 3rd, 1993.
20. M. MINOWA et al., Nucl. Instr. and Methods in Physics Research, A327(1993) 612.
21. M. PAVAN (MILANO Univ. and L.N.G.S.) "Development of low temperature thermal detectors for Dark Matter : first measurements with different types of crystals". Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ. (Roma II) June 1993.
22. O. BESIDA "A time projection Chamber as a Dark Matter Detector" Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ. June 1993; and Doctor Thesis : "Etude d'un détecteur gazeux à ionisation pour la recherche de la Matière Noire de l'Univers." March, 1994
23. J.L. VUILLEUMIER - MUNU (Grenoble, Münster, Neuchâtel, Padova, Zurich Collab.). Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ. (Roma II) June 1993.
24. K. BUCKLAND (San Diego) Berkeley workshop "Strategies for the detection of Dark Matter Particles." February, 1994.



NEW LIMITS ON AXIAL COUPLED WIMPs WITH SCINTILLATORS

Presented by Pierluigi Belli

for the BPRS collaboration

C. Bacci^a, P. Belli^b, R. Bernabei^b, C.J. Dai^c, W. Di Nicolantonio^d, L.K. Ding^c, E. Gaillard-Lecanu^e, G. Gerbier^e, Y. Giraud-Heraud^f, H.H. Kuang^c, A. Incicchitti^g, J. Mallet^e, L. Mosca^e, D. Prosperi^g, C. Tao^f

^a Dip. di Fisica, Universita' di Roma III and INFN, sez. Roma, I-00165 Rome, Italy

^b Dip. di Fisica, Universita' di Roma "Tor Vergata" and INFN, sez. Roma II, I-00173 Rome, Italy

^c IHEP, Chinese Academy, P.O. Box 918/3, Beijing 100039, China

^d Laboratori Nazionali del Gran Sasso (LNGS), S.S. 17/bis Km 18+910, I-67010 L'Aquila, Italy

^e DSM/DAPNIA/SPP, C.E. Saclay, F-91191 Gif-sur-Yvette, France

^f Laboratoire de Physique Corpusculaire, College de France/IN2P3 (CNRS), 11, place M. Berthelot, F-75231 Paris, France

^g Dip. di Fisica, Universita' di Roma "La Sapienza" and INFN, sez. Roma, I-00165 Rome, Italy

and

B. Chambon, V. Chazal, M. De Jesus, D. Drain, Y. Messous, C. Pastor

IPN Lyon, IN2P3 (CNRS), 43 Bd du 11 Novembre 1918, F-69622 Villeurbanne CEDEX, France

Abstract

New results on dark matter direct search with scintillators are presented and compared with previous measurements.

Several materials have been proposed and used to direct search particle dark matter candidates ^{1 to 7}). Our collaboration pointed out its attention to scintillators as target-detectors and, in particular, to NaI(Tl) and CaF₂(Eu). The NaI(Tl) has an high light output and it is well suitable to build large apparatus looking for the annual modulation of the rate for extremely low rate candidates as the neutralino. The CaF₂(Eu) is interesting because of the large cross section on ¹⁹F of axial coupled WIMPs ⁸).

The "quenching factor" of a nucleus recoil (its scintillation efficiency compared to that of an electron with the same kinetic energy) has been measured by our collaboration at Bruyeres Le Chatel with a neutron beam with NaI(Tl) and CaF₂(Eu), obtaining: Q_{Na}=0.25 ⁹), Q_I=0.08, Q_F=0.07 ¹⁰) and Q_{Ca}=0.05 ¹⁰) (see fig. 1). The feasibility of a pulse shape statistical rejection - at low energy - of the γ/e background in NaI(Tl) has been also demonstrated ^{9,14}).

A large number of measurements has been performed in Gran Sasso and Frejus

Laboratories and Mentougou mine with various detectors of different size and composition, made by European, American and Chinese companies.

A first experiment was done with a 760g NaI(Tl) crystal ⁵⁾; in fig. 2 the obtained exclusion plots on I and Na are shown. This result was the first demonstration of the feasibility of a dark matter experiment with scintillators; successively other groups reached similar sensitivities with NaI(Tl) of various size and origin^{11,12,13).}

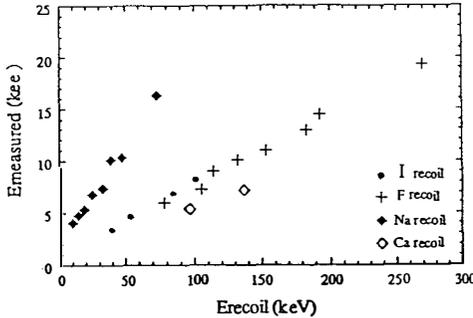


Fig. 1 : Response of NaI(Tl) to Na⁹⁾ and I recoils and of CaF₂(Eu) to F¹⁰⁾ and Ca¹⁰⁾ recoils.

We evaluated the expected rates for the neutralino as a dark matter particle and reported them in ref. ⁶⁾. These calculation showed that we need to improve by at least one order of magnitude the detector quality to reach the most optimistic neutralino allowed region.

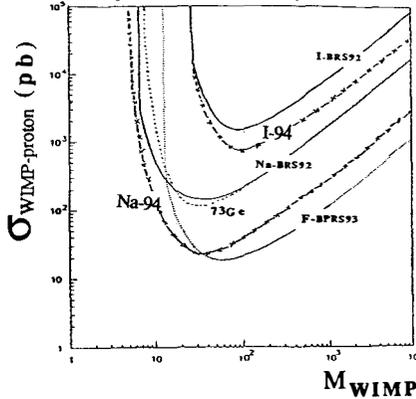


Fig. 2: Exclusion plots for axially coupled WIMPs derived from our measurements: Na and I nuclei 1992 ⁵⁾; new results for Na and I nuclei 1994; first results on F 1993. The exclusion plot on ⁷³Ge - derived from the data of ref. ⁴⁾ - is shown for comparison.

To achieve this result, we carefully measured and selected all the detector components which were used in new detectors. Thorn-EMI provided us an improved low

radioactive background photomultiplier which contributes to the experimental rate a factor 15 to 50 less than previous ones. Detectors with strongly reduced natural impurities have been obtained: $^{238}\text{U}/^{232}\text{Th} \sim 2$ ppt and $^{39}\text{K} < 0.05$ ppm (2σ CL) ¹⁴.

Recently⁺⁾ a new set of four detectors, 7 Kg each, have been measured; three of them were used as anticoincidence for the inner one (reduction rate of about 15 %). Each detector was seen by one PMT on each end through a low activity light guide; a low Z window allowed us a calibration with ^{55}Fe source (5.9 keV X-rays), see fig.3.

A preliminary result - obtained just after underground storage of the detector and with a not fully efficient radon removal - is show in fig. 4; the corresponding exclusion plot is reported also in fig.1. The counting rate is expected to still more decrease by improving the radon removal system and the nearest materials quality and waiting for the decreasing of the short life contaminants. However, we have already reached with these very preliminary measurements an improvement in the exclusion plot of a factor ~ 5 .

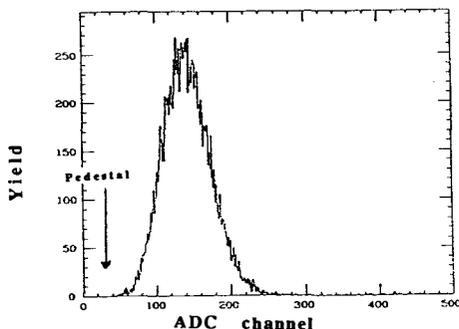


Fig. 3: The ^{55}Fe source as seen - through a low Z window - by the 7 Kg NaI(Tl) detector having a straight light guide and a low activity photomultiplier on each side.

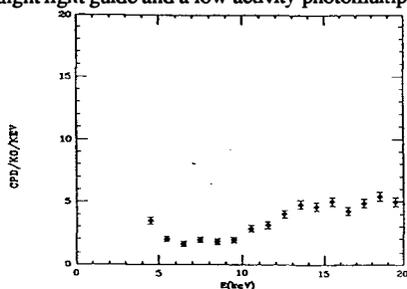


Fig. 4: Preliminary experimental rate of the 7 Kg NaI(Tl) detector at beginning of underground storage. The radon removal and the environmental background are not yet optimized.

⁺⁾ In fact, these results from the 7 Kg detectors have been presented about one month later at Berkeley workshop, for which no proceedings were planned.

A similar sensitivity on axial coupled WIMPs have been also obtained with a 0.37 kg $\text{CaF}_2(\text{Eu})^{10}$ crystal directly coupled on one end to a very low activity EMI photomultiplier. At the other side to a low activity NaI(Tl) was installed as anti-Compton detector (rate reduction $\sim 30\%$ at low energy). The detector has been calibrated with the ^{55}Fe source. From the high energy spectrum we derive a contamination of 0.05 ppb for ^{238}U and 0.16 ppb for ^{232}Th .

The quoted exclusion plots have been evaluated under the following assumptions: the WIMPS are distributed in a spherical halo with a local density in the solar neighbourhood of $0.3 \text{ GeV}/\text{cm}^3$; their velocities follow a Maxwellian distribution in the Galactic frame (truncated at a value V_{esc}) with an average velocity V_{rms} equal to $\sqrt{3/2} V_c$, where $V_c \sim 220 \text{ km/s}$ is the circular solar velocity. Transformation to the laboratory frame is computed with an Earth velocity equal to about 235 km/s . In the calculation of the expected rates, we have included the experimental energy resolution, the spin factors as calculated in the odd group model by Ellis and Flores⁸⁾ and a spin form factor with a Bessel function shape as for the mass form factor of Engel (1991) in reference¹⁵⁾. Ellis and Flores (1991) suggested a correction for the spin radius with respect to the charge radius which they tabulated in ref⁸⁾ for different nuclei. This is taken into account. For ^{73}Ge , the data are from ref.⁴⁾; we should point out that an additional source of difficulty in comparing these data with other nuclei, arises from the fact that its unpaired nucleon is a neutron, while for the other 3 nuclei, the unpaired nucleon is a proton.

We would like to stress the important role played by the form factor correction in the case of heavy nuclei ($A \geq 40$).

In conclusion, we presented here new results on CDM search with scintillators. We plan to improve the anti-Compton rejection and to further work on purification possibilities of the detector materials.

Acknowledgements

It is a pleasure to thank Dr. C. Arpesella for valuable support, Mr A. Bussolotti for technical help and the LNGS and LSM staff for support.

References

1. P. F. Smith and J.D. Lewin, *Phys. Rep.* 187 (1990), 203.
2. S. P. Ahlen et al., *Phys. Lett. B* 195 (1987), 603.
3. D. O. Caldwell et al., *Phys. Rev. Lett.* 61 (1988), 510.
4. D. Reusser et al., *Phys. Lett. B* 255 (1991), 143.
5. C. Bacci et al., *Phys. Lett. B* 293 (1992), 460.
6. A. Bottino et al., *Phys. Lett. B* 295 (1992), 330.
7. P. Belli et al., *Nucl. Instr. & Meth.* A 336 (1993), 336.
8. J. Ellis and R.A. Flores, *Phys. Lett. B* 263 (1991), 259.
9. G. Gerbier et al., *Proceed. of the Moriond Workshop*, Ed. Frontieres (1991).
10. C. Bacci et al., pre-print LNGS 93/83 and LNGS 94/88, to appear on *Astrop. Phys.*
11. K. Fushimi et al., *Phys. Rev. C* 47 (1993) R425.
12. A. Morales et. al., to appear in the vol. "The dark side.", World Sc. ed.
13. T. Sumner et al., these *Proceed.*
14. G. Gerbier et al., to appear in the vol. "The dark side.", World Sc. ed.
15. J. Engel, *Phys. Lett. B* 264 (1991), 114. Engel, Pittel, Vogel, *Int. J. Mod. Phys. E* 1 (1992) 1.

PRELIMINARY MEASUREMENTS ON AXIAL COUPLED WIMPS WITH XENON TARGET AT GRAN SASSO

C. Bacci^a, P. Belli^b, R. Bernabei^b, C.J. Dai^c, W. Di Nicolantonio^d, A. Incicchitti^e, D. Proserpi^e

^a Dip. di Fisica, Universita' di Roma III and INFN, sez. Roma, I-00165 Rome, Italy

^b Dip. di Fisica, Universita' di Roma "Tor Vergata" and INFN, sez. Roma II, I-00173 Rome, Italy

^c IHEP, Chinese Academy, P.O. Box 918/3, Beijing 100039, China

^d Laboratori Nazionali del Gran Sasso (LNGS), S.S. 17/bis Km 18+910, I-67010 L'Aquila, Italy

^e Dip. di Fisica, Universita' di Roma "La Sapienza" and INFN, sez. Roma, I-00165 Rome, Italy

ABSTRACT

Preliminary result on dark matter direct search with a liquid xenon prototype are summarized. A low activity detector is under installation.

Our investigation on direct detection of dark matter search with scintillators ^{1,6-13}) includes the study of Xenon target-detectors ¹⁻⁷). Xenon is a very interesting material for this search, allowing to investigate both vector and axial coupling interactions by comparing results obtained using same set-up with Xenon enriched in even and odd spin isotopes. Furthermore, large mass detectors (to look for annual modulation signature or for candidates with extremely low expected rate) with well reduced dimensions can be realized, while the scintillation properties of the Xenon in gaseous phase allow also to evaluate and subtract most of the residual environmental background.

A ~ 3.5 Kg LXe target-scintillator, whose performances have been described in detail elsewhere ⁵) runned at Gran Sasso National Laboratory ¹⁴); the detector was calibrated at low energy with a ¹⁰⁹Cd source, giving: $\sigma/E=0.05 + 1.63/\sqrt{E(\text{keV})}$ ⁵). The detector efficiency and the response versus the released energy in a WIMP-nucleus elastic scattering, i.e. the quenching factor, have been studied and discussed in ref. ⁵). During the measurements presented here, the detector was shielded by 4 cm of steel and 10 cm of lead. Data have been collected both in the low and high energy range. The high and intermediate energy spectra allowed us to evaluate the U/Th and K contaminations ⁵).

As we already pointed out ⁵) the detector was not selected for low activity, therefore we measured independently the environmental background contribution to the low energy spectrum (15 and 50 keV) by filling the detector with Xenon gas at 2.5 bar, obtaining a new scintillator with much lower xenon mass ⁷). We subtracted the background - by the procedure described elsewhere ⁷) - mainly due to environmental γ rays from the detector walls, being most of γ s coming from outside shielded by the external lead. The residual spectrum can be ascribed only to eventual WIMP-nucleus scatterings or contaminants present in the xenon, whose rate depends on the mass, like e.g. ⁸⁵Kr and radon.

In Fig. 1a) the residual counting rate is shown, while in fig. 1b) the preliminary exclusion plot (2σ C.L.) for spin dependent interaction is reported ⁷), considering the cross sections that WIMPs would have on protons. Spin factors ¹⁵), nuclear form factors

16), energy resolution and detector efficiency have been considered. New measurements of the quenching factor are foreseen to properly verify the evaluation (Lindhard theory 17) corrected for full recombination effect) used here 5); we have to stress the relevance to perform these measurements with same or identical set-up being the light response in LXe strongly dependent on gas purity, UV transmitters and thermodynamical operating conditions, as pointed out also by many authors.

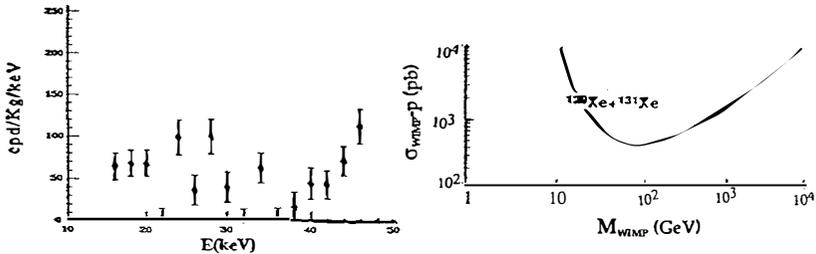


Fig. 1 - a) Residual counting rate; b) exclusion plot for axial coupled WIMPs at 2σ C.L.(see text); for comparison with other nuclei see ref. 18).

We are installing in Gran Sasso a new low activity detector together with an improved shield; we will use natural and enriched (once in odd and once in even spin isotopes) Kr-free xenon gas. Measurements are foreseen looking for annual modulation signature 6).

ACKNOWLEDGMENTS

We gratefully acknowledge Prof. S. d'Angelo for helpful discussions. We wish also to thank Mr. A. Bussolotti for his technical support and LNGS staff for the efficient assistance.

REFERENCES

1. P. Belli et al., *Il Nuovo Cimento* 103A (1990) 767.
2. P. Belli et al., *Nucl. Instr. & Meth.* A299 (1990) 191.
3. P. Belli et al., *Nucl. Instr. & Meth.* A310 (1991) 150.
4. P. Belli et al., *Nucl. Instr. & Meth.* A316 (1992) 55.
5. P. Belli et al., *Nucl. Instr. & Meth.* A336 (1993) 336.
6. P. Belli et al., ROM2F93-42, to appear in the vol. "The dark side..", World Sc. ed.
7. C. Bacci et al., LNGS-93/87 to appear on *Nucl. Phys.*
8. P. Belli et al., *Il Nuovo Cimento* C15 (1992) 475.
9. C. Bacci et al., *Phys. Lett.* B293 (1992) 460.
10. A. Bottino et al., *Phys. Lett.* B295 (1992) 330.
11. C. Bacci et al., LNGS-93/83 (1993) and LNGS-94/88, to appear on *Astrop. Phys.*
12. C. Bacci et al., *Astrop. Phys.* 2 (1994), 13.
13. C. Bacci et al., to appear in the vol. "The dark side..", World Sc. ed.
14. E. Bellotti, *Nucl. Instr. & Meth.* A264 (1988), 1.
15. J. Ellis and R.A. Flores, *Phys. Lett.* B263 (1991).
16. J. Engel, *Phys. Lett.* B264 (1991) 114.
17. J. Lindhard et al., *Mat-Fys. Medd.* 33 (1963) 10.
18. C. Bacci et al., these Proceed.

LATEST LIMITS ON COLD DARK MATTER INTERACTION RATES FROM THE UK UNDERGROUND NaI DETECTOR

T.J. Sumner¹, S.M. Grant¹, J.P. Li¹, J.J. Quenby¹, A. Bewick¹, D. Shaul¹, N.J.T. Smith¹,
W.G. Jones¹, G. Davies¹, C.C. Zammit¹, A.D. Caplin¹, R.A. Stradling¹, C. Lally¹, T. Ali¹,
P.F. Smith², G.J. Homer², G.T.J. Arnison², J.D. Lewin², G.J. Alner², M.v.den Putte^{1,2}, V.
Edwards², M.J. Lea³, P. Stefanyi³, J. Saunders³, N.J.C. Spooner⁴, J.C. Barton⁵ and P.R.
Blake⁶.

¹Blackett Laboratory, Imperial College of Science, Technology and Medicine London SW7 2BZ, UK

²Particle Physics, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, UK

³Physics Department, Royal Holloway & Bedford New College, Egham, Surrey, TW20 0EX, UK

⁴Astrophysics, Nuclear Physics Laboratory, Keble Road, Oxford, OX1 3RH UK

⁵Physics Department, Birkbeck College, Malet Street, London, WC1E 7HX, UK

⁶Physics Department, University of Nottingham, University Park, Nottingham, NG7 2RD, UK



ABSTRACT

New improved limits have now been obtained for the spin-dependent interaction cross-sections in NaI for weakly interacting dark matter particles assuming they are responsible for the dark matter halo of our Galaxy. The underground facility and experimental configuration will be described together with data extraction techniques. The background rate in a 1 kg NaI(Tl) crystal scintillator of low radioactivity construction is reported after exposure in the underground facility. A level of ~ 7 events/kg.day.keV at 4 keV (equivalent photon energy) provides an improved limit for the possible spin-dependent weakly interacting massive particle (WIMP) cross section for elastic scattering from NaI nuclei. A minimum value of $1.5 \cdot 10^{-33} \text{ cm}^2$ is obtained for a 25 GeV WIMP mass. Our results are compared in a consistent model independent way with previous data.

Introduction

First results are presented from room temperature NaI scintillator experiments being carried out by the UK dark matter consortium in a deep underground laboratory in Boulby, U.K. The laboratory is at a depth of 1100 m (3000 m water equivalent) and contains two shielded enclosures. One of these is a 6 m diameter tank filled with recirculated pure water to a depth of 6 m into which experiments can be lowered. The second is a lead and copper castle constructed from materials which have been stored underground for some time. The data used here were acquired from a 1.27 kg NaI(Tl) crystal operating in the castle. The non-zero nuclear spin of both Na and I make such a crystal a useful target for Weakly Interacting Massive Particles (WIMPs) assuming they couple predominantly via an axial spin-dependent scattering process. WIMPs, it has been postulated, may account for the missing matter within the Galaxy, and if this is indeed the case, then experiments of the type reported here currently provide the best limits to their spin-dependent scattering cross-sections and should eventually through improved background suppression lead to direct positive detections. Here the measured background rate in the NaI(Tl) detector has a marginally improved threshold over other experiments and new upper limits for the spin-dependent neutralino elastic scattering cross-sections are derived. Our results are compared in a consistent *model independent* way with previous data.

Experimental details and observations

The shielded 'castle' comprised 10 cm Pb with 10 cm of Cu inside. Observations were carried out with a 1.27 kg NaI(Tl) crystal observed by two EMI 9265A photomultipliers whose 3" diameter matched that of the crystal. The photomultipliers viewed at opposite crystal faces through 30 cm water light guides. Precautions to reduce the contribution to the crystal counting rate from phototube radioactivity via external scattering included use of a C101 copper shield, 3 cm thick, surrounding the crystal. The crystals were prepared using low activity material. Encapsulation was in PTFE with a Cu outer jacket, the window was made of Spectrosil and the glue employed was Epotek, all of known low activity.

Pulse shape recording of the coincident output of the two photomultipliers was performed via a LeCroy 9430 oscilloscope and an analogue pulse height distribution was captured in parallel. The system was calibrated with ^{137}Cs and ^{60}Co sources. An electronic noise threshold corresponding to 2.93 keV was set before waveform capture. The run reported here was of $9.6 \cdot 10^5$ s live time duration, starting on July 29th 1993, and it was possible to scan the pulse profiles of each individual event with an energy loss less than 80 keV. In this manner, spurious electrical noise interference and some forms of spurious photomultiplier pulses could be easily rejected. At this stage no further rise time or pulse shape discrimination was attempted.

Figure 1 shows the counting rate obtained in the high energy part of the background spectrum. The continuum region from 750 keV to 900 keV can be used to put a limit of 0.9 ppm on the internal ^{40}K level from the expected beta-decay spectrum. The photoelectric peak at the 1.4 MeV ^{40}K line gives a 50% higher value and this indicates the presence of an external component. The spectrum obtained at lower energies is shown in figure 2. The peak at 47 keV could be due to ^{210}Pb gamma ray production due to contamination, possibly within the NaI(Tl). Cosmogenic activity due to sea-level exposure to secondary cosmic rays can also produce short lifetime peaks from ^{129}I at the 10 events/kg.day.keV level¹⁾. An average event rate of ~ 7 events/kg.day.keV was achieved at 4 keV equivalent photon energy loss in the crystal. The photoelectron yield from the assembly was 1.6 photoelectrons/keV. The data were binned into channels 0.5 keV wide.

Data analysis and interpretation

To interpret the data of figure 2 and to compare with other work, we adopted a procedure for evaluating the upper limit to a spin-dependent cross section which is virtually independent of the particle physics involved in the scattering process, as there remains significant theoretical uncertainty in the details. We concentrated on the spin-dependent cross section since this is where our work becomes significant. For a given WIMP mass we first predicted the expected form for a measured differential recoil spectrum assuming the WIMPs constitute all the necessary Galactic dark halo mass. The predicted recoil spectrum was then scaled so that it did not exceed the 2σ error bars on the actual measured background spectrum on more than three consecutive data bins²⁾. The maximum allowed interaction rate was given by integrating the scaled predicted spectrum over all energies and an effective cross section was then derived from this rate.

To predict the measured recoil spectrum we started with a Maxwellian dark matter velocity distribution, with $kT = E_0$, modified for the motion of the Earth through the dark matter WIMP population and the Galactic escape velocity cut-off. This gave the WIMP flux passing through the detector. The expected measured differential recoil spectrum in an ideal detector, assuming isotropic centre of mass scattering is then

$$\frac{\partial R}{\partial E_V} = \frac{1}{f'(E_r)E_r + f(E_r)} \frac{R_T}{A} \exp -\frac{E_r}{A} \quad (1)$$

where $A = E_0 r E_c / (E_c + E_0 r)$, r is a kinetic factor which depends on the target nucleus mass, M_T and the WIMP particle mass, M_D , such that $r = \frac{4 M_T M_D}{(M_T + M_D)^2}$, f is the ratio of measured

energy to recoil energy, E_r , i.e. $f = E_V/E_r$, and E_c arises due to a form factor correction. This form factor correction was treated in a relatively simple way at this stage, namely³⁾, with $|F(E_r)|^2 = \exp -E_r/E_c$. Values of E_c for axial spin-dependent coupling were taken from Ellis & Flores⁴⁾. R_T is the total event rate expected in the crystal integrated down to zero energy.

The finite detector resolution, including the Poissonian statistics of the photoelectron generation, was taken account of by convolving the resolution with equation 1. Once this was done we deduced an upper limit to the total interaction rate R_T which was converted into an effective cross-section for each individual target nuclear species, of mass M_T , by

$$\sigma_T = \frac{R_T M_D M_T}{\rho \bar{v}_D} \quad (2)$$

For NaI(Tl) we treated the nuclei of ^{23}Na and ^{127}I separately. The quenching factor, f , is different for each and appears to be independent of energy in the range of interest^{5,6,7)}. We used values of $f_{\text{Na}} = 0.25$ and $f_I = 0.07$. For comparison with Ge results, we used the energy dependent form of f derived from the observations of Sattler et al.⁸⁾, which are lower than those predicted by the commonly used theory of Lindhardt et al.⁹⁾. However this turned out to have a relatively small effect.

The method of Reusser et al.²⁾ was used to establish the exclusion limits on the effective total cross-section and we show the results of this in figure 3. Data from other experiments were treated similarly and are displayed in figure 3. Our results are seen to provide the best cross-section limits for axial-coupling when displayed in this model-independent way. Results from a 6 kg NaI(Tl) crystal will be published separately¹²⁾.

Acknowledgements

SERC funding and studentships (SG, DS, GD, CL) are acknowledged. The ICSTM STARLINK node was used for data analysis.

REFERENCES

- 1) S. Battersby & C.S. Dyer, *Private Communication*.
- 2) D. Reusser et al., Phys. Lett. B. 255 (1991) 143.
- 3) A. Gould, Ap. J. 321 (1987) 571.
- 4) J. Ellis & R.A. Flores, Phys. Lett. B. 263 (1991) 259.
- 5) K. Fushimi et al., Phys. Rev. C 47 (1993) R425.
- 6) C. Bacci et al., Phys. Lett. B 293 (1992) 460.
- 7) N.J.C. Spooner et al., Phys. Lett. B 321 (1994) 156.
- 8) A.R. Sattler, F.L. Vook & J.M. Palms, Phys. Rev. 143 (1966) 588.
- 9) J. Lindhardt et al., Mat-Fys. Medd. 33 (1963) 10.
- 10) C. Bacci et al., LNGS preprint 93/83 (1993).
- 11) D.O. Caldwell et al., Phys. Rev. Lett. 61 (1988) 510.
- 12) J.J. Quenby et al., Phys. Lett. B Submitted (1994).

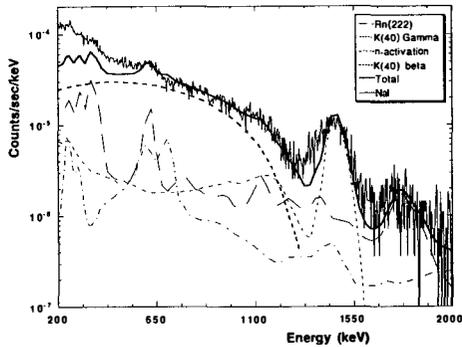


Figure 1: High energy pulse height spectrum from a 1.27 kg NaI(Tl) crystal.

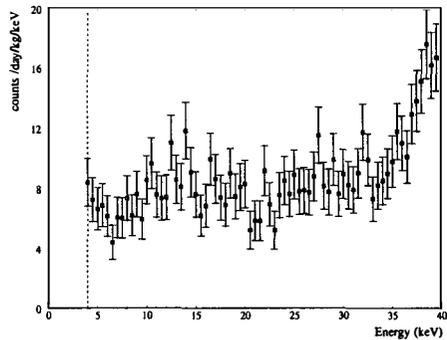


Figure 2: Low energy pulse height spectrum from a 1.27 kg NaI(Tl) crystal. The dashed vertical line is positioned at the lowest energy at which we remain confident of the validity of our pulse scanning technique.

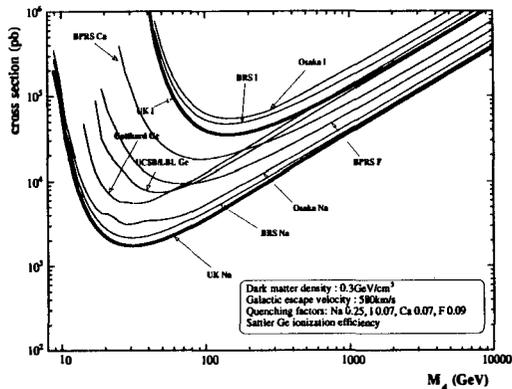
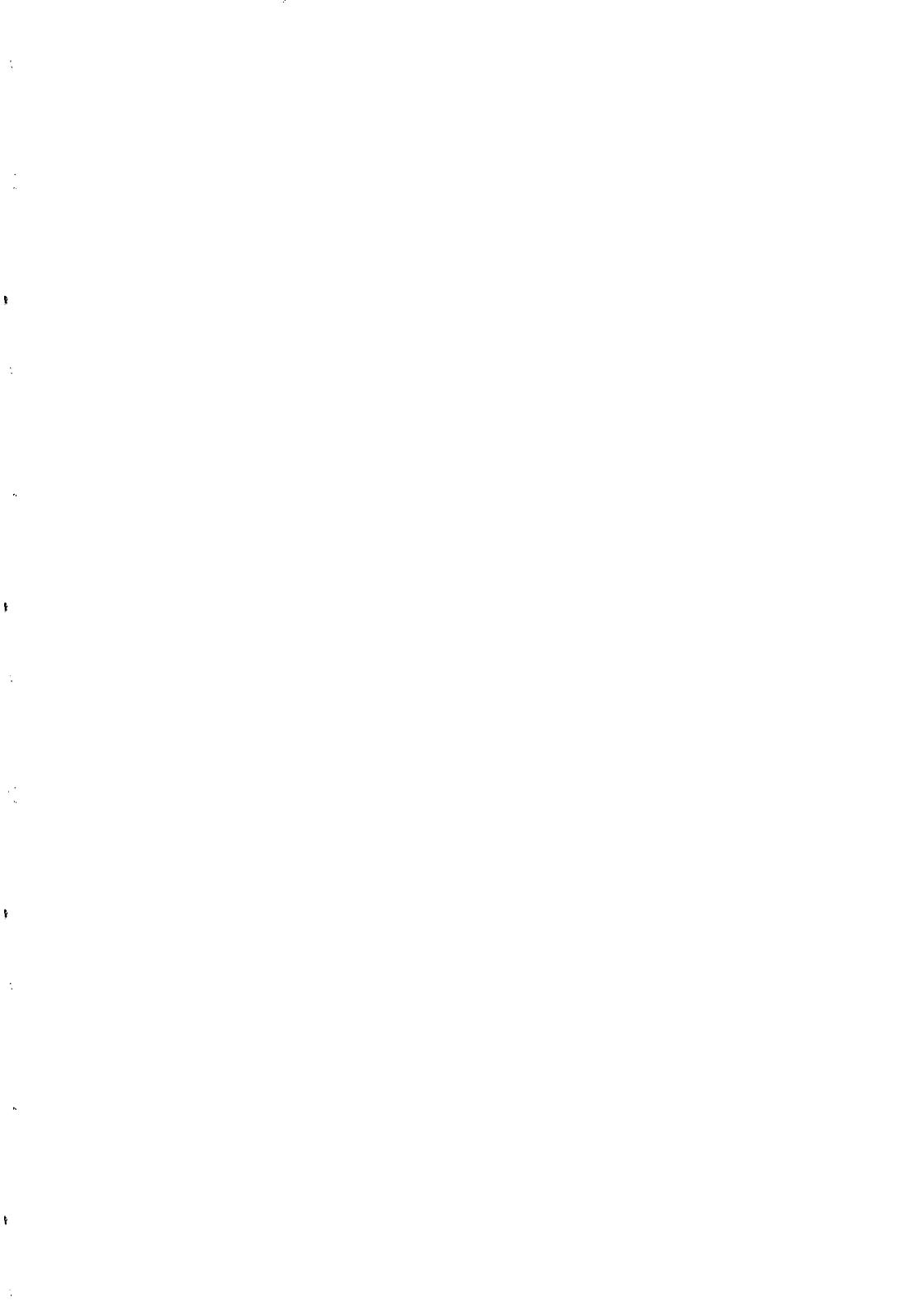


Figure 3: Upper limits to WIMP interaction total cross sections. Na and I curves are calculated from background spectra arising from this work (UK), from the BRS collaboration⁶⁾, and from the Osaka experiment⁹⁾. The Ca and F curves use the data from the BPRS collaboration¹⁰⁾. The ^{73}Ge curves use data from the UCSB¹¹⁾ and Gotthard²⁾ experiments.



INDIRECT DETECTION OF WIMPS

Marc Kamionkowski

School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540



ABSTRACT

I review several proposed techniques for indirect detection of weakly interacting massive particles (WIMPs) in the halo. I focus on distinctive signatures from cosmic-ray positrons, antiprotons, and gamma rays produced by annihilation of WIMPs in the galactic halo.

1. Introduction

There is almost universal agreement on the existence of dark matter in the Universe¹⁾. Luminous matter contributes only a fraction, $\Omega_{\text{LUM}} \sim 0.01$, of critical density. On the other hand, numerous observations suggest that Ω is in fact much larger. A number of theoretical arguments, such as inflation and the Dicke-Peebles timing coincidence suggest that the Universe is actually flat, $\Omega = 1$.

Although there is considerable debate on exactly how much dark matter there is, observations of flat galactic rotation curves provide incontrovertible evidence for the existence of dark matter in galactic halos, including our own. In general, rotation curves seem to remain flat as far out from the galactic center as are observed. Therefore, although it remains unclear exactly how much mass is entrained in galactic halos, it seems that the mass density contributed by halos is at least $\Omega_{\text{halo}} \gtrsim 0.1$. In other words, the dark matter in spiral galaxies outweighs the luminous matter by at least an order of magnitude. Big-bang nucleosynthesis suggests that there are more baryons than are seen, but it also constrains the mass density in baryons to be $\Omega_b \lesssim 0.1$ (Ref. 2). Therefore, it is plausible that there may be some baryonic dark matter in the form of nonluminous massive compact halo objects (MACHOs) such as neutron stars, brown dwarfs, or black holes, but it is difficult to see how baryons could account for all the halo dark matter.

One of the leading candidates for the dark matter is a weakly-interacting massive particle (WIMP). Suppose that in addition to the known particles of the Standard Model there exists a new, yet undiscovered, stable weakly-interacting massive particle, X . It is straightforward to show (see, e.g. Ref. 3) that if such a particle exists, it will have a current cosmological mass density in units of critical density given roughly by $\Omega_X h^2 \simeq \langle \sigma_A v \rangle / (3 \times 10^{-27} \text{ cm}^3 \text{ sec}^{-1})$, where $\langle \sigma_A v \rangle$ is the thermally averaged cross section for annihilation of X 's into all lighter particles times relative velocity v , and h is the Hubble constant in units of 100 km/sec/Mpc.

One can then ask, what annihilation cross section is required to give $\Omega_X \sim 1$? The answer turns out to be a weak scale cross section, i.e., $\sigma_A \sim \alpha^2 / (100 \text{ GeV})^2$, where $\alpha \sim 0.01$. Virtually all particle physicists will agree that there is new physics beyond the Standard Model, and many (if not most) of the best ideas for new physics introduce the existence of a WIMP. For example, a heavy neutrino associated with an extra generation could be the WIMP, but perhaps the most promising WIMP candidate is the neutralino, a linear combination of the supersymmetric partners of the photon, Z boson, and Higgs bosons⁴⁾. Although there can be significant variety in the detailed properties of the WIMP, generically, the interactions of the WIMP are constrained (by $\Omega_X \sim 1$) to be weak scale, and in most models, the mass of the WIMP varies from about 10 GeV to a few TeV.

A number of direct- and indirect-detection schemes are being pursued in an effort to discover WIMPs in the halo. The first class of experiments are laboratory efforts to detect

the recoil energy deposited in a low-background detector when a halo WIMP elastically scatters off a nucleus in the detector⁵). The most promising avenue for indirect detection is observation of energetic neutrinos from WIMP annihilation in the Sun and Earth. WIMPs in the halo which accumulate in the Sun and Earth will annihilate therein and produce energetic neutrinos that can potentially be detected by the many high-energy neutrino telescopes currently in operation or construction. I have reviewed this avenue for detection elsewhere⁶), so in this lecture, I will instead focus on several possible cosmic-ray signatures of WIMPs in the galactic halo.

Although the WIMP is stable, two WIMPs can annihilate into ordinary matter such as quarks, leptons, gauge bosons, etc. in the same way they did in the early Universe. If WIMPs exist in the galactic halo, then they will occasionally annihilate, and their annihilation products will produce cosmic rays. The difficulty in inferring the existence of particle dark matter from cosmic rays lies in discrimination between WIMP-induced cosmic rays and those from standard “background” sources. As will be argued below, it is quite plausible that WIMPs may produce distinctive cosmic-ray signatures distinguishable from background. It should also be made clear that propagation of cosmic rays in the Galaxy is quite poorly understood. Due to these astrophysical uncertainties, it is difficult to make reliable predictions for a given particle dark-matter candidate, so negative results from cosmic-ray searches cannot generally be used to constrain dark-matter candidates. On the other hand, if observed, these cosmic-ray signatures could provide a smoking-gun signal for the existence of WIMPs in the halo.

2. Cosmic-Ray Antiprotons

The best place to look for a distinctive cosmic-ray signature is where the background is smallest. The majority of cosmic rays are protons, and most of the rest are heavier nuclei. Only a very small fraction are antiprotons. Cosmic-ray antiprotons are produced in standard propagation models by spallation of primary cosmic rays on hydrogen atoms in the interstellar medium (ISM)⁷). The exact flux of antiprotons produced by this mechanism actually varies quite a bit in standard propagation models, and the observational situation is equally cloudy. However, there is one feature of the energy spectrum of such secondary antiprotons that is quite generic to standard cosmic-ray models: It is expected that the flux of antiprotons from primary spallation should fall dramatically at low energies, $E_{\bar{p}} \lesssim \text{GeV}$. This is simply because an antiproton at rest must be produced with a large backward momentum in the center-of-momentum frame. This requires a primary cosmic-ray antiproton with a large energy, and the cosmic-ray spectrum falls steeply with energy.

Annihilation of WIMPs, on the other hand, *can* produce low-energy antiprotons⁸). WIMPs will annihilate into quarks, leptons, gauge bosons, etc. which will then hadronize

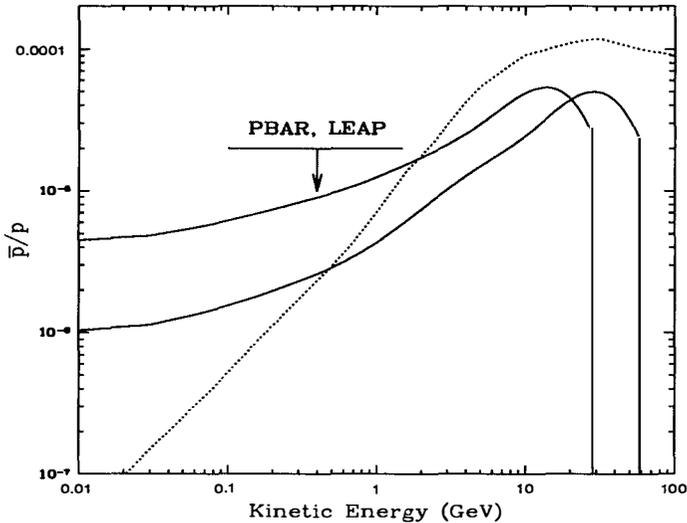


Fig. 1 Observed antiproton/proton ratio as a function of kinetic energy. (From Ref. 8.)

and produce, among other end products, antiprotons. There is no reason why the flux of such antiprotons should decrease dramatically at energies less than a GeV. Therefore, observation of low-energy cosmic-ray antiprotons would provide evidence for WIMPs in the halo.

Calculation of the antiproton flux from WIMP annihilation is straightforward. One assumes that the WIMPs have an isothermal distribution in the halo with a density suitable for accounting for the rotation curves. The flux is proportional to the annihilation rate in the halo. The energy spectrum of the antiprotons is determined by the fragmentation functions for producing antiprotons from the various annihilation products, which are obtained from Monte Carlos and from fits to accelerator data. Propagation of the antiprotons through the interstellar medium and solar modulation must also be considered.

In Fig. 1 are shown the cosmic-ray antiproton spectra expected from models where the dark matter is made up of a B -inos of mass 30 GeV (the upper solid curve) or 60 GeV (the lower solid curve)⁸. For simplicity, we chose the WIMP to be a B -ino and assumed that the WIMPs contribute closure density, $\Omega_{\tilde{\chi}} h^2 = 0.25$ with $h = 0.5$ to fix the annihilation cross section. We also assumed that WIMPs contribute the entire halo density, and used standard confinement times and solar-modulation models. The dotted curve is

the expected background due to spallation in the standard leaky-box model of cosmic-ray propagation. Also shown is the current observational upper limit⁹). As the WIMP mass is increased, the number density in the halo, and therefore the cosmic-ray flux, decrease. As illustrated, observation of low-energy cosmic-ray antiprotons could plausibly provide evidence for the existence of particle dark matter. It should be noted, however, that if the WIMP mass is too large, the antiproton signal would be unobservably small. In addition, even if the WIMP is fairly light, there are considerable astrophysical uncertainties, so it is possible that WIMPs could be the dark matter and still not produce an observable antiproton signal.

3. Cosmic-Ray Positrons

There is also a possibility that annihilation of some WIMP candidates will produce a distinctive cosmic-ray positron signature at high energies. Again, there is a “background” of cosmic-ray positrons from spallation of primary cosmic rays off the ISM. Pions produced when primary cosmic rays interact with ISM protons decay to muons which decay to positrons. The flux of positrons, expressed as a fraction of the flux of electrons, decreases slowly with increasing energies.

The showering of WIMP annihilation products will produce positrons in the same way that antiprotons are produced. The energies of the positrons that come from showering of annihilation products will have a broad energy distribution. The background spectrum of positrons expected from standard production mechanisms is quite uncertain, and precise measurements of the positron energy spectrum are quite difficult, so it is unlikely that positrons from WIMP annihilation with a broad energy spectrum could be distinguished from background.

However, in addition to the positrons that come from decays of hadrons, there is also the possibility that WIMPs may annihilate directly into electron-positron pairs thereby producing a “line” source of positrons. Although propagation through the Galaxy would broaden the line somewhat, the observed positron energy spectrum would still have a distinctive peak at an energy equal to the WIMP mass¹⁰). There are no standard production mechanisms that would produce a positron peak at energies of 10-1000 GeV, so such an observation would be a clear signature of particle dark matter in the halo. It is also interesting to note that some recent measurements of the positron spectrum indicate an increase in the positron fraction at high energies possibly suggestive of WIMP annihilation, although these results are far from conclusive.

Unfortunately, most of the leading WIMP candidates (e.g. neutralinos) are Majorana particles, and such particles do not decay directly into electron-positron pairs. On the other hand, if the WIMP is heavier than the W^\pm boson, it can in some cases (for example,

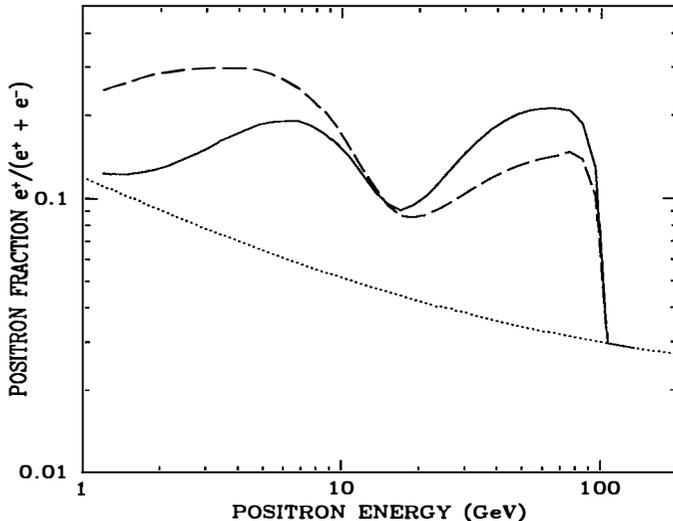


Fig. 2 The differential positron flux divided by the sum of the differential electron and positron fluxes as a function of energy for a neutralino of mass 120 GeV. (From Ref. 11.)

if the WIMP is a higgsino) annihilate into monochromatic W^+W^- pairs, and the W^+ bosons can then decay directly into positrons with a distinctive energy spectrum peaked at roughly half the WIMP mass¹¹). In addition, there will be a continuum of lower energy positrons produced by the other decay channels of the gauge bosons.

Fig. 2 shows the differential positron flux as a ratio of the electron-plus-positron flux as a function of energy for a higgsino of mass 120 GeV for two different models of cosmic-ray propagation (the solid and dashed curves). The dotted curve is the expected background. The peak at higher energies is due to direct decays of gauge bosons produced by WIMP annihilation into positrons, and the broader peak at lower energies comes from the other decay channels of the gauge bosons. The dramatic height of the peak in Fig. 2 is the result of some fairly optimistic, yet reasonable astrophysical assumptions. Again, due to the astrophysical uncertainties, nonobservation of such a signal cannot be used to rule out WIMP candidates.

4. Cosmic Gamma Rays

Cosmic gamma rays will be produced by annihilation of WIMPs in much the same

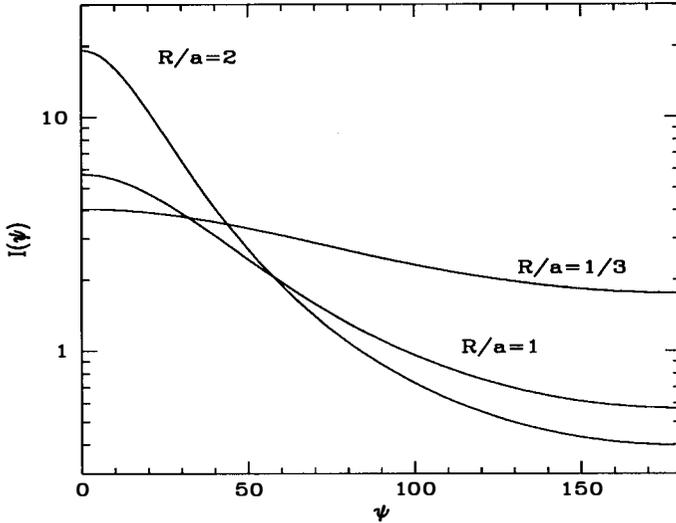


Fig. 3 The intensity of a gamma-ray signal from WIMP annihilation in the halo as a function of the angle between the line of sight and the galactic center. (As in Ref. 12.)

way that antiprotons and positrons are produced. Showering of the annihilation products will produce gamma rays with a broad energy distribution centered roughly around 1/10th the WIMP mass. Such a signal will in general be difficult to distinguish from background. However, there are two possible signatures of WIMP annihilation in the halo.

The first signature will be a distinctive directional dependence of the gamma-ray flux. In the simplest (and most plausible) models that account for galactic rotation curves, WIMPs populate the halo with a spherically symmetric isothermal distribution. Then, the density ρ of WIMPs as a function of distance r from the galactic center is $\rho(r) = \rho_0(R^2 + a^2)/(r^2 + a^2)$, where $R \simeq 8$ kpc the distance between the solar system and the center of the Galaxy, and a is the scale length of the halo. The ratio R/a varies between roughly 1/3 and 2. Given such a distribution, it is straightforward to calculate the angular dependence of the gamma-ray intensity $I(\psi)$ from WIMP annihilation as a function of ψ , the angle between the line of sight and the galactic center. Fig. 3 shows the result for the angular dependence of the gamma-ray flux for three values of the ratio R/a . Observation of such a signal would provide evidence for WIMPs in the halo.

Along similar lines, it has been suggested that there may also be an enhancement in

the dark-matter density in the galactic bulge or in the disk and if this dark matter were made of WIMPs, annihilation could lead to a strong gamma-ray signal from the galactic center or the disk¹³); however, it is difficult to see why WIMPs would accumulate at the galactic center or in the disk. Recently, Gondolo has suggested that the Large Magellanic Cloud could be immersed in a halo of dark matter with a central density 10 times that of our own galaxy, and that annihilation of dark matter therein could lead to a gamma-ray intensity from the LMC roughly ten times stronger than that from our own halo¹⁴).

The other, and very distinguishable, signature is a gamma-ray line from direct annihilation of WIMPs into photons. WIMPs, essentially by definition, have no direct coupling to photons. However, by virtue of the fact that the WIMP must have some appreciable coupling to ordinary matter (or else annihilation in the early Universe would be too weak to provide $\Omega_X h^2 \lesssim 1$), it is almost guaranteed that any realistic WIMP will couple to photons through loop diagrams. Therefore, there will always be some small, but finite, cross section for direct annihilation of two WIMPs into gamma rays. Therefore, WIMP annihilation in the halo can produce a gamma-ray signal that is monochromatic at an energy equal to the WIMP mass. There is no easily imaginable astrophysical source that would lead to a gamma-ray line at an energy between roughly a GeV and a TeV, so discovery of such a line could almost certainly imply the existence of WIMPs in the halo.

The problem with gamma-ray signatures from dark-matter annihilation is that the signals are at best only marginally observable with current detectors even with the most optimistic assumptions. There is, however, hope that heavier WIMPs which couple to the W^\pm boson, such as higgsinos, will annihilate more efficiently into gamma rays¹⁵). Also, there should be substantial improvements in observational high-energy gamma-ray astronomy in the forthcoming years.

5. Conclusions

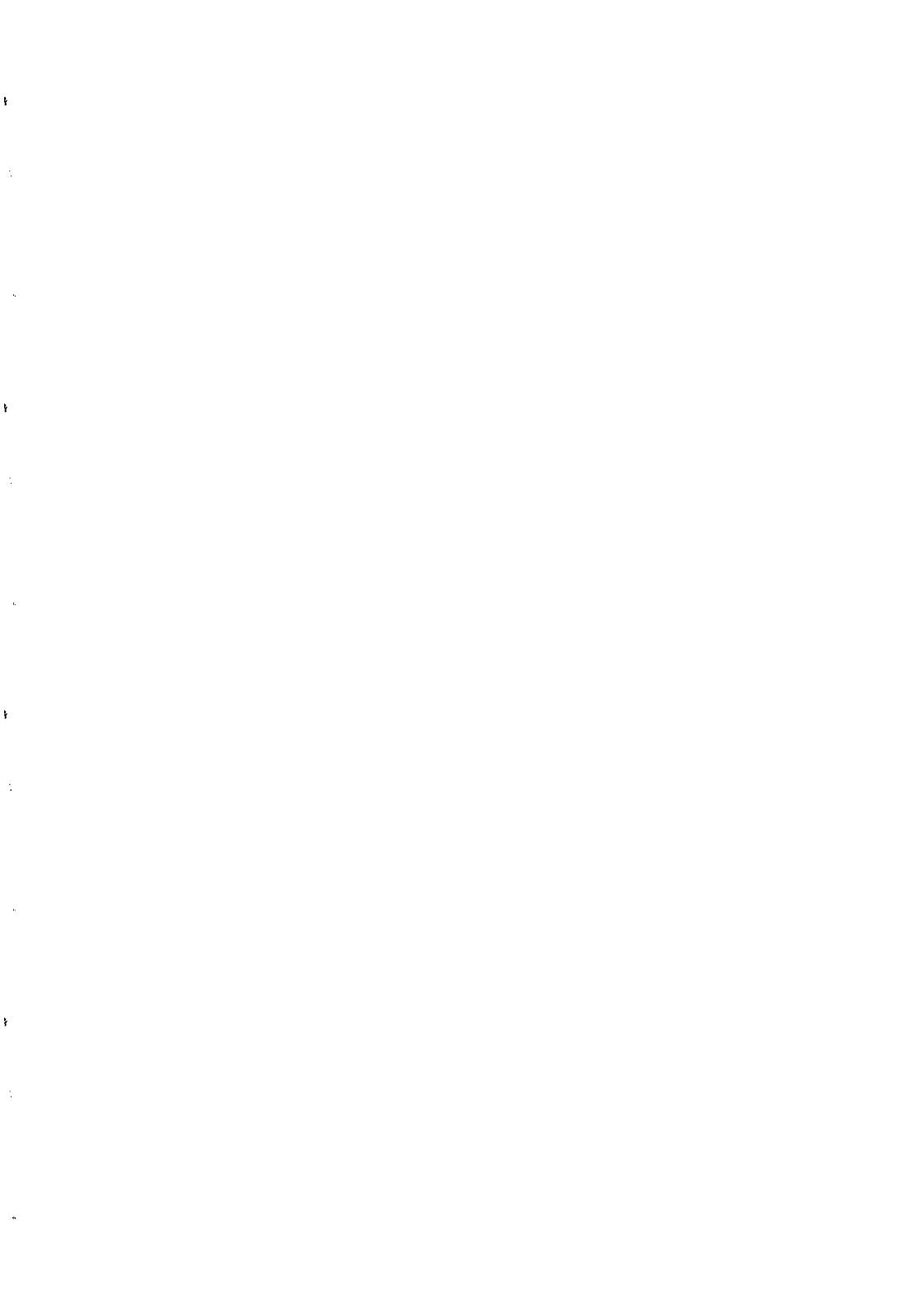
Of the many proposed dark-matter candidates, the WIMP is perhaps the most promising. The rather suggestive result that a stable particle with weak-scale interactions has a cosmological mass density of order unity has spurred tremendous theoretical and experimental activity in an attempt to detect dark-matter particles. The most reliable detection methods involve terrestrial low-background detectors and searches for energetic neutrinos from WIMP annihilation in the Sun and Earth. However, if WIMPs populate the halo, they will annihilate and produce cosmic rays. Although it will generally be difficult to distinguish such cosmic rays from background, WIMP annihilation may possibly lead to distinctive cosmic-ray signatures. Such signatures are by no means guaranteed even if WIMPs are the dark matter, but in many models it is quite plausible that observations of low-energy antiprotons, high-energy positrons, or gamma rays could provide indirect evidence for the existence of particle dark matter in our halo.

6. Acknowledgments

I gratefully acknowledge very productive and enjoyable collaborations with Manuel Drees, Kim Griest, Francis Halzen, Gerard Jungman, Mihoko Nojiri, Tim Stelzer, and Michael Turner. This work was supported by the U.S. Department of Energy under contract DEFG02-90-ER 40542.

References

- 1) *Dark Matter in the Universe*, eds. J. Kormendy and G. Knapp (Reidel, Dordrecht, 1989).
- 2) K. A. Olive, G. Steigman, D. N. Schramm, T. P. Walker, and H. Kang, *Astrophys. J.* **376** (1991) 51.
- 3) E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, 1990).
- 4) H. E. Haber and G. L. Kane, *Phys. Rep.* **117**, 75 (1985).
- 5) J. R. Primack, B. Sadoulet, and D. Seckel, *Ann. Rev. Nucl. Part. Sci.* **B38**, 751 (1988).
- 6) M. Kamionkowski in *High Energy Neutrino Astrophysics*, edited by V. J. Stenger, J. G. Learned, S. Pakvasa, and X. Tata (World Scientific, Singapore, 1992), p. 157.
- 7) T. Gaisser, *Cosmic Rays and Particle Physics*, (Cambridge University Press, Cambridge, 1990).
- 8) See G. Jungman and M. Kamionkowski, *Phys. Rev. D* **49**, 2316 (1994), and references therein.
- 9) S. P. Ahlen *et al.*, *Phys. Rev. D* **61**, 145 (1988); M. Salamon *et al.*, *Astrophys. J.* **349**, 78 (1990); R. E. Streitmatter *et al.*, *Proceedings of the 21st ICRC*, OG 7.3-2, Adelaide **3**, 277 (1990).
- 10) M. S. Turner and F. Wilczek, *Phys. Rev. D* **42**, 1001 (1990); A. J. Tylka, *Phys. Rev. Lett.* **63**, 840 (1989).
- 11) M. Kamionkowski and M. S. Turner, *Phys. Rev. D* **43**, 1774 (1991).
- 12) M. S. Turner, *Phys. Rev. D* **34**, 1921 (1986).
- 13) M. Urban *et al.*, *Phys. Lett. B* **293**, 149 (1992); H. Bloemen and J. Silk, *Astrophys. J. Lett.* **47**, 313 (1987).
- 14) P. Gondolo, astro-ph/9312011 (1993).
- 15) L. Bergstrom and J. Kaplan, USITP-94-03 (1994).



OBSERVING GAMMA LINES FROM HIGGSINOS IN AIR CHERENKOV TELESCOPES

L. Bergström

Department of Physics
Stockholm University
Box 6730, S-113 85 Stockholm
Sweden

and

J. Kaplan

Laboratoire de Physique Théorique et Hautes Energies
Université de Paris VI et VII
Unité associée au CNRS (UA 280)
2 pl. Jussieu, F-752 51 Paris
France

Abstract

We discuss the annihilation processes in the galactic halo $\chi\chi \rightarrow \gamma\gamma$ and $\chi\chi \rightarrow \gamma Z$, where χ is any halo dark matter particle that has W^+W^- as one of the major annihilation modes. Examples of such particles are supersymmetric particles with a dominant pure higgsinos or heavy triplet neutrinos. A substantial branching ratio is found for these modes. We estimate the detection potential of Air Cherenkov Telescopes for monoenergetic γ ray lines that result from such annihilations in the Galactic or LMC halos.

One of the favoured particle dark matter candidates is the lightest supersymmetric particle χ , assumed to be a neutralino, i.e. a mixture of the supersymmetric partners of the photon, the Z^0 and the two neutral CP -even Higgs bosons present in the minimal extension of the supersymmetric standard model (see, e.g. [1]).

A particularly interesting annihilation process in the galactic halo is $\chi\chi \rightarrow \gamma\gamma$ or $\chi\chi \rightarrow Z^0\gamma$. Since these are two-body final states and the annihilating massive particles move with non-relativistic speed in the halo (typically $v/c \sim 10^{-3}$) the produced γ s will be nearly monoenergetic, meaning a γ ray line signature [2].

In [3] the $\gamma\gamma$ process was estimated for the case of a nearly pure Bino. These rates turn out to be very small [3, 4]. Recently the more difficult calculation for a pure Higgsino has been performed, both for the $\gamma\gamma$ and the $Z^0\gamma$ final states [4]. The latter turns out to be the most important one for Higgsinos (and for any other dark matter candidate that has a strong coupling to the W^+W^- final state in its annihilations).

One gets for the imaginary part of the amplitude (see [4] for details) a logarithmic enhancement of the form

$$M_{im} = \beta^2 \log \left(\frac{1 + \beta}{1 - \beta} \right), \quad (1)$$

where $\beta = \sqrt{1 - (m_W/m_\chi)^2}$.

In the limit of large χ masses (above, say, 500 GeV) this becomes

$$M_{im} \sim \log \left(\frac{4m_\chi^2}{m_W^2} \right). \quad (2)$$

In the leading logarithmic approximation, both the real and imaginary parts can be calculated:

$$M_{im}^{LL} = \log \left(\frac{s}{m_W^2} \right), \quad (3)$$

$$M_{re}^{LL} = \frac{1}{2\pi} \log^2 \left(\frac{s}{m_W^2} \right), \quad (4)$$

where $s = 4m_\chi^2$.

The γ ray luminosities from $\gamma\gamma$ and $Z^0\gamma$ should be added together. In fact the line strength from the latter is a few times larger than that from the former.

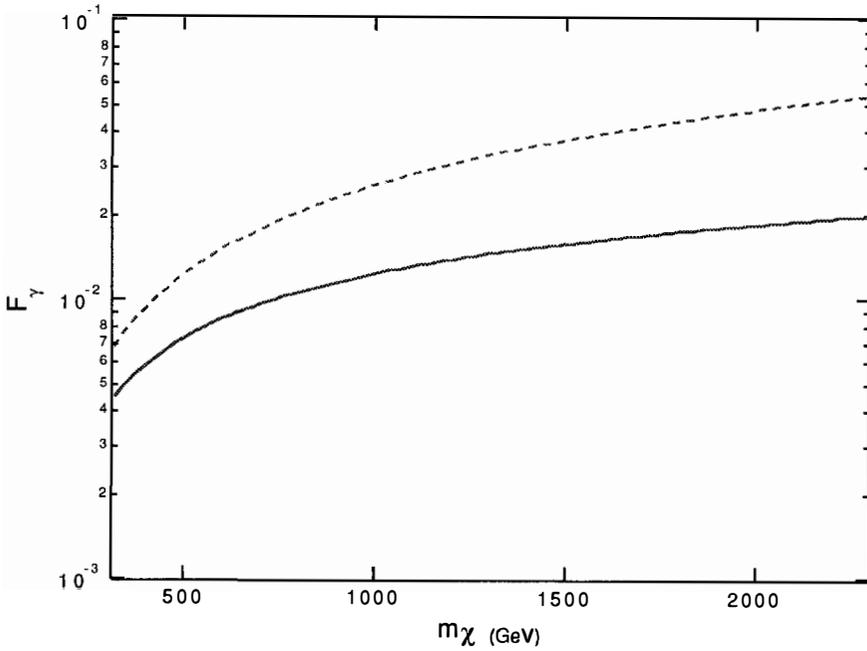


Fig. 1. The average number F_γ (Eq. (5)) of γ line photons from $\chi\chi \rightarrow \gamma\gamma$ and $Z^0\gamma$ normalized to the W^+W^- annihilation rate, as a function of the χ mass. The solid line is the lower bound obtained by using only the imaginary part of the amplitude which is given by unitarity. The dashed line is an estimate of the full rate obtained by taking the leading logarithmic result of Eq. (4) for the real (dispersive) part of the amplitude.

In Fig. 1 we show our lower bound [4] for the average number of γ line photons of energy $\sim m_\chi$ per W^+W^- annihilation:

$$F_\gamma \equiv \frac{2\sigma v(\chi\chi \rightarrow \gamma\gamma) + \sigma v(\chi\chi \rightarrow Z^0\gamma)}{\sigma v(\chi\chi \rightarrow W^+W^-)} = \alpha_{e.m.}^2 (1 + \cot^2 \theta_w) \log^2 \left(\frac{4m_\chi^2}{m_W^2} \right), \quad (5)$$

as well as the value for this rate resulting from the adoption of (4) for the dispersive part of the amplitude. The remarkable feature of our result (5) is its very large size compared to the case of, e.g. a pure Bino [3, 4]. As can be seen in Fig. 1, the effective branching ratio to line photons can easily be as large as 10^{-2} .

For a generic dark matter particle that gives a substantial contribution to Ω of the universe, $(\sigma v)_{tot}$ is of the order of $10^{-26} \text{ cm}^3 \text{ s}^{-1}$. If the annihilation rate of neutralinos into W^+W^- does not vanish nor become small when $\langle v \rangle \rightarrow 0$, $(\sigma v)_{tot}$ is expected to be of the same order of magnitude in the galactic halo. As an example, this happens for a Higgsino of the minimal supersymmetric extension of the Standard Model of a mass around 1 TeV. Since we see from Fig. 1 that the number of line photons per annihilation is roughly of the order of a percent, we can use the value $(\sigma v)_{line} \sim 10^{-28} \text{ cm}^3 \text{ s}^{-1}$ in estimates.

In fact, for the case of a pure Higgsino, we can calculate the absolute value of the lower bound of the line cross section. We find

$$2\sigma v(\chi\chi \rightarrow \gamma\gamma) + \sigma v(\chi\chi \rightarrow Z^0\gamma) > \frac{\pi \alpha_{e.m.}^4}{8m_\chi^2 \sin^6 \theta_w} \log^2 \left(\frac{4m_\chi^2}{m_W^2} \right), \quad (6)$$

which corresponds to $0.6 \cdot 10^{-28} \text{ cm}^3 \text{ s}^{-1}$ for $m_\chi = 1 \text{ TeV}$. Assuming, lacking a full calculation, a contribution from the dispersive part of the amplitudes of the same order of magnitude, we see that in this case the annihilation rate into line photons is indeed around $10^{-28} \text{ cm}^3 \text{ s}^{-1}$.

Using the Ipsier-Sikivie model for the core of the halo [5], a flux of around 200 photons per year is predicted in an ACT of area 20 000 m^2 during a typical observation "year" of $2 \cdot 10^6 \text{ s}$, with an estimated (but poorly known) background about an order of magnitude lower (including misidentified protons and electrons; see [3] for further discussions). This rate for Higgsino-like particles is two orders of magnitude larger than that expected for pure Binos [3, 4]. In the model by Berezhinsky et al [6], the flux can be still at least two

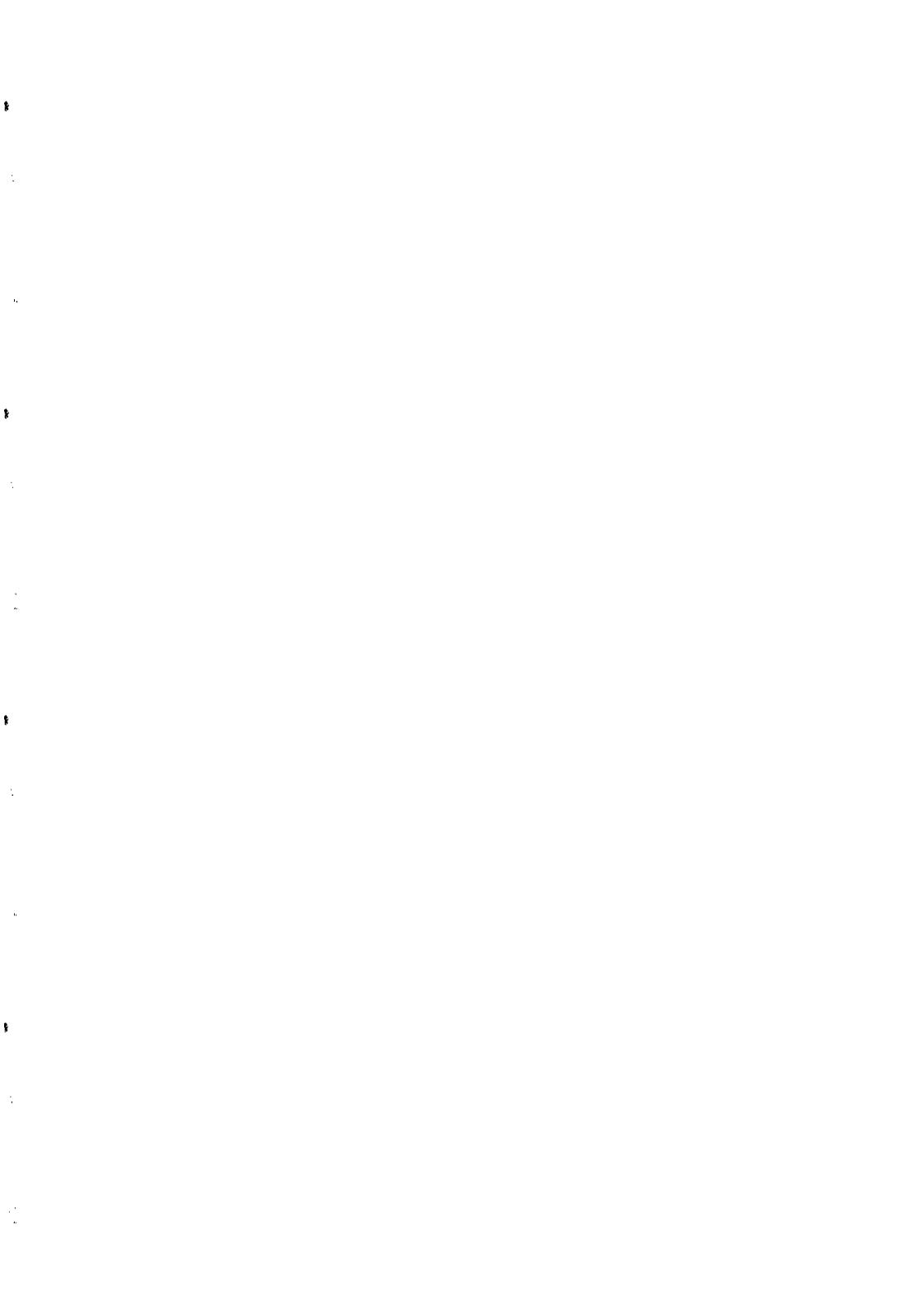
orders of magnitude larger. However, if one assumes a smooth halo distribution $\rho(r) = \rho_{\odot}(r_{\odot}^2 + a^2)/(r^2 + a^2)$ with $a \sim r_{\odot}$, the rate falls below one event per year in our generic ACT detector in the direction of the galactic center. In the LMC model of Gondolo [7] the rate could be enhanced by an order of magnitude making it just about reaching the detectability limit. It should be noted, though, that the high threshold (200-300 GeV) of present day ACTs make these processes rate limited. Going to larger areas and lower thresholds could greatly increase the discovery potential of this type of detector.

If the galactic halo contains regions with higher dark matter density than the local (solar neighbourhood) value, the detection of high energy monoenergetic photons could at the same time determine the mass of the dark matter particle and map its galactic density distribution.

This research has been sponsored by the Swedish Natural Science Research Council and a European Community Twinning grant. We wish to thank P. Gondolo, F. Martin and H. Rubinstein for useful discussions. We are especially grateful to M. Urban, whose idea to detect γ lines from WIMP annihilation using Atmospheric Čerenkov Telescopes triggered this work.

References

- [1] H.E. Haber and G.L. Kane, Phys. Rep. 117 (1985) 75.
- [2] L. Bergström and H. Snellman, Phys. Rev. D37 (1988) 3737; S. Rudaz, Phys. Rev. D39 (1989) 3549; A. Bouquet, P. Salati and J. Silk, Phys. Rev. D40 (1989) 3168, G.F. Giudice and K. Griest, Phys. Rev. D40 (1989) 2549.
- [3] M. Urban, A. Bouquet, B. Degrange, P. Fleury, J. Kaplan, A.L. Melchior and E. Paré, Phys. Lett. B293 (1992) 149.
- [4] L. Bergström and J. Kaplan, preprint USITP-94-03, PAR-LPTHE 94-06, to appear in Astroparticle Physics.
- [5] J.R. Ipser and P. Sikivie, Phys. Rev. D35 (1987) 3695.
- [6] V.S. Berezinsky, A.V. Gurevich and K.P. Zybin, Phys. Lett. B294 (1992) 221.
- [7] P. Gondolo, contributed paper to TAUP 93, Gran Sasso, Italy; to appear in the Proceedings.



CRYOGENIC DETECTOR DEVELOPMENT FOR DIRECT DARK MATTER SEARCH

G. Forster, F. v. Feilitzsch, J. Höhne, E. Kellner, D. Wilhelm
(Physik-Department E 15, Technische Universität München, D-85747 Garching, Germany)
C. Bucci, P. Colling, S. Cooper, P. Ferger, M. Frank, U. Nagel,
A. Nucciotti, F. Pröbst, A. Rulofs, W. Seidel, L. Stodolsky
(Max-Planck-Institut für Physik, Föhringer Ring 6, D-80805 München, Germany)



A fundamental question of astrophysics and cosmology is the nature of "dark matter". Among the possibilities is the existence of weakly interacting massive particles (WIMPs). Cryogenic detectors are especially suited for a direct detection of WIMPs. The recoil energy transferred to a nucleus by elastic scattering of a WIMP is mainly converted into phonons, which are measured by the thermometer of the cryogenic detector. At the Max Planck Institute for Physics and at the Technical University Munich we are developing low temperature calorimetric detectors with large mass, low energy threshold and high energy resolution. A first experiment with a 1 kg detector made from sapphire and with an energy threshold of 500 eV for nuclear recoils is in preparation. The range of sensitivity and the realisation of such an experiment are discussed.

1 Introduction

The search for "dark matter" is currently one of the greatest challenges of particle and astrophysics. Besides baryonic dark matter and massive neutrinos, weakly interacting massive particles (WIMPs) are a possible constituent of dark matter. The lightest supersymmetric particle, the neutralino (χ), is presently a theoretically favored candidate for a WIMP ¹⁾.

Direct detection of WIMPs is possible by their elastic scattering on nuclei. This process transfers a recoil energy to the nucleus:

$$\Delta E = \frac{m_n m_d^2}{(m_n + m_d)^2} v^2 (1 - \cos \theta),$$

where m_n is the mass of the nucleus, m_d the mass of the dark matter particle, v its velocity, and θ the scattering angle in the center of mass system. The velocity of the WIMPs is assumed to follow a Maxwellian distribution ($v_{rms} \approx 300$ km/s) with an upper cutoff at the escape velocity of about 600 km/s, which is modified by the velocity of the solar system relative to the center of the galaxy of 220 km/s. For WIMP masses in the GeV range one expects nuclear recoil energies in the keV range. The present experimental limits are given by low background silicon and germanium ionization detectors ^{2, 3, 4)}.

2 Operating Principles

2.1 Calorimeters

The energy of the recoil nucleus can be measured by various excitations: ionization, scintillation or phonons. The measurement of phonons by a calorimeter has the advantage of being able to measure the full energy of low energy nuclear recoils.

A calorimeter consists of an absorber, in which the particle interaction takes place, and a thermometer attached to the absorber. The particle interaction (i.e. the recoil nucleus) creates high energy phonons, which are absorbed in the thermometer and lead to a temperature rise given by $\Delta T = \Delta E/C$, where C is the heat capacity of the thermometer.

To obtain large signals, the heat capacity has to be reduced by lowering the operating temperature of the calorimeter.

2.2 Phase transition thermometers

To achieve maximum temperature sensitivity, we use as the thermometer an evaporated superconducting film stabilized in its phase transition from the superconducting to the normal state. The very strong dependence of the resistance on the temperature in the limited temperature range of the phase transition makes this kind of thermometer very sensitive.

To obtain phase transition thermometers with low operating temperatures, we use two approaches:

- Tungsten has a bulk critical temperature of 15 mK. By epitaxial evaporation in a designated system, we can produce thin films with a critical temperature near the bulk value ⁵⁾.
- By completely covering a superconducting iridium film with a normalconducting gold film, the critical temperature of the iridium is reduced by the proximity effect. We succeeded in producing thermometers with critical temperatures between 30 and 120 mK ⁶⁾.

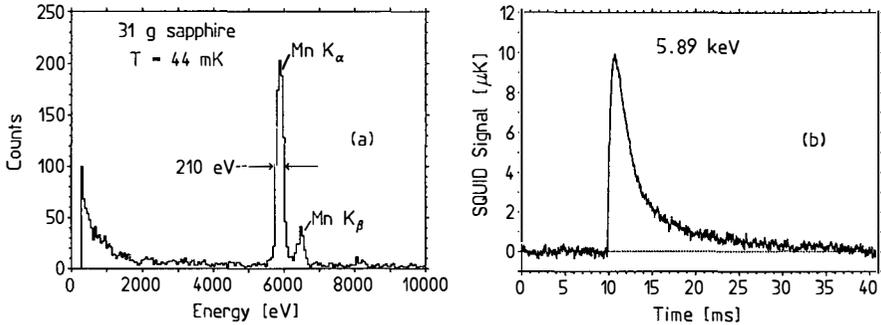


Figure 1: (a) Pulse height spectrum of a ^{55}Fe source. The detector consists of a 31 g sapphire crystal and an iridium-gold thermometer with a critical temperature of 44 mK. (b) Single 5.89 keV event measured with a bandwidth of 15 kHz.

3 Experimental Results

3.1 Iridium-gold bilayer thermometers

A sapphire crystal with a mass of 31 g and an evaporated iridium-gold bilayer thermometer with a critical temperature of 44 mK was irradiated with a ^{55}Fe X-ray source. The energy resolution of this detector is 210 eV for photons of an energy of 5.9 keV (figure 1a)⁸⁾.

Figure 1b shows a single 5.89 keV event exhibiting two pulse components - a pronounced fast one with a decay time of about 2 ms and a small slow one with a decay time of about 13 ms. The fast component results from the absorption of high frequency phonons in the thermometer, and the slow component represents the temperature rise of the absorber crystal. We have developed a model that explains the shape and temperature dependence of the signals⁹⁾ and allows us to optimize our detectors for a dark matter experiment.

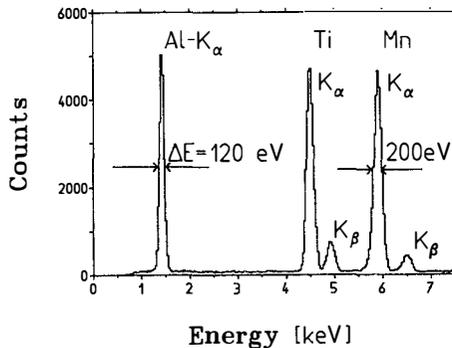


Figure 2: Energy spectrum of a fluorescence X-ray source, observed with a 31 g sapphire crystal with tungsten thermometer. The irradiation time is 12 hours.

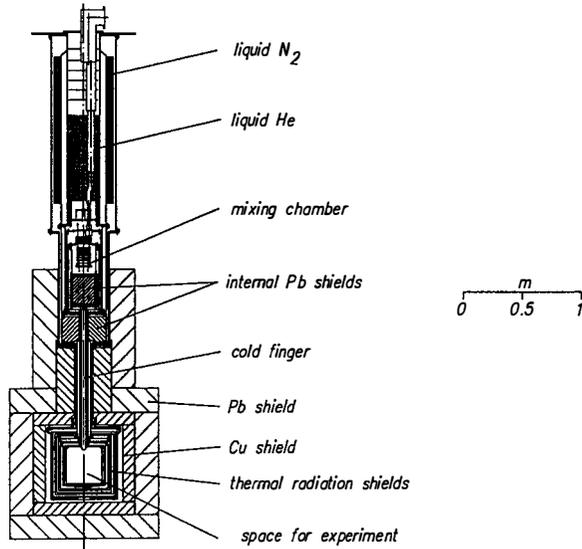


Figure 3: Dilution refrigerator, cold box and shielding.

3.2 Tungsten thermometers

A 31 g sapphire detector with a tungsten thermometer was irradiated with an X-ray fluorescence source. The resulting energy spectrum is shown in figure 2. The resolution for 1.5 keV photons is 120 eV.

By optimizing the thermometer design we want to realize a detector with a mass of 250 g and a threshold of 500 eV. The first phase of our dark matter search will then use four of these detectors.

4 Planned Experiment

4.1 Low-Background Environment

For a search for dark matter, it is of extreme importance to reach the lowest possible radioactive background. To avoid background events induced by cosmic rays the whole experiment needs to be located in an underground laboratory. A request for space to the LNGS ¹⁰⁾ has been approved. A low background cryostat will be installed in Gran Sasso in 1995.

Because not all materials necessary for building a dilution refrigerator are radiopure we decided to divide the experimental setup into two parts: one containing the dilution refrigerator and the other containing the detector surrounded by very low background materials and shielding. The cooling power of the refrigerator will be transferred into this cold box by a cold finger. The space for experiments in the cold box will be about 30 liters. From the mechanical and cryogenic point of view, it is the simplest solution to place the cold box under the dilution refrigerator (figure 3). The line of sight into the cold box will be blocked by an internal lead shield.

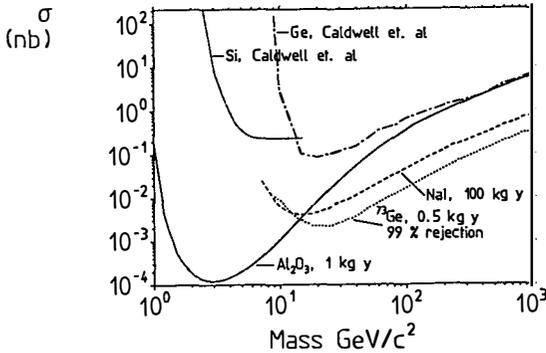


Figure 4: Rough estimate of limits of sensitivity of different experiments in terms of the point-like cross section for the scattering of WIMPs on the corresponding target nucleus, assuming a background of 1 count/(keV kg day) and a measuring time of one year. The sensitivities of a 100 kg NaI, a 0.5 kg ^{73}Ge with 99% background rejection, and a 1 kg sapphire detector are shown. The present limits from Ge and Si ionization detectors are also shown.

4.2 Sensitivity

The present experimental limits for WIMP dark matter are given by low background silicon and germanium ionization detectors. In this section we compare the sensitivity range of planned experiments.

- NaI scintillator ¹¹⁾

A 1 kg prototype of a NaI scintillator has an energy threshold of 4 keV for γ 's, which corresponds to 16 keV for Na nuclear recoils and 60 keV for I recoils due to the poor ionization efficiency for slow recoil nuclei. The energy resolution is 12.5 keV (FWHM) for 88 keV photons. A total detector mass of 100 kg is planned.

- Combined measurement of ionization and heat in a germanium crystal ¹²⁾

This allows discrimination of nuclear recoils from electron or photon induced events in the detector, and may be used to reduce the radioactive background by two orders of magnitude. The threshold of a 60 g detector is about 12.5 keV for nuclear recoils, and the energy resolution is about 1.8 keV (FWHM) for 60 keV photons. The UCSB/LBL/UCB group has enough enriched ^{73}Ge for a 0.5 kg detector.

- Our cryogenic detectors will have low thresholds of about 0.5 keV for nuclear recoils and an excellent energy resolution of about 0.2 keV (FWHM) for 1 keV photons. We plan to use 1 kg of sapphire for an initial experiment. In contrast to the other detectors a wide range of absorber materials can be used.

In figure 4 we show rough estimates of the upper limits on the cross section for WIMP-nucleus scattering as a function of the WIMP mass that different proposed detectors could set assuming a measuring time of one year and a flat background of 1 count/(keV kg day).

For WIMP masses below 15 GeV a detector with a low threshold is needed. For large WIMP masses most of the recoil spectrum is above all thresholds, and the sensitivity is then given mainly by the detector mass, the measuring time and the background rate.

5 Conclusion

We will begin a test phase for a dark matter experiment with a total mass of 1 kg of sapphire in the Gran Sasso underground facility in 1996. The detector should have an energy resolution of 200 eV at 1 keV and an energy threshold of 500 eV. This experiment will be complementary to other planned dark matter searches especially in the low WIMP mass region.

In a later stage of the experiment, we plan to increase the detector mass and to use other absorber materials. The different recoil spectra for different target nuclei may allow us to determine the mass of the WIMPs.

References

- 1) J. Ellis, R. A. Flores, Phys. Lett. **B 263** (1991) 259.
- 2) D. O. Caldwell, Mod. Phys. Lett. **A 5** (1990) 203.
- 3) D. Reusser et al., Phys. Lett. **B 255** (1991) 143.
- 4) D. O. Caldwell et al., Phys. Rev. Lett. **65** (1990) 1305.
- 5) P. Colling et al., in ⁷⁾, 549.
- 6) U. Nagel et. al., in ⁷⁾ 543; U. Nagel et al., submitted to J. App. Phys.
- 7) Proc. 5th Int. Workshop on Low Temperature Detectors (Berkeley, July 1993), J. Low Temp. Phys. Vol. 93, Nos. 3/4 (1993).
- 8) P. Ferger et al., Phys. Lett. **B 323** (1994) 95.
- 9) F. Pröbst, Proc. 7th Int. Conf. on Phonon-Scattering in Condensed Matter (Ithaca, New York, Aug. 1992), eds. M. Meissner and R.O. Pohl, Springer Series in Solid-State Sciences 112; M. Frank et al. in ⁷⁾, p. 213; F. Pröbst et al., in preparation.
- 10) S. Cooper et al., MPI Preprint MPI-PhE/93-29 (1993).
- 11) C. Bacci et al., (BRS coll.), Phys. Lett. **B 293** (1992) 460.
- 12) T. Shutt et al., Phys. Rev. Lett. **69**, (1992), 3425 and 3531.

**PROSPECTS FOR DARK MATTER DETECTION WITH METASTABLE
SUPERCONDUCTORS**

Jl Collar^{a)}, TA Girard^{b)} and G Waysand^{c)}

a) Physics Department, University of South Carolina, Columbia SC 29208 USA.

b) Centro de Física Nuclear, Universidade de Lisboa, 1699 Lisbon Portugal.

c) GPS, Universités Paris 7 et 6, 75251 Paris Cedex 05 France.



ABSTRACT

Recent X-ray and β decay irradiations of detectors based on granular suspensions of metastable Type I superconductors have yielded spectra of the energy loss of the incident radiation. Interpretation of the results suggest the application of these devices to new searches for relic neutrinos and WIMPs, with the possibility of achieving energy resolutions on the order of eV and reasonable event rates.

Current experimental limits on the existence of WIMPs are derived from low background Si and Ge ionization detectors¹. New experiments under development, employing liquid or solid scintillators and cryogenic bolometers, are resumed elsewhere in these Proceedings². All of the new search experiments are based on the detection of nuclear recoils induced by neutral current scattering. The key focus of these new experiments is on the lowering of the recoil energy threshold, and reduction of signal-competing backgrounds.

The use of metastable superconducting detectors in searching for WIMPs was first suggested by Drukier and Stodolsky³. These devices are based on the phase transition of superheated microgranules of Type I superconductors induced by energy loss of incident radiation. Order 10 nS timing resolution with grains of diameter > 8 micron has been demonstrated⁴; smaller diameters however require implementation of SQUID-readout, which is intrinsically slower. XY camera techniques, comprising orthogonal planes of pickup loops, have yielded order 1 mm spatial resolution⁵. Despite having demonstrated sensitivity to a variety of different irradiations⁶, the difficulty with these devices has until recently been in obtaining either energy calibration or particle spectra.

Last year we reported first results⁷ obtained with a Sn detector prototype irradiated by decay electrons of ^{35}S . In Fig. 1, we show these results in comparison with an improved Monte Carlo simulation of the detector response based on the thermodynamics of the hot border model⁸, and experiment detail including the measured distribution of grain diameters. The electrons are randomly emitted from a distributed source in accordance with the decay distribution; tracking through both the grains and dielectric includes Moliere scattering, and the energy loss is accounted for by Bethe-Bloch. Other than the effective source intensity, the simulation contains no free parameters.

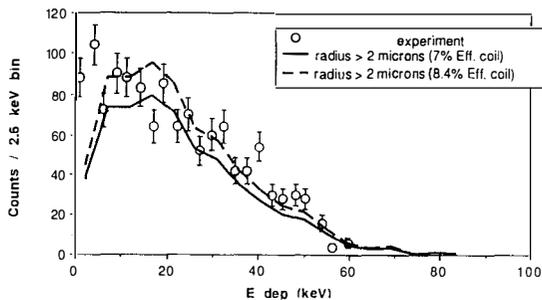


Fig. 1: electron energy loss spectrum obtained with an MS detector

Pertinent features of the measurement were its execution over 0.4 s following a 3 s

pause in magnetic field ramping^{9]} to establish the hot border, and that it was performed from high-to-low energy deposition. More recently, similar irradiation studies of various detector prototypes with 6 keV X-rays of ⁵⁵Fe have yielded energy-resolved spectra of the X-ray absorption and fluorescent activation of the suspension material^{10]}. A typical result, obtained at 400 mK with a composite device of 20-25 μ m diameter Sn and Al grains, is shown in Fig. 2 (the magnetic field was held above 300 G so that the Al was normal).

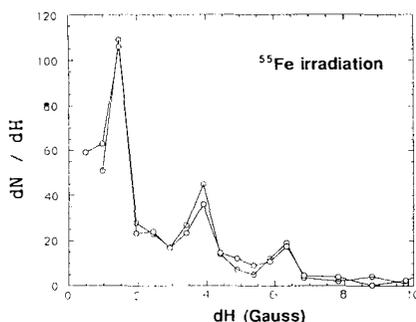


Fig. 2: spectrum of 6 keV X-ray irradiation of an MS detector

For a detector operated at temperature T , the energy calibration is to a first approximation^{8]},

$$\Delta E = C_p [-T + \{ T^2 + 3/2 T_c^2 [H_{sh}(0) - 1] dH \}^{1/2}] \quad (1)$$

where T_c is the critical temperature of the detector material, C_p is its heat capacity, $H_{sh}(0)$ is the superheating critical field at $T=0$; dH is the magnetic field step from the pause field. For the beta decay study, the 0.29 G step-size at 2.3 K was equivalent to 1.2 keV over a range of approximately 100 keV. In the case of the X-ray study, identification of the highest energy peak with the full absorption of the incident photons in the Sn grains implies the intermediate and lowest peaks correspond to the activated Al (1.5 keV) and carbon (277 eV) X-rays, respectively. The FWHM of the carbon line is 100 eV. The energy equivalence of the magnetic stepsize is in principle only a function of the grain specific heat and operating temperature: at 100 mK, with the same suspension, a 0.01 G step would correspond to 0.21 eV.

For WIMP elastic scattering, the maximum recoil energy is dependent on the candidate (m_x) and target nucleus (M) masses:

$$E_{recoil}^{max} [\text{GeV}] = 2m_x^2 M [m_x + M]^{-2} (v/c)^2 \quad (2)$$

Estimates of this energy for a range of candidate detector targets assuming $v/c = 10^{-3}$ are shown in Table 1, as well as corresponding event rate estimates^{11]} assuming $\rho_x = 0.4 \text{ GeV/cm}^3$. For comparison, 10 GeV WIMP values for Ge are $700 \text{ kg}^{-1}\text{d}^{-1}$, $E_{recoil}^{max} =$

2.2 keV. The X-ray results suggest the sensitivity of metastable detectors to energy depositions of well below 1 keV; results obtained in neutron-irradiations of Al and Sn detectors^{12]} similarly indicate a threshold sensitivity of 1 keV or below.

Table 1: WIMP scattering rates and recoil energies for various target nuclei.

m_x [GeV]	--- rate (events/kg day) --			----- $E_{\text{recoil}}^{\text{max}}$ [keV] -----		
	Al	Sn	Pb	Al	Sn	Pb
1	27	170	370	0.076	0.017	0.010
5	95	730	1400	1.4	0.41	0.24
10	150	1400	3400	4.1	1.5	0.93
50	163	4000	9700	22	21	16
100	120	4700	6000	32	50	45

The beta decay results suggest the further possibility to search for evidence of relic neutrinos predicted by standard cosmology^{13,14]}. Studies of nuclear β decay spectra performed in this background are subject to competitive effects from capture reactions, which would yield a peak at the spectrum endpoint with a width determined by the background neutrino spectrum. If massive, these relics might be clustered in galaxies with a density comparable to that estimated for dark matter in our galaxy, and would shift the capture peak $m_\nu c^2$ above the spectrum endpoint. The expected event rate in ^3H for massless neutrinos is $0.15 \text{ yr}^{-1}\text{g}^{-1}$; for 25 eV neutrinos, $31 \text{ yr}^{-1}\text{g}^{-1}$. The ability to perform the experiment from high-to-low energy would obviate to a large extent pileup effects, permitting larger source intensities. Unfortunately, a mg of ^3H equals 10 Ci activity, which *would* heat the detector. An alternative, superconducting, decay possibility is ^{187}Re , with an endpoint of 2.67 keV and rate estimate of $1.27 \times 10^{-2} \text{ g}^{-1}\text{y}^{-1}$ for $m_\nu = 25 \text{ eV}$: the activity of 10 g is 0.5 μC , well below the activity of our higher endpoint ^{35}S experiments. Any plateau in the rate of grain transitions at/above the spectrum endpoint would serve as an indication of the existence of such relics.

The materials-associated radioactive backgrounds of these devices remain largely unexplored. One and only one grain would be expected to transition per event in either search, providing a convenient background rejection criterion; estimates suggest the rejection efficiency in a 10 g Sn-based WIMP experiment to be of order 90%. Another difficulty is the currently small mass-to-readout ratio. While primarily a result of suspending the grains in the dielectric to volume filling factors of $\leq 20\%$, only 10% of the sensitive grains are currently utilized. We show in Table 2 comparative device parameters, based on the current prototype design, yielding a reasonable event rate for each of the searches. One detector element consists of a 3×3

cm² epoxy plane, carrying a single pickup loop on each face; this loop is 700 μm wide (including 200 μm of wire diameter) and serpentine with a pitch of 2000 μm. Each readout channel consists of a pickup loop, transformers, preamp, amp, and discriminator; there are two channels of instrumentation per plane. The suspension thickness is 0.3 mm, so that each detector plane has a height of 0.16 cm.

Table 2: search detector parameters.

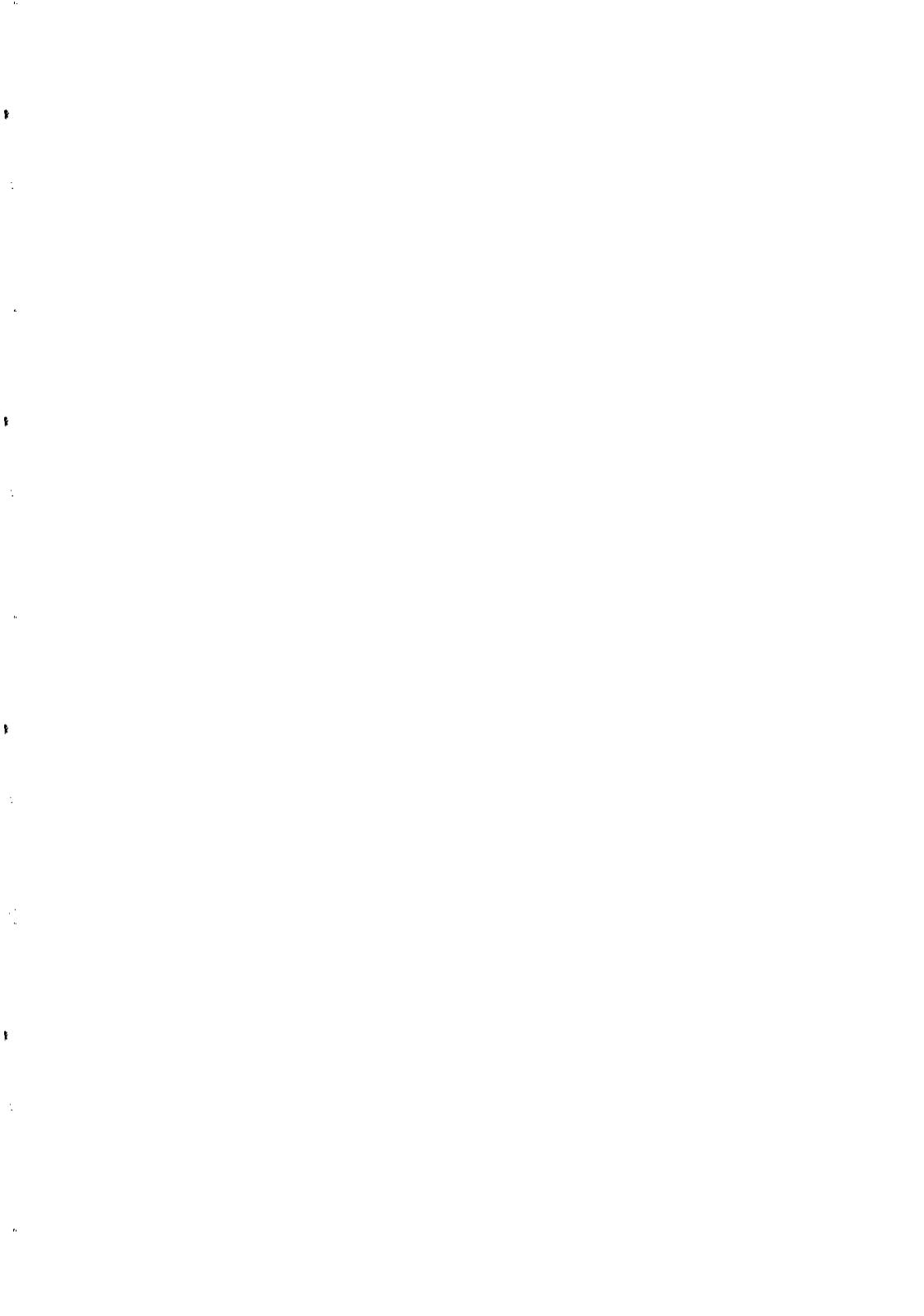
detector	-- relic --	----- WIMP -----	
	Re	Al	Sn
event rate (mo ⁻¹):	1	42	4.2 x 10 ²
active mass (g):	10 ³	10	10
mass/plane (g):	0.823	0.108	0.296
total plane mass (g):	3.59	2.48	2.98
plane no.:	1.2 x 10 ⁴	926	338
stack height (m):	19	1.5	0.54

As evident, either search is electronics-intensive. Eq. (1) suggests a factor 5 or more improvement by simply operating at 100 mK, with a corresponding reduction in the above Table entries; this however remains to be demonstrated. Our work continues.

We thank the remainder of our collaboration -- D. Limagne, V. Jeudy, H. Dubos, and O. Heres -- for the efforts which make this report possible.

References

1. E. Fiorini, J. Low Temp. Phys. 93 (1993) 189;
2. L. Mosca, these proceedings.
3. A.K. Drukier and L. Stodolsky, Phys. Rev. D30 (1984) 2295.
4. A. de Bellefon et al, in Low Temperature Detectors for Neutrinos and Dark Matter II, ed. L. Gonzales-Mestres and D. Perret-Gallix (Ed. Frontieres, Gif-sur-Yvette, 1988), p. 425; M. Furlan et al, in Low Temperature Detectors for Neutrinos and Dark Matter III, ed. N.E. Booth and G.L. Salmon (Ed. Frontieres, Gif-sur-Yvette, 1991), p. 21.
5. A. de Bellefon et al, in Superconductive Particle Detection, ed. A. Barone (World Scientific, Singapore, 1988) p. 241.
6. see for example, TA Girard et al, J. Low Temp. Phys. 93 (1993) 479.
7. TA Girard, in Proc. XXVII Rencontres de Moriond, ed. Tran Tanh Van, T. Damour, E. Hinds and J. Wilkerson (Editions Frontieres, Gif-sur-Yvette, 1993), p. 25.
8. G. Waysand et al, J. Low Temp. Phys. 93 (1993) 485.
9. TA Girard et al, Nucl. Instr. & Meth. A334 (1993) 645.
10. O. Heres et al, J. Low Temp. Phys. 93 (1993) 449.
11. M Abplanalp et al, J. Low Temp. Phys. 93 (1993) 491.
12. M. Goodman and E. Witten, Phys. Rev. D31 (1985) 3059.
13. J.M. Irvine and R. Humphreys, J. Phys. G: Nucl. Phys. 9 (1983) 847.
14. P.F. Smith, Ann. NY Acad. Sci. (1991) 425.



Cosmoparticle physics of family symmetry breaking

Maxim Yu. Khlopov, A.S. Sakharov

CENTER FOR COSMOPARTICLE PHYSICS "COSMION"

Russia, 125047, Moscow, Miusskaya Pl. 4

The models of gauge theory with family symmetry breaking, reproducing the observed properties of quarks and leptons, provide the unified description for the main candidates of dark matter particles: neutrino with the mass 20 eV as hot dark matter, archions (axions, being simultaneously singlet Majorons and familons) as cold dark matter, and unstable neutrino with mass 100 eV, decaying on lighter neutrino and archions with the lifetime 10^{16} s as unstable dark matter. The mechanisms for baryosynthesis and inflation are also provided by the models. The choice between these cosmological models, depending on the unknown a priori scale of the horizontal symmetry breaking can be made unambiguously in the combination of search for rare archion decays and astronomical studies of the dark matter in the Universe.

The foundations of both particle theory and cosmology are hidden at super energy scale and can not be tested by direct laboratory means. Cosmoparticle physics is developed to probe these foundations by the proper combination of their indirect effects, thus providing definite conclusions on their reliability. Cosmological and astrophysical tests turn to be complementary to laboratory searches of rare processes, induced by new physics, as it can be seen in the case of gauge theory of broken symmetry of quark and lepton families, ascribing to the hierarchy of the horizontal symmetry breaking the observed hierarchy of masses and the mixing between quark and lepton families.

The problem of fermion families (generations) remains one of the central problems of particle physics. The standard $SU(3) * SU(2) * U(1)$ model, so as its possible 'vertical' extensions (in the framework of one generation) of the type $SU(5)$, $SO(10)$ etc., does not contain any deep physical grounds both for the existence of mass hierarchy between generations and the observed weak mixing of quarks and leptons owing to arbitrary Yukawa couplings. The identity of quark and lepton families:

$$(u, d, e, \nu_e); \quad (c, s, \mu, \nu_\mu); \quad (t, b, \tau, \nu_\tau) \quad (1)$$

relative to strong and electroweak interactions strongly suggests the existence of 'horizontal' symmetry between these generations. The concept of local horizontal symmetry $SU(3)_H$, first proposed in [1], is reasonable to be considered with left-handed quark and lepton components transforming as $SU(3)_H$ triplets, and the right-handed ones transforming as antitriplets. Their mass terms transform as $3 * 3 = \bar{3} + 6$ and consequently may arise as a result of $SU(3)$ breaking only. (Generalization on the case of n generations, $SU(n)$, is trivial.)

In this approach the hypothesis is reasonable, that the structure of these matrices is determined by the pattern of horizontal symmetry breaking (i.e., by the structure of vacuum expectation values (VEV) of horizontal scalars, maintaining $SU(3)$ breaking) and the mass hierarchy between generations is related to a definite hierarchy in this breaking (the hypothesis of horizontal hierarchy - HHH) [2,3].

The simplest realisation of HHH invokes the introduction of additional superheavy fermions, acquiring their masses via direct coupling with horizontal scalars. The ordinary quark and lepton masses are induced by their "see-saw" mixing [3] with these heavy fermions.

The concept of grand unification (GUT) is another argument in favor of chiral G symmetry. In the GUT models left-handed quarks and leptons are put together with antiparticles of their right-handed component into the same irreducible representations of GUT group G_{GUT} . So in the framework of $G_{GUT} * G_H$ symmetry left-handed and right-handed components must transform as conjugated representations of G_H , i.e., G_H symmetry must be chiral.

One may hope, that complete unification of horizontal and vertical symmetries will be achieved on the base of unifying fundamental symmetry G , including $G_{GUT} * G_H$ in the course of superstring theory development. Though the most elaborated simplest variant of realistic superstring model $E_8 * E'_8$ [4,5] does not leave any room inclusion of horizontal symmetry, such inclusion is possible within the frame of wider class of superstring models, c.f., in $SO(32)$ or in heterotic string models with direct compactification to 4-dimensional spacetime [6]. In the latter case [6] a wide class of GUT groups with the rank smaller than 22 is possible. The analysis of broken horizontal symmetry, given in the present paper, may be useful for the choice of realistic models from this variety of possibilities.

Here we do not consider supersymmetric extensions of the model, in which a number of new particles are predicted, extending the hidden sector of the theory. Properties of such particles depend critically on details of supersymmetry breaking and need special detailed study.

To build the realistic model of broken horizontal symmetry, rather wide set of parameters is to be introduced. But: i) the number of these parameters is smaller than in the realistic models without horizontal symmetry; ii) the bulk of these parameters is fixed by the experimental data on the quark and lepton properties, and, finally, iii) the set of new nontrivial physical phenomena, predicted by the model, provides in principle complete check of the model and determination of all the parameters. These new phenomena arise at a high energy scale of horizontal symmetry breaking $> 10^5 - 10^6 \text{ GeV}$, which can not be achieved even in the far future at accelerators. However combination of experimental searches of their indirect effects in the processes with known particles with the analysis of their cosmological and astrophysical effects makes it possible to study physics, predicted at this scale.

The proposed model satisfies the following naturality conditions:

a) natural suppression of flavour changing neutral currents (FCNC) [7]. There are absent light scalar $SU(2) \times U(1)$ doublets, transforming according to nontrivial representations of G_H (vertical-horizontal fields) and leading to unacceptably strong FCNC. Yukawa couplings, responsible for quark and lepton mass generation, include the only standard $SU(2) \times U(1)$ Higgs doublet, though left-handed and right-handed fermion components transform as conjugate representations of horizontal symmetry group. The price for it is the introduction of additional superheavy fermions, maintaining the hidden sector of the theory. Quark and lepton masses are induced by mixing with these heavy fermions. That is the so called 'see-saw' mechanism [8] is realized not only for neutrinos, but for all the quarks and leptons either. Besides that, our approach provides natural explanation of the hierarchy of electroweak and GUT scales on base of mechanisms suggested for GUT with one generation, c.f., by means of supersymmetric extensions of $SU(5)$, $SO(10)$ etc. [9]. Quark and lepton mass generation by means of vertical-horizontal fields would have needed unnatural fine tuning of their parameters even in the supersymmetric case;

b) natural horizontal hierarchy. The observed mass hierarchy of families (c.f., $m_e : m_\mu : m_\tau = 1 : 200 : 4000$ etc.) is explained by much more moderate hierarchy of horizontal symmetry breaking. The parameters of such breaking are proportional to m for quarks and leptons. So there is no need in special mechanisms to protect the hierarchy from loop corrections;

c) natural solution for QCD CP-violation problem [10]. Through there is the only Higgs doublets present, the theory provides natural inclusion of $U(1)$ Peccei-Quinn symmetry [10], being associated with heavy Higgs fields, breaking G at scale. Breaking of this global $U(1)$ symmetry results in the existence of pseudo-Goldstone boson α of invisible axion type with interaction scale v_H [11]. α has both flavour-diagonal and flavour-nondiagonal coupling with quarks and leptons, i.e., is simultaneously familon [12,13]. Finally, it is related to neutrino Majorana mass generation, being in fact Majoron of singlet type [14].

The model inevitable consequences are:

a) flavour changing neutral transitions, related to axion and horizontal gauge bosons interactions;

b) the existence of neutrino Majorana mass and of the neutrino mass hierarchy of different families;

c) the instability of heavier neutrino relative to axion decay on lighter neutrino;

d) the existence of metastable superheavy fermions.

The presented model may be checked in the combination of laboratory tests (the search for neutrino mass, for neutrino oscillations and for $2\beta_{0\nu}$ decay, the study of $\bar{K}^0 - K^0$ and $\bar{B}^0 - B^0$ transitions, the search for axion decays $\mu \rightarrow e\alpha$, $K \rightarrow \pi\alpha$ etc.) and of analysis of cosmological and astrophysical effects of its predictions. The latter includes study of axion emission effects on stellar evolution, investigation of primordial axion field and massive unstable neutrino effects

on the dynamics of cosmological large scale structure formation, as well as the analysis of the mechanisms of inflation and baryogenesis, based on the hidden sector of the model.

Consider $SU(2) * U(1)$ model with local chiral horizontal symmetry $SU(3)_H$ [1,15] between the families (1). Quarks and leptons are put into the following representations of $SU(2) * U(1) * SU(3)_H$:

$$f_{L\alpha} : \begin{pmatrix} u \\ d \end{pmatrix}_{L\alpha} (2, 1/3, 3), \quad \begin{pmatrix} \nu \\ e \end{pmatrix}_{L\alpha} (2, -1, 3) \quad (2)$$

$$f_R^\alpha : u_R^\alpha(1, 4/3, \bar{3}), \quad d_R^\alpha(1, -2/3, \bar{3}), \quad e_R^\alpha(1, -2, \bar{3})$$

where we retain family ($SU(3)_H$) index: $\alpha = 1, 2, 3$.

We can choose scalars, breaking the horizontal symmetry, as $SU(3)_H$ sextets and triplets. All of them are to be $SU(2) * U(1)$ singlets, to prevent electroweak symmetry breaking of $SU(3)_H$ scale. To generate realistic quark and lepton mass matrices at least three such 'horizontal' scalars are needed. At least one of them with the greatest VEV is to be sextet: $\xi_{\alpha,\beta}^{(0)}$ $\alpha, \beta = 1, 2, 3$. Otherwise, triplet fields only don't generate realistic mass matrices. For the other two scalars $\xi^{(1)}$ and $\xi^{(2)}$ may be concretize further their $SU(3)$ content, mentioning only those cases, when sextet and triplet representations result in different consequences.

Let us introduce additional fermions in the form [3,15]

$$F_L^\alpha : U_L^\alpha(1, 4/3, \bar{3}); \quad D_L^\alpha(1, -2/3, \bar{3}); \quad E_L^\alpha(1, -2, \bar{3}) \quad (3)$$

$$F_{R\alpha} : \quad U_{R\alpha}(1, 4/3, 3); \quad D_{R\alpha}(1, -2/3, 3); \quad E_{R\alpha}(1, -2, 3); \quad N_{R\alpha}(1, 0, 3)$$

Note, that these fermions cancel the $SU(3)_H$ anomaly of quarks and leptons (2). The most general Yukawa couplings allowed by the symmetry are

$$g_f \tilde{f}_{L\alpha} F_{R\alpha} \phi^0 + G_F^{(n)} \tilde{F}_L^\alpha F_{R\beta} \tilde{\xi}^{(n)\alpha\beta} + G_\eta \tilde{F}_L^\alpha f_R^\alpha \eta + h.c.; n = 0, 1, 2 \quad (4)$$

for quarks and leptons [3,15] ($f = u, d, e; F = U, D, E$) and

$$g_\nu \tilde{\nu}_{L\alpha} N_{R\alpha} \phi^0 + G_N^{(n)} N_{R\alpha} C N_{R\beta} \tilde{\xi}^{(n)\alpha\beta} + h.c. \quad (5)$$

for neutrinos [15,16]. Here ϕ^0 is the neutral component of the standard $SU(2) * U(1)$ Higgs doublet $(2, -1, 1)$ ($\langle \phi^0 \rangle = v = (\sqrt{8} G_F) = 250 \text{ GeV}$) and η is real $SU(2) * U(1) * SU(3)_H$ singlet scalar ($(\eta) = \mu/G_\eta$)

Yukawa couplings (4),(5) are invariant relative to global axial $U(1)_H$ transformations:

$$f_L \rightarrow f_L \exp(i\omega), \quad f_R \rightarrow f_R \exp(-i\omega), \quad F_L \rightarrow F_L \exp(-i\omega) \quad (6)$$

$$F_R \rightarrow F_R \exp(i\omega), \quad \phi \rightarrow \phi, \quad \xi^{(n)} \rightarrow \xi^{(n)} \exp(2i\omega); \quad n = 0, 1, 2$$

This $U(1)$ symmetry will be maintained also by the Higgs potential, provided that there are trilinear couplings such as $\Lambda \xi_{\alpha,\beta}^{(0)} \xi^{(1)\alpha} \xi^{(1)\beta} + h.c.$ etc. These couplings are not induced by any other (gauge or Yukawa) interactions. So their absence in the Lagrangian is natural [15,16]

The analysis of the Higgs potential (see [2,15,16] for details) shows, that the VEV matrix can obtain the form:

$$V_H = \langle \xi^{(0)} + \xi^{(1)} + \xi^{(2)} \rangle = \begin{pmatrix} r_1 & p_1 & p_3 \\ +(-)p_1 & r_2 & p_2 \\ +(-)p_3 & +(-)p_2 & r_3 \end{pmatrix} \quad (7)$$

(where '+' and '-' signs correspond to sextet and triplet - and , respectively) with the natural 5-10 fold hierarchy of their values:

$$r_1 > p_1 > r_2 > p_2 > p_3 > r_3 \quad (8)$$

Inserting the scalar VEVs into Yukawa couplings (4),(5) one obtains full 6*6 fermion mass matrices:

$$\begin{pmatrix} \tilde{f}_L \\ \tilde{F}_L \end{pmatrix} \begin{pmatrix} f_R & F_R \\ 0 & g_f v \\ \mu & M_F \end{pmatrix}; \quad \begin{pmatrix} \tilde{\nu}_L \\ \tilde{N}_R \end{pmatrix} \begin{pmatrix} \nu_R & N_R \\ 0 & g_\nu v \\ g_\nu v & M_N \end{pmatrix} \quad (9)$$

Here $M_F = \Sigma \langle \xi^{(n)} \rangle G_F^{(n)}$, $F = U, D, E, N$; and $N_L = C\tilde{N}_R$, $\nu_R = C\tilde{\nu}_L$. Note, that only sextet scalars contribute into Majorana mass matrix M_N . So, one has Dirac "see-saw" mechanism of the quark and lepton mass generation and ordinary Majorana "see-saw" mechanism for neutrino mass term where $N_{R\alpha}$ play the role of right-handed neutrino. The quark and lepton mass matrices obtained from the block-diagonalization of (9) will have the form:

$$m_f = g_f v \mu M_F^{-10}; \quad f = u, d, e; \quad m_\nu = (g_\nu v)^2 M_N^{-1} \quad (10)$$

So the mass hierarchy between the families appears to be inverted with respect to hierarchy of $SU(3)_H * U(1)_H$ symmetry breaking:

$$SU(3)_H * U(1)_H [v_H \sim r_1] \rightarrow SU(2)_H * U(1)'_H [v'_H \sim p_1] \rightarrow U(1)''_H [v''_H \sim p_2] \rightarrow I \quad (11)$$

where. Here the intermediate $SU(2)_H * U(1)'_H$ horizontal symmetry is maintained between the second and the third generations of quarks and leptons: and the remaining global $U(1)''_H$ is appropriate to the third generation only. The considered case is called the inverse hierarchy model in contrast with the direct hierarchy model, in which the quark and lepton mass hierarchy is parallel with hierarchy of $SU(3)_H * U(1)_H$ symmetry breaking [15,17]. The global $U(1)_H$ ($U(1)''_H$) symmetry breaking results in the existence of massless Goldstone boson, a , named archion, having both flavour diagonal and flavour nondiagonal couplings with quarks and leptons and thus being familon of the type [13,18]. For sextet $\xi^{(1)}$ and $\xi^{(2)}$ its main couplings with e.g.charged leptons have the form

$$-i\alpha(g_{\tau\tau}\tilde{\tau}\gamma_5\tau + g_{\tau\mu}\tilde{\tau}\gamma_5\mu + g_{re}\tilde{r}\gamma_5e + g_{\mu\mu}\tilde{\mu}\gamma_5\mu + g_{\mu e}\tilde{\mu}\gamma_5e + g_{ee}\tilde{e}\gamma_5e) + h.c \quad (12)$$

where

$$\begin{aligned} g_{\tau\tau} &= m_\tau/v_H, & g_{\tau\mu} &\approx m_{13}^t/v_H'', \\ g_{re} &\approx m_{13}^t/v_H'', & g_{\mu\mu} &< m_\mu/v_H', \\ g_{ee} &< (m_e/m_\tau)(m_l/v_H'') \\ g_{e\mu} &< (m_\mu/m_\tau)^{1/2}(m_{12}^t/v_H'') \end{aligned}$$

For triplet $\xi^{(1)}$ and $\xi^{(2)}$ flavour nondiagonal coupling are scalars. In the tree approximation the couplings of a with quarks are similar to (12). In our minimal $SU(2) * U(1) * SU(3)_H$ model is the particle of the arion type [19]. It has no couplings $a\gamma\gamma$ and agg induced by triangle diagrams. Its interactions with the ordinary matter are suppressed sufficiently to remove the

strong astrophysical restrictions [20] on the scale v_H^* . So the flavour changing decays may go with noticeable probability [15], e.g.

$$\mu \rightarrow ea, \quad \tau \rightarrow \mu a, \quad K \rightarrow \pi a, \quad B \rightarrow K(K^*)a, \quad D \rightarrow \pi(p)a \quad (13)$$

The search of such decays could provide the valuable information about the structure of fermion mass matrices.

In any realistic extension of our scheme, c.f. in $SU(5) * SU(3)_H$ model triangle diagrams owing to the inevitable presence of additional heavy fermions (GUT) induce $a\gamma\gamma$ and agg vertices, so that $U(1)_H^*$ turns to be Peccei-Quinn symmetry [10] and archion becomes the invisible axion of nearly hadronic type [21]. The scale $v_H^* = v_{PQ}$ is then restricted from below by astrophysical estimations of stellar energy losses due to archion emission: $v_{PQ} > 10^6 \text{ GeV}$ (sun and red giants [22,20,15]) and $v_{PQ} > 10^{10} \text{ GeV}$ (supernova SN1987A [23]). The latter restriction seems to be taken with caution.

According to (10) the hierarchy of neutrino Majorana masses is similar to the ordinary quark and lepton mass hierarchy: $m_{\nu_e} : m_{\nu_\mu} : m_{\nu_\tau} \sim m_e : m_\mu : m_\tau$. For sextet $\xi^{(1)}$ and $\xi^{(2)}$ neutrino mass matrix m_ν (10) is nondiagonal and archion decays are possible with the lifetimes $\tau(\nu_H \rightarrow \nu_L a) = 16\pi/g_{HL}^2 m_H$, where $g_{HL} = g_{\nu_H \nu_L} = m_{HL} v_{PQ}$.

For "small" scale of family symmetry breaking $v_{PQ} \sim 10^6 \text{ GeV}$ the predicted effect of Majorana masses of neutrino is rather close to the modern sensitivity of $2\beta_\nu$ searches.

Since the archion couplings to fermions of the lightest family (u, d, e) are suppressed, the existing constraints on the respective scale are weakened to $v_{PQ} > 10^6 \text{ GeV}$, making the model of archion rather close to the model of hadronic axion [21]. However, it turns out [24], that archion model escapes the serious problem of primordial superheavy stable Q quarks, predicted in the model of hadronic axion [24,21,25]. One can estimate the frozen concentration of Q quarks and respective Q hadrons in the Universe and find it contradicting [24] the upper limits on such concentration, following from the search for anomalous nuclei (so called "crazy isotopes"). So the theory should introduce the mechanism of superheavy quark instability. But the inclusion of the hadronic axion model into GUT models leads inevitably to the existence of superheavy lepton, coupled to axion. Then, the mixing of superheavy quark Q with the light (ordinary) quarks, inducing Q instability, would lead to the existence at the tree level of the axion coupling to leptons, so that axion is not hadronic. In view of these troubles of the model of hadronic axion the model of archion is of special interest, since it naturally provides both superheavy quark instability and the suppression of the axion coupling to leptons.

Since the mass of neutrino $m_\nu \sim v_{PQ}^{-1}$, its lifetime $\tau \sim v_{PQ}^5$ and the density of primordial axion field [26] $\rho_a = \bar{\rho} \sim v_{PQ}$, at larger v_{PQ} axion field is to dominate in the Universe, massive stable neutrino dominancy corresponds to smaller v_{PQ} , and, finally, the smallest possible v_{PQ} correspond to cosmological models with massive unstable neutrino [28,29,24]. So changing the parameter v_{PQ} one reproduces all the main types of cosmological models of the formation of the structure of the Universe. In our approach continuous change of v_{PQ} results in continuous transition from one to another form of dark matter, dominating in the Universe, and in definite predictions of the model for each type of dark matter, corresponding to the combination of respective cosmological, astrophysical and physical constraints.

The total cosmological density ρ_{tot} and the baryon density ρ_B being fixed, relationship [15,24]

$$\rho_a^{prim}(v_{PQ}) + \Sigma \rho_{\nu_i}^{prim}(v_{PQ}) + \rho_a^{dec}(v_{PQ}) + \Sigma \rho_{\nu_i}^{dec}(v_{PQ}) + \rho_B = \rho_{tot} \quad (15)$$

turns to be an equation relative to v_{PQ} .

The solutions of this equation define a discrete set of cosmological models with different types of dark matter, forming the structure of the Universe. In general, there are six different dark matter scenarios [15,24] which may be realized in the framework of the considered model.

1. Cold dark matter (CDM) scenario. The cosmological evolution of axion field after Peccei-Quinn symmetry breaking follows basically in our approach general features of the standard model of 'invisible' axions [26]. $SU(3)_H \times U(1)_H$ symmetry breaking leads $U(1)_H$ to be in fact Peccei-Quinn symmetry of only one quark-lepton generation, what resolves automatically the cosmological θ -domain problem in the axion theory [30]. Stochastic distribution of the axion field: $\theta = a/v_{PQ}$ may lead to its change on 2π along a closed path, resulting in the appearance of string structure, decaying rapidly after the axion mass switches on at $T < 800\text{MeV}$ [31]. Then axion field oscillations with the averaged amplitude $\bar{\theta} \sim 1$ start in the Universe. In the standard model of invisible axion the density decreases in the course of expansion as $\rho_a \sim a^{-3}$, where a is the scale factor, and the modern axion mass density is equal to [26]:

$$\rho_a = (v_{PQ}/4 \cdot 10^{12} \text{GeV}) \rho_{cr}, \quad \rho_{cr} = \frac{3H^2}{8\pi G} \quad (16)$$

According to [31] intensive axion emission by decaying axion cosmic string structure result in the growth of the modern cosmological axion density up to

$$\rho_a = (v_{PQ}/2 \cdot 10^{10} \text{GeV}) \rho_{cr} \quad (17)$$

Comparing these predictions with the total cosmological density, taken to be equal to ρ_{cr} , one obtains the upper limit on v_{PQ} . However, within the frame of inflationary cosmology the situation is possible, when the axion field in the observed part of the Universe has an amplitude $\theta \ll 1$ [27]. Exponential expansion of the region with $\theta \ll 1$ provides also the absence of axion strings in this region, so that there is no increase of the axion density due to axion emission of such strings.

Cosmological upper limits on the scale v_{PQ} seem to be absent in this case, so that this scale may have reached even planckian values $v_{PQ} \sim m_{pl}$. But in our model the scale v_{PQ} may be constrained even in this case, taking into account the condition on the absence of phase transition at the inflationary stage. Based on the chaotic inflation scenario, in order to avoid the peaks in density fluctuations one should exclude the possibility of horizontal symmetry phase transitions at the inflationary stage, what leads to the restriction $v_{PQ} < 210^{10} \text{GeV}$ [32].

It was recently shown [37], that the initial distribution of θ changing on 2π around the string implies the inhomogeneity of the amplitude of coherent axion field oscillations relative to the true vacuum, being proportional to $\theta - \theta_{vac}$, where $\theta_{vac} = 2\pi n$ with n being integer. The large scale distribution of these primordial inhomogeneities named archioles, reflects the vacuum axion walls-surrounded-by-strings structure, formed when the axion mass is "switched on" at $T \sim 800\text{MeV}$. Owing to superweak self-interaction of invisible axion the vacuum walls-strings structure and archioles split and their successive evolution goes separately. The vacuum walls-surrounded-by-strings structure is known to disappear rapidly due to gravitational radiation, and the large scale structure of archioles freezes out at the radiation dominance stage. Archioles reflect the original Brownian nature of axion strings, having at each scale about 80length in the form of infinite string, stretching out the region of the considered size [38]. So the archioles form the fractal structure, causing inhomogeneities at all the scale. Putting aside the small scale evolution of archioles, estimations [37] show, that the large scale inhomogeneities induced by archioles can not be smaller than $\delta \sim 10^{-2} (F/10^{10} \text{GeV})$ causing the serious trouble for the cosmological models with even small axionic dark matter admixtures at $F > 10^8 \text{GeV}$ in view of the observed isotropy of relic radiation. According to [24] at $F < 10^8 \text{GeV}$ coherent

axion field oscillations are thermalized due to $aN(\tilde{N}) \rightarrow \pi N(\tilde{N})$ reactions, so that the archioles structure dissipates. Primordial axion field distribution may in this case induce fractal distribution of baryonic charge.

2. Hot dark matter (HDM) scenario. The dominance of ν_τ with the standard concentration $n_\nu = 3/11n_\gamma$, and the mass $m_{\nu_\tau} = 20\text{eV}$ and the lifetime exceeding the age of the Universe ($t_U < \tau(\nu_\tau \rightarrow \nu_\mu a)$):

$$\rho_{\nu_\tau} = \frac{6,6 \cdot 10^{13} \text{GeV}}{v_{PQ}} (g^2/G) \rho_{cr}, g^2/G \sim 10^{-4} - 10^{-6} \quad (18)$$

3. Relativistic unstable dark matter (UDM) scenario. The dominance in the Universe of relativistic archions and ν_μ , being the products of $\nu_\tau \rightarrow \nu_\mu a$ decay of with the mass $m_{\nu_\tau} = 50 - 100\text{eV}$ and lifetime $\tau(\nu_\tau \rightarrow \nu_\mu a) = 4 \cdot 10^{15} - 10^{16}\text{s}$ ($m_{\nu_\mu} < 5\text{eV}$):

$$\rho_{\nu_\mu+a}^{rel} = \frac{(v_{PQ}/10^{10} \text{GeV})^{3/2}}{x(g^2/G)^{1/2}} \rho_{cr} \quad (19)$$

where $x \sim 1$ is neutrino mixing parameter.

4. Nonrelativistic UDM scenario. The dominance of nonrelativistic ν_μ with the mass $\sim 10\text{eV}$, both primordial and from $\nu_\tau \rightarrow \nu_\mu a$ decay of ν_τ with mass $\sim 100\text{eV}$ and the lifetime $\sim 10^{15}\text{s}$, provided that $\tau(\nu_\mu \rightarrow \nu_e a) > t_U$:

$$\rho_{\nu_\mu} = \frac{0,6 \cdot 10^{12} \text{GeV}}{v_{PQ}} (g^2/G) \rho_{cr} \quad (20)$$

5. Relativistic hierarchial decay (HD) scenario. The dominance in the modern Universe of relativistic archions ρ_a^{dec} and ν_e , from decay of ν_{mu} with the mass $m_{\nu_\mu} = 50 - 100\text{eV}$ and lifetime $\tau(\nu_\mu \rightarrow \nu_e a) = 4 \cdot 10^{15} - 10^{16}\text{s}$, under the condition of rapid decay of ν_τ with the mass $m_{\nu_\tau} \sim (1 - 10)\text{keV}$, $\tau(\nu_\tau \rightarrow \nu_\mu a) < (10^8 - 10^{10})\text{s}$:

$$\rho_{\nu_e+a}^{rel} = \frac{(v_{PQ}/10^8 \text{GeV})^{3/2}}{(g_2/G)^{1/2}} \rho_{cr} \quad (21)$$

6. Nonrelativistic HD scenario. The dominance of nonrelativistic or semirelativistic archions, originated from both early ν_τ decay and successive ν_μ decays, provided that $m_a > m_{\nu_e}$. Or, in the other case ($m_a < m_{\nu_e}$) the dominance of nonrelativistic ν_e both primordial and from ν_τ and ν_μ decays:

$$\rho_{\nu_e} = \frac{3,3 \cdot 10^{10} \text{GeV}}{v_{PQ}} (g^2/G) \rho_{cr} \quad (22)$$

In the former case the main contribution into the inhomogeneous dark matter (in rich galaxy clusters and halos of galaxies) is maintained by both primordial thermal archion background and nonrelativistic archions from early ν_τ decays

$$\rho_a = \frac{9 \cdot 10^4 \text{GeV}}{v_{PQ}} \rho_{cr} \quad (23)$$

So the archion model provides the unique physical basis for CDM (primordial axion condensate) UDM (stable ν_τ), relativistic and nonrelativistic UDM and HD scenarios, being realized as the solutions of the Eq (15). The complete set of the solutions of the Eq. (15) can be realized subject to v_{PQ} for $g^2/G \sim (1,5 - 40)10^{-6}$. On the other hand the set of cosmological and astrophysical constraints leaves only two small intervals near i) $v_{PQ} \sim 10^8\text{GeV}$, in which only HD scenarios 5 and 6 can be realized, and ii) $v_{PQ} \sim 10^{10}\text{GeV}$ where scenarios CDM and HDM or their mixture are possible provided that the archioles problem is solved. The UDM scenarios 3 and 4 are excluded from SN1987A restrictions on v_{PQ} .

Note that HD scenarios 5 and 6 combine the attractive features of HDM, CDM and UDM models. It makes HD scenarios appealing physically relevant theoretical basis for detailed models of the cosmological large scale structure formation and for comparison of the predictions of such models with the astronomical data. Indeed in the HD scenarios the dominance of ν_r with the mass (1-10)keV in the period ($10^8 - 10^{10}$)s induces short wave fluctuations in the spectrum of density perturbations of ν_μ with the mass (50-100)eV. ν_μ from $\nu_r \rightarrow \nu_{\mu a}$ decays enhance in this spectrum (by the factor of 2) the long wave component, inherent to HDM models, providing the formulation of clear cell structure of voids and superclusters. Finally, $\nu_\mu \rightarrow \nu_{e a}$ decays at $t \sim (10^{15} - 10^{16})$ s slow down the rate of the evolution of the structure and provide its survival to the present time. The primordial thermal archion background, being in these models the coldest component of the modern dark matter play the important role in the evolution of the shortest wavelength part of density perturbations, inducing, in particular, the formation of massive halos outside the visible parts of galaxies. One should take in mind, that according to [33], the phase space restrictions on the mass of halo particles [34] can be weakened or even completely removed in the case of Bose gas.

The second possibility $v_{PQ} \sim 10^{10}$ GeV corresponds to more conservative "standard" CDM model of large scale structure formation with possible modification to mixed CDM + HDM scenario. Note that HD scenario also provides effective mixture of CDM + HDM, so that the both possibilities provide natural basis for refined mixed CDM + HDM model of large scale structure formation. The account for archioles problem seem to reduce the space of free parameters of the model to the only possibility of HD scenario.

The recent indications on the existence of the anisotropy of microwave thermal background, claimed in COBE experiment [35] seem to favour such mixed scenario. They also favour "flat" Harrison-Zeldovich spectrum, predicted by simple one-field inflational models (by chaotic inflational scenario, in particular). Such a scenario can find its grounds in the framework of the presented model, since the singlet Higgs field η , determining the flavour independent mass term $\mu = G_\eta(\eta)$ may self consistently play the role of inflaton at $\langle \eta \rangle \sim m_{pl}$ and $G_\eta < v_{PQ}/m_{pl}$. The predicted spectrum of density fluctuations practically coincides with the "flat" one at $v_{PQ} < 10^{10}$ GeV.

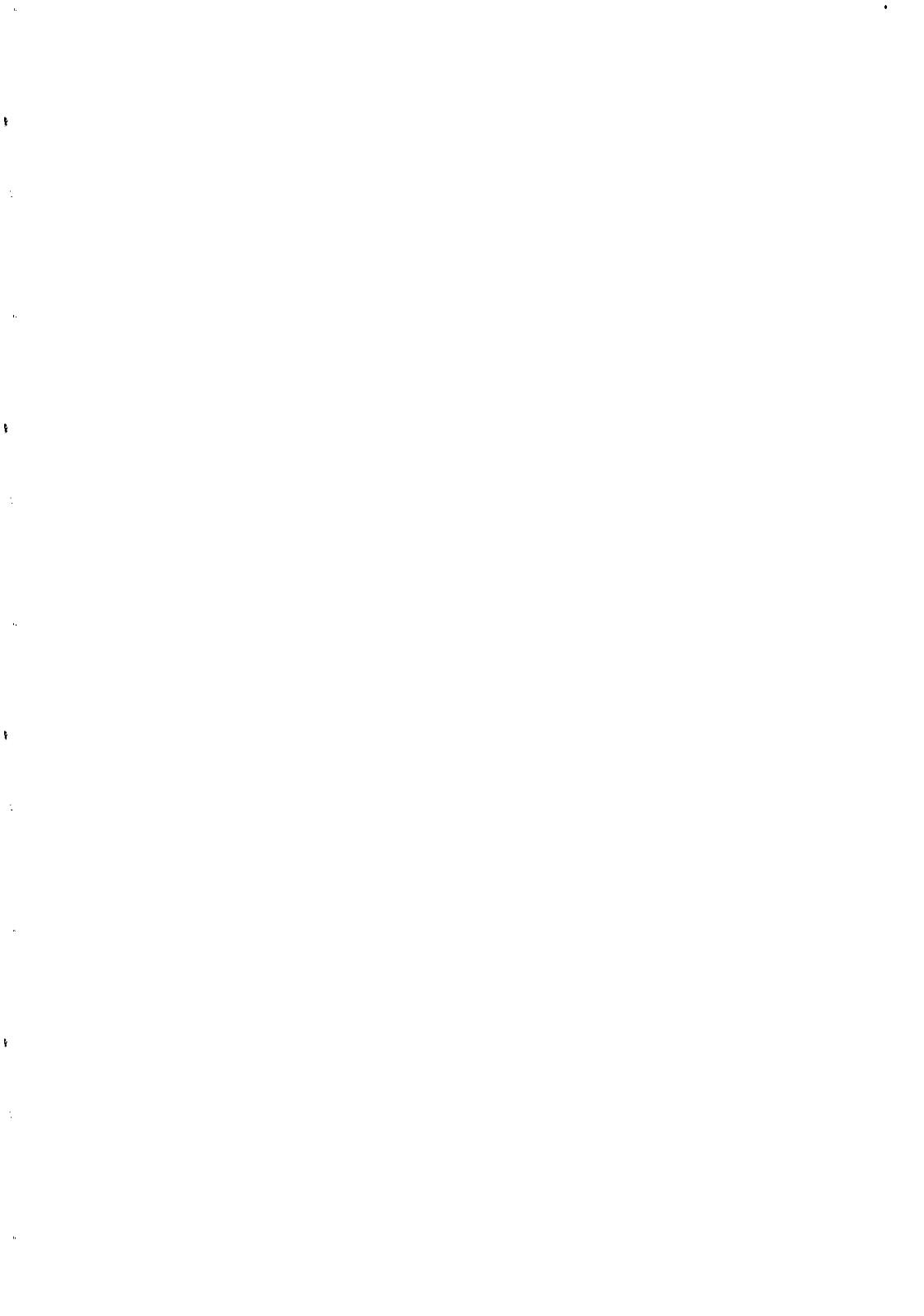
Even at the presented level of "horizontal" $SU(2) \times U(1) \times SU(3)_H$ gauge unification the model provides the mechanism for baryogenesis without GUT-induced baryon nonconservation. The mechanism combines (B+L) nonperturbative electroweak nonconservation at high temperatures with $\Delta L = 2$ nonequilibrium transitions, induced by Majorana neutrino interactions. Estimations [36] show, that the mechanism can, in principle, reproduce the observed baryon asymmetry of the Universe for the allowed parameters of the model. So the proposed model provides unified fundamental basis for theoretical description both of the structure of elementary particles and of the structure of the Universe. Such unified approach to cosmological and particle phenomena is the brightest feature of new science - cosmoparticle physics, forming last years in the confrontation of particle theory and cosmology. Unifying the separate results of studies of partial problems of cosmology and particle physics, the proposed model seems to be the first step on the way towards realistic unified description of unique fundamental grounds of the micro- and macro- world structure on the basis of flavourdynamics.

Our way to the highlights of the theory, based on the detailed elaboration of its 'low energy' basis, may give valuable recommendations for the choice of realistic variant of the complete unified 'theory of everything' (superstring theory, for example), what seems to be of sure importance in view of the existing theoretical uncertainties in the searches for fundamental grounds of physics and cosmology.

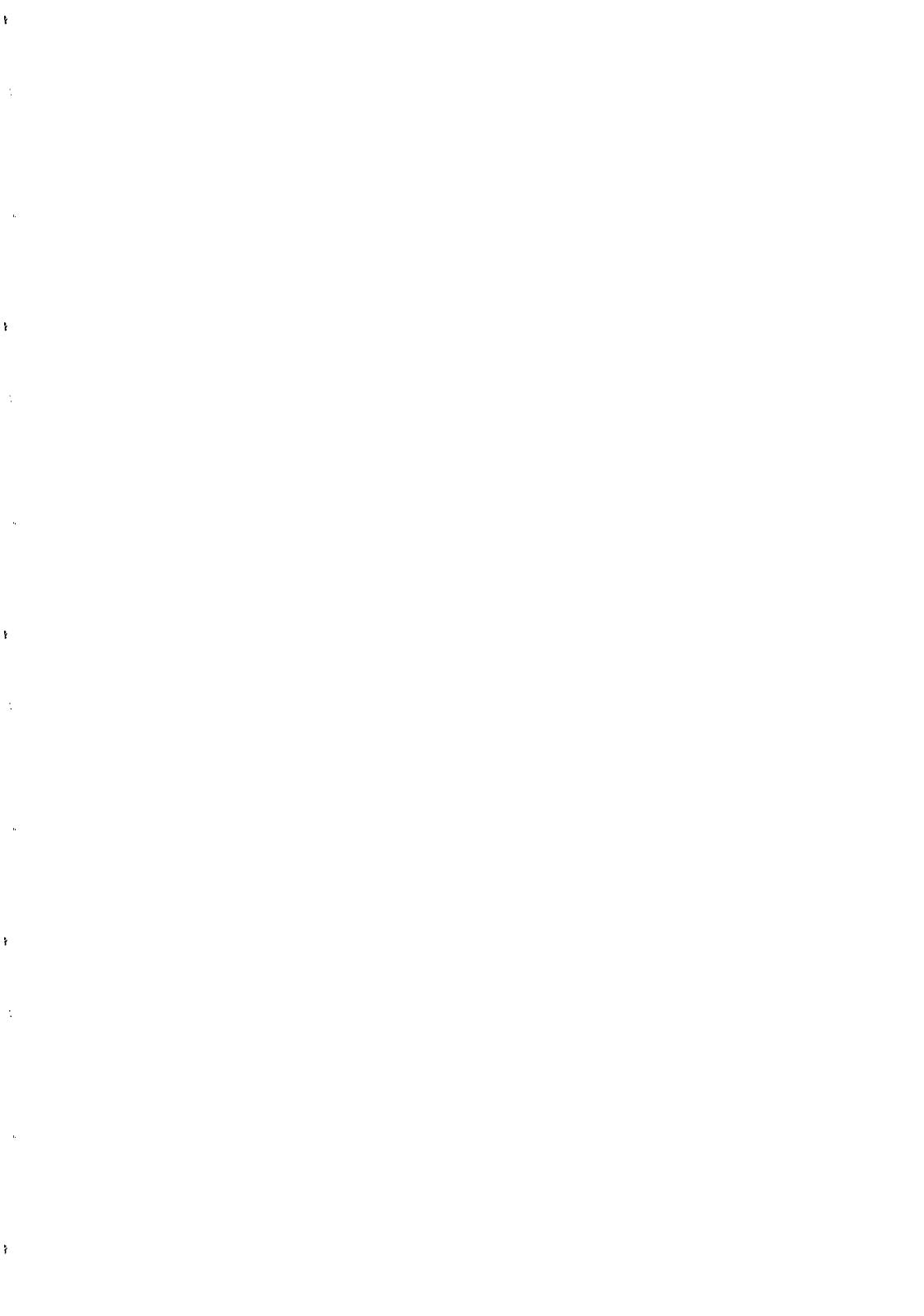
The important epistemological aspect of the presented studies is to be pointed out. We have demonstrated the principal possibility of detailed study of multiparameter "hidden" sector of particle theory. The example of the QFD with low energy scale of family symmetry breaking implies the hope, that multiparameter model of superhighenergy physics, being elaborated in details, will lead to the amount of indirect effects, accessible to experimental and observational tests, exceeding the number of independent parameters of the theory, so that overdetermined system of equations relative to these parameters can be deduced from the set of tests of the predictions of the model considered. So the general approach to the experimental test of the theory, based on the overdetermined system of equations for unknown theoretical parameters, can be realized in the framework of cosmoparticle physics. The analysis of the combination of effects, predicted by the theory provides its detailed study in the case, than direct experimental test is impossible.

- REFERENCES
1. J.K. Chkareuli: Pis'ma ZhETF, 32 (1980), 684.
 2. Z.G. Berezhiani, J.K. Chkareuli: Pis'ma ZhETF, 35 (1982),494; Yadernaya Fizika, 37 (1983), 1043.
 3. Z.G. Berezhiani: Phys.Lett., 129B (1983),99.
 4. P.Candelas, G.Horowitz, A.Strominger, E.Witten: Nucl.Phys., B258 (1985), 46.
 5. E.Witten: Nucl.Phys., B258 (1985), 75.
 6. K.S.Narain: Phys.Lett., 169B (1986), 41; H.Kawai, D.Lewellen, S.H.H.Tye: Phys.Rev.Lett., 57 (1986),1832.
 7. S.L.Glashow, S.Weinberg: Phys.Rev., D15 (1977), 1958.
 8. M.Gell-Mann, P.Ramond, R.Slansky, in: 'Supergravity', eds P. Van Nieuwenhuizen and D.Freedman, North Holland (1979), p.186.
 9. M.I.Vysotski: Uspekhi Fiz.Nauk, 147 (1985), 100; H.P.Nilles: Phys.Rep., 86 (1983), 200.
 10. R.D.Peccei, H.R.Quinn: Phys.Rev., D16 (1977), 1791.
 11. A.R.Zhitnicki: Yadernaya Fizika, 31 (1980), 497; M.Dine, F.Fishler, M.Srednicki: Phys.Lett.,104B (1981) 199; M.Wise, H.Georgi, S.L.Glashow: Phys.Rev.Lett., 47 (1981), 402.
 12. F.Wilczek: Phys.Rev.Lett., 49 (1982), 1549.
 13. A.A.Anselm, N.G.Uraltzev: ZhETF, 84 (1983), 1961.
 14. V. Chikashige, R.N.Mahapatra, R.D.Peccei: Phys.Rev.Lett., 45 (1980), 1926.
 15. Z.G.Berezhiani, M.Yu.Khlopov: Yadernaya Fizika 51 (1990) 1157; 51 (1990) 1479; 52 (1990) 96; Z.Phys.C 49 (1991) 73; Z.G.Berezhiani, M.Yu.Khlopov, R.R.Khomeriki: Yadernaya Fizika 52 (1990) 104,538.
 16. Z.G. Berezhiani, J.L.Chkareuli: Pis'ma ZhETF 37 (1983) 285
 17. Z.G.Berezhiani: Phys.Lett., 150B (1985), 177.
 18. A.A.Anselm, Z.G.Berezhiani: Phys.Lett., 162 (1985), 349.
 19. A.A.Anselm: Pis'ma ZhETF 36 (1982) 46; A.A.Anselm, N.G. Uraltzev: Phys.Lett. 116B (1982) 161
 20. J.E. Kim: Phys.Rep. 150 (1987) 1; H.Y. Cheng: Phys.Rep. 158 (1988) 1
 21. J.E.Kim: Phys.Rev.Lett., 43 (1979). 103;M.A.Shifman, A.I.Vainshtein, V.I.Zakharov: Nucl.Phys., B166 (1980).
 22. M.I. Vysotski, Ya.B. Zeldovich. M.Yu.Khlopov, V.M.Chechetkin: Pis'ma ZhETF, 27 (1978), 533; D.S.Dicus, E.W.Kolb, V.L.Teplitz, R.V.Wagoner: Phys.Rev., D18 (1978), 1829; M.Fukugita, S.Watamura: Phys.Rev.Lett., 48 (1982), 1522.
 23. J.Ellis, K.A.Olive: Phys.Lett. 193B (1987) 525; G.Raffelt, D.Seckel: Phys.Rev.Lett 60 (1988) 1793.
 24. Z.G.Berezhiani, A.S.Sakharov, M.Yu.Khlopov: Yadernaya Fizika 55 (1992) 100; Preprint DFTUZ 91.29, 1991.

25. N.Enrico, R. Esteban: *Phys.Lett.* B245 (1990) 109.
26. L.Abott, P.Sikivie: *Phys.Lett.*, 120B (1983), 133; J.Preskill, M.Wise, F.Wilczek: *Phys.Lett.*, 120B (1983) 127; M.Dine, W.Fishler: *Phys.Lett.*, 120B (1983), 137.
27. A.D.Linde: *Phys.Lett.*, 200B (1988), 272.
28. A.G.Doroshkevich, M.Yu.Khlopov: *Mon.Not.Roy.Astron.Soc.*, 211 (1984),277; *Yadernaya Fizika*, 39 (1984), 869.
29. M.S.Turner, G.Steigman, L.M.Krauss: *Phys.Rev.Lett.*, 92 (1984),2090; G.Gelmini, D.N.Schramm, J.W.F.Valle: *Phys.Lett.*, 146B (1984), 311; A.G.Doroshkevich, A.A.Klypin, M.Yu.Khlopov: *Pis'ma Astron.Zh.* 11 (1985), 483; *Astron.Zh.*, 65 (1988), 248; A.G.Doroshkevich, A.A.Klypin, M.Yu.Khlopov: in 'Large Scale Structures of the Universe', eds. A.Szalay, J.Audouze, D.Reidel (1988), p.293.
30. J.Ipser, P.Sikivie: *Phys.Rev.Lett.* 50 (1983) 925.
31. R.L.Davis: *Phys.Lett.* B180 (1986) 225.
32. A.S.Sakharov, M.Yu.Khlopov: *Yadernaya Fizika* 56 (1993) 220.
33. J.Madsen in: *Particle Astrophysics. The early Universe and cosmic structures* ed. J.M.Alimi, A.Blanchard, A.Bouquet, F.Martin de Volnay, J.Tran Thanh Van. Editions Frontieres, 1990, p.205.
34. S.Tremaine, J.E.Gunn: *Phys.Rev.Lett.* 42 (1979) 407.
35. G.F.Smoot et al: *Astrophys. J.* 300 (1992) L90, E.L.Wright et al: *Astrophys.J.* 300 (1992) L100.
36. M.A.Luty: *Phys.Rev.* D45 (1992) 455.
37. A.S.Sakharov, M.Yu.Khlopov: *Yadernaya Fizika* 58 (1994),N-3, to be published.
38. T.Vachaspati, A.Vilenkin: *Phys.Rev.* D30 (1984) 2036.



GRAVITATIONAL LENSING



MACRO AND MICROLENSING

R. D. Blandford

130-33 Caltech

Pasadena, CA 91125, USA

**Abstract**

A brief introduction to recent advances in gravitational lensing is presented. Multiple-imaged quasars, arc images of distant galaxies elongated by intervening clusters and radio rings formed by individual galaxies are all described. Gravitational lenses may be useful for magnifying distant galaxies, for performing cosmographic measurements and for probing dark matter in our Galactic halo, in distant galaxies and clusters and in the universe at large.

1. Introduction

The subject of gravitational lenses, like much else in physics, has its origin in the work of Isaac Newton. The first query of the third book of Newton's *Opticks* is "Do not Bodies act upon Light at a distance, and by their action bend its Rays; and is not their action (*cæteris paribus*) strongest at the least distance?". The answer, as I hope to demonstrate, is yes.

In his first attempts to construct a theory of gravitation, Einstein ¹⁾ essentially adopted Newton's approach and computed that a stellar image would be deflected through an angle $\hat{\alpha}^N = 2GM_\odot/R_\odot c^2 \sim 0.9''$ near the solar limb. However, in 1915 he demonstrated that the general relativistic prediction ²⁾ included an equal space curvature contribution which doubled the total deflection to

$$\hat{\alpha} = \frac{4GM_\odot}{R_\odot c^2} = 1.75'' \quad (1)$$

The famous eclipse expedition of 1919 ³⁾ measured a deflection more consistent with the relativistic prediction and this success led to wider acceptance of relativity among physicists and brought Einstein himself to public attention. (We now know from VLBI measurements ⁴⁾ that the agreement is at least good to 10^{-3} .) Lodge ⁵⁾ (although he rejected general relativity), is generally credited with being the first to realise that the sun, and consequently other celestial bodies could actually focus a distant source. Despite various prophetic theoretical investigations, sixty years were to elapse before the first *bona fide* example of gravitational lensing, Q0957+561, was discovered by Walsh, Carswell & Weymann (1979) ⁶⁾

2. Simple Lenses

In his attempts to explain relativity to an astronomical establishment somewhat unfamiliar with Riemannian geometry and tensor calculus, Eddington ⁷⁾ pointed out that, in the weak field limit, a stationary gravitational field deflects light as if space were flat but endowed with a refractive index $n = 1 - 2\phi/c^2$, where ϕ is the Newtonian gravitational potential. This prescription is adequate for describing cosmic gravitational lenses and, adopting it, we can see that a self-gravitating, transparent body such as a galaxy or a cluster of galaxies will produce a deflection

$$\hat{\alpha} = 4\pi \left(\frac{\sigma_{\parallel}}{c} \right)^2 \simeq 2'' \sigma_{300}^2 \quad (2)$$

where σ_{\parallel} is the characteristic velocity dispersion and σ_{300} is this quantity measured in units of 300 km s^{-1} . For a galaxy, the deflection is typically a couple of arcseconds; for a cluster of galaxies it is an order of magnitude larger.

If we introduce angular diameter distances, connecting the observer to the deflector, D_d , the observer to the source, D_s , and the deflector to the source, D_{ds} , then a ray connecting the source to the observer must satisfy

$$\beta = \theta - \frac{D_{ds}}{D_s} \hat{\alpha}(D_d \theta) \quad (3)$$

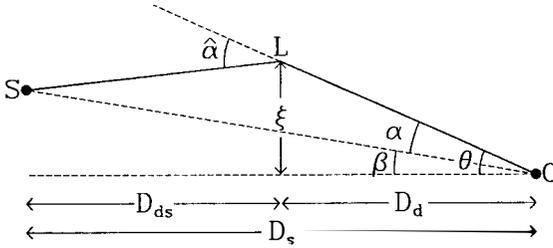


Fig. 1. Basic geometry of gravitational lensing. A ray from the source S is incident upon the deflector or lens L with impact parameter ξ and is deflected through a small angle $\hat{\alpha}$ to reach the observer at O. The apparent position of the image is measured by the angle θ and the true position by β .

where we treat all angles as small and represent them as 2D vectors on the sky.

For a compact lens of mass $M = 10^{12} M_{12} M_{\odot}$, it is useful to introduce the *Einstein radius*

$$\theta_E = \left(\frac{GM}{Dc^2} \right)^{1/2} = 3'' M_{12}^{1/2} D_9^{-1/2} \quad (4)$$

where $D = D_d D_s / D_{ds} = D_9$ Gpc. A point source that lies within a circle centred on the lens of radius θ_E , will produce two comparably bright images. For an extended lens, we can also define a *critical density*

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G D} = 0.35 D_9^{-1} \text{ g cm}^{-2} \quad (5)$$

which is the average mass density in the lens for a background source to create multiple images.

Cosmic lenses differ in two key respects from simple Gaussian lenses in which the deflection of a ray is supposed to be linearly proportional to its impact parameter. The first is that this variation is strongly non-linear. For a point lens, $\hat{\alpha} \propto \theta^{-1}$, for a galaxy, the deflection, although $\propto \theta$ for small impact parameter passes through a maximum and then diminishes to zero. The second difference is that cosmic lenses are non-circular. Both effects contribute to the peculiar properties of observed sources.

3. Types of Gravitational Lenses

3.1 Multiple-Imaged Quasars

The “double quasar”, Q0957+561, was the first example of a multiple-imaged quasar. There are two bright images, separated by $6''$, of a distant quasar with a redshift of $z = 1.4$ formed by an intervening giant galaxy combined with a surrounding cluster with a redshift

$z = 0.36$. Another good example is provided by the $z = 3.3$ radio source 2016+112,⁸⁾ where three images separated by $\sim 3''$ are formed by two intervening galaxies, probably at different redshifts. In total, there are roughly 14 convincing and 8 proposed examples of similar sources with multiple, point-like images. It is often astronomically challenging to decide if lensing really is occurring as spectroscopically similar quasars can occasionally be found in the same cluster. Radio observations are particularly helpful discriminants because gravitational lenses ought to be achromatic and images should show the same flux ratios at all wavelengths. More compelling evidence is provided when four images are located, or when there is resolved structure which is similar, though inverted in two images (as indeed is the case for the VLBI observations of 0957+561)⁹⁾ measuring a correlated variation in two images or locating the lensing mass in the predicted location.

It is easy to show that the probability that a point source is magnified by a factor $> \mu$, is $\propto \mu^{-2}$. As there are many more faint quasars than intrinsically bright quasars, it turns out that many of the brightest gravitational lenses are highly magnified. This effect is known as "magnification bias" and has to be taken into consideration in statistical studies.

3.2 Cluster Arcs and Arclets

The second type of lens is associated with distant rich clusters of galaxies. Yet more distant, individual galaxies, located near their effective foci, can be magnified to form giant luminous arcs. These structures, which have identical surface brightness to an unlensed counterpart but subtend far larger angular sizes so that their fluxes are greatly enhanced, have been found in about 15 clusters as described here by Mellier. (The effect may be simulated by looking at a point source of light axially through a plain wineglass.) Much more numerous than the arcs are the arclets which are formed when a galaxy images are mildly distorted and elongated tangentially by the action of a cluster. Over 50 examples are known associated with the cluster A2218, for example.

3.3 Radio Rings

A related phenomenon is the radio ring. Here the extended object is a radio source and the lens is an individual galaxy. Often, nearly complete "Einstein rings" are formed. 1131+0456 is a prototypical example.¹⁰⁾ In general, there is far more information in a radio ring than in a multiple-imaged quasar and by matching parts of the image with the same surface brightness, it is possible to derive quite accurate models of the lens and source.

4. Uses of Gravitational Lenses

There are some parallels between the study of gravitational lenses and that of pulsars. After a long hiatus, a theoretical idea, neutron stars became observational reality and a decade

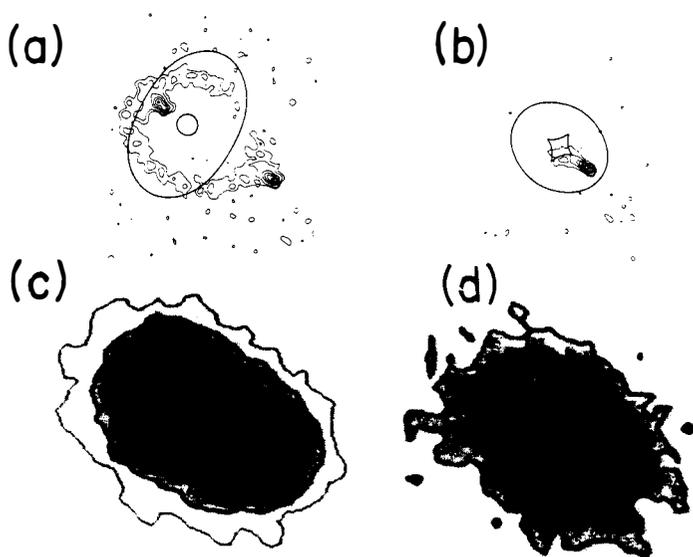


Fig. 2. (a) 15GHz radio image of the radio ring MG1131+0456¹⁰⁾. (b) Reconstructed source model¹¹⁾ (c) 2 μ image showing the two quasar images corresponding to the two compact radio components in (a)¹²⁾ (d) 1 μ image showing the lensing galaxy which absorbs the infrared emission from the quasar.¹²⁾

was spent trying to understand how they worked. In the following decade, the emphasis shifted to their use as astronomical tools. We seem to be entering this last phase with gravitational lenses especially with regard to cosmological applications. Although it still has to be admitted that gravitational lenses have told us so far nothing new about the large scale structure and history of the universe in which we have high confidence, they have corroborated, in quite elegant ways, our prior prejudices and may soon allow us to probe distant galaxies, perform basic cosmography and illuminate the dark matter problem.

4.1 Natural Telescopes

Zwicky¹³⁾ was the first to propose that cosmologically distant clusters of galaxies could act as giant telescopes and magnify background galaxies. This has already happened in the giant arcs where we can see distorted images of ordinary $z \sim 1$ galaxies stretched over tens of arcseconds. This allows us to measure the recessional velocities of the stars and, in this manner, study the dynamics of these galaxies as has been attempted in the case of the "straight arc" in A 2390.¹⁴⁾

A less direct technique for super-resolving source structure is to exploit the phenomenon of "microlensing". Galactic gravitational lenses are not smooth, but are made granular by

their constituent stars. For a source that subtends an angular size large compared with the size of the stellar Einstein rings, ($\sim 10^{15} - 10^{16}$ cm), this is unimportant. However, the inner regions of accretion disks orbiting massive black holes in the nuclei of active galaxies, which are supposed to be responsible for the immense ultraviolet emission from quasars are marginally compact enough to be focused as they pass behind individual stars. It is therefore of considerable interest that the four images of the quasar 2237+0305, that are (rather improbably) formed by a $z = 0.04$ barred spiral galaxy, appear to vary independently, with a variation time ~ 1 month compatible with the expected time for the lens to move across the source. It has been argued that the existence of microlensing variation in this source allows us to deduce that this particular source has to be so compact that we can infer that non-thermal emission is responsible. ¹⁵⁾

4.2 Cosmography

A simple way to understand the formation of images by a single lens is to consider the delay t associated with a geodesic connecting the source to the lens plane and connecting to a second geodesic from the observer. Labelling this pair of rays by the angle θ , we can compute a dimensionless form of this delay using the expression

$$\tau \equiv \frac{cD_{ds}t}{(1+z_d)D_dD_s} = \frac{1}{2}(\theta - \beta)^2 - \psi(\theta) \quad (6)$$

where

$$\psi(\theta) = \frac{2D_{ds}}{c^2D_dD_s} \int \phi ds \quad (7)$$

The first term is simply the geometrical delay associated with the longer path. The second term is the gravitational delay associated with the effective refractive index. (In the context of solar system test of general relativity, this delay is known as the ‘‘Shapiro effect’’.) ψ , is effectively the two dimensional Newtonian potential and satisfies the Poisson equation

$$\nabla^2 \psi = \frac{2\Sigma}{\Sigma_{crit}} \quad (8)$$

Invoking Fermat’s principle, the actual rays that connect the observer with the source are those that extremise τ and from this we trivially recover the lens Eq. (3). However we can also compute the magnification tensor that relates the image shape to the original source shape using the Gaussian curvature of the time delay surface at the image location

$$[\mu] \equiv \frac{\partial \theta}{\partial \beta} = [\nabla \nabla \tau]^{-1} \quad (9)$$

This formalism allows us to understand other properties of lenses. ¹⁶⁾ In particular, the time difference for a pair of images is given applying Eq. (8) to the actual rays. To order of magnitude, this is $\sim \theta^2/H_o \sim 1$ yr, for a galactic lens. Note that, given a good enough

model of the lens, in principle obtainable from the disposition of the galaxies and the image properties, it is possible to measure the Hubble constant, a quantity that is known to no better than fifty percent using traditional astronomical methods. A delay of 1.5 yr has been convincingly measured in one case, the quasar 0957+561. ¹⁷⁾ Unfortunately, the lensing configuration is almost certainly too complex to allow us to infer H_o in this case. There are probably two clusters at different redshifts involved. More promising are the radio rings, where the lensing often appears to be caused by an isolated galaxy. It will undoubtedly be necessary to obtain several concordant determinations of H_o by this method before they are accepted by the general astronomical community.

We can use an alternative physical approach to gravitational lensing to elucidate the formation of cluster arcs. Consider a ray congruence about some fiducial ray propagating backward in time from the observer through the lens to the source. In the paraxial approximation, we can propagate basis vectors e_1, e_2 , perpendicular to the fiducial ray and measure the proper and angular displacement of another ray in the congruence, at the observer, by $\xi(\lambda), \theta(\lambda)$, respectively where $\lambda = \int dt/(1+z)$ is a distance measure known as the affine parameter. ξ, θ are linearly related by a generalised angular diameter distance \mathcal{D} , a complex, two component vector defined by

$$\xi_1 + i\xi_2 = \mathcal{D}_1^* \theta_1 + \mathcal{D}_2 \theta_2^* \quad (10)$$

Now, in the weak field approximation of general relativity, $\mathcal{D}(\lambda)$ propagates according to

$$\dot{\mathcal{D}} = \begin{pmatrix} \mathcal{R} & \mathcal{F}^* \\ \mathcal{F} & \mathcal{R} \end{pmatrix} \mathcal{D} \quad (11)$$

where $\mathcal{R} = -4\pi(1+z)^2 G\rho$ describes isotropic ‘‘Ricci’’ focusing by matter of proper density ρ within the congruence, and $\mathcal{F} = -(1+z)^2(\phi_{11} - \phi_{22} + 2i\phi_{12})$ relates to the anisotropic, ‘‘Weyl’’ focusing by tidal forces associated with matter outside the beam. Eq. (12) is a linear equation which maps $\xi \rightarrow \theta$. This mapping becomes singular at a conjugate point where the fiducial ray is tangent to the caustic surface formed by the rays. In other words, the caustic surface is the envelope formed by all rays propagating backward from the observer.

Now, giant arcs are formed whenever a source galaxy is located close to a cusp line on a caustic surface. As the shape of this caustic surface is slightly sensitive to the world model, it is possible, at least in principle, to measure q_o by measuring the redshifts and locations of these arcs provided that one is able to make an accurate model of the cluster potential. Recently, it has been possible to set limits on extreme values of the cosmological constant using the redshift distribution of multiple-imaged quasars and to rule out some even more radical cosmographic models. ¹⁸⁾

4.3 Probes of Dark Matter

Currently, the greatest interest centres around the use of gravitational lenses as probes of dark matter. Closest to home are the halo dark matter searches. When a star in the Galactic halo (or bulge) passes in front of a more distant star, for example in the Magellanic Clouds, it will magnify its flux by a factor

$$\mu = \frac{1 + \beta^2/2\theta_E^2}{\beta/\theta_E[1 + \beta^2/4\theta_E^2]^{1/2}} \quad (12)$$

where the angular separation of the source varies with time according to $\beta^2 = \beta_{\min}^2 + \dot{\beta}_{\perp}^2 t^2$, with $\dot{\beta}_{\perp}$ its relative proper motion. The EROS, MACHO and OGLE collaborations have all reported candidate microlensing events, the largest with a fitted peak magnification of 6.9, and these will be described here. The expected incidence of halo microlensing depends upon the mass model, but is given to order of magnitude by $\sim (v_{\text{circ}}/c)^2 f_{\star} \sim 10^{-6} f_{\star}$ per star where f_{\star} is the fraction of the halo in stars. The experimental efficiencies are not yet known, but based on conservative estimates, the inferred incidence of low mass stars is surprisingly large, especially in comparison with an extrapolation of the local disk stellar mass function. It seems important to rule out intrinsic variability by identifying good candidate events in real time and obtaining high dispersion spectra of them.*

It is also possible to study distant galaxies using radio rings. So far, it appears that the surface densities derived by inverting the observed images are entirely consistent with those measured by other means in nearby galaxies of matched luminosity, at least within the inner regions. This technique is potentially quite accurate.

A related issue is that of dark matter lenses. Here the evidence is contradictory. A model of a gravitational lens source 1422+231²⁰⁾ ²¹⁾ derived the distribution of mass concentrations and galaxies were subsequently detected at the predicted locations. If this turns out to be common experience, then good limits can be placed upon the incidence of "dark galaxies". Dark matter will either have to be associated with the outer parts of luminous galaxies or more smoothly distributed. In addition, there is a class of large separation double quasars, most notably 1634+267 and 2345+007, that are spectroscopically extremely similar and cannot be lensed by a single galaxy (which in any case is not seen). One possible explanation is that the lensing is caused by a giant mass distribution and that the two quasars are a highly amplified merging pair of images.²²⁾ The test of this that background galaxies

* One explanation that we presumably need not take too seriously is that put forward by Arthur C. Clarke in his 1953 novel *Childhood's End*.¹⁹⁾ In this story, the "accelerations" of a receding, alien space ship distort space and focus the light of distant stars whenever conditions are favourable, "visible proof of relativity – the bending of light in the presence of a colossal gravitational field".

should be distorted tangentially with respect to this mass concentration. In the rich clusters, where a large number of arclets have been identified, it is possible to conclude that the mass distribution is centred upon the luminous galaxies, but is more centrally concentrated. Finally, there have been a searches for weak distortions of faint, background galaxies by the stochastic action of intervening large scale structure in the mass distribution. Simple cosmogonies predict an rms elongation ~ 2 per cent and this is comparable with existing upper limits. ²³⁾

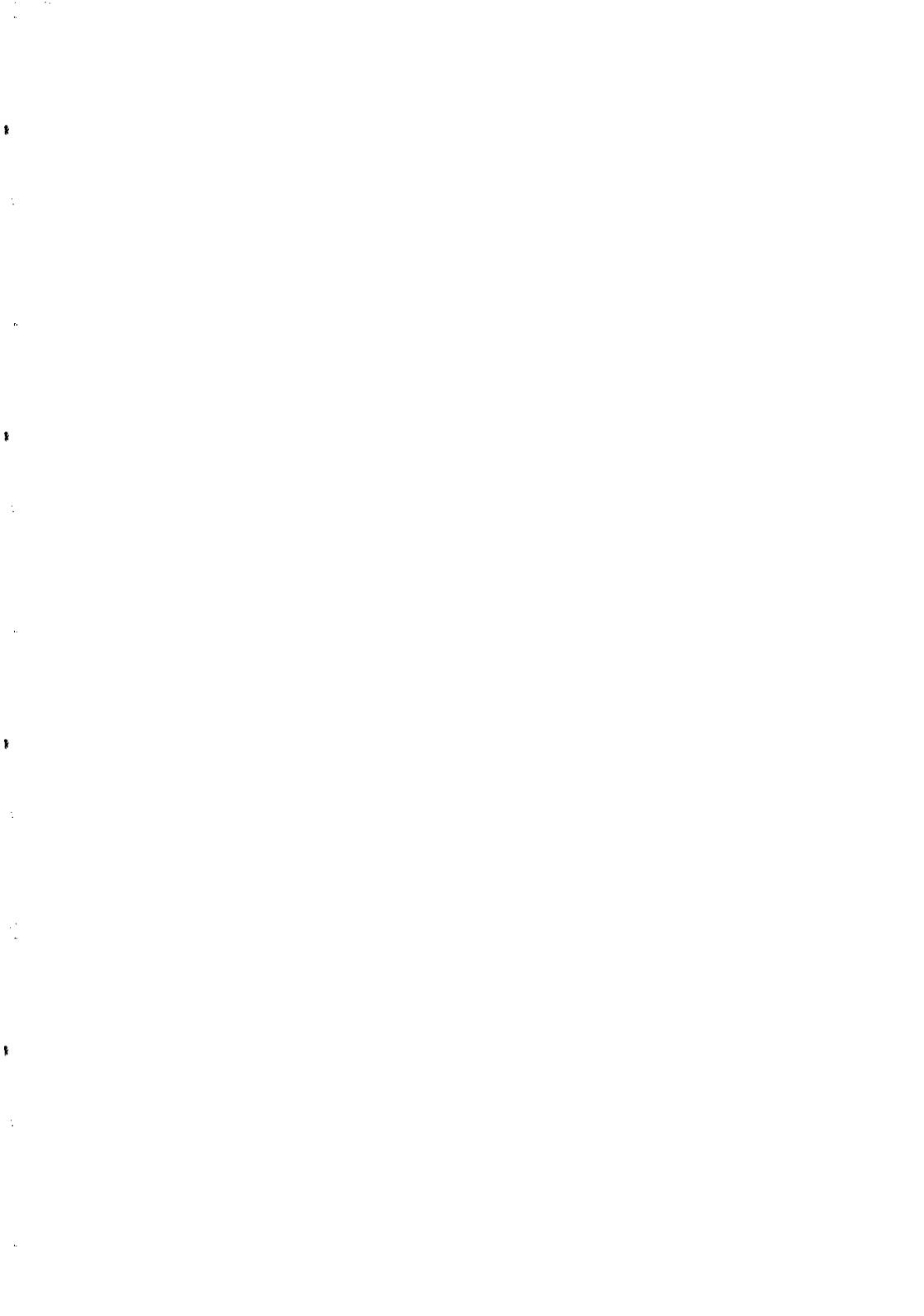
The prospects for gravitational lensing contributing significantly and uniquely to observational cosmology seem good at this time.

Acknowledgements

I acknowledge support under NSF contract 93-23370.

References

1. Einstein, A. 1911 *Ann. Physik* 35 898
2. Einstein, A. 1915 *Sitz. Preuss. Akad. Wiss.* 1 831
3. Dyson, F. W. 1919 *The Observatory* 42 389
4. Robertson, D. S., Carter, W. E. & Dillinger, W. H. 1991 *Nature* 349 768
5. Lodge, O. 1919 *Nature* 104 354
6. Walsh, D., Carswell, R. F. & Weymann, R. J. 1979 *Nature* 279 381
7. Eddington, A. S. 1919 *The Observatory* 42 119
8. Lawrence, C. R. *et al.* 1984 *Science* 223 46
9. Connor, S. J., Lehar, J. & Burke, B. F. 1992 *Ap. J.* 387 L61
10. Hewitt, J. N. *et al.* 1988 *Nature* 333 537
11. Kochanek, C. S., Blandford, R. D., Lawrence, C. R. & Narayan, R. 1989 *MNRAS* 238 43
12. Larkin, J. E. *et al.* 1994 *Ap. J.* 420 L9
13. Zwicky, F. 1937 *Phys. Rev.* 51 290
14. Pello, R. *et al.* 1991 *Ap. J.* 366 405
15. Rauch, K. P. & Blandford, R. D. 1991 *Ap. J.* 381 L39
16. Schneider, P., Ehlers, J. & Falco, E. E. 1982 *Gravitational Lensing*. Berlin: Springer-Verlag
17. Press, W. H., Rybicki, G. B. & Hewitt, J. N. 1992 *Ap. J.* 385 416
19. Clarke, A. C. 1953 *Childhood's End* San Diego: Harcourt-Brace
18. Blandford, R. D. & Narayan, R. D. 1992 *ARAA* 30 311
20. Patnaik, A. R., Browne, I. W. A., Walsh, D., Chaffee, F. H. & Foltz, C. B. 1992 *MNRAS* 259 1p
21. Hogg, D. & Blandford, R. D. 1994 *MNRAS* in press
22. Bonnet, H., Fort, B., Kneib, J. P., Soucail, G. & Mellier, Y. 1994 *Astron. Astrophys.* in press
23. Mould, J. R., Blandford, R. D., Villumsen J., Brainerd, T. G., Smail, I., Small, T. & Kells, W. 1994 *MNRAS* in press



A Search for Low-Mass Objects in the Galactic Halo

François Queinnec

CEA, DAPNIA/Service de Physique des Particules
Centre d'Études de Saclay, F-91191 Gif-sur-Yvette CEDEX

for the

EROS Collaboration



Abstract

We present preliminary results from a search for gravitational microlensing of stars in the Large Magellanic Cloud by low mass objects in the Galactic Halo. The search uses the light curves of over 100,000 stars sampled every 10 minutes for about 300 nights. No light curve exhibits a form that is consistent with a microlensing event of maximum amplification greater than 20 percent on time scales in the range $0.05 < \tau < 10$ days.

Introduction

The presence of large quantities of “dark matter” in the outer regions of spiral galaxies like our own has been inferred from the flat rotation curves obtained from 21 cm line observations [1]. Because they can be easily incorporated into structure formation schemes based on the physics of the early universe, a variety of new weakly interacting elementary particles have been proposed to make up the dark matter. On the other hand, ordinary primordial hydrogen and helium is also a viable candidate if it is in a form that is not easily detected via emission or absorption of photons [2]. Possible forms include compact objects too light to burn hydrogen to ${}^4\text{He}$ ($M < 0.08M_{\odot}$) [2] and cold fractal clouds of helium and molecular hydrogen [3].

We report here results from a search for unseen objects in the Galactic Halo being performed by our collaboration “EROS” (Expérience de Recherche d’Objets Sombres) at the European Southern Observatory at La Silla, Chile [4]. We use the gravitational microlensing effect [5] which would lead to a temporary brightening of stars outside our galaxy as the unseen object passes near the line of sight. The amplification A is given by $A = (u^2 + 2)/[u(u^2 + 4)^{1/2}]$ where u is the undeflected “impact parameter” of the light ray with respect to the unseen object, in units of the “Einstein radius”. For the “standard” isothermal halo model, the probability that a given star in the Large Magellanic Cloud (LMC) is amplified by more than 0.3 magnitudes at any given time is calculated to be about 0.5×10^{-6} [5, 6]. For a deflector of mass M the typical time scale for the amplification is $\tau = 70 \text{ days} \sqrt{M/M_{\odot}}$. The light curve of such an event should be symmetric in time, achromatic, and the event should not be repeated.

The two EROS observing programmes

EROS consists of two complementary programmes. The first uses $5^{\circ} \times 5^{\circ}$ Schmidt plates of the LMC that allow us to monitor about eight million stars with a sampling rate of no more than two measurements per night. This makes this programme primarily sensitive to deflector masses in the range $10^{-4}M_{\odot} < M < 1 M_{\odot}$, corresponding to mean lensing durations in the range $1 \text{ day} < \tau < 100 \text{ days}$. After analysis of 40 percent of the photographic data from this first programme, we have reported two light curves [7] that are consistent with the microlensing hypothesis. Another event has been reported by the MACHO collaboration [8]. The OGLE collaboration [9] has reported the possible observation of the microlensing of a star in the galactic bulge.

EROS’s second programme uses a CCD camera to monitor over 100,000 stars in one field of the LMC bar. The sampling time is 10 minutes making the programme sensitive to deflector masses in the range $10^{-7}M_{\odot} < M < 10^{-3}M_{\odot}$, corresponding to event durations in the range $1 \text{ hour} < \tau < 5 \text{ days}$. We present here the first results from this second programme.

The CCD camera consists of sixteen buttable 576×405 pixels Thomson THX 31157 CCDs covering about 1° by 0.4° . It is mounted on a 40 cm reflector (f/10) constructed by us with the help of the Observatoire de Haute Provence. We have used this setup to observe

one field in the LMC bar from December 1991 to March 1992 and from August 1992 to March 1993. As of March 1993, a total of 8100 exposures have been taken with red and blue filters. (Over 6000 additional exposures are expected from the ongoing 1993-94 observations.) About 100,000 stars are seen on the images with a mean photometric precision of about 6%.

Data reduction

After deflating of the images, one reference image for each color was constructed by combining 50 CCD images taken with good atmospheric conditions. The reference image was subjected to a star finding algorithm in order to establish a star catalog. Next, each image is geometrically aligned with the reference using bright isolated stars. The positions of the stars previously found on the reference image then serve as input to a photometric fitting program to determine the luminosity of each catalog star on the new image. The image is then aligned "photometrically" with the reference image by requiring that the mean luminosity of stars in a given luminosity band be equal to the mean luminosity in the catalog. (The small number of intrinsically variable stars does not affect this procedure.) Successive images then add one point to the blue or red light curve of each star in the catalog. Photometric errors associated with each point on the curve are estimated empirically from the point-to-point variations on a given curve and from the overall image quality.

Search for microlensing light curves

After data reduction, each light curve is subjected to a series of tests designed to isolate microlensing-like events. In this analysis, the 1991-92 and 1992-93 data samples are considered separately. The efficiency of the cuts to accept real microlensing events is estimated with Monte Carlo generated lensing events, superimposed on a random sample of the experimental light curves from both samples.

We start by searching for sequences of 4 points on the red and blue curves that exhibit significant fluctuations from the reference magnitudes defined as the most probable magnitude on the light curve. For the first and second most significant variation in each color, we calculate the mean amplitude, time and χ^2 . Figure 1 shows a scatter plot of the difference Δt between the times of the most significant variations in the red and in the blue, vs. their amplitude, averaged over the red and blue bands. At small amplitude, the red and blue times are uncorrelated indicating that the variations are statistical fluctuations. At amplitudes greater than 1.5, an accumulation at time differences near zero indicates true variations. We retain those stars for which $|\Delta t| < 15$ days; this cut does not affect our detection efficiency.

Having isolated a sample of light curves with at least one significant variation coincident in the red and blue, we proceed to isolate curves with one and only one such variation. We reject stars which show either a second significant variation or activity outside the largest luminosity variation. As an illustration, Figure 2 shows the ratio of the χ^2 's of the second and first most significant variations for the red vs. the same ratio for the blue. Fig. 2a shows the Monte Carlo generated events. The accumulation of events in the lower left is due to events with one and only one significant variation, as expected for microlensing of stable stars. Fig. 2b shows the data. Most of the points are concentrated in the upper right

part of the diagram, indicating that most correspond to stars with more than one significant variation, as expected for certain classes of variable stars. At this stage, a total rejection factor larger than 1000 has been obtained and slightly less than 100 stars are retained.

Their light curves are then examined in detail and fitted for the theoretical light curve. We then reject stars for which the χ^2 per data point in the region of the largest light increase is very large (larger than 10), and those for which the maximum increase, averaged over red and blue, is smaller than 20 % (*i.e.* $u > 1.3$). At this level, most of the selected stars are among the brightest or the reddest in the colour-magnitude diagram. This makes it likely that these stars are intrinsically variable. (Apparent luminosity variations due to microlensing would equally affect all stars, bright or dim.)

Finally, we ask that the times of maximum luminosity be compatible in the red and blue light curves, $|t_{0R} - t_{0B}|/\sqrt{\tau_R + \tau_B} < 0.1$. The effect of this last cut is illustrated in Figure 3. We are left with only two stars passing all the cuts listed above. These two stars are very bright and red. One is selected from the 1991-92 data sample, the other one from the 1992-93 data sample; they are shown in Figure 4. The second one is obviously a variable star that survived all our cuts; this shows how loose these cuts are, especially regarding the activity outside of the main luminosity variation. The first one is compatible with a microlensing light curve, although there is a hint of a deviation before day 760. This same star has been observed to vary in the 1992-93 data sample (its light curve is *not* compatible with microlensing in that period) : it is thus eliminated. We are left with *no candidates*.

We note in addition that these last two stars would correspond to time scales of about 40 days. This is not short enough compared to the duration of an observing season to test safely for the pre- or post-“lensing” stability of the stars. In the microlensing hypothesis, these time scales correspond typically to lensing by objects of $0.3M_{\odot}$, for which we expect less than 0.1 event.

From the absence of microlensing candidates in our sample, we are able to put limits on the presence of low-mass objects in the Galactic Halo. Precise limits are presently under evaluation. From our Monte-Carlo model of the Halo, we expect up to 10 microlensing events in our two data samples if the Halo is fully comprised of these low-mass objects. (This takes into account the effect of LMC star size on the microlensing light curves, which is important for the lowest masses, $10^{-7}M_{\odot}$, and negligible above $10^{-5}M_{\odot}$.) When this analysis is finished, we expect to be able to exclude the presence of a significant amount of objects lighter than $10^{-4}M_{\odot}$ in the Galactic Halo.

As a final remark, we note that a small number of stars that are rejected because their main luminosity variation is smaller than 20 % show light curves that look like microlensing when restricted to the vicinity of the maximum. They are relatively achromatic. They are all very bright or very red, *i.e.* belong to unpopulated regions of the colour-magnitude diagram. All these stars either show activity outside the main variation, or correspond to time scales longer than 10 days. For these reasons, they can be rejected in the analysis and do not at present represent a limiting background in our search for both long and short time scale microlensing phenomena.

Acknowledgements :

This work is based on observations performed at the European Southern Observatory (La Silla, Chile), and is funded by CEA-DSM, CNRS-IN2P3 and CNRS-INSU.

References

- [1] Reviews of dark matter : V. Trimble, *Ann. Rev. Astron. Astrophys.*, 1987; and J.R. Primack et al., *Ann. Rev. Nucl. Sci.*, 1988. A review of the mass estimates of our Galaxy : M. Fich and S. Tremaine, *Ann. Rev. Astron. Astrophys.* 1991, vol.29, 409-45.
- [2] A review of baryonic dark matter : B. J. Carr, *Comm. Astrophys.* **14** (1990) 257-280.
- [3] D. Pfenniger, F. Combes and L. Martinet, to appear in *Astron. and Astroph.*
- [4] E. Aubourg et al., *The Messenger* **72** (June, 1993) 20-27; E. Aubourg, unpublished Ph.D. Thesis (1992), Saclay preprint DAPNIA/SPP 92-22.
- [5] B. Paczyński, *Ap. J.* **304** (1986) 1-5.
- [6] K. Griest et al., *Ap. J.* **372** (1991) L79-82.
- [7] E. Aubourg et al., *Nature* **365** (1993) 623-625; an update on this analysis is given in the contribution of F. Cavalier, these proceedings.
- [8] C. Alcock et al., *Nature* **365** (1993) 621-623.
- [9] A. Udalski et al., *Acta Astronomica* **43** (1993) 283-294.

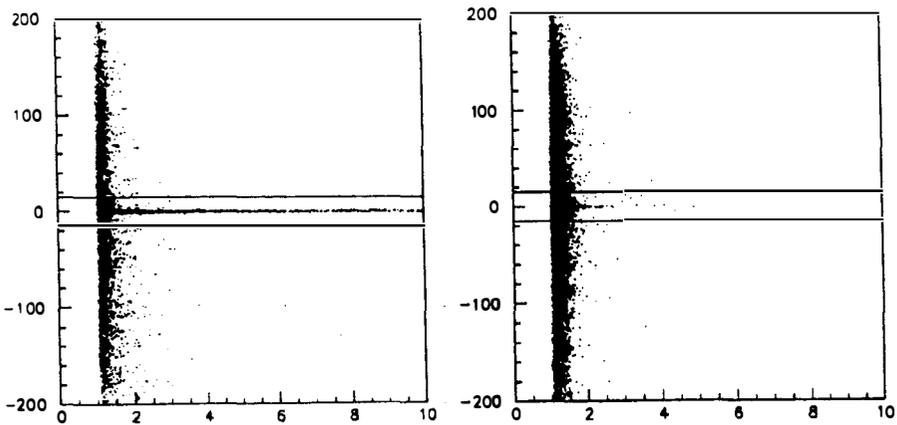


Figure 1. Scatter plot of the difference Δt between the times of the most significant variations in the red and in the blue, vs. their average amplitude. At small amplitude, the red and blue times are uncorrelated indicating that the variations are statistical fluctuations. At amplitudes greater than 1.5, an accumulation at time differences near zero indicates true variations. (Microlensing Monte-Carlo, left; data, right.)

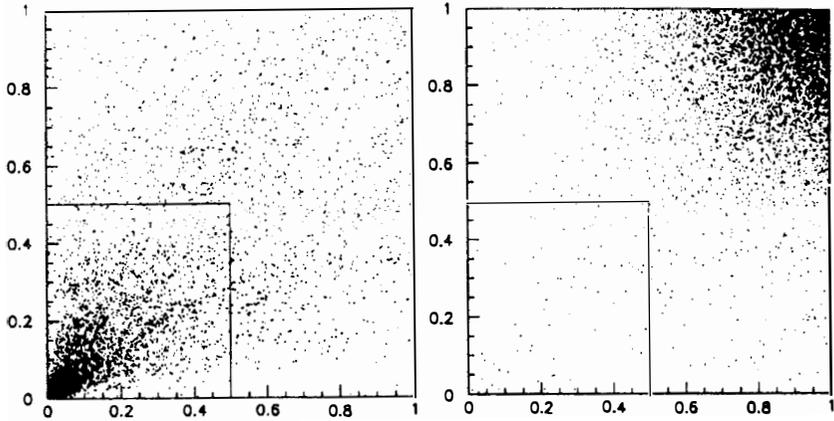


Figure 2. The ratio of the χ^2 's of the second and first most significant variations for the red light curves vs. that for the blue. Fig. 2a shows the Monte Carlo generated events. The accumulation of events in the lower left is due to events with one and only one significant variation, as expected for microlensing of stable stars. Fig. 2b shows the data, for those stars that exhibit at least one significant variation. The accumulation in the upper right corner shows that most stars show more than one such variation.

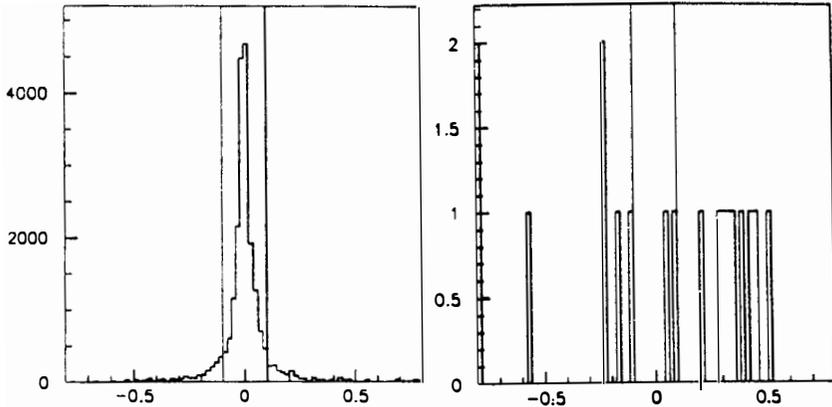


Figure 3. A variable describing the compatibility between the times of maximum light in the red and blue bands, $|t_{0R} - t_{0B}| / \sqrt{\tau_R + \tau_B}$, where t_{0R} and t_{0B} are the fitted times of maximum light and τ_R and τ_B the fitted time scales of luminosity increase in the red and blue. This variable was chosen because it is rather insensitive to the fitted value of τ . Microlensing Monte-Carlo (left) are peaked near zero, contrary to data (right).

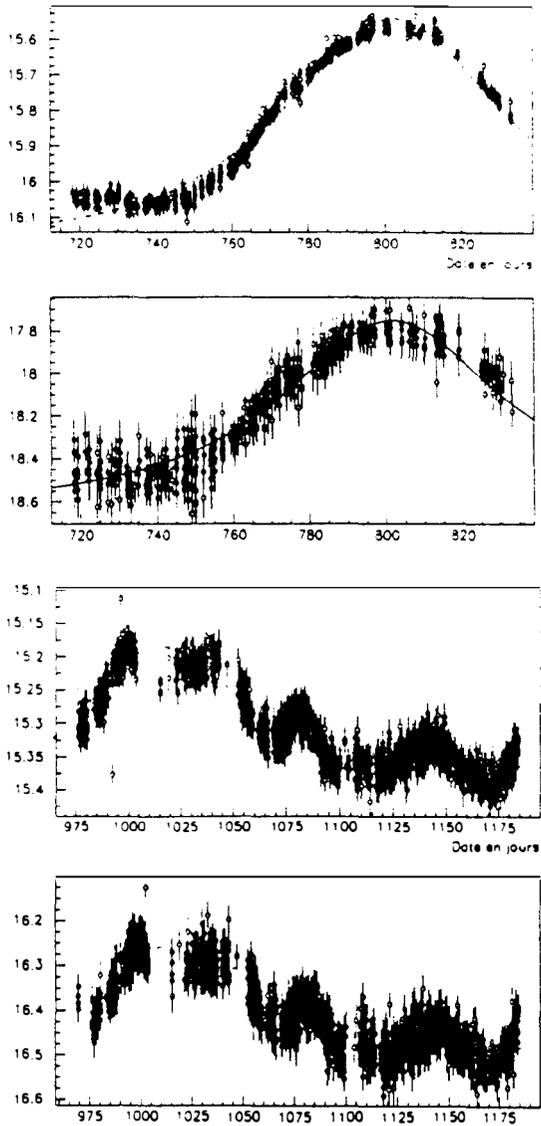
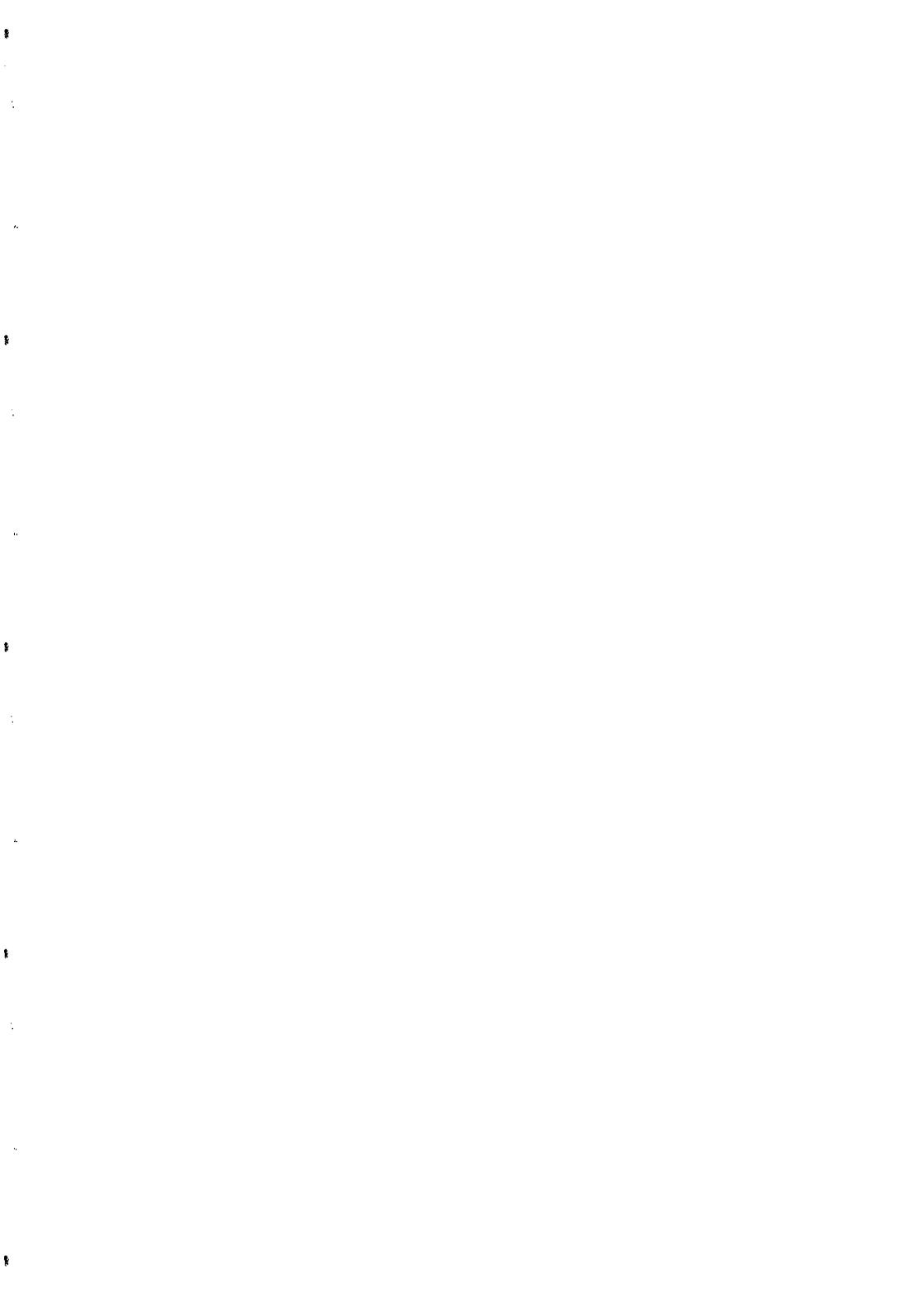


Figure 4. The light curves of the last two selected stars. The first star was observed to vary again in the 1992-93 data sample. The second star shows obvious off-peak activity. Their fitted time scales are respectively 47 and 38 days.



THE OPTICAL GRAVITATIONAL LENSING EXPERIMENT
DISCOVERY OF THE FIRST MICROLENSING EVENTS
IN THE DIRECTION OF THE GALACTIC BULGE¹

A. Udalski, M. Szymański, J. Kałużny, M. Kubiak
Warsaw University Observatory

M. Mateo
Department of Astronomy, University of Michigan

W. Krzemiński, G. Preston
The Observatories of the Carnegie Institution of Washington

B. Paczyński
Princeton University Observatory

Analysis of the data from first two seasons of observations of the Optical Gravitational Lensing Experiment resulted in six microlensing events found (two of these were found after the Moriond Conference). The time scales and magnifications of the events range from 8 to 45 days, and from 1.3 to 11.5, respectively. Gravitational microlensing seems to be the most likely explanation of the observed light variations.

¹Based on observations obtained at the Las Campanas Observatory operated by the Carnegie Institution of Washington

1. Introduction

There is a strong observational evidence that as much as 50–80 percent of the mass of the Galaxy is included in a halo of dark, unseen matter. Theoretical studies (Paczynski 1986, 1991, Griest 1991, Griest *et al.* 1991) led to a conclusion that if this dark matter in the Galactic halo is made of compact objects, such as black holes or brown dwarfs, then the brightness of stars located in the nearby galaxies (e.g. Magellanic Clouds) or in the Galactic Bulge should change because of gravitational microlensing. The probability of microlensing in these directions is very low, $\sim 10^{-6}$ per star, and it is necessary to make many millions of photometric measurements to detect the phenomenon. A huge amount of data should be acquired, reduced, archived and analysed in order to find enough events to draw statistically significant conclusions on the nature and distribution of the dark matter objects.

The Optical Gravitational Lensing Experiment (OGLE) is a long term observational project started in 1992 to search for dark, unseen matter using the microlensing phenomena. The Galactic Bulge has been selected as the first target. After analysing the 1992 and 93 seasons' data we have discovered six events (Udalski *et al.* 1993b, Udalski *et al.* 1994a). Two other teams conduct similar searches in the direction of the LMC: the MACHO and EROS collaborations; both reported the detection of microlensing event candidates (Alcock *et al.* 1993) and the EROS collaboration (Aubourg *et al.* 1993).

2. Observations and Reductions

The OGLE observations are performed with 1-m Swope telescope at Las Campanas Observatory, Chile, which is operated by the Carnegie Institution of Washington. A single chip CCD camera with 2048 by 2048 pixel Ford/Loral detector with the scale of 0.43 arcsec/pixel was used throughout all observations. A total of 150 nights have been allocated in 1992 and 1993 for the OGLE project giving about 2000 frames of the Galactic Bulge. The details of hardware and data pipeline, coordinates of observed fields, error analysis as well as the log of 1992 and 1993 observations can be found in Udalski *et al.* (1992) and Udalski *et al.* (1994b).

The collected frames were reduced almost in real time using the procedure described in detail in Udalski *et al.* (1992). For easy manipulation of the huge amount of data the databases have been built for each field and color (Szymański and Udalski 1993). To select the microlensing event candidates the following procedure has been applied: In 13 fields which were observed in both (1992 and 93) seasons we selected about 1.1 million stars that were constant during 1992. The "constancy" criteria were described in detail in Udalski *et al.* (1993b). Then all observations collected in 1993 were compared with the mean 1992 value and if more than 4 photometric points of a given object deviated by more than 3 sigma from the 1992 mean value, the object was marked as suspected. A variety of filters were then applied to the selected sample of suspected objects leaving a small number of objects which increased their brightness and did not vary randomly. In this way the first microlensing event candidate has been selected (Udalski *et al.* 1993b). Similar procedure was performed on the stars found to be constant during the 1993 season. We checked them for variability in the 1992 data and

found next three microlensing events.

The final photometry of each event was derived by running the DoPhot reduction photometry program (Schechter *et al.* 1993) in the fixed position mode on 150 by 150 pixels subframes centered on the microlensed object. The differential magnitudes were derived with respect to magnitudes of two nearby bright comparison stars which were constant over the entire observing season. We did these additional reductions in order to minimize errors introduced by the variations of the Point Spread Function on the CCD chip. This was normally done by "frame dividing method" (Udalski *et al.* 1992) but in case of microlensing events we wanted to get the photometry as accurate as possible.

3. Microlensing Event Parameters

Table 1 lists the parameters of all microlensing events discovered so far by the OGLE collaboration. For each event we give J2000.0 coordinates of the lensed star as well as its minimal (base) magnitude in *I* band, color index *V* – *I* and the parameters of 4-parameter fit of the theoretical light curve: *A* – maximum magnification, T_{max} – moment of maximum brightness (Heliocentric Julian Date minus 2448000), t_0 – time scale of event and m_0 – minimum (base) *I* magnitude. Figures 1 – 3 show light curves of the discovered events.

Table 1
Microlensing Event Parameters

	OGLE #1	OGLE#2	OGLE#3	OGLE #4	OGLE #5 ¹	OGLE #6 ¹
RA ₂₀₀₀	18 ^h 04 ^m 24 ^s .8	18 ^h 02 ^m 52 ^s .1	18 ^h 04 ^m 43 ^s .4	18 ^h 04 ^m 16 ^s .3	*	*
Dec ₂₀₀₀	–30°05′58″.8	–30°04′21″.5	–30°14′09″.0	–29°51′56″.7	*	*
Mean <i>V</i> – <i>I</i>	1.22	1.56	1.75	1.51	1.57	1.49
Base <i>I</i> mag m_0	18 ^m 82 ± 0.02	19 ^m 05 ± 0.03	15 ^m 89 ± 0.01	19 ^m 27 ± 0.03	17 ^m 88 ± 0.02	18 ^m 07 ± 0.01
Moment of maximum T_0	1153 ^d 9 ± 1.0	803 ^d 4 ± 0.3	831 ^d 5 ± 0.5	807 ^d 3 ± 0.1	824 ^d 33 ± 0.04	818 ^d 90 ± 0.15
Time scale t_0	25 ^d 9 ± 1.7	45 ^d 0 ± 0.9	10 ^d 7 ± 0.7	14 ^d 0 ± 0.8	12 ^d 4 ± 0.4	8 ^d 4 ± 0.3
Magnification <i>A</i>	2.7 ± 0.3	6.5 ± 0.9	1.26 ± 0.02	5.8 ± 1.0	11.5 ± 2.5	6.9 ± 2.5

All events found in the OGLE data fit well the theoretical light curves. Whenever *V*-band observations are available, the light curves are well achromatic (due to limited number of allocated nights, we decided to make observations mostly in *I* band, doing only small number of *V* frames). Also, for at least one observing season the brightness of the lensed stars remained at a constant level suggesting non-repeatability of the events. Still, further observations are required to confirm that fact.

The lensed stars are located in different parts of the color-magnitude diagrams (Udalski *et al.* 1993a). The #3 event star is a red giant, others can be found in different regions of the broad turnover area of the main sequence. It would be difficult to find a physical processes producing identical light curves in so different stars. Secondly, the common shape of the light curves with so different duration

¹ Events #5 and 6 were found after the Moriond Conference

and maximum magnification find the most natural explanation in terms of microlensing. Thirdly, we do not know about any other type of variable stars that would resemble the microlensing variability. Therefore we conclude that we indeed observe microlensing events.

4. Data availability

Photometry of the OGLE microlensing events as well as regularly updated OGLE status report can be found over the Internet network from "sirius.astrouw.edu.pl" host (148.81.8.1) using the "anonymous ftp" service (directory /ogle) or via World Wide Web (WWW): "http://www.astrouw.edu.pl/"

Some of the OGLE data (calibrated FITS images from 1992 season) are accessible for astronomical community from the NASA NSS Data Center.

Acknowledgements. We acknowledge support from the following grants: Polish KBN grants No. 2-1173-91 and BST438/A/93 to M. Kubiak, NFS grants: AST-9023775 and AST-9216494 to B. Paczyński, AST-9216830 to G. W. Preston.

REFERENCES

- Alcock, C. *et al.* 1993, *Nature*, **365**, 621.
 Aubourg, E. *et al.* 1993, *Nature*, **365**, 623.
 Griest, K. 1991, *Astrophys. J.*, **366**, 412.
 Griest, K., *et al.* 1991, *Astrophys. J. Letters*, **372**, L79.
 Schechter, P.L., Mateo, M., and Saha, A. 1993, *P.A.S.P.*, **105**, 1342.
 Paczyński, B. 1986, *Astrophys. J.*, **304**, 1.
 Paczyński, B. 1991, *Astrophys. J. Letters*, **371**, L63.
 Szymański, M., and Udalski, A. 1993, *Acta Astr.*, **43**, 91.
 Udalski, A., Szymański, M., Kałużny, J., Kubiak, M., and Mateo, M. 1992, *Acta Astr.*, **42**, 253.
 Udalski, A., Szymański, M., Kałużny, J., Kubiak, M., and Mateo, M. 1993a, *Acta Astr.*, **43**, 69.
 Udalski, A., Szymański, M., Kałużny, J., Kubiak, M., Krzemiński, W., Mateo, M., Preston, G.W., and Paczyński, B. 1993b, *Acta Astr.*, **43**, 289.
 Udalski, A., Szymański, M., Kałużny, J., Kubiak, M., Krzemiński, W., and Mateo, M. 1994a, *Astrophys. J. Lett.*, in press.
 Udalski, A., Szymański, M., Kałużny, J., Kubiak, M., Krzemiński, W., and Mateo, M. 1994b, *Acta Astr.*, **44**, 1.

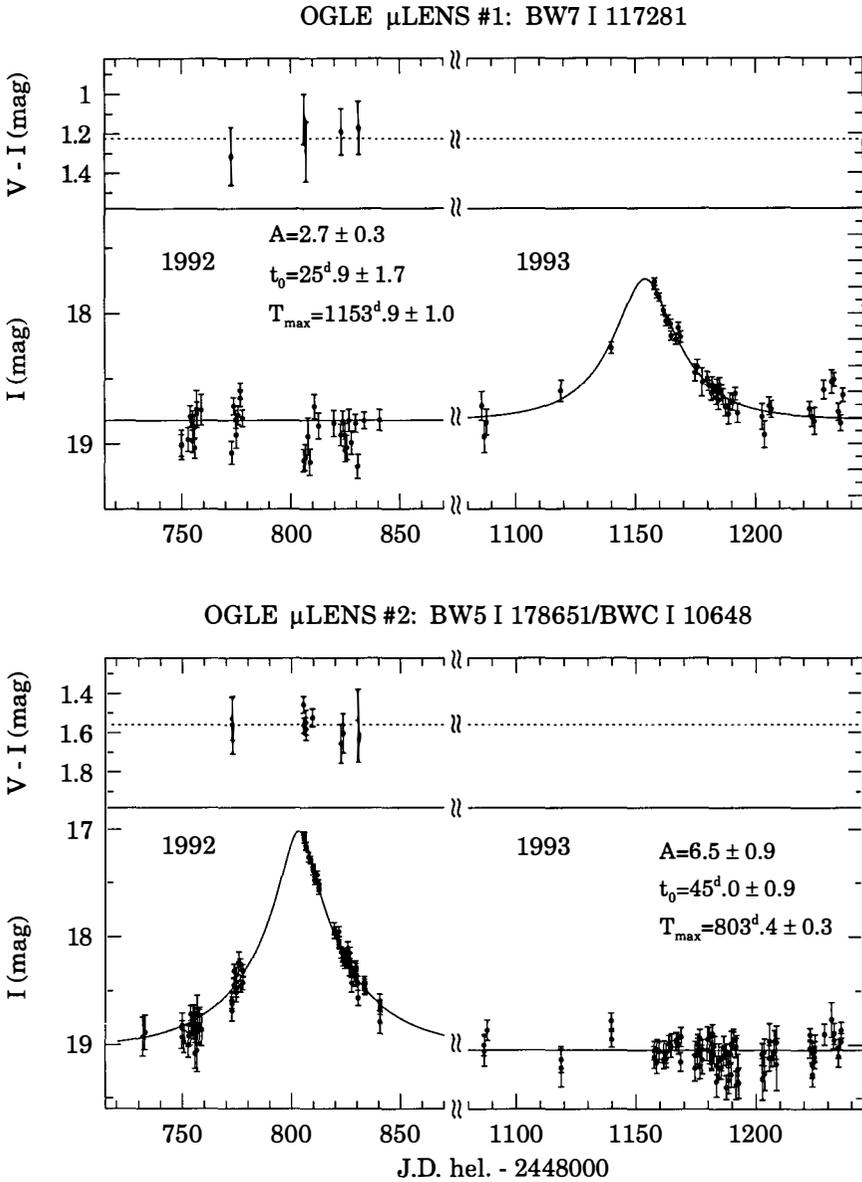


Fig. 1. Light curves of OGLE microlensing events no. 1 and 2.

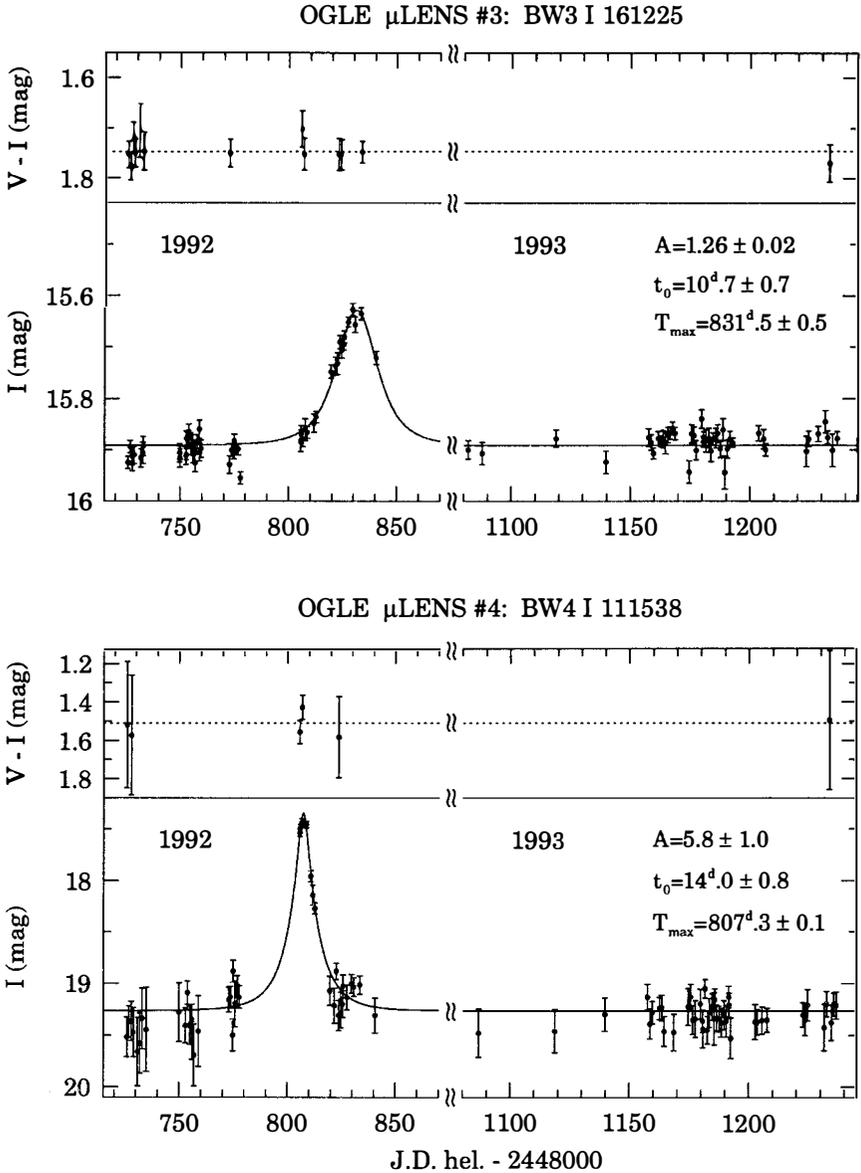


Fig. 2. Light curves of OGLE microlensing events no. 3 and 4.

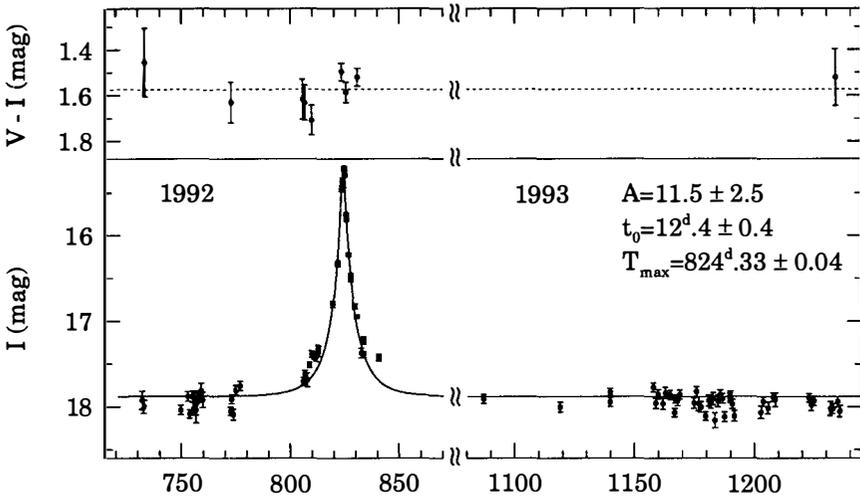
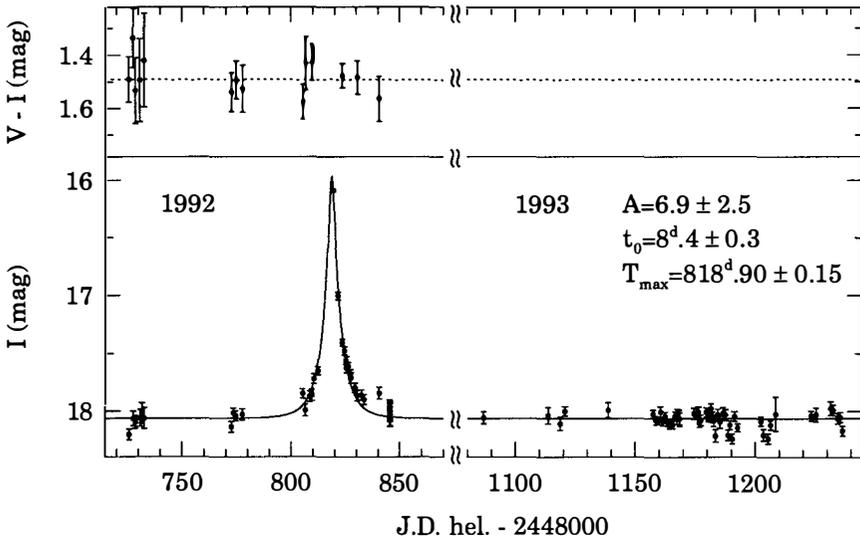
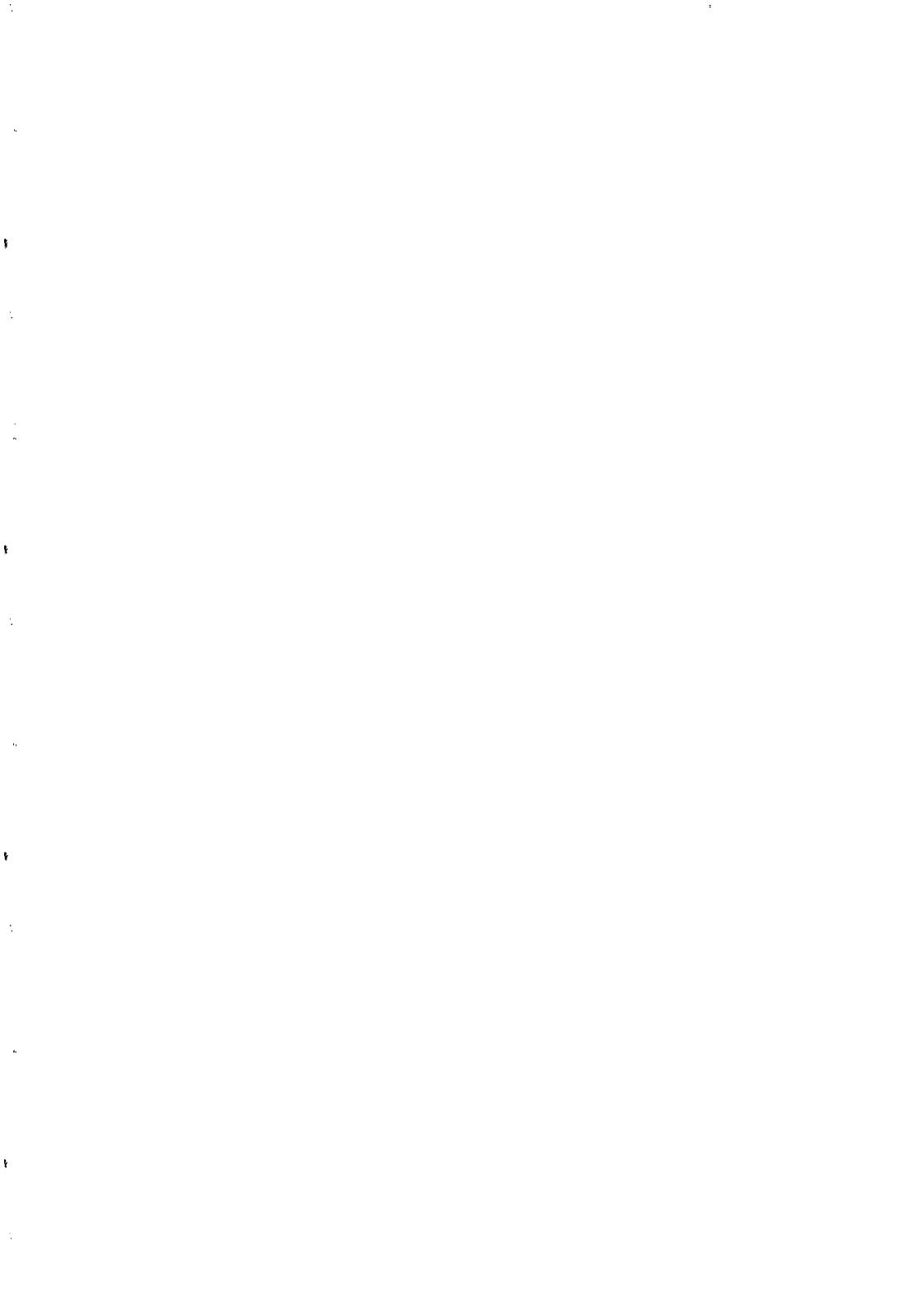
OGLE μ LENS #5: BWC I 120698OGLE μ LENS #6: MM5-B I 128727

Fig. 3. Light curves of OGLE microlensing events no. 5 and 6 (found after the Conference).



LOOKING FOR MICROLENSING OF STARS OF THE ANDROMEDA GALAXY M31 BY MONITORING PIXELS

P. Baillon¹ A. Bouquet² C. Ghesquière³ Y. Giraud-Héraud³
P. Gondolo² J. Kaplan² A.L. Melchior²
¹ CERN Genève, ² LPTHE Universités Paris 6 et 7,
³ LPC Collège de France Paris

Abstract

The galaxy M31 is a very good potential target for microlensing searches. However, only few stars are resolved, and a different method is therefore needed : monitor pixels and not stars. This method and the present status of the study of its feasibility are presented here.

Three experimental teams [2, 3, 4] recently reported the possible detection of microlensings of stars by low mass stellar objects. These observations have been made, as suggested by Paczinsky [1], by looking at well resolved stars in the Large Magellanic Cloud [2, 3] and in the Galactic bulge [4]. We pointed out [5] that it was also possible and even more favourable to look for microlensing phenomena in front of rich fields of unresolved stars such as the M31 galaxy in Andromeda. We describe this project here.

Why M31 ? The Andromeda M31 galaxy is the nearest large galaxy after the Small and Large Magellanic Clouds. It is a giant galaxy, roughly 2 times as large as our Milky Way. M31 probably has its own halo, the brown dwarfs of which, if present, would also give rise to microlensing events. As M31 is tilted with respect to the line of sight, this could give provide to a very interesting signature: assuming an approximately spherical halo for M31, the far side of the disk lies behind a larger amount of M31 dark matter, therefore more microlensing events are expected on the far side of the disk. Such an asymmetry could not be faked by variable stars [6]. The interest of looking for microlensings of stars of M31 has also recently been stressed in reference [7]

In other words, M31 seems very well designed to detect brown dwarfs through microlensing. However, as very few stars of M31 are resolved, we developed an approach to look for microlensing by monitoring the pixels of a CCD, rather than individual stars [5].

Monitoring pixels. In the case of a crowded field such as M31, the light of a pixel comes from the many stars in and around the pixel, plus the sky background. The light of an individual star is spread among all pixels of the seeing spot and only a fraction of this light, the "seeing fraction", reaches any definite pixel. When the star is sufficiently amplified, the brightest pixel may emerge above the background fluctuation. The flux of a pixel is :

$$F_{\text{pixel}} = \left\{ \begin{array}{c} \text{seeing} \\ \text{fraction} \end{array} \right\} F_{\text{star}} + \langle F_{\text{other stars}} \rangle + F_{\text{background}}, \quad (1)$$

where the flux F_{star} of the star to be amplified has been singled out.

If the star is amplified by a factor A , the increase of the pixel flux is

$$\Delta F_{\text{pixel}} = (A - 1) \left\{ \begin{array}{c} \text{seeing} \\ \text{fraction} \end{array} \right\} F_{\text{star}} \quad (2)$$

An event will be called detected if the flux increase is larger than the natural fluctuation of the pixel σ_{pixel} by some factor Q , for at least a few consecutive exposures :

$$\Delta F_{\text{pixel}} > Q \sigma_{\text{pixel}} \quad (3)$$

For example, in the simulation we use to evaluate the number of events we will be able to see, we require that Q be larger than three during 3 consecutive exposures and equal to 5 for at least one of these three exposures.

Pixel fluctuations. The number of expected detectable events depends crucially on the value of the relative fluctuation on the pixel (the root mean square of the fluctuation divided by the flux : σ/F). For example, we ran our Monte-Carlo simulations supposing that we were aiming at M31 with a 1 m telescope, opened at F/4 and equipped with the EROS camera. The result is that we expect one detectable event every 3 days if the relative fluctuation is of the order of 1%, but only one every 10 days if it rises to 2%.

To test the feasibility of the experiment, we have analyzed three series of real data :

- 1) 82 images of the Large Magellanic Cloud (LMC) taken by the EROS collaboration
- 2) 26 images of M31 taken with the one meter telescope at Pic du Midi, in collaboration with F. Colas (Bureau des longitudes, Paris) and J. Lecacheux (DESPA, Meudon)
- 3) 4 images of M31 taken by E. Davoust (OMP, Toulouse) with the 2 m telescope at Pic du Midi. In this case as the angular size of the pixels is very small, we have also rearranged them in 4×4 “superpixels”

The results are given in Table 1 and in Figure 1 where maps of the relative fluctuations are shown in various configurations

Images	mirror size (meter)	pixel size (arcsec)	relative fluctuation
LMC EROS (• 1)	0.4	1.15	3%
M31 Pic (• 2)	1	0.7	1%
M31 Pic (• 3)	2	0.25	0.7%
M31 Pic (• 3)	2	1 (“superpixels”)	0.23%

Table 1: The mean relative fluctuation obtained for the three series of images listed above. The numbers in the first column refer to the numbers in the list.

These numbers clearly show that the required photometric stability of pixels can be reached. Moreover the comparison between the two lower maps of figure 1 shows what we gain by using small pixels regrouped in larger “superpixels” : Pixels small compared to the seeing, allow an efficient matching between different images, whereas, once images are matched, superpixels are more appropriate for a stable photometry.

Discriminating against variable stars. Variable stars will be the main background. The usual tools to discriminate against this background are available (symmetry, unicity, achromaticity of the light curve). Still some points, particular to our approach, are discussed below.

Achromaticity. At first sight, one would think that there is no achromaticity as a star rising above a background of a different color will cause a color variation of the concerned pixels. However, it is easy to show that when a star rises above the background in two color bands (say red and blue) then the ratio

$$\frac{(F_{\text{pixel}} - \langle F_{\text{pixel}} \rangle)_{\text{red}}}{(F_{\text{pixel}} - \langle F_{\text{pixel}} \rangle)_{\text{blue}}} = \frac{F_{\text{star}}|_{\text{red}}}{F_{\text{star}}|_{\text{blue}}} \quad (4)$$

is constant in time.

High amplifications. To rise above the background an unresolved star needs a rather high amplification ($\langle A \rangle \sim 6$), which will exclude most variable stars.

On the other hand small secondary maxima indicating unstable stars with occasional strong flares will not be discriminated. Further study is required in this respect, which will be performed using the EROS variable stars catalog.

Seeing variations. Seeing variations may turn out to be a major problem, as they induce variations of the light flux of pixels near bright stars. This point is under study.

Statistics. To evaluate the number of events we can expect, we have run Monte-Carlo simulations in various configurations. Results are given in Table 2. For the sake of comparison all evaluations are based on the same observation period of 120 consecutive nights. The Brown Dwarf mass is taken to be $0.1 M_{\odot}$, and we assume a usual halo (see reference [5] for instance) with a local dark matter density of $.3 \text{ GeV/cm}^3$ and a core radius of 5 kpc.

Target galaxy	Mirror diameter (m)	field size (' \times ')	pixel size (")	seeing (")	Photometric precision	number of detections
M31	2	60 \times 20 (24fields)	1	1	.23%	110
	1	60 \times 20 (EROS camera)	1.15	1	1%	50
LMC	0.4	60 \times 20 (EROS camera)	1.15	1.5	1%	4

Table 2:

These numbers have to be compared with the total number of events expected by the EROS collaboration during three years of observation, which is of order 5 for brown dwarfs with masses around $0.1 M_{\odot}$.

We thank The EROS collaboration, and E. Davoust who allowed us to use their data, as well as F. Colas, and J. Lecacheux with whom we took data at Pic du Midi.

References

- [1] Paczynsky B., 1986, Apj 304, 1
- [2] Aubourg e. et al., 1993, Nature 365, 623
- [3] Alcock et al., 1993, Nature 365, 621; and these proceedings.
- [4] Udalski A. et al., 1993, Acta astronomica 43, 289.
- [5] Baillon P., Bouquet A., Giraud-Héraud Y., Kaplan J., 1993, A&A 277, 1
- [6] Crotts A. P. S., 1992, Apj 399, L43
- [7] Gould. A., Microlensing by stars in the disk of M31, Ohio State University preprint Osu-TA-5/94

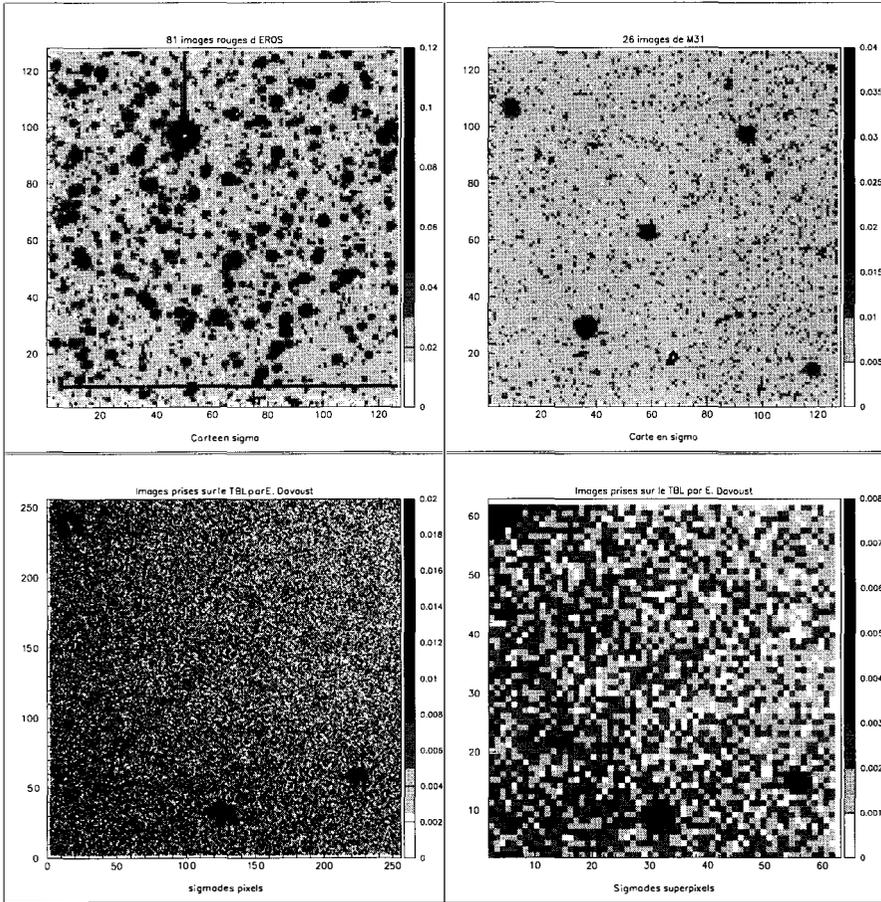
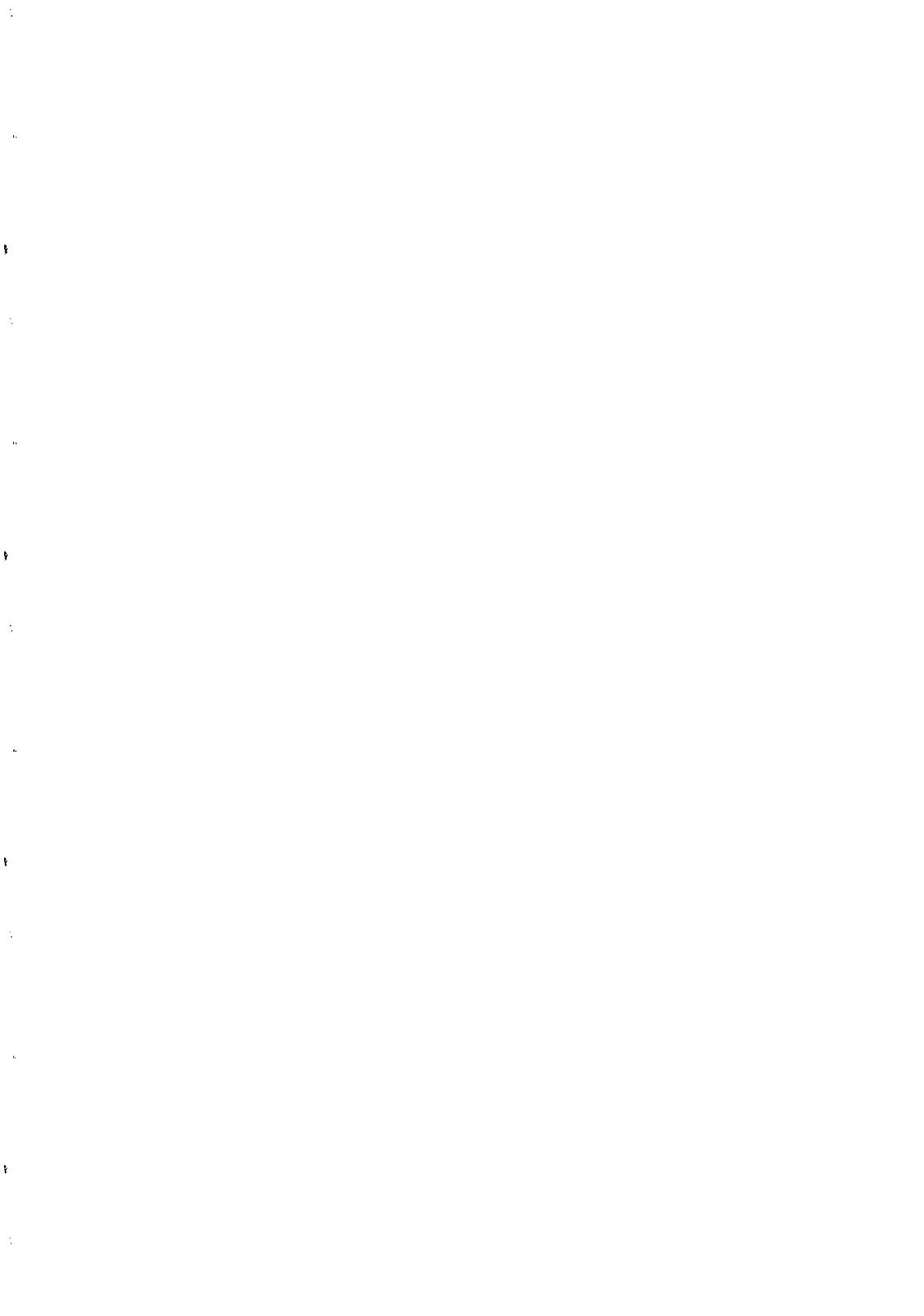
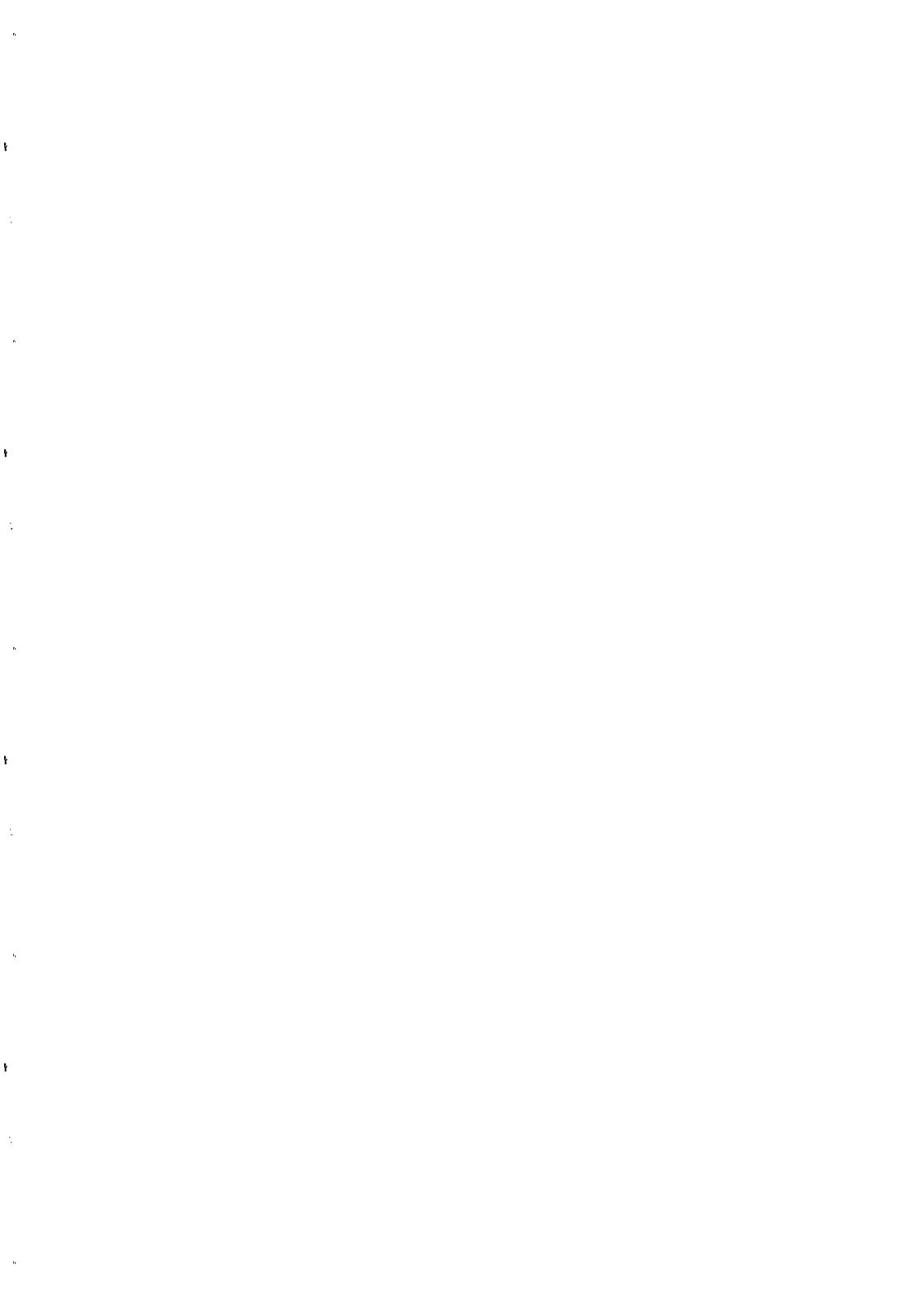


Figure 1: Map of the relative root mean square of the flux fluctuations on various images. Upper-left: 82 images of the LMC taken by the Eros collaboration, 128 by 128 pixels of 1.1 arcsec. Upper-right: 26 Images of M31 taken on the 1 m telescope of Pic du Midi, 128 by 128 pixels of .7 arcsec. Lower-left: 4 images taken on the 2 m telescope at Pic du Midi, 256 by 256 pixels of .23 arc sec. Lower- right, same as lower-left but the pixels have been lumped in 64 by 64 "super-pixels" of .94 arcsec. The scale is indicated on each picture.



NEUTRINOS



KARMEN : NEUTRINO SPECTROSCOPY AT ISIS

Jonny Kleinfeller

*Kernforschungszentrum Karlsruhe, Institut für Kernphysik,
76021 Karlsruhe, Germany*

KARMEN Collaboration*

ABSTRACT

The Karlsruhe-Rutherford Medium Energy Neutrino Experiment KARMEN at the pulsed spallation neutron facility ISIS makes use of ν_μ , ν_e and $\bar{\nu}_\mu$ from the $\pi^+ \rightarrow \mu^+ \rightarrow e^+$ decay chain at rest. Neutrinos are detected in a 56 t high resolution liquid scintillation calorimeter. Consisting entirely of hydrocarbons the calorimeter serves as a massive live target of ^{12}C and ^1H nuclei for the investigation of various ν -induced reactions. The exclusive charged current reaction $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$, which is observed with negligible background and spectroscopic quality, allows to measure for the first time the spectral shape of ν_e emitted in μ^+ -decay. The experimental flux averaged cross sections is compared with theoretical predictions. The measured energy dependence of this cross section is used to deduce the q^2 -dependence of the axial vector form factor, the first direct measurement of a weak nuclear form factor as yet.

*B. Armbruster, G. Drexlin, V. Eberhard, C. Eichner, K. Eitel, H. Gemmeke, W. Grandegger, T. Jannakos, M. Kleifges, P. Plischke, J. Rapp, J. Weber, J. Wochele, J. Wolf, S. Wölfe, B. Zeitnitz: *Kernforschungszentrum Karlsruhe and Universität Karlsruhe*. D. Blaser, B. Bodmann, M. Ferstl, E. Finckh, T. Hanika, M. Hehle, J. Hößl, W. Kretschmer, F. Schilling, H. Schmidt, O. Stumm: *Universität Erlangen-Nürnberg*. R. Maschuw: *Universität Bonn*. I.M. Blair, J.A. Edgington, B. Seligmann: *QMWC London*. N.E. Booth: *Oxford University*. : *Rutherford Appleton Laboratory*.

1. Introduction

Neutrino-induced reactions in nuclei at low and intermediate energies are important in the study of fundamental neutrino properties and ν -interactions with matter, especially in astrophysics. The dynamic processes of core-collapse supernovae where ν -nucleus interactions are believed to contribute to nucleosynthesis¹ is one of the most impressive examples. However, as in solar neutrino physics, the almost total lack of experimental cross section data introduces great uncertainties. Furthermore, ν -nucleus interactions in this energy range are an ideal tool to study the structure of weak interactions. Transitions between *discrete* nuclear states with well defined quantum numbers $J^\pi T$, which can be identified by subsequent characteristic nuclear de-excitation processes, select specific parts of the weak hadronic currents.² In ν -induced transitions of the $^{12}\text{C}_{\text{g.s.}}$ ($0^+ 0$) state, either to $^{12}\text{N}_{\text{g.s.}}$ ($1^+ 1$) through weak charged currents or to $^{12}\text{C}^*$ ($1^+ 1$) through weak neutral currents, the spin flip ($\Delta S = 1$) selects the axial vector amplitudes of the weak hadronic currents, the change of isospin ($\Delta T = 1, \Delta T_3 = 0, 1$) the isovector components. Therefore, this dominant isovector axial vector amplitude can be measured exclusively.

This report describes in detail only results in connection to charged current reactions.

2. The KARMEN Experiment

The ISIS facility, driven by an 800 MeV, 50 Hz proton synchrotron of 200 μA dc-intensity, is the world's most powerful pulsed low energy neutrino source. The successive decays $\pi^+ \rightarrow \mu^+ + \nu_\mu$ and $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$, both π^+ and μ^+ decaying at rest inside the Ta-D₂O beam dump target, are the ν -generating processes of monoenergetic ν_μ (30 MeV) and $\nu_e, \bar{\nu}_\mu$ of energies up to 52.8 MeV. The unique time structure of ISIS (together with the different life times of π^+ (26 ns) and μ^+ (2.2 μs)) defines a duty cycle factor of 10^{-4} , providing high neutrino peak intensity, effective suppression of cosmic background and clear separation of ν_μ -induced reactions from those induced by ν_e or $\bar{\nu}_\mu$.² Neutrinos are detected in a high resolution 56 t liquid scintillation calorimeter³ located at a distance of 17.5 m from the ν -source. A bookshelves structure of totally reflecting double layers of acrylic glass divides the central detector into 512 independent modules with an energy resolution of $\sigma(E)/E = 11.5\% / \sqrt{E(\text{MeV})}$ and a time resolution of 1 ns. The central detector consists almost entirely of hydrocarbons, serving as a massive target of ^{12}C and ^1H nuclei for ν -nuclear interactions and providing excellent calorimetric characteristics for low energy particles at the same time.

3. The Charged Current Reaction $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$

At neutrino energies available at ISIS, charged current reactions in nuclei can only be induced by electron-neutrinos. The signature of the exclusive charged current reaction $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$ is a spatially-correlated delayed coincidence of an electron from the inverse β -decay on ^{12}C within the ν_e time window of 0.5-9.5 μs after beam-on-target, and a positron from the subsequent β -decay of $^{12}\text{N}_{\text{g.s.}}$ back to $^{12}\text{C}_{\text{g.s.}}$ materializing within the following 32 ms. The kinetic energy of the prompt electron is restricted to $E_e < 35.5\text{ MeV}$ due to the negative Q -value of 17.3 MeV in the inverse β -decay. The ^{12}N -decay, characterized by its lifetime of $\tau = 15.9\text{ ms}$ and its end point energy of $E_0 = 16.3\text{ MeV}$, uniquely identifies ν -induced transitions to the ground state of ^{12}N . Only events fully contained within the fiducial volume, with no hit in the veto system and no previous activity in the central detector in a pretrigger interval of 20 μs , were taken into account. After software cuts⁴ on time, energy, position correlation and background subtraction 223 ± 15.3 events remain. The energy distributions (fig. 3) of the electrons and positrons are well matched with simulations of the detector response. The time distributions relative to

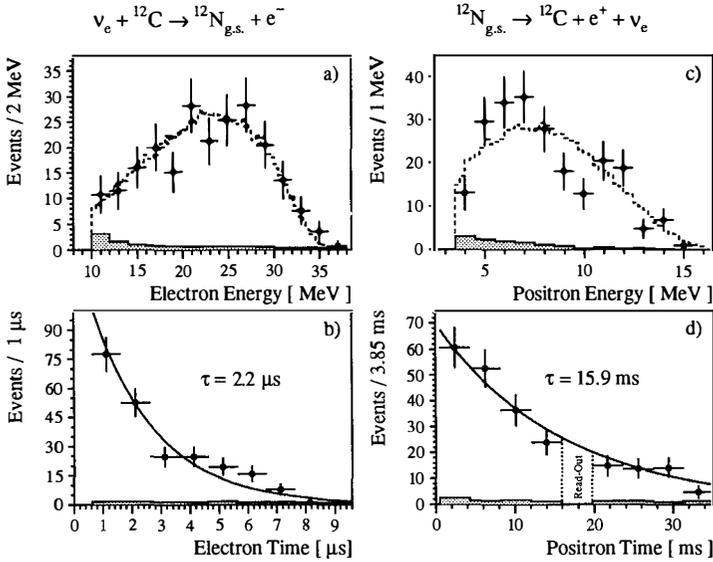


Fig. 1. Energy and time distributions of the $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$ reaction (broken lines GEANT 3 simulations), time distributions with decay curves of μ^+ ($\tau = 2.2 \mu\text{s}$) and ^{12}N ($\tau = 15.9 \text{ms}$) superimposed. Normalized background is shown as shaded area.

beam-on-target reflect the decay time constants of μ^+ -decay and β -decay of $^{12}\text{N}_{\text{g.s.}}$. In particular, the time distribution of the prompt signal (fig. 3(c)), which is fitted by a time constant of $\tau_{\text{exp}} = (2.21 \pm 0.18) \mu\text{s}$, clearly indicates that these events are due to neutrinos from μ^+ -decay. Identical cuts on background data samples of high statistics, taken simultaneously in the time intervals before beam-on-target, result in exponential energy and flat time distributions, demonstrating the random nature of the background coincidences. Background contamination (11.8 events) is almost negligible due to the unambiguous signature. The integral signal to background ratio is 20:1, an advantage a high resolution calorimeter can only gain at a pulsed neutrino source. The cross section for the charged current reaction averaged over the incident ν_e energy distribution (0-52.8 MeV) is:

$$(\sigma_{\text{CC}}(\nu_e))^{\text{exp}} = [8.9 \pm 0.60(\text{stat.}) \pm 0.75(\text{syst.})] \times 10^{-42} \text{cm}^2$$

in good agreement with recent theoretical predictions⁵⁻⁸ of $(8.0-9.4) \times 10^{-42} \text{cm}^2$, confirming that cross sections for ν -induced transitions between *specific* nuclear states can be calculated reliably.

4. Energy Dependence of the Charged Current Cross Section

The excellent calorimetric characteristics of the KARMEN detector made it possible to ascertain the energy dependence of the $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$ cross section for the first time. The initial ν_e energy distribution can be reconstructed from the kinetic energy of the electrons E_{e^-} by $E_\nu = E_{e^-} + 17.3 \text{MeV}$, where the recoil energy transferred to the ^{12}N nucleus is assumed to be

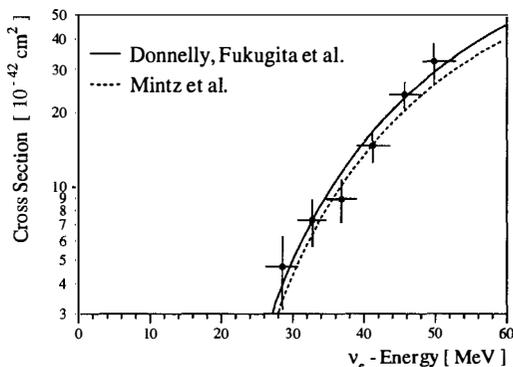


Fig. 2. Energy dependence of the $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$ reaction cross section.

negligible. Thus a precise measurement of the kinetic energy of the electrons determines the initial neutrino energy, allowing the possibility of neutrino spectroscopy. The mean cross sections over six neutrino energy intervals obtained by regularized unfolding⁹ is shown in fig. 4 in comparison with theoretical calculations⁵⁻⁷ of the energy dependence. No further normalization has been applied. The theoretical curves represent the threshold energy dependence $(E_\nu - Q)^2$ modulated by the q^2 -dependence of the axial vector form factor $F_A(q^2)$.

5. q^2 -Dependence of the Axial Vector Form Factor

The energy dependence $\sigma_{CC}(E_\nu)$ can be used to obtain the q^2 -dependence of the weak isovector axial vector form factor $F_A(q^2)$. Experimental information about this important weak form factor has so far been available only at fixed four-momentum transfer q^2 from weak processes such as β -decay of ^{12}N , ^{12}B ($q^2=0$) or from μ^- -capture reactions ($q^2=0.56 m_\pi^2$). To separate the threshold factor $(E_\nu - Q)^2$ from the assumed q^2 -dependence of the axial vector form factor $F_A(q^2)$, an expression for $\sigma_{CC}(E_\nu)$ has been used, where the energy dependence of the $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$ cross section is expressed only in terms of kinematic factors and $F_A(q^2)$, neglecting contributions from $F_M(q^2)$ and $F_T(q^2)$. Assuming $F_A(q^2)$ has the same q^2 -dependence as the weak magnetism form factor $F_M(q^2)$ leads to the quasi-dipole form^{6,7}

$$F_A(q^2) = g(q^2) \times [1 + q^2 / M_A^2]^{-2}$$

The function $g(q^2) = 1 + O(q^4)$ incorporates small corrections in the second and higher power of q^2 . The cross section $\sigma_{CC}(E_\nu, M_A)$ is then folded into the ν_e energy spectrum and the detector response function and M_A determined from a maximum likelihood fit to the electron energy spectrum. The result, using the form of $g(q^2)$ given by Mintz and Pourkaviani,⁷ is:

$$M_A = 180_{-50}^{+240} \text{ MeV} / c^2$$

to be compared with a value of 236 MeV / c^2 derived from inelastic electron scattering.⁷ In fig. 5 the KARMEN result for $F_A(q^2)$ is compared with the form adopted by Mintz and Pourkaviani⁷ and by Fukugita.⁶ This is the first *direct* measurement of a weak nuclear form factor confirming the assumed identical q^2 -dependence of vector and axial vector form factors.

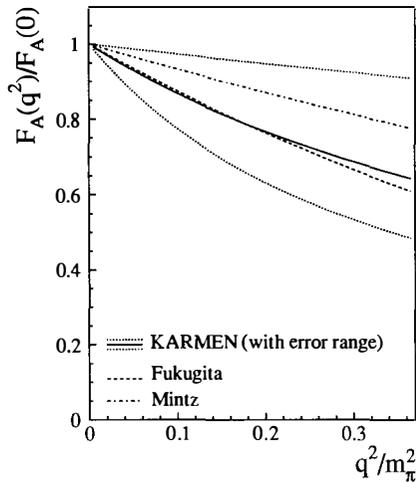


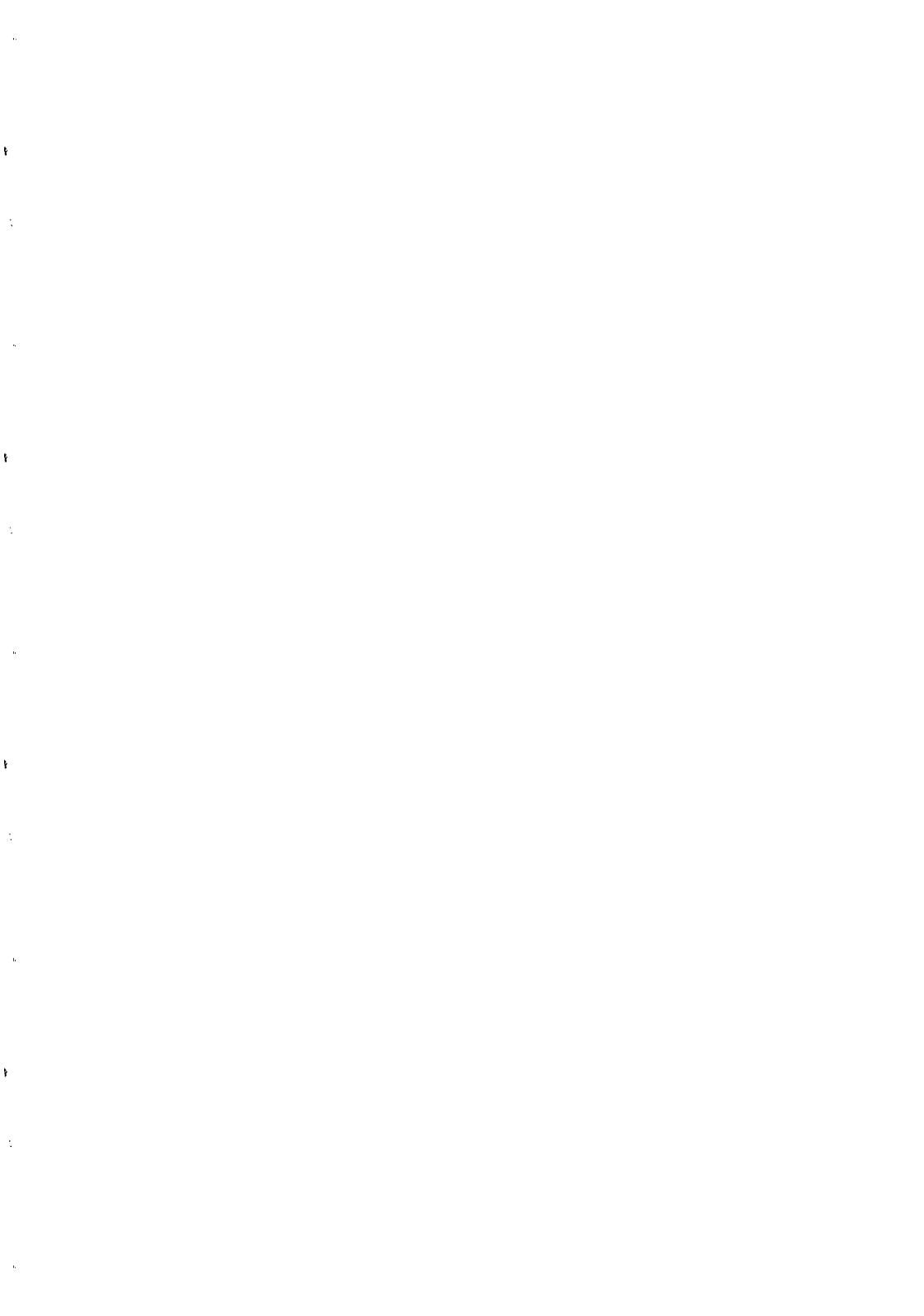
Fig. 3. q^2 -dependence of the weak *isovector axial vector* form factor $F_A(q^2)$ of ^{12}C obtained from the energy dependence of the $^{12}\text{C}(\nu_e, e^-)^{12}\text{N}_{\text{g.s.}}$ reaction cross section (dipole fit with $M_A = 180 \text{ MeV}/c^2$) compared to dipole fits of the electromagnetic *vector* form factor $\mu(q^2)$ of ^{12}C obtained from inelastic electron scattering.

6. Other Results and Outlook

Other main topics of the KARMEN research programme include ν -induced nuclear transition via the weak hadronic neutral current.¹⁰ From the ratio of the neutral current and charged current cross sections a first test of universality of couplings in the weak neutral current at low energies has been derived. A detailed analysis of the ν_e spectrum originating from μ^+ decay, measured directly for the first time, sets limits on non Standard Model contributions in muon decay. The search for lepton number violating pion and muon decay modes tightens the limits on neutrino identity in these decays. A highly sensitive search for appearance type neutrino oscillations between ν_μ and ν_e as well as between the corresponding anti-neutrinos provides reliable limits on neutrino mixing of the first two neutrino generations. The KARMEN experiment is scheduled to continue data taking for another three years. The data sample will then consist of almost 3000 recorded neutrino-nucleus interactions.

References

1. S.E. Woosley *et al.*, *Astrophys. J.* **356** (1990) 272.
2. B. Zeitnitz, *Prog. in Part. and Nucl. Physics*, Vol. 13 (1985) 445.
3. G. Drexlin *et al.*, *Nucl. Instr. and Meth. A* **289** (1990) 490.
4. B. Bodmann *et al.*, *Phys. Lett. B* **280** (1992) 198.
5. T.W. Donnelly, *Phys. Lett. B* **43** (1973) 93, and priv. communication (1992).
6. M. Fukugita, Y. Kohyama and K. Kubodera, *Phys. Lett. B* **212** (1988) 139.
7. S.L. Mintz and M. Pourkaviani, *Phys. Rev. C* **40** (1989) 2458.
8. E. Kolbe, K. Langanke, S. Krewald, CalTech MAP 161, (1993), *subm. to Phys. Rev. C*.
9. V. Blobel, *Proc. 1984 CERN School of Computing*, CERN 85-09 (1985) p. 88.
10. B. Bodmann *et al.*, *Phys. Lett. B* **267** (1991) 321.



**THE HEIDELBERG-MOSCOW $\beta\beta$ EXPERIMENT WITH ^{76}Ge AND THE
PRESENT STATUS IN DOUBLE BETA DECAY RESEARCH**

M. Beck, J. Bockholt, M. Eskef, D. Glatting, J. Hellmig, G. Heusser, H.V. Klapdor-Kleingrothaus*, B. Maier, F. Petry, A. Piepke, E. Scheer, H. Strecker, M. Völlinger
Max-Planck-Institut für Kernphysik, P.O. Box 103980, 69029 Heidelberg, Germany

A. Balysh, S.T. Belyaev*, A. Demehin, A.Gurov, I. Kondratenko,
D. Kotelnikov, V.I. Lebedev
Russian Science Center-Kurchatov Institute, 123 182 Moscow, Russia

A. Müller
Istituto Nazionale di Fisica Nucleare, I-67010 Assergi, Italy

**Spokesmen of the collaboration*

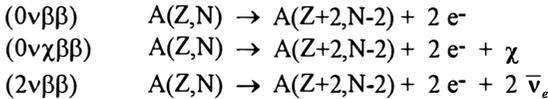
The experimental status of double beta decay is briefly reviewed, with the main emphasis being put on the HEIDELBERG-MOSCOW experiment. In this collaboration at present three enriched (86%) ^{76}Ge detectors with a total active mass of 6.02 kg are operated in the GRAN SASSO laboratory. With a ^{76}Ge exposure of 2924.2 kg·d (91.63 mol·y) we derive the following half life limits for the neutrinoless double beta ($0\nu\beta\beta$) decay: $T_{1/2}(0^+ \rightarrow 0^+) > 1.9 \cdot 10^{24}$ y and $T_{1/2}(0^+ \rightarrow 2^+) > 8.2 \cdot 10^{23}$ y with 90% C.L. By this limit a Majorana mass of the neutrinos larger than 1.1 eV is excluded.

In the first high statistics measurement of the two neutrino double beta ($2\nu\beta\beta$) decay of ^{76}Ge a half life of $T_{1/2} = (1.42 \pm 0.03_{\text{stat}} \pm 0.13_{\text{syst}}) \cdot 10^{21}$ y is extracted. For the neutrinoless double beta decay with majoron emission ($0\nu\chi\beta\beta$) a half life of $1.7 \cdot 10^{22}$ y is excluded, leading to a limit for the neutrino-majoron coupling of $\langle g_{\nu\chi} \rangle < 1.8 \cdot 10^{-4}$ with 90% C.L. First results on a dark matter search improved the existing cross section limits for WIMP masses above ~ 150 GeV and excluded Dirac neutrinos in the range of 26 GeV to 4.7 TeV as the dominant component of the dark halo.

Presented by: B. Maier

1. INTRODUCTION

The main aim of double beta ($\beta\beta$) decay research is to probe physics beyond the standard model by investigating the nature of the neutrino (Majorana or Dirac particle) and the question of a non zero neutrino mass. The constraints derived by $\beta\beta$ decay experiments on the existence of a non zero Majorana neutrino mass and right-handed weak currents (RHC) are the most stringent limits published so far. In $\beta\beta$ decay research usually three decay modes are discussed [1, 2]:



The neutrinoless double beta ($0\nu\beta\beta$) decay and the Majoron-accompanied neutrinoless double beta ($0\nu\chi\beta\beta$) decay are forbidden decay modes in the standard model, since the B-L conservation is violated. With a measured half life limit of the $0\nu\beta\beta$ -decay mode and calculated matrix elements an upper limit for the effective neutrino mass can

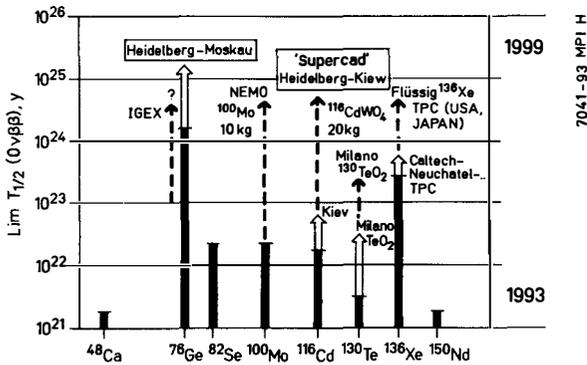


Fig. 1 Present situation and near future of the most promising $\beta\beta$ -experiments. The thick solid bars correspond to the present status in 1993, open bars and dashed lines to 'safe' and 'less safe' expectations for 1999 [2].

be deduced. Since this method relies on theoretically calculated matrix elements (see, e.g. [3]), it is essential to also focus on the two neutrino double beta ($2\nu\beta\beta$) decay, where no unknown particle physics parameters have to be included in the theoretical calculations like in the $0\nu\beta\beta$ -decay. The measurements of $2\nu\beta\beta$ -decay rates can test the predicted half lives in order to enlarge the confidence in the $0\nu\beta\beta$ -matrix elements,

although the calculations show some substantial differences [4]. The fact that out of 35 potential $\beta\beta$ -emitters 15 different isotopes are investigated in more than 50 experiments using 13 different measuring techniques [5] shows the great activity in this field of physics. Experiments working with enriched source materials will set the pace in the near future, especially with the isotopes ^{76}Ge [1,5], ^{136}Xe [6], ^{100}Mo [7] and

^{116}Cd [8]. The present situation and the perspectives of the most promising $\beta\beta$ -experiments are shown in Fig.1, where only isotopes with $0\nu\beta\beta$ half life limits greater than 10^{21} y have been considered.

2. EXPERIMENT

The HEIDELBERG-MOSCOW group is using p-type HP Ge semiconductor detectors, made from isotopically enriched ^{76}Ge [1, 9, 10]. The isotopic abundance of ^{76}Ge is 86% compared to 7.8% in natural Ge. The enrichment was measured for each crystal separately by accelerator mass spectroscopy using residues from the crystal fabrication. This experiment has three main advantages: the excellent energy resolution of the detectors, which favours the search for the discrete spectrum of the $0\nu\beta\beta$ -decay at the Q-value of $Q = 2038.56$ keV, the large size of the detectors, concentrating the background in the peaks, and the fact that the source is equal to the detector allowing large source strengths.

At present three enriched detectors are operated in a common shielding in the Gran Sasso underground laboratory (3500 m.w.e.). This setup of detectors with 6.29 kg total and 6.02 active mass of Ge (taking into account the inactive zones of the diodes) corresponds to a ^{76}Ge source strength of 68.7 mol. Two more detectors with masses of ~ 2.5 and ~ 3.5 kg are in preparation and will be installed soon. In the present setup the three enriched detectors are surrounded by a 30 cm Pb shield, which consists of an inner shielding of 10 cm highly radiopure LC2-grade Pb followed by 20 cm of Boliden Pb. All cryostat parts of the detectors are made of electrolytical Cu. The whole setup is placed in an air-tight steel box and flushed with radiopure nitrogen in order to suppress the Rn-contamination of the air. Since February 1994 the steel box is centered inside a 10 cm bor-polyethylen shielding to decrease the neutron flux from outside. All materials used in the setup are carefully selected and cleaned, while the activation due to cosmic radiation was minimized by storing the materials underground.

Table 1
Parameters of the three enriched operating detectors

Detector	Active mass [kg]	Enrichment [%]	Measuring time [date] [days]		Background B at 2000-2100 keV [counts/keV·y·kg]
Enr. 1	0.920	86	2.8.90 - 3.8.91		0.56
			19.1.92 - 12.8.92	388.6	
Enr. 2	2.758	86	4.10.91 - 12.8.92	223.0	0.29
common shielding of three enriched detectors					
Enr. 1	0.920	86	3.9.92 - 4.10.93	325.9	0.20
Enr. 2	2.758	86	3.9.92 - 4.10.93	325.9	0.22
Enr. 3	2.344	88	3.9.92 - 4.10.93	325.9	0.25

Table 1 shows some of the parameters of the operated detectors and the measured background numbers during the different data acquisition periods. The last period corresponds to the present setup described here. The progress from setup to setup is obvious by the decreasing background numbers. Taking the combined value of B for the present setup with 0.23 [counts/keV.y.kg] we have reached the best background published in literature so far. Accumulating the data of all enriched detectors the ^{76}Ge exposure is 91.634 mol.y or 2924.2 kg.d with a background between 2000 and 2100 keV of 0.29 [counts/keV.y.kg]. The combined spectrum of all detectors is shown in Fig. 2. Due to the already discussed advantages of our detectors (size and energy resolution) it is possible to clearly identify the various background contaminations, like

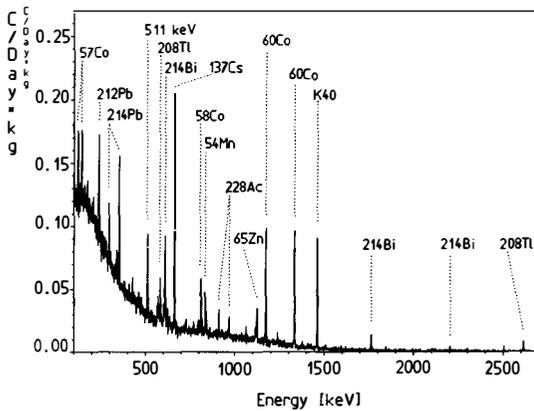


Fig. 2. Combined spectrum of the three enriched detectors with an ^{76}Ge exposure of 91.634 mol.y or 2924.2 kg.d.

the natural decay chains, cosmogenic activities and the activities of ^{137}Cs and ^{40}K . The energy threshold of detector #2 is in the present setup at 13 keV and that of detector #3 at 15 keV. The data taken with these detectors can therefore be used for a dark matter search by looking for scattering of WIMP's at Ge nuclei. The first results of a dark matter evaluation taking the data of detector #2 in the first data acquisition period with a stable threshold of 11 keV are briefly reviewed later.

3. RESULTS

$0\nu\beta\beta$ -decay mode

For the evaluation of the $0\nu\beta\beta$ -decay the data with 91.634 mol.y (2924.2 kg.d) from the accumulated spectrum shown in Fig. 2 are used. The energy resolution in the summed spectrum is 3.7 keV for the ground state transition ($0^+ \rightarrow 0^+$ at $Q=2038.56$ keV), respectively 3.2 keV for the excited state transition ($0^+ \rightarrow 2^+$ at $Q=1479.46$ keV). The FWHM at the decay energies is estimated from a linear fit of the FWHM of the background peaks in the summed spectrum. The region of interest between 2000 and 2080 keV is shown in Fig. 3. The 3σ interval of a hypothetical $0\nu\beta\beta$ -peak for the ground state transition contains 32 events. Using the method recommended by the PDG [11] we can exclude 19.6 (14.5) events with 90 (68) % C.L., leading to a half life limit of $T_{1/2} > 1.9 (2.6) \cdot 10^{24}$ y with 90 (68) % C.L. Using the matrix elements of [3] we

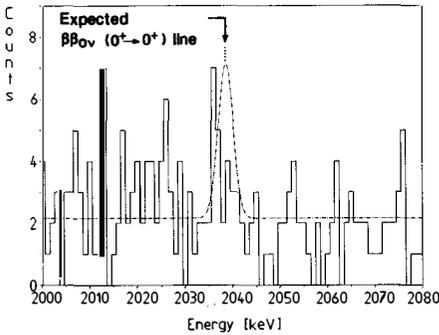


Fig.3. Region of interest for a hypothetical $0\nu\beta\beta$ -peak after 91.634 mol-y. The dotted line represents the signal excluded with 90% C.L.

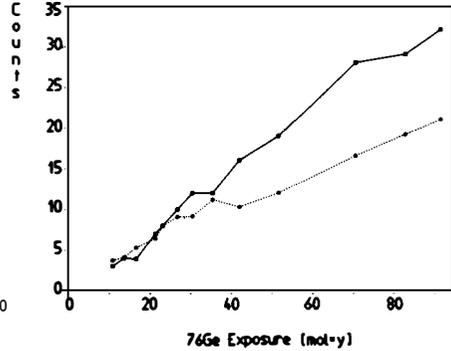


Fig.4. Number of expected (dashed line) and measured counts (solid line) in the 3σ -range of the $0\nu\beta\beta$ -peak as a function of measuring time.

derive neglecting RHC an upper limit for the neutrino mass of $\langle m_\nu \rangle < 1.1$ eV with 90% C.L. Considering RHC we obtain with 90% C.L.: $\langle m_\nu \rangle < 1.3$ eV, $\langle \lambda \rangle < 2.0 \cdot 10^{-6}$ and $\langle \eta \rangle < 1.3 \cdot 10^{-8}$ (effective right-right- and left-right-handed admixture to the weak interaction). These are so far the sharpest limits from detector experiments [1, 2]. Extrapolating the expected background, evaluated in ~ 30 keV ranges left and right from the centroid energy at 2038.56 keV, into the 3σ interval of a hypothetical $0\nu\beta\beta$ -peak we seem to observe a systematic excess of events. In Fig. 4 the expected and measured counts in the 99% peak interval are shown as a function of measuring time. Under the assumption that this excess is background (no-peak hypothesis) the deviation corresponds to $\sim 2.4\sigma$. It has to be mentioned that this effect is very sensitive to single counts in the region of interest, so that in the recent data of the last weeks the deviation decreased to $\sim 2\sigma$. Further measuring time and data evaluation is necessary to observe the evolution of this effect. Especially our developed method of pulse shape analysis [12] will help to increase the efficiency for $\beta\beta$ -events.

For the transition to the first excited state ($0^+ \rightarrow 2^+$) we deduce from 91.634 mol-y of measuring time a half life limit of $T_{1/2} > 8.2$ (13.1) $\cdot 10^{23}$ y with 90 (68) % C.L. This excludes 46.5 (29.1) decays with 90 (68) % C.L. The combined escape probability of 35.5% for the 559.1 keV γ -quanta was calculated with the Monte-Carlo code GEANT3 for each detector separately.

$2\nu\beta\beta$ -decay mode

For the analysis of the $2\nu\beta\beta$ -decay the data from 4.10.91 to 12.8.92 of detector #2 are used, corresponding to a ^{76}Ge exposure of 615 kg-d. For the first time a quantitative background model was developed and applied in such a low-level experiment in order to unfold the various continuous background components [13]. This model is based on

a Monte-Carlo simulation using the CERN-code GEANT3. Only measured quantities are used to normalize the calculated background spectra.

In a first step 27 identified peaks of the background contaminations are removed from the measured data. These measured peak intensities of the identified background isotopes in the raw data are then used to normalize the simulated background spectra. In this way the continuous contributions due to the cosmogenic activities of ^{54}Mn , $^{57,58}\text{Co}$ and ^{65}Zn inside the Ge crystal and the activities of ^{54}Mn , $^{57,58,60}\text{Co}$ in the Cu parts of the cryostat system, identified by their characteristic shifted and unshifted γ -lines, are included in the model. The absolute activity of ^{210}Pb in the lead shielding was measured separately with low-level α -spectroscopy in order to normalize the bremsstrahlung spectrum of its daughter ^{210}Bi in the model. Since several lines of the natural decay chains of ^{232}Th and ^{238}U are identified in the measured spectrum, it was possible using the relative peak intensities to locate these impurities in the Cu parts of the cryostat system and to include their contribution in the model. The small effects of ^{40}K in the lead shielding, of ^{137}Cs in the Cu parts of the cryostat system and a phenomenological constant background due to inelastic neutron scattering complete the background model.

From the difference of measured spectrum and background model we deduce a signal to background ratio of 1.3:1 in the evaluation interval of 500 -1500 keV, containing 72% of the $2\nu\beta\beta$ -intensity. A maximum-likelihood fit yields a half life for the $2\nu\beta\beta$ -decay of ^{76}Ge of $T_{1/2} = (1.42 \pm 0.03 \text{ stat} \pm 0.13 \text{ syst}) \cdot 10^{21} \text{ y}$. The result after subtracting the background model is shown in Fig. 5. The result presented here is probably the first undoubtable evidence for this nuclear decay mode. A comparison with various other results for $2\nu\beta\beta$ -decay measurements is given in [2, 5, 13].

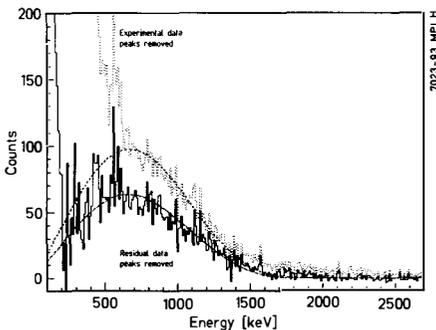


Fig.5. The "stripped" data (dotted histogram) and the residual data after background unfolding (solid histogram). The continuous curves are calculated $2\nu\beta\beta$ -spectra with half lifes of $T_{1/2} = 0.92 \cdot 10^{21} \text{ y}$ [14] and $T_{1/2} = 1.42 \cdot 10^{21} \text{ y}$.

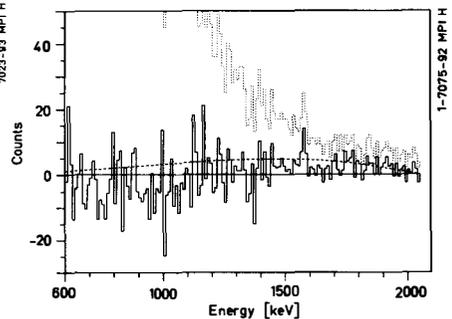


Fig.6. Residual spectrum after subtraction of background model and $2\nu\beta\beta$ -spectrum (solid histogram), the measured raw spectrum (dotted histogram) and a calculated $0\nu\chi\beta\beta$ -spectrum with $T_{1/2} = 1.66 \cdot 10^{22} \text{ y}$ (dashed curve) are shown.

$0\nu\chi\beta\beta$ -decay mode

For the analysis of the $0\nu\chi\beta\beta$ -decay the same set of data as for the $2\nu\beta\beta$ -decay described before has been used. Subtracting the background model and the $2\nu\beta\beta$ -spectrum from the data the $0\nu\chi\beta\beta$ -spectrum should remain (see Fig. 6). In the energy range from 1.1 to 2.05 MeV, containing 74.5 % of the $0\nu\chi\beta\beta$ -spectrum, 208 events are remaining. The analysis of the spectral shape shows a disagreement with the theoretically calculated shape [10]. The deduced half life limit for the $0\nu\chi\beta\beta$ -decay is $T_{1/2} > 1.7 (1.9) \cdot 10^{22}$ y with 90 (68) % C.L. Using matrix elements of [3] we derive an upper limit of $\langle g_{V\chi} \rangle < 1.8 \cdot 10^{-4}$ with 90 % C.L. for the neutrino-majoron coupling constant [10].

Dark Matter

For the first time enriched material was used to search for dark matter [15]. In the first data acquisition period of detector #2 in April 1992 a threshold of 11 keV was achieved. The measuring time of the used data is 165.6 kg-d. The background in the

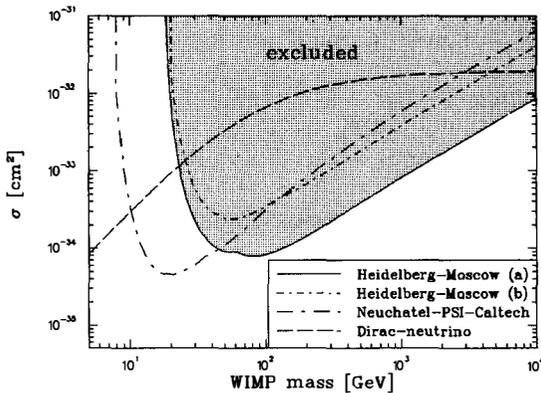


Fig. 7. Dark matter exclusion plot. The Heidelberg-Moscow result a) (shaded area) is not corrected for loss of coherence, while b) is corrected for loss of coherence.

region between 11 keV and 30 keV is (0.106 ± 0.006) [counts/keV·y·kg], which is by a factor of 5 - 7 lower compared with other dedicated dark matter experiments of this type [16]. The parameters for the calculated recoil spectra of the WIMPs (weakly interacting massive particles) are given in [15]. For the evaluation of the data several criteria have been used. The resulting exclusion plot of cross sections and masses for the scattering of spin-independently interacting WIMPs off Germanium is shown in

Fig. 7. The result of the Heidelberg-Moscow experiment is shown 'not corrected for loss of coherence' (shaded area) and 'corrected for loss of coherence' together with the theoretically predicted weak cross section of heavy Dirac neutrinos. The exclusion limit of the Gotthard experiment (geometrically estimated from [16]) is referenced for comparison. The existing cross section limits for WIMP masses above 150 GeV were improved and Dirac neutrinos could be excluded as the dominant component of the dark halo in the mass range between 26 GeV and 4.7 TeV.

4. CONCLUSION

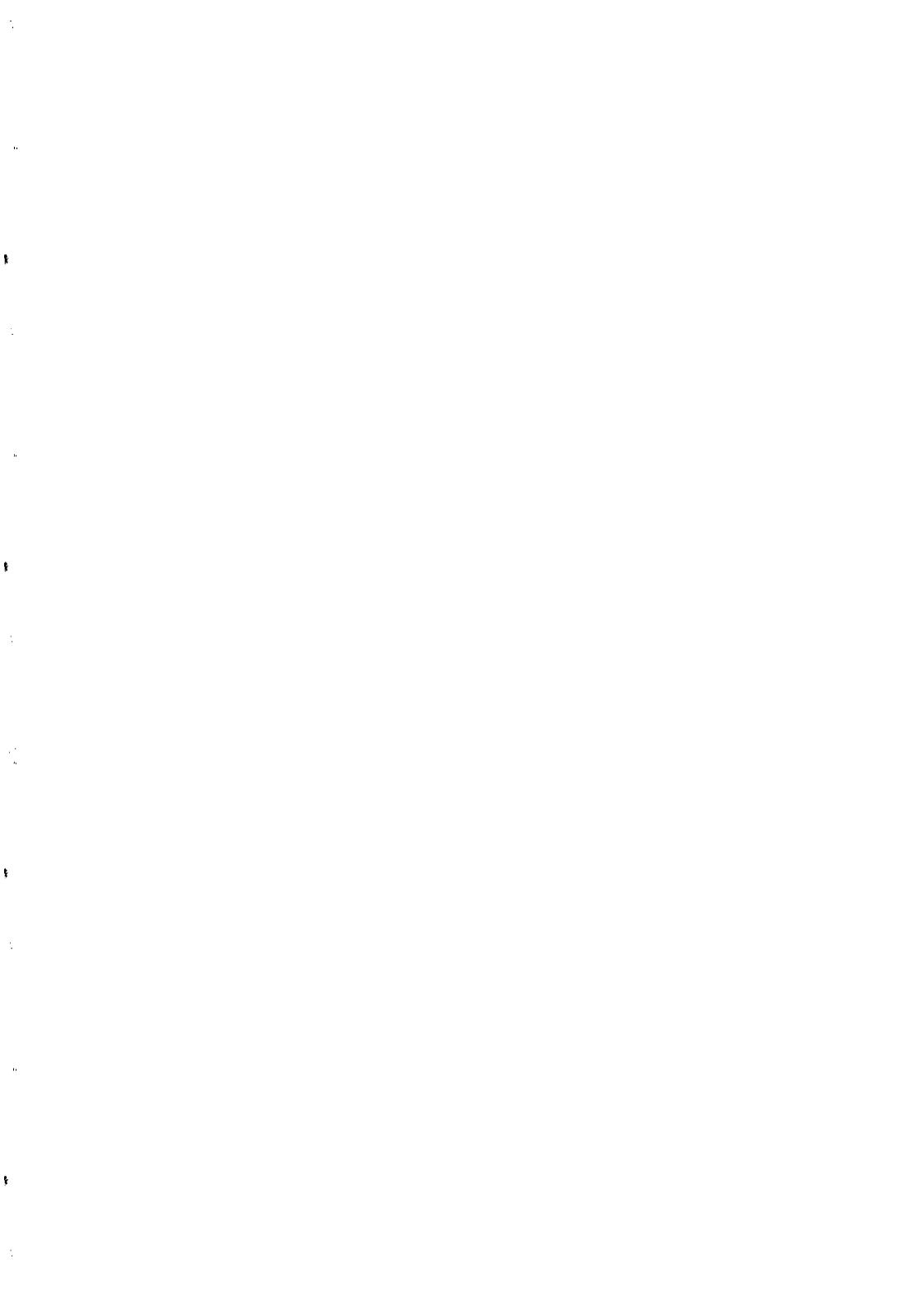
By end of 1994 the experiment will be in full scale operation after the installation of two more crystals in the Gran Sasso laboratory with ~ 130 mol or ~ 12 kg of ${}^{76}\text{Ge}$ -detectors. This will allow to probe Majorana neutrino masses down to ~ 0.2 eV within 5 years of measurement.

We would like to thank Professors Bellotti, Cabibbo, Maiani and Monacelli for their continuous generous support of the experiment. We also thank the German Ministry of Research and Technology (BMFT) and the Russian Atomic Energy Committee for their substantial support.

REFERENCES

- [1] H. V. Klapdor-Kleingrothaus, Proc. III. Int. Symp. on WEAK and Electromagnetic Interactions in Nuclei, WEIN'92, Dubna, 1992 p. 201, World Scientific, Singapore.
- [2] H. V. Klapdor-Kleingrothaus, in: Progress in Particle and Nuclear Physics Vol. 32, 1994, Pergamon Press, Oxford.
- [3] A. Staudt, K. Muto and H.V. Klapdor-Kleingrothaus, Europhys. Lett. **13** (1990) 31.
- [4] K. Muto, E. Bender, H.V. Klapdor, Z. Phys. A **334** (1989) 177, 187.
- [5] M.K. Moe, Preprint UCI-Neutrino 93-1 (1993).
- [6] H.T. Wong, J.Phys. G **17** (1991) Suppl.
- [7] H.Ejiri *et al.*, Phys. Lett. B **258** (1991) 17.
- [8] Yu. Zdesenko, in: Progress in Particle and Nuclear Physics Vol. 32, Erice, 1994 to be published, Pergamon Press, Oxford.
- [9] A. Balysh *et al.* (HEIDELBERG-MOSCOW collaboration), Phys. Lett. B **283** (1992) 32.
- [10] M. Beck *et al.* (HEIDELBERG-MOSCOW collaboration), Phys. Rev. Lett. **70** (1993) 2853.
- [11] Particle Data Group, J.J. Hernandez *et al.*, Review of particle properties, Phys. Lett. B **239** (1990) p. III. 36.
- [12] F. Petry *et al.*, Nucl. Instrum. Meth. AA **332**, (1993) 107.
- [13] A. Balysh *et al.* (HEIDELBERG-MOSCOW collaboration), Phys. Lett. B **322** (1994) 176.
- [14] F.T. Avignone *et al.*, Phys. Lett. B **256** (1991) 559.
- [15] M. Beck *et al.* (HEIDELBERG-MOSCOW collaboration), to be published
- [16] D. Reusser *et al.*, Phys. Lett. B **255** (1991) 143.

ATOMIC PHYSICS



ATOMIC ELECTRIC DIPOLE MOMENTS AND CP VIOLATION

S.M.Barr
Bartol Research Institute
University of Delaware
Newark, DE 19716
USA

Abstract

The subject of atomic electric dipole moments, the rapid recent progress in searching for them, and their significance for fundamental issues in particle theory is surveyed. In particular it is shown how the edms of different kinds of atoms and molecules, as well as of the neutron, give vital information on the nature and origin of CP violation. Special stress is laid on supersymmetric theories and their consequences.

I. INTRODUCTION

In this talk I am going to discuss atomic and molecular electric dipole moments (edms) from a particle theorist's point of view. The first and fundamental point is that permanent electric dipole moments violate both P and T. If we assume, as we are entitled to do, that CPT is conserved then we may speak equivalently of T-violation and CP-violation. I will mostly use the latter designation.

That a permanent edm violates T is easily shown. Consider a proton. It has a magnetic dipole moment oriented along its spin axis. Suppose it also has an electric edm oriented, say, parallel to the magnetic dipole. Under T the electric dipole is not changed, as the spatial charge distribution is unaffected. But the magnetic dipole changes sign because current flows are reversed by T. Thus T takes a proton with parallel electric and magnetic dipoles into one with antiparallel moments. Now, if T is assumed to be an exact symmetry these two experimentally distinguishable kinds of proton will have the same mass. But we know that there is only one type of proton in the world, as otherwise the chart of nuclides would look very different. Thus either the proton has no edm or T is not an exact symmetry. The same argument applies to any other elementary particle (that is not "CP doubled", which includes all the known ones). The same kind of argument shows that edms violate P.

Now, CP violation is not well understood. That it is violated has been known since the classic experiment of Christenson, Cronin, Fitch and Turlay in 1964¹⁾ for which Cronin and Fitch received the Nobel Prize in 1980. What that experiment showed was that CP was violated in the neutral Kaon system; in particular that K_L decayed about two times in a thousand into two pions rather than three. It is remarkable that from that time until now – thirty years – CP violation has been observed in the laboratory in no other system than the neutral Kaons. I put in the qualifier "in the laboratory" because according to the currently standard picture of cosmology the cosmological baryon asymmetry (CBA), which is the fancy name for the fact that the matter in the universe is almost entirely matter rather than antimatter, is due to CP-violating processes in the early universe.²⁾ However, this fact, while it almost certainly points to CP-violating physics beyond the standard model³⁾, is of little help in understanding CP violation because there are so many possible scenarios of cosmological baryon generation.

From the neutral kaon system there are two CP-violating quantities which can be measured, ϵ and ϵ' . ϵ has been known for thirty years ($|\epsilon| = 2.3 \times 10^{-3}$), while the situation with ϵ' remains murky, with one experiment showing a definitely non-zero value⁴⁾ and the other⁵⁾ being consistent with zero. The Particle Data Book gives the average of the two as $|\epsilon'/\epsilon| = 2.2 \pm 1.1 \times 10^{-3}$.

With so little data to work with it is not surprising that CP violation is poorly understood. Our other information about CP violation is in the form of null results. The most important of these is the limit on the neutron edm,⁶⁾ which is $d_n \leq 1.2 \times 10^{-25}$ e-cm. As we shall see, this leads to a major puzzle for particle theory called the Strong CP Problem. As we shall also see, the limits on the edms of atoms and molecules are also theoretically very important and will become greatly more important as they are pushed further, unless, of course atomic or molecular edms are seen, which would be epoch-making.

It should be emphasized that the subject of CP violation is not some obscure corner of particle physics. It is a central topic of inquiry. It touches on such diverse issues as the edms of elementary particles and the origin of matter in the universe (as already noted), as well as on the nature of dark matter ("axions" are a prime candidate), the origin of galaxies, and the structure of physics beyond the so-called Standard Model of particle physics, as we shall see. And, of course, it is a question of fundamental interest in itself whether the laws of physics are T invariant and if not why. Fortunately, there are good prospects of learning more about CP violation over the next ten years. The main hope lies in the direction of further progress

in atomic and molecular edm experiments and in the study of neutral B mesons (a heavier analogue of the neutral kaons) at either “B factories” or at the LHC. There is also a large and promising effort being made to detect dark-matter axions by a variety of techniques, as is being discussed in other talks at this conference.

II. CP VIOLATING PHASES

CP violation arises, essentially, because of the presence in the underlying particle theory of non-trivial, physical, complex phases. Not all phases are physical. Some can be “rotated away”; that is, they can be absorbed by field redefinitions.

Take, for example, QED, the theory of photons and electrons. In that theory the electron mass parameter can be complex: $m_e e^{i\theta} \bar{e}_L e_R + \text{h.c.}$. The phase θ is not physical as it is possible to eliminate it by redefining the righthanded electron field: $e'_R \equiv e^{i\theta} e_R \Rightarrow m_e \bar{e}_L e'_R + \text{h.c.}$

In simple theories there are often few enough parameters that all the phases can be rotated away, as in QED. Such theories automatically conserve CP. But as a theory becomes more complicated the number of CP-violating phases tends to proliferate rapidly.

III. CP VIOLATION IN THE STANDARD MODEL

The Standard Model of particle physics is so simple that there are only two possible CP-violating phases in it. They are called δ (or δ_{KM}) and $\bar{\theta}$. δ is the Kobayashi-Maskawa phase which appears in the charged weak interactions, while $\bar{\theta}$ is the QCD vacuum angle which appears in the theory of the strong interactions. Let us discuss the latter first.

A. The Strong CP Problem

The Strong CP Problem, or θ Problem, arises due to the fact that the neutron edm gets a contribution of order $(10^{-16} \text{ e-cm}) \times \sin \bar{\theta}$. The experimental limit, therefore, implies that $\bar{\theta} \lesssim 10^{-9}$. The problem is to explain why, given the presumed fact that the weak interactions badly violate CP (suggested by the neutral Kaon system), the strong interactions conserve CP to at least a part in a billion. Any explanation of this requires going beyond the Standard Model.

There are two main approaches to solving this problem, the Peccei-Quinn mechanism⁷⁾ and models with spontaneous CP breaking⁸⁾.

The Peccei-Quinn mechanism involves the existence of the particles called “axions”. (Alternatively there could be massless quarks. For example, it is much discussed whether the up quark could in actuality be massless.) In the axion models $\bar{\theta}$ comes out to be of order 10^{-15} , generally. (In exotic variants it could be different.) This would give a value of order 10^{-31} e-cm for the $\bar{\theta}$ contribution to d_n , far too small to be interesting experimentally. Axion models are not without difficulties. They are highly constrained by laboratory searches, and astrophysical and cosmological limits. Moreover there are potential problems with cosmic domain walls and with possible quantum gravity effects that have to be faced in constructing such models. Nevertheless, the axion idea still has to be considered the most attractive solution to the strong CP Problem.

The alternative approach is based on the idea that $\bar{\theta}$ is small as the consequence of some symmetry. The most likely possibility is that this symmetry is CP invariance itself. The idea is that CP is a good symmetry of the Lagrangian which is “spontaneously broken” by a CP-non-invariant ground state. In this way some CP-violating parameters may be small as a consequence of the underlying symmetry of the theory, while others may feel the spontaneous

CP breaking more directly and have larger values. In many of these models there is some difficulty in making $\bar{\theta}$ come out as small as 10^{-9} . This may be even more of a problem in the context of supersymmetry. However, this idea has attractive features and quite viable and relatively simple models exist. One would expect that if this solution of the Strong CP Problem is the correct one that $\bar{\theta}$ would be not far below present limits.

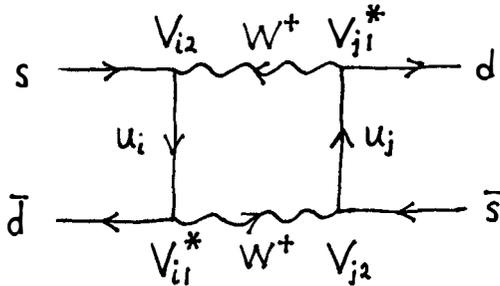
We see, therefore, that edm experiments already have had a major impact on particle theory. And it should be emphasized that the limits on $\bar{\theta}$ coming from searches for the edms of diamagnetic atoms and molecules are competitive with that coming from the neutron, and it is quite likely that $\bar{\theta}$ may eventually first be seen in atomic experiments.

If an edm of the neutron or a diamagnetic atom or molecule is observed, how will we know that it is a consequence of $\bar{\theta}$? With just this kind of experiment we will not. There are many kinds of new CP-violating physics that could give rise to these effects. The answer will probably come from experiments on paramagnetic atoms or molecules. These are sensitive to the edm of the electron. Most sources of CP violation contribute significantly to both the electron edm and to hadronic effects such as the neutron edm. A distinctive feature of $\bar{\theta}$ is that it contributes strongly to the latter but utterly negligibly to d_e , since the electron has no strong interactions.

B. Weak Interaction CP Violation in the Standard Model

In the charged weak interactions of quarks, a W^+ boson can be absorbed by a down-type quark, $d_j \equiv (d, s, b)$, and be converted into an up-type quark, $u_i \equiv (u, c, t)$. This, and the related processes, are governed by a matrix, V_{ij} , called the Kobayashi-Maskawa matrix. Being a 3×3 unitary matrix it is specified by 9 independent quantities. Five of these are phases which can be absorbed by redefining the relative phases of the six quark “flavors”. Remaining as physical parameters are three mixing angles (essentially Euler angles in generation-space) and one CP-violating phase, which is δ .

Thus, weak-interaction CP violation only shows up in “flavor-changing” processes in which one kind of quark turns into another. It can show up, therefore, in the oscillations of neutral Kaons, $\bar{K}^0 \rightleftharpoons K^0$, where effectively $\bar{d} + s \rightleftharpoons \bar{s} + d$. This occurs through the well known “box diagram” shown here:



Notice that because there is a net flavor change the phases in the product of factors $V_{i1}^*V_{i2}V_{j1}^*V_{j2}$ do not cancel. This is the mechanism that is generally thought to give rise to the ϵ parameter in the Kaon system. A crucial fact, however, is that the Kobayashi-Maskawa phase contributes negligibly to electric dipole moments. Because the edm operator conserves the flavor of the particle the KM phase cancels in low-order diagrams. For example, a one loop diagram for the edm of the d quark would involve the d emitting a W^- to become a $u_i \equiv (u, c, t)$ and reabsorbing it to become a d again. This depends on the product $V_{i1}^*V_{i1}$, which is obviously

real. In fact it was shown by Donoghue⁹⁾ for the electron and independently by Shabalin⁹⁾ for quarks that even the two loop Weak contributions to the edm vanish. In fact it has recently been shown by Khriplovich et al.⁹⁾ that the three-loop Weak contributions vanish as well.

The weak-interaction contribution to d_n is expected to be of order 10^{-31} e-cm if δ_{KM} is of order unity. d_e is expected to be many orders of magnitude less than that, though no reliable estimate exists. In other words, there is no prospect of seeing these contributions.

C. Beyond the Standard Model: edms an ideal way to search

Because the Kobayashi-Maskawa phase contributes negligibly to edms they are an ideal place to search for CP violating effects that go beyond the Standard Model. Any edm that is observed necessarily tells us about physics beyond the Standard Model. If it is due to $\bar{\theta}$ that is so because any resolution of the Strong CP Problem must involve new physics, whether axions or the machinery associated with spontaneous CP violation. And the observation of a neutron or atomic edm due to $\bar{\theta}$ would provide evidence against the axion solutions, which, as we have noted, tend to predict $d_n \sim 10^{-31}$ e-cm. If $\bar{\theta}$ is not involved, then any edm observed must originate from new sources of CP violation that lie beyond the Standard Model.

There are many compelling reasons to believe that the Standard Model is not the ultimate theory of non-gravitational particle interactions. The subject of CP violation provides us with at least two reasons, the Strong CP Problem, and the necessity (as it presently appears) to have a new source of CP violation to generate the baryon asymmetry. There are a number of other reasons as well. There is strong circumstantial evidence that there is a grand unification of electroweak and strong interactions. There is the technical issue of the "Gauge Hierarchy Problem" that cannot be resolved within the Standard Model. And there is the problem of understanding the spectrum of quarks and leptons.

What is exciting is that, for reasons explained above, most extensions of the Standard Model to incorporate new physics tend to bring in new CP violating phases, and richer possibilities for edms. Some of the more frequently discussed extensions are supersymmetry (SUSY), multi-Higgs models, and Left-Right models. I will discuss the first two, which I believe to be the best motivated. But first I will discuss very briefly the subject of atomic and molecular edms.

IV. ATOMIC EDMs

Atomic and molecular permanent electric dipole moments can arise from a variety of sources: an edm of the electron, an edm of the neutron or proton, T-odd and P-odd electron-nucleon interactions, and T-odd and P-odd nucleon-nucleon interactions.

A well-known fact in this subject is Schiff's Theorem, which states that a nonrelativistic assemblage of point charges in equilibrium under the action of electrostatic forces alone will not have any permanent edm even if the constituent particles do.¹⁰⁾ This is fairly easy to prove and can be understood even more easily as follows. Consider an atom in which the nucleus has an edm. The nucleus, being charged, will drift to a region where it feels no net electric field, no matter what external electric field may be applied to the atom. Thus, there is no $\vec{E} \cdot \vec{d}_{nuc}$ interaction energy; and the nuclear edm does not contribute to the atomic edm. In other words the electron cloud acts to screen the nuclear edm.

In this example it is clear that we have assumed that the nucleus is moving slowly, is acted upon only by the electrostatic forces, and is a point particle. In actual atoms the finite size effect of the nucleus, non-electrostatic forces, and relativistic effects would allow an atomic edm to exist if CP violation is present. However, in many cases the atomic edm will have a "Schiff suppression", so that $d_{atom} \ll d_e, d_{nucleus}$.

In large atoms, where relativistic effects can be quite important, if there is an unpaired electron there can actually be a large **enhancement**. Typically, in such “paramagnetic” atoms $d_{atom} \sim (Z^3 \alpha^2) \cdot d_e$. (One factor of Z comes from the coulomb field of the nucleus. The other factors of $Z\alpha$ can be understood as relativistic in origin.) When Z is of order 100 then enhancements of order 100 to 1,000 are possible. For example,¹¹⁾ for Cesium ($Z = 55$) $|d_{Cs}| \simeq 120 |d_e|$; and for atomic Thallium ($Z = 81$) $|d_{Tl}| \simeq 600 |d_e|$.

In “diamagnetic” atoms, where electrons are paired, there is of course a suppression of the d_e contribution. For example,¹¹⁾ $|d_{Hg}| \simeq 1.2 \times 10^{-2} |d_e|$.

As it turns out, an edm of a paramagnetic system is most likely to come from the edm of the electron, while an edm of a diamagnetic system is most likely to come from the nucleus (in particular the so-called “Schiff moment” of the nucleus – a finite-size effect whereby there is a mismatch between the charge distribution and the edm distribution inside the nucleus). The Schiff moment, in turn, would come primarily from the T-odd and P-odd nucleon-nucleon interactions. Essentially, diamagnetic systems look at the same underlying CP violating effects at the quark level as does the neutron edm. For some quark-level operators the neutron may be more sensitive, while for others the atoms may be. Both approaches are important.

It should be emphasized that diatomic molecules have further enhancements that make them very useful. For example, the molecule TlF , although it has paired electrons, has $|d_{TlF}/d_e| \simeq 80$; and it has been estimated that paramagnetic diatomic molecules can have enhancements of even 10^7 . If the latter can be exploited, it may be possible to see an electron edm as small as 10^{-29} e-cm!

V. MULTI-HIGGS MODELS

Several contexts in particle theory call for a doubling of the Higgs sector of the theory: certain grand unified theories, like $SO(10)$ (though both Higgs need not be light), certain axion models (in particular DSFZ models), and supersymmetry. Such a doubling makes possible in some cases CP violation in the Higgs sector.

Consider the Higgs potential of the Standard Model.

$$V_{SM} = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (1)$$

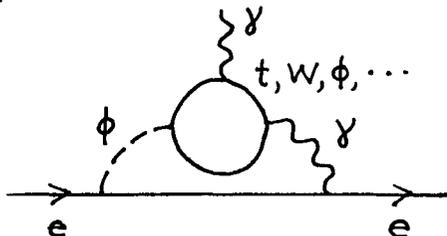
Since the potential energy is real and the operator $\phi^\dagger \phi$ is hermitian, the parameters λ and μ^2 are real quantities. However, consider the Higgs potential of a model with two Higgs doublets.

$$\begin{aligned} V_{2-Higgs} = & \mu_1^2 (\phi_1^\dagger \phi_1) + \lambda_1 (\phi_1^\dagger \phi_1)^2 \\ & + \mu_2^2 (\phi_2^\dagger \phi_2) + \lambda_2 (\phi_2^\dagger \phi_2)^2 \\ & + (\mu_{12}^2 \phi_1^\dagger \phi_2 + \mu_{12}^{*2} \phi_2^\dagger \phi_1) \\ & + (\lambda_{12} (\phi_1^\dagger \phi_2)^2 + \lambda_{12}^* (\phi_2^\dagger \phi_1)^2) \\ & + \text{other terms} \end{aligned} \quad (2)$$

Here we see that there are several terms that involve nonhermitian operators and hence can have complex coefficients, for example μ_{12}^2 and λ_{12} . One complex phase can be absorbed into the relative phase of ϕ_1 and ϕ_2 , so that if there is more than one complex parameter in the Higgs potential CP violation results.

This CP violation shows up, among other places, in the propagator of the neutral Higgs; for instance, if a ϕ_1^0 makes a transition into a ϕ_2^0 . In 1989 Weinberg¹²⁾ pointed out that such an effect could lead to a measurable edm for the neutron through a two-loop diagram which induces a T-odd and P-odd three gluon operator. In 1990, Zee and I observed¹³⁾ that a different class of two-loop diagrams can, in these models, induce directly edms for the electron and quarks as

well as “chromo-electric dipole moments” for the quarks. The Weinberg three-gluon operator, the edms of the quarks, and the chromo-edms of the quarks should all contribute roughly comparably to d_n , with the chromo-edms perhaps dominating. A diagram that contributes to d_e is the following.



This diagram gives, typically, if the CP violating phases are of order unity and the Higgs masses are of the order of 100 GeV, $d_e \simeq 10^{-27}$ to 10^{-26} e-cm. This should be compared to the present best limit on d_e coming from the Thallium experiment of the Berkeley group: $d_e = (-2.7 \pm 8.3) \times 10^{-27}$ e-cm. Clearly these experiments are able to tell us important things about the structure of the Higgs system!

It is worth emphasizing to those who have not followed these developments that in the last several years there has been remarkable progress in the search for d_e as can be seen in the following table listing some of the limits obtained.¹⁴⁾

		$d_e(e - cm)$
1968	<i>M.C.Weisskopf et al.</i>	$\leq 3 \times 10^{-24}$
1970	<i>H.Gould</i>	$(-2.6 \pm 4.8) \times 10^{-24}$
1987	<i>S.K.Lamoreaux et al.(Hg)</i>	$(-0.5 \pm 1.1) \times 10^{-24}$
1989	<i>S.A.Murthy et al.(Cs)</i>	$(-1.5 \pm 5.5 \pm 1.5) \times 10^{-26}$
1989	<i>D.Cho et al.(TlF)</i>	$(-1.4 \pm 2.4) \times 10^{-25}$
1990	<i>K.Abdullah et al.(Tl)</i>	$(-2.7 \pm 8.3) \times 10^{-27}$

VI. SUPERSYMMETRY

There is a growing belief among particle theorists that at energy scales above a few hundred GeV or a few TeV nature will exhibit a symmetry called supersymmetry (“SUSY”). There are at least three reasons for this.

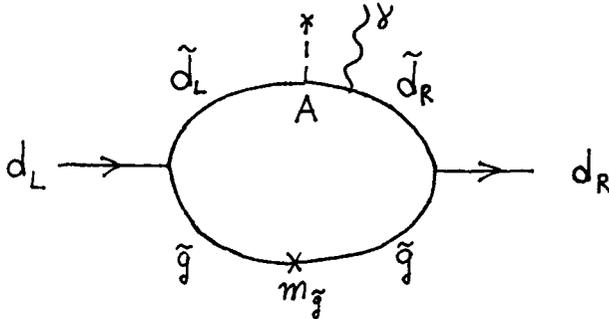
First, there is a fundamental problem in particle physics called the Gauge Hierarchy Problem, which is the problem of understanding how the scale of the weak interactions ($M_W \cong 80$ GeV) can be so much smaller than either the Planck scale (10^{19} GeV) or the scale of grand unification (10^{16} GeV). The problem is that without fine-tuning of parameters to an accuracy of 26 decimal places radiative corrections should drive the weak scale up to the unification scale. Supersymmetry seems to be an essential ingredient in a natural solution to this difficulty.

Second, there are strong reasons to suppose that superstring theory may be the correct theory that unifies gravity with the other forces. In particular, superstrings appear to tame the severe problems that arise in quantizing General Relativity. And supersymmetry is a feature of superstring theories.

And, third, there is the fact that the three gauge couplings of the Standard Model (which are functions of energy) converge to a point when extrapolated to high energy if supersymmetry is assumed above the weak scale, whereas they do not so converge in the non-supersymmetric Standard Model. Conversely, if supersymmetry and unification of couplings are assumed there is a "prediction" for the ratio of electroweak couplings measured at low energies: $\sin^2 \theta_W = 0.2334 \pm 0.0036$ (SUSY GUTS), $\sin^2 \theta_W = 0.2324 \pm 0.0003$ (from experiment).

A notable example of expectations among theorists is provided in remarks made by S. Weinberg in his summary talk at the XXVIth ICHEP meeting in Dallas in 1992. In projecting the future shape of particle physics he said "My ... guess ... is that the hierarchy problem will be solved by supersymmetry. I say this for a number of reasons ..."

What is exciting for us is that supersymmetry can give rise to large edms for both the neutron and the electron.¹⁵⁾ A diagram that gives quark edms and thus a neutron edm is the following.



There are other diagrams as well. \tilde{d}_L and \tilde{d}_R are the scalar partners under supersymmetry of the left- and right-handed down quarks, respectively. \tilde{g} is the gluino, the spin-1/2 partner of the gluon. There can be CP-violating phases in the parameter A and in the gluino mass, $m_{\tilde{g}}$. There are similar graphs for the electron edm with photinos, neutralinos, or charginos instead of the gluino, and scalar electrons instead of scalar quarks.

If one assumes that the relevant CP-violating phases are of order unity, and that the unknown masses of the various superparticles is of order 100 GeV, then one finds that typically $d_n \sim 10^{-23}$ to 10^{-22} e-cm - that is, 10^2 to 10^3 times the experimental limit! Likewise, typically one finds $d_e \sim 10^{-25}$ e-cm, which is about 10 times the present limit. In other words it is a bit of a puzzle why edms of the neutron electron and atoms have not been seen already.

There are several possibilities for what might be going on.

(A) The reason may simply be numerical. That is the real world may correspond to a point in parameter space where the edms come out smaller than their typical values. This would require that the CP-violating phases be somewhat small, the scalar quarks and leptons be quite heavy, or, probably, both. If this is the case, then edms cannot hide for much longer. They should be "around the corner".

(B) There may be some suppression mechanism, dynamical or symmetry-based, that makes the new CP-violating phases of supersymmetry, such as the phases or the gluino mass, the photino mass or the A -parameter, very small. One idea for suppressing these phases involves CP being a spontaneously broken symmetry. We saw that ideas such as this have been invoked to suppress the θ parameter. It should be noted that these new phases of supersymmetry appear in the so-called soft supersymmetry-breaking terms. So, another idea is that the breaking of supersymmetry occurs in such a way that these terms conserve CP. Perhaps this can happen in certain schemes of dynamical supersymmetry breaking.

If there is such a suppression mechanism at work, it would be a model dependent question by how much the edms are suppressed below their "typical" values. Some models exist in the literature¹⁶⁾ in which this suppression is of order 10^{-2} to 10^{-3} . This is a reasonable expectation if the phases arise at one-loop order. In this case, the edms of atoms may be just around the corner. On the other hand, it is also possible that the suppression could be much greater, in which case edms would not be observed.

(C) Supersymmetry may not exist.

If supersymmetry does exist, and that should be discoverable at the LHC, if not sooner, then the search for edms will – one way or the other – provide information about CP-violation of enormous importance. If edms are seen, of course it will be epoch-making. Even if not seen it will tell us something crucial either about the way CP is violated or about the way supersymmetry is broken, or both.

It is anyone's guess whether atomic edms will be seen. I will close with another quote from the same Dallas talk by Weinberg: "Also endemic in supersymmetry are CP violations that go beyond the CKM [Cabbibo-Kobayashi-Maskawa] matrix, and for this reason it may be that the next exciting thing to come along will be the discovery of a neutron or atomic or electron electric dipole moment. These electric dipole moments were just briefly mentioned at this conference, but they seem to me to offer one of the most exciting possibilities for progress in particle physics."

REFERENCES

- (1) J.H. Christenson, J.W. Cronin, V.L. Fitch, and R. Turley, *phys. Rev. Lett.* **13**, 138 (1964).
- (2) A.D. Sakharov, *JETP Letters* **5**, 24 (1967). For a review see chapter 6 of *The Early Universe*, E.W. Kolb and M.S. Turner (Addison-Wesley, Redwood City, Calif. 1989).
- (3) P. Huet and E. Sather, SLAC-PUB-6479 (1994).
- (4) NA31 group at CERN: B. Renk, in *Proc. of the XXVIth Int. Conf. on High Energy Physics*, p. 537, 1992, Dallas, Texas, edited by J.R. Stanford; H. Burkhardt et al. *Phys. Lett. B* **206**, 169 (1988).
- (5) B. Winstein, in *Proc. of Lepton-Photon Symposium*, Geneva, 186, 1991; J.R. Patterson et al., *Phys. Rev. Lett.* **64**, 1491 (1990).
- (6) K.F. Smith et al., *Phys. Lett. B* **234**, 191 (1990); I.S. Altarev, et al., *JETP Lett.* **44**, 460 (1986).
- (7) R.D. Peccei and H. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977).
- (8) For a review see talk of S.M. Barr in *CP Violation in Particle Physics and Astrophysics*, edited by J. Tran Thanh Van (Editions Frontières, 1990). Also see S.M. Barr and G. Segre, *Phys. Rev. D*, 1992.
- (9) J. Donoghue, *Phys. Rev. D* **18**, 1632 (1978). E.P. Shabalin, *Sov. J. Nucl. Phys.* **28**, 75 (1978). I.B. Khriplovich, private communication.
- (10) L.I. Schiff, *Phys. Rev.* **132**, 2194 (1963). For a review of Atomic edms see S.M. Barr, *Int. J. of Mod. Phys. A* **8**, 209 (1993).
- (11) I.B. Khriplovich, *Parity Non-Conservation in Atomic Phenomena* (Gordon and Breach, Amsterdam 1991) and references therein.
- (12) S. Weinberg, *Phys. Rev. Lett.* **63**, 2333 (1989).
- (13) S.M. Barr and A. Zee, *Phys. Rev. Lett.* **65**, 21 (1990).
- (14) M.C. Weisskopf et al., *Phys. Rev. Lett.* **21**, 1645 (1968); H. Gould, *Phys. Rev. Lett.* **24**, 1091 (1970); S.K. Lamoreaux, J. Jacobs, B. Heckel, F. Raab, and E.N. Fortson, *Phys. Rev.*

- Lett. **59**, 2275 (1987); S.A. Murthy, D. Krause Jr., Z.L. Li, and L.R. Hunter, Phys. Rev. Lett. **63**, 965 (1989); D. Cho, K. Sangster, and E.A. Hinds, Phys. Rev. Lett. **63**, 2559 (1989); K. Abdullah, D. Krause, Jr., E.D. Commins, H. Gould, and S. Ross, Phys. Rev. Lett. **65**, 2347 (1990).
- (15) J. Ellis, S. Ferrara, and D.V. Nanopoulos, Phys. Lett. **B114**, 231 (1982); W. Buchmuller and D. Wyler, Phys. Lett. **B121**, 321 (1982); J. Polchinski and M.B. Wise, Phys. Lett. **B125**, 393 (1983); E. Franco and M. Mangano, Phys. Lett. **B135**, 445 (1984).
- (16) K.S. Babu and S.M. Barr, Bartol preprint BA-93-48 (1993), to be published in Phys. Rev. Lett.; R. Garisto, TRIUMF preprint, TRI-PP-93-86 (1993).

A TEST OF TIME REVERSAL SYMMETRY USING ^{199}Hg

Steve K. Lamoreaux
University of Washington
Physics Department, FM-15
Seattle, WA 98195 USA

An experimental measurement of the ^{199}Hg atomic electric dipole moment (EDM) has yielded the null result $d(^{199}\text{Hg}) < 1.3 \times 10^{-27}$ e cm. This is the smallest limit ever set on a system electric dipole moment, and sets the most stringent limits to date on several sources of time-reversal violation, and gives limits comparable to those set by the neutron EDM. A brief review of the technical innovations which made such precision possible will be given.

I. Introduction

The nature of the interaction which leads to the observed violation of CP symmetry (combined operation of charge conjugation and parity inversion) in a rare decay mode of the K_0 meson remains an enigma.¹⁾ Since $CPT = 1$, the observed CP implies time reversal symmetry violation. Theories put forward to explain the observed CP violation predict the existence of permanent electric dipole moments (EDMs) of elementary particles (e.g., the electron) and more complex structures (e.g., neutron, atoms).

A system with angular momentum I which has an EDM interacts with electric and magnetic fields as described by the following Hamiltonian:

$$H = -(d\vec{E} + \mu\vec{B}) \cdot \vec{I}/I \quad (1)$$

where d and μ represent the electric and usual magnetic dipole moments, \vec{E} and \vec{B} are the applied electric and magnetic fields. The Hamiltonian is written in this form because any vector expectation value associated with the system must lie along \vec{I} , otherwise additional quantum numbers would be required to describe the system; that there are only two states associated with the neutron (ground state) is well-verified by the observation that nucleons in composite nuclei obey the Pauli exclusion principle.

An important feature of Eq. (1) is that the EDM is proportional to I , just as the magnetic moment is. The implication is that an EDM is the result of internal (time-dependent) dynamics, and represent it as a simple charge separation along I isn't correct. When the system is reflected in the mirror, the EDM direction must reverse with I , in order for there to be an observable which manifests T violation. This can be seen from Eq. (1); under time reversal, $\vec{E} \rightarrow \vec{E}$, $\vec{B} \rightarrow -\vec{B}$, while under parity, $\vec{E} \rightarrow -\vec{E}$, $\vec{B} \rightarrow \vec{B}$, and in order for the eigenvalues of Eq. (1) to show parity or time reversal effects, the EDM must be proportional to \vec{I} .

To measure the EDM of an object, one in principle could simply look for the precession of the spin about an applied electric field. In the case of spin $1/2$, any spin dependent interaction with the applied field is strictly forbidden (Kramer's theorem, which is a statement about time reversal symmetry). For spin $> 1/2$, there can be an electric polarizability effect, which should be the same independent of electric field direction, unlike a true EDM. Perfect reversal of an electric field is difficult in practice.

By Eq. (1), if parallel electric and magnetic fields are applied, on reversal of \vec{E} relative to \vec{B} , there will be a shift in Larmor frequency

$$\omega = 2dE/\hbar. \quad (2)$$

Equation (1) is also applicable to the internal interactions within a molecule, and if one of the atoms of a polar molecule has an EDM, the effects can be evident in the molecular hyperfine structure.

II. Theoretical Interpretation of Atomic EDM Measurements

A splendid review of the fundamental processes which can lead to an EDM is given by S. Barr,²⁾ and a review of the calculation of atomic EDMs is given by A.-M. Martensson.³⁾ For the present discussion, I will compare limits set by our ^{199}Hg EDM experiment⁴⁾ to those of the neutron.⁵⁾ A full discussion of the ^{199}Hg result is given in Ref. 4).

It has been long known that a neutral bound system of point particles interacting electrostatically cannot manifest the effects of an edm in the nonrelativistic limit. This is simply because under equilibrium conditions, the average electric field at each constituent particle is zero. Thus, one might expect that to search for an atomic EDM is a bit pointless. Schiff discussed possible ways around this problem;⁶⁾ in the case of atoms, in addition to electrostatic interactions, there is also the magnetic fine structure and hyperfine interactions. Furthermore, the electrons are relativistic near the nucleus, and the nucleus is of finite extent. Thus, we can expect a non-zero atomic EDM if, for example, the electron has an EDM. In the case of the alkali atom Cs, the atomic EDM is actually about 120 times the electron EDM.⁷⁾ For ^{199}Hg , the effect is less favorable, where the atomic EDM is about 1/100 of the electron EDM. The difference between these two atoms is the electron spin: Hg in the ground state has none, and an electron EDM couples in third order through the hyperfine interaction.

In the case of a nuclear EDM, the atomic effect can be estimated, following arguments used in a related calculation of an intrinsic proton EDM:⁸⁾

$$d_{atom} \approx Z^2 K_r (R_{nucleus}/R_{atom})^2 d_{nucleus} \quad (3)$$

where Z is the atomic number and K_r is a relativistic enhancement factor of order 10 for Hg; we thus see an atomic EDM about a factor of 100 less than the nucleus. If we think of the nuclear EDM as being that of the valence nucleon, we can thus think of our atomic Hg experiment having an inherent factor of 100 reduction in sensitivity over the neutron. In fact, our measurement⁴⁾

$$d(^{199}\text{Hg}) < 1.3 \times 10^{-27} \text{e} \cdot \text{cm}$$

(95% confidence) limits the QCD θ parameter to

$$\theta_{QCD} < 7 \times 10^{-10}$$

whereas the neutron limit

$$d(n) < 1.2 \times 10^{-25} \text{e} \cdot \text{cm}$$

limits

$$\theta_{QCD} < 4 \times 10^{-10}.$$

Limits on T -odd quark-quark interactions are also comparable between the two.⁹⁾

III. Experimental Technique

Our first measurement of the ^{199}Hg EDM was completed in 1987 set a limit of 3×10^{-26} e cm;¹⁰⁾ our recently completed work has improved this limit by a factor of 25. This work is fully described in Refs. 4) and 11)

The major factors contributing to the increased sensitivity are as follows: use of isotopically enriched ^{199}Hg , improved optical pumping cells, and the use of a transverse pumping scheme (also, considerably more data was taken with the new apparatus). The transverse pumping scheme, first suggesting in 1987 and described in Ref. 12), eliminates possible effects due to the AC Stark shift of the optical pumping light since the light propagation direction and magnetic fields are perpendicular. In addition, fluctuations in this "light shift" seemed to be a source of excess noise in the earlier version of the experiment. The transverse pumping scheme gives a factor of two increase in inherent sensitivity.

We opted to use atomic oscillators as opposed to a pump/probe free precession technique because the latter requires switching static magnetic fields within the shields. At our level of sensitivity, magnetic field drifts associated with such switching would be unacceptable. However, I have proposed a new technique which gets around this problem, and has a number of advantages, as will be described later.

The optical pumping cells are the heart of the experiment, and took about 7 years to develop. The spin relaxation lifetime is about 70 seconds, electric fields up to 10kV/cm can be applied, and about 5×10^{12} atoms are contained in the cell (of total volume 5 cc). Two cells are used in the experiment; the electric field is applied oppositely between the two cells, and an EDM would be evident by a relative phase change between the respective atomic oscillator signals. This differential measurement discriminates drifts in the static homogeneous field, which tend to be about a factor of 10 greater than drifts in the gradient. So far, our result is consistent with the shot noise limit, however, a future increase in precision will require better field stability, or auxiliary measurements.

To search for an EDM, one simply periodically reverses the direction of the electric field; a correlation with the atomic oscillator frequency with electric field direction would be evidence of an EDM. The experiment operates with a 1/8 duty cycle for the following reasons: the atomic oscillators take about 20 seconds to respond, displacement currents associated with field-reversal charging currents can give a transient shift so we give the system time to respond, and data is taken with no voltage applied (between field reversals) to test for systematic effects. The electric field is reversed every 1000 seconds, and about 1/8 to 1/16 of the data can be used with confidence.

IV. Statistical Limits to Sensitivity

The basic experimental idea is to measure the spin precession frequency difference for

parallel and antiparallel magnetic fields. We have N uncorrelated atoms per cell, and on reversal of the electric field, for each cell, from Eq. (2), a net frequency uncertainty of

$$\delta\omega = \sqrt{2} \frac{1}{\tau} \frac{1}{\sqrt{N}} \quad (4)$$

is expected by the Heisenberg relation over an observation time $\tau \approx 25$ sec, the effective spin lifetime. After many such measurements, over a time T , the final uncertainty is

$$\delta\omega = \frac{\sqrt{2}}{\tau} \frac{1}{\sqrt{N}} \sqrt{\frac{\tau}{T}} = \sqrt{\frac{2}{\tau TN}}. \quad (5)$$

Taking an electric field of 10 kV/cm, and taking a difference between two cells (a net factor $\sqrt{2}$ increase in sensitivity over a single cell), we might expect a daily uncertainty of

$$\delta d \approx 3.4 \times 10^{-30} \text{ e} \cdot \text{cm}.$$

However, there are a number of reduction factors. The spin precession is detected by the absorption of photons. First, the signal modulation is only a fraction of the total transmitted photon flux detected (1/8). Next, the light collection efficiency is 25%, the photomultiplier quantum efficiency is 10%, and the net transmitted photon flux is $e^{-1.8} = .17$, and the incident flux is determined by N/τ . Finally, the duty cycle is 1/8. The last four factors enter as the square root, whereas the overall signal to noise factor 1/8 enters linearly (in addition, the effective linewidth of the atomic oscillator is twice the Heisenberg limit because not all atoms live for τ , this also enters linearly); we thus have a reduction factor

$$\frac{\sqrt{0.25 \times 0.10 \times 0.17 \times \frac{1}{8}}}{16} = 1.4 \times 10^{-3}$$

giving a daily statistical accuracy of about 2.5×10^{-27} e cm, which is experimentally observed (at 10kV/cm).

V. Systematic Effects Considered

A "quadratic" effect due to changes in the light shift of the pumping light (1mHz) with application of high voltage through the usual Stark effect is possible; field perpendicularity and polarization averaging to zero reduces the effective light shift to $< 1\mu\text{Hz}$, and the high voltage symmetry, certainly better than 10%, gives an EDM signal $< 10^{-28}$ e cm.

Leakage currents in the steady state of order 1 pA, by flowing in a loop around the cell, can give rise to an EDM of 5×10^{-29} e cm.

Displacement currents from the high voltage application can permanently magnetize the shield. Simulating these currents with wires placed in the shields but with current increase by 100-1000 times showed no remanance effects.

In addition, the zero-voltage data (i.e., atomic oscillator phase after each high voltage application) yielded no resolved effect.

Magnetic fields associated with the high voltage supply are $1\mu\text{G}$ 1 meter from the shields; the shielding factor on the gradient is $20 \times 70,000$ which gives an EDM of 10^{-28}e cm .

Changes in spatial dependence of the atomic polarization due to application of the high voltage was explored by applying a magnetic field gradient 50 times larger, along various directions, than is normally present, and no effects were observed.¹³⁾

The time dependence of the atomic oscillator signals, after application and removal of the high voltage, showed no resolved systematic effects.

The possibility of a $v \times E$ systematic magnetic field exists due to the difference between the spatial average of pumping vs. detection; this is estimated to be $< 10^{-28}\text{e cm}$.

VI. Future Prospects

In order to reduce dead-time effects and increase the signal-to-noise, we are now considering a pump/probe type of experiment. Furthermore, background effects inherent in the atomic oscillator effects will be reduced, and higher quantum efficiency photomultipliers (which have a light polarization dependent sensitivity) can be used, and the collection efficiency increased. In total, a factor of about 3 increase in sensitivity is expected.

Since the switching of static fields must be avoided, a coherent precessing polarization must be produced by other means. One possibility is to pump the atoms along the fixed static field direction, then use an RF pulse to rotate the polarization by $\pi/2$, thereby creating the precessing spin polarization. Unfortunately, the electrodes of the cells are not transparent to the 254 nm ultraviolet pumping light, so this is not possible; also, the application of the RF magnetic field could lead to drifts.

This led to the proposal of a "pulse pump" system; pumping light is incident on the cells from a direction perpendicular to the applied static magnetic field, and this light is pulsed at the Larmor frequency. A net spin polarization builds up, because each time the polarization precesses a cycle, it gets reinforced.

A Zeeman-shifted high intensity pump light has been constructed; the idea is that non-resonant light from ^{202}Hg can be Zeeman shifted to resonance, and the light parallel to the magnetic field is circularly polarized. Thus, the pumping light needs no additional polarizer. The light is chopped by a toothed disk turning at a frequency such that the pulses occur at the Larmor frequency. The pump light is only used for polarization, after which it is blocked. An atomic polarization of 80% has been produced by this method.

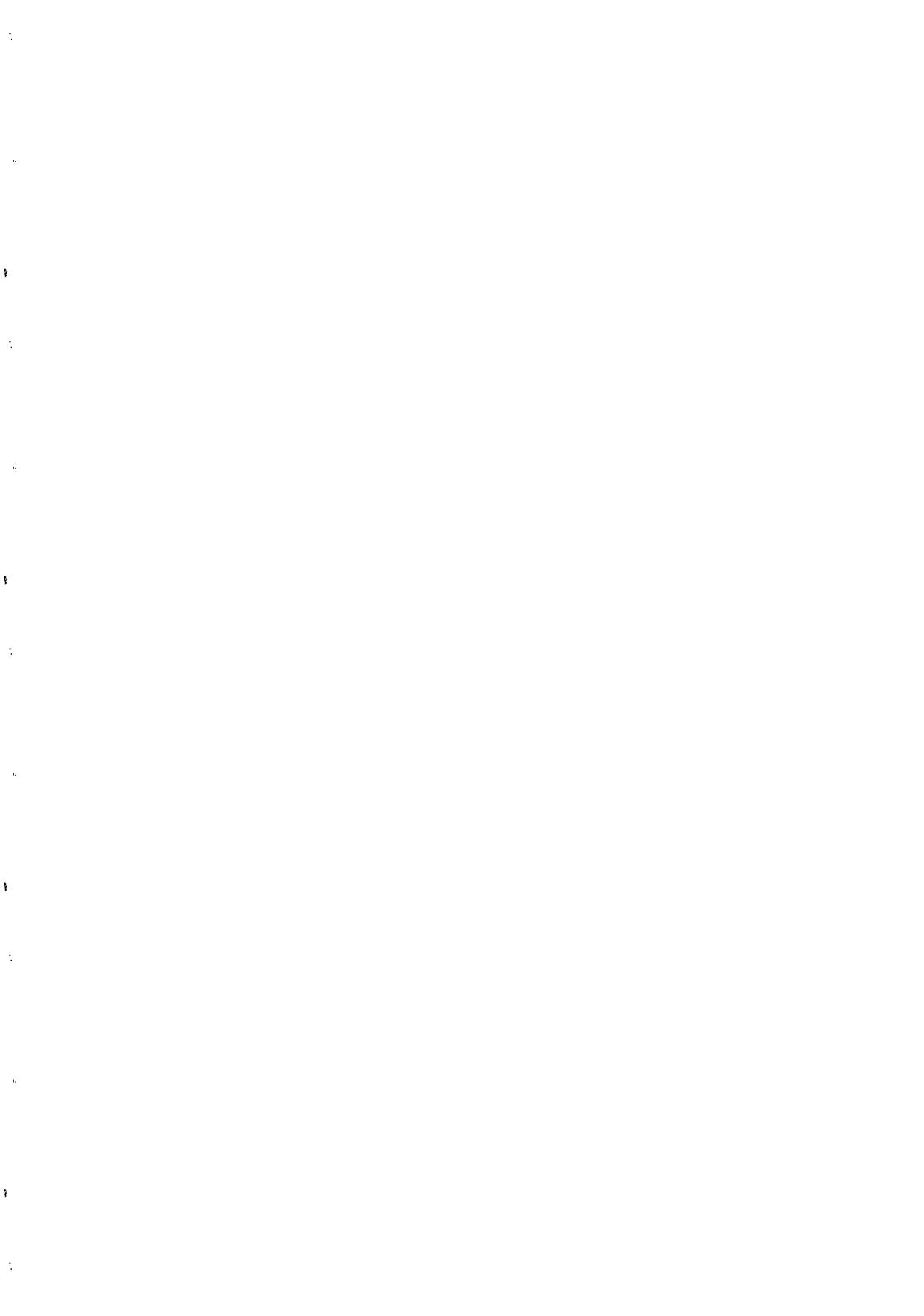
The probe light has its polarization modulated at high frequency by use of a photoelastic Lame plate modulator. Thus, the light shift is again reduced by this, and by having the incident probe light perpendicular to the applied field.

The gain in sensitivity comes about from an increased atomic polarization, decreased dead time (pump time is of order 10 sec, no waiting for atomic oscillator response or displacement current effects), and increased detection efficiency. At least a factor of three increase in sensitivity is expected, and systematic effects associated with the atomic oscillators will be eliminated, although the new technique will certainly have its own set of systematic problems.

We also hope to fabricate cells with a longer spin lifetime, and to incorporate either a Cs or Squid magnetometer system to monitor systematic and background magnetic fields, which are anticipated to be a problem at the new level of sensitivity. High T_c Squid magnetometers, or a High T_c flux transformer coupling to a conventional Squid outside the magnetic shield are being considered.

VII. References

- 1). J.H. Christenson, J.W. Cronin, V.L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).
- 2). S.M. Barr, Intl. Jour. Mod. Phys. A **8**, 209 (1993); S.M. Barr, these proceedings.
- 3). A.-M. Martensson, Meth. in Comp. Chem. **5**, 99 (1992).
- 4). J.P. Jacobs et al., Phys. Rev. Lett. **71**, 3782 (1993).
- 5). For a recent review of the neutron EDM, see R. Golub and Steve K. Lamoreaux, Phys. Rep. **237**, 1 (1994).
- 6). L.I. Schiff, Phys. Rev. **132**, 2194 (1963).
- 7). P.G.H. Sandars, Phys. Lett. **14**, 194 (1965); Phys. Lett. **22**, 290 (1966).
- 8). E.A. Hinds and P.G.H. Sandars, Phys. Rev. A **21**, 480 (1980).
- 9). V.M. Khatsymovsky and I.B. Khriplovich, Phys. Lett. B **296**, 219 (1992).
- 10). S.K. Lamoreaux et al., Phys. Rev. Lett. **59**, 2275 (1987).
- 11). J.P. Jacobs, Ph.D. thesis, Univ. Wash., Seattle (1991) (unpublished); J.P. Jacobs et al., in preparation.
- 12). S.K. Lamoreaux, NIM A **284**, 43 (1989).
- 13). S.K. Lamoreaux, in *Proceeding of Wein '92*, ed. Ts. D. Vylov, p. 440 (World Scientific, Singapore, 1993).



THE BERKELEY SEARCH FOR THE ELECTRON ELECTRIC DIPOLE MOMENT

Stephen B. Ross, Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748
Garching b. München, Deutschland



Abstract

We present the status of the Berkeley measurement of the electron electric dipole moment. The electron electric dipole moment can only be non-zero if the fundamental symmetries C and T are violated. To date, no non-vanishing dipole moment has been measured.

Although the Standard Model predicts a value of $d_e \cong 10^{-37} e \text{ cm}$, approximately ten orders of magnitude smaller than we can measure, extensions of the model predicts non-zero EDM's easily within experimental reach.¹⁾

We shall describe the basic elements of our experiment, and then provide the current expectations of machine performance.

1 Introduction

Our electron electric dipole moment (EDM) experiment is a classic Ramsey interference experiment. Fig. 1 is an experimental schematic. We make two vertical, counterpropagating beams of atomic thallium with effusive atomic beam ovens. The beams are not run simultaneously and are not cooled. We first optically pump the atoms, aligning the ground state. They then undergo a pulse of radio frequency, rotating the alignment into a superposition of two polarizations. They then enter the true interaction region, a large electric field.

In this experiment the interaction of interest is: $-\mathbf{d}_{atom} \cdot \mathbf{E}$, and this enters as a phase of opposite sign in the two polarization components. Atomic theory predicts that the atomic dipole moment of the ground state of thallium is -585 times larger than the electron dipole moment, so we can infer the electron EDM from this atomic measurement. This magnification is called the “enhancement factor.” As with all Ramsey interference experiments, we then pulse the atoms with a second rf region, which is phase related with the first. This allows us to extract $-\mathbf{d}_{atom}$ linearly with a second optical pumping region. The invariant associated with the measurement is: $\mathbf{B} \cdot \mathbf{E}$; a true EDM will change sign with electric and magnetic field reversal.

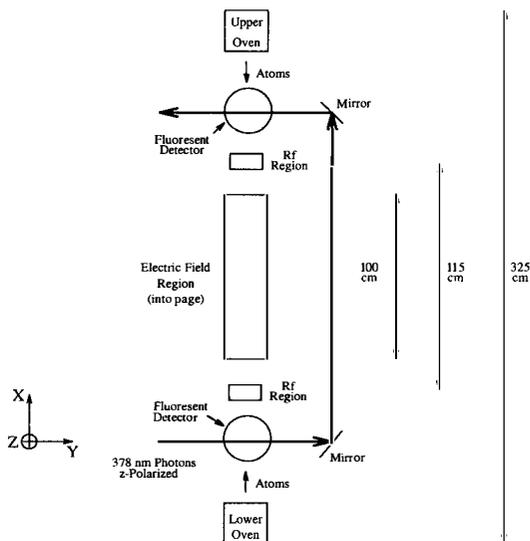


Figure 1: The necessary elements of the EDM experiment. Not drawn is a magnetic field in the z direction, which is constant over the entire apparatus except for the oven regions.

As already mentioned, this experiment measures the dipole moment of the ground state of atomic thallium, *not* the electron. However, we will discuss the measurement as an electron measurement, the true quantity of interest.

The most recent results of the two experiments most sensitive to the electron electric dipole moment are: L. Hunter and collaborators measured the electron EDM in cesium, and found, $d_e = -1.5 \pm 5.5 \pm 1.5 \times 10^{-26} e \text{ cm}$.²⁾ Most recently, our group measured: $d_e = -2.7 \pm 8.3 \times 10^{-27} e \text{ cm}$ in thallium.³⁾

2 An Experimental Overview

We chose the thallium atom because of its large enhancement factor and spectroscopic convenience. Fig. 2 is a level diagram of the three lowest levels. The nuclear spin is $1/2$, so the $6^2P_{1/2}$ and $7^2S_{1/2}$ states are each split by the hyperfine interaction. The $6^2P_{3/2}$ state is metastable, $\tau \cong 0.3 \text{ s}$; its hyperfine structure is irrelevant to the experiment. The $7^2S_{1/2}$ state decays with roughly equal probabilities via E1 transitions to the $6^2P_{1/2}$ and the $6^2P_{3/2}$ states.

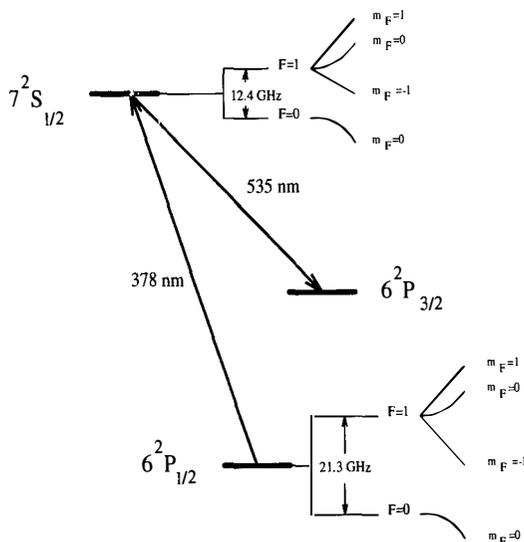


Figure 2: The low lying energy levels of thallium 205, $Z=81$. The figure includes the Zeeman splittings at the right. Isotope 205 is used because it is 70% abundant.

We pump the $6^2P_{1/2} F = 1 \rightarrow 7^2S_{1/2} F = 1$ transition, cycling it hundreds of times. We

collect and count the $7^2S_{1/2} \rightarrow 6^2P_{3/2}$, 535 nm light. The Clebsch-Gordan coefficient between the $m_F = 0$ states vanishes, and, therefore, the pumping has three effects: 1) depopulates the $6^2P_{1/2} F = 1, m_F = \pm 1$ states, 2) populates the $6^2P_{1/2} F = 1, m_F = 0$ state, and 3) measures the populations of the $6^2P_{1/2} F = 1, m_F = \pm 1$ states.

Keeping the effects of the optical pumping in mind, we can follow a thallium atom up through the apparatus as it moves from the lower oven to the final detection region. First, the atom is optically pumped into the $6^2P_{1/2} F = 1, m_F = 0$ state. Then, it is pulsed with rf tuned to the (M1) Zeeman transitions, depopulating the $m_F = 0$ state and populating the $m_F = \pm 1$ states equally. The atom passes through the electric field region; if there is a non-zero EDM, the $m_F = \pm 1$ states acquire a phase of opposite sign and the magnitude:

$$\epsilon_0 = \frac{-d_e |e| R E L}{\hbar v} \tag{1}$$

v is the atomic velocity, 4.4×10^4 cm/s, E is the electric field strength, 120 kV/cm, R is the enhancement factor, -585, e is the unit charge in cgs, and L is the length of the electric field region, 100 cm. For an EDM of 10^{-27} e cm, $\epsilon_0 \cong 10^{-7}$. Note that this phase changes sign with E . Like all Ramsey interference experiments, the atoms are then pulsed again with rf, phase different from the first interaction region, which brings the EDM phase down from the exponent. We then reprobe the atoms with the same optical pumping light in the second optical interaction region. This extracts a signal proportional to the populations of the $m_F = \pm 1$ transitions, and, therefore, the phase, ϵ_0 . Reversing the electric field and constructing the signal (S) asymmetry:

$$A = \frac{S_{+E} - S_{-E}}{S_{+E} + S_{-E}} \tag{2}$$

constructs a number proportional to the electron EDM. If we scan the rf through the transition, we see the well known Ramsey fringes. These are shown in Fig. 3.

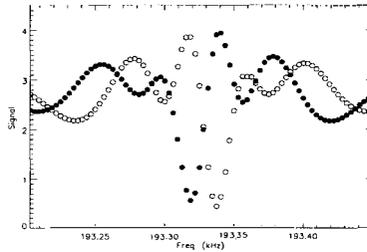


Figure 3: Two measured double loop resonances, central 250 Hz. The phase difference between the rf regions is, $\pm 3\pi/4$.

When we reverse the sign of the electric field, we move off the center of this fringe by an amount equal to the about 10^{-4} Hz for an EDM of 10^{-27} e cm . Because the central fringe has a large slope, the number of 535 nm photons changes with the electric field reversal.

3 Systematic Uncertainties

Critical to the success of this experiment is the control and understanding of its systematics. In general, any non-random phase which rotates the $m_F = \pm 1$ components of the wave function in opposite directions, that is not due to a real atomic EDM, is a systematic. This unwanted phase is most likely due to the $-\boldsymbol{\mu} \cdot \mathbf{B}$ interaction, it can also be topological in origin. To remove the unwanted phase, we reduce it manually with the apparatus, or make certain by calculation that it is so small that it isn't mimicking a true EDM. We discuss the dominant contributions only.

3.1 The ' $\mathbf{E} \times \mathbf{v}$ ' Effect

Our atomic EDM measurement is severely complicated by the large magnetic moment of the $6^2P_{1/2} F = 1, m_F = \pm 1$ states: $\mu_B/3$ in magnitude. Any magnetic field which changes synchronously with the electric field could create a false EDM; in particular, the magnetic field includes relativistic terms from the thallium atom's passage through the intense electric field, $\mathbf{B} = \mathbf{E} \times \mathbf{v}/c$.

A detailed calculations shows that the magnitude of this effect is:

$$d_{false} = 7.9 \times 10^{-21} \left(\frac{\delta \bar{v}_y}{|\bar{v}_x|} \left\langle \frac{B_x}{|B_z|} \right\rangle + \frac{\delta B_y}{\langle |B_z| \rangle} + \frac{\delta E_y}{\langle E_z \rangle} \frac{B_z}{|B_z|} \right) e \text{ cm} \quad (3)$$

The δ 's indicate up/down atomic beam difference. Note that if the beams overlap exactly and have the same velocities, the uncertainty vanishes. This is our motivation for using two opposing atomic beams. We are able to estimate and minimize the contribution from each of these terms. Note that strict atomic beam overlap is essential for the experiment's success; if the apparatus were not vertical, the orbits would be affected by gravity and such good overlap impossible.

We have recently noticed that there could be a contribution to the $\mathbf{E} \times \mathbf{v}$ effect from a difference in atomic beam overlap as a function of \hat{x} if a large electric gradient is present. However, this must be investigated further.

3.2 The Geometric Phase Effect

If the \hat{x} direction of magnetic field is different at the two ends of the magnetic field, then there is an additional systematic uncertainty that is topological in origin. It is described in detail elsewhere.⁴⁾ The $m_F \pm 1$ components of the wavefunction will accumulate a phase of opposite sign and magnitude:

$$\Omega = 2v \frac{(B_{x1} - B_{x2})}{cB_z^2} E \quad (4)$$

where 1 and 2 refer to the two ends of the electric field region.

If the \hat{x} gradient is due to the \hat{z} coils, the systematic is B -even (unlike a true EDM) and varies like B_z^{-1} . If the gradient is from an external source, then the effect is B -odd but scales like B_z^{-2} . We know the magnitude and the sign of the effect at a variety of magnetic field magnitudes, and may be able to subtract it from our result.

3.3 Charging and Leakage Currents

Charging currents may generate an EDM-like asymmetry if they have not stopped flowing when data taking commences, so we allow a generous delay before taking data.

Leakage currents are a very difficult problem. We cannot calculate the effect of them, and they can be very hard to remove. During runs, we measure that they were less than 10 nA over the entire 500 cm² surface of the plates and recondition the plates in an argon atmosphere during the running.

3.4 Systematics Summary

The following is a summary of the two major systematics of our experiment. The important experimental contributions are listed, followed by their values and a conversion coefficient, and the final uncertainty. These are the expected values; they might be larger or smaller during the final data runs. Note the experimental values of the magnetic gradients and the up/down atomic beam spacial overlap.

' $\mathbf{E} \times \mathbf{v}$ '

$$\begin{aligned} \frac{\delta v_y}{|v_x|} \left\langle \frac{B_x}{|B_z|} \right\rangle &= 7.9 \times 10^{-21} \cdot 8 \times 10^{-5} \cdot \frac{1.7 \times 10^{-4}}{0.409} \\ &< 2 \times 10^{-28} \text{ e cm} \\ \frac{\delta B_y}{(|B_x|)} &= 7.9 \times 10^{-21} \cdot \left\{ \frac{\delta z \frac{\partial B_y}{\partial z} + \delta y \frac{\partial B_y}{\partial y}}{|B_z|} \right\} < 7.9 \times 10^{-21} \cdot \left\{ \frac{10^{-4} \cdot 6 \times 10^{-6} + 10^{-3} \cdot 6 \times 10^{-6}}{0.409} \right\} \\ &< 2 \times 10^{-28} \text{ e cm} \\ \frac{\delta E_y}{(E_z)} \frac{B_x}{|B_z|} &= 7.9 \times 10^{-21} \cdot \frac{1}{E_z} \cdot \left\{ \frac{\partial E_y}{\partial z} \delta z \right\} < 7.9 \times 10^{-21} \cdot \frac{1}{400} \cdot 0.36 \cdot 10^{-4} \end{aligned}$$

$$< 7 \times 10^{-28} e \text{ cm}$$

(This uncertainty is removed completely by reversing the electric field plates.)

$$\frac{1}{2c^2} \frac{v_z^2 E^2}{|B_z|}$$

$$< 7.6 \times 10^{-28} e \text{ cm}$$

(This uncertainty is \mathbf{B} – even and cannot mimic a true EDM.)

Geometric Phase

$$\frac{2v_z(B_{x1}-B_{x2})E}{cB_z^2}$$

$$= 1.4 \times 10^{-23} (B_{x1} - B_{x2})$$

$$< 3 \times 10^{-27} e \text{ cm at } B_z = 0.17 \text{ G}$$

$$< 7 \times 10^{-28} e \text{ cm at } B_z = 0.4 \text{ G}$$

Total Systematic Uncertainty: $7 \times 10^{-28} e \text{ cm}$

If we measure the geometric phase systematic well, this uncertainty could drop to $4 \times 10^{-28} e \text{ cm}$.

4 Counting Time

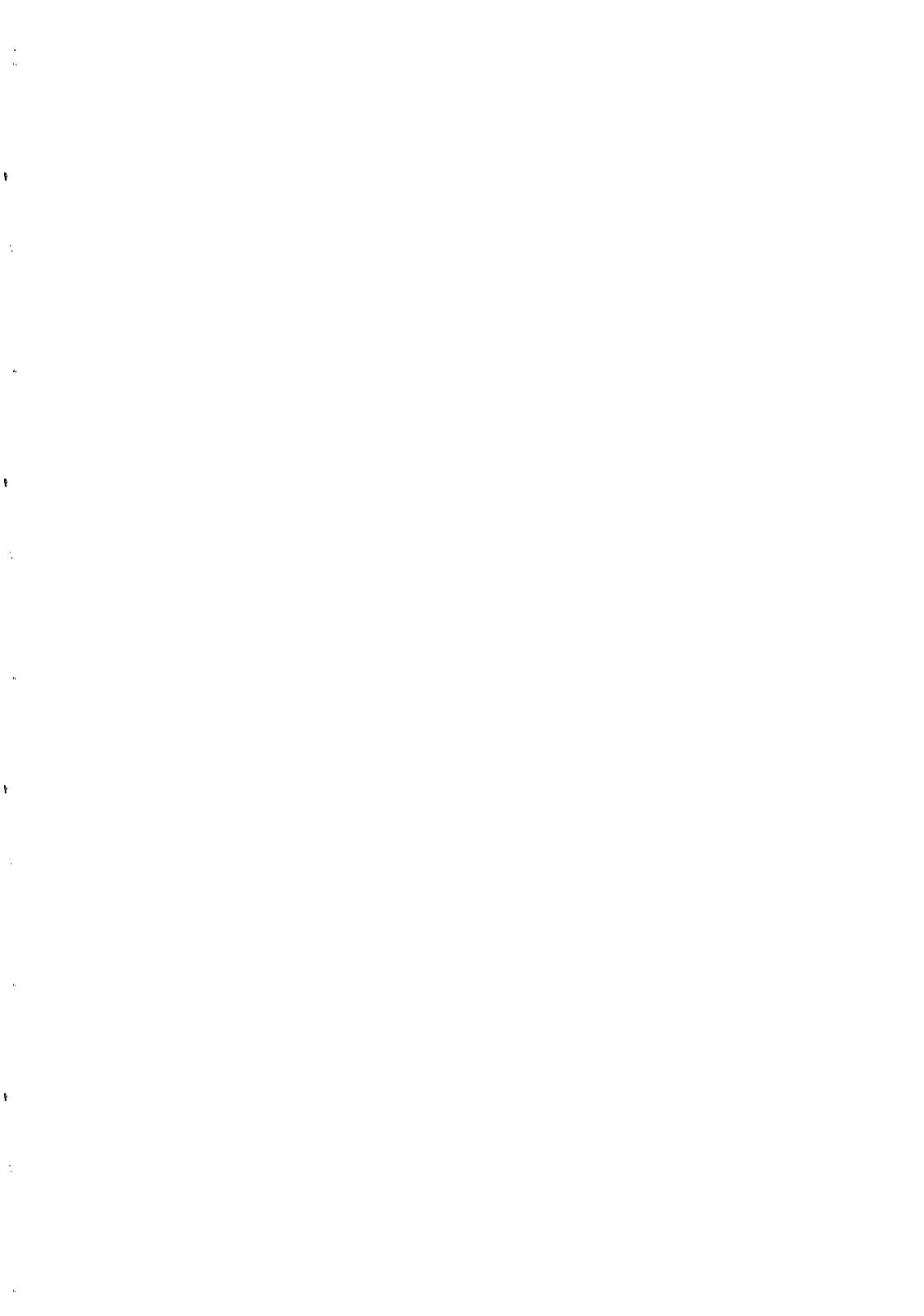
From shot noise considerations alone, we expect a mere 7.5 hours to count to $10^{-27} e \text{ cm}$. However, there is additional noise which limits us to roughly $5.3 \times 10^{-25} e \text{ cm}$ per second, or 80 hours to achieve $10^{-27} e \text{ cm}$.

5 Conclusion

We have designed an experiment to measure the electron electric dipole moment that should be sensitive to a moment in the neighborhood of $10^{-27} e \text{ cm}$.

6 References

- 1) Bernreuther, W. and M. Suzuki, "The Electric Dipole Moment of the Electron", *Rev. of Mod. Phys.*, **63**, 1991, 313.
- 2) Murthy, S. A. et. al., "New limit on the electron electric dipole moment from cesium", *Phys. Rev. Lett.*, **63**, 1989, 965.
- 3) Abdullah, K. A. et. al., "New Experimental limit on the electron electric dipole moment", *Phys. Rev. Lett.*, **65**, 1990, 2347.
- 4) Commins, E. D., "Berry's geometric phase and motional fields", *Am. J. Phys.*, **59**, 1991, 1077



**The Search for a Permanent Electric Dipole Moment
Using Dual Spin Exchange Pumped ^{129}Xe and ^3He Zeeman Masers**

Timothy E. Chupp², Rohan J. Hoare², Eduardo R. Oteiza^{2,3}

Ronald Walsworth³, Jeffrey T. Wright², Bo Wu²

¹ *Harvard University, Cambridge, Massachusetts 02138 USA*

² *University of Michigan, Ann Arbor, Michigan 48109 USA*

³ *Smithsonian Astrophysical Observatory, Cambridge, Massachusetts 02138 USA*



Abstract

Time reversal and parity non-invariant interactions within an atom naturally give rise to an atomic permanent electric dipole moment (EDM). For noble gas atoms, the size of such an EDM scales as Z^2 and higher powers of Z depending on the actual manifestation of T non-invariance, most importantly a distribution of electric dipole moment within the nucleus (Schiff Moment) and a T-odd tensor interaction between the nucleus and atomic electrons. We have developed a noble gas Zeeman maser that can simultaneously measure the EDMs of ^{129}Xe and ^3He in the same maser cell in order to mitigate systematic effects due to leakage currents and common mode problems such as magnetic field and time base noise. The philosophy of our approach is that the EDM of ^3He is negligible compared to that of ^{129}Xe , and thus we use the ^3He as a “magnetometer” and monitor of systematic effects. The maser’s measured precision for an integration time τ has been observed to be $10^{-2}\tau^{-3/2}$, and we anticipate EDM sensitivity better than 10^{-27} e-cm for a one day run.

Introduction

The observation of an Permanent Electric Dipole Moment (EDM) would reveal interactions that are not invariant under time reversal (T) and parity (P) transformations.¹ Such interactions, described in the language of modern elementary particle physics, also violate invariance under the transformation of CP (C is the charge conjugation transformation). The question of the symmetry of physics under the CP and T transformations is a natural one and independent of the first observation of CP non-invariance in the neutral Kaon system nearly thirty years ago.² In atomic systems, the searches for an EDM are currently motivated most strongly by the opportunity to discover physics beyond the Standard Model. This is because the current Standard Model picture of elementary particle interactions can accommodate CP violation in the Kaon system but predicts EDMs much smaller than will be detectable in the near future in any system. For example, the neutron EDM searches currently set upper limits near 10^{-25} e-cm while the Standard Model predicts values near 10^{-32} e-cm. Thus we hope to discover physics beyond the Standard Model by measuring EDM's much larger than the Standard Model prediction or to clarify the theoretical picture by setting upper limits on CP non-invariant contributions which eliminate or constrain alternatives to the Standard Model. In fact several extensions of the standard model can accommodate electron and neutron EDMs of the order of current experiments' sensitivity.^{3,4} (We refer the reader to the review by S. Barr presented at this Workshop.)

In an atom, a detectable EDM arises because the atom is polarized by T non-invariant interactions. In general, heavier atoms are more polarizable and the size of the EDM of the atom grows as Z^2 and higher powers of Z .⁵ Thus the heaviest atoms, Xe, Cs, Tl and Hg have generally been the experimenters' first choice. However, certain systematic effects may be limiting current experiments' sensitivity to an atomic EDM in the range of 10^{-27} e-cm. The comparison of ^{129}Xe with $Z = 54$ and ^3He with $Z = 2$ sampling the same volume at the same time is designed to greatly reduce sensitivity to the systematic effects that are common mode including bulk magnetic fields, especially those created by leakage currents present when the electric field is applied and reversed. Other common mode noise sources such as time base drift are also reduced.

In Table 1. we list the several sources of an EDM in ^{129}Xe . These include 1) the possibility of an electron EDM (d_e), 2) T-noninvariant nuclear forces such as those inducing the neutron's EDM (d_n) or a nuclear distribution of electric dipole moment different from the distribution of electric charge leading to a finite Schiff moment (Q_S), and 3) T-noninvariant weak neutral currents between the electrons and nucleus which may have tensor-pseudotensor character (labeled C_T) or scalar-pseudoscalar character (labeled C_S). In terms of these sources, the atomic EDM of ^{129}Xe in units of e-cm may be written

$$d_{Xe} = 1 \times 10^{-3} d_e + 2.7 \times 10^{-18} \frac{Q_s}{fm^3} + 5.2 \times 10^{-21} C_T + 7.5 \times 10^{-23} C_S$$

In Table 1 we also show the Z dependence for each source of atomic EDM and the limits established by recent experiments. These current limits establish $d_{Xe} < 10^{-27}$ e-cm, which sets the scale for our measurement goals.

Table 1. Sources of atomic EDM and current limits

Source	Z dependence	Current Limit	system
Electron EDM	$Z^3 \alpha^2$ for $J \neq 0$	$d_e = -3 \pm 8 \times 10^{-27}$ e-cm	Tl- beam ⁶
T-nonconserving		$d_n = -3 \pm 5 \times 10^{-25}$ e-cm	neutron bottle ⁷
nuclear interactions	$Z^2 A^{2/3}$	$Q_S = -6 \pm 13 \times 10^{-12}$ e-fm ³	Hg - cell ⁸
T-nonconserving	Z^2	$C_T = -4 \pm 9 \times 10^{-9}$	Hg - cell ⁸
neutral currents	Z^3	$C_S = 2 \pm 6 \times 10^{-7}$	Hg - cell ⁸

The Free Induction Decay Experiment

Our initial approach was a Free Induction Decay (FID) technique that simultaneously measured the precession frequencies of ^{129}Xe and ^3He in a single cell in the presence of parallel or anti-parallel applied magnetic and electric fields. The apparatus is shown in figure 1. The ^{129}Xe and ^3He were contained in a cell which also contains a small amount of Rb and several hundred torr of N_2 . The Rb was optically pumped with light from a Ti-sapphire laser and polarization was transferred from the Rb to the noble gas nuclei by spin exchange.⁹ Spin exchange is mediated by a contact hyperfine interaction which is enhanced during a collision due to effects which are much stronger for the ^{129}Xe -Rb collisions than for the ^3He -Rb collisions.¹⁰ The cells consisted of glass cylinders about 0.8 cm long (made of Schott 8290) with glass end plates (Schott Tempax) and Molybdenum wire mesh electrodes sandwiched between the end plates and cylinder. The cylinder walls were coated with octadecyltrichlorosilane and the end caps and mesh were attached with epoxy. The cells must satisfy the following requirements: (1) they must hold off several kV; (2) the ^3He and ^{129}Xe relaxation times must be 10 hours and 0.3 hours respectively for T_1 and at least 0.3 hours for T_2 ; (3) they must remain mechanically intact and maintain the relaxation times during thermal cycling between 120 C and 80 C; (4) all materials must be inert to Rb in the operating temperature range. Detailed information on cell construction can be found in reference 8.

The laser polarized the Rb and the noble gas nuclei parallel (or anti-parallel) to \vec{E}_{applied} . We began a run cycle by pumping for several hours at a temperature of 120C in order to build up the ^3He polarization to about 10%. Then the cell was cooled to 80C, a temperature at which the ^3He polarization relaxation time was about 10 hours and the ^{129}Xe relaxation time was about 0.3 hours (1000 seconds). A measurement was initiated by applying resonant oscillating magnetic field pulses simultaneously to both species (about 3 kHz and 10 kHz

(no magnet control) and closed (magnetic field locked to maintain constant ^3He frequency). With the loop open, the frequency modulation is 10 mHz and with the loop closed, a limit of $10 \mu\text{Hz}$ demonstrated a factor of greater than 1000 rejection of magnetic field variations. If we use the typical leakage current of 1 pA and assume this current completes one full loop around our cell, the resulting magnetic field would produce a false EDM frequency shift corresponding to an EDM of $\approx 10^{-28}$ e-cm. With the 1000-fold rejection provided by ^3He magnetometry, this false EDM signal is reduced to 10^{-31} e-cm! This clearly demonstrates the importance of our two species approach.

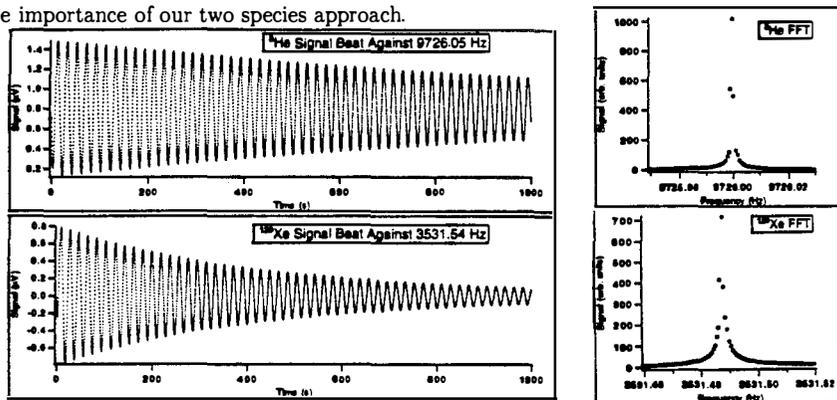


Figure 2. Beat signals for ^3He and ^{129}Xe in the time and frequency domains. The magnetic field is controlled to maintain the ^3He precession frequency at 9726 Hz.

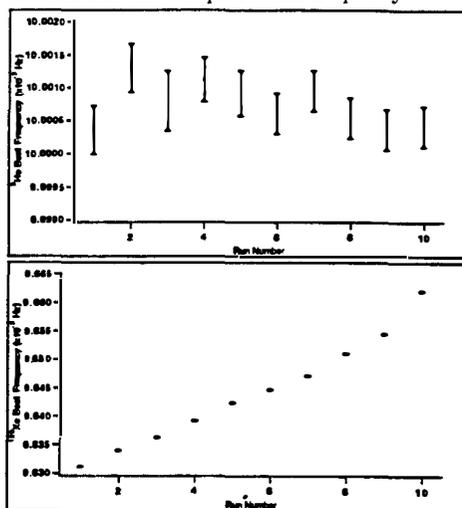


Figure 3. Extracted ^3He and ^{129}Xe frequencies for a cycle of 12 runs.

Frequency Shifts

Frequency shifts of the ^{129}Xe and ^3He limit the sensitivity to an EDM. Two sorts of shifts must be considered: shifts that correlate with the electric field reversal and thus appear as a false EDM signal and shifts that are not correlated, but are not stable and therefore produce frequency noise that limits the precision of our measurement. The magnetic field is locked to maintain the ^3He free precession frequency constant and therefore any frequency shifts of either the ^3He or ^{129}Xe appear in the extracted ^{129}Xe frequency as

$$(\vec{A} \cdot \hat{B})(k_{Xe} - \frac{\mu_{Xe}}{\mu_{He}} k_{He})$$

where \vec{A} represents any field which couples to the noble gas spin with strength k . (Both the EDM coupling and leakage current effects have this form, however for a leakage current, $k_{Xe}/k_{He} = \mu_{Xe}/\mu_{He}$, and the shift vanishes.)

The cases of $k_{Xe}/k_{He} \neq \mu_{Xe}/\mu_{He}$, in order of decreasing importance are: 1) The longitudinal magnetizations of the ^{129}Xe and ^3He produce frequency shifts estimated to be as large as 250 μHz for the ^{129}Xe and 30 μHz for the ^3He ; 2) Rb is polarized by spin exchange with ^{129}Xe and ^3He and the hyperfine contact interaction leads to shifts $\leq 10 \mu\text{Hz}$; 3) Any rotating magnetic field, B_T with frequency ω_T causes a shift proportional to $(\mu_{Xe} B_T)/(\omega_{Xe} - \omega_T)$ for ^{129}Xe and similarly for ^3He . This is a generalization of the Bloch-Siegert shift for which $\omega_T = -\omega_{Xe}$. These fields include the oscillating Rb quenching fields and the rotating fields due to the precessing ^{129}Xe and ^3He magnetization. The estimated shifts are about 0.2 μHz and $\ll 1$ nHz respectively for the two sources; 4) The wall interactions of ^{129}Xe and ^3He are temperature dependent, and we have determined that the shifts are less than 0.1 μHz for a temperature change of 0.1 K; 5) Gradients of static and rotating magnetic fields also induce shifts, which we estimate to be less than about 10 nHz. None of these sources of frequency shift lead to a false EDM signal, however they do introduce frequency noise. For the FID technique this severely limits the EDM measurement because the shifts dependent on the ^3He or ^{129}Xe magnetization change abruptly every time a run is initiated. These shifts are effectively mitigated with the dual noble gas maser in which the population inversion is produced by spin exchange. The maser technique provides greater signal to noise and CW stability of the ^3He and ^{129}Xe magnetizations.

The Spin Exchange Pumped Noble Gas Maser

A noble gas Zeeman maser can be effected by spin exchange pumping a population inversion and coupling the precessing coherence to a resonant cavity, or in our case a tank circuit formed with the pick-up coils. In a two cell maser, illustrated in figure 4, conditions for both species can be optimized. A pumping cell (labeled V_P) and a measurement or maser

cell (labeled V_M) are connected by a transfer tube¹³ The precessing magnetization induces a current in the pick-up coil and therefore a resonant oscillating magnetic field, which produces a torque on the spins that changes the magnetization. This self-coupling can lead to stimulated or super-radiant Rabi precession of the spins and to steady state maser oscillations with stable longitudinal and transverse magnetizations. Spins in the pumping cell do not take part in the maser oscillation because the transverse component of magnetization is rapidly damped by spin exchange and by the magnetic field gradients.

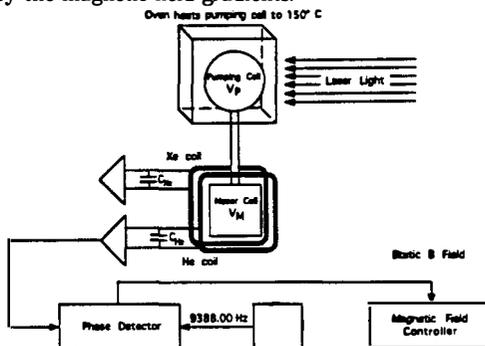


Figure 4. The spin exchange pumped noble gas maser apparatus. Data acquisition and electronics are similar to the FID set-up shown in Figure 1.

The maser operation can be understood by writing the equations of motion for the noble gas polarization (\vec{P}) in the frame rotating at the noble gas precession frequency¹⁴ and assuming the tank-circuit resonant frequency is equal to the precession frequency. In this frame the magnetization is directed along the x -axis. Neglecting the high frequency (*e.g.* Bloch-Siegert) components of the induced magnetic field

$$\dot{P}_x = P_x \left(\frac{P_z}{P_0 T_{RD}} - \frac{1}{T_2} \right)$$

$$\dot{P}_z = P_P G_M - \frac{P_x^2}{P_0 T_{RD}} - P_z \left(\frac{1}{T_1} + G_M \right)$$

where T_1 and T_2 are the phenomenological relaxation times for P_z and P_x , P_P is the pump cell polarization, G_M is a rate of polarization transfer to the maser cell¹³, P_0 is the equilibrium value of P_z in the absence of spin-cavity coupling, and $P_x \leq P_0$. $1/T_{RD}$, often called the radiation damping rate, is a measure of the torque that is self-induced by the spin-coil coupling:

$$\frac{1}{T_{RD}} = \frac{1}{2} \gamma_{ng}^2 \eta Q \mu_0 \frac{\hbar}{2} [ng] P_0.$$

η is a geometric factor that accounts for the difference in volume of the pick-up coil and the magnetized sample, and $Q = \omega_0 L/R$ is the quality factor of the pick-up coil tank circuit.

For a single cell maser, equations (3) can be modified by setting $P_P = P_{Rb}$ (the Rb polarization) and $G_M = \gamma_{SE}$ (the spin exchange rate)⁹. The solution of these coupled equations depends on initial conditions. The conditions for steady state maser oscillation, found by setting $\dot{P}_x = \dot{P}_z = \dot{P}_P = 0$, are (1) $P_z > 0$ (*i.e.* population inversion), (2) $\tau_{RD} < T_2$, and (3) $P_x > 0$. The last condition (3) is required for the maser to start and can be induced by spontaneous emission, black body radiation, or with a pulse of applied oscillating magnetic field.

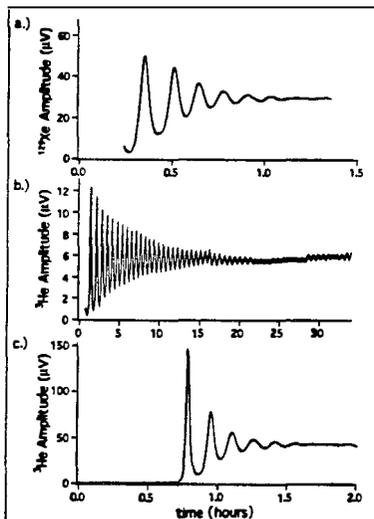


Figure 5. Observed start-up transients and approach to steady state for the single cell ¹²⁹Xe (a.) and the ³He (b.) masers and the two cell ³He maser (c.).

In figure 5 we show the measured maser signal (proportional to M_T) for the start up and approach to steady state of separate single cell masers using ¹²⁹Xe (a.) and ³He (b.) and for a two cell ³He maser (c.). The optical pumping was initiated at $t \approx 0$. The single cell is spherical in shape with volume approximately 1 cm³ and is coated with a siliconizing agent which increases the wall relaxation time for ¹²⁹Xe¹¹. The double cells, each spherical with volume ≈ 3 cm³ are uncoated. The T_2 's for the maser are limited by the polarization transfer times $1/G_M$ and are typically much shorter than for free induction decay. The illustrated transient behavior can be understood as follows: spin exchange produces polarization P_x , which builds up and increases the radiation damping torque. When that torque exceeds the threshold given by $\tau_{RD} < (P_x/P_0)T_2$, P_x can increase as the polarization rotates. Thus P_z decreases, decreasing the radiation damping torque until a temporary equilibrium is attained from which the process repeats. Eventually, the transients decay, and steady state oscillation proceeds. One interesting feature of these transients is the damping time. For the single

cell maser, the rate at which the transients damp out is determined by γ_{SE} which is several minutes for ^{129}Xe and several hours for ^3He . For the two cell ^3He maser, the transients damp out with a rate characteristic of the spin transfer rate $G_M \approx 1/(3\text{min})$.

The precision of the maser for the EDM search is illustrated in an experiment that uses the maser precession frequency to lock the magnetic field while monitoring the maser oscillation frequency. In this experiment, the maser serves as a phase locked local oscillator which is compared to a master clock as shown in figure 4. This measurement is not sensitive to the intrinsic frequency stability of the maser for observation times long compared to the time constants of the phase locked loop. The frequency was measured by beating the maser signals against a reference oscillator shifted from the lock frequency by 0.05 Hz (*i.e.* for the single cell ^3He maser, the reference oscillator is at 9388.05 Hz). This beat signal was averaged with a 1.25 s time constant and sampled at 1 Hz.

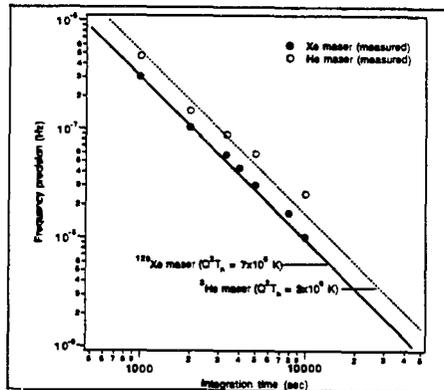


Figure 6. Frequency precision of the single cell ^3He and ^{129}Xe masers. In both cases, the maser oscillation frequency is locked to an external oscillator by feed back to the static magnetic field.

We extract the best fit frequency, which can be thought of as a linear fit of the phase as a function of time. In our case, the phase is determined by the measured signal voltage with the relationship $v(t_i) = v_0 \sin(\phi(t_i))$, where v_0 is the amplitude of the measured voltage. The expected optimal maser frequency performance, the precision for a given measurement interval τ , is limited by the error on the slope of ϕ as a function of time. For our measurements, with τ between 1,000 and 10,000 seconds, this is dominated by the white voltage noise from the pick-up coil at temperature T and noise from the electronics which cause white phase noise of uniform density. Fluctuations of the maser frequency on time scales short compared to

the inverse of the 0.05 Hz beat frequency also appear as phase noise and cannot be separated from the voltage noise. For white phase noise, the error on the measured frequency is

$$\sigma_\nu = \frac{1}{2\pi} \sqrt{\frac{12 \phi_{rms}^2 / B}{\pi \tau^3}} \propto \tau^{-3/2}$$

where we assume that B is matched to the sample rate so that the number of samples is $N = 2\pi B\tau$. In figure 6 we show the results of our measurements for the single cell ^3He and ^{129}Xe masers. Each data point represents the standard deviation of a series of measurements for each measurement interval τ . The data demonstrate that $\sigma_\nu \propto \tau^{-3/2}$ (solid and dashed lines) as expected for the phase locked maser. The intrinsic frequency noise of the maser includes unstable collisional and wall induced shifts and imperfect operation of the magnetic field lock loop. White frequency noise causes a random walk of the phase that accumulates as $\phi_{rms}^2 \propto \tau^{15}$, and for τ sufficiently large so that white frequency noise dominates, we would expect $\sigma_\nu \propto \tau^{-1/2}$. The best strategy is to choose the electric field flipping rate to limit τ so that white phase noise dominates, which we anticipate to be $\tau \approx 10,000\text{sec}$ and therefore in one day of data taking, frequency precision of 3×10^{-9} Hz is anticipated. For a 5 kV/cm electric field, this would correspond to EDM sensitivity of 1×10^{-27} e-cm.

Conclusion

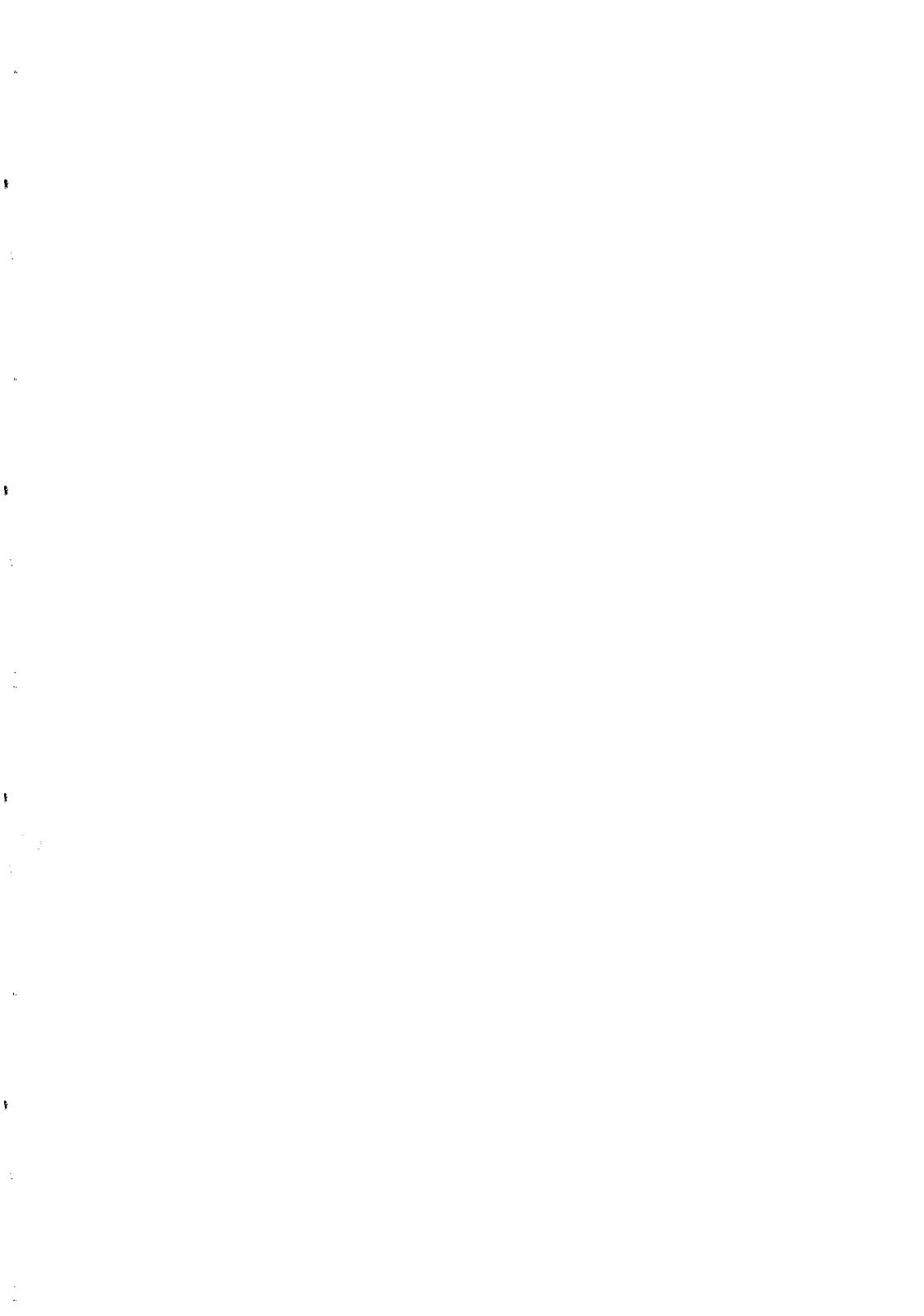
Measurements with two species are expected to be crucial for future EDM measurements. We have developed extremely precise free induction decay techniques with ^3He and ^{129}Xe have the limitation of frequency shifts due, most importantly, to changing noble gas magnetization during the pulse and FID cycles. The recently developed spin exchange pumped masers, which have somewhat better signal to noise characteristics and the important advantage of CW operation. The performance achieved to date for the single cell, single species, phase locked ^{129}Xe maser can be expressed as $\sigma_\nu < 10^{-2} \tau^{-3/2}$ Hz for measurement intervals $\tau = 1,000$ to $10,000$ seconds. We anticipate significant improvement of this performance in the optimized two cell configuration. The two cell arrangement also eliminates the frequency shift due to polarized Rb, the need for quenching RF, and the maser cell is cooled so that higher voltage can be applied. Also note that, since maser performance is currently dominated by pick-up coil Johnson noise and other sources of white voltage noise, a DC squid detecting the precessing magnetization could also lead to improved performance. Operation with other noble gas species, including ^{21}Ne , is straight forward, providing opportunities for a wide array of applications.

Acknowledgements

The authors are grateful for the efforts of and discussions with Mark Wagshul, Miriam Vu, and Reginald Jaynes. This work was supported by the University of Michigan, the Smithsonian Institution's Scholarly Studies Program, and a NIST Precision Measurement Grant.

References

1. R.G. Sachs, **The Physics of Time Reversal**, Univ. of Chicago Press, Chicago, (1987).
2. J.H. Christenson, J.W. Cronin, V.L. Fitch, R. Turlay, *Phys. Rev. Lett.* **13** 138 (1964).
3. S.M. Barr and W.J. Marciano, in **CP Violation**, edited by C. Jarlskog, World Scientific, Singapore (1989).
4. L.R. Hunter, in **Atomic Physics 12**, edited by J. Zorn and R. Lewis, AIP, New York, 429 (1991).
5. P.G.H. Sandars, in **Atomic Physics 9**, edited by R.S. van Dyck and E.N. Fortson, World Scientific, New Jersey, (1985).
6. K. Abdullah, C. Carlberg, E.D. Commins, H. Gould, S. Ross, , *Phys. Rev. Lett.* **65** 2347 (1990).
7. K. Smith, *et al.*, *Phys. Lett.* **B234**, 191 (1990). I.S. Altarev *et al* *JETP Lett* **44** 460 (1986).
8. J.P. Jacobs, W.M. Klipstein, S.K Lamoreaux, B.R. Heckel, E.N. Fortson, *Phys. Rev. Lett.* **71**, 3782 (1993).
9. T.E. Chupp, M.E. Wagshul, K.P. Coulter, A.B. McDonald, W. Happer, *Phys. Rev.* **C36**, 2244 (1987).
10. X. Zeng, A. Wu, T. Call, E. Miron, D. Schreiber, W. Happer, *Phys. Rev.* **A31**, 260 (1985).
11. E.R. Oteiza, PhD Dissertation, Harvard University (1992).
12. T.E. Chupp, E.R. Oteiza, J.M. Richardson, T.R. White, *Phys. Rev.* **A38**, 3998 (1988); T.E. Chupp, R.J. Hoare, R.A. Loveman, E.R. Oteiza, J.M. Richardson, M.E. Wagshul, A.K. Thompson, *Phys. Rev. Lett.* **63**, 1541 (1989).
13. T.E. Chupp, R.A. Loveman, A.K. Thompson, A.M. Bernstein, D.R. Tieger, *Phys. Rev.* **C45**, 915 (1992)
14. T.E. Chupp, R.J. Hoare, B. Wu, R.L. Walsworth, *Phys. Rev. Letters* **72**, 2363 (1994).
15. D.W. Allen, *Proc. IEEE* **54**, 221 (1966).



CP, T, AND CPT VIOLATION IN THE NEUTRAL KAON SYSTEM AT THE CPLEAR EXPERIMENT

The CPLEAR Collaboration

R. Adler², T. Alhalel², A. Angelopoulos¹, A. Apostolakis¹, E. Aslanides¹¹, G. Backenstoss², C.P. Bee¹¹, O. Behnke¹⁷, J. Bennet⁹, V. Bertin¹¹, F. Blanc^{7,13}, P. Bloch⁴, Ch. Bula¹³, P. Carlson¹⁵, M. Carroll⁹, J. Carvalho⁵, E. Cawley⁹, S. Charalambous¹⁶, M. Chardalas¹⁶, G. Chardin¹⁴, M.B. Chertok³, M. Danielsson¹⁵, A. Cody⁹, S. Dedoussis¹⁶, M. DeJardin⁴, J. Derre¹⁴, M. Dodgson⁹, J. Duclos¹⁴, A. Ealet¹¹, B. Eckart², C. Eleftheriadis¹⁶, I. Evangelou⁸, L. Faravel^{7,11}, P. Fassnacht¹¹, J.L. Faure¹⁴, C. Felder², R. Ferreira-Marques⁵, W. Fetscher¹⁷, M. Fidecaro⁴, A. Filipčić¹⁰, D. Francis³, J. Fry⁹, E. Gabathuler⁹, R. Garnet⁹, D. Garreta¹⁴, T. Geralis¹³, H.-J. Gerber¹⁷, A. Go³, P. Gumplinger¹⁷, C. Guyot¹⁴, A. Haselden⁹, P.J. Hayman⁹, F. Henry-Couannier¹¹, R.W. Hollander⁶, E. Hubert¹¹, K. Jansson¹⁵, H.U. Jhoner⁷, K. Jon-And¹⁵, P.R. Kettle¹³, C. Kochowski¹⁴, P. Kokkas⁸, R. Kreuger⁶, T. Lawry³, R. Le Gac¹¹, F. Leimgruber², A. Liolios¹⁶, E. Machado⁵, P. Maley⁹, I. Mandić¹⁰, N. Manthos⁸, G. Marel¹⁴, M. Mikuz¹⁰, J. Miller³, F. Montanet¹¹, T. Nakada¹³, A. Onofre⁵, B. Pagels¹⁷, P. Pavlopoulos², F. Pelucchi¹¹, J. Pinto da Cunha⁵, A. Policarpo⁵, G. Polivka², H. Postma⁶, R. Rickenbach², B.L. Roberts³, E. Rozaki¹, T. Ruf⁴, L. Sacks⁹, L. Sakeliou¹, P. Sanders⁹, C. Santoni², K. Sarigiannis¹, M. Schäfer¹⁷, L.A. Schaller⁷, A. Schopper⁴, P. Schune¹⁴, A. Soares¹⁴, L. Tauscher², C. Thibault¹², F. Touchard⁴, C. Touramanis⁹, F. Triantis⁸, D.A. Tröster², E. Van Beveren⁵, C.W.E. Van Eijk⁶, S. Vlachos², P. Weber¹⁷, O. Wigger¹³, C. Witzig¹⁷, M. Wolter¹⁷, C. Yeche¹⁴, D. Zavrtanik¹⁰ and D. Zimmerman³.

¹University of Athens, ²University of Basle, ³Boston University, ⁴CERN, ⁵LIP and University of Coimbra, ⁶Delft University of Technology, ⁷University of Fribourg, ⁸University of Ioannina, ⁹University of Liverpool, ¹⁰J. Stefan Inst. and Phys. Dep. University of Ljubljana, ¹¹CPPM, IN2P3-CNRS et Université d'Aix-Marseille II, ¹²CSNSM, IN2P3-CNRS, ¹³Paul-Scherrer-Institut (PSI), ¹⁴DAPNIA/SPP CE Saclay, ¹⁵KTH Stockholm, ¹⁶University of Thessaloniki, ¹⁷ETH-ITP Zürich.

Presented by D. Garreta
DAPNIA/SPP, CE Saclay, Gif-sur-Yvette, France

Abstract

The essential characteristics of the neutral kaon system and the way CP, T, and possible CPT violations may be observed in it are recalled. The principle of the CPLEAR experiment is presented. CPLEAR experimental results in the semi-leptonic decay channels are given and discussed. It is shown, in particular, that direct time reversal invariance violation will be experimentally observed for the first time.

1 INTRODUCTION

CP violation was first observed in the long-lived neutral kaon decaying¹⁾ into $\pi^+\pi^-$. We will first recall the essential characteristics of the neutral kaon system. Then we will show how CP, T, and CPT symmetries come in and how their eventual violation can be observed. The principle of the CPLEAR experiment will be explained²⁾. Finally CPLEAR experimental results in the semi-leptonic decay channels will be given and discussed. We will show, in particular, that in the CPLEAR experiment, for the first time, direct T violation will be experimentally observed.

2 NEUTRAL KAON SYSTEM

We first describe the neutral kaon system assuming that CP symmetry is valid.

Neutral kaons are produced under the form $|K^0\rangle = |\bar{s}d\rangle$, $S = +1$, $|\bar{K}^0\rangle = |s\bar{d}\rangle$, $S = -1$ with a well defined strangeness S .

Once strong interaction cannot do anything more to the neutral kaon, weak interaction has time to induce, among other things, $\Delta S = 2$ transitions that produce $K^0 \longleftrightarrow \bar{K}^0$ oscillations (see Fig. 1).

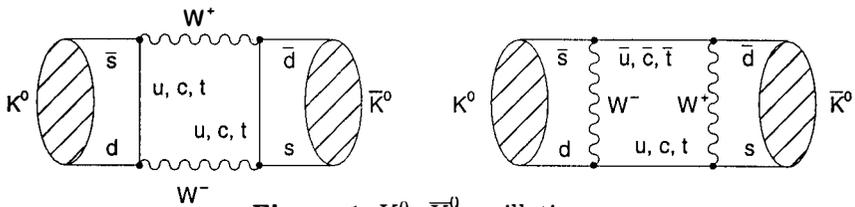


Figure 1: K^0, \bar{K}^0 oscillations

The two neutral kaon physical states are K_S and K_L with very different lifetimes, due to phase-space conditions. By looking at their decay channels (see Table 1), they can be shown to be the symmetric and antisymmetric combinations of K^0 and \bar{K}^0 , that is the CP eigenstates K_1 and K_2 with eigenvalues $+1$ and -1 , $|K_1\rangle = (|K^0\rangle + |\bar{K}^0\rangle)/\sqrt{2}$, $|K_2\rangle = (|K^0\rangle - |\bar{K}^0\rangle)/\sqrt{2}$.

Table 1 Neutral kaon main decay channels

$ K_S\rangle$	$\rightarrow \pi^+\pi^-$ 69%	} CP = +1	$M_S \simeq 498$ MeV $\tau_S = 1/\Gamma_S \simeq 0.09$ ns
	$\rightarrow \pi^0\pi^0$ 31%		
$ K_L\rangle$	$\rightarrow \pi^+\pi^-\pi^0$ 12.4%	} CP = -1 ($\ell = 0$)	$M_L \simeq 498$ MeV $\tau_L = 1/\Gamma_L \simeq 52$ ns
	$\rightarrow \pi^0\pi^0\pi^0$ 21.6%		
	$\rightarrow \pi^+\mu^-\bar{\nu}_\mu\rangle, \pi^-\mu^+\nu_\mu\rangle$ 27%		
	$\rightarrow \pi^+e^-\bar{\nu}_e\rangle, \pi^-e^+\nu_e\rangle$ 39%		

Once a K^0 is produced, its K_S and K_L components evolve according to their masses and widths. The probabilities of finding a K^0 or a \bar{K}^0 at a time t after production can easily be calculated.

$$\mathcal{P}(K^0, t) = \frac{1}{4} \{ \exp[-\Gamma_S t] + \exp[-\Gamma_L t] \pm 2 \exp[-(\Gamma_S + \Gamma_L)t/2] \cos[\Delta m t] \}$$

$$\mathcal{P}(\bar{K}^0, t)$$

An interference term clearly shows up between the K_S and K_L decay terms. The essential characteristics of the neutral kaon system is that the $\Delta m = M_L - M_S$ term that drives this interference is about the same as the decay constant $\Gamma_S/2$. This means that this interference pattern has time to show up before it is killed by the decay. We will see that this effect is experimentally observable in the semi-leptonic decay charge asymmetry of a K^0 (or a \bar{K}^0) beam because of the $\Delta S = \Delta Q$ rule that relates directly the charge of the lepton to the strangeness of the kaon (see Fig. 2).

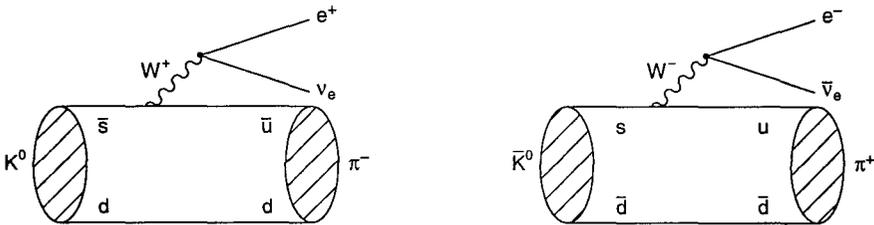


Figure 2: K^0, \bar{K}^0 semileptonic decay

3 CP, T, AND CPT VIOLATION

The fact that $2 \times 10^{-3} K_L$ also decayed into $\pi^+ \pi^{-1}$ showed that CP was violated somewhere. This can happen in two places:

$$\langle 2\pi | H | K_L \rangle = \begin{matrix} \text{CP} & +1 = +1 & +1 \neq -1 \\ \varepsilon_L \langle 2\pi | H | K_1 \rangle & + & \langle 2\pi | H | K_2 \rangle \end{matrix} \quad \begin{matrix} \text{(I)} \\ \text{(II)} \end{matrix}$$

- (I) In the production: K_S and K_L may no longer be the CP eigenstates K_1 and K_2 but there could be a small admixture of the other component. $|K_S\rangle = |K_1\rangle + \varepsilon_S |K_2\rangle$; $|K_L\rangle = \varepsilon_L |K_1\rangle + |K_2\rangle$; $|\varepsilon_S|, |\varepsilon_L| \ll 1$; $\varepsilon_S = \varepsilon + \delta$; $\varepsilon_L = \varepsilon - \delta$; ε and δ , sometimes referred to as ε_T and δ_{CPT} , can be shown³⁾ to correspond to T and CPT violations respectively.
- (II) In the decay: The main component K_2 of K_L can also decay directly into 2π through a CP violating transition (penguin diagram). This direct CP violating term is parametrized through a coefficient ε' , which is measured^{4,5)} to be at most a few 10^{-3} of ε .

3.1 CP-violating amplitude measurements

The CP-violating amplitude to a final CP eigenstate F is parametrized by a complex number η_F which is its ratio to the CP conserving amplitude. The decay rate of a $K^0(\bar{K}^0)$ beam into $F = 2\pi$ shows again an interference term between the allowed K_S decay and the CP forbidden K_L decay:

$$R(K^0, F; t) = C_{\mp}(e^{-\Gamma_S t} + |\eta_F|^2 e^{-\Gamma_L t} \pm 2|\eta_F| e^{-(\Gamma_S + \Gamma_L)t/2} \cos(\Delta m t - \phi_F))$$

$R(\bar{K}^0, F; t)$

where $C_{\mp} = \frac{1 \mp 2 \operatorname{Re}(\varepsilon)}{2} |\langle F | H | K_S \rangle|^2$.

A first method to get information about η_F is to compare the K_S and K_L decay rates. It gives $|\eta_F|^2$ and has been used by NA31⁴⁾ at CERN and E731⁵⁾ at Fermilab to measure ε'/ε .

A second method used by CPLEAR consists of measuring the decay rate asymmetry of K^0 and \bar{K}^0 in the interference region, which provides not only $|\eta_F|$ but also ϕ_F . It can be calculated that this asymmetry can go up to 60% at $t \sim 13.5 \tau_S$, which means that at this lifetime the rate of K^0 decay to $\pi^+\pi^-$ is four times larger than that of \bar{K}^0 . Starting from $|\eta_F| \sim 2 \times 10^{-3}$ the observable effect, the ratio of rates between initial particles and antiparticles, goes up to four, typical of interferometry where a small signal is strongly enhanced by interfering with a large one. In addition CPLEAR uses the K_S decay region to calibrate the respective K^0 and \bar{K}^0 detection efficiencies.

Coming back to the semi-leptonic decay charge asymmetry, in addition to the interference pattern due to $K^0 \leftrightarrow \bar{K}^0$ oscillations, CP violation adds, for large decay times, a residual asymmetry $2 \operatorname{Re}(\varepsilon)$ which has the same sign for K^0 and \bar{K}^0 .

This leads to an absolute definition of what we call matter and antimatter. In a world made of matter this residual asymmetry has the sign of the nuclei of the chemical elements; in a world made of antimatter it would be the opposite.

4 CPLEAR EXPERIMENT

The CPLEAR experiment uses the annihilation at rest of the low-energy antiproton beam of LEAR at CERN on a hydrogen target. It selects the annihilation channels leading to a neutral kaon and a pair of charged K and π mesons. By detecting and measuring the momenta of the charged mesons tagged neutral kaons with known momentum and initial strangeness are produced.

Two essential features of the CPLEAR experiment are the following: K^0 and \bar{K}^0 are produced with the same rates and they have low momenta (≤ 800 MeV/c), which gives a typical decay length of ~ 2.5 cm for the K_S so that the interference pattern is fully contained inside the detector. The simultaneous detection of the K^0 and the \bar{K}^0 together with their decay products to many channels allows for, by using rate asymmetries between \bar{K}^0 and K^0 , the cancellation of many systematic errors.

The decay channels studied are: 2π ; 3π ; semi-leptonic $\pi^+\ell^-\bar{\nu}_\ell$ and $\pi^-\ell^+\nu_\ell$.

5 CPLEAR EXPERIMENTAL RESULTS IN THE SEMI-LEPTONIC DECAY CHANNELS

Four decay rates are measured:

$$\begin{aligned} R^+ &= R(K^0 \Rightarrow \pi^-\ell^+\nu_\ell, t), & \bar{R}^- &= R(\bar{K}^0 \Rightarrow \pi^+\ell^-\bar{\nu}_\ell, t) \\ R^- &= R(K^0 \Rightarrow \pi^+\ell^-\bar{\nu}_\ell, t), & \bar{R}^+ &= R(\bar{K}^0 \Rightarrow \pi^-\ell^+\nu_\ell, t) \end{aligned}$$

The last two are directly forbidden by the $\Delta S = \Delta Q$ rule implying, to be non zero, that at least one $K^0 \longleftrightarrow \bar{K}^0$ transition has occurred.

A possible violation of this rule is characterized by the parameter:

$$x = \langle \pi^-\ell^+\nu_\ell | H | \bar{K}^0 \rangle / \langle \pi^-\ell^+\nu_\ell | H | K^0 \rangle.$$

From the four rates the following four asymmetries can be built:

$A_1(t) = \frac{(R^+ - R^-) - (\bar{R}^+ - \bar{R}^-)}{(R^+ + R^-) + (\bar{R}^+ + \bar{R}^-)}$ and $A_T(t) = \frac{\bar{R}^+ - R^-}{\bar{R}^+ + R^-} = 4 \text{Re}(\varepsilon_T)$ for $x = 0$, are shown in Fig. 3.

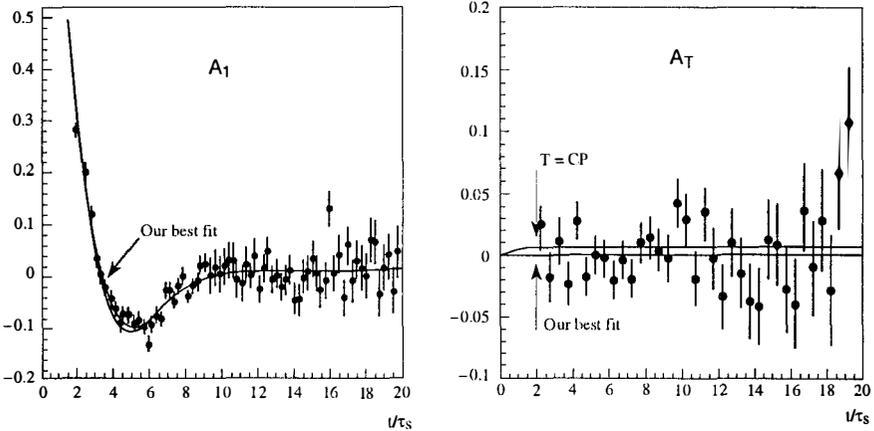


Figure 3: A_1 , A_T semi-leptonic asymmetries

- $A_1(t)$ is sensitive to Δm and $\text{Re}(x)$. Since it is roughly the charge asymmetry of a K^0 beam minus that of a \bar{K}^0 beam, the interference pattern adds up and is clearly visible and the residual asymmetry $2\text{Re}(\varepsilon)$ cancels out.
- $A_T(t)$ is expected to be $\sim 6 \times 10^{-3}$. If the $\Delta S = \Delta Q$ rule holds it can only be different from zero through a first $K^0 \iff \bar{K}^0$ step. Therefore it is a direct measurement of T violation.
- $A_2(t) = \frac{(\bar{R}^+ + \bar{R}^-) - (R^+ + R^-)}{(\bar{R}^+ + \bar{R}^-) + (R^+ + R^-)}$ is sensitive to $\text{Re}(\varepsilon_S)$ and $\text{Im}(x)$.
- $A_{\text{CPT}}(t) = \frac{\bar{R}^- - R^+}{\bar{R}^- + R^+} \rightarrow 4 \text{Re}(\delta)$ for $t \gg \tau_S$ is related to CPT violation.

6 CONCLUSION

Table 2 shows the statistical precision (1σ) of the parameters measured at CPLEAR together with the precision on these parameters quoted by the Particle Data Group (PDG)⁶⁾ before CPLEAR.

Table 2 Present and future performances of CPLEAR.

Parameter	PDG ('92)	'92	'93	'94 & '95
$\Delta m [10^{10} \hbar/s]$	0.0024	0.006	0.003	0.002
$\text{Re}(x) [10^{-3}]$	18	20	10	5
$\text{Im}(x) [10^{-3}]$	26	8	4	2
$\text{Re}(\varepsilon_S)[10^{-3}]$	-	1.4	0.7	0.4
$A_T[10^{-3}]$	-	4	2	1

It can be seen that at the completion of the experiment in 1995, taking into account systematic errors, the accuracy on $\text{Re}(x)$ and $\text{Im}(x)$ should be improved by factors of 3 and 10 respectively. Another striking feature is that a non-zero value of time reversal invariance violation should be directly measured at a four-standard deviation level.

References

- [1] J.H. Christenson, J.W. Cronin, V.L. Fitch and R. Turlay, Phys. Rev. Lett. **13** (1964) 138.
- [2] E.Gabathuler and P. Pavlopoulos, Proc. of the Workshop on Physics at LEAR with Low-Energy Cooled Antiprotons, ed. U.Gastaldi and R.Klapisch (Plenum, New York, 1982), p. 747.
- [3] J.S. Bell and J. Steinberger, Proc. Conf. on Elementary Particles, Oxford 1965 (Rutherford Lab., Chilton, Didcot, 1966), p. 195.
- [4] G.D.Barr et al., Phys. Lett. **B317** (1993) 233.
- [5] L.K.Gibbons et al., Phys. Rev. Lett. **70** (1993) 1203.
- [6] Particle Data Group, Phys. Rev. **D45** (1992).

RECENT RESULTS OF THE CPLEAR EXPERIMENT

The CPLEAR Collaboration

R. Adler², T. Alhalel², A. Angelopoulos¹, A. Apostolakis¹, E. Aslanides¹¹, G. Backenstoss², C.P. Bee¹¹, O. Behnke¹⁷, J. Bennet⁹, V. Bertin¹¹, F. Blanc^{7,13}, P. Bloch⁴, Ch. Bula¹³, P. Carlson¹⁵, M. Carroll⁹, J. Carvalho⁵, E. Cawley⁹, S. Charalambous¹⁶, M. Chardalas¹⁶, G. Chardin¹⁴, M.B. Chertok³, M. Danielsson¹⁵, A. Cody⁹, S. Dedoussis¹⁶, M. Dejardin⁴, J. Derre¹⁴, M. Dodgson⁹, J. Duclos¹⁴, A. Ealet¹¹, B. Eckart², C. Eleftheriadis¹⁶, I. Evangelou⁸, L. Faravel^{7,11}, P. Fassnacht¹¹, J.L. Faure¹⁴, C. Felder², R. Ferreira-Marques⁵, W. Fetscher¹⁷, M. Fidencaro⁴, A. Filipčić¹⁰, D. Francis³, J. Fry⁸, E. Gabathuler⁹, R. Gamet⁹, D. Garreta¹⁴, T. Gerasis¹³, H.-J. Gerber¹⁷, A. Go³, P. Gumplinger¹⁷, C. Guyot¹⁴, A. Haselden⁹, P.J. Hayman⁹, F. Henry-Couannier¹¹, R.W. Hollander⁶, E. Hubert¹¹, K. Jansson¹⁵, H.U. Johner⁷, K. Jon-And¹⁵, P.R. Kettle¹³, C. Kochowski¹⁴, P. Kokkas⁸, R. Kreuger⁶, T. Lawry³, R. Le Gac¹¹, F. Leimgruber², A. Liolios¹⁶, E. Machado⁵, P. Maley⁹, I. Mandić¹⁰, N. Manthos⁸, G. Marel¹⁴, M. Mikuž¹⁰, J. Miller³, F. Montanet¹¹, T. Nakada¹³, A. Onofre⁵, B. Pagels¹⁷, P. Pavlopoulos², F. Pelucchi¹¹, J. Pinto da Cunha⁵, A. Policarpo⁵, G. Polivka², H. Postma⁶, R. Rickenbach², B.L. Roberts³, E. Rozak¹, T. Ruf⁴, L. Sacks⁹, L. Sakeliou¹, P. Sanders⁸, C. Santoni², K. Sarigiannis¹, M. Schäfer¹⁷, L.A. Schaller⁷, A. Schopper⁴, P. Schune¹⁴, A. Soares¹⁴, L. Tauscher², C. Thibault¹², F. Touchard⁴, C. Touramanis⁹, F. Triantis⁸, D.A. Tröster², E. Van Beveren⁵, C.W.E. Van Eijk⁶, S. Vlachos², P. Weber¹⁷, O. Wigger¹³, C. Witzig¹⁷, M. Wolter¹⁷, C. Yeche¹⁴, D. Zavrtanik¹⁰ and D. Zimmerman³.

¹University of Athens, ²University of Basle, ³Boston University, ⁴CERN, ⁵LIP and University of Coimbra, ⁶Delft University of Technology, ⁷University of Fribourg, ⁸University of Ioannina, ⁹University of Liverpool,

¹⁰J. Stefan Inst. and Phys. Dep., University of Ljubljana, ¹¹CPPM-IN2P3-CNRS et Université d'Aix-Marseille II, ¹²CSNSM-IN2P3-CNRS, ¹³Paul-Scherrer-Institut(PSI), ¹⁴DAPNIA/SPP CE Saclay,

¹⁵KTH Stockholm, ¹⁶University of Thessaloniki, ¹⁷ETH-ITP Zürich

Presented by R. LE GAC

IN2P3-CPPM, 163 av. de Luminy, case 905, F-13288 MARSEILLE CEDEX 9



Abstract

The CPLEAR experiment running at CERN, measures CP T and CPT,violating parameters in the neutral kaon system by measuring time dependent decay asymmetries between K^0 and \bar{K}^0 . Physics results corresponding to about 25% of the presently available statistics are presented for the $\pi^+\pi^-$ and the $\pi^+\pi^-\pi^0$ decay channels. The present precisions are already at the level of the world average value for ϕ_{+-} and one order of magnitude better than the published results for η_{+-0} . Moreover, for the first time the amplitude for the CP allowed K_S decay into $\pi^+\pi^-\pi^0$ has been evaluated.

INTRODUCTION

The CPLEAR experiment aims to measure the CP, T violation effects through the observation of particle-antiparticle decay rate asymmetries by detecting initial K^0 and \bar{K}^0 states and their decay products: $\pi^+\pi^-$, $\pi^+\pi^-\pi^0$ and $\pi\nu l$.

The decay rate of initial $K^0(\bar{K}^0)$ into $\pi^+\pi^-$ assuming $\epsilon_S = \epsilon_L = \epsilon$, is given by (see the previous presentation for the definition of the parameters):

$$\frac{R_{K^0}(t)}{R_{\bar{K}^0}(t)} \propto [1 \mp \text{Re}(\epsilon)] [e^{-\gamma_S t} + |\eta_{+-}|^2 e^{-\gamma_L t} \pm 2 |\eta_{+-}| e^{-\frac{(\gamma_S + \gamma_L)t}{2}} \cos(\Delta m t - \phi_{+-})]$$

The mass difference Δm between K_S and K_L is approximately equal to half of the decay width of the K_S and the interference amplitude is observable between 5 and 15 τ_S . CPLEAR measures the time dependent asymmetry between \bar{K}^0 and K^0 :

$$A_{\pi^+\pi^-}(t) = \frac{R_{\bar{K}^0 \rightarrow \pi^+\pi^-}(t) - R_{K^0 \rightarrow \pi^+\pi^-}(t)}{R_{\bar{K}^0 \rightarrow \pi^+\pi^-}(t) + R_{K^0 \rightarrow \pi^+\pi^-}(t)}$$

which is sensitive to Δm and the CP violation parameters $|\eta_{+-}|$ and ϕ_{+-} :

$$A_{\pi^+\pi^-}(t) \approx 2\text{Re}(\epsilon) - \frac{2 |\eta_{+-}| e^{\frac{(\gamma_S - \gamma_L)t}{2}} \cos(\Delta m t - \phi_{+-})}{1 + |\eta_{+-}|^2 e^{(\gamma_S - \gamma_L)t}} \quad (1)$$

The detector efficiencies, geometrical acceptance and residual background are identical for K^0 and \bar{K}^0 decay and therefore cancel in the asymmetry. This measurement allows to evaluate both the amplitude and the phase of the CP violation parameter in the $\pi^+\pi^-$ decay channel. Moreover, the conservation of CPT can be tested by comparing ϕ_{+-} with the superweak angle ϕ_{SW} defined as $\arctan(2\Delta m/(\gamma_S - \gamma_L))$. If CPT is a good symmetry $\phi_{+-} \approx \phi_{\text{SW}}$. The present value for ϕ_{SW} is $(43.7 \pm 0.2)^\circ$ [1].

Time dependent asymmetries probe the CP violation even if the final state is not a pure CP eigenstate and allow to measure η_{+-0} in the $\pi^+\pi^-\pi^0$ decay.

1 THE CPLEAR EXPERIMENT

The CPLEAR detector, shown on Figure 1, allows to tag the strangeness of the neutral kaon and to measure its decay product. It has a cylindrical geometry and is mounted inside a solenoid of 3.6 m length and 1 m radius, which produces a magnetic field of 0.44 T parallel to the antiproton beam. Antiprotons coming from LEAR with a momentum of 200 MeV/c annihilate at rest, at the center of the detector, in a spherical target filled with hydrogen at a pressure of 16 bars. K^0 and \bar{K}^0 are produced in $\approx 4 \cdot 10^{-3}$ of the annihilations through the reactions:

$$p\bar{p} \rightarrow K^+\pi^-\bar{K}^0 \text{ or } K^-\pi^+K^0 \quad (2)$$

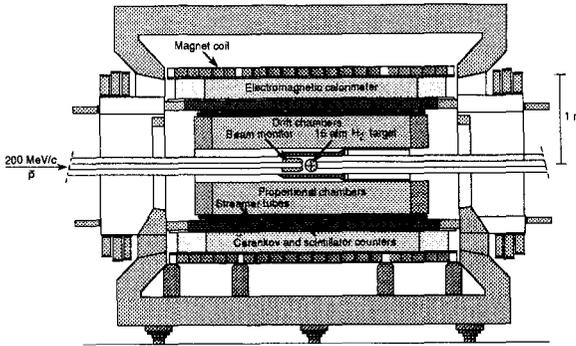


Figure 1: CPLEAR detector

In these reactions, the identification of the charged kaon and of its charge sign determines the strangeness of the associated neutral kaon.

The target is surrounded by a tracking device consisting of 2 proportional chambers, 6 drift chambers and 2 layers of streamer tubes. This detector allows the reconstruction of the charged particle tracks and the measurement of their momentum. It is followed by a particle identification device, made of a threshold Cherenkov sandwiched between 2 layers of scintillators, which separates charged kaons (which do not produce light in the Cherenkov) from charged pions. Finally an 18-layer gas-sampling calorimeter is used to detect gammas. With its high granularity, the impact points of the showers are reconstructed accurately.

To face the high rate (1 MHz) required to measure accurately the CP violation parameters and to select the reaction (2) among the much more numerous multipionic annihilations, a sophisticated multi-level trigger is required. It is based on the recognition of the associated charged $K\pi$ pair at the annihilation vertex. The first requirement of the trigger is that there are at least two hits in the inner scintillator layer and at least one charged kaon candidate in the Cherenkov. The next trigger stages place kinematical constraints on the event firstly by applying a cut to the transverse momentum of the kaon candidate, and then by requiring 2 or 4 tracks after a full event reconstruction and track parameterization. The following processor improves the particle identification using the energy loss of the particles in the scintillators, the time of flight difference between kaons and pions and the number of photo-electrons in the Cherenkov. The last trigger stage requires a minimum number of clusters in the calorimeter for two-tracks events. The maximum trigger decision time is around $34 \mu\text{s}$ and its reduction factor is about 1000.

2 K^0 DECAY INTO $\pi^+\pi^-$

The identification of the K^0 decay into $\pi^+\pi^-$ is based on constrained fits. The constraints are the following:

- the missing mass to the $K^\pm\pi^\mp$ pair at the annihilation vertex should be equal to the K^0 mass.
- the total energy of the 4 particles should be equal to twice the proton mass and the total momentum should be equal to zero.
- the K^0 flight direction should agree with the direction between the annihilation and the decay vertex.

The present analysis used data recorded up to the end of 1992. After the fits, $2 \cdot 10^6$ events are selected with a decay time above $2 \tau_s$. In the final sample the remaining background is very small. It is only present at large decay time and comes mainly from the K_L decay into $\pi\nu l$. It has been evaluated to be about 1.21 ± 0.14 larger than the CP violating K_L decay into $\pi^+\pi^-$.

Even if K^0 and \bar{K}^0 are produced symmetrically, the number of reconstructed K^0 and \bar{K}^0 is not equal because the detector efficiencies of K^+ and K^- differ in the detector due to their different nuclear cross-section. This effect modifies the relative proportion of K^0 and \bar{K}^0 in our sample but we have checked that this proportion does depend neither on the K^0 momentum nor on the K^0 decay time. It is taken into account by introducing a normalization factor 1.1187 ± 0.0026 measured on $K^0 \rightarrow \pi^+\pi^-$ events with a decay time between 2 and $4 \tau_s$ where the CP violating effects are very small.

The reconstructed time dependent asymmetry is shown on Figure 2a together with the results of the fit. The fit of the asymmetry to expression (1), using a fixed value of γ_S, γ_L and Δm ($0.5351 \pm 0.0024 \cdot 10^{10} \bar{h}/s$) from reference [1] and taking into account the remaining background, gives the following values for the CP violation parameters:

$$\begin{cases} \eta_{+-} &= (2.25 \pm 0.07_{\text{stat}} \pm 0.02_{\text{sys}}) 10^{-3} \\ \phi_{+-} &= (44.7 \pm 1.3_{\text{stat}} \pm 0.7_{\text{sys}} \pm 0.7_{\Delta m})^\circ \end{cases}$$

The quoted systematic errors are preliminary. They take into account the uncertainties linked to the correction for the regeneration effects and to the K^0, \bar{K}^0 normalization. The error related to the uncertainty on Δm will be reduced in the future due to the measurement of Δm which will be obtained in our experiment.

3 K^0 DECAY INTO $\pi^+\pi^-\pi^0$

In this analysis, events are selected by using the following criteria:

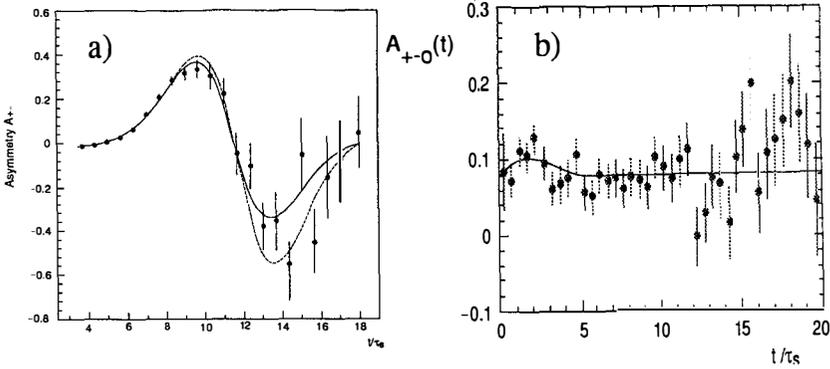


Figure 2: a) Time dependent asymmetry for K^0 decay into $\pi^+\pi^-$. The solid (dotted) line is the result of a fit with (without) taking into account the residual three-body background. b) Time dependent asymmetry for K^0 decay into $\pi^+\pi^-\pi^0$.

- the missing mass to the $K^\pm\pi^\mp$ pair should be equal to the K^0 mass.
- the missing mass to the 4 detected charged particles should be equal to the π^0 mass.
- the K^0 flight direction should agree with the direction between the annihilation and the decay vertex.
- the event should not satisfy the kinematics of $K^0 \rightarrow \pi^+\pi^-$.
- at least one photon should be detected in the calorimeter.

After selection, about 36 000 events remain with a background level of 15%. The time dependent asymmetry A_{+-0} allows a measurement of $\text{Re}(\eta_{+-0})$ and $\text{Im}(\eta_{+-0})$:

$$A_{+-0}(t) \approx 2\text{Re}(\epsilon_S) - 2(\text{Re}(\eta_{+-0}) \cos(\Delta mt) - \text{Im}(\eta_{+-0}) \sin(\Delta mt)) e^{-\frac{(\gamma_S - \gamma_L)t}{2}}$$

The fit of the time dependent asymmetry shown on Figure 2b gives the following results:

$$\begin{cases} \text{Re}(\eta_{+-0}) = 0.002 \pm 0.016_{\text{stat}} \\ \text{Im}(\eta_{+-0}) = 0.044 \pm 0.026_{\text{stat}} \end{cases}$$

The systematic errors, still under study, are estimated to be smaller than the present statistical error.

By studying separately the time dependent asymmetry for events with $x < 0$ or $x > 0$, where $x = (m_{\pi^+\pi^0}^2 - m_{\pi^-\pi^0}^2)/m_\pi^2$, the amplitude of the CP allowed K_S decay into $\pi^+\pi^-\pi^0$ can be measured (Figure 3). The fit performed simultaneously on both asymmetries leads to the following CP allowed decay width of the K_S :

$$\frac{\Gamma(K_S \rightarrow \pi^+\pi^-\pi^0)_{CP=+1}}{\Gamma(K_L \rightarrow \pi^+\pi^-\pi^0)_{CP=-1}} = (2.1 \pm 1.2) 10^{-3}$$

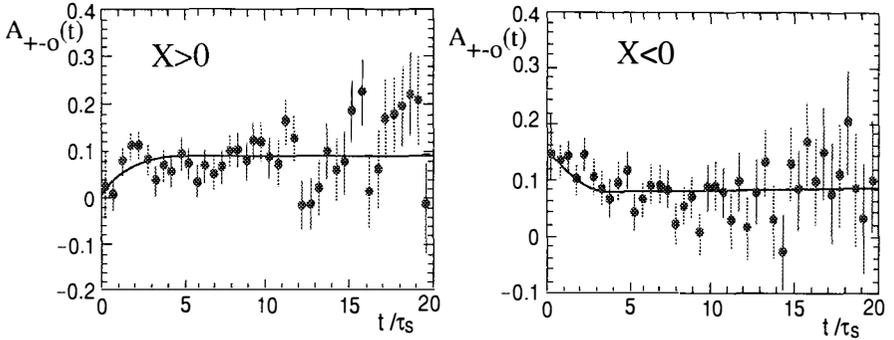


Figure 3: Time dependent asymmetry for K^0 decay into $\pi^+\pi^-\pi^0$ with $x > 0$ and $x < 0$

	PDG '92	CPLEAR '92	CPLEAR '95
$ \eta_{+-} [10^{-3}]$	2.279 ± 0.022	2.25 ± 0.07	± 0.02
$\phi_{+-} [^\circ]$	46.5 ± 1.2	44.7 ± 1.3	± 0.4
$\text{Re}(\eta_{+-0})$	0.05 ± 0.17 [2]	0.002 ± 0.016	± 0.004
$\text{Im}(\eta_{+-0})$	0.15 ± 0.33 [2]	0.044 ± 0.026	± 0.006
$\frac{\Gamma(K_S \rightarrow \pi^+\pi^-\pi^0)_{CP=+1}}{\Gamma(K_L \rightarrow \pi^+\pi^-\pi^0)_{CP=-1}}$		$(2.1 \pm 1.2) 10^{-3}$	$\pm 0.3 10^{-3}$

Table 1: Present and future performances of CPLEAR

CONCLUSIONS

The CPLEAR experiment measures CP violating parameters in the K^0 system by measuring time dependent rate asymmetries between initial K^0 and \bar{K}^0 . The present results are summarized in table 1 together with the world average values given in the PDG 92 [1]. The data recorded in 1993 are being analyzed and should bring an improvement of factor 4 in statistics. Running up to 1995 should allow a statistical improvement by a factor of ≈ 16 . The expected statistical precisions at the end of the experiment are listed in table 1. The improved accuracy on ϕ_{+-} will allow a precise test of CPT.

References

- [1] Part. Data Group, Phys. Rev. **D45**, 1 (1992)
- [2] V. V. Barmin et al., Nuovo Cimento **A85**, 703 (1985)

**PARITY VIOLATION IN ATOMS,
THE PARIS EXPERIMENTS ON CESIUM**

Marie-Anne BOUCHIAT
Laboratoire Kastler Brossel de l'ENS
24 rue Lhomond, 75231 Paris Cedex 05

ABSTRACT

Parity violation (PV) in stable atoms arises from Z^0 exchange between electrons and nucleons. Measurements of atomic PV provide low energy tests of the electroweak interaction complementary to high energy experiments because in the atomic case all nucleons interact coherently with the Z^0 . The cesium atom is of particular interest since atomic calculations have reached 1% accuracy, while work is underway to push the accuracy of the experimental measurements to the same level. This would constrain further electroweak theories while the comparison of PV from different hyperfine components should allow one to identify a nuclear spin dependent PV interaction.

1. Introduction

Parity violation occurs in atoms because, in addition to the electromagnetic interaction between electrons and nucleons (exchange of photons), there exist weak neutral currents (Z^0 exchange). The hamiltonian describing the interaction is, in the non-relativistic approximation,

$$H^{PV} = \frac{G_F}{4\sqrt{2}} \left[\frac{\boldsymbol{\sigma}_e \cdot \mathbf{p}_e \delta(\mathbf{r}) + h.c.}{m_e} \right] Q_W \quad (1)$$

where G_F is the Fermi constant and $\boldsymbol{\sigma}_e \cdot \mathbf{p}_e / |\mathbf{p}_e|$ the electron helicity. The delta function $\delta(\mathbf{r})$ indicates the short range of the interaction associated with the heavy boson mass, very short indeed on the atomic scale. The parameter Q_W is the weak nuclear charge. It represents the combined effect of the weak charges of all the quarks of the nucleus :

$$Q_W = (2Z + N) Q_W(u) + (Z + 2N) Q_W(d) \quad (2)$$

Z and N being the atomic and neutron number. As far as atomic physics is concerned, the weak interaction H^{PV} mixes states of opposite parity. This leads to the apparition of an electric dipole transition moment between states of the same parity; in other words, usual atomic selection rules are violated. However, since T-invariance is preserved, there are no energy level shifts at first order in H^{PV} . If the atomic proportionality factor has been calculated, then by measuring a parity violating transition amplitude one can extract a value for Q_W .

The possibility of observing PV in heavy atoms was originally suggested by Claude Bouchiat and myself in 1973¹⁻². An experiment begun in Paris at that time led in 1982-83 to the first quantitative determination of PV in an atom. The result reached 15% accuracy

this being the quadrature sum of the experimental uncertainty and theoretical one at the moment³. The principle of the experiment performed in atomic Cs was to excite a forbidden magnetic dipole transition by counter-propagating laser beams of opposite helicity and to search for a circular polarisation of the fluorescent light emitted perpendicular to the direction of excitation. A forbidden electromagnetic transition was chosen in order that the weak interaction assume a greater relative contribution. This reduces the risk of systematic error. In fact, because the transition is so forbidden a static electric field was applied to render it slightly allowed and so optimise the signal-to-noise ratio. In addition, since the PV signal changes with the sign of the electric field orthogonal to the direction of the beams, the reversal of this field provided a further discriminant against systematic effects. Indeed the dominant contribution to the experimental uncertainty was of statistical origin.

In this paper we shall focus our attention on the work done in Cesium. This atom even today still appears the best compromise between a high value of Z which considerably enhances the parity violating effects in atoms² and the simplicity of the atomic structure which helps one to perform reliably the computations required for the interpretation. The experimental and theoretical work concerning the thallium atom and the optical rotation of atomic vapors like bismuth and lead has been discussed in various review articles⁴ and the most recent work is presented in this volume, as well as the very interesting dysprosium work.

2. The M_1 Forbidden Transitions and the First Generation of Experiments in Cs and Tl

In order to discuss electroweak parity-violation in highly forbidden transitions it is convenient to introduce an effective dipole operator. This is a hybrid vector-type quantity which behaves as a matrix element in the radial atomic coordinate space and as an operator in the spin space: σ is the spin operator. It contains the various components which govern the interaction with the radiation field exciting the $nS_{1/2} - n'S_{1/2}$ (or $nP_{1/2} - n'P_{1/2}$) transition⁵:

$$\begin{aligned} \mathbf{d}(n', n) = & -\alpha(n', n)\mathbf{E} - i\beta(n', n)\boldsymbol{\sigma} \times \mathbf{E} \\ & + M_1(n', n)\boldsymbol{\sigma} \times \mathbf{k} \\ & - iImE_1^{PV}(n', n)\boldsymbol{\sigma} \\ & + \dots(\text{nuclear spin dependent terms}) \end{aligned} \quad (3)$$

The first term corresponds to the Stark-induced dipole, it involves a scalar and a vector transition polarizability. The next term represents the normal magnetic dipole, \mathbf{k} is the direction of propagation of the light beam. The last component directed along the electron spin is the result of the PV weak interaction.

When the transition is excited the radiation field is absorbed coherently by these transition dipoles. Interferences take place between the various dipole amplitudes and lead to new optical pumping effects^{5,6}. An interesting effect of this kind occurs in the electronic polarization created in the final state when the transition is excited in a transverse electric field with circularly polarized light (ξ represents the incident photon helicity). From the usual optical pumping rules we expect a longitudinal component such that the atomic angular momentum

is proportional to the photon angular momentum. It is indeed present and takes the form : $\mathbf{P}_{\parallel} = -\frac{\beta}{\alpha}\xi\mathbf{k}$. However we can predict also a transverse component, far more unusual :

$$\mathbf{P}_{\perp} = -2\frac{M_1 + \xi ImE_1^{PV}}{\alpha E^2}\mathbf{k} \times \mathbf{E}$$

The first term proportional to M_1 is present even if the incident beam is linearly polarized and it is parity conserving : indeed the quantity $\mathbf{k} \times \mathbf{E}$ transforms under P like an angular momentum. It results from a breaking of the revolution symmetry by the E field. The second term proportional to ImE_1^{PV} involves the same quantity multiplied by the beam helicity ξ , therefore it behaves as a polar vector and this is a manifestation of parity-violation. From these considerations results the basic principle of our experiment in 1982 at ENS^{3,7}. We used counterpropagating beams whose contributions to \mathbf{P}_{\perp} via the M_1 amplitude cancel each other. Hence the ratio P_{\perp}/P_{\parallel} directly yields the PV parameter of interest : $ImE_1^{PV}/\beta E \sim 1.5 \times 10^{-5}$ at a typical field strength of 100 V/cm. The electronic polarization in the $7S$ state was detected by measuring the circular polarization ratio of the fluorescence light emitted in the radiative decay of the $7S_{1/2}$ state via the allowed $7S_{1/2} - 6P_{1/2}$ transition (1.36 μ).

A schematic view of this experiment is shown in Fig. 1. The forward and backward beams propagate parallel to the capacitor plates and one observes the fluorescence light emitted in the direction orthogonal to the beams and to the field through a circular polarization analyzer. In this way \mathbf{P}_{\perp} is monitored continuously. The detection of \mathbf{P}_{\parallel} is performed during calibration sequences during which a 10G magnetic field is applied along E : the component of \mathbf{P}_{\parallel} appearing in the observation direction is thus detected with the same optics and electronics as \mathbf{P}_{\perp} . Two key-components of this experiment were the cesium cell with multipassages of the laser beam between internal mirrors (Fig.2) and the triple modulator providing complete characterization of the beam polarization. In this experiment the shot-noise limit was reached. The control of geometrical alignment, stray E - fields and light polarization defects was made from parity-conserving signals⁷. About thirty such signals were recorded simultaneously with the PV signal and exploited in order to reconstruct in real time the systematic asymmetry resulting from the combination of two defects. The effective averaging time lasted 300 hours. The experiment was performed twice, on two different hf s components, a $\Delta F = 0$ and a $\Delta F = 1$ transition, where the observation conditions are very different. The two results were in agreement with each other.

The Paris 82 experiment which reached 12 % experimental accuracy led to a new quantitatively interpretable test of electroweak theory at low energy. There was room for improvement however, mainly because of the limited angular acceptance of our detection system. Therefore, we kept thinking about various possibilities of improvements. In fact, in 1978 we had proposed new experiments in static transverse electric and magnetic fields where a pseudoscalar this time appears in the transition probability⁸.

In the first one the PV effect is a chiral absorption of plane polarized light involving the pseudoscalar quantity $(\boldsymbol{\varepsilon} \cdot \mathbf{B})(\boldsymbol{\varepsilon} \cdot \mathbf{E} \times \mathbf{B})$. This effect was observed later on in the $6P_{1/2} - 7P_{1/2}$ transition in Tl in a beautiful experiment conducted in 1984-86 by E. D. Commins at Berkeley⁹.

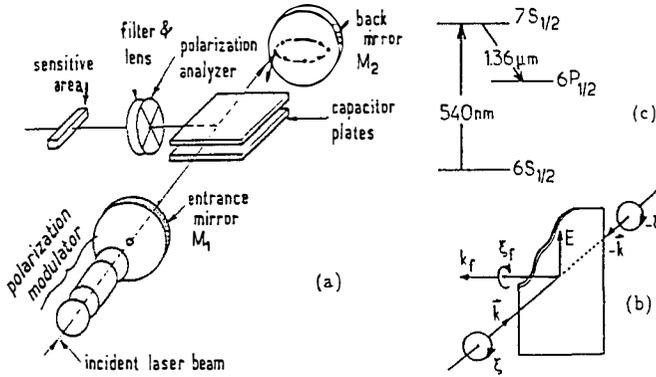


Fig. 1 : The Paris 82-83 experiment. (a) The main elements of the apparatus. Mirrors M_1 and M_2 inside the cesium cell allow about 60 double passages of the beam through the vapor with small direction dispersion $\lesssim 1^\circ$ (beam impacts shown on mirror M_2). (b) Mirror-symmetry of the experiment broken by the circular polarization ξ_f of the fluorescence intensity detected in the direction k_f contained in the symmetry plane. (c) Cesium energy levels involved in the parity-violation experiment.

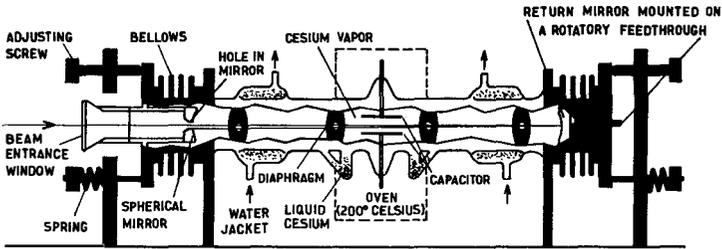


Fig. 2 : The Cesium cell with multipassages of the beams between internal mirrors ⁷.

In our second suggestion the PV effect is a circular dichroism involving the pseudoscalar quantity $\xi \mathbf{k} \cdot \mathbf{E} \times \mathbf{B}$. This is the basic principle of the experiment started in Cs by C. E. Wieman at Boulder which yielded results in 1985¹⁰ and 88¹¹ and presently provides the most precise PV measurements in Cs. For this experiment the Zeeman structure of the $6P_{1/2} - 7P_{1/2}$ transition must be resolved, otherwise the effect cancels out. For this reason the Boulder group has chosen to work with an atomic beam that they irradiate at right angles. Thus Doppler broadening is avoided and the field strength required is less than 100G. Since it is no longer necessary to analyse the polarization of the fluorescence light they detect all the light emitted in the red in the second step of the $7S_{1/2}$ state radiative decay, at the price of a certain loss in selectivity. The accuracy achieved in 1988 reached 2.5 %. There is still room for improvement by using a polarized Cs beam.

The upgraded version of the Boulder experiment is discussed elsewhere in this volume and a totally new version of the Paris experiment in Cs, at an advanced stage is presented later on. Before this, we are going to examine the present results of experiments and atomic theory in Cs and their implications for electroweak theory. This will make more explicit our motivations for improving the experimental accuracy.

3. Present Status of the Results in Cesium

We turn now to the atomic physics calculations necessary to extract the weak charge from the experimental data¹². We can write E_1^{PV} as the product of $(Q_W - N)$ by a purely atomic quantity, $\mathcal{E}_1^{PV} = E_1^{PV} / \left(\frac{Q_W}{-N} \right)$ (the neutron number is introduced here just for convenience : $Q_W / -N$ being close to unity). \mathcal{E}_1^{PV} can be considered as an infinite sum over the intermediate P states admixed with the S states by the parity-violating interaction :

$$\mathcal{E}_1^{PV} = \left(-\frac{N}{Q_W} \right) \sum_n \frac{\langle 7S_{1/2} | d_z | nP_{1/2} \rangle \langle nP_{1/2} | V_{PV}^{(1)} | 6S_{1/2} \rangle}{E(6S_{1/2}) - E(nP_{1/2})} + \text{crossed terms} \quad (4)$$

The atomic orbitals and the valence-state energies are perturbed by many-body effects. There are two different approaches, one starting from first-principles and the second semi-empirical. In the first case the Dirac equation is solved for one electron in an Hartree-Fock potential and the many-body corrections are computed perturbatively (with resummation of specially chosen sets of diagrams). The second approach incorporates in a consistent way much empirical information available from the numerous spectroscopic measurements performed in Cs, namely information regarding energies of the valence states, allowed-dipole matrix elements and hfs splittings of the $S_{1/2}$ and $P_{1/2}$ states. In this way atomic many-body perturbation theory is needed only to compute small correcting terms.

Table 1 : Summary of the results in Cesium

A - Experiments ^(*)	B - Atomic Theory Computations
$ImE_1^{PV} (10^{-11} e a_0)$	$(-N/Q_W) E_1^{PV} (10^{-11} e a_0)$
-0.803 ± 0.095 ³ Paris 82 - 83	"First Principles" :
-0.872 ± 0.069 ¹⁰ Boulder 85	-0.908 ± 0.010 ¹³ Novosibirsk 89
-0.833 ± 0.020 ¹¹ Boulder 88	-0.905 ± 0.009 ¹⁴ Notre Dame 90
*using $\beta (a_0^3)$ from S.E. determinations	Semi - Empirical :
27.19 ± 0.4 ¹⁸	$-0.935 \pm 0.02 \pm 0.03$ ¹⁵ Paris 86
27.17 ± 0.35 ¹⁹	-0.904 ± 0.02 ¹⁶ Oxford 90
	-0.895 ± 0.02 ¹⁷ Paris 91

$$A/B \Rightarrow Q_W^{ex} = -71.5 \pm 1.8$$

The results in cesium are presented in Table 1. Experimental values of ImE_1^{PV} are deduced from the measured PV parameter ImE_1^{PV}/β , combined with the value of β obtained

from two independent semi-empirical determinations (Table 1). The accuracy has improved from 12 % in our first 82-83 experiment to 2.5 % in the last Boulder measurements but all results are in agreement with each other to within the quoted uncertainty. Concerning atomic physics theoretical results there are two computations using the first-principles approach^{13,14} which agree with each other, each one reaching an accuracy of one per cent. This is a real "tour de force" and in this respect we should stress the pioneering work of the Novosibirsk group¹³, the first to show that relativistic many-body computation in Cs could be pushed to the 1 % precision level using techniques inspired from field theory. The work of the Notre Dame group was much more than a mere confirmation. By using very advanced numerical methods they arrived at an exact treatment of two electron-hole pair excitation. The calculations by the semi-empirical approach^{15,16} now reach an accuracy of 2 % which is limited by the uncertainty relative to the empirical inputs (the allowed E_1 amplitudes). This should improve once new precise measurements have been performed. The agreement obtained by the two approaches still reinforces our confidence in the final result. Finally, by combining the experimental and the atomic theoretical results we obtain an empirical value of Q_W whose accuracy at present is limited essentially by experiment :

$$Q_W^{\text{ex}} = -71.5 \pm 1.8 \quad (5)$$

4. Implications for Electroweak Theory

4.1. Model-Independent Interpretation

The most direct implication of a precise determination of the weak nuclear charge concerns the determination of the fundamental constants $C_u^{(1)}$, $C_d^{(1)}$ appearing in the parity-violating Lagrangian :

$$\mathcal{L}_{PV} = -\frac{G_F}{\sqrt{2}} \sum_{\alpha} \bar{e} \gamma^{\mu} \gamma_5 e C_{q_{\alpha}}^{(1)} \bar{q}_{\alpha} \gamma_{\mu} q_{\alpha} \quad (6)$$

where q_{α} represent the various quarks (u , d , etc...) of the nucleus. In a model independent analysis, in the electron-hadron sector the cesium result defines an allowed band in the $C_u^{(1)}$, $C_d^{(1)}$ plane (Fig. 3), $C_u^{(1)} = -\frac{1}{2}Q_W(u)$ and $C_d^{(1)} = -\frac{1}{2}Q_W(d)$ are directly related to the weak charges of the quarks.

The high-energy SLAC experiment observing highly inelastic polarized electron scattering on deuterons²⁰ define another allowed band. Its slope is totally different because, the nuclei and the nucleons themselves being broken, the quarks act incoherently and the linear combination of $C_u^{(1)}$ and $C_d^{(1)}$ determined from the measurement happens to be nearly orthogonal to that obtained for atomic physics. This illustrates the complementarity of the two types of experiments. The high-energy band is also larger, this is partly so because the result is obtained from a two-parameter fit : indeed both the vector and the axial vector hadronic current contribute to the asymmetry. We expect SLC and HERA will soon improve upon the precision of the SLAC results.

The predictions of the Standard Model for the various values of $\sin^2\theta_W$ are represented in Fig. 3 by a segment. It is noteworthy that it intercepts the intersection region for a value of $\sin^2\theta_W$ which is close to the world average value.

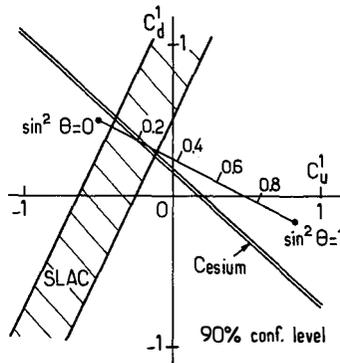


Fig. 3 : Experimental, model-independent determination of the weak charges of the u and d quarks. The two bands represent the domains allowed by the high-energy SLAC experiment²⁰ and by the cesium experiments (Table 1). The graduated segment represents the prediction of the Standard Model for values of the parameter $\sin^2\theta$ from 0 to 1.

4.2. Comparison to the Standard Model Prediction

Since Q_W is a fundamental quantity it must be compared to the value predicted by the Standard Model. In order to incorporate radiative electroweak corrections, the theory being renormalizable, it is necessary to fix a certain number of constants, three in fact, for a complete definition of the theory. We choose α , G_F and as the third constant the most natural choice is the mass of the Z_0 boson, M_Z . With such a choice the sensitivity of the radiative corrections to the top mass, via vacuum polarization loops which involve $(m_t/m_W)^2$ terms (Veltman effect), practically disappears. It even cancels accidentally for the particular ratio of protons to neutrons involved in natural cesium²¹. The theoretical value of Q_W is :

$$Q_W^{\text{th}} = -73.20 \pm 0.13 \quad (7)$$

the uncertainty coming from hadronic physics. We must note that the difference with the experimental value

$$\delta Q_W = Q_W^{\text{ex}} - Q_W^{\text{th}} = 1.7 \pm 1.8 \quad (8)$$

does not exceed one σ . This is a remarkable result if one thinks of the difference of some 10 orders of magnitude between the values of q^2 at which the determinations of M_Z and Q_W were carried out. Such a result demonstrates the success of the Standard Model over this extensively wide range. If as the third constant we choose M_W/M_Z instead of M_Z , the determination of Q_W^{th} would then involve the top quark mass. As early as 1989 this was used to place a limit on $m_t < 190 \text{ GeV}/c^2$ (at the 90 % confidence level). Even if there are now more stringent compatible limits coming from LEP and recently more direct indications coming from Fermi Lab, it is satisfactory to note the consistency between results of very different physical origin.

4.3. Constraints Regarding an Extra Z_0 Boson

It presently looks beyond the reach of atomic physics results to compete with LEP data in a test of q^2 – independent radiative corrections of pure electroweak origin²². However the major merit of atomic physics results is their ability to test q^2 – dependent effects. In particular most alternatives to the Standard Model make the assumption of a second neutral vector boson Z'_0 . While the LEP data are sensitive to the Z'_0 only through its coupling with the Z_0 boson, the atomic physics data receive a contribution whatever the degree of $Z_0 - Z'_0$ mixing and even when there is no mixing whatsoever. The corresponding modification of Q_W is proportional to $(M_Z/M_{Z'})^2$. Several theorists²³ have analyzed the present results. In a specific model G. Altarelli et al. find that the lower limit to $M_{Z'}$ coming from Q_W^{ex} with the present uncertainty is $360 \text{ Gev}/c^2$ while the direct CDF measurements lead to a somewhat smaller limit $288 \text{ Gev}/c^2$. A more detailed analysis is presented elsewhere²⁴.

4.4. The Nuclear Spin-Dependent Electron-Nucleus PV Interaction

The study of the nuclear spin-dependent e-nucleus PV interaction constitutes another motivation for precise PV atomic measurements. Indeed from measurements performed on different hfs lines, the spin-dependent and spin-independent contributions can be separated unambiguously. The separation of the effects does not require any improvement in atomic theory.

In the case of Cs it is convenient to introduce the measurable physical quantity :

$$\tau_{hf} = \frac{E_1^{PV}(6S - 7S, \Delta F = -1)}{E_1^{PV}(6S - 7S, \Delta F = +1)} - 1 \quad (9)$$

A non zero experimental value of τ_{hf} can serve as a diagnosis of a nuclear spin-dependent contribution. The Boulder group has announced a 2σ effect¹¹ :

$$\tau_{hf} = (8.3 \pm 4.4) \times 10^{-2} \quad (10)$$

Obviously more measurements are necessary to arrive at a definite conclusion.

The dominant physical origin is actually considered as being not the $e -$ nucleus neutral currents but in fact the hadronic weak interactions taking place inside the nucleus. The parity-violation in the nucleus contaminates the atom via the conventional photon exchange (see Fig. 4). Actually, this kind of interaction takes the same structure as the axial coupling of the Z_0 to the nucleus but, for cesium, it is about 4 times larger. It was first discussed by Zel'dovich, who introduced the relevant concept of the nuclear anapole moment²⁵. The magnitude of the total nuclear spin-dependent effect was computed by the Novosibirsk group who predicted for Cs²⁶ :

$$\tau_{hf} = 3.2 \times 10^{-2} \quad (11)$$

More detailed computations by the Paris group included nuclear configuration mixing effects and gave an estimate of the uncertainties originating from both nuclear and hadronic physics²⁷ :

$$\tau_{hf} = (1.7 \text{ to } 2.3) \times 10^{-2} \quad (12)$$



Fig. 4 : (Left), schematic representation of the dominant nuclear spin-dependent parity-violating electron-nucleus interaction. (Right), chiral distortion of the nuclear magnetization resulting from parity-violating interactions between the nucleons. The nuclear spin \mathbf{I} is perpendicular to the figure.

From these evaluations (Eq. 11 and 12) we note that theory predicts values substantially smaller than the central experimental value (Eq. 10).

The Paris group has also given a physical interpretation of the nuclear anapole effect in the frame of a realistic nuclear model. The parity-violating nuclear interactions generate a chiral distortion \mathcal{M}^{PV} in the normal spin magnetization. This distortion lies in the plane perpendicular to the nuclear spin \mathbf{I} and is shown in Fig. 4. The nuclear anapole moment \mathcal{A} is the first moment of this chiral distribution :

$$\mathcal{A} = 2\pi \int \mathbf{r} \times \mathcal{M}^{PV}(\mathbf{r}) d^3r \quad (13)$$

\mathcal{M}^{PV} being deduced from the normal spin magnetization $\mathcal{M}_I(\mathbf{r})$ by an infinitesimal rotation β around the radius vector \mathbf{r} :

$$\mathcal{M}^{PV}(\mathbf{r}) = \beta \frac{\mathbf{r}}{|\mathbf{r}|} \times \mathcal{M}_I(\mathbf{r}) \quad (14)$$

It can be seen that \mathcal{A} is an axial vector directed along \mathbf{I} . In terms of this anapole moment the electron-nucleus PV hamiltonian can be written :

$$H^{PV} \propto \boldsymbol{\alpha} \cdot \mathcal{A} p(r_e) \quad (15)$$

$\boldsymbol{\alpha}$ being the Dirac velocity operator and $p(r_e)$ the nuclear spin distribution.

5. The Second Generation of Cesium Experiments in Paris with Detection via Stimulated Emission

The basic principle in our new method in Paris is to replace the fluorescence method of detection by one using the observation of the stimulated emission caused by a second laser beam tuned at the resonance frequency of the $7S_{1/2} - 6P_{3/2}$ transition. The detection efficiency is greatly improved since all atoms are forced to emit in the direction of the probe beam. To take full advantage of the method the only constraint is that the probe amplification reach or exceed unity : otherwise the atomic signal would be overwhelmed by the probe intensity and the asymmetry appearing in the transmitted probe intensity would be

seriously diluted. In practice this condition is achieved by employing intense and short pulses of excitation light to create a large $7S - 6P_{3/2}$ population inversion in a time short compared with the $7S$ lifetime. In addition, the interaction region can be made long by choosing the excitation and probe beams to be colinear. An electric field, colinear with the beams, controls the transition probability and hence the probe amplification.

In this experiment the right-left asymmetry manifests itself as a change of the principal axes of the refractive index of the excited vapor²⁸. (Note that the imaginary part of the refractive index is nothing but the gain). With linearly polarized excitation beam the excited vapor is prepared with two planes of symmetry. However, as a result of parity violation, its principal axes do not lie in those symmetry planes. They deviate from them by a small angle θ^{PV} and this angle $\theta^{PV} = ImE_1^{PV}/\beta E$, odd under reversal of the electric field, is the PV parameter to be determined experimentally. As a consequence of this deviation angle, the gain of a linearly polarized probe beam is different according to whether its polarization ϵ_{pr} lies at $+45^\circ$ or -45° to the excitation beam polarization ϵ_{ex} . In other words, the two mirror-image experimental configurations represented at the bottom of Fig. 5 lead to different gains. This reflects the existence of a pseudoscalar contribution to the gain $\propto ImE_1^{PV} \cdot \beta E(\epsilon_{ex} \cdot \epsilon_{pr})(\mathbf{E} \cdot \epsilon_{ex} \times \epsilon_{pr})$.

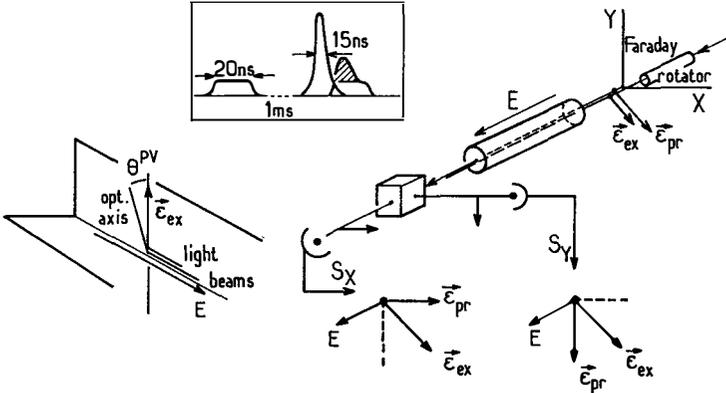


Fig. 5 : The second generation Paris experiment. (Left), the planes of symmetry of the experiment and the deviation of the principle axes from those planes. (Bottom right), the two mirror-image configurations leading to a different amplification of the probe. (Top right), optical system giving access to the deviation angle θ^{PV} at every individual excitation pulse. Inset : Experimental timing.

This is the very principle of the measurement which consists precisely of a comparison of the optical gains for two situations which only differ by the plus or minus sign of the 45° angle between the excitation and probe polarizations. To this end we irradiate the vapor with a coherent linear superposition of these two configurations : i.e. ϵ_{pr} and ϵ_{ex} at the cell entrance are taken either parallel or perpendicular. At the output a two channel polarization analyzer splits the probe beam in two orthogonally directed beams, each one being polarized at $+45^\circ$ and -45° to the incoming beam. In this way the information relative to each one of

the two-configurations to be compared is switched towards two distinct photodetectors (see Fig. 5). For each reference probe pulse, with no $7S$ excitation, exact balance between the two signals S_1 and S_2 is verified. For every probe pulse immediately following an excitation pulse, we measure the imbalance $R = (S_1 - S_2) / (S_1 + S_2)$. We retain the part odd under electric field reversal. This method²⁹ ensures complete rejection of the pulse-to-pulse intensity fluctuations common to both channels. Furthermore, it is a doubly differential technique, since it compares two channels with and without $7S$ excitation. Thereby, it presents specific advantages. Firstly, since it involves the polarization properties of the stimulated emission, it guarantees that only $7S$ atoms excited via the $6S-7S$ transition contribute, so it is highly selective. Secondly, the imbalance R is kept always close to zero, that is to say θ^{PV} is detected via a dark field technique. If we write $R(\pm E) = \theta_0 \pm \theta^{PV}$, the offset θ_0 can be kept below the noise level, providing some careful optical adjustments are made prior to data acquisition. To this end, as well as for precise calibration of the measured value of R in terms of an absolute angle, we use a Faraday rotator to make fine angle tuning of the direction of ϵ_{ex} ²⁹. We note that the present Boulder experiment misses those attractive features : selectivity and zero-measurement technique. However we shall not enter the discussion concerning the advantages and drawbacks of the two parallel approaches at Boulder and Paris, very different in their conception and in the Physics involved. The most important point is that they will provide truly independent measurements.

With the signal to noise ratio presently achieved in our experiment, we can anticipate a precision on θ^{PV} close to 1 % for an effective averaging time comparable to that of the 82–83 experiment. We now face the crucial problem of systematics. For this new experiment we want to stick to the same philosophy adopted in our 82 – 83 experiment⁷. Then, our efforts succeeded in having the various causes determined before beginning long runs, reducing the effects individually (below the expected final statistical uncertainty) and controlling them during acquisition. In this way the final result was practically equal to the raw result : no systematic correction was necessary. Of course the realization of a new experimental situation such as the one chosen, which requires high excitation energy density, means the apparition of new physical effects. Complete understanding of the subjacent atomic processes is required and, we believe, is well on the way to being achieved.

6. Conclusion

Low-energy atomic physics experiments have still an important role to play in the verification of the Standard Model and the exploration of possible alternatives to it. Any gain in the experimental determination of Q_W will directly improve the test of electroweak interactions in the electron-hadron sector, in a region of four-momentum transfers not accessible to large accelerators. In this way they will place more significant constraints on alternative models, in particular those involving an extra neutral vector boson : optical experiments can help to understand Physics at the Tev mass scale !

With the expected gain in precision the nuclear spin-dependent contribution to the electron-nucleus PV interaction will show-up by comparing different *hfs* components. In

addition by comparing different isotopes this should open up the path to further implications for nuclear physics such as detection of an anapole nuclear moment but also variations in the neutron distribution along a string of isotopes ³⁰.

This is why it is important that independent projects develop simultaneously with different new techniques being tried.

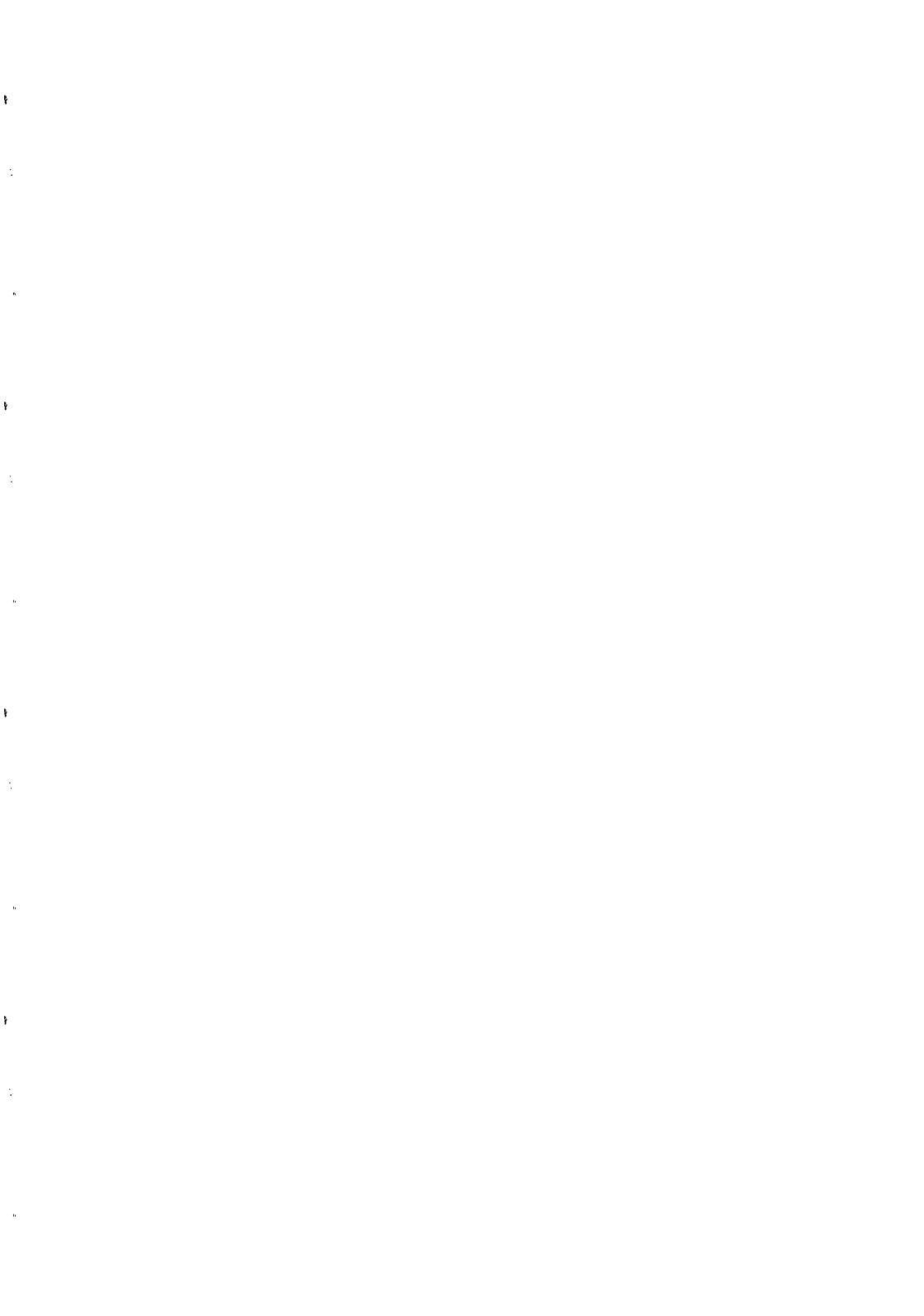
Acknowledgements

The present text owes much to many stimulating discussions with Claude Bouchiat. I am profoundly grateful to the members of our first PV team, Lionel Pottier, Jocelyne Guéna, Larry Hunter for their unique contribution which made the realization of our 82-83 experiment quite an exciting adventure. I express my thanks to all the members of the new PV team, Jocelyne Guéna, Philippe Jacquier and Michel Lintz, whose participation is essential to the accomplishment of our new project.

References

1. M. A. Bouchiat and C. C. Bouchiat, *Phys. Lett.* **48B** (1974) 111.
2. M. A. Bouchiat and C. Bouchiat, *J. Phys. (France)* **35** (1974) 899.
3. M. A. Bouchiat, J. Guéna, L. Hunter and L. Pottier, *Phys. Lett.* **117B** (1982) 358 ; *ibid.* **134B** (1984) 463.
4. E.D. Commins, *Phys. Scr.* **36** (1987) 468 ; E.A. Hinds, in *Atomic Phys.* **11**, ed S. Haroche, J.C. Gay and G. Grynberg (World Scientific, Singapore, 1988) p. 151 ; D.N. Stacey, in *Atomic Phys.* **13**, ed. H. Walther and T.W. Hänsch (American Institute of Physics, New York, 1993) p.46.
5. M. A. Bouchiat and C. Bouchiat, *J. de Phys.* **36** (1975) 493.
6. M. A. Bouchiat, *Ann. Phys. (Paris)* **10** (1985) 923.
7. For a detailed account of the cesium experiment in Paris, including theory, procedure and apparatus, see M. A. Bouchiat, J. Guéna, L. Pottier, *J. Phys. (France)* **46** (1985) 1897 ; for systematic effects *ibid.* **47** (1986) 1175 ; for results and implications, see M. A. Bouchiat, J. Guéna, L. Hunter and L. Pottier, *ibid.*, p. 1709.
8. M. A. Bouchiat, M. Poirier and C. Bouchiat, *J. Phys. (France)* **40** (1979) 1127.
9. P. S. Drell and E. D. Commins, *Phys. Rev. Lett.* **53** (1984) 968 ; *Phys. Rev.* **A32** (1985) 2196 and C. E. Tanner and E. D. Commins, *Phys. Rev. Lett.* **56** (1986) 332.
10. S. L. Gilbert, M. C. Noecker, R. N. Watts, C. E. Wieman, *Phys. Rev. Lett.* **55** (1985) 2680 ; S. L. Gilbert and C. E. Wieman, *Phys. Rev.* **A34** (1986) 792.
11. M. C. Noecker, B. P. Masterson and C. E. Wieman, *Phys. Rev. Lett.* **61** (1988) 310.
12. For a review of atomic calculations in Cesium see S. A. Blundell, A. C. Hartley, Z. Liu, A. M. Martensson-Pendrill and J. Sapirstein, *Theor. Chim. Acta* **80** (1991) 257.
13. V. A. Dzuba, V. V. Flambaum, A. Ya. Kraftmakher and O. P. Sushkov, *Phys. Lett.*

- A141** (1989) 147 and V.A. Dzuba, V.V. Flambaum, and O.P. Sushkor, *Phys. Lett. A* **141** (1989) 147.
14. S. A. Blundell, W. R. Johnson and J. Sapirstein, *Phys. Rev. Lett.* **65** (1990) 1411 ; *Phys. Rev.* **D45** (1992) 1602.
 15. C. Bouchiat and C. A. Piketty, *Europhys. Lett.* **2** (1986) 511.
 16. A. C. Hartley and P. G. H. Sandars, *J. Phys.* **B23** (1990) 2649.
 17. C. Bouchiat, Private Communication.
 18. J. Hoffnagle, V. L. Telegdi and A. Weiss, *Phys. Lett.* **A86** (1981) 457 ; C. Bouchiat and C. Piketty, *Phys. Lett.* **B128** (1983) 73 and *Nucl. Phys.* **B282** (1987) 745 ; Z. Liu and W. R. Johnson, *Phys. Lett.* **A119** (1987) 407.
 19. M. A. Bouchiat and J. Guéna, *J. Phys. (France)* **49** (1988) 2037.
 20. C. Y. Prescott et al., *Phys. Lett.* **77B** (1978) 347 and **84B** (1979) 524.
 21. W. Marciano and A. Sirlin, *Phys. Rev.* **D29** (1984) 75.
 22. G. Altarelli, R. Barbieri and S. Jadach, *Nucl. Phys.* **B369** (1992) 3.
 23. G. Altarelli et al., *Phys. Lett.* **261** (1991) 146 ; P. Langacker and M. Luo, *Phys. Rev.* **D45** (1992) 278.
 24. P. Binétruy, in *Neutral Currents Twenty Years Later*, Paris (1993), Proceedings edited by U. Nguyen-Khac.
 25. Ya. B. Zel'dovich, *Sov. Phys. JETP* **6** (1957) 1184.
 26. V. V. Flambaum and I. B. Khriplovich, *JETP* **52** (1980) 835 ; V. V. Flambaum, I. B. Khriplovich and O. P. Sushkov, *Phys. Lett.* **B146** (1984) 367.
 27. C. Bouchiat and C. A. Piketty, *Z. Phys.* **C49** (1991) 49 ; *Phys. Lett.* **B269** (1991) 195.
 28. M. A. Bouchiat, Ph. Jacquier, M. Lintz and L. Pottier, *Opt. Commun.* **56** (1985) 100.
 29. M. A. Bouchiat, J. Guéna, Ph. Jacquier, M. Lintz and L. Pottier, *Opt. Commun.* **77** (1990) 374.
 30. E. N. Fortson, Y. Pang and L. Wilets, *Phys. Rev. Lett.* **65** (1990) 2857.



**POLARIZED BEAM MEASUREMENT
OF PARITY NONCONSERVATION IN ATOMIC CESIUM**

D. Cho⁺, C. S. Wood, S. C. Bennett, B. P. Masterson
C. E. Tanner and C. E. Wieman
Joint Institute for Laboratory Astrophysics and Department of Physics,
University of Colorado, Boulder, Colorado 80309-0440



An improved measurement of parity nonconservation in cesium is being performed. We measured the forbidden electric dipole transition amplitude between 6S and 7S states of cesium. We employed optical pumping technique to produce a spin polarized cesium atomic beam, and achieved five times better signal-to-noise ratio than our 1988 measurement. So far we have reproduced the 1988 result and plan to measure the parity nonconserving amplitude within 0.5%. By studying four different hyperfine transitions between 6S and 7S states we expect to measure the anapole moment of the cesium nucleus as well.

I. INTRODUCTION

Since the proposals by the Bouchiat¹⁾ to measure the parity-nonconserving neutral current interactions in heavy atoms, parity nonconservation (PNC) has been measured in Bi, Tl, Pb and Cs with varying degrees of precision. During the 80's when these experiments measured the PNC within roughly 10%, the main interest was to show the existence of the neutral current interaction at this very low energy scale. Following the 2% measurement on cesium PNC by the Boulder group in 1988²⁾, however, atomic cesium has become one of the most important systems to precisely study the standard model and look for a new physics beyond it. It should also be emphasized that we need to know atomic wave function with corresponding precision to interpret the experimental result in atomic PNC in terms of more fundamental parameters of the interaction. The groups at Notre Dame³⁾ and Novosibirsk provided just such calculations for cesium with 1% precision. It is relatively simple structure of cesium atom - one valence electron outside a closed core - that made the calculation tractable and calculation for other heavy atoms with a comparable precision is much more difficult, if not completely impossible.

Since our 1988 measurement we have implemented optical pumping and new detection schemes that gave us a big improvement in the signal-to-noise ratio. Recently we successfully reproduced the 1988 result with only one day of data taking and are currently working to bring the precision down to 0.5%.

II. CESIUM ATOM and PNC MIXING

Heavy atoms are preferred in PNC measurement because size of the PNC is scaling as a cube of an atomic number (Z^3) in addition to relativistic enhancement. For cesium, which is the heaviest stable alkali atom ($Z = 55$), added advantages include relative ease in generating high-density atomic beam, availability of lasers at all the relevant transitions, and the wealth of precise spectroscopic information due to its role in atomic clock.

Ground state cesium is in $6S_{1/2}$ state and Cs^{133} has nuclear spin of $2/7$, leading to the hyperfine structure of $F=3$ and 4 (see Fig. 1). The hyperfine separation is 9.2 GHz, the famous clock transition. First excited states consist of fine structure of $6P_{1/2}$ and $6P_{3/2}$. Transitions between $6S$ and $6P$, known as alkali D lines, are electric dipole (E1) allowed and very strong. They are used for optical pumping or cooling and trapping of cesium atoms. In our PNC measurement we are interested in the transition between $6S$ and $7S$ states, which is forbidden for E1 amplitude if the parity is conserved symmetry. Given the parity nonconserving

nature of the neutral current interaction between electrons and quarks mediated by Z bosons (Fig. 2) we may rewrite the nS eigenstate as

$$|nS\rangle = |nS\rangle + \sum_{n'} \frac{\langle n'P | H_{PNC} | nS \rangle}{E_{nS} - E_{n'P}} |n'P\rangle \tag{1}$$

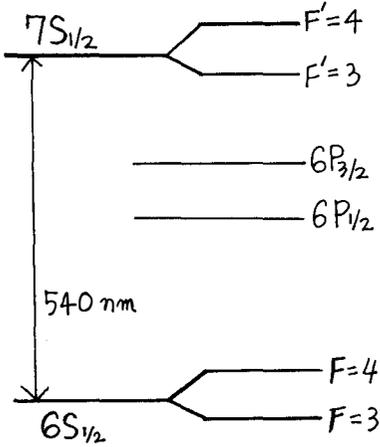


Fig. 1 Cesium Energy Levels

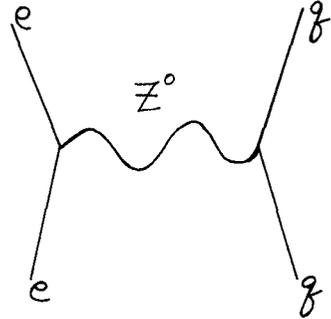


Fig. 2 Diagram for PNC Hamiltonian

III. OUTLINE OF THE EXPERIMENT

Evaluation of the matrix elements in (1) shows that the amplitude of P-state admixture to S is 10^{-11} . This is hopelessly small quantity to measure directly. Instead, we applied dc electric field to produce Stark induced E1 amplitude,

$$A_{E1} = A_{ST} + A_{PNC} \tag{2}$$

which can interfere with each other to produce E1 rate

$$R_{E1} = |A_{ST}|^2 + 2RE (A_{ST}^* A_{PNC}) + |A_{PNC}|^2 \tag{3}$$

The interference term can be made much larger than the PNC term by applying suitable dc electric field. In the real experimental set up (Fig.3) we have E field normal to the k vector of 540-nm laser photons, and magnetic field is also applied perpendicular to both E and k. We

may regard these three vectors as defining our coordinate system. Reversal of any one of E, B or helicity of the photons constitutes a parity transformation and introduces sign change to the interference term. From this point of view we define PNC modulation as

$$dPNC = \frac{R_{E1}(RHC) - R_{E1}(LHC)}{R_{E1}(RHC) + R_{E1}(LHC)}, \quad (4)$$

where, RHC stands for right handed coordinate system, LHC for left handed coordinate. At typical electric field, $E = 500\text{V/cm}$, $dPNC$ is about 6×10^{-6} , which is large enough to be measured precisely.

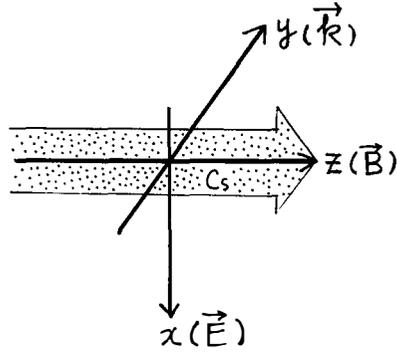


Fig. 3 Experimental Configuration

The interference term in (3) vanishes when averaged over ground state Zeeman levels, and to measure PNC we have to study 6S-7S transition between well defined hyperfine and Zeeman levels. In 1988 experiment we accomplished it by applying large (70 Gauss) magnetic field to resolve transitions between different Zeeman levels. For the present experiment we employed optical pumping technique. We have two diode lasers tuned to cesium D lines to pump almost all the atoms into a single state, e.g. $|6S, F=3, m_F=3\rangle$. This implies 16 fold increase in signal size. Optical pumping also allowed us to use a novel scheme to detect 6S-7S transition.

While optical pumping increased effective atomic beam density, power buildup cavity of 540nm laser beam increased number of photons that interact with atoms. Use of power buildup cavity is not new in Cs PNC experiment, but its improvement helped to increase signal size. Overall, we achieved 25 times larger signal (or 5 times signal-to-noise ratio), and we can measure $dPNC$ to within 10% of its size in less than half an hour.

IV. PRELIMINARY RESULT

The improvement in sensitivity presented us with both opportunity and challenge. We can now measure $dPNC$ to 0.5% precision, which will provide very stringent test on many ideas of new physics. However, higher precision also requires correspondingly better control of various systematics. One class of systematics concerns the issue of calibration. Imperfect

optical pumping, presence of nearby unwanted 6S-7S transitions or saturation of the main transition can complicate interpretation of dPNC. Another group of systematics is from those rate terms that can mimic the Stark-PNC interference. Most problematic in this regard is the presence of magnetic dipole transition amplitude and its mixing with the Stark amplitude. Also some type of misalignment in E or B can couple with ambient field to produce PNC-like modulation on 6S-7S transition rate.

After two years of intensive study all of these problems are well understood and we can either measure or control them with enough precision and accuracy. We made a series of runs in October and December of 1993 to measure PNC. Our new result was in good agreement with the 1988 result and presently we are taking care of final loose ends before we further accumulate data.

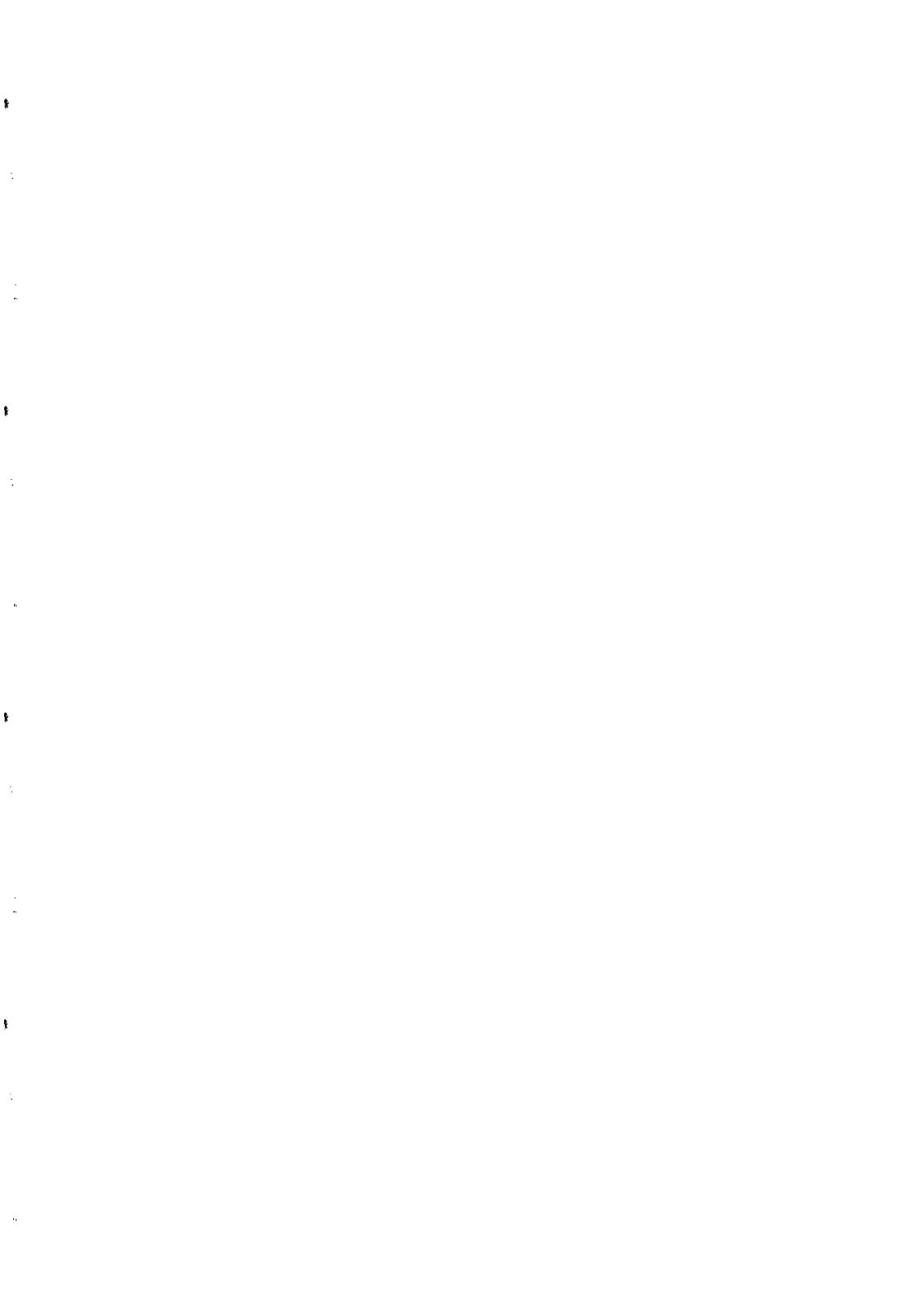
V. CONCLUSION

So far we have focussed our attention on a single hyperfine transition, namely 6S $F=3$ to 7S $F=4$. By simply tuning the lasers to proper transitions we should be able to study all 4 lines between 6S and 7S. PNC measurement on $\Delta F=0$ transitions is new, because application of magnetic field does not remove degeneracy among Zeeman levels when $\Delta F=0$. Comparison among the PNC numbers from different hyperfine transitions are important because they should provide information on parity violating anapole moment of cesium nucleus.

During the last decade there have been revolutionary developments in cooling and trapping atoms. Alkali atoms are the most extensively studied species in this field, and this opens up new experimental possibilities in atomic PNC measurements. High atomic density and pure polarization should give higher signal-to-noise ratio than beam experiment. It will also allow us to make measurements on a string of cesium isotopes. Comparison of PNC among isotopes implies cancellation of atomic matrix element whose uncertainty will soon limit the cesium PNC measurement. We may also measure PNC from francium ($Z=87$) which has 20 times larger PNC effect than cesium, but does not have a stable isotope.

+ current address: Department of Physics, Korea University, Seoul 136-701, Korea

- 1) M.A.Bouchiat and C.Bouchiat, *J. Phys. (Paris)* **35**, 899 (1974); **36**, 493 (1975)
- 2) M.C.Noecker, B.P.Masterson, and C.E.Wieman, *Phys.Rev.Let.* **61**, 310 (1988)
- 3) S.A.Blundell, J.P. Sapiirstein, and W.R.Johnson, *Phys.Rev.D* **45**, 1602 (1992)
- 4) B.P.Masterson, C.Tanner, H.Patrick, and C.E. Wieman, *Phys.Rev.A* **47**,2139 (1992)



HIGH ACCURACY MEASUREMENT OF WEAK INTERACTION INDUCED
OPTICAL ACTIVITY IN ATOMIC LEAD AND THALLIUM

Steve K. Lamoreaux
University of Washington
Physics Department, FM-15
Seattle, WA 98195 USA

A brief review our recent measurements of weak interaction effects in heavy atoms is presented, with emphasis on the technical innovations which have allowed us to obtain this unprecedented precision.

As was originally theoretically shown by Bouchiat and Bouchiat, the weak interaction between electrons and nucleons can lead to observable effects in heavy atoms.¹⁾ This is because the mixing of opposite parity states increases as the atomic number as Z^3 . Each of the following contribute one factor of Z to the enhancement: momentum of the electron at the nucleus, overlap of the wavefunction at the nucleus, and the number of nucleons (for the spin independent part of the weak interaction).

The weak interaction leads to observable effects in the optical properties of atoms since opposite parity states are mixed into $S_{1/2}$ and $P_{1/2}$ states, which have a non-zero overlap with the nucleus. This leads to, for example circular dichroism: the absorption of light depends on the light helicity. By the Kramers-Kronig relations, a helicity dependence of the index of refraction is expected also. Since linearly polarized light is a superposition of left- and right-handed circular polarization, on propagating through a vapor, the linear polarization vector will rotate on propagation through a heavy metal vapor. This effect is enhanced at weak transitions, i.e., magnetic dipole, since the rotation comes about by an interference between the different possible electromagnetic transition coupling the mixed parity state with the excited state. For an allowed $M1$ atomic transition, makes an $E1$ transition between these levels also possible. The experimental observable is

$$R = \text{Im} \frac{E1_{pnc}}{M1} \quad (1)$$

which gives the rotation angle per optical absorption length of atomic vapor. The rotation is frequency dependent (dispersive about line center) and R is a constant which simply sets the amplitude of the dispersive lineshape, all other parameters of this line shape being determined by the absorption profile, in principle. For heavy atoms such as Pb and Tl, the effect is of order 10^{-7} rad/ absorption length.

R is measured by simply directly measuring the rotation angle of plane polarized light as a function of frequency on propagation through a dense atomic vapor. Such measurements have been performed on Bismuth and Lead,²⁾ and more recently Thallium.³⁾ Precision measurements (of order 1% accuracy) on Bismuth⁴⁾ at Oxford, and on Lead⁵⁾ at the University of Washington have recently been completed. A high precision measurement of Thallium is currently in progress at the University of Washington. The technical innovations which led to the precision of our Lead and Thallium work will be the main focus of this paper. Further details concerning this work can be found in Ref. 5).

II. Interpretation of R

The magnitude of R is proportional to the weak charge,

$$Q_W \approx -N + (1 - 4 \sin^2 \theta_W)Z \quad (2)$$

where N and Z are the neutron and proton numbers, and θ_W is the weak angle. There are, in addition, radiative correction and possible new physics contained in Q_W , as we will address shortly.

In order to determine Q_W from an experimental determination of R , accurate atomic calculations are required. Currently, the best heavy atom calculations are in Cesium, which are now better than 1%, while the experimental accuracy is about 2%. In Lead and Thallium, our experimental results are better than 1%, while the atomic theory is at the 8% and 3% level, respectively.⁶⁾ There is hope to improve the Thallium calculations, as we will discuss.

One of the most useful ways to compare the low-energy tests of the Standard Model which atomic parity experiments provide with High Energy results is through modifications of Q_W parameterized as follows:⁷⁾

$$Q_W = (.9857 \pm .0004)(1.00782T)(-N + Z[1 - (4.012 \pm .010)]\bar{x}) + Q_{new}^{tree} \quad (3)$$

$$\bar{x} = 0.2323 \pm .0007 + .00365S - .00261T \quad (4)$$

where S and T represent weak isospin conserving and breaking effects, respectively, and possible new tree level interactions are included. In the case of heavy atoms, where $Z \approx \frac{2}{3}N$,

$$Q_W \approx 0.06T + 1.20S - N \quad (5)$$

$$\frac{\partial Q_W}{\partial N} = .00365S - .00261T. \quad (6)$$

A 1% accuracy measurement plus theory of atomic weak interaction effects determines S to an accuracy ± 0.9 .

In the atomic theory calculations, nuclear size effects enter at the 0.3% level.⁸⁾ It has been suggested that a comparison of a string of isotopes can be compared⁹⁾ and the atomic theory uncertainties be eliminated, however, the nuclear structure uncertainty remains at some level. In addition, much higher experimental accuracy is required for deviations from the Standard Model enter as a fractional difference in Q_W between isotopes. An isotope comparison determines S and T in nearly the same linear combination as results of high energy experiments⁷⁾ whereas a single isotope plus theory result determines S , as can be seen by Eqs. (5) and (6). So far, atomic parity experiments and high energy experiments show $S < 0$ at about 1σ confidence; the Standard Model and extension thereof predict $S \geq 0$.

There are also nuclear spin dependent weak interaction effects. In this case, the electron-nucleon direct weak interaction is on order $(1 - 4 \sin^2 \theta_W)/Z$ times smaller than the spin-independent part. However, the weak interaction within the nucleus, between the nucleons, can give rise to an internal electromagnetic moment of the nucleus, the so-called anapole moment,¹⁰⁾ with which the electron interacts. Such an effect would appear as a hyperfine-level dependence on the observed value of R . The effect is expected to be relatively large in Thallium¹¹⁾, and a primary goal of our work is a reliable measurement of this.

III. Recent Technical Innovations

A. The Laser

The heart of our experiment is an external cavity semiconductor diode laser. Since both the Thallium and Lead $6p_{1/2} - 6p_{3/2}$ magnetic dipole transition frequencies are near 1280 nm, the same laser is used for both elements.

The first version of the laser used a Mitsubishi ML7011R InGaAsP diode, rated at 5 mW output power with wavelength 1300 ± 20 nm. The units we obtained have a center frequency of 1310 nm.

The laser can was carefully opened, and the output facet of the semiconductor chip was antireflection coated with antimony trioxide, which has a nearly perfect index (1.86) for such a coating. Although the coating is less robust than, for example, sapphire or silicon oxide, the evaporation temperature of 700 C makes the coating process much simpler, reliable, and safer. The spontaneous output power from the opposite facet is monitored during the evaporation process, and a minimum indicates a perfect quarter wave coating (at the center of the gain curve). Measurements of the spontaneous emission spectra of lasers coated by this technique indicate a residual facet reflectivity of less than 1%.

Since the ML7011R laser is contained in a can, light from only one facet is available. A Melles-Griot 6 mm focal length, $f=0.7$ lens is used for collimation. The external cavity is 6 cm long, and the light is fed back with a 1200 line/mm grating (750 nm blaze) in the Littrow configuration. Light is extracted from the cavity with a 10% cube beamsplitter. Light from the direction of the laser chip is used in the experiment; tuning of the laser, accomplished by a simultaneous displacement and rotation of the grating, could in principle, result in a movement of the grating-reflected beam. The diode provides a fixed point diffraction limited light source.

In this version of the laser, the diode is cooled to -100 C to shift the gain curve to 1280 nm. The diode and collimating lens are contained in a vacuum vessel, with a window at Brewster's angle to pass the light out of the vessel, and the grating and beamsplitter are outside the vessel. The cavity is constructed of Invar rods. The net output power is about 0.5mW, and is continuously tunable over about 20 GHz. The intrinsic linewidth is less than 1 MHz, however, boiling of the liquid nitrogen used for cooling causes low frequency vibration, resulting in a technical linewidth of about 5MHz, with the deviations at relatively low frequency.

The second version of the laser is built around a STC Defence Systems LYC7*M2 semiconductor diode which had a peak wavelength of 1305 nm as obtained from the company. For this laser, both facets are available for output light; one facet is slightly blocked by the chip mounting, however, this only results in a factor of two loss, and the beam splitter in the

external cavity is no longer required. Amazingly, we were able to tune this laser to 1280 nm at room temperature without a crippling loss of power. The net output power is about 1 mW (as measured from the partially blocked facet) when the light is collected with an f2.8 lens; this indicates an internal power of about 5 mW, to be compared with the original power of about 10 mW at the same current. This laser is single-mode continuously tunable over 30 GHz and has a total linewidth of less than 2 MHz.

Since the piezoelectric transducers used for tuning are nonlinear and have considerable hysteresis, a Fabry-Perot cavity with 15 cm mirror spacings is used to provide frequency markers as the laser is scanned over the atomic transitions. In the data analysis, these markers are used to determine the frequency axis through a linear and quadratic fit.

Since the laser linewidth is so narrow, its effect on the line shape is insignificant, and the data analysis is thus simplified.

B. Detection Electronics

A major advance over all previous measurements of parity violation in this wavelength range was the substitution of Germanium photodetectors with InGaAsP PIN photodiodes. The Germanium detectors were slow, noisy, and had to be cooled, whereas the PIN photodiodes are very quiet, extremely fast, and of extremely low noise. We are currently using a Mitsubishi PD7006 PIN diode which has a quantum efficiency of nearly 100%.

The photocurrent must be converted to a voltage; Johnson noise in the resistor used to do this is an important issue, as is the bandwidth of the amplifier. With a feedback resistor of 50 M Ω , Johnson noise and shot noise contributions to the angle noise are about equal when about 60% of the incident light is absorbed by the atoms. At higher intensities, the noise is dominated by vibration of the optical table and air currents.

In order to have a frequency response greater than 2kHz, the input/output capacitance of the amplifier must be less than 2pF; this is much smaller than can be obtained in any integrated circuit amplifier, so we built our own discrete component amplifier. A low-noise FET cascode amplifier was found to be most effective, with the input FET a 2SK117, and the follower a 2N4416, and is based on a design by F. Walls of NIST, Boulder. The noise was measured to be about $1 \text{ nV}\sqrt{\text{Hz}}$, $0.5 \text{ fA}\sqrt{\text{Hz}}$, with an input capacitance of less than 20 pF, an unmeasurable input/output capacitance, and input resistance of over $10^{12}\Omega$.

Another important feature of the detection is the elimination of black-body light from the oven. This is accomplished by use of a diffraction grating; the laser beam is diffracted with high dispersion after the polarization analyzer; 60% of the incident light is diffracted, after which it travels about 1 m to a spherical mirror; the spherical mirror brings the beam to a focus at 1 m, where there is a 1 mm pinhole. The net effect is a filter with about 1 nm width and 60% transmission efficiency, which reduces blackbody light to less than 5% of the total signal. This also reduces the detected laser spontaneous emission light.

C. Background Rotation Elimination

Due presumably to multiple paths of reflected and scattered light, coupled with residual birefringence of the optical elements, the light polarization angle shows a background frequency dependence of magnitude $1 \mu\text{rad}$ which must be eliminated from the atomic signal if an accurate value of R is to be obtained. Since the metal and metal vapor is contained in a tube, heated at the center and cooled at the ends (heat-pipe oven), and there are no windows on this tube, the metal vapor can be easily moved out of the laser beam without disturbing any of the optical elements, and the background rotation can be easily measured and subtracted from the atomic signal.

For our apparatus, I proposed a "Gatling Gun" configuration, which allows the possibility of comparing several isotopes. A number of vapor tubes, each with a particular isotope, could be measured in sequence, by simply rotating consecutive "barrels" into the beam. So far, we have only employed a "double-barrel" configuration, with one tube empty, which is used for background measurement. The motivation for measuring separated isotopes in Lead or Thallium seems to be waning.

D. Calibration

We have two independent means of calibrating the apparatus: first, we have very accurately measured the rotation angle as a function of current applied to a coil on the Faraday modulator glass. This was done mechanically. Second, there is a very accurately wound coil on the vapor tube, which allows a precisely known magnetic field to be applied to the atoms (better than 0.2% accuracy). The vapor tube coil is constructed of 1 mm diameter Molybdenum wire, cemented in place with an alumina-based ceramic adhesive. Expansion of the tube on heating is taken into account. The resulting atomic Faraday signal from the accurately known magnetic field provides a calibration. We find that the atomic calibration is more reliable than the mechanical; however, an important systematic check is the consistency of R using the two calibrations. The mechanical calibration results show a larger scatter in R , but with unshifted central value. We have found that the mechanical calibration is sensitive to the optical path, hence is less reliable—the atomic Faraday calibration gives the true effect of atoms on the light, in which case effects of the optical elements are irrelevant.

IV. Future Prospects

We are considering a new generation rotation experiment based on using an atomic beam. The idea is to send the laser light through the beam, transverse to the velocity. This eliminates the first order doppler shift to some degree of precision, hence the absorption line narrows, and the absorption, for a given number of atoms, increases. The principal advantages are the line narrowing, which might allow the various isotopes present in natural Pb or Tl to

be independently resolved, and the optical path length is considerably reduced, which might increase the system stability. The background rotation angle effects are over rather larger frequency scales, so problems associated with such might be substantially reduced. In order for a beam technique to be useful (high enough density) several supersonic Tl sources would have to be used.¹²⁾

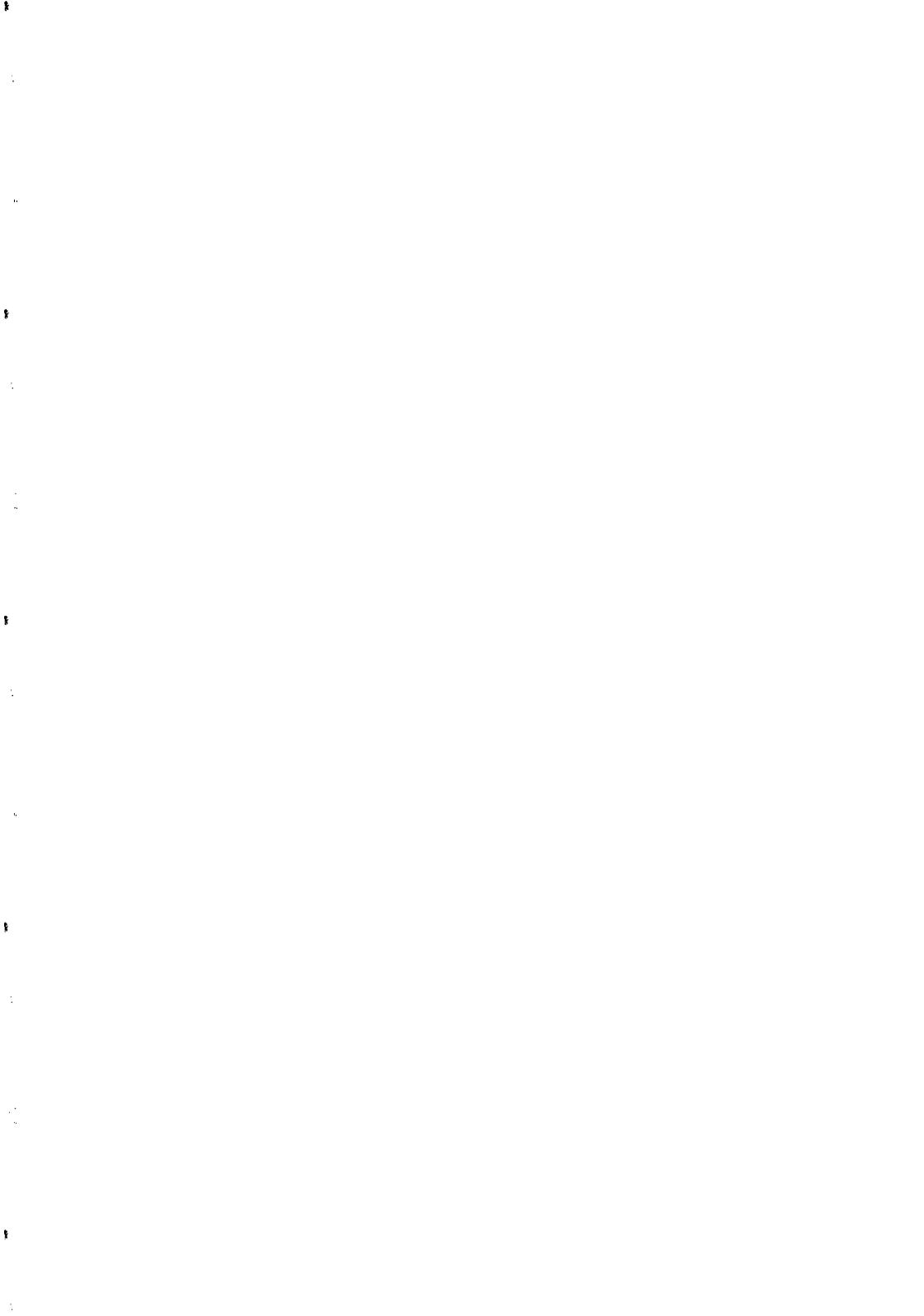
V. Conclusion

We have completed a measurement of R for natural Lead with 1% accuracy; this is currently the most accurate measurement of parity violation in any atomic system. We have nearly completed a measurement in Thallium, and we expect slightly better accuracy. Atomic theory for Lead is at the 8% level and there is little hope of improvement here. For Thallium, theory is at the 3% level and an improvement is anticipated. Already, A.-M. Martensson has calculated the electric quadrupole amplitude, which agrees to better than 1% with our experimental measurement of this parameter.

Most interesting is the possibility to measure the Thallium anapole moment, expected to be relatively large. Thallium has the additional advantage over Lead in that both naturally occurring isotope have nuclear spin 1/2, whereas Lead has only 207 with nuclear spin 1/2, all other natural Lead isotopes being even. Even so, the limit on a spin dependent effect in Lead has already determined $g_n < 4$.¹¹⁾

V. References

- 1). M.A.Bouchiat and C.C.Bouchiat, Phys. Lett **48B**, 111 (1974). M.A. Bouchiat, these proceedings.
- 2). For a review of the work as of 1984, see E.N. Fortson and L.L. Lewis, Phys. Rep. **113**, 290 (1984).
- 3). T.D. Wolfenden, P.E.G. Baird, and P.G.H. Sandars, Europhys. Lett. **15**, 7321 (1991).
- 4). M.J.D. Macpherson et al., Phys. Rev. Lett. **67**, 2784 (1991).
- 5). D.M. Meekhof et al., Phys. Rev. Lett. **71**, 3442 (1993).
- 6). V.A. Dzuba, V.V. Flambaum, P.G. Silvestrov, and O.P. Sushkov, Europhys. Lett. **7**, 413 (1988).
- 7). M.E. Peskin and T. Takeuchi, Phys. Rev. D **46**, 38 (1992).
- 8). S.J. Pollock, E.N. Fortson, and L. Wilets, Phys. Rev. C **46**, 2587 (1992).
- 9). D. Cho, these proceedings.
- 10). Ya. B. Zel'dovich, JETP **36**, 964 (1959).
- 11). I.B. Khriplovich, these proceedings.
- 12). D. DeMille, these proceedings.



Nearly-Degenerate Opposite-Parity Levels in Atomic Dysprosium: A Novel System for the Study of Parity Non-Conservation

D. DeMille, D. ~~Budker~~, E. Commins and M. Zolotarev*

Physics Department, University of California, Berkeley, Berkeley, CA USA 94720

*Stanford Linear Accelerator Center, Stanford, CA 94309

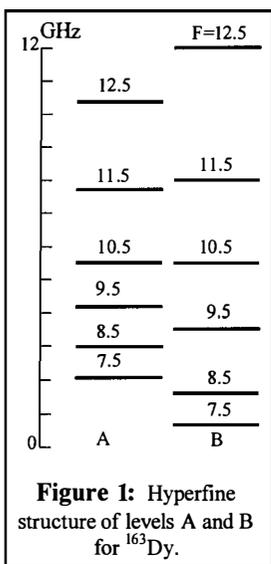
Abstract:

There exists in atomic dysprosium a pair of nearly-degenerate levels of opposite parity. These levels are of potential interest for the study of parity non-conservation because the small energy separation enhances level mixing due to weak-interaction processes. In this paper, we review the unique properties of these levels which make them attractive for measurements of parity non-conservation, and outline the principle of such an experiment now underway.

Introduction

The study of parity non-conservation (PNC) in atoms is now a mature field. Measurements of PNC effects have reached a precision of 2% or better in several elements^{1,2,3}. Measurements with higher precision are of interest for several reasons. First, it should be possible to observe the nuclear anapole moment, which arises from PNC in the nucleus and induces a nuclear spin-dependent component of atomic PNC⁴. In addition, if PNC can be measured to high precision in a chain of isotopes of the same element, information can be obtained about neutron distributions in the nucleus⁵. Finally, if some other method of obtaining neutron distributions is available, measurement of PNC in a chain of isotopes can serve as a high-precision test of the standard model of weak interactions^{6,7}.

Measurement of PNC in the system of nearly-degenerate levels of opposite parity in atomic dysprosium ($Z=66$) offers the possibility in principle to make such a measurement at a level of precision orders of magnitude better than that achieved so far in other elements. This possibility arises because the small energy splitting enhances the weak-interaction induced mixing between the levels. We describe here the unique properties of this system, and will outline the effort now underway to make a first observation of PNC in dysprosium.



The nearly-degenerate levels A (even parity) and B (odd parity) both have angular momentum $J=10$ and lie 19797.96 cm^{-1} above the ground level ($J=8$). (The level structure of atomic Dy is tabulated in Ref. 8.) Dysprosium has both even and odd neutron-number stable isotopes ranging from $A=156$ to $A=164$. The magnitude of the energy splitting between A and B is of the order typically associated with hyperfine and isotope shifts⁹, and thus varies greatly among the various HF and IS components. For the even isotopes (which have nuclear spin $I=0$, and thus no hyperfine structure) this splitting ranges from 235 MHz for ^{162}Dy to 4200 MHz for ^{156}Dy . The two odd isotopes (^{161}Dy and ^{163}Dy) both have nuclear spin $I = 5/2$. Fig. 1 shows the hyperfine structure of levels A and B in ^{163}Dy . The hyperfine components with $F=10.5$ of ^{163}Dy are the closest pair, with a separation of only 3.1 MHz. This pair of levels is used in the current search for PNC.

Outline of the PNC Experiment

In the PNC experiment now underway, we observe Stark-induced quantum beats between levels A and B and look for interference between the Stark amplitude and the much smaller PNC amplitude connecting the two levels. A magnetic field is applied in order to bring Zeeman sublevels of A and B with the same value of m_F to near crossing, and thus to enhance the PNC mixing between them. (Fig. 2 shows some of the Zeeman structure of the ^{163}Dy $F=10.5$ levels of A and B in

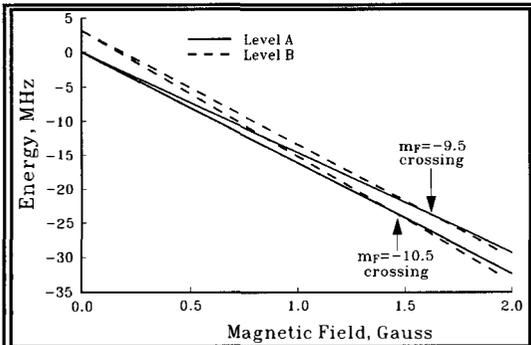


Figure 2: Partial Zeeman structure of the ^{163}Dy $F=10.5$ sublevels of A and B.

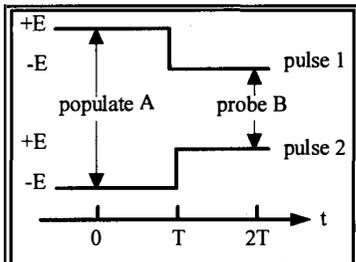


Figure 3: Electric field pulses used in the nonadiabatic switching scheme. The PNC signal is the fractional difference between the signals obtained with these two pulses.

a weak magnetic field.) An electric field is applied parallel to the magnetic field to induce $\sim 100\%$ amplitude of the quantum beats. The Stark and PNC amplitudes have a relative phase of $\pi/2$ and so do not interfere if the electric field is DC. In order to adjust the phases in such a way as to produce interference between these two amplitudes, the polarity of the electric field is rapidly

(non-adiabatically) switched. This scheme of PNC detection was originally proposed in a slightly different form for use in the $2s-2p$ system in hydrogen¹⁰.

Quantum beats are observed by populating level A instantaneously at time $t=0$ and probing level B at some later time $t=2T$. The electric field polarity is switched between the population and probe pulses at $t=T$ (see Fig. 3). Level A is populated by two-step excitation from the ground state via an intermediate odd-parity level with $J=9$; the corresponding E1 transitions are excited by two

consecutive laser pulses at 626 nm and 2614 nm. Level B is probed by excitation to a high-lying even-parity level with another pulsed laser at 571 nm and detection of the subsequent fluorescence. The PNC signal is the fractional difference between the signals from two consecutive population-probe sequences with opposite ordering of the E-field polarity (Fig. 3).

Calculation of the Quantum Beats and the Stark-PNC Interference

The Hamiltonian for the 2-level system formed by the nearly-crossed components of A and B (with the same value of F and m_F) in the presence of an electric field E parallel to the magnetic field is:

$$H = \begin{pmatrix} \frac{-i\Gamma_A}{2} & dE + i\delta \\ dE - i\delta & \Delta - \frac{i\Gamma_B}{2} \end{pmatrix}. \quad (1)$$

Here, Γ_A (Γ_B) is the natural width of level A(B), Δ is the energy splitting in the absence of mixing (determined by the applied magnetic field), $i\delta$ is the PNC matrix element (pure imaginary due to T-reversal invariance), and d is the electric dipole matrix element between these sublevels. The experimentally determined values of Γ_A , Γ_B and d are⁹:

$$\Gamma_A = 20 \text{ kHz}; \quad (2)$$

$$\Gamma_B < 1 \text{ kHz}; \quad (3)$$

$$|d(F=10.5, m_F=10.5)| = 4 \text{ kHz}/(\text{V/cm}). \quad (4)$$

In addition, the PNC matrix element δ has been estimated¹¹ using multiconfiguration Hartree-Fock wavefunctions for Dy:

$$|\delta| = 40 \text{ Hz}. \quad (5)$$

The nonzero value of δ arises from configuration mixing and core-polarization effects: the dominant configurations of A and B differ by the exchange of an f electron for a d electron, whereas the PNC Hamiltonian mixes only s and p states. As a comparison, it can be noted that the PNC matrix element between the $6p_{1/2}$ and $7s_{1/2}$ states of Tl has magnitude ≈ 100 kHz. The electric dipole matrix element is also suppressed, because E1 selection rules are not satisfied in the transition between the dominant terms of A and B; a typical value for an allowed E1 amplitude is $e a_0 \approx 1 \text{ MHz}/(\text{V/cm})$.

The complex eigenvalues $\lambda_{1,2}$ and corresponding eigenstates $\begin{pmatrix} a_{1,2} \\ b_{1,2} \end{pmatrix}$ of the Hamiltonian in eqn. 1 are found by solving the characteristic equation. The linear combination of these eigenstates which corresponds to state A is excited at time $t=0$:

$$\psi(t=0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c_1 \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + c_2 \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}. \quad (6)$$

Just before the electric field switching, the wavefunction has evolved to:

$$\psi(t = T) = c_1 e^{-i\lambda_1 T} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + c_2 e^{-i\lambda_2 T} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}. \quad (7)$$

In the above wavefunction there appears no Stark-PNC interference term, since the PNC matrix element is pure imaginary and the Stark matrix element is real. Non-adiabatic switching of the polarity of the electric field can adjust the relative phases of the PNC and Stark mixing terms to produce interference. After the electric field has been switched, there are new eigenvalues $\Lambda_{1,2}$ and

eigenfunctions $\begin{pmatrix} A_{1,2} \\ B_{1,2} \end{pmatrix}$. Thus, just after the switching, the wavefunction can be written in the form

$$\psi'(t = T) = C_1 \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} + C_2 \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \quad (8)$$

where C_1 and C_2 are determined by setting $\psi'(t = T) = \psi(t = T)$. At the time $t = 2T$ of the probe pulse, the wavefunction is:

$$\psi'(t = 2T) = C_1 e^{-i\Lambda_1 T} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} + C_2 e^{-i\Lambda_2 T} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \quad (9)$$

Since the probe pulse couples only to state B, the signal has the form

$$S \propto \left| C_1 e^{-i\Lambda_1 T} B_1 + C_2 e^{-i\Lambda_2 T} B_2 \right|^2. \quad (10)$$

Fig. 4 shows the measured and expected dependence of the signal as a function of the applied magnetic field near a level crossing.

For $dE \gg \delta$, The signal asymmetry

$$\mathcal{A} = \frac{S(E+ \rightarrow E-) - S(E- \rightarrow E+)}{S(E+ \rightarrow E-) + S(E- \rightarrow E+)} \quad (11)$$

is linear in δ . With T , E , and the magnetic field chosen so that $(1/T) \approx dE \approx \Delta \approx \Gamma_A$, the population of state B at the time of the probe pulse is a considerable fraction of the initial population of state A, and the asymmetry \mathcal{A} is on the order of $(\delta/\Delta) \approx 2 \times 10^{-3}$. This asymmetry changes sign with the overall applied magnetic field and with the detuning Δ ; it has the signature of the P-odd, T-even invariant

$$(\mathbf{E}_i - \mathbf{E}_f) \cdot (\mathbf{B} - \mathbf{B}_c), \quad (12)$$

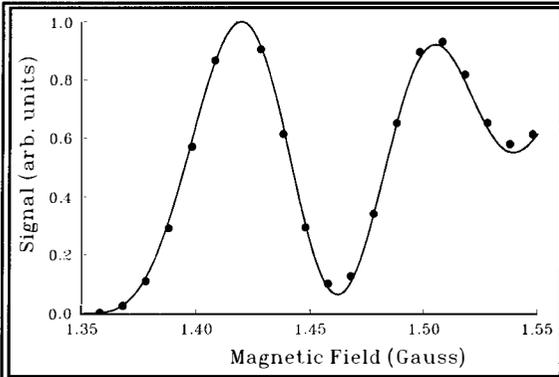


Figure 4: Magnetic field dependence of the nonadiabatic-switching quantum beat signal near the ^{163}Dy $F = 10.5$, $m_F = 10.5$ level crossing, which occurs at 1.46 G. The line shows the expected dependence; the points are experimental data. The conditions are: $E = \pm 7$ V/cm, $T = 4$ μs . The lineshape is asymmetric because of contributions from the nearby $m_F = 9.5$ level crossing.

where the subscripts f and i denote final and initial, and B_c is the magnetic field required to produce an exact level crossing.

Systematics and Sensitivity

A full analysis of possible systematic effects in the PNC measurement is beyond the scope of this paper. Briefly, however, several points can be made. First, it can be noted that the availability of multiple field reversals to change the sign of the asymmetry leads to rejection of

systematic effects due to stray and nonreversing fields. In particular, we find that the use of balanced E-field switching (i.e., with both polarities for each pulse, as in Fig. 3) greatly suppresses spurious effects due to nonreversing E-fields. In addition, the presence of spurious fields can be detected and corrected for by means of numerous auxiliary measurements using the Dy atoms themselves. It is also of interest to note that the magnitude of a stray or non-reversing electric field required to produce mixing as large as the expected PNC mixing is macroscopic: $|\delta/d| \approx 10$ mV/cm. For these reasons, we believe that systematic effects can be controlled to a high level of precision.

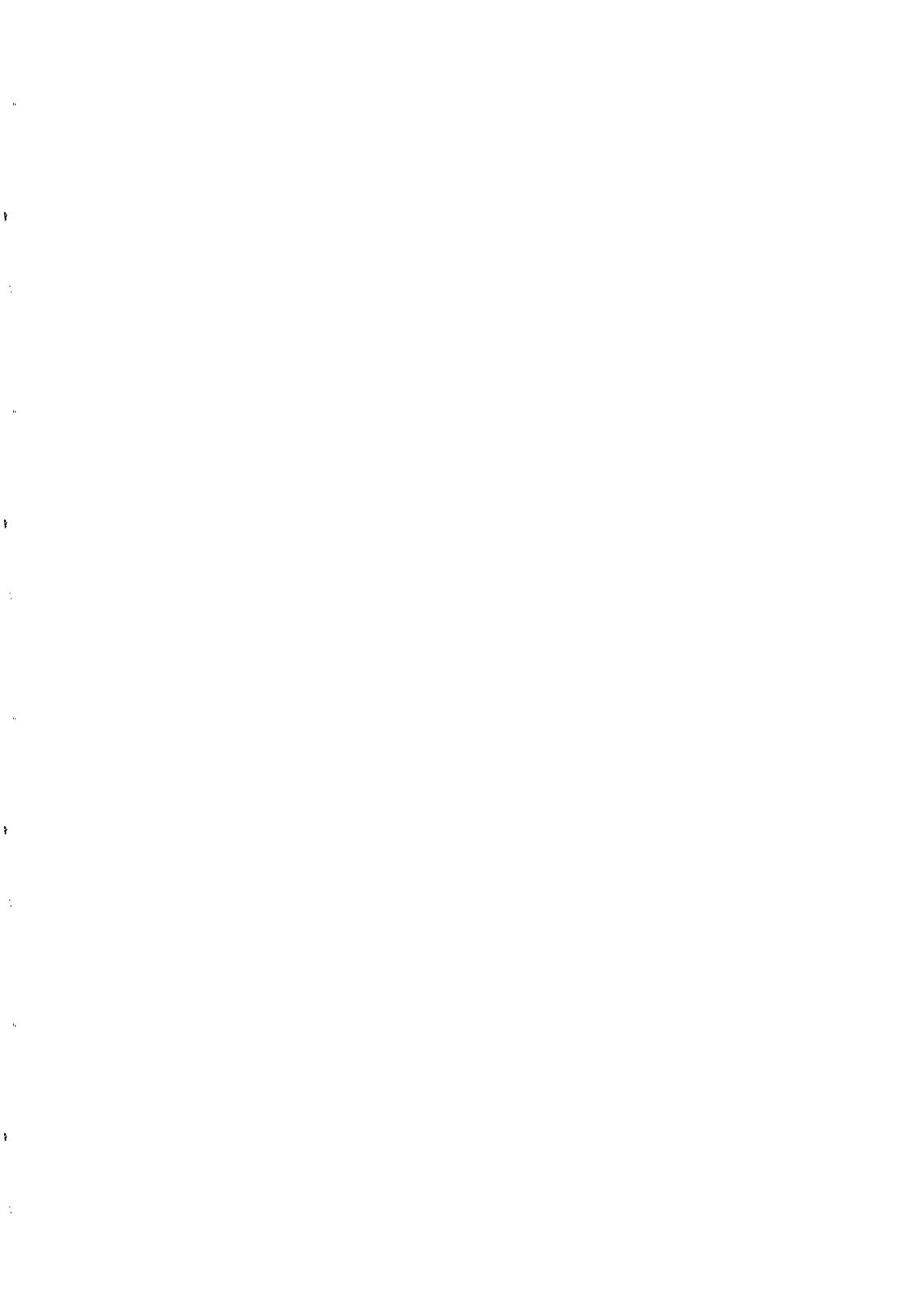
Study of PNC in the nearly-degenerate levels offers the potential for unprecedented levels of statistical precision in the measurement of PNC effects. This potential arises both because the predicted PNC asymmetry for these levels is large ($\mathcal{A} \approx 2 \times 10^{-3}$) and because this asymmetry is obtained using E1 transitions in every stage of the experiment, which means that high counting rates can be achieved. Both properties are important because, in the shot-noise limit, the fractional uncertainty in the PNC matrix element δ is given by $(\mathcal{A}\sqrt{N})^{-1}$, where N is the total number of detector counts.

Current Status and Conclusion

The current statistical sensitivity of our apparatus to the asymmetry \mathcal{A} is $\sim 5 \times 10^{-2} / \sqrt{\text{Hz}}$. This sensitivity is limited by two primary difficulties. First, our laser system is far from optimal for the PNC measurement; the pulsed lasers have a repetition rate of 10 Hz, which gives an effective duty cycle of $\sim 10^{-4}$. Second, pulse-to-pulse fluctuations in the laser intensities and spectral profiles lead to noise far above the shot noise limit. Nevertheless, it should be possible to unambiguously detect PNC at the predicted level with only a few hours of integration time. Before this is attempted, however, we are investigating in detail all possible sources of systematic effects and attempting to reduce the stray and nonreversing fields now present in our apparatus. With control over these fields at levels similar to those obtained in other PNC experiments and with an optimized laser system, it is a reasonable goal to measure PNC in the nearly-degenerate levels of Dy to a precision of 10^{-4} or better.

References

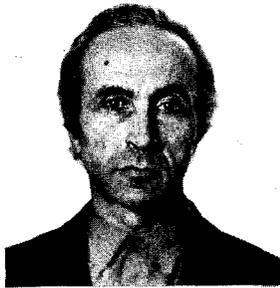
- ¹ M.C. Noecker, B.P. Masterson and C.E. Wieman, *Phys. Rev. Lett.* **61**, 310 (1988).
- ² M.J.D. Macpherson, et. al., *Phys. Rev. Lett.* **67**, 2784 (1991).
- ³ D.M. Meekhof, et. al., *Phys. Rev. Lett.* **71**, 3442 (1993).
- ⁴ V.V. Flambaum and I.B. Khriplovich, *JETP* **52**, 835 (1980); V.V. Flambaum, I.B. Khriplovich and O.P. Sushkov, *Phys. Lett.* **B146**, 367 (1984).
- ⁵ E.N. Fortson, Y. Pang and L. Wilets, *Phys. Rev. Lett.* **65**, 2857 (1990); S.J. Pollock, E.N. Fortson and L. Wilets, *Phys. Rev.* **C46**, 2587 (1992).
- ⁶ V.A. Dzuba, V.V. Flambaum and I.B. Khriplovich, *Z. Phys.* **D1**, 243 (1986).
- ⁷ D.N. Stacey, in *Atomic Physics 13*, eds. H. Walther, T.W. Hänsch and B. Neizert (AIP, New York, 1993), p. 46.
- ⁸ W.C. Martin, R. Zalubas and L. Hagan, *Atomic Energy Levels-The Rare Earth Elements* (National Bureau of Standards, Washington, DC., 1978).
- ⁹ D. Budker, D.DeMille, E.D. Commins and M.S. Zolotarev, submitted to *Phys. Rev. A*; *Phys. Rev. Lett.* **70**, 3019 (1993).
- ¹⁰ R.R. Lewis and W.L. Williams, *Phys. Lett.* **B59**, 70 (1975).
- ¹¹ M. Kozlov, V. Dzuba and V. Flambaum, submitted to *Phys. Rev. A*.



**ATOMIC PHYSICS:
QED, FUNDAMENTAL SYMMETRIES**

I.B. Khriplovich

Budker Institute of Nuclear Physics, 630090 Novosibirsk, Russia



Abstract

This is a summary of the talks on the mentioned subjects presented at the Workshop.

The time allotted for the summary is short, and the number of the talks to be covered is large. I like all of them. I enjoyed those by my theoretical friends, *S. Barr* and *J. Sucher*. Could not but admire the experimental ingenuity with which modern traps for positrons and hydrogen are constructed (*K. Abdullah* and *T. Hijmans*). As to the spin-polarized hydrogen (*I. Silvera*), my fascination by those experiments lasts already for many years.

But I don't wish to make the summary talk a complete commonplace. So, with the hope to be able to say at least something of interest to the audience, I'll confine to comments on three subjects:

1. Precision measurements in hydrogen and positronium
2. Parity nonconservation in atoms
3. CP-violation

1. Precision measurements constitute for a long time the essential feature of atomic physics. Such measurements in hydrogen involving high levels and 1s - 2s transitions were described in the talks by *L. Julien* and *C. Zimmerman* respectively. Besides being a tool for metrology, those experiments are a source of valuable information on the electron-proton mass ratio and on the proton charge radius, they constitute a sensitive test of quantum electrodynamics. And who can predict at all what surprises await us at the next level of precision?

Such a surprise has been brought indeed by precise measurements of the orthopositronium decay rate ($oPs \rightarrow 3\gamma$), as discussed in the talk by *R. Conti*. The experimental value of this decay rate

$$\Gamma_{exp}^{oPs} = 7.0482(16)\mu s^{-1} \quad (1)$$

is by six standard deviations above the theoretical prediction

$$\Gamma_{th}^{oPs} = m\alpha^6 \frac{2(\pi^2 - 9)}{9\pi} \left[1 - 10.28 \frac{\alpha}{\pi} - \frac{1}{3} \alpha^2 \log \frac{1}{\alpha} \right] = 7.03830\mu s^{-1}. \quad (2)$$

And positronium is presumably pure QED system!

To reconcile the theory with experiment staying within QED, one should assume that the correction $\sim (\alpha/\pi)^2$ should be

$$250(40)(\alpha/\pi)^2.$$

which may look unreasonably large. We believe nevertheless that the natural scale for the factor at $(\alpha/\pi)^2$ is about 100. The argument is as follows (I.B. Khriplovich, A.I. Milstein, A.S. Yelkhovsky, *Phys.Scr.* **46** (1993) 252). The large, -10.28 , factor at the α/π correction to the decay rate (see (2)) means that the typical magnitude of the factor at the α/π correction to the decay amplitude is roughly 5. Correspondingly, this correction squared contributes about $25(\alpha/\pi)^2$ to the decay rate. Indeed, numerical calculations by Burichenko (*Yad.Fiz.* **56** (1993) 123 [*Sov.J.Nucl.Phys.* **56** (1993)]) have given factor 28.8(0.2) at $(\alpha/\pi)^2$ in this contribution.

Then, it is only natural to expect that the interference of the second-order radiative correction to the amplitude with the zeroth-order amplitude should contribute about twice as much to the decay rate as the square of the first-order correction. In other words, the natural scale for the total second-order radiative correction to the decay rate should be close to

$$100(\alpha/\pi)^2. \quad (3)$$

One more class of large corrections is of relativistic origin. As *Conti* claimed in his talk,

$$\alpha = \beta \quad (\beta = v/c).$$

But of course it is other way around:

$$\beta = \alpha.$$

And even this is not always true. In particular, in the problem discussed

$$\beta = \alpha i \sqrt{3/4}.$$

Coming back to more serious tone, I wish to say that a preliminary result of calculating relativistic corrections to the orthopositronium decay rate gives for their contribution (I.B. Khriplovich, A.I. Milstein, A.S. Yelkhovsky)

$$\sim 5\alpha^2.$$

As distinct from usual radiative corrections, the relativistic ones do not contain π in denominator (π is almost a dimensional number!). In the common "radiative" form, with π in denominator, this correction constitutes

$$\sim 50(\alpha/\pi)^2. \quad (4)$$

Having in mind estimates (3) and (4), we believe that it is premature to talk about the contradiction between the QED and the experimental result (1) now, until the complete calculation of the second-order correction to the decay rate is done.

In the conclusion of this section it is worth mentioning that terms of relativistic origin absolutely dominate the corrections of the relative order α^4 to the positronium nP levels (I.B. Khriplovich, A.I. Milstein, A.S. Yelkhovsky, Phys.Rev.Lett. **71** (1993) 4323). By the way, in this problem

$$\beta = \alpha/2n.$$

2. Parity nonconservation in atoms has gone already a long way,
from

the pioneering works by Bouchiat and Bouchiat (1974) which transformed the subject from science fiction into science attracting attention to heavy atoms, pointing out the P-odd correlations of experimental interest in strongly forbidden M1 transitions;

from

the first experimental observation of electron-nucleon weak interaction due to neutral currents, by measuring optical activity of bismuth vapour near normal M1 transitions (Novosibirsk, 1978)

to

present precision tests of the electroweak theory: the experimental accuracy in cesium has reached now 2% (Boulder, 1988), the theoretical one is about 1% (Novosibirsk, 1989; Notre Dame, 1990).

Recent precision measurements of the optical activity of lead and thallium vapours with the accuracy 1% in Pb and about 0.5% in Tl were described by *S. Lamoreaux*.

It is certainly a challenge for atomic theory to reach the corresponding accuracy. That would mean not only new independent tests of the standard model, but a new level of precision for those tests. And at present the accuracy of atomic calculations of parity-nonconserving effects is no better than 8% in lead and 3% in thallium.

However, from the talk by *A.-M. Pendrill* we have learned about the impressive progress achieved in the theoretical description of the thallium hyperfine structure constants and E2 amplitude for $6p_{1/2} - 6p_{3/2}$ transition in Tl. This progress makes the accuracy about 1% in calculating parity-nonconserving effects in thallium sufficiently realistic. But to guarantee such a precision the theoretical calculations should be accompanied by as precise measurements of E1 amplitudes in thallium.

New generation of the experiments with cesium, in preparation now, are discussed in the talks by *M.-A. Bouchiat* and *D. Cho*.

Quite curious situation takes place in dysprosium where there are opposite parity levels separated by tiny energy interval. Ingenuous spectroscopic investigations described by *D. DeMill* have shown that the interval between hyperfine components belonging to those levels is in some cases as small as 3 MHz (compare with 10^3 MHz in the hydrogen Lamb shift!). The experiment is being prepared now to investigate parity-nonconserving effects in this system. Theoretical uncertainties do not allow to predict reliably the expected magnitude of those effects. So, the success of such an experiment here is to a large degree a matter of good luck. But at least it is a first-rate atomic spectroscopy.

One more aim of the experiments discussed is the search for nuclear anapole moments. Anapole is a new electromagnetic moment arising if parity is not conserved. Nuclear anapole moment is induced by P-odd nuclear forces. Atomic P-odd effects dependent on nuclear spin are dominated by contact electromagnetic interaction of electrons with nuclear anapole moment.

The problem of experimental observation of a new physical phenomenon, nuclear anapole moment, is by itself a fascinating one. Moreover, if the theoretical results of one-particle approximation, which are remarkably stable by nuclear standards, will be supplemented by a serious treatment of many-body effects, those experimental investigations will give reliable quantitative information on P-odd nuclear forces.

3. Up to now CP-violation has been observed in K-meson decays only. That is why the discussion in the talks by *D. Garreta* and *R. Legac* of CP-odd effects in kaon decays, as studied at LEAR, is so interesting.

The searches for electric dipole moments of neutron, atoms and molecules constitute another source of information on CP-violation. In the talk by *S. Lamoreaux* a record-breaking upper limit on electric dipole moment (EDM) of anything was reported. The measurements of atomic EDM of a mercury isotope result in

$$d(^{199}\text{Hg})/e < 1.3 \cdot 10^{-27} \text{ cm.} \quad (5)$$

The measurements in atomic thallium, as reported by *S. Ross*, brought the following upper limit on electron EDM:

$$d(e)/e < 10^{-26} \text{ cm.} \quad (6)$$

New experiment with xenon is being prepared by *T. Chupp*. I learned also, but in private discussions only, from *E. Hinds* and *A. Weiss* about their new projects. Let us hope that at the next Moriond Workshop we will hear exciting news from these experiments.

But what is the real significance of the limits obtained? How do they relate to, say, the results of studying the neutron EDM? The answer to this question is given in Table where the limits on the constants of an effective Hamiltonian

$$H = \frac{G}{\sqrt{2}} \sum_i k_i O_i$$

($G = 10^{-5}/m_p^2$ is the Fermi constant, m_p is the proton mass) for the CP-odd interaction of u-, d-quarks and gluons are presented (V.M. Khatsymovsky, I.B. Khriplovich, Phys.Lett. **B296** (1992) 219).

Clearly, the atomic experiments are as informative as the neutron ones. In fact, they are complementary to each other. But still, what is the real value of both?

$k_i O_i$	$d(n)/e < 10^{-25} \text{ cm}$	$d(^{199}\text{Hg})/e < 1.3 \cdot 10^{-27} \text{ cm}$
$k_s(\bar{q}_1 i \gamma_5 q_1)(\bar{q}_2 q_2)$	$ k_s < 3 \cdot 10^{-5}$	$ k_s < 3 \cdot 10^{-6}$
$k_s^c(\bar{q}_1 i \gamma_5 t^a q_1)(\bar{q}_2 t^a q_2)$		
$q_1 = q_2$	$ k_s^c < 10^{-4}$	$ k_s^c < 10^{-5}$
$q_1 \neq q_2$	$ k_s^c < 10^{-4}$	$ k_s^c < 2 \cdot 10^{-3}$
$k_i(1/2)\epsilon_{\mu\nu\alpha\beta}(\bar{u}\sigma_{\mu\nu}u)(\bar{d}\sigma_{\alpha\beta}d)$	$ k_i < 10^{-5}$	$ k_i < 2 \cdot 10^{-4}$
$k_i^c(1/2)\epsilon_{\mu\nu\alpha\beta}(\bar{u}\sigma_{\mu\nu}t^a u)(\bar{d}\sigma_{\alpha\beta}t^a d)$	$ k_i^c < 10^{-5}$	$ k_i^c < 2 \cdot 10^{-4}$
$k_q^g m_p \bar{q} \gamma_5 \sigma_{\mu\nu} g G_{\mu\nu}^a t^a q$	$ k_q^g < 2 \cdot 10^{-7}$	$ k_q^g < 5 \cdot 10^{-8}$
$k^g(1/6)\epsilon_{\mu\nu\alpha\beta} f^{abc} G_{\mu\nu}^a G_{\alpha\beta}^b G_{\beta\rho}^c$	$ k^g < 5 \cdot 10^{-5}$	$ k^g < 3 \cdot 10^{-2}$
$\theta(\alpha_s/8\pi)(1/2)\epsilon_{\mu\nu\alpha\beta} G_{\mu\nu}^a G_{\alpha\beta}^a$	$ \theta < 3 \cdot 10^{-10}$	$ \theta < 10^{-9}$

Table

Although dipole moments have not been discovered up to now, the limits obtained in those experiments have played an extremely important role, allowing one to exclude a number of the models of CP-violation.

In particular, to explain CP-violation in kaon decays by the so-called θ -term (see the last line of Table) one should assume $\theta \sim 10^{-3}$ which is definitely excluded.

The model where spontaneous breaking of CP-violation in the Higgs sector is the only, or main, source of CP-odd effects in kaon decays (Weinberg model of CP-violation) is also excluded. Moreover, present upper limits on $d(n)$ and $d(e)$ are close to the predictions of the model of spontaneous breaking of CP-violation in the Higgs sector at the "natural" choice of its parameters.

The conclusion is clear. Atomic experiments give valuable information on elementary particle physics NOW, and will give even more valuable one in future.

WHAT DO WE KNOW ABOUT THE TL NUCLEUS FROM ATOMIC SPECTROSCOPY?

Ann-Marie Mårtensson-Pendrill, Department of Physics
Chalmers University of Technology and Göteborg University
S-412 96 Göteborg, Sweden, e-mail: f3aamp@fy.chalmers.se

ABSTRACT

The calculation of accurate electronic properties for low-lying states of Tl would be helpful for the interpretation of experimental results, e.g. of isotope shifts and parity non-conserving effects. Such calculations are, however, quite challenging, not only due to the high Z , but also because of the valence structure: Although the 6p and 7s states studied in this work can formally be described as a valence electron outside a closed shell core, the presence of two 6s electrons in the core leads to very large perturbations, which are not adequately described by low-order perturbation theory. The coupled-cluster approach provides a systematic way to sum certain terms in perturbation theory to infinite orders. In this work single and double excitations involving the core orbitals 4f, 5p, 5d and 6s have been included. The wavefunctions have been used to calculate hyperfine structure including anomalies, field isotope shift constants, electric dipole and quadrupole transition matrix elements. The extracted properties include changes in charge and magnetic radii, from isotope shifts and hyperfine anomalies. It was found, e.g., that the change $\delta\langle r_m^2 \rangle$ in magnetic radius between the two stable isotopes ^{203}Tl and ^{205}Tl is significantly larger than the corresponding change in charge radius. This may have some implications for P and T violating Schiff moments, discussed elsewhere in these proceedings.

1. Probing the nucleus

Is it not fantastic that a valence electron outside the 80 core electrons in Tl can probe the nucleus? Of course, the valence electron isn't always outside the core electrons, and it even has a finite probability of being within the nucleus, in particular if it has $j=1/2$, like the $6p_{1/2}$ ground state of Tl. Also valence electrons with $j>1/2$, which have no significant penetration of the nucleus, can perturb the core electrons that do and in this way probe the nucleus. The correlation between valence and core electrons must thus be well described in order to extract nuclear properties from spectroscopic measurements.

The study of *nuclear effects in atomic structure* has a long tradition, particularly in the case of hyperfine structure (hfs) and isotope shift¹⁾ and more recently involving parity non-conserving effects^{2,3)}, with or without accompanying T violation. There is accordingly a long tradition of using semi-empirical electronic factors to extract the nuclear properties from the atomic spectroscopy data¹⁾. However, the development of relativistic atomic many-body techniques⁴⁻⁷⁾ makes it possible to perform accurate *ab initio* calculations of such factors and also to examine the validity of the semi-empirical evaluations^{8,9)}. The improved accuracy enables a more reliable extraction of e.g. nuclear charge radii and quadrupole moments from optical isotope shifts and electric hfs. Since magnetic moments can often be determined independently¹⁰⁾, the magnetic hfs can instead be used to test the quality of the wavefunctions.

In this work we describe the application of relativistic "*coupled-cluster*" calculations to a few low-lying states of Tl, following the procedures applied earlier^{11,12)} to calculations for Ba⁺ and Yb⁺. The calculations presented here include single and double excitations with respect to the unperturbed core of Tl⁺ involving excitations of the core orbitals 4f, 5p, 5d and 6s up to a maximum angular momentum of 6 for the core-valence pair excitations and a maximum angular momentum of 4 for the core-core excitations. The orbitals were expanded in a logarithmic grid with 76 points in the range $\exp(-6.6)/Z$ to $\exp(8.4)/Z$. The energy eigenvalues obtained in this procedure are $\epsilon_{6p_{1/2}} = -0.22458$ a.u., $\epsilon_{6p_{3/2}} = -0.18980$ a.u. and $\epsilon_{7s} = -0.10514$ a.u. (including Dirac-Fock-Breit contributions of 0.00052 a.u., 0.00015 a.u. and 0.00005 a.u. respectively) compared to the experimental values $\epsilon_{6p_{1/2}} = -0.22446$ a.u., $\epsilon_{6p_{3/2}} = -0.18896$ a.u. and $\epsilon_{7s} = -0.10382$ a.u. The resulting wavefunctions were used to evaluate properties, such as electric dipole and quadrupole transition rates, hyperfine structure, isotope shifts and hyperfine anomalies.

Table 1. Comparison between calculated hyperfine structure constants A for ^{205}Tl and electronic factors F for the field isotope shift with experimental values.

	F($6p_{1/2}$) (GHz/fm ²)	A($6p_{1/2}$) (GHz)	F($6p_{3/2}$) (GHz/fm ²)	A($6p_{3/2}$) (GHz)	F(7s) (GHz/fm ²)	A(7s) (GHz)
This work						
DF	-2.70	17.68	0	1.304	-4.36	7.78
BO	-0.34	2.24	0	0.21	-1.25	2.22
RPA	13.60	-1.98	11.40	-1.64	0.69	2.88
Corr	-0.38	-1.04	0.05	0.39	0.69	-0.20
Total:	10.10	20.86	11.35	0.256	-4.24	12.67
+inner RPA ^{a)}	10.13	21.43	11.43	0.317	-4.22	12.92
+Finite size	9.52	21.3	10.74	0.34	-3.97	12.8
Other calculations						
SCDF ^{b)}	12.32	18.73	13.21	1.381	-4.37	
MCDF ^{c)}	15.77	20.32 ^{c)}	17.31	1.485 ^{c)}		
g-Hartree ^{d)}		20.89		0.895		
BO+RPA ^{e)}	16.95 ^{f)}	24.06 ^{g)}	17.04 ^{f)}	-1.885 ^{g)}	-5.08 ^{f)}	13.06 ^{g)}
		21.77 ^{h)}		-1.919 ^{h)}		12.47 ^{h)}
+ "int. corr." ⁱ⁾		21.32 ^{h)}		0.600 ^{h)}		12.56 ^{h)}
Experiment (GHz)	10.6(6) ^{j)}	21.3108	11.4 (6) ^{j)}	0.2651	-3.64(3) ^{j)}	12.297

a) This line gives an estimate of the RPA correction from core orbitals not included in the coupled-cluster calculation. b) Hermann *et al*¹⁵⁾.

c) Grexa *et al*¹⁶⁾, The MCDF hfs values include increments from separate MCDF results.

d) Millack¹⁷⁾ e) Brueckner orbitals + RPA corrections

f,g) Hartley and Mårtensson-Pendrill, refs 14,18 h) Dzuba *et al*⁴⁾ i) BO+RPA+"internal correlation"

j) The estimated field isotope shifts, 1.216(0)(58) GHz, 1.324(1)(49)GHz, and -0.419(2)(27)GHz, respectively, are taken from Ref 14, using the experimental level shifts by Grexa *et al*¹⁶⁾ The first error bar gives the experimental error, the second reflects the uncertainty in the SMS. The "experimental" F values were obtained by combining these estimates with the muonic value¹⁹⁾ $\delta\langle r^{-2} \rangle_{203,205} = 0.115 (3) \text{ fm}^2$.

2. Hyperfine Structure and Isotope Shifts

The magnetic interaction between the electrons and the nuclear magnetic moments gives rise to the magnetic *hyperfine structure*. In the case of the hydrogen ground state, this results in the well-known 21-cm line observed abundantly in radio-astronomy. The hyperfine structure of the ground state of ^{133}Cs constitutes the

definition of the second and many other hyperfine structures, e.g. that of the Tl $6p_{1/2}$ ground state¹³), have been determined very accurately using radio-frequency techniques.

Different isotopes of an element have slightly different transition energies. For light nuclei, this "*isotope shift*" is mainly due to the kinetic energy associated with the nuclear motion. The "reduced mass effect", leading to the "Bohr" or "normal" mass shift is directly proportional to the transition energy and is thus easily evaluated. This is the shift observed, e.g. between the $Ly\alpha$ lines for H and D. For atoms with more than one electron, the nuclear kinetic energy leads also to a cross terms between individual electron momenta, the so-called "mass polarisation" or the "specific mass shift", proportional to $\sum_{i \neq j} \mathbf{p}_i \cdot \mathbf{p}_j$. This is a two-body operator, which is found to be very sensitive to correlation effects¹⁴). It is nearly as large as the normal mass shift for the 2s-2p transitions in ${}^6\text{Li}$ and ${}^7\text{Li}$.

For heavy atoms, the isotope shift is instead dominated by the "*field shift*" or "volume shift", $\delta v^{FS} = F\kappa \delta \langle r_c^2 \rangle = F\lambda_c$ arising due to differences in the charge distribution for different isotopes of an element, and one reason for studying isotope shift in series of isotopes is then to obtain information about the nuclear charge distribution. In the case of Tl, the mass dependent effects account for only about a percent of the observed isotope shift¹⁴), but is still an important source of uncertainty in the extraction of charge radii. The calculated electronic factors then provide a scale relating observed isotope shift to changes in nuclear charge radii, as discussed in detail in our earlier studies of $\text{K}^{(8)}$, $\text{Ca}^{+(9)}$, $\text{Cs}^{(14)}$, $\text{Ba}^{+(11)}$, and $\text{Yb}^{+(12)}$.

The contributions to the calculated hyperfine structure A-factors and electronic factors, F, for the field isotope shifts are shown in Table 1 together with other theoretical^{4,14-18}) and experimental^{13,15-16,19-20}) results. A detailed description of the terms included in the individual terms can be found in Ref. 21. The calculated $A(6p_{1/2})$ factor is very close to the experimental result. The $A(7s)$ factor, on the other hand is overestimated by nearly 4%. The $6p_{3/2}$ state is more problematic: Many body corrections reduce the $6p_{3/2}$ hfs to about 1/5 of its Dirac-Fock value. It is then hardly surprising that the final result for this state is quite uncertain. The estimate of the contribution from perturbation of inner core orbitals amount to nearly 20%, thereby destroying the previously good agreement. Although the result obtained here seems to be the best calculated value to date, an even more detailed treatment is required, probably involving either three-particle excitations or a more general model space as a starting point. Our calculated F factors for the field shift are essentially consistent with the change in charge radius $\delta \langle r_c^2 \rangle = 0.115(3)\text{fm}^2$ found in experiments on muonic atoms.¹⁹) Due to the uncertainty in the calculated electronic factors, in particular for the specific mass shift, the values extracted from accurate laser spectroscopic experiments cannot yet

compete with the muonic determination for the stable Tl isotopes. Nevertheless, the calculated F factors provide a scale for extracting $\delta\langle r_c^2 \rangle$ from the isotope shift measurements on chains of radioactive isotope Tl isotopes.²²⁾

3. Hyperfine anomalies

Changes in the distribution of charge, but also of magnetic moments cause a shift, $\Delta_{12}=(A_1g_2)/(A_2g_1)-1$, in the ratio between the measured hfs A-factor and the nuclear g-factor ($g_1=\mu_1/I$), arising from changes in the wavefunction normalization close to the nucleus, and from modified expressions for the hyperfine interaction, respectively. The form of the interaction is different for magnetic moments arising from nuclear spin and orbital moment, as first analysed in detail by Bohr and Weisskopf²³⁾, and later, e.g., by Stroke *et al.*²⁴⁾ We find that the anomaly can be written as $\Delta_{12}=b_{2s}\lambda_{c,m}=b_{2s}(\lambda_m\pm 1.9(1)\lambda_c)$, where $\lambda_c=\kappa_c\delta\langle r_c^2 \rangle\approx 0.94\delta\langle r_c^2 \rangle$ for Tl²⁵⁾ was introduced in Sect. 2 and

$$\lambda_m=\kappa_m\delta\langle r_m^2 \rangle=\delta\langle r_m^2 \rangle(1+(b_{4s}/b_{2s})\delta\langle r_m^4 \rangle/\delta\langle r_m^2 \rangle+\dots)$$

with $\kappa_m\approx 0.81$ for a magnetic moment arising mainly from a nuclear spin, as expected for the two stable Tl isotopes, which both have an unpaired $3s_{1/2}$ proton in the ground configuration. The numerical Dirac-Fock orbitals used in our evaluation were obtained in the potential from a homogeneous nuclear charge density and give $b_{2s}(6p_{1/2})\approx -2.26\cdot 10^{-4}/\text{fm}^2$ and $b_{2s}(7s)\approx -7.95\cdot 10^{-4}/\text{fm}^2$. To a good approximation, these coefficients depend only on the angular momentum, but are independent of the orbital energy. The correction due to higher moments of the distribution are essentially the same: for both these states we find $b_{4s}/b_{2s}\approx -0.0032/\text{fm}^2$ and $b_{6s}/b_{2s}\approx 8.8\cdot 10^{-6}/\text{fm}^2$.

In Table 2, we have given the correction (expressed in terms of b_{2s}) due to a distributed nuclear moment for the various contribution to the hyperfine structure (shown in Table 1) for the 6p and 7s states. These corrections reflect the interplay between admixtures of various angular momenta into the wavefunction. For the $6p_{1/2}$ state, the admixture of s orbitals with larger anomalies gives a 10% increase of the final hfs anomaly, whereas for the $6p_{3/2}$ state, where the individual corrections to the hfs are many times larger than the final result, the final hfs anomaly is even larger than for the s state (and with opposite sign). These results can be combined with experimental data for $^{203,205}\text{Tl}$, as shown in Table 2, and with the known change in charge radius, to extract a change in magnetic radius, $\delta\langle r_m^2 \rangle\approx 0.26(2)\text{fm}^2$, between the two stable Tl isotopes. The error bar includes an uncertainty due to the correction of higher moments of the distribution. We see that

this change is more than twice as large as that for the charge distribution. The distribution of a possible P and T violating nuclear *electric* dipole moment may be similar to that of the magnetic dipole distribution, and its difference from the charge distribution gives essentially the "Schiff moment".³⁾ The results shown here give no information about the size of a possible Schiff moment, but indicate a difference between the two stable Tl isotopes and may, in this way, provide a calibration for nuclear physics calculations.

Table 2. Hyperfine anomalies: The theoretical values give the relative effect of the magnetic moment distribution on the various terms shown in Table 1 and are given in terms of b_2 factors, i.e. Δ/λ .

Calculated ($10^{-4}/\text{fm}^2$)	$6p_{1/2}$ $\Delta(6p_{1/2})/\lambda$	$6p_{3/2}$ $\Delta(6p_{3/2})/\lambda$	$7s$ $\Delta(7s)/\lambda$
DF(BO)	-2.26	0	-7.95
RPA	-4.89	-5.02	-5.86
Corr	-4.12	-4.24	-1.63
Total	-2.48	43.0	-7.62
Experiment (GHz)	$A(6p_{1/2})^a$	$A(6p_{3/2})^b$	$A(7s)^c$
^{203}Tl ($\mu=1.62225787\mu_N$)	21.105447(5)	0.2620300(1)	12181.6(22)
^{205}Tl ($\mu=1.63821461\mu_N$)	21.310835(5)	0.2650383(1)	12297.2(16)
Anomaly $\Delta(6p_{1/2})$ (10^{-4})	$\Delta(6p_{3/2})$ -1.04	$\Delta(7s)$ 16.26	-3.4 \pm 1.8
λ (fm^2)	0.42	0.38	0.44(23)

a) Lurio and Prodell (1956)¹³⁾ b) Gould (1956)²⁰⁾ c) Hermann et al (1993)¹⁵⁾

4. Discussion and Outlook

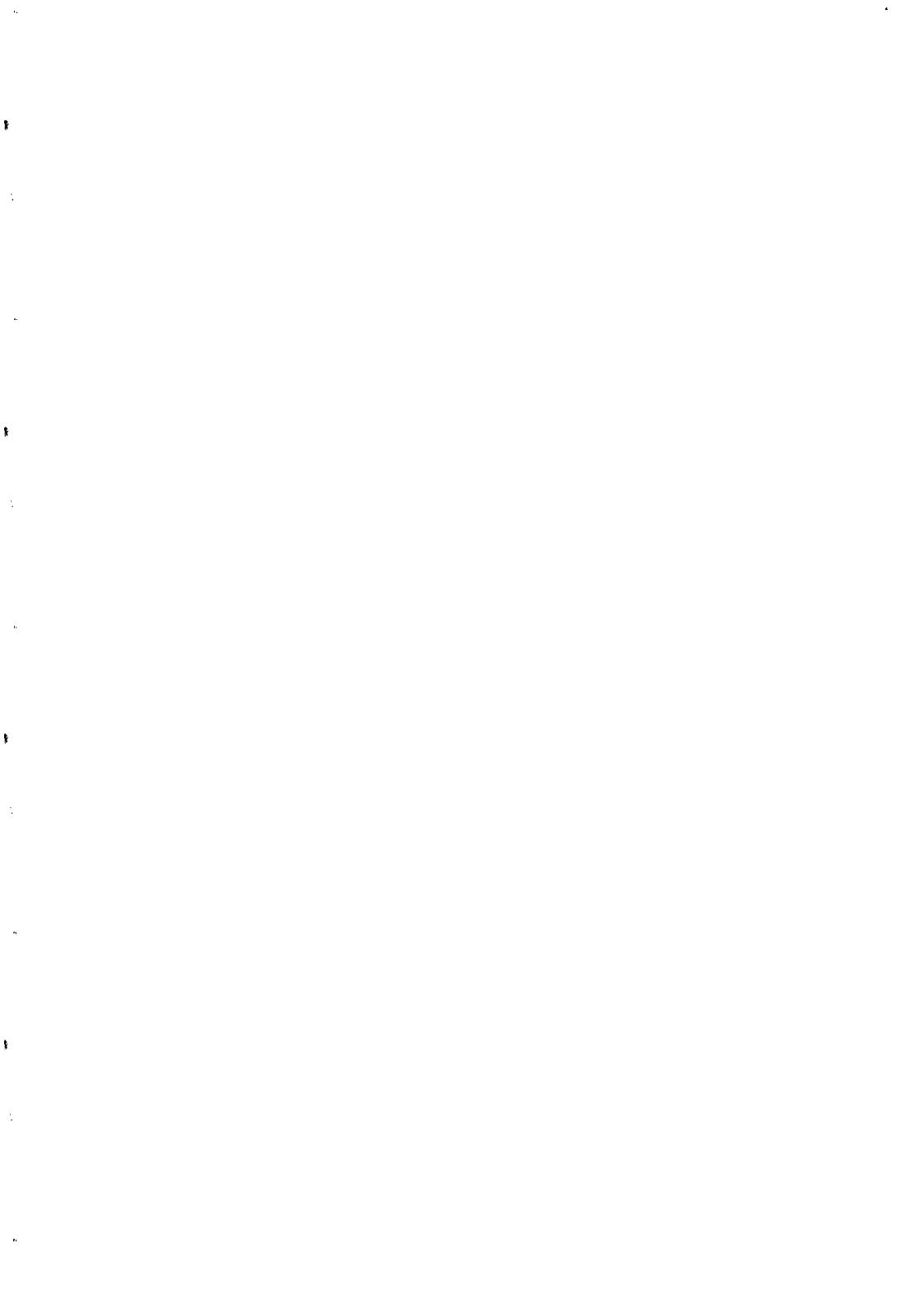
Atomic theory is capable of giving information about nuclear structure. The capacity is not yet fully utilised, as seen e.g. from the extraction of changes in magnetic radii from hyperfine anomaly measurements in the 50's. The relatively good accuracy obtained in these coupled-cluster calculation gives hope for more accurate calculations of parity non-conserving properties.

ACKNOWLEDGEMENTS

Financial support by the Swedish Natural Science Research Council is gratefully acknowledged. The calculations presented here were performed using computer programs developed over a long time in the the Göteborg atomic physics group under professor Ingvar Lindgren and it is a pleasure to acknowledge stimulating collaboration also with Eva Lindroth, Anders Ynnerman, Sten Salomonson, Per Öster and Adam Hartley. I would also like to thank Peter Hannaford for suggesting the application of our methods to the study of hyperfine anomalies.

REFERENCES

- 1 Kopfermann H, Nuclear Moments (1958), (New York, Academic)
2. see e.g the contributions by Bouchiat M-A, by Lamoreaux S and by D Cho, this volume, and Blundell S A, Hartley A C, Liu Z W, Mårtensson-Pendrill A-M and Sapirstein J, *Theor. Chim. Acta* **80** 257 (1991)
3. see e.g. contributions by Lamoreaux S, by Ross S and by Chupp T, this volume, and Mårtensson-Pendrill A-M, *Methods in Computational Chemistry* **5** 99 (Plenum, Ed. S Wilson, 1992)
4. Dzuba V A, Flambaum VV, Silvestrov P G and Sushkov O P *J Phys. B* **18** 597 (1985), *J. Phys. B* **20**, 1399 (1987) *Phys. Lett. A* **131** 461 (1988),
- 5 Dzuba V A, Flambaum VV and Sushkov O P, *Phys. Lett. A* **140** 493(1989), *ibid* **A142** 373 (1989)
6. Johnson W R, Blundell S, Liu Z W, Sapirstein J, *Phys. Rev. A* **37** 1395 (1988) Blundell S A, Johnson W R and Sapirstein J, *Phys. Rev. A* **37** 307 (1988)
7. Salomonson S and Öster P, *Phys. Rev. A* **40** 5548, 5559(1989)
8. Mårtensson-Pendrill A-M, Pendrill L, Salomonson S, Ynnerman A and Warston H, *J. Phys. B* **23** 1749-61 (1990)
9. Mårtensson-Pendrill A-M, Ynnerman A, Warston H, Vermeeren L, Silverans R E, Klein A, Neugart R, Schulz C, Lievens P and the ISOLDE Collaboration *Phys. Rev. A* **45** 4675 (1992)
10. Raghavan P, *At. Data Nucl. Data Tables*, **42** 189 (1987)
11. Mårtensson-Pendrill A-M and Ynnerman A, *J. Phys. B* **25** L551 (1992)
12. Mårtensson-Pendrill A-M, Gough D S and Hannaford P, *Phys. Rev. A*, in print
13. Lurio A and Prodell A G, *Phys. Rev.* **101** 79 (1956)
14. Hartley A C and Mårtensson-Pendrill A-M, *J. Phys. B* **24** 1193 (1991)
15. Hermann G, Lasnitschka G and Spengler D, *Z. Physik D* **28** 127 (1993)
16. Grexa G, Hermann G, Lasnitschka G and Fricke B, *Phys. Rev. A* **38** 1263 (1988)
17. Millack T, *Z. Phys. D* **8** 119 (1988)
18. Hartley A C and Mårtensson-Pendrill A-M, *Z. Phys. D* **15** 309 (1990)
19. Engfer R *et al*, *At. Data Nuclear Data Tables*, **14** 509 (1974)
20. Gould G, *Phys. Rev.* **101** 1828 (1956)
21. Mårtensson-Pendrill A-M and Ynnerman A, *Physica Scripta* **41** 329 (1990)
22. Bounds J A *et al*, *Phys. Rev. C* **36** 2560 (1987) Menges *et al*, *Z. Phys. A* **341** 475 (1992)
23. Bohr A and Weisskopf V F, *Phys. Rev.* **77** 94 (1950)
24. Stroke H H, Blin-Stoyle R J and Jaccarino V, *Phys. Rev.* **123** 1326 (1961)
25. Torbohm G, Fricke B and Rosén A, *Phys. Rev. A* **31** 2038-53 (1985)



HIGH RESOLUTION SPECTROSCOPY OF ATOMIC HYDROGEN ; A TOOL FOR METROLOGY

L. JULIEN and F. BIRABEN

Laboratoire Kastler Brossel, Université Pierre et Marie Curie,
4 place Jussieu, Case 74, 75252 Paris Cedex 05, FRANCE

Introduction

Modern Doppler-free laser techniques have allowed the resolution obtained in optical spectroscopy to be improved by a huge factor. The application of these techniques to atomic hydrogen makes it possible to test theoretical calculations on this simple system with a very high degree of accuracy.

The energy levels of atomic hydrogen

The hydrogen atom consists only of one proton (mass M , charge e) and one electron (mass m , charge $-e$). To a rough approximation, the energy depends only on the principal quantum number n and is given by the simple formula :

$$E_n = - h c R_H (1/n^2) \quad (1)$$

with R_H the Rydberg constant for hydrogen, h Planck's constant and c the velocity of light.

From the above expression, one can derive the Balmer-Rydberg formula which gives the transition wavelength between two different levels n and p :

$$1/\lambda = R_H (1/n^2 - 1/p^2) \quad (2)$$

The hydrogen atom has transitions in a wide range of wavelengths from the UV to microwaves, and the Rydberg constant R_H can be deduced from any wavelength or frequency measurement between two levels having different principal quantum numbers.

A knowledge of the electron-to-proton mass ratio allows one to deduce the Rydberg constant R_∞ for the case of an infinite nuclear mass : $R_\infty = R_H (1 + m/M)$.

R_∞ is a fundamental constant combining m , e and h :

$$R_\infty = m e^4 / 8 \varepsilon_0^2 h^3 c. \quad (3)$$

In fact, in order to obtain the exact energy levels, some corrections must be added to this very simple treatment :

- relativistic corrections, calculated from Dirac theory, which are responsible for the fine structure of the levels,
- hyperfine corrections, due to the magnetic moment of the proton ($I = 1/2$), which split all the levels into two components,
- radiative corrections which result from the interaction of the electron with the quantized electromagnetic field and are calculated using the theory of quantum electrodynamics (QED),
- the nuclear size corrections arising from the finite volume of the proton.

The last two corrections are especially important for the S levels : taken together they give the so-called Lamb shift. The ground state (1S) Lamb shift is about 8 GHz ; the uncertainty in this quantity limits the accuracy with which the theoretical value of the 1S energy itself is known, to a few parts in 10^{11} .

Finally, although the listed corrections cause the energy to depend on four quantum numbers (n, l, j, F), the Rydberg constant remains an overall scaling factor.

Motivation for high precision measurements of the Rydberg constant

Although the Rydberg constant is the most precisely known physical constant, with a present accuracy of 2 parts in 10^{11} , there is still a vigorous competition between several groups aiming to improve the precision still further.

One motivation to continue is the prospect of more stringent tests of QED. The experimental precision is now close to the precision of calculations giving the energy levels of hydrogen and simple atomic systems. To carry out QED tests on these systems, one needs a precise value of the scaling factor hcR_∞ , their natural unit

of energy. We quote two examples : the 1S-2S two-photon transition of hydrogen at 243 nm which gives the 1S Lamb shift, and the $1^3\text{S}-2^3\text{S}$ two-photon transition of the positronium atom at 486 nm ¹⁾ (positronium is of special interest as it is a purely leptonic system).

Another motivation is that the Rydberg constant gives a relation between such fundamental constants as the electron mass and charge. It can also be expressed in terms of the fine structure constant α :

$$R_\infty = (mc/2h) \alpha^2 \quad (4)$$

As a consequence it plays a key role in the least-squares adjustment of the fundamental constants where it is used as an auxiliary constant (that is, with a fixed value) ²⁾. It then gives a test of the consistency between various areas of physics.

Finally, R_∞ can be deduced from any transition frequency in hydrogen and gives a means of connecting microwave and optical frequencies. The methods used to study some optical transitions are reported in the next section. In the microwave range, the most convenient transitions are between adjacent circular states as suggested by Hulet and Kleppner ³⁾. Circular states have maximum orbital angular momentum ($l = n-1$), so that - in a classical description of the atom - the electron is orbiting in a circle centred on the nucleus. For these states, the probability for the electron to be close to the nucleus is very low and radiative and nuclear size corrections are at a minimum. Moreover they have a lifetime much longer than the other Rydberg states. The study of these transitions with n of the order of 25 or 30, in order to deduce the Rydberg constant in the microwave domain, has been carried out on the lithium atom in Paris ⁴⁾ and on hydrogen at M.I.T.

The hydrogen atom gives then a means to test the $1/r$ dependence of the Coulomb potential ⁵⁾. If this dependence is supposed exact, the hydrogen atom can provide a check of frequency chains between the optical and the microwave ranges (specifically, the Cs clock at 9 GHz) and may even be used in the future to give a *mise en pratique* of the meter in the visible range.

Spectroscopic measurements in atomic hydrogen

In the optical range, the Balmer single-photon transitions in atomic hydrogen have been studied extensively. The most precise result, with a precision of 3 parts in 10^{10} ⁶⁾ was obtained at Yale on the 2S-4P transition. It was limited by the natural width of the P levels which is responsible for a relative linewidth of a few parts in

10^8 . Indeed, the P levels are the broadest in hydrogen since they can radiate directly to the ground state.

The advantage of the two-photon transitions is to avoid these P states. The 1S-2S transition has an extremely small natural width (5 parts in 10^{16}) and lies in the UV range. It is presently being studied in Garching (see the paper of C. Zimmermann in this volume) and at Yale. The study of the 2S-nS/nD, carried out in Paris, is described in the following.

The Paris experiment

The 2S-nS/nD two-photon transitions ($n = 6$ to 12) lie in the near infrared (750-820 nm) and have relative linewidths of a few parts in 10^{10} . In order to obtain very narrow signals, we use the following experimental conditions ⁷⁾ :

- Metastable atoms are produced in an atomic beam to avoid collisional broadening.
- They are optically excited by two counterpropagating laser beams so that the first-order Doppler effect is cancelled.
- In the interaction region, the atomic beam is collinear with the two laser beams so that the transit time broadening is reduced.

The metastable atoms are produced by dissociation of molecular hydrogen in a radio frequency discharge followed by atomic excitation to the 2S state by electronic impact. The deviation of the atomic beam due to inelastic collisions with electrons is used to align the 2S beam with the laser beams. After passing through the interaction region (length ≈ 50 cm), the metastable atoms are detected by quenching in an applied electric field and measurement of the induced Lyman- α fluorescence by two photomultipliers.

The light source is a Ti-sapphire laser. To enhance the two-photon transition probability, the whole apparatus is placed inside a Fabry-Perot cavity. The length of this cavity is locked to the laser wavelength so that the light power experienced by the atoms can be up to 100 W in each direction of propagation. After optical excitation to the nS or nD states, most of the atoms decay to the ground state ; the two-photon absorption signal can then be detected through the associated decrease of the 2S beam intensity. An example of a recorded signal is shown in figure 1.

A numerical calculation of the signal profiles has been done taking into account light shift and saturation ⁸⁾. For each recording, the fit of the theoretical curve to the experimental one gives the line position corrected for light shift with respect to a fringe of a reference Fabry-Perot cavity.

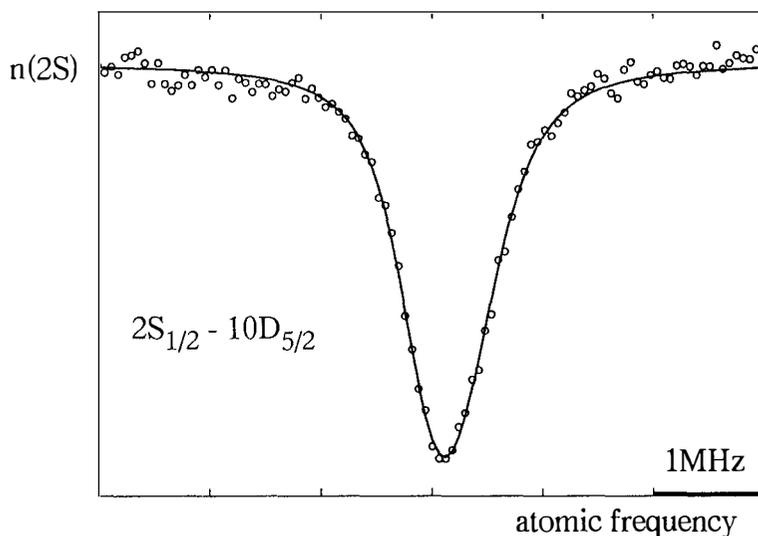


Figure 1. Signal obtained for the $2S_{1/2} - 8D_{5/2}$ transition. The circles are the experimental data and a fitted calculated profile is superimposed; the light power experienced by the atoms is here 48 W, the relative decrease of the metastable yield 11.5% and the observed linewidth 950 kHz.

The calibration of the transition wavelength or frequency of the transitions is done by comparison with a standard laser, as for example an I_2 -stabilized standard He-Ne laser at 633 nm. Two methods can be used for this absolute measurement.

In our previous measurement of the Rydberg constant, performed in 1988, the transition wavelength was compared with the standard one in a high stability Fabry-Perot interferometer⁹⁾. We estimate the intrinsic limit of this interferometric method to be about 2 parts in 10^{11} . The physical reason is that in the interferometer we compare two beams having different wavelengths and hence different spatial extensions : a small imperfection on a mirror surface can be "seen" differently by the two beams and so induce an error of this magnitude in the wavelength comparison.

To avoid this limitation we have now replaced the wavelength measurement by a direct frequency comparison of the $2S - 8S_{1/2}$, $8D_{3/2}$ and $8D_{5/2}$ lines with two optical standards¹⁰⁾. The scheme used for this comparison takes advantage of a near coincidence (89 GHz) between the frequency of the I_2 -stabilized He-Ne standard at 633 nm and the frequency sum of the CH_4 -stabilized He-Ne laser at 3.39 μm and the excitation laser at 778 nm. Two Ti-sapphire lasers are used, one for

the excitation of the atomic transitions and the other for the frequency sum in a LiIO_3 crystal with an auxiliary He-Ne at $3.39 \mu\text{m}$, frequency-locked to the He-Ne/ CH_4 standard laser. The synthesized beam generated at 633 nm is heterodyned with the He-Ne/ I_2 standard laser. The frequency difference between the two Ti-sapphire lasers is measured with a Schottky diode using a Gunn diode as an auxiliary source at 89 GHz . The frequencies of both standard lasers have been previously measured by comparison with the Cs clock through frequency chains ¹¹⁾.

Using this method, we have measured the frequencies of three two-photon transitions in atomic hydrogen : the $2S_{1/2} - 8S_{1/2}$, $2S_{1/2} - 8D_{3/2}$ and $2S_{1/2} - 8D_{5/2}$ transitions. We obtained in this way three independent determinations of the Rydberg constant, which are in good agreement. Our weighted mean is :

$$R_\infty = 109\,737.315\,683\,4\,(24) \text{ cm}^{-1}$$

This result, with an uncertainty of 2.2 parts in 10^{11} is the most precise one at the present time. It is in very good agreement with that obtained by T.W. Hänsch and his collaborators from the frequency measurement of the 1S-2S transition with a completely different frequency chain ¹²⁾.

Table I : Uncertainty budget.

Contribution	Parts in 10^{11}
Statistical uncertainty	0.7
Fits to theoretical profile	0.5
Standard at 633 nm	0.9
Standard at 3.39 mm	0.3
Second-order Doppler effect	0.3
Stark effect	0.2
Lamb shift and QED calculations	1.3
Electron-to-proton mass ratio	1.1
Global uncertainty	2.2

The various contributions to our uncertainty are listed in table I. One can distinguish two types of contribution :

- those due to our experimental method, which give altogether a contribution of 1.3 parts in 10^{11} ; they come either from our signal and its study, or from the standards and their intercomparison, and could be reduced in the future.

- those not inherent in our experiment; they arise from other data needed to deduce the Rydberg constant from our measurements, for example the electron-to-proton mass ratio, the 2S Lamb shift and QED calculations. They give altogether an uncertainty of 1.7 parts in 10^{11} .

Prospects

To improve further the accuracy of the Rydberg constant R_∞ , a new measurement of the electron-to-proton mass ratio, presently known with a precision of 2 parts in 10^8 , is highly desirable. Until we have such a measurement, increasing the precision of frequency measurements in hydrogen will lead only to improved values of R_H .

It is also the case, however, that new Lamb shift measurements are of great interest. As proposed by T.W. Hänsch, a convenient means for the determination of the 1S Lamb shift is the comparison of two frequencies having an almost integer ratio, such as the 1S-2S and 2S-4S/4D two-photon transitions studied in Garching¹³⁾. A similar experiment is currently underway prepared in Paris based on the 1S-3S and 2S-6S/6D two-photon transitions at 205 and 820 nm.

Finally, the two recent results obtained in Paris and Garching for the 2S-8D and 1S-2S transitions, with respective uncertainties 12 and 45 kHz, can be combined together to deduce the 1S Lamb shift: $L = 8172.815 (70)$ MHz. Moreover, if one assumes that the Lamb shift varies as $1/n^3$, a linear combination of these two transition frequencies allows one to eliminate the 1S and 2S Lamb shifts and then directly provides the R_H constant. The uncertainty would be 2 parts in 10^{12} in R_H if the two frequencies were known at a level 10^{-12} , which can be expected in the near future.

Conclusion

The development of tunable lasers has allowed an advance of three orders of magnitude in the precision of the Rydberg constant in less than twenty years. The experimental precision is presently 2 parts in 10^{11} and will probably be a few parts in 10^{12} in the near future. This precision is now challenging the theoretical calculations.

In atomic hydrogen two types of spectroscopic measurements can be distinguished :

- *Absolute measurements of a given transition frequency.* The difficulty of this type of measurement is the need for a good frequency standard and for a frequency chain to connect the studied transition to the standard.
- *Comparisons of two different transition frequencies,* one involving the ground state. Such a measurement provides a determination of the 1S Lamb shift.

Both are complementary and give together on one hand the Rydberg constant and on the other hand QED and proton size corrections^{1,4)}.

References

- 1) M.S. Fee, S. Chu, A.P. Mills Jr., R.J. Chichester, D.M. Zuckerman, E.D. Shaw and K. Danzmann, Phys. Rev. A **48**, 192 (1993).
- 2) E.R. Cohen and B.N. Taylor, Rev. Mod. Phys. **59**, 1121 (1987).
- 3) R.G. Hulet and D. Kleppner, Phys. Rev. Lett. **51**, 1430 (1983).
- 4) A. Nussenzweig, J. Hare, A.M. Steinberg, L. Moi, M. Gross and S. Haroche, Europhys. Lett. **14**, 755 (1991).
- 5) D.F. Barlett and S. Lögl, Phys. Rev. Lett. **61**, 2285 (1988) (see also E. Fishbach in this volume).
- 6) P. Zhao, W. Lichten, H.P. Layer and J.C. Bergquist, Phys. Rev. Lett. **58**, 1293 (1987); Phys. Rev. A **39**, 2888 (1989).
- 7) J.C. Garreau, M. Allegrini, L. Julien and F. Biraben, J. Phys. France **51**, 2263 (1990).
- 8) J.C. Garreau, M. Allegrini, L. Julien and F. Biraben, J. Phys. France **51**, 2275 (1990).
- 9) F. Biraben, J.C. Garreau, L. Julien and M. Allegrini, Phys. Rev. Lett. **62**, 621 (1989).
- 10) F. Nez, M.D. Plimmer, S. Bourzeix, L. Julien, F. Biraben, R. Felder, Y. Millerioux and P. De Natale, Europhys. Lett. **24**, 635 (1993).
- 11) O. Acef, J.J. Zondy, M. Abed, D.G. Rovera, A.H. Gérard, A. Clairon, Ph. Laurent, Y. Millerioux and P. Juncar, Opt. Commun. **97**, 29 (1993).
- 12) T. Andreae, W. König, R. Wynands, D. Leibfried, F. Schmidt-Kaler, C. Zimmermann, D. Meschede and T.W. Hänsch, Phys. Rev. Lett. **69**, 1923 (1992) (see also C. Zimmermann in this volume).
- 13) M. Weitz, F. Schmidt-Kaler and T.W. Hänsch, Phys. Rev. Lett. **68**, 1120 (1992); M. Weitz, A. Huber, F. Schmidt-Kaler D. Leibfried and T.W. Hänsch, Phys. Rev. Lett. **72**, 328 (1994).
- 14) B. Cagnac, M.D. Plimmer, L. Julien and F. Biraben, Reports Prog. Phys. (in press).

HIGH RESOLUTION SPECTROSCOPY OF ATOMIC HYDROGEN

C. Zimmermann, A. Huber, W. König, D. Leibfried, F. Schmidt-Kaler, M. Weitz,
and T. W. Hänsch.

Max-Planck-Institut für Quantenoptik, 85740 Garching, Germany,
and Sektion Physik der Universität München, 80799 München, Germany



Abstract

Two-photon spectroscopy of the extremely narrow 1S-2S transition has reached a resolution of better than 1 part in 10^{11} . Recent precision measurements of the 1S-Lamb shift, the hydrogen deuterium isotope shift, and the Rydberg constant are reported.

Introduction

As the simplest of stable atoms hydrogen may be studied, theoretically and experimentally with extraordinary accuracy. This offers unique opportunities for stringent tests of fundamental physics such as the theory of quantum electrodynamics (QED) as well as for ultra precise measurements of fundamental constants.⁽¹⁾ The experimental resolution of an optical atomic resonance is ultimately limited by its natural linewidths which for most transitions is typically on the order of several MHz. However, the hydrogen atom features an exceptionally narrow transition between the 1S ground state and the metastable 2S state. The

provided by a single mode dye laser near 486nm whose frequency is doubled in a nonlinear crystal (β -Barium Borate). Sufficient conversion efficiency is obtained by placing the crystal in a passive resonator which enhances the fundamental light power by a factor of 50 resulting in about 6 mW ultraviolet radiation. The dye laser is electronically stabilized to a carefully designed ultrastable reference resonator made from dielectric mirrors with gyroscopic quality which are optically contacted to the front facets of a 40 cm long Zerodur spacer. The spacer is suspended by soft springs and placed inside a thermally stabilized vacuum chamber. A wooden cover improves thermal and acoustic isolation. The laser bandwidth of about 1kHz is determined by residual frequency fluctuations of the reference resonator due to seismic and acoustic noise.

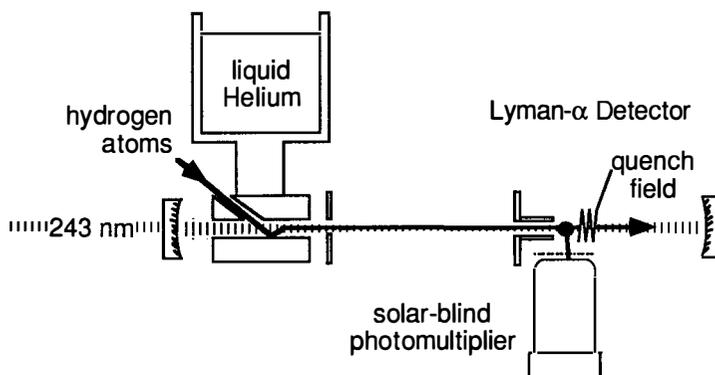


Fig. 1 Atomic beam apparatus for resolving the hydrogen 1S-2S transition with relative linewidth of better than 1 part in 10^{11} .

The spectra of the $F=1$ hyperfine component for different temperatures are shown in fig. 2. At room temperature, the asymmetric line shape is explained by the relativistic Doppler effect which shifts the atomic resonance due to time dilation of special relativity. The asymmetry reflects the thermal velocity distribution of the atomic beam. At lower temperature the Doppler broadening vanishes and at 8.6 K the resonance shape is almost entirely determined by transit broadening. The line width of 12 kHz in the ultraviolet corresponds to a resolution of 1 part in 10^{11} . For the heavier deuterium isotope, an even narrower line of 9 kHz has been observed.

lifetime of the 2S state - which for parity reason may not decay to the ground state by a single photon electric dipole transition - is determined by a two photon decay and limits the resolution to 1.3 Hz. The transition may be excited by two photon absorption from counterpropagating laser beams⁽²⁾. This has the advantage that the transition frequency is almost independent of the atomic velocity since the linear Doppler effect cancels for absorption from opposite directions. In addition the two photon transition splits the necessary photon energy and shifts the excitation wavelength into a range (243 nm) where continuous wave laser sources are available.

In this article we describe recent results of the Garching hydrogen project, all based on high resolution spectroscopy of the 1S-2S transition. A comparison of the 1S-2S transition frequency with the 2S-4S two photon transition has led to an improved value for the 1S Lamb-shift which for the first time exceeds the accuracy of the best reported measurements of the "classical" 2S Lamb shift.⁽³⁾ Initially, the result did not agree with theory but now new improved calculations of higher order QED corrections⁽⁴⁾ seem to solve the discrepancy. In a second experiment the 1S-2S transition frequency has been compared to the cesium clock by means of a phase coherent frequency chain and an intermediate frequency standard near 3.39 μ m.⁽⁵⁾ The result was used to derive an improved value of the Rydberg constant. Finally, the isotope shift between hydrogen and deuterium has been measured with 20-fold improved accuracy providing information about the charge radii of the proton and the deuteron.⁽⁶⁾

Two Photon Spectroscopy of the 1S-2S Transition

To minimize line broadening due to collisions and transit time effects, the hydrogen atoms are observed by longitudinal excitation of a cold atomic beam⁽⁷⁾ (fig.1). The hydrogen atoms are produced by dissociation of H₂ in a radio frequency gas discharge and enter the vacuum chamber through a nozzle which is mounted at the bottom of a cryostat. The atoms thermalize by wall collisions inside the nozzle and form a atomic beam parallel to the axis of a standing wave resonator. The resonator builds up the incident ultraviolet laser light to about 20-50 mW circulating power and provides the two counterpropagating laser beams for Doppler free two photon excitation. At the end of the 15 cm long atomic beam, the excited atoms are detected by mixing the 2S and the 2P state with a small electric field (10V/cm) which stimulates a rapid radiative decay. The emitted Lyman- α photons are counted by a photomultiplier. The ultraviolet light near 243nm is

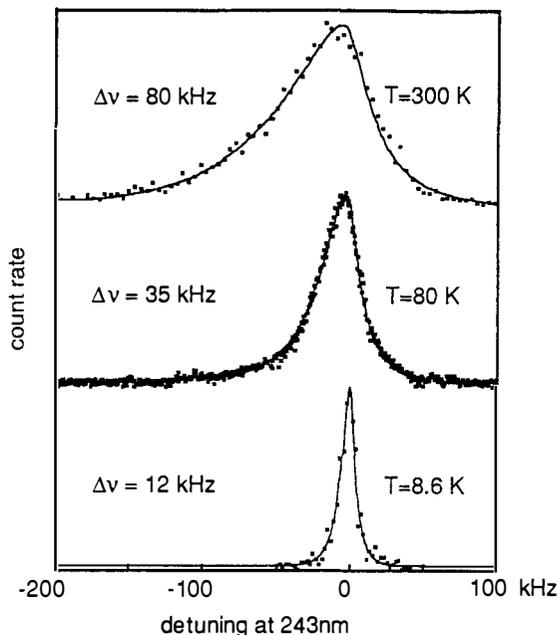


Fig.2 Spectra of the $F=1$ hyperfine component of the hydrogen $1S-2S$ two photon transition at different temperatures.

The $1S-2S$ energy interval is determined by the Dirac energy, by the $1S$ Lamb shift, by well understood corrections due to the motion of the nucleus, by the $2S$ Lamb shift and by the hyperfine interaction. The latter two contributions are both precisely measured. Thus, the $1S-2S$ transition mainly provides information about the $1S$ Lamb shift and the Dirac energy. An independent determination of both entities is possible by observing a linear combination of two hydrogen transition frequencies. If for instance the $1S-2S$ transition frequency is subtracted from four times the $2S-4S$ transition frequency all energy contributions proportional to n^{-2} cancel and with it the main part of the Dirac energy (n is the principle quantum number). The compound frequency which is on the order of several GHz is therefore dominated by quantum electrodynamic corrections and a precise value for the $1S$ Lamb shift may be derived. In combination with an absolute measurement of the $1S-2S$ frequency interval the precise knowledge of the compound frequency also allows to determine a new value for the Rydberg constant.

Precision Measurement of the 1S Lamb shift

The 2S-4S transition is observed by longitudinal excitation of a metastable hydrogen beam. The hydrogen atoms are excited into the 2S state by electron collisions which in addition deflect the atoms onto the axis of a standing wave resonator with 50 W circulating power near 972 nm. At the end of the beam the metastable atoms are observed by applying an electric field and observing the Lyman- α fluorescence with a channeltron detector. If the laser is tuned on resonance some of the atoms are excited into the 2S-4S from where they decay rapidly into the ground state via the 2P or 3P level and a drop of the metastable flux is observed. With an additional fluorescence detector we also record the emitted Balmer- β photons during the decay into the 2P state. The excitation light is provided by a single mode titanium sapphire laser which is electronically stabilized to an ultrastable reference resonator. Part of the laser output is frequency doubled in a potassium niobate crystal and superimposed with light from the dye laser. The resulting beat signal is recorded with fast photodiode and an electronic counter.

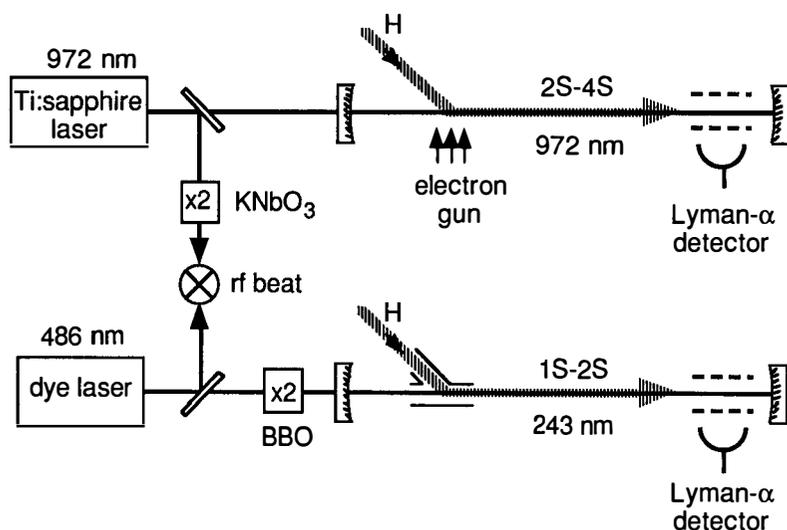


Fig. 4 Experimental apparatus for comparison of the 1S-2S two photon transition frequency with the 2S-4S transition frequency.

From the measured beat note of 4797.340(11) GHz we find a 1S Lamb shift of 8172.86(5) GHz with an accuracy, which for the first time exceeds the best measurements of the "classical" 2S Lamb shift.⁽³⁾ Good agreement with theory is obtained if the Mainz value⁽⁸⁾ for the proton charge radius is assumed to be correct and very recent new calculations of quantum electrodynamic two loop corrections⁽⁴⁾ are taken into account. The result of these calculations have been somewhat surprising since they reveal an unexpected large QED contribution which in the past has been underestimated by about an order of magnitude.

Transition Frequency of the 1S-2S Transition and a New Rydberg Constant

For an absolute measurement of the 1S-2S frequency interval the dye laser is electronically locked to the 1S-2S resonance and its frequency is compared to the cesium clock by means of an intermediate frequency standard near 3.39 μm . This standard was built in the laboratory of V. Chebotayev in Novosibirsk and is based on a helium neon laser which is locked to the Lamb dip of a methane resonance observed in an intracavity gas cell. The standard was calibrated with a reproducibility of better than 10^{-12} at the Physikalisch Technische Bundesanstalt (Braunschweig, Germany) in direct comparison with a cesium frequency standard. We synthesize a violet reference frequency near 424 nm by generating the 8th harmonic in three steps of second harmonic generation. Because of the poor conversion efficiency after each step a phase locked laser transfer oscillator boosts the power before the light is sent into next doubling stage. In this way the full accuracy of the helium neon standard is transferred to the violet frequency at 707 THz which by now represents the highest optical frequency that has been synthesized with an accuracy of better than 10^{-12} . Simultaneously, we generate the sum frequency between the dye laser and the helium neon standard in an additional nonlinear crystal resulting in a frequency which by coincidence is close to the 8th harmonic of the helium neon standard. A remaining frequency mismatch of about 2 THz is bridged by 6372 longitudinal modes of the dye laser reference resonator which mode spacing has been calibrated with an accuracy of better than 1 Hz. As result we find a 1S-2S frequency interval of 2466 061 413.182(45) Hz with an 18 fold improved accuracy compared to the best previous measurement.⁽⁹⁾ From comparison with theory together with the measured 1S-Lamb shift a new value for the Rydberg constant of $109\,737.315\,684\,1(31)\text{ cm}^{-1}$ has been derived with a relative uncertainty of $2.8 \cdot 10^{-11}$. In the meantime our measurement has been fully confirmed by a recent measurement carried out at the Ecole Normale Supérieure

in Paris.⁽⁹⁾ As the best known fundamental constant the Rydberg constant not only serves as an important corner stone for the adjustment of fundamental constants but also establishes an entire system of accurately known reference frequencies from the vacuum ultraviolet to the visible, infrared and microwave region, because the frequencies of all other hydrogen transitions can now be predicted with improved precision, provided that QED is correct.

The hydrogen deuterium isotope shift

The 1S-2S energy interval of deuterium is shifted relative to hydrogen by 672 nm due to its heavier nucleus and its different nuclear charge radius. Since the proton/deuteron mass ratio is known very accurately a precise measurement of the isotope shift may yield to an improved value of the electron/proton mass ratio provided that the nuclear size corrections are known e. g. from measurements of the 1S Lamb-shift. To measure the isotope shift we tune the dye laser half way between the hydrogen and the deuterium 1S-2S resonance. A novel fast electrooptic modulator⁽¹⁰⁾ generates optical sidebands at ± 84 GHz which serve as frequency markers close to both atomic resonances. A second identical stabilized dye laser system is alternately locked to the maximum of the hydrogen and the deuterium resonance and its frequency is compared to the sideband marker frequencies by recording the beat signal on a fast photodiode. The measured isotope shift of 670 994 337 (22) kHz is almost 25 times more accurate than the best previous measurement and so far limited by slow frequency drifts of the reference laser.⁽⁶⁾ Comparison with the theoretical value of 670 994 414 (22) kHz reveals a discrepancy of 3.5 standard deviations which is yet to be explained. A recent analysis of K. Pachucki et al.⁽¹¹⁾ which includes contributions due to the deuteron polarizability reduces the 1S-2S interval by about 20 kHz.

Tabletop spectroscopic experiments are thus beginning to reach to effects of nuclear structure and dynamics in a regime of low energy and momentum transfer which is inaccessible to experiments with large colliders. High resolution spectroscopic experiments of hydrogen and hydrogen like atoms most notably antihydrogen⁽¹²⁾ continue to hold fascinating challenges. Perhaps the biggest surprise in the endeavour would be if we found no surprise.

This work has been supported in part by the Deutsche Forschungsgemeinschaft.

References

- 1) The Hydrogen Atom, eds. G. F. Basini, M. Inguscio, and T. W. Hänsch (Springer Verlag, Berlin, Heidelberg, New York 1989).
- 2) L. S. Vasilenko, V. P. Chebotaev, and A. V. Shishaev, JETP Lett. 12, 113 (1979).
- 3) M. Weitz, A. Huber, F. Schmidt-Kaler, D. Leibfried, and T. W. Hänsch, Phys. Rev. Lett. 72, 328 (1994).
- 4) K. Pachucki, Phys. Rev. A 48, 2609 (1993), and K. Pachucki submitted for publication in Phys. Rev. A.
- 5) T. Andreae, W. König, R. Wynands, D. Leibfried, F. Schmidt-Kaler, C. Zimmermann, D. Meschede, and T.W. Hänsch, Phys. Rev. Lett. 68 (1992) 1923.
- 6) F. Schmidt-Kaler, D. Leibfried, M. Weitz, and T. W. Hänsch Phys. Rev. Lett. 70, 2261 (1993).
- 7) C. Zimmermann, R. Kallenbach, T. W. Hänsch, Phys. Rev. Lett. 65, 571 (1990).
- 8) G. G. Simon, Ch. Schmitt, F. Borkowski und V. H. Walther, Nucl. Phys. A333, 381 (1980).
- 9) Nez, F., et al. Phys. Rev. Lett. 69 (1992) 2326, and L. Julien in this volume.
- 10) D. Leibfried, F. Schmidt-Kaler, M. Weitz, and T. W. Hänsch, Appl. Phys. B 56, 65 (1993).
- 11) K. Pachucki, D. Leibfried and T. W. Hänsch, Phys. Rev. A 48, R1 (1993), and K. Pachucki, M. Weitz, and T. W. Hänsch submitted to Phys. Rev. A.
- 12) Antihydrogen, Hyperfine Interactions, 76, 1ff (1993).

Trapped Positrons for Ion Cooling and Antihydrogen

K. Abdullah, L. Haarsma, and G. Gabrielse

Department of Physics, Harvard University, Cambridge, MA 02138

Positrons from a 12 mCi ^{22}Na source are slowed by a W (110) reflection moderator and then introduced into a Penning trap, where their motion is coupled to a damping circuit. Because of the stability of the Penning trap and the cryogenic ultra-high vacuum environment, we anticipate that positrons can be accumulated and stored indefinitely. The first demonstration has a continuous loading rate of $0.14\text{ e}^+/\text{sec}$ for 32 hours, yielding more than 1.6×10^4 trapped positrons, with improvements expected. The extremely high vacuum is required for compatibility with an existing antiproton trap, which has already held more than 10^5 particles, for producing antihydrogen at low temperatures. It may also prove useful for cooling highly stripped ions.

I. INTRODUCTION

Cryogenically cooled Penning traps should provide an excellent environment for the production and study of antihydrogen. By mixing cold (4 K), positrons with antiprotons, antihydrogen can possibly be formed at a high instantaneous rate [1], at energies suitable for magnetic trapping, in the way already done with hydrogen [2]. More than 10^5 antiprotons have already been trapped under these conditions [3] by sending a pulse of 10^8 antiprotons, delivered from the low energy antiproton ring at CERN, through a degrader foil which slows the antiprotons collisionally to below 3 keV. These antiprotons are then further cooled by collisions with cold electrons in the trap [4], and by coupling their motion to a damping circuit, to 4 K. Storage times of several months have been demonstrated [5].

Positrons trapped in a high vacuum are the other ingredient for antihydrogen production. Less than 100 positrons were once confined for long periods of time in an ultrahigh, cryogenic environment [6, 7], but with much lower numbers and trapping rates than are required for antihydrogen production. Large numbers of positrons have already been trapped using buffer gasses [8] to slow the positrons. However, the background gas atoms in these traps unacceptably limit the antimatter storage times. It may eventually be possible to improve a third technique which utilizes pulsed electrodes [9], to accumulate small numbers of positrons in relatively high pressure, but this seems difficult.

We report here a different technique [10, 11] which allows positrons to be loaded continuously from a radioactive source in the same vacuum as the antiproton trap. The use of a moderator [12-14] to slow the positrons is a critical feature. Trapping rates exceeding $0.2 e^+/\text{sec}$, and total numbers of 1.6×10^4 trapped e^+ , have already been achieved, and there are several significant possibilities for considerably increasing this rate. Because of the stability of Penning traps and the ultrahigh vacuum, the

number of trapped positrons may be limited only by the patience of the experimenter and the trap capacity.

II. SLOWING THE POSITRONS

The Penning trap used in this experiment is shown in Fig. 1. Positrons from a 12 mCi ^{22}Na source located in the fringing field of a superconducting solenoid ($B_s = 1.6$ Tesla) follow the magnetic field lines to the W (110) moderator through loading tubes and small holes in the trap electrodes. Only 90% of the decays generate positrons, only half of which travel in the forward direction towards the trap, and an estimated half of these in turn are absorbed within the source container itself, yielding a total source activity $A = 1 \times 10^8$ e⁺/sec, over an active area with radius $r_s = 1$ mm. The trap is located at the center of the solenoid's magnetic field ($B_t = 5.4$ T). Positrons emitted with a forward angle $\theta_c \geq \arcsin(B_s/B_t)^{1/2}$ with respect to the field lines magnetically "bounce" back toward the source and never reach the moderator. In addition, the active area of the beam shrinks by a factor of B_s/B_t before the positrons reach the holes in the trap electrode ($r_h = 0.5$ mm). The average cyclotron radius of these energetic positrons reduces the effective aperture to $r_e \approx 0.25$ mm. Therefore, the total flux on the moderator is

$$F = A \frac{r_e^2 B_t}{r_s^2 B_s} \left[1 - \sqrt{\left(1 - \frac{B_s}{B_t}\right)} \right] \approx 3 \times 10^6 \text{ e}^+/\text{s}. \quad (1)$$

We detect this flux by mechanically chopping the beam at ≈ 1 Hz and measuring the current on the moderator with a lock-in amplifier.

Some positrons are backscattered epithermally from the moderator, but most rapidly thermalize within the crystal. A fraction of these will diffuse to the surface, where they experience a negative (for positrons) work function ϕ_+ and are ejected nearly monoenergetically, if the crystal surface is clean and well-annealed. (See for example reference [14].) We chose to use a W (110) reflection moderator since these

had already been shown to be capable of good efficiencies ($\eta \approx 10^{-3}$) and energy widths of $\Delta E \leq 65$ meV at temperatures approaching 4 K [12, 13]. A small energy spread in the re-emitted positrons (compared to the several hundred keV energy spread from the radioactive source) is critically important, since only those positrons in a very narrow energy band will be trapped. The use of a moderator also avoids the problem of high-energy positrons being relativistically shifted out of resonance with the trap's damping circuit [6, 7].

For this experiment, the tungsten crystal was mechanically polished and then electrochemically etched to remove the damaged surface layers. It was annealed at temperatures exceeding 2000 K for four hours at 10^{-7} Torr and then heated to ≈ 900 K in 10^{-6} Torr oxygen for two hours in an effort to remove some of the interstitial carbon near the surface. The crystal was then transferred into the vacuum enclosure of the Penning trap and an electron field emission point array was installed immediately below the moderator (see Fig. 1). The array produces more than 5 mA of beam current, allowing us to heat the moderator (mounted on thin tungsten wires) for several minutes *in situ* by electron bombardment, while keeping the rest of the trap and vacuum enclosure, which is heat sunk to a liquid helium reservoir, cold and under high vacuum.

The positron trapping rate as a function of the moderator bias voltage is shown in Fig. 2. Before the first heating, the energy spread of positrons emerging from the moderator was $\Delta E \approx 3$ volts, presumably due to positrons scattering off adsorbed surface material. After the first cleaning, the distribution changed dramatically. The FWHM of the peak dropped to $\Delta E = 700$ meV, with a low-energy tail of approximately twice the area of the peak. We hope that moderator performance will improve with further treatment, but these studies are still underway.

III. POSITRON TRAPPING AND ACCUMULATION

The crystal is biased to nearly the work function, with respect to the endcap electrodes, so that moderated positrons will re-enter the Penning trap with very little kinetic energy. The off axis loading tube (see Fig. 1) is negatively biased to create a slight “dip” in the trap’s electric potential near the entrance aperture, so that the total kinetic plus potential energy of positrons entering the trap is less than the total trap depth. The ring and compensation electrodes are biased so that the positron’s axial motion is coupled to an LRC damping circuit [15], shown in Fig 3.

During the positron’s first axial orbit inside the trap there is insufficient time for this damping to be effective. Fortunately, during this first axial orbit, the positron with also $\mathbf{E} \times \mathbf{B}$ drift laterally in its magnetron orbit [7, 15] a distance

$$\delta_{\text{mag}} = 2\pi \rho_{\text{mag}} \left(\frac{\nu_{\text{mag}}}{\nu_z} \right) \approx 5 \mu\text{m}, \quad (2)$$

where $\nu_z = 70$ MHz is the axial motion frequency, $\nu_{\text{mag}} = 16$ kHz is the magnetron frequency, and $\rho_{\text{mag}} = 0.36$ cm is the distance between the axis of the loading tube and the central symmetry axis of the trap. The “dip” in the trap’s electric potential near the entrance aperture is strongest along the central axis of the tube [7, 11], so those positrons whose magnetron drift carries them *away* from the central tube axis will essentially “climb a potential hill.” A very narrow band (typically ~ 1 meV) will have insufficient energy, and therefore remain in the trap for at least one complete magnetron orbit before returning to the entrance tube. During this orbit they can lose energy to the damping circuit

$$\delta E = \gamma_z e V_0 / \nu_m \approx 33 \text{ meV} \quad (3)$$

where $V_0 = 12\text{V}$ is the trap depth and $\gamma_z/2\pi \approx 7$ Hz for this trap, given the observed quality factor ($Q = 1050$) of our LRC damping circuit [15, 16]. In order for a positron

to remain trapped after its first magnetron orbit, it must enter the trap with no more than δE energy greater than the lowest “dip” potential it experiences for all points in its magnetron orbit. The total trapping efficiency is calculated in more detail in reference [11], and can be summarized

$$\begin{aligned} \epsilon &= \frac{2\delta_{\text{mag}}\delta E}{\pi r_e \Delta E} && \text{if } \delta E < \Delta E, \\ &= \frac{2\delta_{\text{mag}}}{\pi r_e} && \text{otherwise.} \end{aligned} \quad (4)$$

In order for the positron’s axial motion to damp, the frequency of its oscillations must be near the resonance frequency of the LRC circuit. This is accomplished by adjusting the potentials applied to the ring and the compensation electrodes (see Fig. 1). Figure 4 shows the positron loading rate as a function of these voltages. The shape of this dependence is expected from the Lorentzian shape of the damping circuit’s frequency response. The potentials measured to bring the axial motion into resonance match those calculated by modeling the trap and using relaxation techniques to compute the electric potentials along the entrance tube axis [17].

Combining equations 1 and 4 yields an overall trapping rate $R = F\eta\epsilon$ of

$$\begin{aligned} R &= \eta A \frac{B_t}{B_s} \left[1 - \sqrt{\left(1 - \frac{B_s}{B_t}\right)} \right] \left(\frac{2r_e \delta_{\text{mag}}}{\pi r_s^2} \right) \left[\frac{\delta E}{\Delta E} \right] && \text{if } \delta E < \Delta E, \\ &= \eta A \frac{B_t}{B_s} \left[1 - \sqrt{\left(1 - \frac{B_s}{B_t}\right)} \right] \left(\frac{2r_e \delta_{\text{mag}}}{\pi r_s^2} \right) && \text{otherwise.} \end{aligned} \quad (5)$$

Solving for our measured trapping rate, source activity, and ΔE , we calculate that our moderator efficiency $\eta \approx 4 \times 10^{-4}$.

Once positrons lose energy to the damping circuit and “survive” their first magnetron orbit, they can be held indefinitely without loss. After loading for some time, the positrons can be moved to the center of the trap [15, 18] where their number can be determined by their interaction with the damping circuit [16, 19], as shown in Fig. 5. Figure 6 shows the results of one run in which positrons were accumulated

for more than 30 hours. The initial trapping rate is $0.14 \text{ e}^+/\text{s}$ and no discernable interruption or decrease in the loading rate was observed.

IV. IMPROVEMENTS AND USES

Now that this technique has proved successful, Eq. 5 suggests several ways to increase the trapping rate R . One obvious method—provided appropriate precautions are taken for the health of the experimenter—is simply to increase the source activity A (without increasing r_s), either by using a source with a shorter half-life or by placing a more intense source and moderator outside of the strong magnetic field and focussing the beam onto a remoderator near the trap [11]. An additional factor of seven could be gained by moving the source closer to the trap and out of the solenoid's fringing field, to eliminate magnetic "bouncing," while simultaneously increasing the size of the trap's loading apertures. Of critical importance is the moderator efficiency η and the energy width ΔE of the moderated positrons. Reducing ΔE to 65 meV, presumably through better preparation of the crystal, would increase the trapping rate by a factor of 30. In addition, the trapping rate scales linearly with the quality factor Q of the damping circuit, so long as $\delta E < \Delta E$. Quality factors as high as $Q = 2700$ have been achieved on a similar Penning trap used for antiprotons.

With such a trap installed in the same vacuum enclosure as the antiproton trap currently working at LEAR, it would only take a few hours or a few days to accumulate sufficient positrons to produce antihydrogen [1]. The two trapped species could be made to overlap, either by using a nested pair of Penning traps in which the antiprotons oscillate slowly through a central potential well filled with positrons [20], or by superimposing a radiofrequency signal on the Penning electrodes, which is capable of confining both charged species in the same mechanism as a Paul trap. Under these conditions, antihydrogen would be radiately produced at a rate [1, 21]

$$R_{\bar{H}} = 3 \times 10^{-11} \left(\frac{4.2}{T} \right)^{1/2} N_{p^-} n_{e^+} \text{ s}^{-1}. \quad (6)$$

For a positron density $n_{e^+} = 10^7 / \text{cm}^{-3}$ (typical for these traps) and a number of trapped antiprotons $N_{p^-} = 10^5$, at 4.2 K, this yields a production rate $R_{\bar{H}} = 30/\text{s}$.

The three-body recombination mechanism $p^- + e^+ + e^+ \rightarrow \bar{H} + e^+$ may well be more efficient for antihydrogen production by many orders of magnitude. Although the initial positron capture in this case occurs within a few kT of the ionization limit, producing highly-excited Rydberg atoms, the antihydrogen atoms will be moving slowly, and could possibly be held by their large magnetic moment long enough for collisional de-excitation down to a state where spontaneous emission de-excitation would dominate. The rate for the equivalent matter recombination has been calculated in various ways [22], including under conditions of high magnetic field [23], giving

$$R_{\bar{H}} \approx 10^{-12} \left(\frac{4.2}{T} \right)^{9/2} N_{p^-} n_{e^+}^2 \text{ s}^{-1}. \quad (7)$$

Note the extremely sensitive dependence upon temperature, and the squared dependence upon positron density. For the identical conditions as previous, this would produce antihydrogen at a rate $R_{\bar{H}} \approx 10^7 / \text{s}$ (provided the newly formed atoms remain in the positron cloud long enough to collisionally de-excite). Of course, N_{p^-} sets a limit on the total *number* of antihydrogen atoms produced.

This technique of trapping positrons could also be useful for trapping and studying highly stripped ions. Ions would cool from keV energies to sub-eV energies in just a few seconds [4] via collisions with the trapped positrons. The ultra-high cryogenic vacuum would prevent loss of ions due to charge-transfer processes with neutral background atoms.

We are grateful to B. Brown for contributions to earlier stages of this work, and

to K. Canter and K.G. Lynn for useful discussions regarding positron moderators.

This research was supported by the Office of Naval Research of the US.

REFERENCES

- [1] G. Gabrielse, S.L. Rolston, L. Haarsma, and W. Kells, *Phys. Lett. A* **129**, 38 (1988).
- [2] H.F. Hess, G.P. Kochanski, J.M. Doyle, N. Masuhara, D. Kleppner, and T.J. Greytak, *Phys. Rev. Lett.* **59**, 672 (1987).
- [3] G. Gabrielse, X. Fei, L.A. Orozco, R.L. Tjoelker, J. Haas, H. Kalinowsky, T.A. Trainor, and W. Kells, *Phys. Rev. Lett.* **63**, 1360 (1989).
- [4] S.L. Rolston and G. Gabrielse, *Hyperfine Interactions* **44**, 233 (1988).
- [5] G. Gabrielse, X. Fei, L.A. Orozco, R.L. Tjoelker, J. Haas, H. Kalinowsky, T.A. Trainor, and W. Kells, *Phys. Rev. Lett.* **65**, 1317 (1990).
- [6] P.B. Schwinberg, R.S. Van Dyck Jr., and H.G. Dehmelt, *Phys. Lett. A* **81**, 119 (1981).
- [7] P.B. Schwinberg, Ph.D. thesis, University of Washington, 1979.
- [8] C.M. Surko, M. Leventhal, and A. Passner, *Phys. Rev. Lett.* **62**, 901 (1989); T.J. Murphy and C.M. Surko, *J. Phys.* **B23**, L727 (1990).
- [9] R.S. Conti, B. Ghaffari, and T.D. Steiger, in *Proceedings of the Antihydrogen Workshop, Munich, 1992*, edited by J. Eades (CERN, Geneva, Switzerland, 1992).
- [10] G. Gabrielse and B. Brown, in *The Hydrogen Atom*, edited by G.F. Bassani, M. Inguscio, and T.W. Hansch (Springer, Berlin, 1989) p. 196.

- [11] L. Haarsma, K. Abdullah, and G. Gabrielse, *Hyperfine Interactions* **76**, 143 (1993).
- [12] D.A. Fischer, Ph.D. thesis, State University of New York at Stony Brook, 1984; D.A. Fischer, K.G. Lynn, and D.W. Gidley, *Phys. Rev.* **B33**, 4479 (1986).
- [13] B.L. Brown, W.S. Crane, and A.P. Mills Jr., *Appl. Phys. Lett.* **48**, 739 (1986).
- [14] P.J. Schultz and K.G. Lynn, *Rev. Mod. Phys.* **60**, 701 (1988).
- [15] L.S. Brown and G. Gabrielse, *Rev. Mod. Phys.* **58**, 233 (1986).
- [16] H.G. Dehmelt and F.L. Walls, *Phys. Rev. Lett.* **21**, 127 (1968); D.J. Wineland and H.G. Dehmelt, *J. Appl. Phys.* **46**, 919 (1975).
- [17] G. Gabrielse, *Phys. Rev.* **A27**, 2277 (1983).
- [18] D.J. Wineland and H.G. Dehmelt, *Intl. J. Mass Spectrom. Ion Phys.* **16**, 338 (1975); R.S. Van Dyck, Jr., P.B. Schwinberg, and H.G. Dehmelt, in *New Frontiers in High Energy Physics*, edited by B. Kursunoglu, A. Perimutter, and L. Scott (Plenum, New York, 1978).
- [19] J.N. Tan, Ph.D. thesis, Harvard University, 1992.
- [20] W. Quint, R. Kaiser, D. Hall, and G. Gabrielse, *Hyperfine Interactions* **76**, 181, (1993).
- [21] D.R. Bates and A. Dalgarno, *Atomic and Molecular Processes* (Academic Press, New York, 1962) p. 245.
- [22] D.R. Bates, A.E. Kingston, and R.W.P. McWhirter, *Proc. R. Soc.* **A267**, 297 (1962); J. Stevefelt, J. Boulmer, and J.F. Delpech, *Phys. Rev.* **A12**, 1246 (1975).
- [23] M.E. Glinsky and T.M. O'Neil, *Phys. Fluids* **B3**, 1279 (1991).

FIGURES

FIG. 1. Cross section of Penning trap and moderator for continuous trapping of positrons.

FIG. 2. Positron trapping rate measured before ● and after ○ the moderator was cleaned by heating to ~ 1600 K *in situ* for three minutes.

FIG. 3. Penning trap and idealized picture of damping circuit. The trap electrodes themselves form the capacitor of the tuned circuit.

FIG. 4. Positron trapping rate as a function of ring electrode voltage for compensation voltages (a) $V_c \approx -36$ volts, (b) $V_c \approx -6$ volts, and (c) $V_c \approx +24$ volts. Loading rates are maximum when the positron axial frequency matches the resonance frequency of the damping circuit.

FIG. 5. Frequency spectrum of thermal noise in LRC circuit (a) with trap empty and (b) with ~ 16200 e^+ .

FIG. 6. Positron accumulated as a function of time.

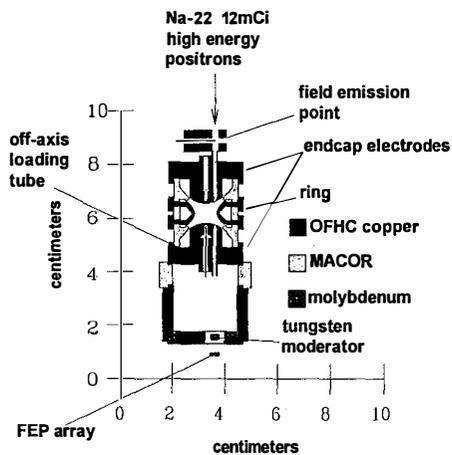


Figure 1

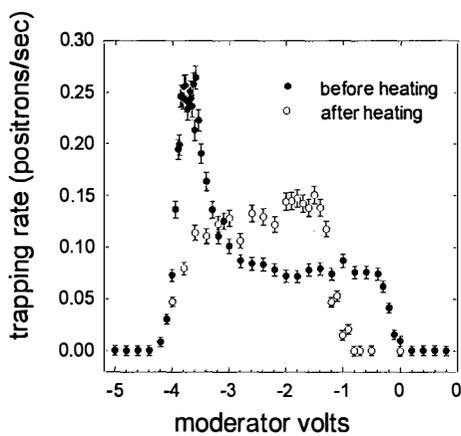


Figure 2

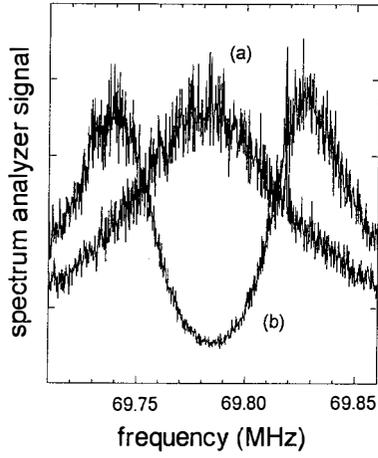


Figure 5

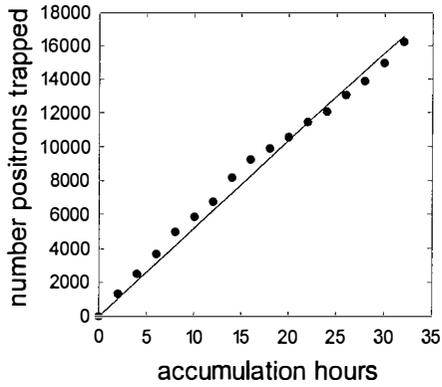


Figure 6

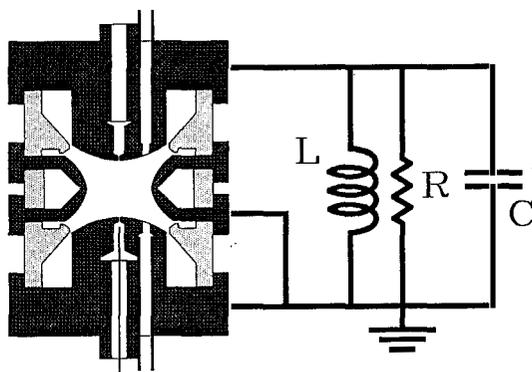


Figure 3

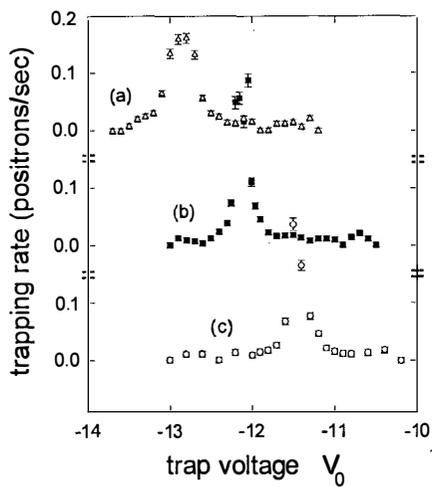
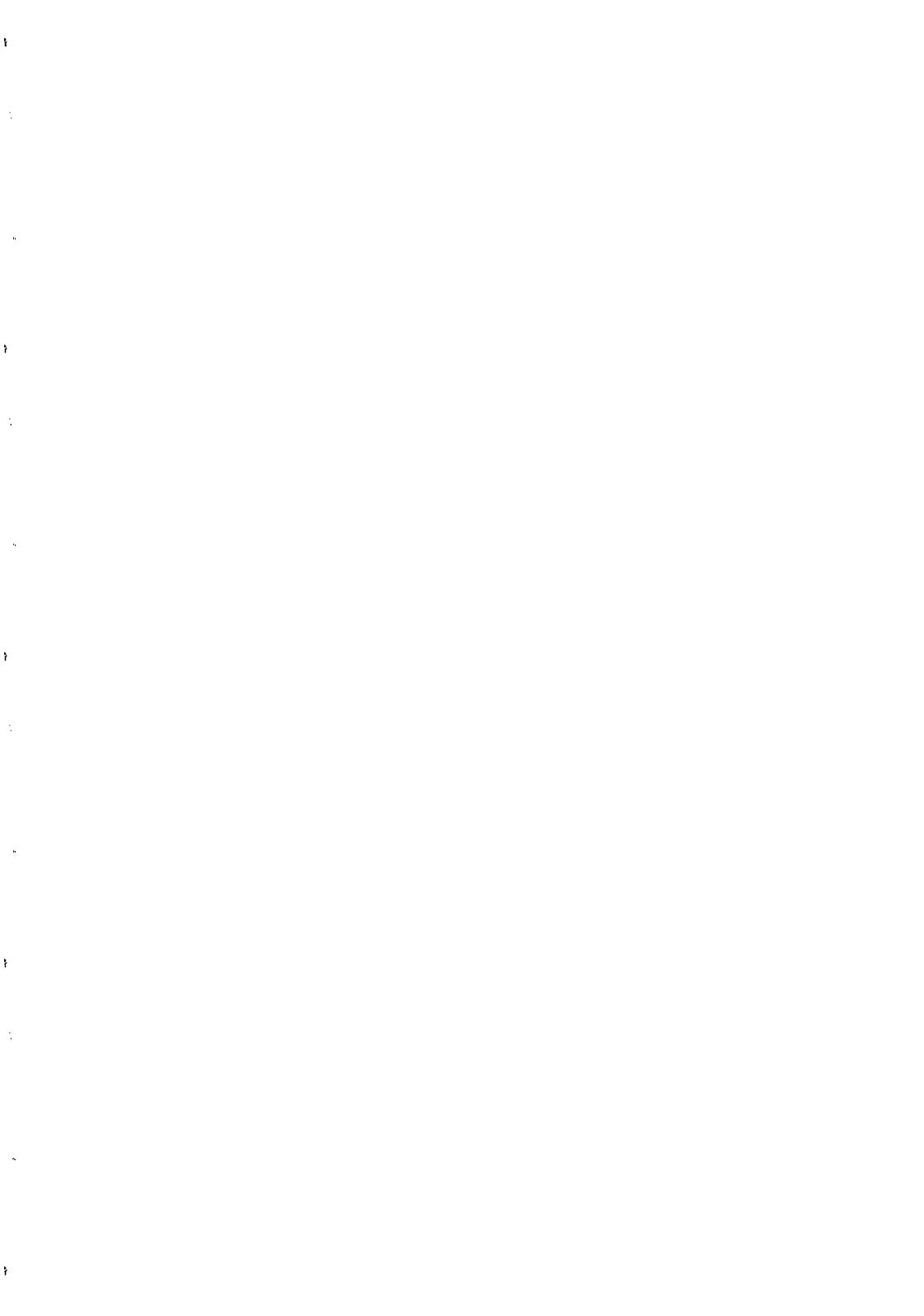


Figure 4



POTENTIAL CONCEPTS, LONG-RANGE FORCES AND BOUND-STATE QED

J. Sucher

Center for Theoretical Physics and Department of Physics,
University of Maryland, College Park, MD, 20742

Abstract

The effects of one- and two-photon exchange on the energy levels of composite systems have been studied for many years, especially in the case of two-body bound states. However, a systematic analysis of the extent to which such effects can be described as resulting from the action of configuration-space potentials, determined by field theory but acting directly between the constituents, appears not to have been carried out. A method for studying this problem will be presented and the results obtained to date will be summarized, with emphasis on the long-range character and spin-dependence of such potentials, as well as on conceptual aspects; there are some surprises here. Some comments will also be made on the long-range forces arising from two-neutrino exchange, which have recently been the object of renewed interest, and on aspects of the long-range force between a hydrogen and helium atom, which involves a competition between a repulsive two-photon exchange force and the attractive force of gravity.

I. Introduction

The main purpose of this talk is to discuss some aspects of the use of effective potentials in few-body problems and to report results of a recent paper on this topic, in the context of the long-range forces arising from photon exchange.¹⁾ I will also make some brief remarks on gravity and (virtual!) neutrinos, major themes at this meeting. As befits the pre-dinner time slot, my sermon will be very informal and start with an ever popular type of tale: a referee story. The first referee on the above-mentioned paper indicated that he might be old-fashioned but he thought it should not be published because it does not contain a calculation whose results could be immediately compared with experiment. This was a criterion I had not encountered before; it would certainly lead to a much thinner journal! A second referee was very positive, agreeing that the issues raised were important. I conclude that the latter is not only modern in outlook but, if male, both a gentleman and a scholar.

By way of an aperitif, here's a quiz. Recall that the usual starting point for a relativistic theory of hydrogen or H-like ions is the Dirac equation:

$$h(1)\psi(1) \equiv [\alpha_1 \cdot p_1 + \beta_1 m + V_{\text{ext}}(1)]\psi(1) = E\psi(1) \quad (1)$$

Now consider the statement: "With $U_c = e_1 e_2 / 4\pi r$, a reasonable starting point for a relativistic theory of helium or He-like ions is the equation:

$$h(1,2)\psi(1,2) \equiv [h(1) + h(2) + U_c]\psi(1,2) = E\psi(1,2)." \quad (2)$$

Question 1: The last statement is a) true, b) false, c) not well posed.

Having given you five seconds to think about it, I will tell you that (c), which I would normally chose in such a pressure situation (one can always argue about what is meant by "reasonable"), is incorrect. By any standard, the answer is (b), because (1) has no normalizable solutions associated with bound states.

Another example: In the late 1920's Breit and Gaunt independently considered the question of the leading correction to the Coulomb interaction, in the context of a Dirac description of electrons, arriving at different results:

$$U_b = -\frac{1}{2}(\alpha_A \cdot \alpha_B + \alpha_A \cdot \mathbf{r} \alpha_B \cdot \mathbf{r}) U_c, \quad U_g = -\alpha_A \cdot \alpha_B U_c. \quad (3)$$

Question 2: Which one, if any, is right?

I will return to this later. The point is that in the context of relativistic quantum field theory (RQFT) the concept of an

effective two-body interaction operator or potential V is subtle and beset with ambiguity.²⁾ In RQFT the interaction Lagrangian is primary; V is secondary and requires sharp definition. Historically, effective potentials have often "emerged" in the context of level-shift calculations for a specific physical system, initially in the context of time-independent perturbation theory, exemplified by Tamm-Dancoff (TD) type of calculations, and later from four-dimensional Bethe-Salpeter (BS) type of equations, via a relatively messy reduction to equal times.³⁾ However such potentials really merit *a priori* definitions and delineation of their use.

II. A different approach to effective potentials and bound states

I want to briefly sketch another approach, whose genesis is in work done long ago with the late Gary Feinberg on the quantum theory of long range forces (LRF). Using the techniques of particle theory (Lorentz and gauge invariance, analyticity and unitarity) we studied, in particular, the LRF arising from photon exchange between two composite neutral spinless systems. I review some of the results very briefly here, in order to write down some formulas which are needed later. As one would expect, the potential $V_{1\gamma}$ associated with one-photon exchange turns out to be short-range. However, the potential $V_{2\gamma}$ from two-photon exchange between two such systems, A and B , is long-range. For large r and low energies one finds that

$$V_{2\gamma} \approx -D/r^7, \quad D \equiv (23/4\pi)(\alpha_E^A \alpha_E^B + \alpha_M^A \alpha_M^B) - (7/4\pi)(\alpha_E^A \alpha_M^B + \alpha_M^A \alpha_E^B), \quad (4)$$

where the α 's denote electric and magnetic polarizabilities. The purely electric terms were first obtained in the classic work of Casimir and Polder.⁴⁾ Application of the same techniques to the case of a neutral composite A and a charged elementary particle B ($V_{1\gamma}$ is then still short-range) yields

$$V_{2\gamma}(r) = (e_B^2/4\pi) [(-1/2)\alpha_E r^{-4} + (11/4\pi)\alpha_E r^{-4}(\lambda_B/r) + (5/4\pi)\alpha_M r^{-4}(\lambda_B/r) + \dots] \quad (5)$$

where $\lambda_B = m_B^{-1}$ and the dots denote terms which fall off as $1/r^7$ or faster; these terms are however important in applications of the results to the fine structure of He Rydberg levels.⁵⁾

Returning to our problem, both for practical reasons and to gain insight, it is obviously desirable, on the one hand, to try (i) to describe the interaction of particles in terms of potentials which can be used in 3-dimensional equations and to define such potentials directly. On the other hand, we wish (ii) to retain the enormous

simplification achieved by the use of Feynman graphs and techniques in the computation of higher-order effects. Of course we want (iii) to avoid any *a priori* nonrelativistic approximations and if possible, (iv) to avoid any approximations which destroy gauge-invariance (GI). In contrast, the BS equation involves a kernel K which must be truncated in practice; this destroys GI in gauge theories such as QED. Further, in bound-state problems the use of Coulomb gauge is a practical necessity, which destroys manifest Lorentz invariance. While these are not fatal flaws, the approach I will sketch has, at a minimum, some conceptual advantages; unlike TD or BS, it retains both Lorentz and gauge invariance at any stage of approximation and has some other practical advantages. In any case, I believe it has a higher ISQ (intellectual satisfaction quotient) than the traditional approaches, which to some extent have the character of a black box.

I will focus on the two-body problem. The basic idea is quite simple,⁶⁾ a sort of a geometric mean between TD and BS. Somewhat paradoxically, we first consider the scattering problem and the associated two-body transition amplitude T . We then ask to what extent T can be regarded as arising from an effective two-body potential, to be used in a Schroedinger type of equation. To be more explicit, we define an interaction operator V , acting directly in configuration space, as a Fourier transform of an on-shell amplitude, obtained from gauge-invariant subsets of Feynman diagrams, modified by appropriate subtractions to avoid double counting; V is constrained by the requirement that when used in a specified type of center-or momentum system (c.m.s.) relativistic Schroedinger equation it reproduce $T_{c.m.}$, the value of T in the c.m.s. The equation has the natural form

$$h_{op} \phi = W\phi, \quad h_{op} = h_{op}^{(0)} + V \quad (6a)$$

with $h_{op}^{(0)}$ defined by

$$h_{op}^{(0)} = E_A^{op} + E_B^{op}, \quad [E_i^{op} \equiv (m_i^2 + \mathbf{p}_{op}^2)^{1/2}, \quad \mathbf{p}_{op} = -i\partial/\partial\mathbf{r}]. \quad (6b)$$

The associated potential theory transition amplitude T_{pot} is given by

$$T_{pot} = \langle \mathbf{p}' | V + V(W - h_0^{op} - V + i\epsilon)^{-1} V | \mathbf{p} \rangle. \quad (7)$$

The field-theory transition amplitude T is given, in the c.m.s. by

$$T_{c.m.} = M(s, t) / 4E_A E_B \quad (8)$$

where $M(s, t)$ denotes the invariant Feynman amplitude and $s \equiv (\mathbf{p}_A + \mathbf{p}_B)^2$, $t \equiv (\mathbf{p}_A - \mathbf{p}_A')^2$. The constraint on V , which in general will depend parametrically on s , is then simply that $T_{pot} = T_{c.m.}$, a condition which

is to be satisfied order-by-order in perturbation theory. To apply this to bound states, we look for normalizable solutions of (6); these correspond to poles of M at values of s below $s_0 = (m_A + m_B)^2$.

III. LRF between charged particles: Beyond the Coulomb potential

For concreteness, consider two point-like spin-0 particles, with charges e_A and e_B , and confine attention to the so-called generalized ladder approximation to $M(s,t)$, i.e. to graphs which involve photon exchange only between the particles. Consider first the one-photon exchange potential $V_{1\gamma}$. If one uses Feynman gauge in writing down the (gauge invariant) one-photon exchange amplitude $M_{1\gamma}$, one is led⁶⁾ to a Feynman-gauge inspired (FGI) potential $V_{1\gamma}^{\text{FGI}}$,

$$V_{1\gamma}^{\text{FGI}} = z'_{\text{op}} U_C z'_{\text{op}} + Y_{\text{op}} (\mathbf{p}_{\text{op}} \cdot U_C \mathbf{p}_{\text{op}} / 2m_A m_B) Y_{\text{op}} \quad (9a)$$

where

$$z'_{\text{op}} \equiv (1 + \mathbf{p}_{\text{op}}^2 / 2E_A^{\text{op}} E_B^{\text{op}})^{1/2}, \quad Y_{\text{op}} = (m_A m_B / E_A^{\text{op}} E_B^{\text{op}})^{1/2}. \quad (9b)$$

The corresponding Coulomb-gauge inspired (CGI) one-photon exchange potential $V_{1\gamma}^{\text{CGI}}$ is given by¹⁾

$$V_{1\gamma}^{\text{CGI}} \equiv Y_{\text{op}} [\{E_A^{\text{op}}, \{E_B^{\text{op}}, U_C\}\} + (1/2) \{\mathbf{p}_i^{\text{op}}, \{\mathbf{p}_j^{\text{op}}, (\delta_{ij} + \hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) U_C\}\}] Y_{\text{op}} / 4m_A m_B \quad (10)$$

In the n.r. limit (9) yields as the leading correction to U_C an orbit-orbit interaction U_{o-o} of the form

$$U_{o-o}^{\text{FGI}} = \{\mathbf{p}_i^{\text{op}}, \{\mathbf{p}_j^{\text{op}}, \delta_{ij} U_C\}\} / 4m_A m_B. \quad (11)$$

whereas (10) yields

$$U_{o-o}^{\text{CGI}} = (1/2) \{\mathbf{p}_i^{\text{op}}, \{\mathbf{p}_j^{\text{op}}, (\delta_{ij} + \hat{\mathbf{r}}_i \hat{\mathbf{r}}_j) U_C\}\} / 4m_A m_B. \quad (12)$$

This is a manifestly hermitian form of the orbit-orbit interaction U_{o-o} familiar from atomic physics, usually described as arising from reduction of the Breit operator (3) to n.r. form. But, of course, spin has nothing to do with it!

Some insight into the difference between the two choices comes from examining the potential $V_{2\gamma}$ from 2γ exchange. Surprisingly, the computation of this is more difficult than in the case where at least one of A or B is neutral, because of the presence of infrared divergences (IR); the cure for these turns out to be precisely the subtractions necessary anyhow to avoid double counting.⁶⁾ *These subtractions depend on the choice of $V_{1\gamma}$ and turn out to affect even the asymptotic form of $V_{2\gamma}$.* At low energies and large r one finds, with $k \equiv e_A e_B / 4\pi$, that¹⁾

$$V_{2\gamma} = c_2 r^{-2} + c_3 r^{-3} + \dots \quad (13)$$

where $c_3 = -7k^2 / 6\pi m_A m_B$ in both cases, but

$$c_2^{\text{FGI}} = k^2 / 2(m_A + m_B), \quad c_2^{\text{CGI}} = 0. \quad (14)$$

This observation resolves a long-standing puzzle in the literature and shows that in the case of two charged particles the concept of the asymptotic form of the effective potential has an unexpected ambiguity. Further, as was noted some time ago by L. Spruch, c_2^{FGI} is classical in character. It turns out that this can also be understood, by an extension of the classic work of Darwin.⁷⁾

IV. Inclusion of spin

Similar results hold when A and B are point particles but B, say, has spin-1/2. The FGI $V_{1\gamma}$ is then given by⁸⁾

$$V_{1\gamma}^{FGI} = (Y_A^{op} \Lambda_+^{op}) U^{(2)} (Y_A^{op} \Lambda_+^{op}), \quad U^{(2)} \equiv \{E_A^{op} - \alpha \cdot \mathbf{p}_{op}, U_C\} / 2m_A, \quad (15)$$

where $\Lambda_+^{op} = (E_B^{op} + h_B^{op}) / 2E_B^{op}$ is a Casimir positive-energy projection operator and $Y_A^{op} = (m_A / E_B^{op})^{1/2}$. The corresponding CGI $V_{1\gamma}$ is

$$V_{1\gamma}^{CGI} = (Y_A^{op} \Lambda_+^{op}) (U_C' + U_T') (Y_A^{op} \Lambda_+^{op}) \quad (16a)$$

where

$$U_C' \equiv \{E_A^{op}, U_C\} / 2m_A, \quad U_T' \equiv -(1/2) \{\mathbf{p}_i^{op} \alpha_j, (\delta_{ij} + \mathbf{r}_i \mathbf{r}_j) U_C\} / 2m_A. \quad (16b)$$

Reduction of (15) and (16) to the n.r. limit yields the same spin-dependent (s.d) interaction operator $-[1 + 2m_B/m_A] U_C \sigma \cdot \boldsymbol{\ell} / 4m_B^2 r^2$, while the s.i. potentials differ in the same way as when B has spin-0. If $V_{1\gamma}^{FGI}$ is used, the two-photon exchange yields a s.i. term with c_2 as in (14) and a s.d. correction given at large r by⁸⁾

$$V_{2\gamma;pt}^{s.o.} \approx -K^2 [(3m_A + 5m_B) / m_A (m_A + m_B)] (\sigma \cdot \boldsymbol{\ell} / 4m_B^2 r^4). \quad (17)$$

If A has structure, $V_{2\gamma}$ also contains a spin-orbit polarizability potential

$$V_{2\gamma;pol}^{s.o.} = (e_B^2 / 4\pi) [(\alpha_E^A m_B + \alpha_M^A (m_A + m_B)) / 2m_A m_B^2] (\sigma \cdot \boldsymbol{\ell} / r^6) + O(r^{-7}). \quad (18)$$

There are a number of physical situations in which it may be possible to detect the effects of $V_{2\gamma}^{s.o.}$. Typically these involve measurements of bound state energies in exotic atoms, where one particle has spin 1/2 and another has spin 0. Examples include anti-protonic atoms with a spin-0 nucleus, such as p-He⁴, pionic atoms with a spin-1/2 nucleus, such as pionic hydrogen, and the pi-muon bound state known as pi-muonium. Certain aspects of $V_{2\gamma}^{s.o.}$ may be observable in Rydberg states of helium-like ions whose nuclei have spin 1/2. For details see Ref.8.

For two-spin 1/2 particles one finds, with Λ_{++} a projection operator product,

$$V_{1\gamma}^{FGI} = \Lambda_{++} (U_C + U_G) \Lambda_{++}, \quad V_{1\gamma}^{CGI} = \Lambda_{++} (U_C + U_G) \Lambda_{++}. \quad (19)$$

Thus the answer to the question posed earlier is "neither". The computation of $V_{2\gamma}$ for this case is a major undertaking, which is

currently in progress.⁹ When this is completed one will be able to reanalyze the spin-dependent level structure of a number of physical systems and gain new insight into some aspects of QED.

IV. Some exotic long-range forces

Since both gravity and neutrinos are topics of this meeting I thought it would be of interest to discuss briefly some unusual aspects of long-range forces, which touch on these areas.

A. *The long-range H-He force and gravity.* From (4) we can see that, contrary to folklore, there are cases in which $V_{2\gamma}$ is repulsive at large distances. Indeed, although the coefficient D in (4) is necessarily positive if $A = B$, already the very simplest departure from this, viz. $A = H$ and $B = He^4$ gives a negative D. This is because $\alpha_H \ll \alpha_E$ for He while for H, $\alpha_H \approx 129a^3$, which is about 30 times larger than $\alpha_E \approx 9a^3/2$ ($a = \text{Bohr radius}$) and more than compensates for the ratio $23/7 \approx 3$. Thus $D \approx -(7/4\pi)\alpha_H^H\alpha_E^{He} < 0$ and $V_{2\gamma}^{as} \equiv -D/r^7$ is repulsive. Since $V_{2\gamma}$ is attractive at distances of a few Bohr radii, when graphed as a function of r it must cross the $V = 0$ axis for large enough r and then approach this axis from above. If we now add the one-graviton exchange potential $V_{1g} = -Gm_Hm_{He}/r$ to $V_{2\gamma}$ the total potential $V_{tot} = V_{2\gamma} + V_{1g}$ must eventually approach the axis from below. It is possible that there is a little potential well at some sufficiently large r ? If so, perhaps something could be made of it. Indeed the sum $V_{tot}^{as} \equiv V_{2\gamma}^{as} + V_{1g}$ vanishes at $r_0 \approx [(7/4\pi)\hbar c\alpha_H^H\alpha_E^{He}/Gm_Hm_{He}]^{1/6}$ and its derivative vanishes at $r_1 = 7^{1/6}r_0 \approx 1.4r_0$. With $\alpha_E^{He} \approx 1.4a^3$ one gets $r_0 \approx 4 \times 10^6 a \approx 2 \times 10^{-2}$ cm. However, one must now ask whether $V_{2\gamma}$ is well approximated by $V_{2\gamma}^{as}$ for such values of r . This requires a detailed knowledge of $V_{2\gamma}$ for the case at hand. Recently Chi Kwan Au and I decided to have a closer look at this.¹⁰⁾

We may write, in an obvious notation, $V_{2\gamma} = V_{EE} + V_{ME} + V_{EM} + V_{MM}$. Although V_{EM} and V_{MM} can be neglected, analysis shows that the actual V_{ME} is canceled by V_{1g} at a much smaller value of r . To be quantitative, for values of r such that $2\omega_H r/c \ll 1$, where $\hbar\omega_H$ is the hyperfine splitting of the hydrogen ground state, one finds that

$$V_{ME} \approx (5/8\pi r^5)\alpha^3 a^2 \alpha_E^{He} (\hbar\omega_H). \quad (20)$$

Then $V_{ME} + V_{1g} = 0$ for $r_0' = [(5/8\pi)\alpha^3 a^2 \alpha_E^{He} (\hbar\omega_H)/Gm_Hm_{He}]^{1/4} \approx 2 \times 10^{-3}$ cm and the derivative vanishes for $r_1' = 5^{1/4}r_0' = 1.5r_0'$. Since V_{EE} is well approximated by its asymptotic form for $r \gg \alpha^{-1}a \approx 10^{-6}$ cm, we have

$$-V_{ME}/V_{EE} \approx (5/46)\alpha^3 (2\pi a^2/\Lambda_H \alpha_E^H) r^2 \approx .015\alpha^3 (r^2/\Lambda_H a) \quad (22)$$

in the region of interest, with $\lambda_H = 21$ cm. For $r = r_0'$ the ratio (22) is of order 10^{-6} and so is much less than unity. The V_{EE} potential therefore dominates in this region and there is no minimum. Thus, this intriguing possibility for detecting an interplay between electromagnetism and gravity in an atomic system appears to be out of reach for the foreseeable future.

B. *The two-neutrino exchange force.* The exchange of neutrino-antineutrino pair between two spin-1/2 particles also gives rise to a long-range potential, $V_{2\nu}$, which falls off as r^{-5} , as shown long ago.¹¹⁾ My interest in this subject was reawakened by a preprint in of Hsu and Sikivie.¹²⁾ These authors had come across some lectures of Feynman in which he studied the question of whether neutrino-pair exchange forces might be responsible for gravity.¹³⁾ In this connection Feynman considered the possibility that the two-neutrino exchange potential might fall off as r^{-3} . On dimensional grounds the simplest form of $V_{2\nu}$ compatible with this is $V_{2\nu} \propto G_F^2 m_A m_B / r^3$, in which case some experimental tests might be feasible. Because the results of Ref. 12 disagreed with those found in Ref. 11 (in the coefficients of both the s.i. and s.d. parts of $V_{\nu\nu}$), I was led to check the old work in a more general way, without making any n.r. approximations. The result turns out to be of some interest, especially in view of Feynman's considerations.

Consider a current-current interaction of Dirac fields ψ_A and ψ_B with a massless neutrino field ψ_ν of the form

$$L_{\text{eff}} = -G_A (\bar{\psi}_A \Gamma_A^\sigma \psi_A) (\bar{\psi}_\nu \Gamma_\nu^\sigma \psi_\nu) - (A \rightarrow B) \quad (22)$$

where the Γ 's have the generic form

$$\Gamma_A^\sigma = \gamma^\sigma (1 + \xi_A \gamma_5), \quad \Gamma_B^\sigma = \gamma^\sigma (1 + \xi_B \gamma_5), \quad \Gamma_\nu^\sigma = \gamma^\sigma (1 + c \gamma_5) \quad (23)$$

and G 's denote effective Fermi coupling constants. One then finds, on use of the methods of Ref. 11, that the long-range part of $V_{\nu\nu}$ is given by $V_1 + V_2$, where, with $G^2 \equiv (1+c^2) G_A G_B$,

$$V_1 \equiv (G^2/4\pi^3) (\gamma_A^0 \gamma_B^0) (\Gamma_A \cdot \Gamma_B r^{-5}), \quad V_2 \equiv (3/2) (G^2/4\pi^3) (\gamma_A^0 \gamma_B^0) (m_A m_B \gamma_A^5 \gamma_B^5 / r^3). \quad (24)$$

Thus, within a Dirac description of spin-1/2 particles, there is indeed a term proportional to $m_A m_B r^{-3}$, but it comes accompanied with a γ^5 factor for each particle. Reduction to Schroedinger-Pauli form of V_2 then yields, in the low-energy limit, only spin-dependent terms of order r^{-5} . To be precise,

$$(V_1)_{\text{red}} = (G^2/4\pi^3 r^5) (1 - \xi_A \xi_B \sigma_A \cdot \sigma_B), \quad (V_2)_{\text{red}} = (G^2/4\pi^3 r^5) (\xi_A \xi_B / 2) (5\sigma_A \cdot \hat{\mathbf{r}} \sigma_B \cdot \hat{\mathbf{r}} - \sigma_A \cdot \sigma_B)$$

so that

$$(V_{\nu\nu})_{\text{red}} = (G^2/4\pi^3 r^5) [1 + (\xi_A \xi_B / 2) (5\sigma_A \cdot \hat{\mathbf{r}} \sigma_B \cdot \hat{\mathbf{r}} - 3\sigma_A \cdot \sigma_B)]. \quad (25)$$

For $G_i = G_f/2^{1/2}$ and $c = -1$, $\xi_i = -1$, as in the old four-fermion interaction for charged leptons, this reduces to the results of Ref. 11. In the standard model, we have $G_i = c_\nu G_f/2^{1/2}$, $\xi_i = -c_A/c_\nu$, with $c_\nu - 1 = 2\sin^2\theta_w - 1/2$, $c_A - 1 = -1/2$, so that the s.i. part of $(V_w)_{red}$ becomes $(2\sin^2\theta_w + 1/2)^2 G_f^2/4\pi^3 r^5$, in agreement with Ref. 2.

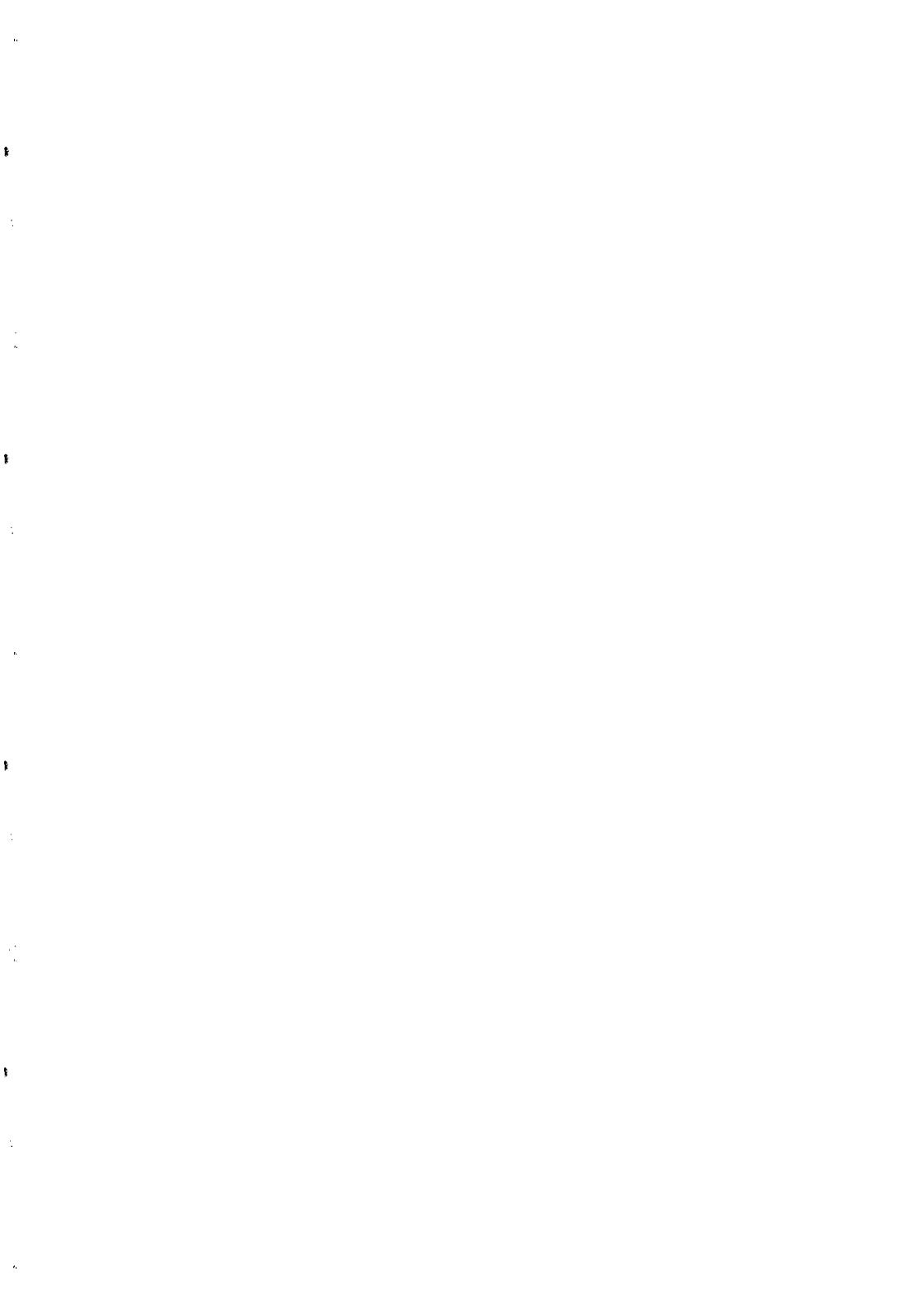
It is amusing to note that if one had two macroscopic bodies which contained an appreciable fraction of relativistic polarized electrons, so that $\langle \gamma_s \rangle \approx 1$, one could indeed generate an interaction of the Feynman type! Unfortunately such systems are hard to come by.

Acknowledgement

This work was supported in part by the National Science Foundation.

References

1. J. Sucher, to appear in *Phys. Rev. D*, April, 1994.
2. For a discussion of such ambiguities see J. Sucher, in *Proceedings of the Program on Relativistic, Quantum Electrodynamical and Weak Interaction Effects in Atoms*, ITP, Santa Barbara (AIP Conf. Proc. No. 189, 1989), edited by W.R. Johnson, P. Mohr, and J. Sucher (AIP, New York, 1989), p. 337.
3. See, e.g., C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw Hill, New York, 1980), p. 481.
4. For a review see G. Feinberg, J. Sucher, and C.K. Au, *Phys. Rep.* **180**, 85 (1989).
5. See reviews in *Long-Range Casimir Forces: Theory and Recent Experiments in Atomic Systems*, edited by F.S. Levin and D.A. Micha (Plenum, New York, 1993).
6. G. Feinberg and J. Sucher, *Phys. Rev. D* **38**, 3763 (1988); see also G. Feinberg and J. Sucher, "A gauge-independent approach to the two-body problem" (UMDTR, 1989; unpublished).
7. C.G. Darwin, *Philos. Mag.* **39**, 537 (1920); for a discussion of this point see J. Sucher, "Long-range electromagnetic forces in quantum field theory: 23 skidoo, 7 come 11", to appear in *Comments on Atomic and Molecular Physics*, 1994.
8. G. Feinberg and J. Sucher, *Phys. Rev.* **45**, 2493 (1992).
9. G. Gilbert and J. Sucher (in preparation).
10. C.K. Au and J. Sucher (unpublished).
11. G. Feinberg and J. Sucher, *Phys. Rev.* **166**, 1638 (1965).
12. S.D.H. Hsu and P. Sikivie, Harvard University preprint, 92-A041. There are several errors in this paper; a corrected version is in agreement with Ref. 11 (P. Sikivie, private communication).
13. R.P. Feynman, *Lectures on Gravitation*, 1962-63. notes by F.B. Morinigo and W.W. Wagner (California Institute of Technology, 1971)



ATOMS IN INTENSE LASER FIELDS

C. J. Joachain

Physique Théorique, Université Libre de Bruxelles, Belgium



ABSTRACT

A review is given of recent progress made in studying the interaction of atomic systems with intense laser fields. Following a survey of multiphoton ionization, harmonic generation and laser-assisted electron-atom collisions, several non-perturbative theories of laser-atom interactions are discussed. Future developments of this rapidly expanding area of atomic physics are also considered.

1. INTRODUCTION

The study of atoms interacting with intense laser fields is one of the most rapidly growing areas of atomic physics. Advances in laser technology have led to the discovery of new phenomena in laser-atom interactions, mainly via the investigation of multiphoton processes. Lasers are now available, which are capable of producing oscillating electric fields of strength \mathcal{E}_0 larger than the atomic unit of electric field $e/a_0^2 \simeq 5.1 \times 10^9 \text{ Vcm}^{-1}$, the corresponding intensity being $I \simeq 3.5 \times 10^{16} \text{ Wcm}^{-2}$.

In this article we shall review some recent progress made in studying the interaction of atoms and ions with intense laser fields. We begin in Section 2 by giving a survey of several important multiphoton processes: multiphoton ionization, harmonic generation and laser-assisted electron-atom collisions. In Section 3, we discuss non-perturbative theories which have been developed to analyze these phenomena. Finally, in Section 4, we examine some possible future developments. Comprehensive reviews of the physics of atoms in intense laser fields have been written recently by Burnett et al ¹⁾ and Joachain ²⁾. Detailed articles covering various aspects of the subject may also be found in a volume edited by Gavrila ³⁾.

2. MULTIPHOTON PROCESSES

The multiphoton phenomena occurring in laser-atom interactions are conveniently divided in two classes: laser-induced and laser-assisted processes. Laser-induced processes can only occur significantly in the presence of the laser field, while laser-assisted processes can take place in the absence of the laser field, but are modified in its presence. In what follows we shall consider two types of laser-induced processes: multiphoton ionization of atoms and ions and harmonic generation. Laser-assisted processes, on the other hand, will be illustrated by examining electron-atom collisions in the presence of a laser field.

2.1 Multiphoton ionization of atoms and ions

We begin by considering the multiphoton (single) ionization (MPI) reaction



where $A(i)$ is an atom (ion) in state i and $A^+(f)$ the corresponding ionized system in state f . This process was first observed in 1965 by Voronov and Delone ⁴⁾, who used a ruby laser to induce seven photon ionization from the ground state of xenon, and by Hall et al. ⁵⁾, who recorded two-photon ionization from the negative ion I^- . Important results were then obtained by several groups concerning the dependence of the ionization yields on the laser intensity, the resonant enhancement of MPI and multiphoton multi-ionization.

A crucial step forward in the understanding of MPI was made when experiments detecting the ionized electrons were performed. Thus Agostini et al ⁶⁾, who measured the energy of the photoelectrons produced by the MPI reaction, discovered that the ejected electron could absorb

photons in excess of the minimum required for MPI. The study of this excess-photon ionization, currently known as "above threshold ionization" (ATI) has been one of the central themes of multiphoton physics in recent years. The ATI photoelectron energy spectra consist of several peaks, separated by the photon energy $\hbar\omega$. As the laser intensity increases, the number of these peaks becomes larger, the ratios of their intensities do not follow the predictions of perturbation theory, and the lowest order peaks are reduced and eventually suppressed. The reason for this peak suppression is that the energies of the atomic states are shifted in the presence of a laser field. For low frequencies (e.g. the Nd-YAG laser, for which $\hbar\omega = 1.165$ eV) the dynamic Stark shifts for the lowest bound states (in particular the ground state) are small in magnitude. On the other hand, the shifts of the Rydberg and continuum states are essentially given by the electron ponderomotive energy U_p , which is the kinetic energy due to the quiver motion of the electron in the laser field, averaged over a laser cycle. This ponderomotive energy is given by

$$U_p = \frac{2\pi e^2 I}{m c \omega^2} \quad (2)$$

where m is the electron mass, I is the laser intensity and ω its angular frequency. For example, in the case of the Nd-YAG laser, $U_p = \hbar\omega = 1.165$ eV when $I \simeq 10^{13}$ Wcm $^{-2}$. Since the energies of the Rydberg and continuum states are shifted upwards relative to the lower bound states by about U_p , there is a corresponding increase in the ionization threshold. If this threshold increase is such that $E_i + n\hbar\omega - U_p < 0$, where E_i is the unperturbed (field free) energy of the initial state, then the peak in the ATI spectrum corresponding to ionization by n photons will be suppressed. Interesting effects in ATI spectra due to the influence of the laser pulse duration have also been reported ^{7,8)}.

When the ponderomotive energy U_p is much larger than the ionization potential, ionization occurs by the tunneling mechanism, i.e. the electron tunnels through the barrier formed by the atomic potential and the laser electric field ⁹⁾, provided that the field is of low frequency and not so strong that field ionization occurs. In this case the regular structure of ATI peaks disappears. Field ionization ¹⁰⁾ takes place when the laser intensity reaches a critical value I_c (which is about 1.4×10^{14} Wcm $^{-2}$ for the ground state of atomic hydrogen) such that the saddle point of the barrier is so lowered that the electron can "flow over the top"; in this regime the atom ionizes in about one orbital period.

If both the intensity and the frequency of the laser field are large, it is predicted that the atom should become stable against ionization, and that the ionization rate should decrease as a function of the intensity in the high frequency, high intensity regime ^{11,12)}. This "stabilization" of atoms at super-intensities is currently attracting considerable attention.

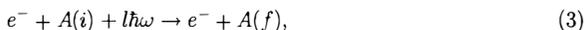
2.2 Harmonic generation

Atoms irradiated by laser light can emit radiation, whose frequency is an odd multiple $N\omega$ of the laser frequency ω because of parity conservation. This harmonic generation process has been the subject of active investigations for many years. Third harmonic generation in a rare gas medium was observed for the first time in 1967 by New and Ward ¹³⁾ while fifth - and

higher - order harmonics were seen a few years later. Recently, it has been shown¹⁴⁻¹⁶⁾ that very high (odd) harmonics could be generated when atoms interact with intense laser pulses. For example, in the experiments of L'Huillier and Balcou¹⁶⁾, performed with a Nd-YAG laser of photon energy $\hbar\omega = 1.165$ eV, the harmonic order $N = 133$ was observed in the case of neon. High order harmonic generation has attracted considerable interest as a potential source of short pulse, high frequency coherent radiation. The harmonic intensity distribution consists of a rapid decrease over the first few harmonics, followed by a plateau of approximately constant intensities and a cut-off. This general structure is in good qualitative agreement with calculated single-atom spectra, indicating that propagation effects in the medium play a minor role, or affect all the harmonics in the same way. The existence of a plateau in the harmonic intensity distribution is a non-perturbative effect which appears to be a very general characteristic of strongly driven non-linear systems.

2.3 Laser-assisted electron-atom collisions

Electron-atom collisions in the presence of laser field (also called free-free transitions) have attracted a great deal of attention not only because of the importance of these processes in applied areas (such as plasma heating), but also in view of their interest in fundamental atomic collision theory. In the first experiments performed by Andrick and Langhans¹⁷⁾ with low-intensity lasers, the transfer of only one photon between the electron-atom system and the laser field was observed. Subsequently, however, several experiments¹⁸⁻²⁰⁾ have been carried out in which the transfer of several photons has been observed in laser-assisted collisions of the type



where positive integer values of l correspond to photon absorption (inverse bremsstrahlung), negative integer ones to photon emission (stimulated bremsstrahlung) and $l = 0$ to a collision process without net absorption or emission of photons but in the presence of the laser field.

It is interesting to note that the observation of multiphoton phenomena in free-free transitions requires much lower intensities than for laser-induced processes. For example, significant rates for the emission or absorption of 10 laser quanta have been observed¹⁹⁾ in laser-assisted "elastic" electron-argon collisions at a laser intensity of about 10^8 Wcm⁻².

3. NON-PERTURBATIVE THEORIES

Most of the early theoretical treatments of multiphoton processes relied on the use of perturbation theory. However, in the case of strong laser fields, non-perturbative methods are required. We shall discuss below two of these methods: the Floquet theory and the direct numerical integration of the time-dependent Schrödinger equation.

3.1 Floquet theory

Although more general laser fields can be considered (for example "two-color " fields) we shall assume that we are dealing with a laser field wich is treated classically as a spatially homogeneous, monochromatic electric field of angular frequency ω . Then, in both the length or the velocity gauge, the Schrödinger equation describing the atomic system in the presence of the laser field is

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = [H_{at} + H_{int}(t)] |\Psi(t)\rangle \quad (4)$$

where H_{at} is the field free atomic Hamiltonian and the laser-atom interaction term $H_{int}(t)$ can be written in the form

$$H_{int}(t) = H_+ e^{-i\omega t} + H_- e^{i\omega t} \quad (5)$$

where H_+ and H_- are time-independent operators. The Hamiltonian of the system, $H(t) = H_{at} + H_{int}(t)$, is periodic , i.e. $H(t+T) = H(t)$, where $T = 2\pi/\omega$. The Floquet method can then be used to write the state vector $|\Psi(t)\rangle$ in the form

$$|\Psi(t)\rangle = e^{-iEt/\hbar} |F(t)\rangle \quad (6)$$

where the "quasi-energy" E does not depend on time and $|F(t)\rangle$ is periodic in time, with period T , so that it can be expressed as the Fourier series

$$|F(t)\rangle = \sum_{n=-\infty}^{+\infty} e^{-in\omega t} |F_n\rangle \quad (7)$$

The $|F_n\rangle$ are called the harmonic components of $|F(t)\rangle$. Using equations (4)-(7) one obtains for the harmonic components $|F_n\rangle$ the time-independent infinite system of coupled equations

$$(E + n\hbar\omega - H_{at})|F_n\rangle = H_+ |F_{n-1}\rangle + H_- |F_{n+1}\rangle \quad (8)$$

These equations, together with appropriate boundary conditions, form a problem which in general must be solved by keeping only a finite number of harmonic components, i.e. by truncating the sum on n in eq. (7) .

Two non-perturbative methods have been used successfully to solve the coupled equations (8). The first one, wich in practice is limited to one-electron systems, is the Sturmian-Floquet method²¹⁾. In this approach, each harmonic component in position space $F_n(\mathbf{r}) \equiv \langle \mathbf{r} | F_n \rangle$ is expanded on a discrete basis set consisting of spherical harmonics and complex Sturmian radial functions. This method has been applied extensively to study multiphoton ionization and harmonic generation in atomic hydrogen. A review of this work has been given by Potvliege and Shakeshaft²²⁾. We note in particular the good agreement between the Sturmian-Floquet calculations of Dörr et al.²³⁾ and the experimental data of Rottke et al.²⁴⁾ for multiphoton ionization of H(1s).

The second non-perturbative method is a new approach - the R-matrix-Floquet theory - which has been proposed recently²⁴⁾ to analyze multiphoton processes. It combines the powerful R-matrix theory with the Floquet method to treat multiphoton ionization, harmonic generation and laser-assisted electron-atom collisions in an unified way. It is completely ab initio and is

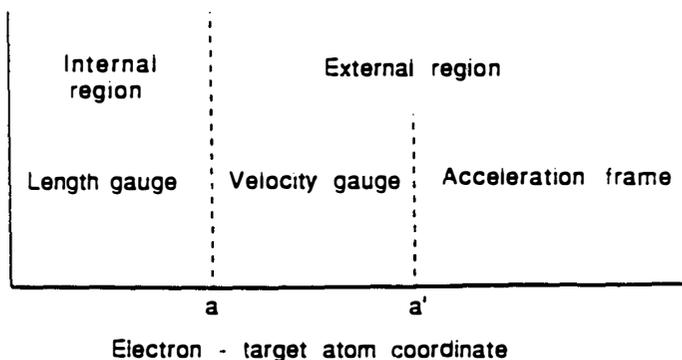


Figure 1: Partitioning of configuration space in the R-matrix-Floquet theory.

applicable to an arbitrary atom or ion. According to the R-matrix method, configuration space is subdivided into two regions (see Fig. 1). The internal region is defined by the condition that the coordinates r_i of all N electrons of the atomic system are such that $r_i \leq a$, where the sphere of radius a envelops the charge distribution of the target atom(ion) states retained in the calculation. Hence in this region exchange effects involving all N electrons are important. The external region is defined so that one of the N electrons lies on or outside the sphere of radius a , while the remaining $(N - 1)$ "target" electrons are confined within this sphere. Thus in this region exchange effects between this one electron and the remaining $(N - 1)$ "target" electrons are negligible. The Schrödinger equation (4) is solved in these two regions separately by using the Floquet method, and the solutions are then connected on the boundary at $r = a$. In the internal region it is convenient to use the length gauge since in this gauge the laser-atom coupling tends to zero at the origin. In the external region it is advantageous to use the velocity gauge out to a radius a' which may extend to infinity. In certain cases it is useful to transform at $r = a'$ to the Kramers acceleration frame which enables simple asymptotic boundary conditions to be defined. The R-matrix-Floquet theory has already been applied with great success to analyze multiphoton processes in atomic hydrogen^{25,26}, in the two-electron systems H^- and He^{27} and in more complex atoms such as neon and argon²⁸. Branching ratios into the ATI channels have been obtained, as well as angular distributions for the photoelectrons. Atomic stabilization at high frequencies and high intensities has been studied, and novel non-perturbative correlation effects²⁷ and laser-induced degeneracies²⁸ have been found.

3.2. Numerical integration of the time-dependent Schrödinger equation

One of the major non-perturbative techniques which has been used in recent years to study atoms in strong laser fields is the direct, numerical solution of the corresponding time-

dependent Schrödinger equation. This approach, which is particularly useful for very short laser pulses (where the Floquet method is not applicable), has been used in two different ways. Firstly, "numerical experiments" of multiphoton processes have been conducted in studies using one dimensional potentials²⁹⁾. Secondly, numerical methods have been used to solve the time-dependent Schrödinger equation in the dipole approximation for "realistic" (three-dimensional) atoms in pulsed laser fields³⁰⁾. Most of these computations have been single electron calculations that are "exact" for hydrogenic systems, but only approximate for atoms or ions with more than one electron. Very recently, atomic hydrogen in a super-intense, high frequency field has been studied beyond the dipole approximation³¹⁾ by solving numerically the time-dependent Schrödinger equation. In addition, the coupling of the magnetic field to the spin of the electron has also been included by solving the time-dependent Pauli equation. Adiabatic stabilization was obtained at high intensities, and corrections to the dipole approximation were found to modify the ionization probability only slightly.

4. FUTURE DEVELOPMENTS

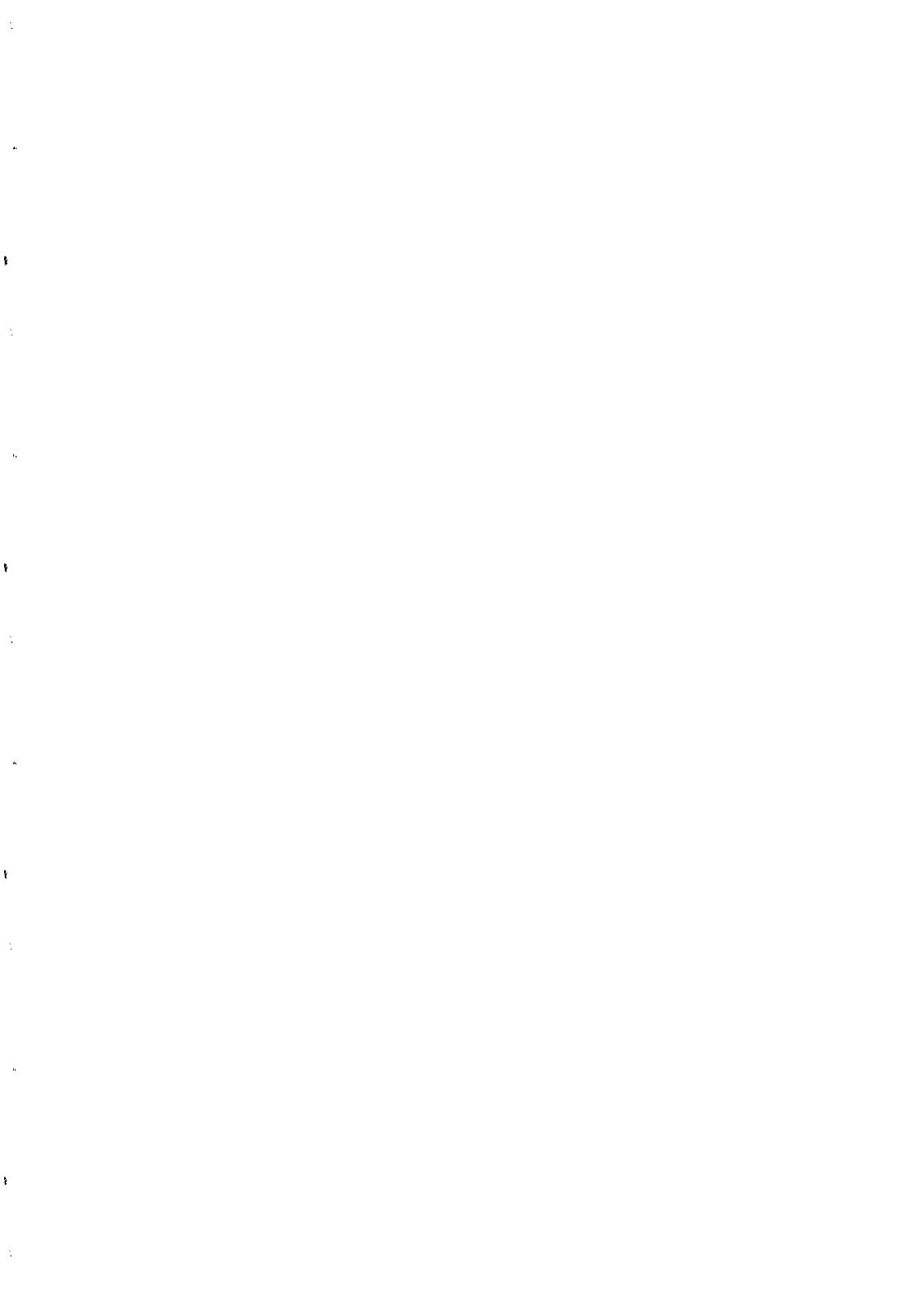
Lasers will soon become available, which can deliver ultra-intense (10^{20} W cm⁻²) and ultra short (femtosecond) pulses. The study of laser-atom interactions under these conditions will require the development of non-perturbative theories taking fully into account relativistic effects. New phenomena will also become accessible to investigation, such as the excitation and ionization of inner shell atomic electrons, electron-positron pair production and multiphoton Compton scattering, leading to very high harmonic generation. Clearly, this area of research holds great potential for future developments.

References

1. K. Burnett, V.C. Reed and P.L. Knight, *J. Phys. B* 26, 561 (1993).
2. C.J. Joachain, "Theory of Laser-Atom Interactions", in "Laser Interactions with Atoms, Solids and Plasmas" (Plenum, New York, 1994).
3. M. Gavrilá (ed), *Adv. At. Mol. Phys. Suppl.* 1 (1992).
4. G.S. Voronov and N.B. Delone, *JETP Lett.* 1, 66 (1965).
5. J.L. Hall, E.J. Robinson and L.M. Branscomb, *Phys. Rev.Lett.* 14, 1013 (1965).
6. P. Agostini, F.Fabre, G.Mainfray, G. Petite and N. Rahman, *Phys. Rev. Lett.* 42, 1127 (1979).
7. R.R. Freeman, P.H. Bucksbaum, H. Milchberg, S. Darack, D. Schumacher and M.E. Geusic, *Phys. Rev. Lett.* 59, 1092 (1987).
8. G. Petite, P. Agostini and F. Yergeau, *J.Opt.Soc.Am. B* 4, 765 (1987).
9. L.V. Keldysh, *Sov. Phys. JETP* 20, 1307 (1965).

10. S. Augst, D. Strickland, D.D. Meyerhofer, S.L. Chin and J.H. Eberly, *Phys.Rev.Lett.* 63, 2212 (1989).
11. M. Janjusevic and M.H. Mittleman, *J.Phys.* B 21, 2279 (1988).
12. M. Pont and M. Gavrilu, *Phys.Rev.Lett.* 65, 2362 (1990).
13. G.H.C. New and J.F. Ward, *Phys.Rev.Lett.* 19, 556 (1967).
14. A. McPherson, G. Gibson, H. Jara, U. Johann, T. S. Luk, I. McIntyre, K. Boyer and C. K. Rhodes, *J. Opt. Soc. Am. B* 4, 595 (1987).
15. M. Ferray, A. L'Huillier, X. F. Li, L.A. Lompré, G. Mainfray and C. Manus, *J.Phys.* B 21, L31 (1988).
16. A. L'Huillier and P. Balcou, *Phys.Rev.Lett.* 70, 774 (1993).
17. D. Andrick and L. Langhans, *J.Phys.* B 9, L459 (1976).
18. A. Weingartshofer, J. Holmes, G. Caudle, E. Clarke and H. Krüger, *Phys.Rev.Lett.* 39, 269 (1977).
19. A. Weingartshofer, J. K. Holmes, J. Sabbagh and S. L. Chin, *J.Phys.* B 16, 1805 (1983).
20. B. Wallbank, J. K. Holmes and A. Weingartshofer, *Phys. Rev. A* 40, 5461 (1989).
21. A. Maquet, S. I. Chu and W. P. Reinhardt, *Phys. Rev. A* 27, 2946 (1983).
22. R. M. Potvliege and R. Shakeshaft, *Adv. At. Mol. Phys. Suppl.* 1, 373 (1992).
23. M. Dörr, R. M. Potvliege and R. Shakeshaft, *Phys. Rev. Lett.* 64, 2003 (1990).
24. P. G. Burke, P. Francken and C. J. Joachain, *Europhys. Lett.* 13, 617 (1990); *J. Phys.* B 24, 761 (1991).
25. M. Dörr, M. Terao-Dunseath, J. Purvis, C. J. Noble, P. G. Burke and C. J. Joachain, *J. Phys.* B 25, 2809 (1992).
26. M. Dörr, P. G. Burke, C. J. Joachain, C. J. Noble, J. Purvis and M. Terao-Dunseath, *J. Phys.* B26, L275 (1993).
27. J. Purvis, M. Dörr, M. Terao-Dunseath, C. J. Joachain, P. G. Burke and C. J. Noble, *Phys. Rev. Lett.* 71, 3943 (1993).
28. O. Latinne, N. Kylstra, M. Dörr, J. Purvis, M. Terao-Dunseath, C. J. Joachain, P. G. Burke and C. J. Noble, to be published.
29. J. H. Eberly, J. Javanainen and K. Rzazewski, *Phys. Rep.* 204, 331 (1991).
30. K. C. Kulander, K. J. Schafer and J. L. Krause, *Adv. At. Mol. Phys. Suppl.* 1, 247 (1992).
31. O. Latinne, C. J. Joachain and M. Dörr, *Europhys. Lett.* (in press).

GRAVITATION



THE STRING DILATON AND THE EQUIVALENCE PRINCIPLE

Thibault Damour

Institut des Hautes Etudes Scientifiques

91440 Bures sur Yvette, France and,

DARC, CNRS-Observatoire de Paris, 92195 Meudon, France

It is pointed out that string-loop effects may generate matter couplings for the dilaton allowing this scalar partner of the tensorial graviton to stay massless while contributing to macroscopic gravity in a way naturally compatible with existing experimental data. Under a certain assumption of universality of the dilaton coupling functions, the cosmological evolution drives the dilaton towards values where it decouples from matter. At the present cosmological epoch, the coupling to matter of the dilaton should be very small, but non zero. This provides a new motivation for improving the experimental tests of Einstein's Equivalence Principle.

String theory is, at present, the only scheme promising to provide a combined quantum theory of gravity and of the gauge interactions. It is striking that, within string theory, the usual Einsteinian tensor graviton is intimately mixed with a scalar partner: the dilaton. The matter couplings of the dilaton *a priori* generate drastic deviations from general relativity, notably violations of Einstein's Equivalence Principle (EP): universality of free fall, constancy of the constants, . . . This is why it is generally assumed that the dilaton will acquire a (Planck scale) mass due to some yet unknown dynamical mechanism.

An alternative possibility has been recently suggested [1]: string-loop effects (associated with worldsheets of arbitrary genus in intermediate string states) may naturally reconcile the existence of a massless dilaton with existing experimental data if they exhibit the same kind of universality as the tree level dilaton couplings.

More precisely, let us assume that the effective action for the string massless modes (considered directly in four dimensions) takes the form

$$S = \int d^4x \sqrt{\hat{g}} B(\Phi) \left\{ \frac{1}{\alpha'} [\hat{R} + 4\hat{\square}\Phi - 4(\hat{\nabla}\Phi)^2] - \frac{k}{4} \hat{F}^2 - \overline{\Psi} \hat{D}\Psi + \dots \right\}, \quad (1)$$

where the function of the dilaton Φ appearing as a common factor in front is given by a series of the type

$$B(\Phi) = e^{-2\Phi} + c_0 + c_1 e^{2\Phi} + c_2 e^{4\Phi} + \dots \quad (2)$$

The first term on the right-hand side of Eq. (2) is the string tree level contribution (spherical topology for intermediate worldsheets) which is known to couple in a universal multiplicative manner [2] [3] [4]. The further terms represent the string-loop effects: the genus- n string-loop contribution containing a factor $g_s^{2(n-1)}$ where $g_s \equiv \exp(\Phi)$ is the string coupling constant. Apart from the fact that Eq. (2) is a series in powers of g_s^2 , little is known about the global behaviour of the dilaton coupling function $B(\Phi)$. For the cosmological attractor mechanism of Ref. [1] to apply, it is sufficient that $B(\Phi)$ admit a local maximum. [The universality of the factor $B(\Phi)$ seems necessary to ensure the observed smallness of possible deviations from general relativity. Ref. [1] suggests to use this universality as a criterion for selecting a preferred class of string models].

It is convenient to transform the action (1) by introducing several Φ -dependent rescalings. In particular, one replaces the original "string-frame" metric $\hat{g}_{\mu\nu}$ by a conformally related "Einstein-frame" metric $g_{\mu\nu} \equiv C B(\Phi) \hat{g}_{\mu\nu}$, and the original dilaton field Φ by a canonical scalar field φ . The transformed action reads

$$S = \int d^4x \sqrt{g} \left\{ \frac{1}{4q} R - \frac{1}{2q} (\nabla\varphi)^2 - \overline{\Psi} D\Psi - \frac{k}{4} B(\varphi) F^2 + \dots \right\}, \quad (3)$$

where $q \equiv 4\pi\overline{G} \equiv \frac{1}{4} C\alpha'$ denotes a bare gravitational coupling constant and $B(\varphi) \equiv B[\Phi(\varphi)]$.

The basic clue followed in Ref. [1] to relate string models to the observed low-energy world is the dilaton dependence of the gauge coupling constants apparent in Eq. (3): $g^{-2} = kB(\varphi)$. This dependence implies that the QCD mass scale Λ_{QCD} is given by

$$\Lambda_{QCD}(\varphi) \sim C^{-1/2} B^{-1/2}(\varphi) \exp[-8\pi^2 b_3^{-1} k_3 B(\varphi)] \widehat{\Lambda}_s, \quad (4)$$

where b_3 is a (rational) one-loop coefficient associated with the scale dependence of the SU(3) coupling constant, and where $\widehat{\Lambda}_s \simeq 3 \times 10^{17} \text{GeV}$ [5] is the (string-frame) string unification scale $\propto \alpha'^{-1/2}$. The Einstein-frame mass of hadrons is essentially some pure number times $\Lambda_{QCD}(\varphi)$. More generally, under the assumption of a universal $B(\varphi)$, the masses of all the particles will depend on φ only through the function $B(\varphi)$:

$$m_A(\varphi) = m_A[B(\varphi)]. \quad (5)$$

When studying the cosmological evolution of the graviton-dilaton-matter system, one finds that the dilaton vacuum expectation value φ is dynamically driven toward the values φ_m corresponding to a local maximum of $B(\varphi)$, i.e. a local minimum of all the various mass functions $m_A(\varphi)$. [With some important physical differences, this cosmological attractor mechanism is similar to the one discussed in Ref. [6] which concerned metrically-coupled tensor-scalar theories]. The main parameter determining the efficiency of the cosmological relaxation of φ toward φ_m is the curvature κ of the function $\ln B(\varphi)$ near the maximum φ_m :

$$\ln B(\varphi) \simeq \text{const.} - \frac{1}{2} \kappa (\varphi - \varphi_m)^2. \quad (6)$$

Because of the steep dependence of $m_A(\varphi)$ upon $B(\varphi)$ [illustrated by Eq. (4)], each ‘‘mass threshold’’ during the radiation-dominated era [i.e. each time the cosmic temperature T becomes of order of the mass m_A of some particle] attracts φ towards φ_m by a factor $\sim 1/3$ (see Fig. 1 of [1]). In the subsequent matter-dominated era, φ is further attracted toward φ_m by a factor proportional to $Z_0^{-3/4}$ where $Z_0 \simeq 1.3 \times 10^4$ is the redshift separating us from the end of the radiation era. Finally, in the approximation where the phases of the ten or so successive relaxation oscillations around φ_m undergone by φ during the cosmological expansion are randomly distributed, one can estimate that the present value φ_0 of the dilaton differs, in a *rms* sense, from φ_m by

$$(\varphi_0 - \varphi_m)_{rms} \sim 2.75 \times 10^{-9} \times \kappa^{-3} \Omega_{75}^{-3/4} \Delta\varphi, \quad (7)$$

where $\Omega_{75} \equiv \rho_0^{\text{matter}}/1.0568 \times 10^{-29} \text{g cm}^{-3}$ and where $\Delta\varphi$ denotes the deviation of φ from φ_m at the beginning of the (classical) radiation era. The estimate (7) assumes that $\kappa \gtrsim 0.5$. Actually, a numerical calculation of $\varphi_0 - \varphi_m$ (assuming a φ -dependence of the particle masses

of the type shown in Eq. (4) finds that the attraction power of the radiation era may be more effective than that assumed in the analytical estimate (7): see Fig. 2 of [1].

The scenario of Ref. [1] predicts the existence of many small, but non zero, deviations from general relativity. Indeed, a cosmologically relaxed dilaton field couples to matter around us with a strength (relative to usual gravity)

$$\alpha_A = \left. \frac{\partial \ln m_A(\varphi)}{\partial \varphi} \right|_{\varphi_0} \simeq \beta_A(\varphi_0 - \varphi_m) \quad (8)$$

with $\beta_A \simeq \beta_3 \equiv 40.8\kappa$ for hadronic matter. Therefore, all deviations from Einstein's theory contain a small factor $(\varphi_0 - \varphi_m)^2$ coming from the exchange of a φ excitation. More precisely, the post-Newtonian deviations from general relativity at the present epoch are given by the Eddington parameters

$$1 - \gamma_{\text{Ed}} \simeq 2(\beta_3)^2(\varphi_0 - \varphi_m)^2, \quad (9)$$

$$\beta_{\text{Ed}} - 1 \simeq \frac{1}{2}(\beta_3)^3(\varphi_0 - \varphi_m)^2, \quad (10)$$

while the residual cosmological variation of the coupling constants is at the level

$$\frac{\dot{\alpha}}{\alpha} \simeq -\kappa \left[\omega \tan \theta_0 + \frac{3}{4} \right] (\varphi_0 - \varphi_m)^2 H_0, \quad (11)$$

$$\frac{\dot{G}}{G} \simeq -2\beta_3^2 \left[\omega \tan \theta_0 + \frac{3}{4} \right] (\varphi_0 - \varphi_m)^2 H_0, \quad (12)$$

where $\omega \equiv \left[\frac{3}{2} \left((\beta_3 - \frac{3}{8}) \right) \right]^{1/2}$, and where θ_0 denotes the phase of the matter-era relaxation toward φ_m , while H_0 denotes the present value of Hubble's "constant".

The most sensitive way to look for the existence of a weakly coupled massless dilaton is through tests of the universality of free fall. The interaction potential between particle A and particle B is $-G_{AB}m_A m_B / r_{AB}$ where $G_{AB} = \bar{G}(1 + \alpha_A \alpha_B)$. Therefore two test masses, A and B , will fall in the gravitational field generated by an external mass m_E with accelerations a_A and a_B differing by

$$\left(\frac{\Delta a}{a} \right)_{AB} \equiv 2 \frac{a_A - a_B}{a_A + a_B} \simeq (\alpha_A - \alpha_B) \alpha_E. \quad (13)$$

The difference $\alpha_A - \alpha_B$ introduces a small factor proportional to the ratio $m_{\text{quark}}/m_{\text{nucleon}}$ or to the fine structure constant α . Finally, Ref. [1] predicts an equivalence-principle violation of the form ⁽¹⁾

$$\left(\frac{\Delta a}{a} \right)_{AB} = \kappa^2 (\varphi_0 - \varphi_m)^2 \left[C_B \Delta \left(\frac{B}{M} \right) + C_D \Delta \left(\frac{D}{M} \right) + C_E \Delta \left(\frac{E}{M} \right) \right]_{AB}, \quad (14)$$

⁽¹⁾ The consequences of the form (14) for the choice of materials in Equivalence Principle experiments is discussed in the contribution of T. Damour and J.P. Blaser to these proceedings.

where $B \equiv N + Z$ is the baryon number, $D \equiv N - Z$ the neutron excess, $E \equiv Z(Z - 1)/(N + Z)^{1/3}$ a Coulomb energy factor and M the mass of a nucleus. The only coefficient in Eq. (14) which can be reliably estimated is the last one: $C_E \simeq 3.14 \times 10^{-2}$. The largest $\Delta a/a$ will arise in comparing Uranium ($E/M \simeq 5.7$) with Hydrogen (or some other light element). For such a pair Eq. (14) yields

$$\left(\frac{\Delta a}{a}\right)_{\max} = 0.18 \kappa^2 (\varphi_0 - \varphi_m)^2. \quad (15)$$

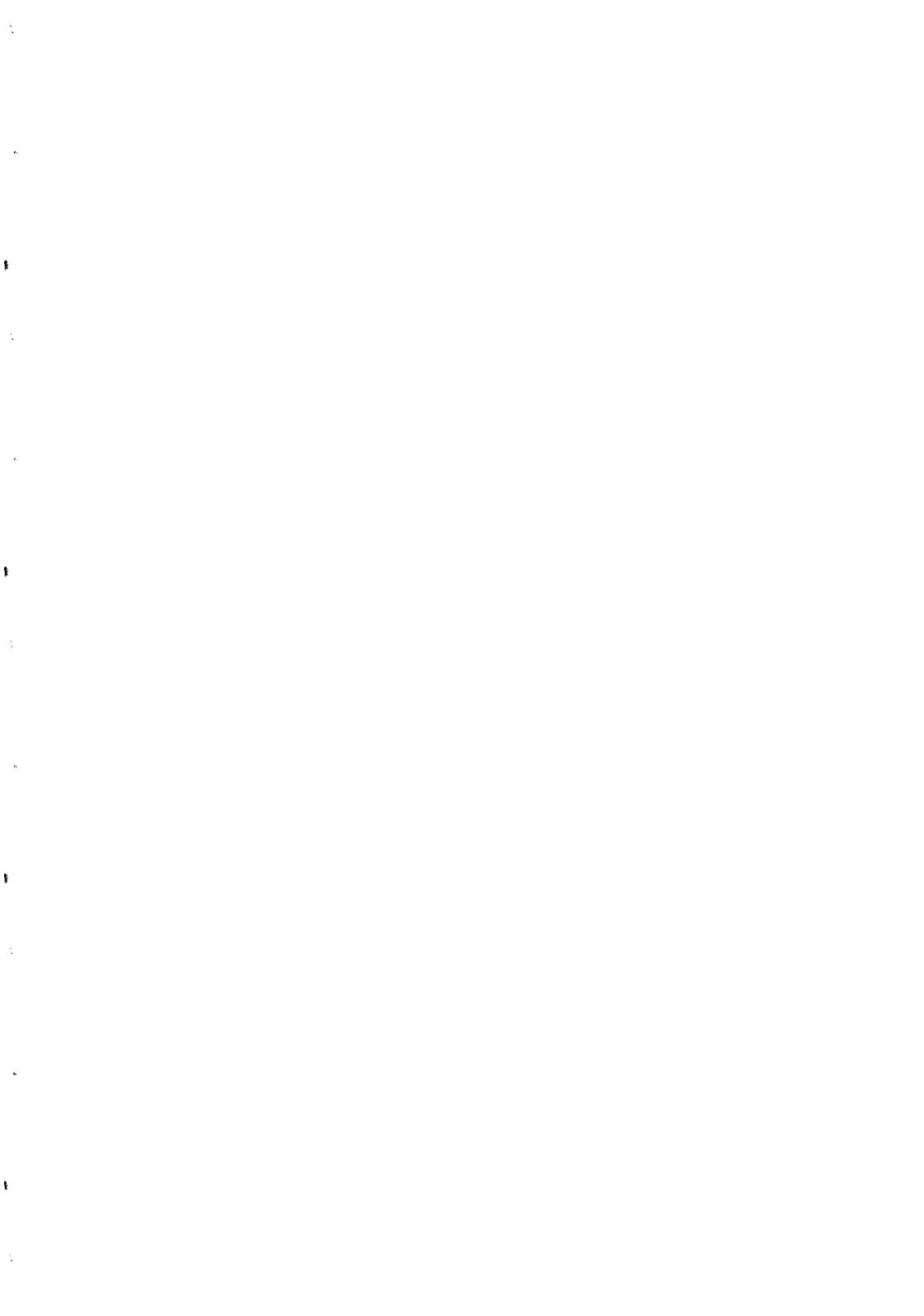
Inserting the analytical estimate (7) [which is probably optimistic in view of the smaller result obtained from numerical calculations] leads to

$$\left(\frac{\Delta a}{a}\right)_{rms}^{\max} = 1.36 \times 10^{-18} \kappa^{-4} \Omega_{75}^{-3/2} (\Delta\varphi)^2. \quad (16)$$

The results (9)-(12) and (14)-(16) provide a new motivation for trying to improve by several orders of magnitude the experimental tests of general relativity, notably the tests of the equivalence principle (universality of free fall, constancy of the constants, . . .). Although the results of Ref. [1] summarized here do not provide a precise target to shoot for, they remove of the *a priori* objections to spending time and money on improving tests of the equivalence principle: the fact that it is already tested at the $\sim 10^{-12}$ level [7] [8]. Indeed, the scenario of Ref. [1] gives an example of a well-motivated theoretical model containing no small parameters and naturally predicting very small deviations from general relativity at the present epoch. In this model, high-precision tests of the equivalence principle can be viewed as low-energy windows on string-scale physics: not only could they discover the dilaton, but, by measuring the ratios C_B/C_E , C_D/C_E in Eq. (14) they would probe some of the presently most obscure aspects of particle physics: Higgs sector and unification of coupling constants.

REFERENCES

- [1] T. Damour and A.M. Polyakov, preprint IHES/P/94/1, submitted to Nucl. Phys. B.
- [2] E.S. Fradkin and A.A. Tseytlin, Phys. Lett. **B158**, 316 (1985).
- [3] C.G. Callan, D. Friedan, E.J. Martinec and M.J. Perry, Nucl. Phys. **B262**, 593 (1985).
- [4] C.G. Callan, I.R. Klebanov and M.J. Perry, Nucl. Phys. **B278**, 78 (1986).
- [5] V.S. Kaplunovsky, Nucl. Phys. **B307**, 145 (1988).
- [6] T. Damour and K. Nordtvedt, Phys. Rev. Lett. **70**, 2217 (1993); Phys. Rev. D **48**, 3436 (1993).
- [7] J.O. Dickey et al., Science (1994) in press.
- [8] G. Smith, contribution to these proceedings.



NEW TESTS OF THE EQUIVALENCE PRINCIPLE

G.L. Smith

Nuclear Physics Lab, GL-10, University of Washington, Seattle, WA 98195

The Eöt-Wash group at the University of Washington has tested the equivalence principle by comparing the accelerations of beryllium, aluminium, and copper in the field of the earth. These tests would be sensitive to equivalence principle violating interactions with ranges $1\text{m} \leq \lambda \leq \infty$ and a broad regime of charges. Our results for the equivalence principle parameter η are: $\eta(\text{Be} - \text{Cu}) = (-1.9 \pm 2.5) \times 10^{-12}$ and $\eta(\text{Be} - \text{Al}) = (-0.2 \pm 2.8) \times 10^{-12}$,

We have also compared the accelerations of our 'ordinary' matter test bodies toward the unidentified, and potentially exotic, 'dark' matter at the center of our galaxy. No violation of the equivalence principle is observed, providing laboratory evidence that acceleration towards dark matter is predominately gravitational.

1 Introduction

General relativity and all metric theories of gravitation include the equivalence principle, EP, as a basic postulate. The EP demands that the motion of a test mass in an accelerating frame, which is governed by its inertial mass, be indistinguishable from the motion of the test mass in a gravitational field, which is governed by its gravitational mass.

An EP violating interaction can be expressed by a Yukawa potential:

$$V = \alpha_5 \left(\frac{q_5}{\mu} \right)_1 \left(\frac{q_5}{\mu} \right)_2 \frac{e^{-r/\lambda}}{r}, \quad (1)$$

where $\alpha_5 = \pm gGm_1m_2/(4\pi)$, g is the EP-violating coupling constant, G is the Newtonian gravitational constant, q_5 is the 'charge' of this unknown interaction, μ is the mass of each object in amu, and λ is the range of the interaction.

The EP parameter η describes the ratio of inertial to gravitational mass, or the size of the EP violation, between two materials

$$\eta = \frac{\left[\frac{(m_i/m_g)_1 - (m_i/m_g)_2}{(m_i/m_g)_1 + (m_i/m_g)_2} \right]}{\frac{\delta a}{a_g}}, \quad (2)$$

where δa is the differential acceleration of the two bodies and a_g is the gravitational acceleration. The classic tests of the EP by the Princeton[1] and Moscow[2] groups constrained η for ranges greater than an astronomical unit: $\eta = (-1.3 \pm 1.5) \times 10^{-11}$ for Au-Al[1], and $\eta = (3.0 \pm 4.5) \times 10^{-13}$ for Pt-Al[2]. We have developed an apparatus which complements these experiments and tests the EP over a broader regime of ranges, $1\text{m} \leq \lambda \leq \infty$, and with sensitivity to a wider range of possible interaction charges.

2 Experimental details

A complete description of our apparatus exists in the literature[3]. We compare the horizontal accelerations of two materials in the field of the earth, and search for a difference ascribable to their differing compositions. We use the Eöt-Wash torsion balance located on the University of Washington campus.

The body of the 4-fold symmetric torsion pendulum, which hangs from an 80 cm long 25 μm diameter tungsten wire, contains two pairs of test bodies arranged in a composition dipole. Any differential acceleration between the two materials causes a torque about the fiber axis, which is opposed by the restoring torque of the fiber, with torsional spring constant of $k \approx 0.033$ erg/rad. We modulate the proposed compositional-dependent torque by rotating our balance slowly about the fiber axis and probe for a sinusoidal variation in the equilibrium position of the pendulum.

Of course the trick is isolating effects due to 'normal physics' from that due to an EP violating interaction. A complete description of our systematic error suppression and analysis is given in ref[3]. In brief, the pendulum is placed inside a vacuum vessel evacuated to ≈ 1 torr, and is centered within a 3-layer magnetic shield, a set of gravity gradient compensators, and layers of both active and passive heat shields. The test bodies have identical outside dimensions, surface coatings (Au), masses, and vanishing gravitational quadrupole moments. The vessel and magnetic shielding sit upon a high quality turntable, and are rotated at $1/8$ the free torsional frequency of the pendulum, or about every 1.5 hours. Approximately every 25 seconds (every $1/32$ of a torsional period) we

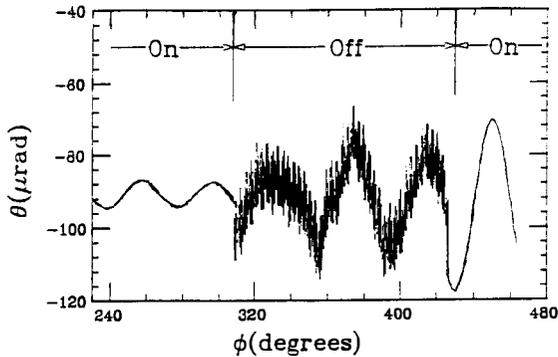


Figure 1: Effect of the rotation rate feedback system. The system was turned off and then back on; its effect can be seen on the measurement of the pendulum's deflection angle. Each high-frequency oscillation is due to one rotation of the worm gear in the turntable drive.

record the angular position of the turntable, the deflection of the pendulum in the rotating frame, and 20 other quantities used to monitor systematic effects.

Improvements implemented since ref[3] are described below.

2.1 Feedback system in the turntable drive.

Any short duration (compared to the ≈ 12 min period of the pendulum) variation in the angular velocity of the turntable generates an apparent torque on the pendulum, because the autocollimator measuring the deflection of the pendulum is rigidly attached to the turntable, while the pendulum is nearly inertial because of the weak torsional spring constant of the fiber. The turntable rotation rate feedback system consists of 900,000 pulse/revolution optical shaft encoder, whose output is locked to a temperature-controlled crystal oscillator. As shown in Fig.1 turning off the feedback system introduces high frequency noise to the deflection signal.

Fig.2 shows the Fourier transforms of the deflection signal when the turntable is rotating and when it is stationary. Note that with the feedback system operating the rotation does not introduce any significant additional noise. The solid line in Fig.2 shows the theoretical noise in an oscillator with our resonant frequency and Q value in thermal equilibrium at room temperature, indicating that we are operating very close to the thermal limit for an ideal oscillator with our Q value.

2.2 Gravitational gradient compensators

Fig.3 shows the effect of rainwater soaking into the surrounding hillside after a dry summer. This changed the mass distribution of the hillside and thus the size of the gravitational gradients at the location of the test bodies. This effect is less than 2% of the hillside's gradient before we flattened the field at the location of the balance. As we couldn't control the weather, we made our pendulum more perfect by adding two orthogonal pairs of 0.041 g Al screws, on both the top and bottom counterweights of the pendulum. These were then adjusted to reduce our sensitivity to the leading order gravitational gradient by more than a factor of 7.

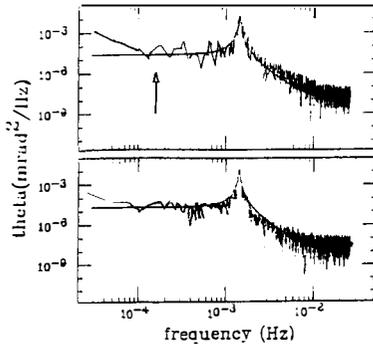


Figure 2: Spectral power densities of the angular deflection signal. Top panel: turntable rotating; the arrow shows the frequency of our signal. Center panel: turntable stationary. The smooth curves in the upper two plots show the predicted thermal noise levels plus the noise 'floor' from the electronics

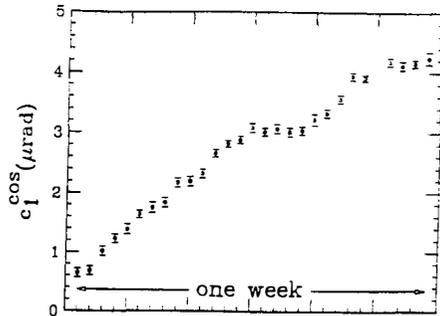


Figure 3: Rainwater saturating the ground as seen by our gradiometer test bodies.

2.3 Analysis

Our analysis procedure has been modified since the publication of ref[3]. We start with raw measurements of the deflection of the pendulum as a function of the turntable orientation, with 32 measurements being taken per free torsional cycle. We first apply a digital notch filter to eliminate the free torsional motion of the pendulum by averaging points $1/2$ torsional period apart. We next apply another digital filter to eliminate the thermally driven linear drift and, as a byproduct, the even harmonics of the turntable rotation frequency, by subtracting points that are $1/2$ turntable period apart. Then, to simplify the subsequent analysis, we average 6 adjacent points. We finally divide the data into 'cuts' of exactly 3 turntable rotations, and fit this filtered and averaged data to odd harmonics of the rotation frequency — the fundamental being the frequency of interest — and Legendre polynomials to account for the nonperiodic 'drift' terms.

After roughly 30 'cuts' we interchange the test bodies on the pendulum and repeat our measurements. The phase of an EP violating interaction must similarly change to follow the direction of the dipole on the pendulum, while most systematic effects (those not associated with the test bodies themselves) will not. In this way we are able to separate instrumental offsets from any composition-dependent torque.

Evidence that this analysis system works well is seen in our sensitivity test, Fig.4, where we have exploited the 4-fold symmetry of the pendulum to put a known gravitational torque on the pendulum. This torque was provided by four 32.5 gm masses placed 32.2 cm from the center of the balance. Because of the 4-fold symmetry the torque is modulated at 4 times the rotation frequency. The observed amplitude of $(1.5 \pm 0.3) \times 10^{-9}$ dyne-cm agrees well with the calculated amplitude of

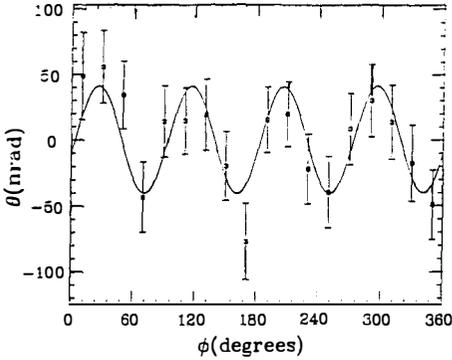


Figure 4: Sensitivity test with a 4-fold symmetric gravitational source which generated a 4ω signal of the same amplitude as an EP-violating interaction with $\eta=2.1\times 10^{-11}$. The best fit signal, shown as a smooth curve, has an amplitude and phase, (44 ± 8) nrad and $(45\pm 3)^\circ$, in excellent agreement with the expected values of 38 nrad and 45° .

$(1.25 \pm 0.01) \times 10^{-9}$ dyne-cm. Note that a torque of this magnitude (if it were at the fundamental frequency) would correspond to an EP-violating signal so small that it would be consistent with the null result quoted by the Princeton group[1].

3 Results for δa towards ordinary matter

Our 1σ results are

$$\eta(\text{Be} - \text{Cu}) = (-1.9 \pm 2.5) \times 10^{-12} \quad (3)$$

$$\eta(\text{Be} - \text{Al}) = (-0.2 \pm 2.8) \times 10^{-12},$$

indicating no EP violation to this high level of sensitivity.

Fig.6 shows our constraints on α_5 as described in Eqn.1, interpreted for $q_5 =$ baryon number and for $q_5 = (B - L)/\sqrt{2}$. We see that this experiment provides the most stringent constraints on new interactions over 9 orders of magnitude. The 'gap' where we do not quote results arises because our earth models are inadequate at these ranges.

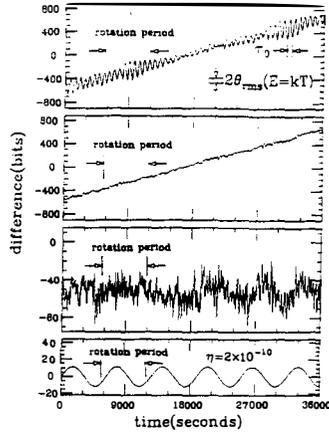


Figure 5: Top panel: raw autocollimator readings. Upper middle panel: after applying the filter that suppressed free torsional oscillations. Lower middle panel: after applying the filter that suppresses fiber drift. Bottom panel: expected signal from an EP-violating interaction with $\eta=2\times 10^{-10}$

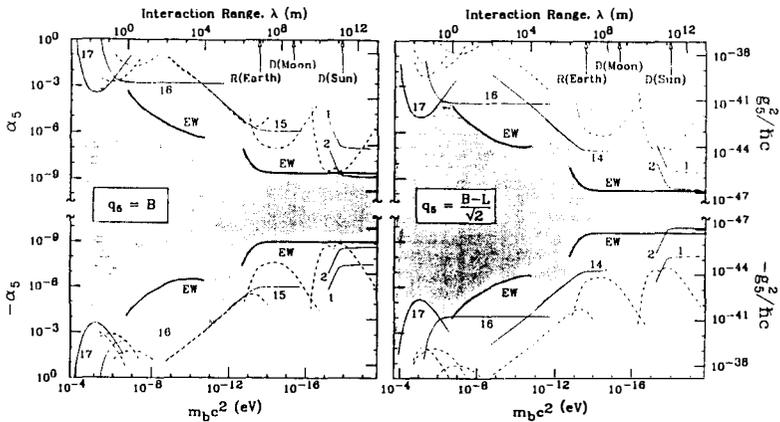


Figure 6: 2σ upper limits on the coupling strength of a new Yukawa interaction as a function of the interaction range λ . Results are shown for possible charges of $q_5 = B$ (baryon number) and for $q_5 = (B - L)/\sqrt{2}$. Heavy solid lines labeled EW are from this work. Light solid curves are labeled by reference numbers in ref [4]. The experimentally allowed region is shaded.

4 Results for δa towards dark matter

Astronomical studies of rotation curves in spiral galaxies, as well as studies of the motions of galactic clusters, show luminous objects (stars or galaxies) experience a greater acceleration than can be accounted for by the amount or distribution or other luminous matter. This extra acceleration is usually attributed to a gravitational field of nonluminous 'dark matter'. This unidentified dark matter is thought to make up 90% of the total matter in the universe. It is interesting to see what laboratory experiments can say about the nature of the acceleration towards the potentially exotic dark matter. Using the analysis described in ref [5] we have performed an EP test for three 'normal-matter' composition dipoles, Be-Cu, Be-Al and Si/Cu-Al, falling toward the dark matter at the center of our galaxy. ($\approx 25\%$ of our sun's galactocentric acceleration is thought to be due to dark matter[7].)

As discussed in ref[5], significant vector interactions between dark and ordinary matter are ruled out by existing experiments and the requirement that the dark matter halo be stable. Scalar interactions, however, are not similarly excluded. Calculation of possible scalar charges are highly model dependent, so we have used a tree level approximation

$$q_5 = \frac{Z(q_p - q_e) + (A - Z)q_n}{A} \quad (4)$$

where q_p , q_e , and q_n are the scalar charges of the free proton, electron and neutron respectively.

Our limits on the possible non-gravitational (EP violating) acceleration of neutral hydrogen falling toward the dark matter, as a function of the possible scalar charge are shown in Fig.7.

One can see that even where the scalar charge is proportional to the masses of the free particles, where our sensitivity is weakest, the acceleration of neutral hydrogen toward dark matter must be predominately gravitational. Nordtvedt[6] has pointed out that the Earth-Moon differential

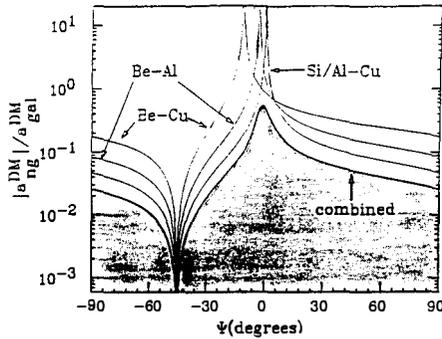


Figure 7: 1σ constraints on the non-gravitational acceleration of neutral hydrogen due to a hypothetical long-range interaction with dark matter. The vertical axis shows the ratio of the anomalous non-gravitational acceleration to the total acceleration towards the dark matter. The horizontal axis describes the possible scalar charge, parametrized with $\psi = \arctan((q_5^+ + q_5^p - q_5^-)/q_5^0)$

acceleration, measured with very high precision in lunar ranging experiments, provides a stronger constraint on the EP violating acceleration of neutral hydrogen for most values of ψ . However, because the lunar experiment has no sensitivity at a point near $\psi = 0$, where the scalar charge difference between the earth and moon vanishes, the lunar results do not significantly improve the ‘worst case’ scenario shown in Fig.7.

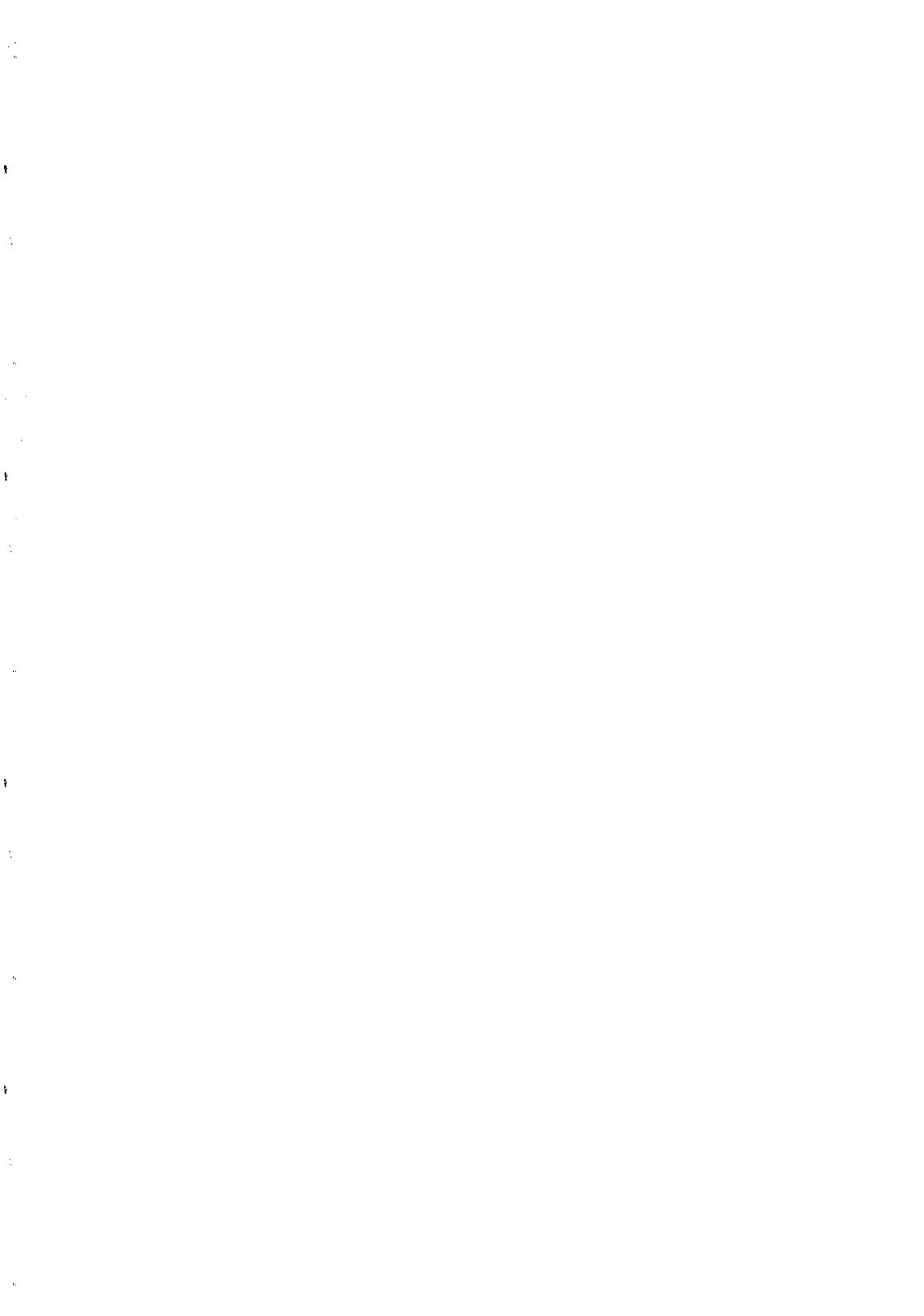
A more thorough presentation of this experiment will be published shortly[4]

5 Acknowledgements

I would like to acknowledge the contribution of all the members of the Eöt-Wash collaboration and the financial support of the National Science Foundation, grant PHY9104541.

References

- [1] P.G. Roll, R. Krotkov and R.H. Dicke, *Ann. Phys. (N.Y.)* **26**, 442 (1964).
- [2] V.B. Braginski and V.I. Panov, *Zh. Eksp. Teor. Fiz.* **61**, 873 (1971) [*Sov. Phys. JETP* **34**,463 (1972)].
- [3] E.G. Adelberger *et al.*, *Phys. Rev. D* **42**, 3267 (1990).
- [4] Y. Su, *et al.*, *Phys. Rev. D*, in press.
- [5] G.L. Smith, E.G. Adelberger, B.R. Heckel and Y. Su, *Phys. Rev. Lett.* **70**,123 (1993).
- [6] K.L. Nordtvedt, private communication (1994).
- [7] C.W. Stubbs, *Phys. Rev. Lett* **70**, 119 (1993).



EÖTVÖS, a new concept for low disturbance measurements in space

Jean-Pierre Blaser
Paul Scherrer Institute
CH 5232 Villigen, Switzerland

Ho Jung Paik
University of Maryland
College Park, MD 20742, USA

(paper presented by Thibault Damour)



Summary

In space, for measurements like the test of the equivalence principle, a limiting source of error on differential accelerometers is the residual noise of the drag-free system. To allow a further increase in sensitivity, it is suggested to make a spin-stabilised instrument virtually drag-free by placing it free inside an outer spacecraft (s/c) which follows it by servoed thrusters. For cryogenic instruments cooling is done by gas conduction. The mechanical separation of the instrument from the outer s/c allows the use of cryo-coolers instead of liquid Helium. Two temperature levels of the instrument allow to dispose of electronics heat dissipation of the order of Watts. For a normal temperature instrument radiation cooling could remove tens of Watts.

1. Introduction

A number of fundamental measurements like testing the equivalence principle, the inverse-square law, a "fifth force", the gravitomagnetic force, as well as determining the gravitational constant G and gradiometry for geodesy, require the measurement of extremely small accelerations. Experiments on Earth use torsion balances to compensate gravity. One of the limitations is the acceleration noise due to microseism. Going to space allows to work in weightless state and to benefit from a quieter environment. For high sensitivity, however, even space is not free from disturbing accelerations, and these residual forces (air drag, radiation pressure and electromagnetic disturbances) must be compensated by a drag-free system.

Future missions like GP-B and STEP will use proportional Helium thrusters to compensate drag accelerations. However, though such servo systems can be made extremely sensitive at a given frequency, their residual noise increases away from this optimised frequency. As an example, in STEP, compensation is possible down to $E-12 \text{ m/s}^2/\text{Sqrt(Hz)}$ at $3E-4 \text{ Hz}$ but rises by a factor 1000 towards the frequency of $E-2 \text{ Hz}$, required for geodesy signals. Therefore, an ultimate factor limiting the precision is the coupling of residual drag noise through the differential accelerometer alignment errors.

2. Concept of EÖTVÖS

To overcome the limitations due to the residual drag, the only possibility is to have the instrument block (called the "instrument" in what follows) carrying the accelerometers freely floating and shielded from all outside disturbances by an outer spacecraft. Their relative position is measured optically and a servo system actuates thrusters in order to have the outer s/c follow the instrument. The latter acts as the reference proof mass and, in principle, follows an orbit determined uniquely by the Earth gravitational field. The main advantage of the concept is the possibility of using a conventional drag-free system correcting displacement to a few tenths of a millimeter, without having to compensate acceleration to a very high level.

The price to pay for these great advantages is twofold. First, the instrument must be supplied with power and signals exchanged without mechanical effects. Secondly, the instrument must be thermally controlled, which, in the case of a cryogenic instrument, is not trivial.

In what follows, the case of an experiment complement like that of STEP is discussed as an example. The basic detectors for testing the equivalence principle (EP) are differential accelerometers composed of two concentric cylinders. The instrument makes essential use of superconductivity for levitation, for position measurements by SQUIDS, as well as for shielding against magnetic fields. This requires temperatures below about 5K.

3. Mechanical design

It is suggested to have the instrument slowly rotating, with some 1000 s period. This frequency ω_s , added to the orbit frequency ω_o , determines the frequency of a possible EP violation signal ($\omega_s \pm \omega_o$). It can be chosen arbitrarily and changed to eliminate errors correlating with the orbit revolution period or other disturbances.

The orientation of the accelerometers must be known and adjustable. Coarse sensing from the outer s/c can be done by star trackers. The three axis geodesy gradiometers give very sensitive orientation signals relative to the Earth field. The instrument behaving like a gyroscope, its orientation must be changed by applying torques, which make it precess. This can be done by the system of electrostatic electrodes shown in figure 1.

The rotation of the instrument itself could be adjusted by means of induction motors. The Helium gas will slow down this rotation in days or weeks if the outer s/c is in a fixed orientation. If on the other hand it is co-rotating, the rotation period will stay constant. This latter scheme is adopted here as it has further advantages as discussed below. It requires, however, that the center of mass of the outer s/c be adjustable to fall exactly on the axis of free rotation of the instrument.

4. Thermal design

A *cryogenic* instrument like STEP could be cooled by thermal conduction in dilute Helium gas at a pressure slightly above the molecular regime. It is clear that only an instrument having two temperature levels is possible (see figure 1). At the superconductivity level T_1 only very few milliwatts of dissipation can be tolerated. The main dissipation of the electronics which absolutely must be located on the instrument has to be removed at a higher temperature level T_2 , where the Carnot efficiency and gas conduction are much higher. Temperatures T_2 of 25 to 50 K are suggested, where most electronics would work. The critical point is the heat leaks between the two parts at T_1 and T_2 . There are three: heat conduction through the support axle, heat flow through the gas (molecular regime) and the thermal conductivity of the wires connecting the electronics (these could possibly be HiTc superconductors). One should be able to keep the total leak to a few mW.

The actual dissipation of the low temperature electronics is therefore the crucial element and determines the cooling schemes which could be considered. Three variants are:

- Completely mechanical cooling by Stirling cycle cryo-coolers down to T_2 and by combined Stirling/Joule-Thompson units down to T_1 . Typical cooling powers of existing, partly already space-qualified units are 2 W and 10 mW, respectively. Any Helium tides are eliminated.
- Stirling cryo-cooler down to T_2 and a Helium Dewar for the low temperature part. Several schemes involving superfluid as well as supercritical Helium are conceivable.

The detailed problems of gravitational disturbances by Helium tides must be taken into account, though quite small quantities of Helium should suffice.

- A fully Helium-based cooling is only practical for dissipations below the Watt level and for short mission duration. It would require rather large Helium quantities and the tide problems remain.

The second scheme seems particularly attractive.

For *non-cryogenic* accelerometers (e.g. using electrostatic sensors) the single instrument block could be in vacuum and radiation cooled, a few tens of Watts being removable at temperatures of 250-300K.

5. Position measurement, power and signal transmission

The position of the rotating instrument relative to the outer s/c must be measured continuously in all six degrees of freedom. Modern micro-optoelectronic techniques should permit measurements to a few microns, this being especially easy in the case of co-rotation.

Power of a few Watts must be supplied to the instrument without any mechanical disturbance and with no significant heat input. This could be achieved by a microwave link, quite easily in the case of co-rotation. Superconducting devices are, however, very sensitive to microwaves so a high shielding factor is necessary.

Both command and measuring signals must be transmitted wireless between the instrument and the outer s/c. The low data rates needed should be easy to transmit by digital IR links.

6. Orbit

The choice depends on the actual goals of a mission. Cryo-coolers need a few hundred Watts of power, so sufficiently illuminated solar panels are required. Eclipses and unfavourable panel orientation can be critical. In the case of an equivalence principle test, a sun-synchronous orbit and a mission of less than six months is suitable. The orbit precession would then require periodic or continuous readjustments of the rotation axis of the instrument. For a geodetic mission, an exactly polar orbit of longer duration would be best, raising the problems of panel illumination mentioned above.

7. Disturbances and error budget

The instrument rotating rather fast will give rise to inertial forces: centrifugal accelerations and the effects of a possible nutation. However, most gravity and acceleration errors are modulated at frequencies other than the EP signal frequency ($\omega_s \pm \omega_e$). In particular, the centre of mass matching requirement, arising from coupling to the Earth gravity gradient, is reduced by two orders of magnitude from the non-spinning case. This in turn relaxes the charge control requirement.

The instrument being the reference proof mass, its centre of mass will be inertial. Therefore, the accelerometers located at a distance from this point will experience tidal forces due to the gradients of the Earth gravity field. If these forces have to be rejected

to $E-18 \text{ g}$ ($g=9.81\text{m/s}^2$), the sensitive axes of the accelerometers must be aligned to $5E-6 \text{ rad}$. The uneven mass distribution in the outer *s/c* and its relative movements due to the drag-free operation will exert gravitational forces on the instrument. These will, on one hand, lead to a self-acceleration of the whole, which cannot be exactly known. On the other, it will give rise to time-varying gradients disturbing the geodesy gradiometry. These gravitational forces will also exert small torques on the instrument and make it precess. This small effect is absent in the case of co-rotation.

In the case of a three-axis gradiometer, the two axes oriented perpendicular to the spin axis will experience a large centrifugal acceleration in addition to a fully modulated Earth gradient. The co-rotation will stabilize the spin frequency and keep the centrifugal acceleration constant. A time-varying part of this acceleration can be removed to the first order by using the gradiometers themselves as attitude sensors and imposing the tracelessness condition for the gravity gradient tensor.

The full modulation of the Earth gravity requires a dynamic range of $2E7$ for gradient sensitivity of $E-4 \text{ E}$ ($1 \text{ Eötvös}=E-9 \text{ s}^{-2}$). The SQUID electronics can provide such a dynamic range at the low frequencies involved, but it poses a great challenge to *AVD* converters and scale factor linearity. These problems can of course be alleviated if one sticks to a single axis configuration, as baselined for STEP.

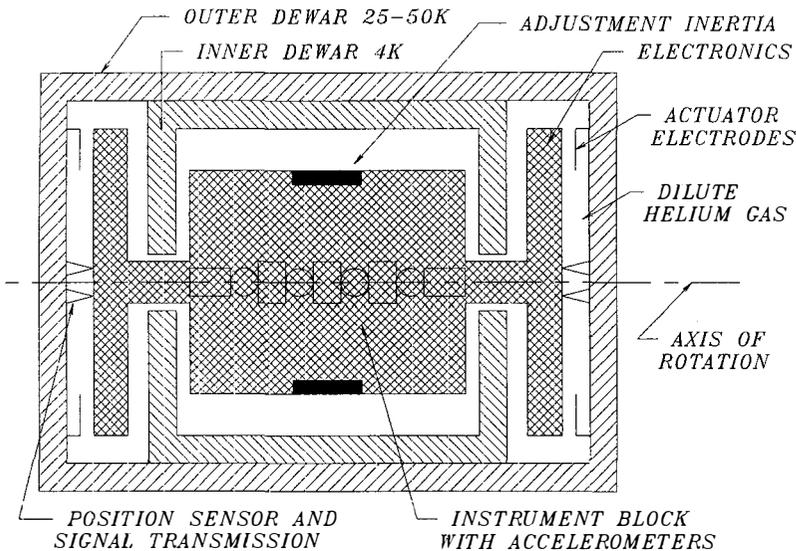
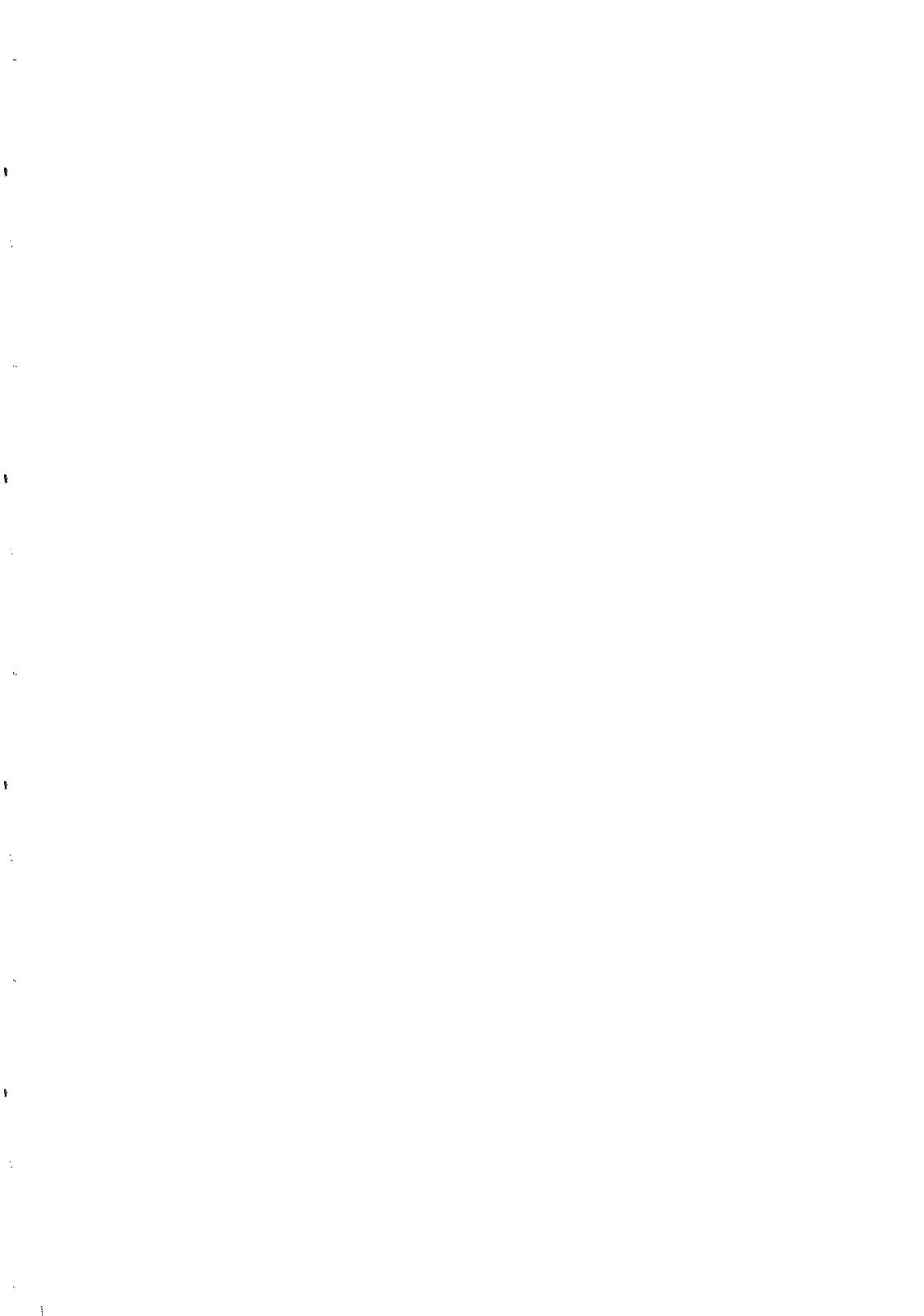


Figure 1: Schematic structure of a cryogenic spacecraft with fully inertial instrument



**OPTIMIZING THE CHOICE OF MATERIALS
IN EQUIVALENCE PRINCIPLE EXPERIMENTS**

Thibault Damour

Institut des Hautes Etudes Scientifiques

91440 Bures sur Yvette, France, and

DARC, CNRS-Observatoire de Paris, 92195 Meudon, France

Jean-Pierre Blaser

Paul Scherrer Institute

CH-5232 Villingen, Switzerland

A strategy is proposed for choosing the materials tested in equivalence principle experiments. This strategy combines a theory-independent approach with a theory-dependent one. The theory-dependent approach is conveniently formulated in geometrical terms: the possible materials are represented as a cloud of points in a multi-dimensional space of “elementary charges”. The optimum choices of material configurations are those which maximize the volume spanned by the “vectors” connecting the elements.

1. Introduction

This paper takes the optimistic view that presently planned improved Equivalence Principle (EP) tests will give positive, i.e. non null, results. In such a situation, it seems important to choose the pairs of materials tested so as to maximize, at once: (i) our confidence in the reality of the EP violation signals, (ii) the quantity of information that we can extract, independently of any theoretical model, from the EP data, and (iii) the precision with which we can measure the coupling parameters appearing in the presently most plausible theoretical models of EP violation. The first two aims must be reached within a model-independent approach, while the third one will be based on a specific, though rather general, assumption about the composition-dependence of the EP violating signals.

In all cases, we shall assume that we are looking for EP violation signals caused by some *long-range* interaction. In the non-relativistic approximation, the effect of a long-range interaction is to add to the usual gravitational interaction potential $-Gm_A m_E / r_{AE}$ between body A and the external body E (e.g. the Earth), a supplementary contribution $-HQ_A Q_E / r_{AE}$, where H is a new coupling constant (which can be positive or negative) and Q_A the “charge” of body A to which the new interaction is coupled. The effective (composition-dependent) gravitational constant for the (AE) pair is

$$G_{AE} = G + H \frac{Q_A Q_E}{m_A m_E} = G \left[1 + \frac{H}{Gu^2} \frac{Q_A Q_E}{\mu_A \mu_E} \right], \quad (1)$$

where u denotes one atomic mass unit, and $\mu_A \equiv m_A / u$. The fractional difference in free fall acceleration of the pair (AB) , $(\Delta a/a)_{AB} = 2(a_A - a_B)/(a_A + a_B)$ is given by

$$\left(\frac{\Delta a}{a} \right)_{AB} \simeq \frac{G_{AE} - G_{BE}}{G} = q_A - q_B, \quad (2)$$

in which we have defined

$$q_A \equiv \frac{H}{Gu^2} \frac{Q_E Q_A}{\mu_E \mu_A} \equiv h \widehat{Q}_E \widehat{Q}_A, \quad (3)$$

with $h \equiv H/(Gu^2)$ and $\widehat{Q}_A \equiv Q_A/\mu_A$.

In this work we assume that the external source E is fixed and we focus on the A -dependence of the “effective charge” q_A .

2. Model-independent approach

Let \mathcal{C} denote the “configuration” of materials being tested, i.e. the set of all pairs (AB) whose gravitational acceleration is compared. We consider that the Equivalence Principle data consist of a set of (noisy) differential acceleration measurements m_{AB} indexed by the elements of \mathcal{C} :

$$\mathcal{D}_{\mathcal{C}} = \{m_{AB}; (AB) \in \mathcal{C}\} \quad (4)$$

measure the combination (3) modulo an arbitrary additive constant. This is not even enough information to determine the sign of H . [Even if one can include the external material E among the tested materials, the additive constant in q_A prevents one from finding the sign of h , let alone its magnitude].

3. Model-dependent approach

Let us now assume some theoretical model predicting a composition dependence of the reduced charge $\widehat{Q}_A \equiv Q_A/\mu_A$ of the form

$$\widehat{Q}_A = \beta_0 + \sum_{i=1}^n \beta_i \xi_A^i, \quad (6)$$

where β_0, β_i are some unknown coupling parameters, and ξ_A^i some known (reduced) “elementary charges”. For instance, if we assume that the apparent EP violation is due to a long-range vector field (an extra $U(1)$ gauge field [1]) we expect to have only two independent elementary charges [2], say $\xi^1 = (N + Z)/\mu$ and $\xi^2 = (N - Z)/\mu$ (where N =neutron number, Z =proton number=atomic number), and $\beta_0 = 0$. In the case of the EP violation due to the dilaton scalar field of String Theory [3], we expect to have three independent elementary charges,

$$\xi^1 = (N + Z)/\mu, \quad (7a)$$

$$\xi^2 = (N - Z)/\mu, \quad (7b)$$

$$\xi^3 = E/\mu \simeq Z(Z - 1)/((N + Z)^{1/3}\mu), \quad (7c)$$

with E denoting a contribution proportional to the Coulomb interaction energy of a nucleus. Moreover, in that case we expect to have a non zero composition-independent contribution: $\beta_0 \neq 0$ in (6). From Eqs. (3) and (6) we conclude that the theoretically expected composition dependence of q_A is of the form

$$q_A = \alpha_0 + \sum_{i=1}^n \alpha_i \xi_A^i \quad (8)$$

where $\alpha_0 \equiv h(\beta_0 + \sum_j \beta_j \xi_E^j)\beta_0$, $\alpha_i \equiv h(\beta_0 + \sum_j \beta_j \xi_E^j)\beta_i$.

The A -independent contribution α_0 in Eq. (8) is not accessible from the measurements (5). We are looking for the choice of material configuration \mathcal{C} which optimizes the measurement of the n effective coupling parameters α_i .

Once the α_i are known, it will be possible to determine the original coupling parameters h, β_0 and β_i (modulo the evident rescaling freedom $\beta_0 \rightarrow \lambda\beta_0, \beta_i \rightarrow \lambda\beta_i, h \rightarrow \lambda^{-2}h$) if and only if one knows something about the relative value of β_0 with respect to the β_i 's.

Each measurement in the data set \mathcal{D}_C is a sum

$$m_{AB} = q_A - q_B + n_{AB}, \quad (5)$$

where q_A is the theoretical quantity (3) and n_{AB} is some noise contribution.

The noise n_{AB} will contain both statistical and systematic contributions. The first include thermal noise, inherent SQUID noise, and mechanical noise due to the drag free system. Systematics will come from the structure of the differential accelerometers with test masses of widely different densities, mechanical misalignments, electrostatic charging, etc... In evaluating the performances of the various checks discussed below (cyclic conditions, connectedness of elements) the constraints coming from noise considerations will have to be taken into account.

The problem we discuss in this section is that of choosing the configuration C so as to maximize the quality and quantity of information one can extract from \mathcal{D}_C without assuming any specific law for the composition-dependence of the EP violation model for the new long-range interaction, i.e. any particular form for the composition-dependence of the charge Q_A , i.e. for the A -dependence of the effective charge q_A .

In absence of theoretical assumptions about the A -dependence of q_A , the maximum information we can hope to extract from Eq. (5) is the value of q_A , *modulo an unknown additional constant*, for some elements A . This suggests that an optimum strategy will be to restrict oneself to *connected* configurations C , i.e. set of pairs $(A_i A_j)$ such that any two elements A_1, A_n can be connected by a sequence of pairs belonging to C . Indeed, the choice of disconnected configurations (e.g. $\{(AB), (BC), (DE)\}$) will introduce more than one unknown additional constant in the determination of the possible values of q_A .

Moreover, as one wishes to establish as firmly as possible the reality of a violation of the EP, i.e. the fact that $q_A - q_B$ is non zero for some pairs, it seems important to include *topological loops* in the configuration C . The simplest way to do it would be to include a null pair (AA) . However, the corresponding check, namely $m_{AA} \simeq 0$, is not very convincing as it does not involve any "positive" measurement. The next simplest way is to include two similar pairs. $\{(AB), (BA)\}$. The corresponding check being $m_{AB} + m_{BA} \simeq 0$, and involving (hopefully) non null results. Higher order loops can be triangular $\{(AB), (BC), (CA)\}$, quadrangular $\{(AB), (BC), (CD), (DA)\}$, etc... Evidently, one will have to weigh the conceptual advantages of long loops against their higher sensitivity to partial failures. From this point of view, binary loops $\{(AB), (BA)\}$ seem very safe ways of confirming the reality of an EP violation.

To end this section, let us note that a model-independent approach cannot give access to the basic theoretical quantities H and Q_A (or better HQ_E^2 and Q_A/Q_E). One can, at best,

If σ_{AB} denotes the standard deviation of the measurement m_{AB} , the likelihood function of the theory parameters (given some data set \mathcal{D}_C) is

$$\chi^2(\alpha_i; \text{given } \mathcal{D}_C) = \sum_{AB \in \mathcal{C}} \sigma_{AB}^{-2} (q_A - q_B - m_{AB})^2 = \sum_{AB} \sigma_{AB}^{-2} \left(\sum_i \alpha_i \xi_{AB}^i - m_{AB} \right)^2, \quad (9)$$

where we have introduced the short-hand notation

$$\xi_{AB}^i \equiv \xi_A^i - \xi_B^i. \quad (10)$$

The best-fit values of α_i are those that minimize $\chi^2(\alpha_i)$. Here we are interested in the precision with which the α 's are determined. If we assume that the measurements errors are independent gaussian variables, it is well known that the deviations $\bar{\alpha}_i = \alpha_i^{\text{best fit}}(\mathcal{D}_C) - \alpha_i^{\text{true}}$ will be zero-mean gaussian variables with distribution function $\propto \exp[-\Delta\chi^2/2]$ where $\Delta\chi^2 = \chi^2(\alpha) - \chi_{\min}^2$ can be written as

$$\Delta\chi^2 = \sum_{AB \in \mathcal{C}} \sigma_{AB}^{-2} \left[\sum_i \bar{\alpha}_i \xi_{AB}^i \right]^2 = g_{\mathcal{C}}^{ij} \bar{\alpha}_i \bar{\alpha}_j, \quad (11)$$

where we have defined

$$g_{\mathcal{C}}^{ij} \equiv \sum_{AB \in \mathcal{C}} \sigma_{AB}^{-2} \xi_{AB}^i \xi_{AB}^j. \quad (12)$$

The most naturally defined confidence regions for the $\bar{\alpha}$'s are quadrics (generalized ellipsoids) $g_{\mathcal{C}}^{ij} \bar{\alpha}_i \bar{\alpha}_j < \Delta\chi^2 = \text{const.}$, where the value of $\Delta\chi^2$ depends both on n and on the chosen level of confidence (e.g. if $n = 2$ the ellipsoids $\Delta\chi^2 = 2.3$ and $\Delta\chi^2 = 6.2$ correspond to 68% and 95% confidence regions, respectively). The optimum choice for the configuration of materials \mathcal{C} is that which leads to the smallest possible errors on the theory parameters α_i , and to the least possible correlations between the measurements of the different α_i 's. It is useful to formulate this requirement in *geometrical terms*:

Let us introduce the n -dimensional ξ -space of the (reduced) elementary charges: e.g. in the dilatonic case the 3-space spanned by the coordinates (7a)-(7c). In this space, the set of all possible materials defines a cloud of points. Choosing a configuration \mathcal{C} means choosing a geometrical figure made of a set of vectors $\xi_{AB}^i = \xi_A^i - \xi_B^i$ connecting two points in the cloud. To any such set of vectors ξ_{AB}^i we associate the tensor $g_{\mathcal{C}}^{ij}$ defined by Eq. (12). In turn this tensor defines an ellipsoid in the α -space (dual of the ξ = space) by the formula $g_{\mathcal{C}}^{ij} \bar{\alpha}_i \bar{\alpha}_j = \text{const.}$ We wish to find the configurations \mathcal{C} which are associated to the smallest and "roundest" ellipsoids in α -space. A general rule is that the more extended in ξ -space is the figure made by the connecting vectors ξ_{AB}^i , the smaller is the associated ellipsoid of errors in α -space. [For simplicity, we henceforth assume that all the measurement standard

deviations σ_{AB} are equal. This leads to a purely geometrical criterion for choosing \mathcal{C}]. More precisely, if we consider the minimal configurations made of just n vectors (smaller configurations leads to degenerate ellipsoids of errors which are infinite in some directions), it is easy to prove that the volume of the ellipsoid of errors is inversely proportional to the volume defined by the n vectors, i.e. the quantity $\epsilon_{ij\dots l} \xi_1^i \xi_2^j \dots \xi_n^l$. [In 3-space this is the “mixed” product $(\vec{\xi}_1, \vec{\xi}_2, \vec{\xi}_3) = \vec{\xi}_1 \cdot \vec{\xi}_2 \times \vec{\xi}_3$, which measures – with a sign depending on the orientation – the ordinary volume of the parallelogram built on the three vectors $\vec{\xi}_1, \vec{\xi}_2, \vec{\xi}_3$ parallelly transported so as to have the same origin]. The same inverse proportionality of volumes holds in the case where one has $n + 1$ vectors linked by one loop condition. For instance, in the two-dimensional [extra $U(1)$] case a triangular configuration $\vec{AB}, \vec{BC}, \vec{CA}$ in ξ -space gives one loop check $m_{AB} + m_{BC} + m_{CA} \simeq 0$ and defines (through Eq. (12) where the sum is taken over all three links in the configuration) an ellipse of errors in α -space whose volume is inversely proportional to the volume of the triangle ABC in ξ -space. [It should be noted that the geometry of ξ -space that we use is purely *affine*, i.e. it does not use the concepts of (euclidean) length or angle. In other words, our criteria are invariant under arbitrary rescalings and linear combinations of the elementary charges.]

4. Application

Table 1 gives the values of the reduced elementary charges (7) for a selection of materials. The quantity ξ^3 was computed by means of the approximate analytical (Bethe-Weizsäcker) formula given in Eq. (7c). [We have checked that a more accurate estimate of E/μ based on the nuclear charge density profiles probed by electron scattering experiments generally agrees with the analytical estimate within $\lesssim 20\%$].

Figure 1 represents the cloud of materials of Table 1 in the two-dimensional ξ -space appropriate to the assumption of an extra $U(1)$ vector field. An eye inspection of Fig. 1 suggests that the minimal optimum choice would be the two pairs $\{(LiH, V), (CH_2, U)\}$. However, if we consider that it may be unpractical to admit elements such as Uranium, Lithium Hydride or Polyethylene in a high-precision EP test, and if we bring in the requirements of the model-independent approach of Section 2, we might settle for three pairs based on the elements Be, Si and Au: e.g. $\{(Be, Si), (Si, Be), (Be, Au)\}$ or $\{(Be, Si), (Si, Au), (Au, Be)\}$.

Figure 2 represents the cloud of materials of Table 1 (minus Li H and CH_2) in the three-dimensional ξ -space appropriate to a dilaton-like theoretical model. If we comply with the model-free requirements of Section 2, the minimal choice will include four pairs of materials (containing one loop). There are several ways of realizing the loop: repeated segment, triangle, quadrangle. Moreover, one may wish to discard the requirement of connectedness in favour of an augmentation of the volume spanned by the three independent connecting

vectors in ζ -space. We leave to the reader the responsibility of choosing among the many possible options, using the criterions given here as guides. We note however in Fig. 2 the special position of Beryllium and the fact that most of the other elements are more or less lying in a plane. (The latter fact gives rise to correlations which strongly increase the errors).

It is a pleasure to thank Eric Adelberger, Ho Jung Paik and Chris Stubbs for useful comments.

Table 1. A selection of possible proof mass materials and the corresponding elementary charges. Neutron numbers and masses are averages weighted with natural isotope abundances. For the Hydrogen compounds, only the heavy atom is counted for the electrostatic energy.

elem	Z	N	$\left(\frac{N+Z}{\mu} - 1\right)10^3$	$\frac{N-Z}{\mu}$	$\frac{Z(Z-1)}{(N+Z)^{3/2}\mu}$
Li6H	4.	3.	-3.26644	-0.142391	0.893229
Be	4.	5.	-1.35175	0.110961	0.640133
C	6.	6.011	-0.003072	0.000916	1.09064
CH2	8.	6.	-1.12017	-0.142697	1.65779
Mg	12.	12.3202	0.62322	0.013174	1.87451
Al	13.	14.	0.684212	0.037062	1.92724
Si	14.	14.1087	0.825719	0.003870	2.13129
Ti	22.	25.93	1.0772	0.082083	2.65644
V	23.	27.9975	1.09987	0.098103	2.67853
Cu	29.	34.6166	1.11663	0.088387	3.20096
Ge	32.	40.71	1.07046	0.119919	3.27228
Zr	40.	51.3184	1.0387	0.124073	3.7975
Ag	47.	60.9632	0.881352	0.129447	4.20924
Sn	50.	68.8079	0.822075	0.158435	4.19819
Ba	56.	81.4216	0.689875	0.185118	4.3462
Ta	73.	108.	0.287415	0.193425	5.13502
W	74.	109.898	0.266057	0.195257	5.16696
Pt	78.	117.116	0.18295	0.200511	5.30813
Au	79.	118.	0.169856	0.198003	5.37659
Bi	83.	126.	0.093913	0.205761	5.48788
U	92.	146.	-0.213316	0.226842	5.67502

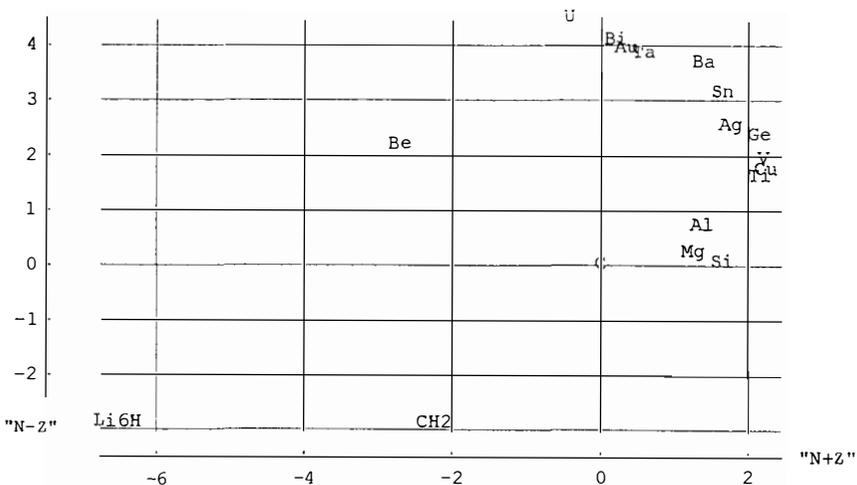


Figure 1. Two-dimensional element distribution. Scale factors have been applied here and in fig. 2 to improve the readability.

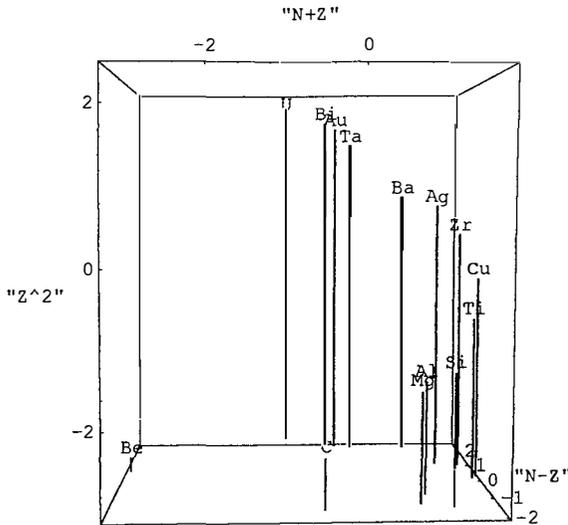


Figure 2. Three-dimensional element cloud (excluding the exotic materials Li6H and CH2). Table 2 gives some examples of the sensitivities obtained for various accelerometer arrangements

Table 2. Volume of the parallelepiped tangent to the ellipsoid of errors in α -space (= error product) for some typical configurations. In many cases correlations between the unknowns are present, leading to large error products for the coupling parameters.

number of differential accelerometers, material pairs	cyclic conditions	materials connected	error product
4 accelerometers:			
Be-Al/Al-Cu/Cu-Au/Au-Be	quadrangle	yes	33400
Be-U/U-C/C-Be/Be-V	triangle	yes	5020
CH2-Be/Be-Si/Si-CH2/Si-U	triangle	yes	253
Be-CH2/CH2-U/U-Be/Li6H-Cu	triangle	partly	150
6 accelerometers:			
Be-Si/Si-C/C-Be/Au-C/Au-Be/Au-Si	full tetrahedron	yes	64500
Mg-V/V-Au/Au-Mg/Mg-Ag/Ag-V/Ag-Au	full tetrahedron	yes	10600
Be-Zr/Zr-C/C-Be/U-CH2/U-Be/U-Si	full tetrahedron	yes	2350
Be-Si/Si-CH2/CH2-Be/U-CH2/U-Be/U-Si	full tetrahedron	yes	149
Be-V/V-CH2/CH2-Be/U-CH2/U-Be/U-V	full tetrahedron	yes	118

REFERENCES

- [1] P. Fayet, Phys. Lett. B 171, 261 (1986); Nucl. Phys. B 347, 743 (1990).
- [2] A. De Rújula, Phys. Lett. B 180, 213 (1986).
- [3] T. Damour and A.M. Polyakov, submitted to Nucl. Phys. B ; see also T. Damour, these proceedings.

PRELIMINARY RESULTS FOR ANOMALOUS SPIN COUPLING TO THE DARK
MATTER IN OUR GALAXY

Rogers C. Ritter, Linda I. Winkler*, George T. Gillies
*Department of Physics, University of Virginia
Charlottesville, VA 22901, U.S.*

**Dept. of Physics and Astronomy, Appalachian State University
Boone, NC 28608, U.S.*

ABSTRACT

Special, intrinsically compensated masses having $\sim 10^{23}$ polarized electrons, but with no externally measurable magnetic moment, have been fabricated to probe anomalous spin interactions. Previously, these were used to investigate the existence of weak, hypothetical spin-dependent forces. In the present experiment, two such masses are located on the horizontal beam of a sensitive torsion pendulum, with the spin axes horizontal, perpendicular to the supporting cross-beam, and anti-parallel to each other. Ordinary gravitational interactions at a distance will yield no torque on the pendulum beyond that due to the unavoidable very small mass difference and/or location imbalance of these two masses. Rotation of the Earth provides a scan of the sky for regions of stronger anomalous spin interaction with the masses. An essentially fixed interactive source such as a potential dark matter distribution concentric with the center of our galaxy would provide a torque on the pendulum with a 12 - sidereal hour period. Twenty nine 24-hour runs with the pendulum have been analyzed up to this time. Fitting the observed torques to an appropriate sidereal time function sets a limit of $3.5 \times 10^{-9} \text{ cm-s}^{-2}$ for the acceleration of one of these masses towards the center of the galaxy.

1. INTRODUCTION

Phillips¹ and Ni² have performed torsion pendulum tests for an anisotropy in the interaction of a laboratory spin mass and the cosmic background, essentially macroscopic versions of the Hughs-Drever anisotropy tests. The present experiment provides a more specific search for anisotropy, towards a hypothetical dark matter distribution centered at the center of our galaxy. This particular motivation follows the suggestion by Stubbs³ and experiments by the University of Washington group⁴ seeking a long range, composition-dependent interaction of ordinary matter falling toward dark matter. This dark matter distribution is inferred from the visibly missing matter needed to account for the observed dynamics of stellar motion about our galaxy's center⁵. Evidence increases for the dark matter to consist of particles outside the Standard Model of particle physics⁶, thus making it a likely source for anomalous interactions with local detectors. The experiment presented here combines features of the fifth force and spatial anisotropy searches, using a torsion pendulum to test for a possible long-range anomalous spin interaction with the nonluminous matter.

2. THE SPIN - DARK MATTER EXPERIMENT

Properties of our torsion pendulum, as used for two previous anomalous spin-dependent tests at ranges of a few centimeters, have been outlined in earlier publications^{7,8}. The primary experimental feature was the use of special Dy₆-Fe₂₃ masses having $\sim 10^{23}$ electrons aligned in intrinsic spin, while the orbital spin cancels their magnetic moment to a high degree at room temperature, thus providing a "macroscopic electron" with huge intrinsic spin but little external magnetic moment. In the present experiment the spin masses are arranged differently from the previous two cases. Now the spin masses are mounted with axes horizontal on the ends of, and perpendicular to, the support rod attached at its center to the fiber. With the mass spins antiparallel to each other this forms a spin couple. A force of the $\sigma \cdot r$ form between a distant nonpolarized source and this spin couple will create a torque about the fiber, the sense of which depends on the sign of the postulated interaction.

A major difference from the previous two experiments also lies in the fact that in those the pendulum was operated in the dynamic mode, in which the change in its oscillation period with spin direction was the signal of interest, while we now use it in the static mode, seeking patterned angular deflection of the fiber as a signal. The fiber has a torsion constant of $0.9 \text{ dyne-cm-rad}^{-1}$, leading to a period with the present mass

arrangement of 649 s.

The pendulum operates at atmospheric pressure, with a natural decay time of 6503 s, resulting in a Q of 63. A magnetic damper at the top of the fiber⁷ dissipates the simple pendulum modes but has insignificant effect on the torsional mode. Although moderately low, the damping time still allows the pendulum to execute oscillations driven by noise of several types⁹. To precondition the digitally-sampled signal (at 1000 s intervals), we pass it through a low-pass analog filter with 0.001 Hz cutoff frequency and a roll-off of 48 db per octave above cutoff. These conditions make it relatively easy to avoid aliasing from higher-frequency noise, while not damping any significant components of the postulated dark matter signal. Digital filtering acts subsequently, by appropriate, successive smoothing of the signal with a hanning shape. Figure 1 exhibits a representative, unfiltered signal, along with the (unnormalized) pattern predicted for a σ - r interaction towards the center of our galaxy.

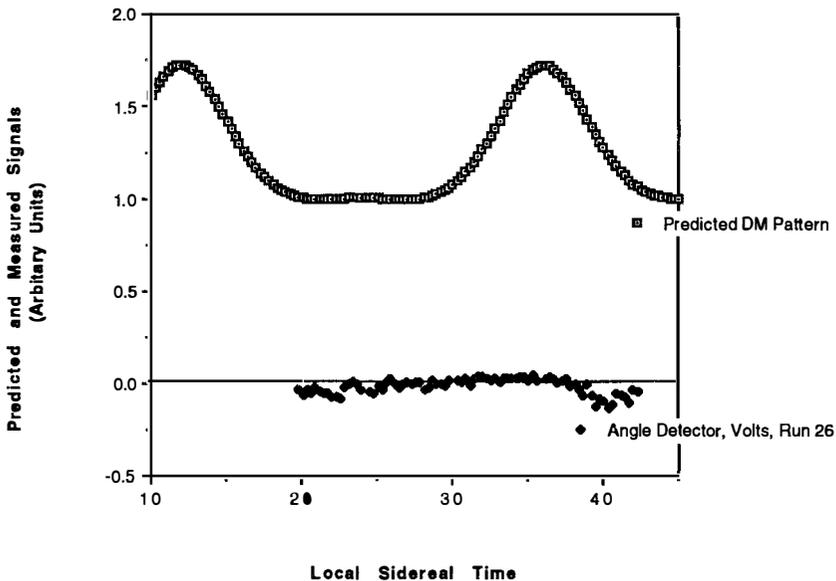


Figure 1. Predicted sidereal pattern for acceleration towards galactic center of the interaction form σ - r , along with raw torque pattern for one 24-hour run.

3. EXPERIMENTAL RESULTS

The experimental runs are fit to the sidereal pattern shown in Figure 1. A theoretically predicted axionic σ - r interaction¹⁰ cannot have a long enough range to act

in this situation, in view of astrophysical limits on the axion mass¹¹. A more likely candidate would be the arion of Anselm¹². (Many other models of the possible interaction between dark matter and the observable matter in the universe exist. See the preliminary bibliography by Gillies¹³ for citations of some of them.) Consequently we calculate the possible amplitude of the *unnormalized* pattern in the presence of experimental noise, at the 1- σ limit, to constitute the individual result of each run, and express it in terms of the acceleration on the spin masses. The present limit for the mean pendulum response in the direction of the galactic center at the 1- σ level is 5.3×10^{-4} degree. Taken with the fiber sensitivity and the spin mass locations 3.4 cm from the fiber, this yields an acceleration limit for the masses of 3.5×10^{-9} cm-s⁻² towards the center of our galaxy. The acceleration of gravity to the galactic center is 1.85×10^{-8} cm-s⁻², so the present $\sigma \cdot r$ limit is about 0.2 of gravity.

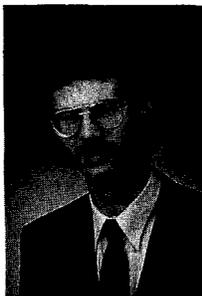
REFERENCES

1. Phillips, P.R., Phys. Rev. Lett. **59**, 1784 (1987).
2. Ni, W.-T. *et al*, Mod. Phys. Lett A, in press.
3. Stubbs, C.W., Phys. Rev. Lett. **70**, 119 (1993).
4. Smith, G., *et al*., Phys. Rev. Lett. **70**, 123 (1993).
5. Binney, J. and Tremaine, S., *Galactic Dynamics* (Princeton Univ. Press, Princeton, 1987).
6. Caldwell, D.O., in *Perspectives in Neutrinos, Atomic Physics and Gravitation*, Ed. J. Tran Thanh Van, T. Damour, E. Hinds and J. Wilkerson, (Editions Frontieres, Gif-sur-Yvette Cedex - France, 1993), p. 187.
7. Ritter, R.C., *et al*, Phys. Rev. D **42**, 977 (1990).
8. Ritter, R.C., *et al*, Phys. Rev. Lett. **70**, 701 (1993).
9. Gillies, G.T. and Ritter, R.C., Rev. Sci. Instrum. **64**, 283 (1993).
10. Moody, J.E. and Wilczek, F., Phys. Rev. D **30**, 130 (1984).
11. Turner, M.S., Phys. Rep. **197**, 67 (1990).
12. Anselm, A.A., Pis'ma Zh. Eksp. Teor. Fiz. **36**, 46 (1982).
13. Gillies, G.T. "A Preliminary Bibliographic Survey of Recent Studies of Dark Matter in the Universe," U.Va./SEAS Technical Report No. UVA/124620/MANE94/107 (January 1994), 55 pp.

THE GRAVITATIONAL FORCE AT MASS SEPARATIONS FROM
0.6 m TO 2.1 m AND THE PRECISE MEASUREMENT OF G

H. Walesch, H. Meyer, H. Piel, J. Schurr

Fachbereich Physik, Bergische Universität Gesamthochschule Wuppertal
Gaußstraße 20, D-42119 Wuppertal, Germany



Abstract

We present the results of a gravitational experiment which is based on a microwave resonator. The gravitational force of a test mass acting on the resonator is measured as a function of distance. No deviation from Newton's law has been found and the gravitational constant G has been determined with a relative accuracy of 2.2×10^{-4} to be $G = (6.6724 \pm 0.0015) \times 10^{-11} \text{ Nm}^2 \text{ kg}^{-2}$ in good agreement with the presently accepted value.

with ω_0 the eigenfrequency of the pendula. The measured shift Δf of the resonance frequency can be converted into a shift Δb of the mirror separation using the conversion factor df/db based on standard theory of resonator performance [7].

The horizontal gravitational acceleration of each pendulum is calculated by a numerical integration of Newton's inverse square law over the mass distributions of the test mass (M) and the resonator (m_1, m_2) using a Gaussian integration formula. This leads to our basic relation:

$$\Delta f(r) = \frac{df}{db} \omega_0^{-2} GM \left[\left(\frac{1}{r^2} - \frac{1}{(r+b)^2} \right) K(r) - \left(\frac{1}{r_{\text{ref}}^2} - \frac{1}{(r_{\text{ref}}+b)^2} \right) K(r_{\text{ref}}) \right] \quad (2)$$

The terms in parentheses correspond to the difference of the gravitational force between the test mass and each pendulum. The function K takes the finite dimensions of the masses into account ($K=1$ corresponds to point masses). $\Delta f(r)$ is the measured frequency shift of the resonator obtained by moving the test mass periodically between a position r and a reference position r_{ref} . Different separations r between the test mass and the resonator from 0.6m to 2.1m have been chosen to measure the gravitational force as a function of distance. This measurement is used to test the inverse-square law and it is a powerful tool to detect otherwise hidden systematic errors in order to increase the precision of G .

Experimental Set-Up

A schematic sketch of the experimental set-up is shown in Fig.2. Its main part, the Fabry-Perot resonator, consist of two spherical mirrors. They are fabricated from OFHC copper and the roughness of their diamond machined surfaces is about 50nm. The diameter of the circular mirrors is 192mm, their radius of curvature is 580mm and they are separated by a distance b of 240mm. In this case stable electro-magnetic modes exist in the Fabry-Perot resonator [9].

The resonator is suspended in loops of tungsten wire mounted to a special suspension platform. The mirror separation is held constant by means of a quartz plate in order to reduce thermal drift of the resonance frequency of the Fabry-Perot resonator.

To damp the pendulum oscillations caused by microseismic noise we use eddy-current brakes which consists of iron plates supporting an array of permanent magnets with a chess-board pattern. The brakes can be adjusted

in situ in order to synchronize the oscillation of the two pendula and therefore to minimize the oscillation of their separation.

The resonator with the eddy-current brakes and the suspension platform are placed inside a vacuum tank. To avoid dielectric effects and convection in the residual gas which disturb the resonance frequency of the Fabry-Perot resonator the vacuum pressure is kept below 10^{-4} mbar.

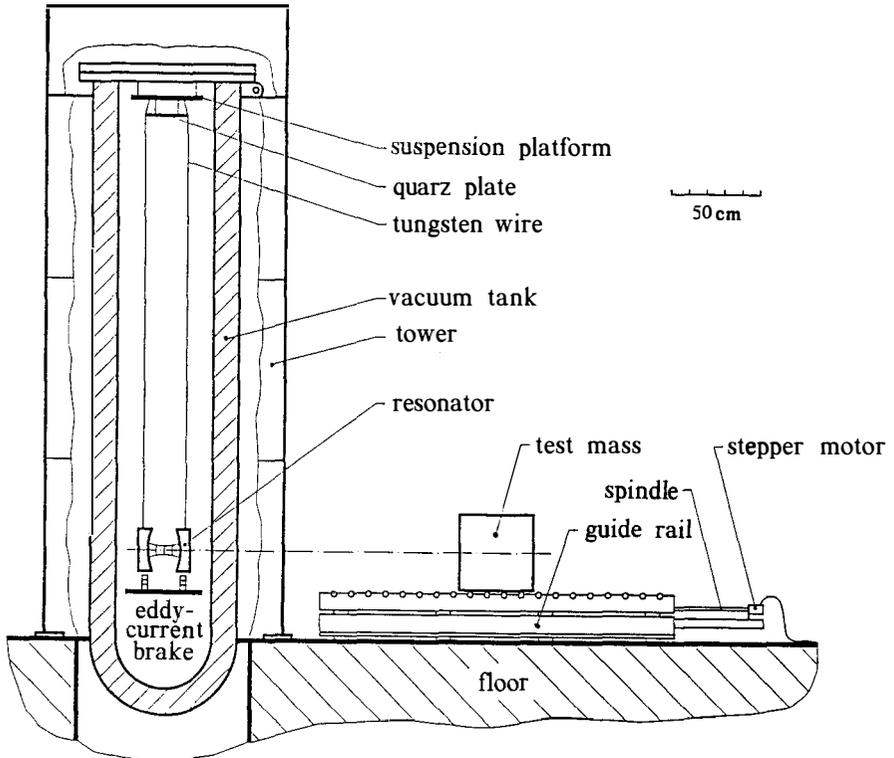


Fig. 2: The experimental set-up (schematic)

To keep thermal expansion effects small a good insulation of the vacuum tank is necessary. In our case we use a vacuum tank with an additional vacuum insulation and with additional layers of superinsulation inside.

The tank is mounted into a supporting steel trestle (called tower) to align the resonator on the same height as the test mass. The tower is build from

strong steel girders to obtain such a rigidity that the lowest eigenfrequencies of the tower are much higher than the eigenfrequencies of the pendula.

The test mass outside the trestle rests on a special guide rail and glides on rollers which are mounted on ball bearings around axels fixed to this rail. The test mass is a cylinder with a diameter of 440mm and a length of 430mm. Its dimensions are chosen in a way that the gravitational force between test mass and resonator nearly behaves like point masses, e.g. the correction funktion $K(r)$ which has been introduced in equation (2) takes values between 0.98 and 1.001. In order to avoid magnetic forces between the test mass and the resonator a special brass was chosen for the material of the test mass. The material is an alloy of 90% copper and 10% zinc with a magnetic susceptibility smaller than 10^{-5} . Thus the magnetic forces between the test mass and the resonator are well below the detection level.

The test mass is positioned precisely ($\sim 1\mu\text{m}$) by means of a spindle driven by a stepper motor. The stepper motor itself is controlled by a computer and allows an exact periodic motion of the test mass.

Experimental Results

The gravitational force between the test mass and the resonator is measured by moving the test mass periodically between a position r and a reference position r_{ref} .

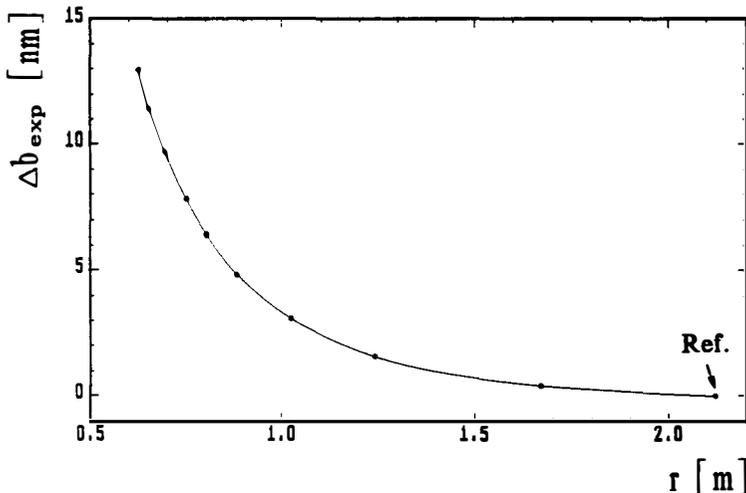


Fig. 3: The measured shift of mirror separation Δb due to the gravitational force versus the distance r .

This leads to a square wave modulation of the resonator frequency. The modulation amplitude Δf is determined from the data by means of a demodulation technique with high precision. Therefore disturbing thermal drift of the mirror separation and random noise is strongly suppressed and currently leads to a resolution of the change in mirror separation of about 1×10^{-12} m.

This procedure has been repeated for different positions of the test mass but always with the same reference position. The number of positions is usually chosen to be 9 where the number of cycles per position is about 12. In Fig.3 the shift Δb of the mirror separation is plotted versus the distance r between the resonator and the test mass.

The residuals of Δb are shown in Fig.4 in more detail. The data points are normally distributed and no significant deviations from the inverse-square law are observed.

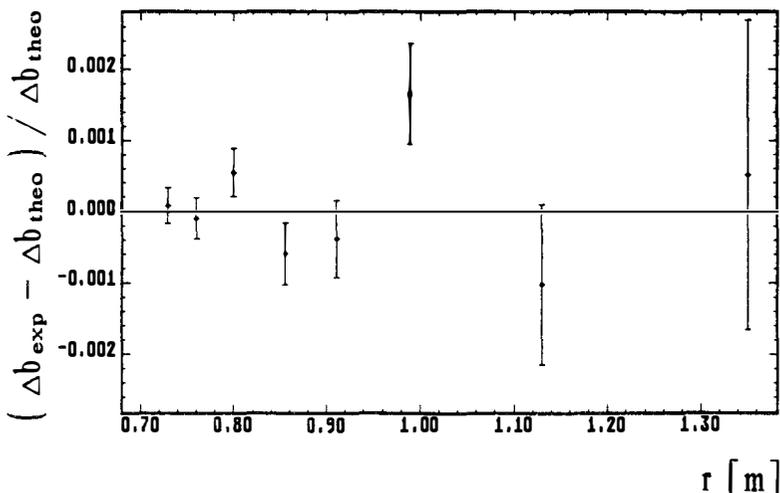


Fig. 4: The measured deviations from the inverse-square law.

The full line in Fig.3 is a least-square fit to the data. From this fit we determined the gravitational constant with a relative accuracy of 2.2×10^{-4} to be:

$$G = (6.6724 \pm 0.0015) \times 10^{-11} \text{ Nm}^2 \text{ kg}^{-2}$$

The quoted error includes both statistical (100ppm) and systematical (200ppm) errors. The measured value of G is consistent with the CODATA value [2].

The precision achieved is not inherent to the chosen experimental method. Limits of the precision by which G can be determined caused by the finite Q ($\sim 2.4 \times 10^5$) of our normal conducting Fabry-Perot resonator are on a level of 5×10^{-5} . This limit can in principle be overcome by using a superconducting resonator instead of a normal conducting one. In an earlier experiment [9] it was demonstrated, that at the temperature of liquid helium (4.2K) a Q value of 1.8×10^7 can be obtained using niobium mirrors instead of copper mirrors.

The precision achieved so far is limited by systematic uncertainties and we try to get further improvements in our future work.

This work was funded in main part by the Deutsche Forschungsgemeinschaft under grant number Pi 118/4-2.

References

- [1] G.G. Luther, W.R. Towler, Phys. Rev. Lett. **48**, 121-123 (1981).
- [2] E.R. Cohen, B.N. Taylor, Reviews of Modern Physics **59**, 1121-1148 (1987).
- [3] G.T. Gillies, Metrologia **24**, 1-56 (1987)
- [4] E. Fishbach et al., Metrologia **29**, 213-260 (1992).
- [5] J. Schurr, N. Klein, H. Meyer, H. Piel, H. Walesch, Metrologia **28**, 397-404 (1991)
- [6] J. Schurr, H. Meyer, H. Piel, H. Walesch, Lecture Notes in Physics **410**, Proc., Bad Honnef, Germany 1991, Springer-Verlag (1992).
- [7] K.M. Luk, P.K. Yu, IEE Proc. **132**, 105-113 (1985)
- [8] H. Kogelnik, T. Li, Appl. Opt. **5**, 1550-1566 (1966)
- [9] N. Klein, thesis, Universität Wuppertal, Germany, WUB-DIS 89-3 (1989)



DETERMINATION OF THE GRAVITATIONAL CONSTANT WITH A LAKE EXPERIMENT

B. Hubler, A. Cornaz*, and W. Kündig
Physik-Institut, Universität Zürich, 8057 Zürich, Switzerland



Abstract

The Newtonian gravitational constant has been determined at an effective interaction distance of 112 m. A high-precision balance was used to compare the weights of two 1-kg stainless steel masses located above and below the variable water level of a pumped-storage lake. Water-level changes up to 43 m produced a maximum weight difference of 1290 μg , which could be measured with a resolution of $< 1 \mu\text{g}$. The data yield a value for the gravitational constant G of $6.6700(54) \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ (1σ) in agreement with laboratory determinations. New limits for the strength of possible new intermediate-range forces are placed.

1. Introduction

Though Newton's inverse-square law of gravity is well tested at astronomical distances, deviations from Newtonian gravity can not be excluded for distances less than 1000 km. Possible deviations may be described with an intermediate-range "fifth force" [1].

In a lake experiment [2] the gravitational constant G was measured at an effective distance of 112 m. Together with the laboratory value [3] new limits can be placed on the strength of a composition-independent fifth force in the range 5 cm to 100 m.

2. The Gigerwald experiment

The experimental site is at the Gigerwald lake, a pumped-storage reservoir at an altitude of 1335 m above sea level. The lake is confined by a 147-m-high concrete dam of parabolic shape allowing maximum water-level changes of 90 m. In the central plumb shaft of the dam a balance (a modified Mettler-Toledo mass comparator) measures the weight difference of two test masses of stainless steel located above and below the variable water level (see Fig. 1). The upper mass (of 1.11445 kg) is hanging just below the balance, whereas the lower mass (of 1.09987 kg) is suspended by a 100- μm -diam tungsten wire at the bottom of the shaft. Their center of mass separation is determined to be 103.822(2) m. To avoid air convection and variable buoyancy the balance and the test masses are held in vacuum (5×10^{-3} mbar).

The weight difference is determined by alternately weighing each test mass over a time of 3 min using sophisticated exchange and suspension devices. After a mass, say the upper mass, is weighed, it is detached from the balance by slightly lifting it (≈ 1 mm), simultaneously the other mass is lowered and suspended to the balance. During this exchange the load is held constant within 1 g to avoid relaxation effects in the balance. The weight difference is then calculated by linear interpolation between two successive measurements of the upper mass at the time the lower mass was measured. After 60 mass exchanges a calibration of the scale sensitivity is carried out by adding a small mass of 0.999993 g to the balance. In fact, no significant change of scale sensitivity was observed. With a three-stage temperature control system the temperature of the balance is held constant within 1 mK over several weeks.

Since the weight difference is measured in a short time (≈ 12 min) balance drifts cancel. Time-variable gravity effects originating from distances much larger than the separation of test masses completely vanish (e.g. tides). By comparing the weight differences at several water

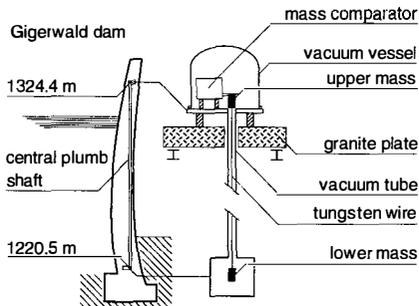


Figure 1. Scheme of the Gigerwald experiment.

levels even the static local gravity from the surroundings cancels. Finally, the recorded gravity signal is just due to the interaction between the locally moved mass (water and air) and the test masses.

3. Lake survey

Since the gravity effect of a infinite horizontal sheet of thickness h and density ρ is proportional to ρh and approximately remains true in our case of a finite lake, it is advantageous to determine the water level via the water pressure instead via floats or similar methods. We used a high-precision pressure transducer (Rittmeyer W1Q) with a resolution of 1 mm and an absolute accuracy of 1 cm (see Fig. 2). The instrument automatically corrects for the ambient air pressure.

The shore and dam contours were determined by air photogrammetry and conventional surveying, respectively. Data processing yielded 2-m-equidistant contour lines. The dam shape is known to better than 1 cm. The coordinates of the shoreline have a random uncertainty of 30 cm. Their systematic uncertainty is much less than 5 cm as a result of a well surveyed control network. The coordinates of test masses are uncertain by less than 3 mm. Not the whole shore is bounded by rocks, there are also layers of scree, where water is seeping in. From geological surveys based on numerous drill holes the solid rock boundary is known. The porosity of the scree is estimated by geologists to be 0.30 ± 0.03 , what enabled us to calculate the effect of water seepage.

Water density was measured in the laboratory and was found to be slightly denser than pure water by $1.2(2) \times 10^{-4} \text{ g cm}^{-3}$. Temperature profile measurements in the lake revealed no significant change of temperature with depth except for a thin boundary layer below the surface. During the measuring period water temperature varied from 3°C to 6°C. In this range water-density variations are negligible. The density which enters the calculation of the gravity effect is the density difference between water and air. The mean density of air is determined to be $0.00108 \text{ g cm}^{-3}$.

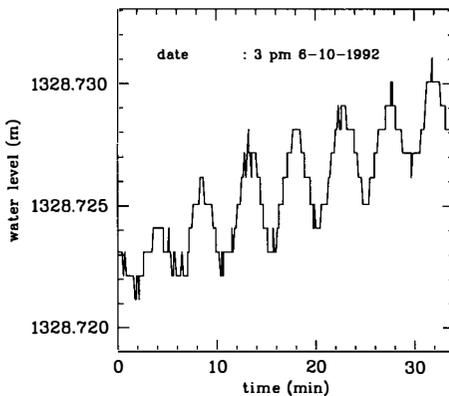


Figure 2. Water level measurements. The observed fundamental oscillations (“seiches”) in the Gigerwald lake have a period of 4.7 min. The steps are due to the digital resolution of the pressure transducer.

4. Data analysis and results

To calculate the Newtonian gravity signal the method of Talwani and Ewing [4] was used. They give an exact formula for the vertical gravity of a flat horizontal lamina whose contour is represented by an arbitrary n -sided polygon. The effect of Earth curvature is negligible. The effective interaction distance r_{eff} is:

$$r_{\text{eff}} = (\int r_1 dF_1^z - \int r_2 dF_2^z) / (\int dF_1^z - \int dF_2^z),$$

where the r_i 's and the dF_i^z 's are the distances, respectively the vertical gravitational forces between a water element and the two test masses. The integration over the Gigerwald lake yielded $r_{\text{eff}} = (112 \pm 2)$ m depending on the water level.

Figure 3 shows the digitally recorded raw data of water level and weight difference. In 1993 water-level changes up to 43 m produced a maximum weight signal of 1290 μg , which could be measured with a short-term resolution of 0.5 μg . As long as water-level changes are well between the two test masses, the weight difference varies almost linearly by 33 $\mu\text{g m}^{-1}$. Whenever water level approaches the height of the upper mass at 1324.359 m, the weight signal diminishes and becomes even negative for higher levels. The measured and calculated weight difference as a function of water level is shown in Fig. 4.

The weight differences at same heights must be equal for all times. Any discrepancies are interpreted as change of mass distribution in the environment, e.g. change of soil moisture. Variable atmospheric pressure have an effect of $< 0.1 \mu\text{g}$ on the weight signal. Observed lake oscillations ("seiches") with amplitudes of 2 mm also have no effect. Dam movements originating from temperature and water pressure changes do not produce a significant uncertainty, since both lake water and test masses are moved the same way. Effects of balance tilts due to dam movements cancel by measuring the weight difference and by repeated scale calibration. Ground vibrations were not observed.

In order to minimize the effect of soil moisture in the weight signal, the following model equation for the measured weight difference is taken:

$$\Delta F(t) = a\Delta F_{\text{Newton}}(t) + b_i t + c_i$$

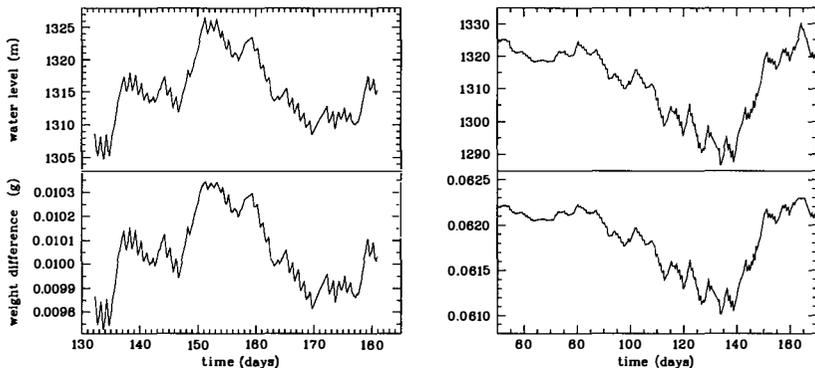


Figure 3. Raw data of water level and weight difference of the 1992 (left) and the 1993 measurement (right). The peaks are caused by daily, weekly and seasonal water-level changes depending on power consumption (the local gravity at the balance is $9.804\,208 \text{ m s}^{-2}$).

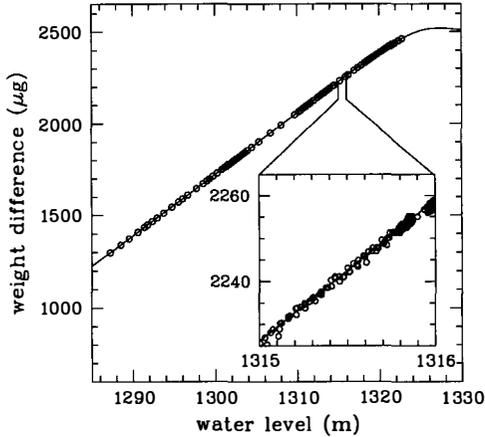


Figure 4. The solid curve is the calculated weight difference of the two test masses as a function of the water level following pure Newtonian gravity (the origin is set at 1240 m for an empty lake). The measured gravity data of 1993 (open circles), corrected only with a linear drift, fit well the predicted gravity effect. Each circle represents the average of 100 measurements. The insert shows a typical region where all gravity data are presented.

The parameter a represents the ratio G/G_{lab} over the whole measuring period, where G_{lab} denotes the currently accepted laboratory value of the gravitational constant [3]. b_i and c_i are the coefficients of linear drifts in time periods of about 5 days. The drifts are $< 0.5 \mu\text{g}$ per day and vary in sign. The value for a determined from data is 0.99961(18), which is the weighted mean of the 1992 and 1993 measurement (see Table 1). For a consistency check the 1993 data were subdivided into two data sets of measurements with water level above and below 1305 m. The results are in reasonable agreement with each other. The final uncertainty in this determination of G is obtained by taking the root-mean-square of the uncertainties listed in Table 2.

The largest component of the error budget is the uncertainty of water seepage in the scree. It turns out to be the limiting element of this experiment. The resulting value of the gravitational constant is

$$G = (6.6700 \pm 0.0054) \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2},$$

which is in agreement with laboratory values.

Table 1. Results.

Data		$a = G/G_{\text{lab}}$
1992	overall	1.00017(76)
1993	> 1305 m	0.99929(26)
1993	< 1305 m	0.99981(33)
1993	overall	0.99958(19)
1992+1993		0.99961(18)

Table 2. Error budget.

Source	Uncertainty (parts per 1000)
Dam contour	0.30
Shore contour	0.37
Porosity of scree	0.61
Test mass positions	0.04
Water level	0.12
Water density	0.05
Random experimental error	0.18
Total (1σ)	0.81

Together with the laboratory value G_{lab} the Gigerwald experiment sets useful limits on the strength of a fifth force described by the energy potential

$$V(r) = -G_{\infty} \frac{m_1 m_2}{r} (1 + \alpha e^{-r/\lambda}),$$

where α is the strength of the new force relative to gravity and λ the Compton wavelength of the exchange particle. Hence it follows a force

$$F(r) = G(r) \frac{m_1 m_2}{r^2}, \quad G(r) = G_{\infty} [1 + \alpha(1 + r/\lambda)e^{-r/\lambda}],$$

and the gravitational “constant” becomes distance-dependent.

A measured value of the ratio $G(r_1)/G(r_2) = \beta$ at the distances r_1 and r_2 constrains the strength $\alpha(\lambda)$ by varying β within the limits of experimental errors [8]. Fig. 5 illustrates the

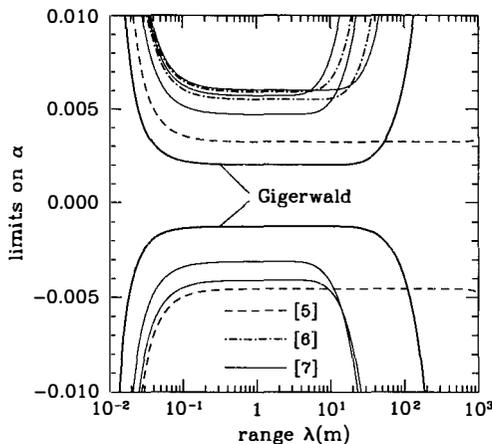


Figure 5. Excluded strengths α and ranges λ for a single Yukawa model at the 2σ level arising from experiments measuring directly the gravitational constant at geophysical distances (the allowed region lies between the curves). Constraints from other experiments are not shown. Numbers in brackets refer to the references in the text.

constraints placed by G experiments at geophysical distances on such an interaction with 95% confidence.

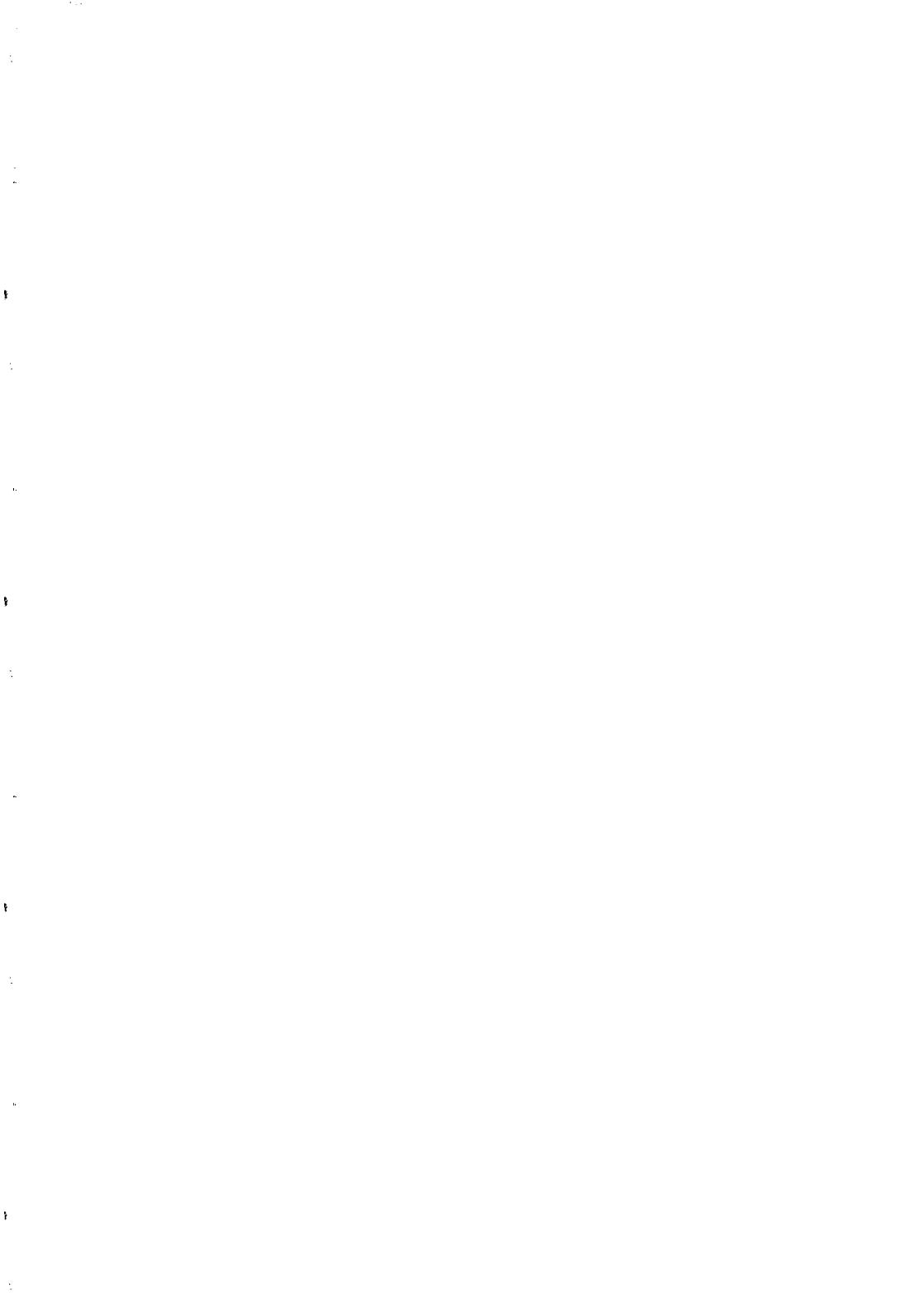
Acknowledgments

We thank W. Grünenfelder of the Kraftwerke Sarganserland KSL and Chr. Venzin of the Nordostschweizerische Kraftwerke NOK for the hospitality in the Gigerwald dam and helpful technical support. We like to thank M. Baumeler of Mettler-Toledo AG in Greifensee, Switzerland, for providing the mass comparator and for the crucial collaboration to enhance its performance. We are indebted to Auer+Clement AG for air photogrammetry and to the Ingenieur-Büro Schneider for survey data. We gratefully acknowledge important contributions from A. Geiger, T. Kersten and E. Klingele of the Institut für Geodäsie und Photogrammetrie der ETH Zürich and from H. Stüssi of the Physik-Institut der Universität Zürich. This work was supported by the Schweizerische Nationalfonds and the Dr. Tomalla-Stiftung.

* Lecture was given by A. Cornaz.

References

- [1] For a review, see E. G. Adelberger *et al.*, *Annu. Rev. Nucl. Part. Sci.* **41**, 269 (1991), or E. Fischbach and C. Talmadge, *Nature* **356**, 207 (1992).
- [2] A. Cornaz, B. Hubler, and W. Kündig, *Phys. Rev. Lett.* **72**, 1152 (1994).
- [3] G. G. Luther and W. R. Towler, *Phys. Rev. Lett.* **48**, 121 (1982).
- [4] M. Talwani and M. Ewing, *Geophysics* **25**, 203 (1960).
- [5] M. A. Zumberge *et al.*, *Phys. Rev. Lett.* **67**, 3051 (1991).
- [6] G. Müller *et al.*, *Phys. Rev. Lett.* **63**, 2621 (1989); G. Müller *et al.*, *Geophys. J. Int.* **101**, 329 (1990).
- [7] M. Oldham *et al.*, *Geophys. J. Int.* **113**, 83 (1993).
- [8] G. W. Gibbons and B. F. Whiting, *Nature* **291**, 636 (1981).



APPLICATIONS OF SUPERCONDUCTIVITY TO GRAVITATIONAL EXPERIMENTS IN SPACE

Saps Buchman

W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, U.S.A.



ABSTRACT

We discuss the superconductivity aspects of two space-based experiments — Gravity Probe B (GP-B) and the Satellite Test of the Equivalence Principle (STEP). GP-B is an experimental test of General Relativity using gyroscopes, while STEP uses differential accelerometers to test the Equivalence Principle. The readout of the GP-B gyroscopes is based on the London moment effect and uses state-of-the-art dc SQUIDs. We present experimental proof that the London moment is the quantum mechanical ground state of a superconducting system and we analyze the stability of the flux trapped in the gyroscopes and the flux flushing techniques being developed. We discuss the use of superconducting shielding techniques to achieve dc magnetic fields smaller than 10^{-7} G and an ac magnetic shielding factor of 10^{13} . SQUIDs with performance requirements similar to those for GP-B are also used for the STEP readout. Shielding for the STEP science instrument is provided by superconducting shells, while the STEP test masses are suspended in superconducting magnetic bearings.

1. INTRODUCTION

Low temperature techniques have been widely used in high precision experiments due to their advantages in reducing thermal and mechanical disturbances, and the opportunity to utilize the low temperature phenomena of superconductivity and superfluidity. Gravitational experiments have reached the stage at which both low temperature technology and ultra low gravity space environments are needed to meet the requirements for measurement precision. In this paper we present some of the applications of superconductivity to two space based gravitational experiments — Gravity Probe B¹⁾ (GP-B) and the Satellite Test of the Equivalence Principle²⁾ (STEP).

General Relativity predicts the geodetic and frame dragging effects to be 6.6 arcsec/yr and 0.042 arcsec/yr in a 650 km polar orbit³⁾. The GP-B experiment will measure these effects with an accuracy of at least 0.3 marcsec/yr, by determining the precession of the local frame, given by gyroscopes, with respect to the fixed frame of the stars, given by a telescope pointed to HR5110. Figure 1 shows the GP-B experimental concept.

GP-B uses superconductivity for the implementation of the London moment gyroscope readout⁴⁾. The London moment is aligned with the angular momentum of the gyroscopes, and is measured using ultra low noise dc SQUIDS. Two techniques of flux reduction in superconducting films insure a minimal level of trapped vortexes. Superconducting shielding is used to achieve both low static magnetic fields and high attenuation of variable magnetic fields. The following sections discuss the London moment readout, trapped flux reduction, and superconducting magnetic shielding.

STEP is designed to improve the measurement of the Equivalence Principle from the present accuracy of 10^{-12} to about 10^{-17} . This cryogenic experiment uses differential accelerometers (with test masses of different materials) placed in a drag free satellite. Figure 2 shows the STEP experimental concept. The position of the test masses is sensed with SQUID magnetometers with 10^{-2} Å precision, while magnetic shielding is provided by superconducting enclosures; both adaptations of GP-B developed techniques. Radial suspension for the test masses is provided by superconducting magnetic bearings.

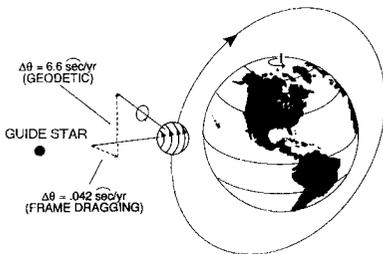


Figure 1. GP-B experimental concept

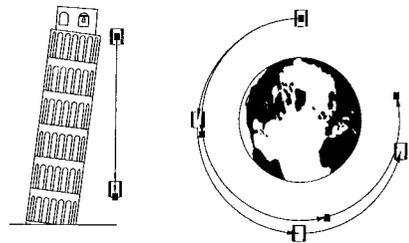


Figure 2. STEP experimental concept

III. TRAPPED FLUX REDUCTION

The Meissner effect predicts that a perfect superconductor will expel all magnetic flux when placed in a field small compared to the critical field. However, the $2.5\ \mu\text{m}$ niobium film covering the 3.8 cm diameter quartz gyroscopes has enough defects to actually trap a density of flux lines roughly equivalent to low ambient fields in the range 10^{-3}G - 10^{-7}G . The London moment read-out can be complicated by the polhoding induced motion of this trapped flux, therefore making it desirable to minimize the density of fluxons.

Two methods for trapped flux reduction are used at GP-B. The first approach is based on the controlled thermal cycling of the superconductor, thus minimizing the thermally induced currents which generate the trapped flux. Using this method we achieve a dipole moment trapped in the sphere corresponding to a uniform field of about $0.2\ \mu\text{G}$, or a total of about 50 trapped fluxons⁶⁾. The second method makes use of the fact that in a spherical shell the number of right-handed and left-handed flux vortices is equal. A normal spot, produced by an infrared laser beam, is used to trap and sweep the flux lines until they close on themselves⁷⁾. The motion of the spot is produced by the slow rotation and precession of the gyroscope. Adjustable experimental parameters are the infrared power

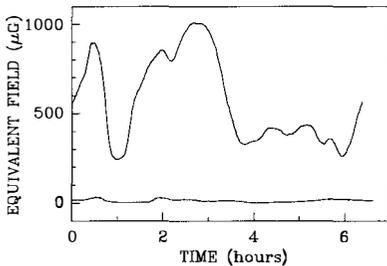


Figure 4. Evolution of trapped magnetic dipole moment before (upper trace) and after flux flushing (lower trace). Dipole magnitude expressed in units of equivalent uniform field in the sphere.

absorbed in the superconducting film ($\sim 0.3\ \text{mW}$), the pressure of the helium exchange gas ($6 \times 10^{-3}\ \text{Pa}$), and the spin speed of the gyroscope ($\sim 0.5\ \text{Hz}$), resulting in a normal spot of about $50\ \mu\text{m}$ radius. Using this technique the dipole moment corresponding to a field of $1\ \text{mG}$ has been reduced to about 6% of its original level. Figure 4 shows the dipole moment component normal to the pickup loop, as a function of time, before and after the laser annealing. Further experiments, aimed at extending this technique to use in ultra low magnetic fields, are currently underway in an ambient field of $10^{-7}\ \text{G}$.

IV. MAGNETIC SHIELDING

The London moment based readout requires shielding from external ac magnetic fields by thirteen orders of magnitude, and a dc field of less than 10^{-7}G needed to insure low trapped flux in the gyroscopes. This level of shielding is achieved by using a system of superconducting magnetic shields, in conjunction with controlling the magnetic

properties of the probe materials. A ferromagnetic shield provides a 10mG magnetic field region within which the main field reduction is achieved by the expansions of superconducting lead foil cylinders. The foil cylinder, folded to minimize the cross-sectional area, is cooled through its superconducting transition temperature. Cooling rates are controlled in order to minimize thermal gradient induced currents. Once the lead cylinder is superconducting, it is mechanically unfolded to its maximum diameter. This operation reduces the magnetic flux in the cylinder by roughly the ratio of the folded

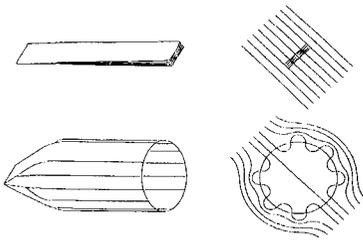


Figure 5. Lead cylinder expansion.

to the expanded cross-sectional areas, a factor of about 20 in practice. Figure 5 represents schematically the expansion process. This process is repeated with additional lead cylinders being expanded one at a time inside the reduced field produced by the previous expansions. The lowest magnetic field values achieved are about 5×10^{-8} G; with further reduction prevented by thermoelectric effects in the lead foil⁸⁾.

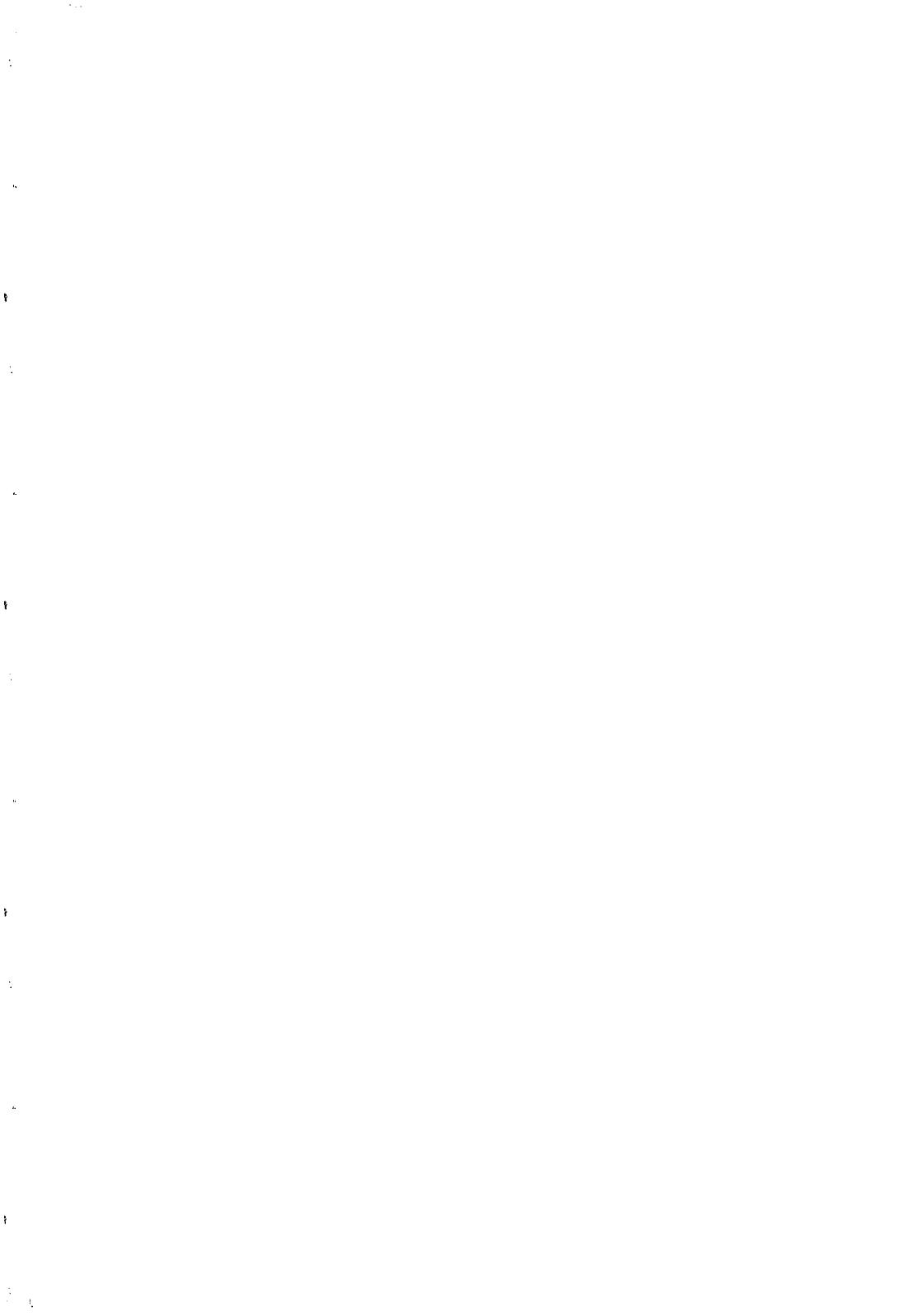
The ferromagnetic shield and lead foil superconducting cylinder provide an ac shielding factor of about 10^9 , while a superconducting cylinder surrounding the gyroscope increases the shielding factor to 10^{12} . Finally, the superconducting gyroscope shields the readout loop from external fields by an additional factor of 10. The performance of the ac shielding scheme has been verified to the 2×10^{11} level, the limitation being the noise performance of the rf SQUID detectors used⁸⁾. Measurements with improved resolution are currently being prepared.

ACKNOWLEDGMENTS

This work was supported by NASA contract No. NAS8-36125. The author wishes to thank the members of the Stanford University GP-B group for many stimulating discussions.

REFERENCES

- 1) J.P.Turneure *et al.*, *Adv. Space Res.* **9**, 29 (1989).
- 2) Satellite Test of the Equivalence Principle, *ESA Phase A Study Report #SCI (93)4*.
- 3) L.I.Schiff, *Proc. Natl. Acad. Sci.* **46**, 871 (1960).
- 4) F.London, *Superfluids* (Dover, New York, 1961), Vol. I.
- 5) H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1972).
- 6) Y. Xiao, S. Buchman, G.M. Keiser, *et al.*, *to appear in Physica B* (1994).
- 7) R. Brumley, S. Buchman, and Y. Xiao, *to appear in Physica B* (1994).
- 8) M. Taber, D. Murray, J.M. Lockhart, *et al.*, *Adv. Cryogenic Eng.* **39** (1993).



**PROSPECTS FOR MEASURING γ TO A PART IN 10^5
IN THE GRAVITY PROBE B EXPERIMENT**

C.W.F. Everitt, Saps Buchman

W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, U.S.A.



ABSTRACT

Gravity Probe B is a space based test of gravitational theory, making use of gyroscopes in a 650 km polar orbit to measure the precessions due to the geodetic and frame dragging effects. The present requirement for experimental accuracy is 0.3 marcsec/yr. We discuss improvements in the areas of gyroscope performance, gyroscope readout, and star tracking which will increase the measurement accuracy to about 0.07 marcsec for the 1.5 years experiment duration. This translates in a measurement of the parametrized post-Newtonian (PPN) parameter γ (mass produced curvature of three dimensional space) to one part in 10^5 , extending the search for a possible scalar in gravity by two orders of magnitude, and allowing a test of the critically damped version of the Damour-Nordtved "attractor mechanism".

I. INTRODUCTION

Gravity Probe B (GP-B)¹⁾, also known as the Relativity Gyroscope experiment, is designed to measure very precisely the frame dragging and geodetic effects predicted by Einstein's General Relativity theory. Leonard Schiff²⁾ has calculated the relativistic precession $\vec{\Omega}$ of a gyroscope in a circular Earth orbit to be:

$$\vec{\Omega} = \frac{3GM}{2c^2R^3} (\vec{R} \times \vec{v}) + \frac{GI}{c^2R^3} \left[\frac{3\vec{R}}{R^2} \cdot (\vec{\omega}_e \cdot \vec{R}) - \vec{\omega}_e \right] \quad (1)$$

where \vec{R} and \vec{v} are the location and the orbital velocity of the gyroscope, and I , M , and $\vec{\omega}_e$ are the moment of inertia, the mass, and the angular velocity of the Earth.

The first term represents the geodetic effect, which produces a precession with a rate of 6.6 arcsec/yr for a 650 km polar orbit, and is due to the motion of the gyroscope through the curved space-time around the Earth. This effect will be measured to about one part in 10^5 , providing the most precise test to date of any of the positive predictions of General Relativity. The second term represents the effect is due to the dragging of the inertial frame by the rotation of the Earth, and its rate is 0.042 arcsec/yr. GP-B will measure the frame dragging effect for the first time and with a precision of about three parts in 10^3 .

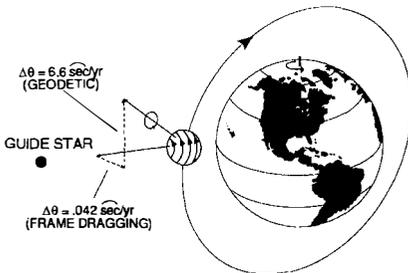


Figure 1. Geodetic and frame dragging precessions predicted by General Relativity for a 650 km polar orbit

The two precessions are measured by referring the local frame, determined by the gyroscopes, to the universal inertial frame, determined by a telescope pointed to a remote star, namely HR5110. Note that the proper motion of the star must be determined in a separate experiment. The experiment will be placed in a drag free satellite³⁾. As shown schematically in Figure 1, the polar orbit will result in orthogonal geodetic and frame dragging precessions, thus significantly simplifying the data analysis.

Recent work by Damour and Nordtvedt⁴⁾ shows that tensor-scalar theories of gravity can result in deviations from zero of the PPN parameter γ of up to $\gamma - 1 \leq 4 \times 10^{-5}$. These theories contain a mechanism which drives the world towards pure General Relativity, with the above deviation a result of the finite time elapsed since the end of the radiation-dominated era. This provides additional incentive to increase the measurement accuracy in order to add the test of this prediction to the GP-B tests⁵⁾. Expressed in terms of the PPN parameters, the relativistic precessions are:

$$\tilde{\Omega} = \left(\gamma + \frac{1}{2} \right) \frac{GM}{c^2 R^3} (\bar{R} \times \bar{v}) + \left(\gamma + 1 + \frac{\alpha_1}{4} \right) \frac{GI}{2c^2 R^3} \left[\frac{3\bar{R}}{R^2} \cdot (\bar{\omega}_e \cdot \bar{R}) - \bar{\omega}_e \right] \quad (2)$$

The following sections of this paper describe progress recently made at GP-B, in the areas of gyroscope performance and readout, and of guide star tracking, which resulted in the increase of the measurement accuracy from 0.3 marcsec/yr to about 0.07 marcsec for 1.5 years, thus offering the means of measuring γ to one part in 10^5 .

II. GYROSCOPE PERFORMANCE

Four high precision, cryogenic, electrostatically suspended gyroscopes determine the reference frame in the vicinity of Earth. Residual torques are reduced to a minimum by using a drag free satellite and by carefully controlling the sphericity of the gyroscope and its housing. The ratio of the surface non-uniformity, with respect to the 1.9 cm spherical radius, is about 10^{-6} for the gyroscope and 10^{-5} for its housing. The drag free technique uses the thrust of helium, boiloff from the experiment's dewar, to maintain the spacecraft centered around a drag free sensor which floats in vacuum near the center of mass of the satellite. This assures a residual acceleration level of about $10^{-9} g$ ($g = 9.81 \text{ m/s}^2$) at the drag free sensor, which is averaged by the satellite roll to better than $10^{-12} g$ transverse acceleration. Due to their location away from this sensor, the individual gyroscopes will experience a residual acceleration of about $10^{-8} g$ caused by the Earth's gravity gradient. The main residual torques on the gyroscopes are caused by the interaction of the electrostatic suspension system (needed to compensate for the gravity gradient acceleration) with the imperfections of the gyroscope surface.

Table 1. Gyroscope disturbance precessions

DISTURBANCE TYPE	GYROSCOPE SUPPORT	
	Supported (marcsec/yr)	Unsupported (marcsec/yr)
Mass Unbalance (25nm)	< 0.014	< 0.002
Electrostatic Suspension	< 0.140	< 0.010
Residual He Gas(10^{-11} torr)		
Differential Damping	< 0.006	< 0.006
Brownian Motion	< 0.001	< 0.001
Rotor Charge (10pC)	< 0.010	< 0.010
Gravity Gradient	< 0.001	< 0.001
Cosmic Radiation	< 0.001	< 0.001
Magnetic	< 0.001	< 0.001
Photon Gas	< 0.001	< 0.001
ROOT SUM SQUARE	< 0.140	< 0.016

One of the fourfold redundant gyroscopes can be used as the drag free sensor, thus removing the suspension, and therefore the main contributions to the torques. The drift

rate is reduced from 0.14 marcsec/yr, for the supported gyroscopes, to 0.016 marcsec/yr for the unsupported gyroscope. Table 1 summarizes the disturbance precessions for supported and unsupported gyroscopes, assuming 130 Hz spin speed, 10^{-12} g average transverse acceleration, and 10 arcsec misalignment of the roll and spin axes. The disturbance precession increases with time, and is 0.024 marcsec for 1.5 years for the unsupported gyroscope, and 0.12 marcsec for the combined three supported gyroscopes.

III. GYROSCOPE READOUT

The gyroscope readout is based on the detection of the London dipole moment, created by a spinning superconductor. Precession of the angular momentum is mirrored by the precession of the magnetic dipole, and induces variations of the flux threading a loop concentric to the gyroscope, which are then measured by a SQUID magnetometer. Rolling of the satellite around the line of sight to the guide star, (also the gyroscope spin axis), allows readout at roll rate, thus avoiding zero frequency measurements.

The noise in the readout system is dominated by $1/f$ noise, making it desirable to increase the roll rate. This however has a number of drawbacks, the main one being the centrifugal force on the gyroscopes caused by misalignment of their centers of mass to the roll axis. Improvements in the experimental probe have recently allowed the increase of the roll rate from 1.7 mHz to 5.5 mHz thus significantly reducing the expected readout noise. Coupled with reductions in the noise level of the dc SQUID system, this has reduced the expected error in the readout system to the level of about 0.11 marcsec/yr. Figure 2 shows the results of a Monte Carlo simulation (0.09 marcsec/yr) and of the calculated readout error for a one year experiment. Each gyroscope has a dual readout, consisting of two parallel SQUID systems, thus allowing for a reduction of the error to 0.78 marcsec/yr. Furthermore, the readout error decreases with the square root of time, resulting in an error of 0.064 marcsec for 1.5 years, (9.9 arcsec geodetic precession).

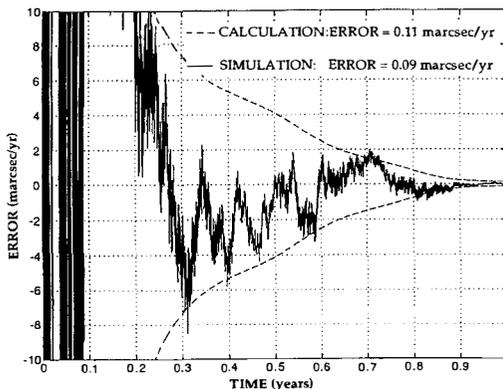


Figure 2. Monte Carlo simulation (solid line) and theoretical calculation (dashed lines) of readout error for one year experiment.

IV. GUIDE STAR REFERENCE

The third and last major component of the experiment is the optical telescope which gives the reference to the guide star. Originally the guide star was Rigel with a declination of $\text{DECL} = -8^{\circ}13'$ and luminosity $m_v = -0.1$. This choice was dictated by the sensitivity of the optical detection system: photo multipliers, choppers, fiber optics, and optical windows. A new detection system using cryogenic photo detectors and JFET preamplifiers now allows the telescope readout to utilize stars down to magnitude $m_v \sim 6$. Consequently, guide star candidates now include stars with radio emission, making it possible to use VLBI to provide accurate measurements of the proper motion of these stars. Two promising candidates are the close in binary systems (< 1.6 marcsec) HR1099: $\text{DECL} = 0^{\circ}36'$, $m_v = 5.7$, and HR5110 $\text{DECL} = 37^{\circ}10'$, $m_v = 5.0$, with HR5110 being the present choice. The present accuracy of the measurement of the proper motion of HR5110 is 0.1 marcsec, and is expected to be improved to 0.025 marcsec in the next 3 years.

V. CONCLUSIONS

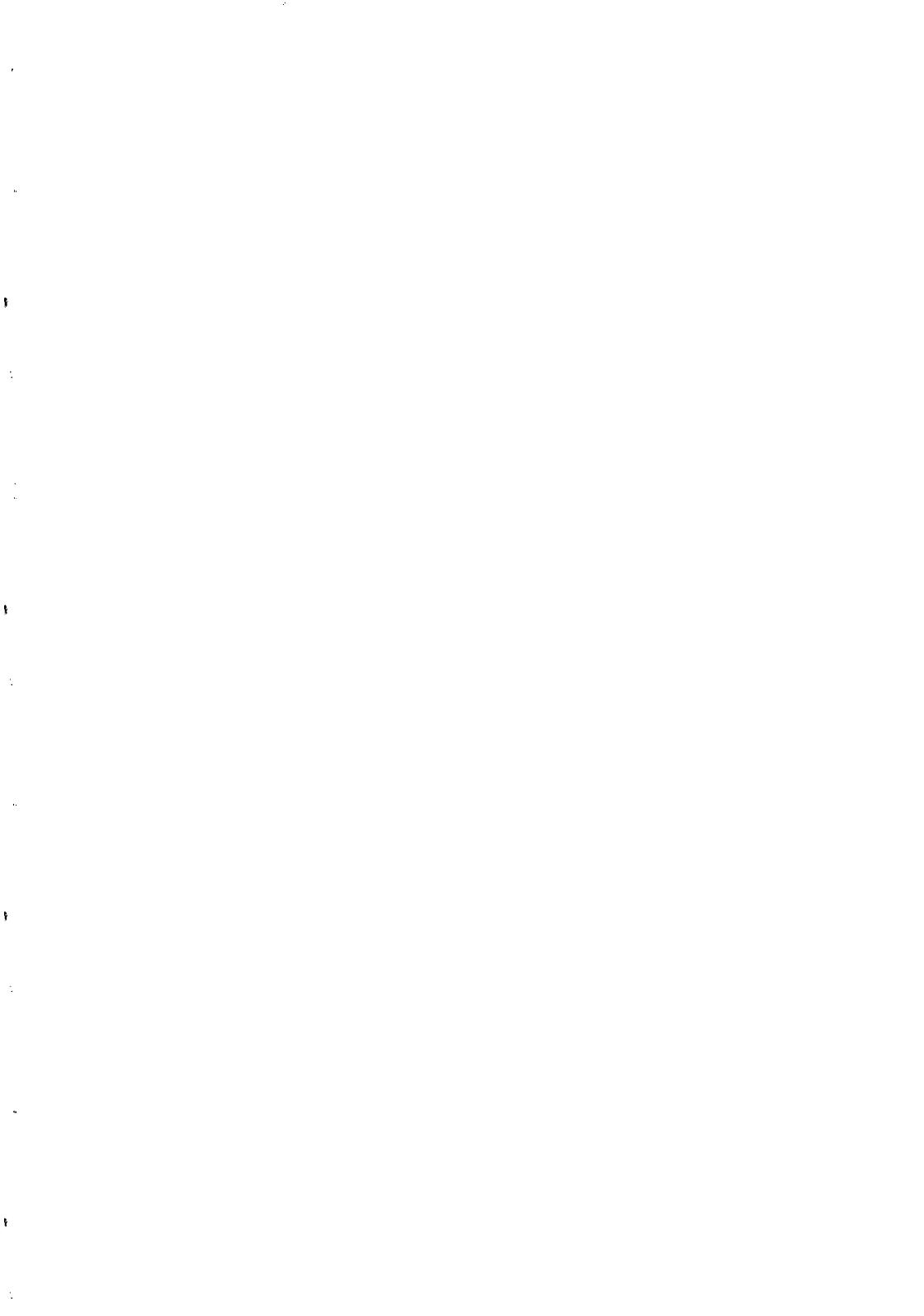
The errors in the three major areas of the GP-B experiment have been reduced to 0.064 marcsec, 0.025 marcsec, and 0.024 marcsec for the readout, star tracking, and gyroscope respectively, resulting in a total expected error of 0.073 marcsec for 1.5 years, and a measurement of γ to 1.1 parts in 10^5 . Future possible improvements are: 1) lower gyroscope charge and position measurement voltages, and higher vacuum resulting in an unsupported gyroscope error of 0.012 marcsec, 2) reduction in the SQUID noise, and increased roll and spin rates, reducing the readout error to 0.040 marcsec, 3) determination of the proper motion of the guide star to the 0.01 marcsec level, and 4) modeling of the disturbance torques of the supported gyroscopes to the 10% level, reducing the error of their ensemble to 0.012 marcsec. If realized, the first three improvements will reduce the total error of 0.045 marcsec, and allow a measurement of γ to seven parts in 10^6 , while torque modeling will further reduce the error to 0.032 marcsec, and result in a measurement of γ to five parts in 10^6 .

ACKNOWLEDGMENTS

This work was supported by NASA contract No. NAS8-36125. The authors wish to thank John Turneure, Mac Keiser, and Jeremy Kasdin, of the Stanford University GP-B group, for their help with the preparation of some of the data in this paper.

REFERENCES

- 1) J.P. Turneure *et al.*, Adv. Space Res. **9**, 29 (1989).
- 2) L.I. Schiff, Proc. Natl. Acad. Sci. **46**, 871 (1960).
- 3) D.B. DeBra, *DISCOS Description* (private communication, Stanford University, 1970).
- 4) T. Damour, K. Nordtvedt, Phys. Rev. Lett. **70**, 2217 (1993).
- 5) C.W.F. Everitt. Proc. VI Marcell Grossmann Meeting, Kyoto, Japan (1991).



Quantum Limits for Measurements on Macroscopic Bodies : a Decoherence Analysis

Carlos O. Escobar, Léa Ferreira dos Santos

and

Paulo C. Marques Filho

Instituto de Física da Universidade de São Paulo

01498-970 C.P. 20516-São Paulo, SP

Brazil



Abstract

We consider in this article the quantum limits for measurements on macroscopic bodies, which are obtained in a novel way, employing the concept of decoherence coming from an analysis of the quantum mechanics of dissipative systems. Two cases are analysed, the free particle and the harmonic oscillator and for both systems, we compare our approach with previous treatments of such limits.

One of the outstanding problems in contemporary physics is the detection of the gravitational waves predicted by Einstein in his General Relativity Theory [1]. The detection of such waves presents a challenge to the experimentalists who have to monitor extremely small fluctuations in the position of macroscopic bodies [2]. The sensitivity planned for the future generation of detectors, will bring us close to the limits imposed by quantum mechanics [3].

It is our purpose in this letter to examine in detail the obtention of quantum limits for free masses and for the harmonic oscillator, subjected to thermal fluctuations. We rely on the methods developed by several authors to analyse the quantum mechanics of a system in interaction with its environment [4]. We differ in this respect from Braginsky [3] who uses a two step procedure in order to obtain the quantum limits, firstly considering the limits imposed by the uncertainty principle on an isolated system, not subjected to thermal fluctuations and then, in the second step, looking at the same system, this time classically behaved but interacting with the environment. This procedure is, in our view, unacceptable since it ignores that the system is always in contact with the environment. There is no need for this asymmetry, treating the system classically when subjected to thermal fluctuations, but ignoring the latter when viewing it quantum mechanically.

Following the approach of references [4], we are led to consider the master equation for the evolution of the reduced density matrix ρ in the coordinate representation, describing the state of the system under consideration

$$\frac{d\rho}{dt} = -\frac{i}{\hbar} [H, \rho] - \gamma(x - x') \left(\frac{\partial \rho}{\partial x} - \frac{\partial \rho}{\partial x'} \right) - \frac{2m\gamma k_B T}{\hbar^2} (x - x')^2 \rho \quad (1)$$

For our purposes the relevant term in Eq.(1) is the last one. We see that this term contains the thermal fluctuations in a quantum mechanical context, entailing a characteristic scale which gives the decoherence time. The decoherence time characterizes the vanishing of the off-diagonal elements of the density matrix, which provide a measure of the quantumness of the system [4]. From Eq.(1) the decoherence time scale is given by

$$\tau_D = \frac{\hbar^2}{2m\gamma k_B T (\Delta x)^2} \quad (2)$$

In (1) and (2) γ^{-1} is the relaxation time of the system, m the mass and T the temperature. A proper interpretation of Eq. 2 indicates that $(\Delta \mathbf{x})^2$ is the mean square deviation of the position of the system. We will come back to this important point at the end of this letter.

Let us now apply the above concept in order to obtain the quantum limits for two systems, the free particle relevant to laser interferometric antennas, and the harmonic oscillator, relevant to mechanical bars.

Free Particle The quantum limit is obtained when $(\Delta \mathbf{x})^2$ in Eq.(2) becomes of the same magnitude as the uncertainty arising from Heisenberg's principle, when applied to successive measurement of the position of the free particle separated by a time interval τ . This quantum mechanical uncertainty is given by [3]

$$(\Delta \mathbf{x})^2 \approx \frac{\hbar \tau}{m} \quad (3)$$

Inserting (3) into the expression for τ_D (Eq.(2)), we obtain

$$\tau_D = \frac{\hbar}{2\gamma k_B T \tau} \quad (4)$$

If τ_D is greater than the time interval between two successive measurements, τ , the system must be treated quantum mechanically. When this happens we obtain the quantum limit

$$\hbar > 2\gamma k_B T \tau^2 \quad (5)$$

Since γ^{-1} is the relaxation time of the system (τ^*), Eq.(5) can be rewritten in the more usual form

$$\hbar > 2k_B T \frac{\tau^2}{\tau^*} \quad (6)$$

which is the limit obtained by Braginsky [3], who arrived at this result by a somewhat obscure path, as we will make explicit later on.

Harmonic Oscillator For a harmonic oscillator of fundamental frequency ω , the Heisenberg uncertainty principle gives

$$(\Delta x)^2 \approx \frac{\hbar}{2m\omega} \quad (7)$$

for a measurement time of the order of the period of the harmonic oscillator ($\tau \approx \frac{2\pi}{\omega}$).

Replacing (7) into (3) leads to

$$\tau_D = \frac{\hbar\omega}{\gamma k_B T} \quad (8)$$

Imposing $\tau_D > \tau$ gives the quantum limit for the oscillator

$$\hbar > \gamma \frac{k_B T \tau}{\omega} = \frac{k_B T}{\omega} \frac{\tau}{\tau^*} \quad (9)$$

A similar result was obtained by Braginsky [3], only differing from (9) by a numerical factor.

Having shown how to obtain the quantum limits in an internally consistent way, which takes into account both quantum and thermal fluctuations, we now raise one further objection to the derivation by Braginsky of the quantum limit for a free particle.

The authors of reference [5] use as the starting point, Nyquist's theorem which gives for the spectral density of the fluctuating force the result

$$\langle F_{fl}^2 \rangle_\omega = 4k_B T \frac{\gamma}{m} \quad (10)$$

Then they use this fluctuating force properly integrated over a range of frequencies $\Delta\omega \approx \tau^{-1}$, to find the displacement of the particle, $\mathbf{x} = \mathbf{x}_0 + \frac{1}{2}at^2$ under a constant acceleration given by

$$a = \frac{\sqrt{\langle F_{fl}^2 \rangle_\omega \tau^{-1}}}{m} \quad (11)$$

which then gives for the *displacement*

$$\Delta x = \sqrt{\frac{k_B T}{m} \frac{\tau^3}{\tau^*}} \quad (12)$$

This result is in disagreement with a standard treatment of the Brownian classical particle [6]. Of course the latter treatment refers to the mean square deviation of the position of the particle, which as a matter of fact is what is monitored in gravitational wave antennas [2], while Braginsky considers a displacement, to be later on compared with a fluctuation (from Heisenberg's principle). By sheer coincidence his result is the same as the one obtained by us. Finally we remark that when Braginsky considers the harmonic oscillator, the classical part is treated in the right way, as a fluctuating Brownian particle in the potential of a harmonic oscillator.

We finish this letter with a comment to the respect of the validity conditions of the master equation (1). It is normally referred in the literature as been derived in the high-temperature limit. The results of Zurek [7] indicate however its validity outside this domain. Next we plan to attack the problem of quantum non-demolition measurements [8] taking into account dissipative effects.

ACKNOWLEDGMENTS

L.F.dos Santos and P.Marques Filho would like to thank Fundação de Amparo à Pesquisa do Estado de São Paulo for financial support.

REFERENCES

- [1] A. Einstein, Sitzungsber Preuss.Akad.Wiss., 688 (1916).
- [2] For a review of the current status of G.W. detectors see, "The detection of Gravitational Waves" ed. D.G. Blair, Cambridge University Press, 1991.
- [3] V.B. Braginsky and Yu.I. Vorontsov, Sov. Phys Usp. **17**, 644 (1974); V.B.Braginsky, Sov. Phys. Usp. **31**, 836 (1988).
- [4] A.O. Caldeira and A.J. Leggett, Physica A **121**, 587 (1983). W.G. Unruh and W.H. Zurek, Phys.Rev. D **40**, 1071 (1989). W.H. Zurek. Phys.Today **44**,No 10, 36 (1991) (for an introduction).
- [5] V.B. Braginsky and A.B. Manukin, "Measurements of Weak Forces in Physical Experiments", Univ. of Chicago Press, 1977.
- [6] F. Reif, " Fundamentals of Statistical and Thermal Physics", McGraw-Hill, 1965.
- [7] W.H. Zurek, S. Habib and J.P. Paz, Phys.Rev.Lett., **70**, 1187 (1993).
- [8] J.N. Hollenhorst, Phys.Rev. D **19**,1669 (1979).

NEW PROPOSED EXPERIMENT ON TIME-VARIABILITY OF THE FINE-STRUCTURE CONSTANT

Yasunori Fujii*
Nihon Fukushi University
Okuda, Chita-gun, Aichi, 470-32 Japan
and
Kazuaki Kuroda and Nobuyuki Kanda
Institute of Cosmic Ray Research
University of Tokyo
Tanashi, Tokyo, 188 Japan



(Presented by Y. Fujii)

ABSTRACT

Solving the cosmological constant problem in the unified theories by means of a scalar field results in the gauge coupling constants which are time-dependent beyond the observational upper bounds by many orders of magnitude. We propose a remedy by exploiting “hesitation behavior” of the scalar field, a highly nontrivial solution of the cosmological equations. In this connection we also propose a new type of experiment to probe $\dot{\alpha}/\alpha$, time variability of the fine-structure constant, by using a high-finesse Fabry-Perot interferometer.

*e-mail address: ysfujii@tansei.cc.u-tokyo.ac.jp

In spite of remarkable achievements for an ultimate theory supposed to unify particle physics and gravitation, we find the cosmological constant still remaining to be a major problem; the theoretical prediction is larger than the observational upper bound by as much as 120 orders or so. So serious that, unfortunately or understandably, only few people have tried to face it. But we have to do something. Otherwise the whole program for unification will be seriously undermined. Probably the only promising way out is to *abandon the very notion that the cosmological constant is a true constant.*

More specifically, one would expect that $\Lambda(t)$ decays like $\sim t^{-2}$ as a function of the cosmic time t . This is nice because it gives us a simple and natural way of understanding a small number like 10^{-120} . The point is that the present age of the universe $t_0 \sim 10^{10}y$ is of the order of 10^{60} in units of the Planck time $\sim 10^{-43}sec$, which is the fundamental time unit in the physics of unification. Its inverse square gives naturally 10^{-120} . In this scenario of "a decaying cosmological constant," *today's cosmological constant is unusually small only because our universe is old*, not due to any unnatural fine-tuning of parameters.

The simplest way to implement this scenario, as a dynamical effect starting out from a truly constant Λ , is to introduce a scalar field, as in the Jordan-Brans-Dicke theory [1-2]. It is amusing to find that in most of the theoretical models of unification we find some candidates of the scalar field of this type. The most probable candidate is the "dilaton" field, having some relevance to two-dimensional conformal invariance of the theory.

If we look into some details of this theory, we find that the scalar field must grow with time *steadily* without turning back. This is a condition necessary for the success of the scenario, as we can see from Weinberg's argument [3]. We should accept a scalar field which *keeps changing with time even today.*

Also there is a difference from the original JBD theory. The dilaton field, or almost any other candidate scalar field, couples to matter fields, particularly to the gauge fields. Combining this with the above result we come to conclude that the *observed* gauge coupling constants, including the usual fine-structure constant, must keep changing with time even today. Most naively, we expect a power-law behavior of $\alpha(t)$, then $\dot{\alpha}/\alpha \sim 10^{-10}y^{-1}$ at the present time.

How about the observation? Here is a brief summary of the past results.

Table I: Observational upper bounds on $\dot{\alpha}/\alpha$ and $\dot{\alpha}_s/\alpha_s$, for the electromagnetic(E) and strong interaction(S), respectively.

Source	Interaction	Upper bound (y^{-1})
Very long-lived nuclei [4]	E	3×10^{-13}
Primordial nucleosynthesis [5]	S	2×10^{-12}
Stellar nucleosynthesis [6]	S	10^{-13}
Distant QSO [7]	E	4×10^{-12}
Oklo phenomenon [8]	S	5×10^{-19}
Comparison with atomic clocks [9]	E	3×10^{-13}

Results for the electromagnetic and strong interactions should be interpreted as essentially the same from the point of view of unified theories. As we see, these are all the upper bounds;

no evidence for the time-variability has been ever reported at the level of 10^{-10}y^{-1} . The most stringent constraint gives a number as small as $\sim 10^{-19}\text{y}^{-1}$, nearly 9 orders smaller than what we expect naively. This seems to be another serious confrontation between theory and observation.

I propose a remedy, taking advantage of what I call a “hesitation” behavior of the scalar field [2]. Recently we discovered a *highly non-trivial solution* of the cosmological equations in which the scalar field, which basically grows with time, may stay nearly constant for some time. This is due to a subtle competition between the “force” driving the scalar field and the cosmological “friction.” According to this mechanism, the observed no time-variability of the coupling constants would be because we at the present time happen to live during this hesitation period. I also emphasize that this behavior is a rather *common phenomenon* sharing essentially the same origin as what is widely known as the “relaxation oscillation” in nonlinear systems. Incidentally, the cosmological constant had decayed sufficiently fast before the onset of the hesitation.

The obvious drawback of this approach is that the theory is so flexible that it lacks predictive power; we have no unique prediction on how much $\dot{\alpha}/\alpha$ should be. At the same time, however, we have no theoretical reason why $\dot{\alpha}/\alpha$ should be far below the upper bounds obtained so far. It may be waiting for to be discovered right there. In this sense, searching for $\dot{\alpha}/\alpha$ with a better accuracy might be justified. In what follows, we suggest a possible new type of experiment.

As an example, consider an alkali atom with a fine-structure doublet. For the two transitions ($i = 1, 2$), the frequencies or the wave-numbers are given by

$$k_i = k_0 \left(1 + \beta_i \alpha^2 \right), \quad (1)$$

where $\alpha = e^2/\hbar c \approx 1/137$. Also $k_0 = (\mu c Z^2 \alpha^2 / 2\pi^2 \hbar)$ and

$$\beta_i = \frac{Z^2}{n^2} \left(\frac{n}{|K_i|} - \frac{3}{4} \right), \quad (2)$$

where $K_i = -\ell - 1$ or ℓ , according to $j_i = \ell + 1/2$ or $\ell - 1/2$, respectively. Notice that the ratio of the two wavelengths depends only on α^2 and a pure constant. Measure this ratio at different occasions separated by a year, for example. We then may be able to probe a possible change of α .

For this purpose we prepare a Fabry-Perot cavity as an ultra sensitive spectrometer, as shown in Fig. 1. A Fabry-Perot cavity or resonator consists of a completely reflecting mirror and a partially reflecting mirror with the reflection coefficient R , separated by a distance L . Suppose a laser beam is injected, taking aside the complication due to two beams for the moment. Major part of the beam is reflected back on the surface of the first mirror but the remaining part goes into the inside, going back and forth, re-emerging and producing an interference pattern at the photodetector.

The re-emerging beam would be strong enough only at resonances defined by $kL = 2\pi \times$ integer. On the other hand, the resonance width with respect to kL is $1 - \sqrt{R} \approx \frac{1}{2}(1 - R)/\sqrt{R} = (\pi/2)\mathcal{F}^{-1}$, where \mathcal{F} is called “finesse.” The closer R to 1, the larger \mathcal{F} , and hence giving the better resolving power of the spectrometer. Now suppose the length L is changed continuously.

Then at the detector we observe “dark fringes” appearing successively as L passes through the resonance positions.

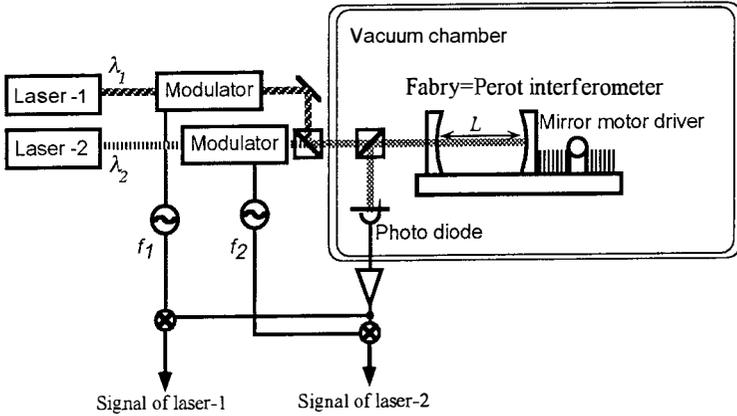


Figure 1: A schematic illustration of the apparatus.

In fact we inject two laser beams corresponding to the two transitions simultaneously and coaxially. But they are distinguished from each other by applying different modulation frequencies. Make certain adjustment such that we observe dark fringes simultaneously for the two beams. Starting from this position of coincident dark fringes, move the mirror. We would observe no coincident dark fringes any more because the wavelengths are different. But after N_1 (N_2) dark fringes have passed in the beam 1 (2), we may find another coincident or near coincident dark fringes again for $N_1/N_2 \approx k_1/k_2$. We may think of a vernier. Of course, the coincidence should be approximate in practice.

Repeat the same experiment a year later ($\Delta t = 1y$). Namely first find a position of a coincident dark fringes for certain length. Starting from this position change the length L . After having passed N_1 dark fringes in the beam 1, check if we find the coincident dark fringes again in the beam 2. If we do as before, α must be the same as a year before. If we find instead a shift of the positions of the two dark fringes, it would indicate the change of α .

This shift can be expressed in terms of the phase difference $\delta\Phi$ given by

$$\delta\Phi = \frac{\mathcal{F}}{\pi} 8(\Delta\beta)\alpha^2 \frac{\dot{\alpha}}{\alpha} (\Delta t) k_0 (\Delta L), \quad (3)$$

where ΔL is the change of distance, and $\Delta\beta$ is the difference of β_i . The phase difference is naturally proportional to $\dot{\alpha}/\alpha$ and the sensitivity, namely the finesse.

As we learn, finesse as large as 10^6 is now available [10]. To measure the phase difference, we lock one of the wavelengths to a stabilized standard laser with its accuracy $\sim 10^{-13}$. This implies $\delta\Phi \sim 10^{-7}$. Using these values together with $\Delta\beta \sim 10^2$ and $\Delta L \sim 10\text{cm}$, we find that we can probe $\dot{\alpha}/\alpha$ to $10^{-17}y^{-1}$. We may expect to improve the result even further. In spite of

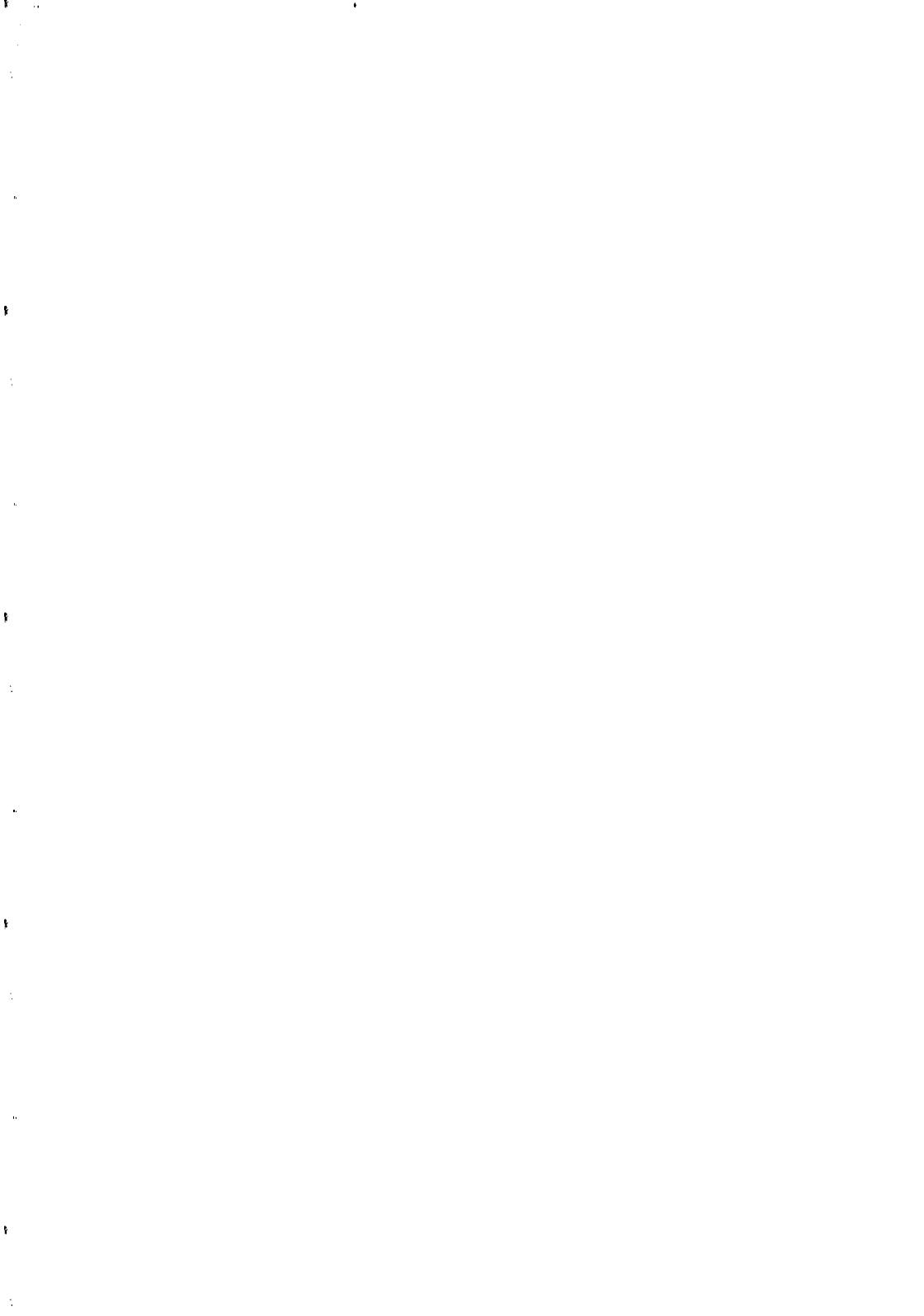
many technical difficulties ahead, I believe this is a promising way, because this is a *controlled laboratory experiment* unlike most of the past attempts.

To conclude, I add two remarks. First, the scalar field may show up as the fifth force [11]. Although without any experimental evidence so far [12], *the theoretical motivation for this phenomenon is still strong*. We have some good reasons why the strength would be weaker than what had been suspected earlier [2]. We still encourage the experimentalists to continue their efforts whenever a new technology becomes available for the better accuracy. I myself propose a new experiment by using an ultra-sensitive laser interferometer with high-finesse Fabry-Perot cavities, as an improvement of my past suggestion [13].

Secondly, and finally I add [14] that the hesitation behavior of the scalar field may allow us to understand a possible *small but nonzero cosmological constant* recently suggested by a number of cosmological analyses [15].

REFERENCES

- [1] A.D. Dolgov, in *The Very Early Universe*, Proc. Nuffield Workshop, 1982, eds., G.W. Gibbons and S.T. Siklos, (Cambridge University Press, 1982): Y. Fujii, Phys. Rev. **D26** (1982), 2580; L.H. Ford, Phys. Rev. **D35**, 2339 (1987); Y. Fujii and T. Nishioka, Phys. Rev. **D42**, 361 (1990).
- [2] Y. Fujii, M. Omote and T. Nishioka, Time-Dependent Coupling Constant in Unified Theories? Preprint.
- [3] S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989).
- [4] F.J. Dyson, Phys. Rev. Lett. **19**, 1291 (1967).
- [5] P.C.W. Davies, J. Phys. **A5**, 1296 (1972).
- [6] F. Hoyle, *Galaxies, Nuclei and Quasars*, Heinemann, London, 1965.
- [7] A.M. Wolfe, R.L. Brown and M.S. Roberts, Phys. Rev. Lett. **37**, 179 (1976).
- [8] A.I. Shlyakhter, Nature, **264**, 340 (1976).
- [9] A. Godone, C. Novero and P. Tavella, Phys. Rev. Lett. **71**, 2364 (1993).
- [10] G. Rempe *et al*, Optics Letters, **17**, 362 (1992).
- [11] Y. Fujii, Nature Phys. Sci, **234**, 5 (1971); E. Fischbach, *et al*, Phys. Rev. Lett. **56**, 3 (1986); Y. Fujii, Int. J. Mod. Phys. **A6**, 3505 (1991).
- [12] See, for example, E. Fischbach and C. Talmadge, Nature, **356**, 207 (1992); M.V. Moody and H.J. Paik, Phys. Rev. Lett. **70**, 1195 (1993).
- [13] Y. Fujii, Phys. Lett. **B255**, 439 (1991).
- [14] Y. Fujii and T. Nishioka, Phys. Lett. **B254**, 347 (1991).
- [15] See, for example, M. Fukugita *et al*, Ap. J. **361**, L1 (1990); E. Bertschinger, Nature, **348**, 675 (1990).



ON THE DETECTABILITY OF THE POST-NEWTONIAN EFFECTS IN
THE GRAVITATIONAL WAVE EMISSION OF A COALESCING BINARY

Kostas D. Kokkotas

Department of Physics, Aristotle University of Thessaloniki, Macedonia, Greece

Andrzej Królak

Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8,

00-950 Warsaw, Poland

Gerhard Schäfer

Max-Planck-Gesellschaft, Arbeitsgruppe Gravitationstheorie

an der Friedrich-Schiller-Universität, Max-Wien-Platz 1, 7743 Jena, Germany



ABSTRACT

The effect of the post-Newtonian corrections on the detectability of the gravitational-wave signal from a coalescing binary is investigated. We show that the post-Newtonian signal can be identified with a high probability using the Newtonian waveform as a filter.

1 Introduction

It is currently believed that the gravitational waves that come from the final stages of the evolution of compact binaries just before their coalescence will be the most likely waves to be detected by long arm laser interferometers ¹⁾. (The two projects to build such detectors - LIGO and VIRGO - are already approved and they are rapidly progressing.) The reason is that in the case of binary systems we can predict the gravitational waveform very well; and the amplitudes are reasonably high for sources at distances out to 200 Mpc so that several events per year can be expected. The waveform derived using the quadrupole formula has been known for quite some time ²⁾. A standard optimal method to detect the signal from a coalescing binary in a noisy data set and to estimate its parameters is to correlate the data with the filter matched to the signal and vary the parameters of the filter until the correlation is maximal. The parameters of the filter that maximise the correlation are estimators for the parameters of the signal. The detailed algorithms and the performance of the matched-filtering method has been investigated in several references ^{3,4,5)}. It has been recently realized ⁷⁾ that the correlation is very sensitive to even very small variations of the phase of the filter. Consequently the addition of small corrections to the phase of the signal due to the post-Newtonian effects decreases the correlation considerably. Thus the post-Newtonian effects in the coalescing binary waveform can be detected and estimated to a much higher accuracy than it was thought before ⁶⁾. This opens new prospects but also considerable data analysis challenges for the LIGO and VIRGO projects. It was also found ⁷⁾ that the post-Newtonian series is not converging rapidly for a binary near coalescence. Hence higher post-Newtonian corrections will affect the correlation. Only the first two post-Newtonian corrections to the quadrupole formula are presently known and the calculation of further corrections is a major problem. In this article we analyse the detectability of the post-Newtonian signal using the Newtonian waveform as a filter. This tool was suggested as a possible search template by Cutler et al. ⁷⁾. The units are chosen such that $G = c = 1$.

2 Post-Newtonian effects

Let us first give the formula for the gravitational waveform of a binary with the two currently known post-Newtonian corrections. We make two further approximations to the waveform. We assume that the orbit of the binary is quasi-circular, which is justified due to the effect of rapid circularization of the orbit by radiation reaction, and we only include the post-Newtonian corrections to the phase of the signal keeping the amplitude in its Newtonian form; this is because the effect of the phase on the correlation is dominant. With these approximations the waveform, as a function of time, is given by the following expression,

$$(1) \quad h(t) = \frac{8}{5} \pi^{2/3} \frac{\mu m^{2/3}}{R} f(t)^{2/3} \cos[2\pi \int_{t_m}^t f(t') dt' + \phi],$$

where ϕ is an arbitrary phase, μ and m are the reduced and the total mass of the binary, respectively; t_m is some time parameter and R is the distance to the source. The characteristic time for the evolution of the binary to the currently known 3/2 post-Newtonian order is given by ⁷⁾

$$(2) \quad \tau_{PN3/2} := \frac{f}{df/dt} = \frac{5}{96} \frac{1}{\mu m^{2/3}} \frac{1}{(\pi f)^{8/3}} \times \\ \left[1 + \left(\frac{743}{336} + \frac{11}{4} \frac{\mu}{m} \right) (\pi m f)^{2/3} - (4\pi - s)(\pi m f) \right]$$

where s is a spin parameter.

To calculate the expectation value of the correlation of the noisy data with the matched filter it is convenient to go to the frequency domain. The expression for the Fourier transform of our signal in the stationary phase approximation is given by

$$(3) \quad \tilde{h} = \tilde{A} f^{-7/6} \exp[i2\pi f t_m + \phi - \pi/4 + a(f)k + a_1(f)k_1 + a_{3/2}(f)k_{3/2}], \text{ for } f > 0,$$

(and by the complex conjugate of the above expression for $f < 0$) where

$$(4) \quad \tilde{A} = \frac{1}{(30)^{1/2}} \frac{1}{\pi^{2/3}} \frac{\mu^{1/2} m^{1/3}}{R},$$

$$(5) \quad k = \frac{1}{\mu m^{2/3}}, \quad k_1 = \frac{1}{\mu} \left(\frac{743}{336} + \frac{11}{4} \frac{\mu}{m} \right), \quad k_{3/2} = \frac{m^{1/3}}{\mu} (4\pi - s),$$

$$(6) \quad a(f) = \frac{1}{128} \left(\frac{3}{(\pi f)^{5/3}} + \frac{5\pi f}{(\pi f_m)^{8/3}} - \frac{8}{(\pi f_m)^{5/3}} \right),$$

$$(7) \quad a_1(f) = \frac{5}{96} \left(\frac{1}{\pi f} + \frac{\pi f}{(\pi f_m)^2} - \frac{2}{\pi f_m} \right),$$

$$(8) \quad a_{3/2}(f) = -\frac{1}{32} \left(\frac{3}{(\pi f)^{2/3}} + \frac{2\pi f}{(\pi f_m)^{5/3}} - \frac{5}{(\pi f_m)^{2/3}} \right)$$

hold and where $f_m = f(t_m)$. Notice $a(f_m) = a_1(f_m) = a_{3/2}(f_m) = 0$.

The stationary phase approximation, eq. (3), is an excellent approximation for frequencies f that are somewhat larger than the initial frequency f_i where the signal starts. In the following calculation we shall choose $f_i = 10\text{Hz}$.

The expression for the expectation value of the correlation is given by

$$(9) \quad H(\Delta t, \Delta\phi, \Delta k, \Delta k_1, \Delta k_{3/2}) = 4\tilde{A}^2 \int_{f_i}^{\infty} \frac{df}{S_h(f) f^{7/3}} \cos[2\pi f \Delta t + \Delta\phi + a(f)\Delta k + a_1(f)\Delta k_1 + a_{3/2}(f)\Delta k_{3/2}],$$

where Δt means the difference in the time parameters of the signal and the filter and $S_h(f)$ is the spectral density of the noise. The performance of the matched filtering technique is determined by the following two quantities, signal-to-noise ratio $d = \sqrt{H(0, 0, 0, 0, 0)}$ and the Fisher information matrix $\Gamma = \partial^2 H / \partial \Theta_S \partial \Theta_F$ ($\Theta_S = \Theta_F$), where S refers to the parameters of the signal and F to the parameters of the filter³⁾. The inverse of the matrix Γ gives bounds on the covariances of the parameters of the signal.

In this investigation we shall use a simple approximation for the spectral density $S_h(f)$ of the noise in the advanced LIGO detectors, devised by Cutler and Flanagan⁵⁾,

$$(10) \quad S_h(f) = S_o((f_o/f)^4 + 2(1 + (f/f_o)^2))/5,$$

where $f_o = 70\text{Hz}$ and $S_o = 3 \times 10^{-48} \text{Hz}^{-1}$. Then for a binary at the distance of 100Mpc with components of 1.4 solar mass (M_\odot) each and spin parameter $s = 0.024$ we get for the signal-to-noise ratio and the accuracy of the mass parameters

$$(11) \quad d = 29.6,$$

$$(12) \quad \frac{\delta\mu}{\mu} = 3.39 \times 10^{-2}, \quad \frac{\delta m}{m} = 5.06 \times 10^{-2}.$$

We see that the accuracy with which the post-Newtonian effects can be estimated is high.

3 The Newtonian filter

The high performance of the filtering will only be possible if the filter can match the post-Newtonian effects accurately. We demonstrate that this can be done just with the Newtonian filter. Let us consider the correlation of the post-Newtonian signal with the Newtonian filter. Then the argument Φ of the cosine in the integrand of the correlation, eq. (9), takes the form,

$$(13) \quad \Phi(f) = 2\pi f \Delta t + \Delta\phi + a(f)\Delta k + a_1(f)k_1 + a_{3/2}(f)k_{3/2}.$$

Let us examine the functional behaviour of Φ , eq. (13), around the frequency f_m . We find

$$(14) \quad \Phi(f) = 2\pi f \Delta t + \Delta\phi + \frac{5}{96}(f/f_m - 1)^2 \left[\frac{\Delta k}{(\pi f_m)^{5/3}} + \frac{k_1}{\pi f_m} - \frac{k_{3/2}}{(\pi f_m)^{2/3}} \right] + O[(f/f_m - 1)^3].$$

The amplitude of the integrand $\frac{1}{S_h(f)^{f/3}}$ in eq. (9) has a maximum at a frequency f'_o where $f'_o = 46.57\text{Hz}$ for the frequency f_o given above. The expectation value of the correlation is maximal when the absolute value of the phase function is as small as possible near the frequency f'_o . From eq. (14) we get the following solution for the parameters of the filter,

$$(15) \quad \Delta t_{max} = t_{mFmax} - t_{mS} \simeq 0, \quad \Delta\phi_{max} = \phi_{Fmax} - \phi_S \simeq 0,$$

$$(16) \quad \Delta k_{max} = k_{Fmax} - k_S \simeq -k_1(\pi f_m)^{2/3} + k_{3/2}(\pi f_m).$$

For the frequency $f_m = 46.57\text{Hz}$ this gives $k_{Fmax} - k_S = -0.01509M_\odot^{-5/3}$. We have also investigated the problem numerically, using expression (13). For convenience we have chosen the upper bound of integration to be 200Hz and found the maximum to be located at

$$(17) \quad \Delta t_{max} = -3.727 \times 10^{-3}\text{sec}, \quad \Delta\phi_{max} = 1.491, \quad \Delta k_{max} = -0.01558M_\odot^{-5/3}.$$

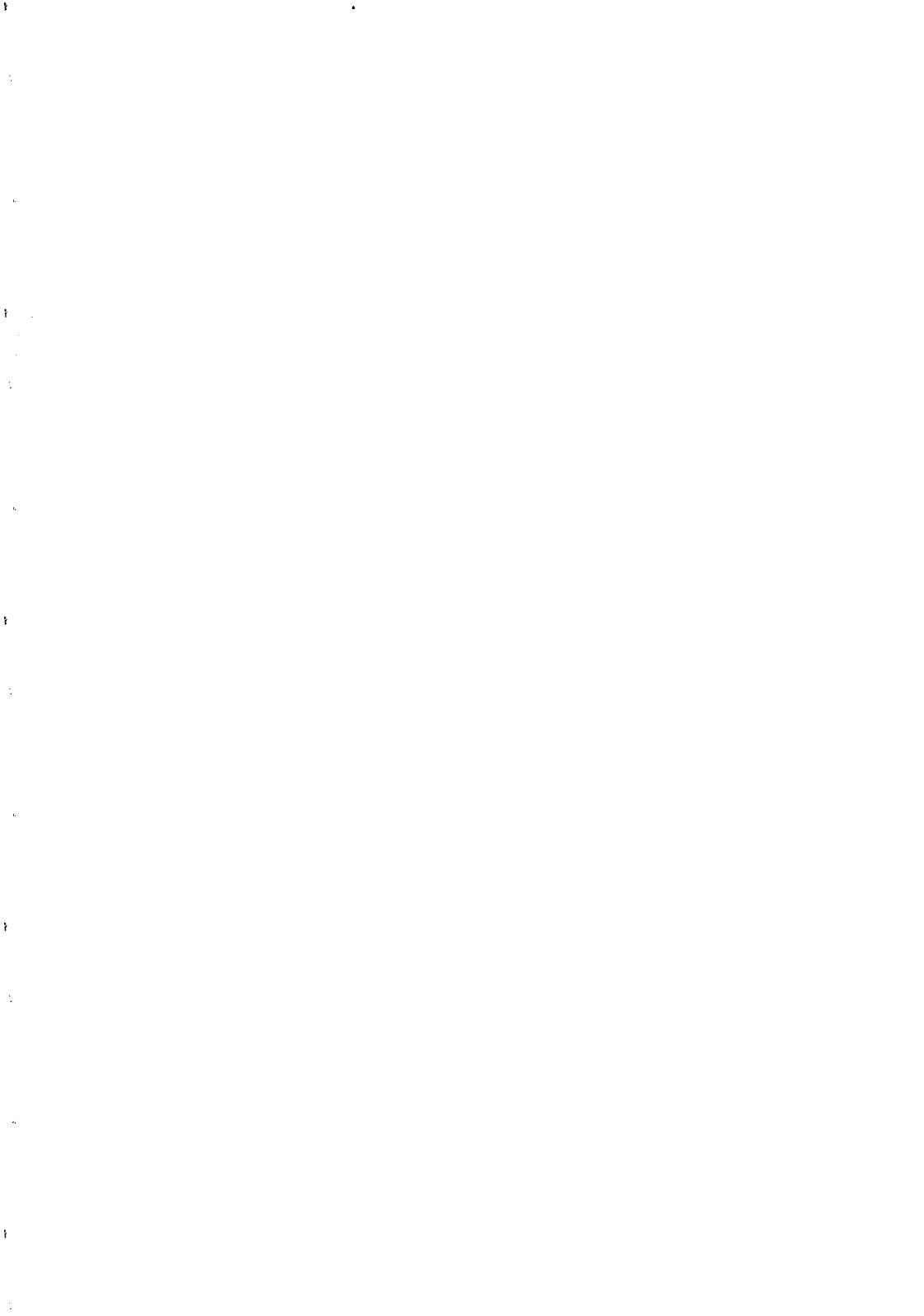
The signal-to-noise d_{max} ratio was found to be 26.5, i.e. only 10.5% less than the value for the post-Newtonian filter. Thus, the detectability with the Newtonian filter is degraded very little. Moreover, the estimated value of the Newtonian parameter k will be shifted by an amount determined by the post-Newtonian corrections.

Acknowledgment One of us (A. Królak) would like to thank the Max-Planck-Gesellschaft for support and the Arbeitsgruppe Gravitationstheorie an der Friedrich-Schiller-Universität in Jena for hospitality during the time this work was done. This work was also supported in part by grant KBN 2 P302 076 04.

References

- [1] K.S. Thorne, in *300 Years of Gravitation*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987), pp. 330-458.
- [2] P.C. Peters, Phys. Rev. **136**, B1224 (1964).
- [3] A. Królak, J.A. Lobo, and B.J. Meers, Phys. Rev. D, **48**, 3451 (1993).
- [4] L.S. Finn and D.F. Chernoff, Phys. Rev. D **47**, 2198 (1993).
- [5] C. Cutler and E. Flanagan, Caltech Preprint GRP-369 (1993), submitted to Phys. Rev. D.
- [6] A. Królak, in *Gravitational Wave Data Analysis*, edited by B.F. Schutz (Kluwer, Dordrecht, 1989), pp. 59-69.
- [7] C. Cutler *et al.*, Phys. Rev. Lett. **70**, 2984 (1993).

COMMON SESSION WITH MESOSCOPIC SYSTEMS



THE GRAVITATIONAL ACCELERATION OF ANTIMATTER

M. H. Holzscheiter
Los Alamos National Laboratory
Los Alamos, NM 87545, U.S.A.



ABSTRACT

We have proposed measuring the acceleration of antiprotons in the Earth's gravitational field by launching antiprotons from a thermal distribution at 4 K upwards against the force of gravity and measuring their time-of-flight (TOF). The TOF distribution thus obtained will exhibit a cut-off representing the minimum kinetic energy necessary to reach the detector at the top of the experiment. The cut-off time is independent of the inertial mass of the particles and is a direct measure of g for the particles studied. We propose to compare the cut-off time, and thereby g , of negative hydrogen ions and antiprotons. The single most difficult problem to be solved for this method consists of shielding all stray-electric fields to a level where the force of gravity is the dominating force acting on the particle. Alternative methods for reducing the effect of stray-electric fields are discussed and a brief analysis of experimental possibilities using neutral antihydrogen atoms is presented.

1. THE FREE FALL EXPERIMENT - PS200

In 1986 we have proposed¹⁾ measuring the acceleration, g , of antiprotons in the Earth's gravitational field by launching antiprotons from a thermal distribution at 4 K in groups of approximately 100 particles upwards against the force of gravity and measuring their time-of-flight (TOF) for a 1 m flight path. Due to the distribution of initial energies, a time of flight distribution will be observed. This TOF distribution will exhibit a cut-off representing the minimum kinetic energy necessary to reach the detector at the top of the experiment. This cut-off time is independent of the particle's inertial mass and of the detailed shape of the thermal distribution the particles are launched from, and is a direct measure of g for the particles under study. We plan to compare this cut-off time, and thereby g , of negative hydrogen ions and antiprotons to a level of accuracy of better than 0.1 %.

1.1 Capture and cooling of antiprotons

An important part of experiment PS200 is the supply of a large number of ultra-low energy antiprotons. Numerical models have shown that launching 10^6 to 10^7 particles from a thermal distribution at 4 K will allow a measurement of g at the 0.1 % level (on statistical grounds only). Many repetitive shots will be necessary to study systematic effects, but a minimum of 10^6 particles is needed in the trap at one time to collect a single time-of-flight spectrum with meaningful statistics.

The lowest energy at which antiprotons are currently available is 5.9 MeV at the Low Energy Antiproton Ring (LEAR) at CERN. This is much too high for this experiment and a staged deceleration scheme has been developed and tested over the past years to bridge this gap of nine orders of magnitude. High energy antiprotons are slowed down by passing them through materials, due to collisions with the nuclei in the material antiprotons lose energy and eventually come to rest and annihilate. If the material is chosen thin enough, antiprotons can escape before annihilating, and the number of antiprotons escaping with kinetic energies below a specific value can be optimized. We have found experimentally, that the yield of antiprotons emerging at energies below 30 keV (50 keV) is maximized around 2.5% (5%) by setting the thickness of the "degrader" material such that 50 % of the incident antiprotons are

transmitted²⁾. Particles at these energies can be captured in a combination of electrostatic and magnetic fields, commonly known as Penning traps³⁾. Once captured, they can be cooled by interaction with a cold gas of electrons stored in the same volume⁴⁾.

For ease of injecting external beams, as well as to allow a large ratio between axial and radial dimensions of the trap which is necessary to match the trap phase space to the accelerator output, we have chosen an open end-cap design consisting of a series of cylindrical electrodes with specific length to radius ratios⁵⁾. The trap consists of two nested system: (a) a long, non-harmonic portion comprised of the degrading foil and a cylindrical electrode which are separated by approximately 50 cm distance and can support up to 60 kV electrical potential with respect to ground, and (b) a central portion consisting of 5 cylinders with lengths carefully adjusted to provide a harmonic potential near the center of the trap. The entire system is situated in the cryogenic bore of a horizontal, super conductive magnet system which provides an axial magnetic field of up to 6 Tesla.

The principal design parameter for the trap is the length necessary to capture all particles of energies up to 30 keV kinetic energy emerging from the degrading foil during a pulse from LEAR of typically 200 ns duration. At a 1 m round trip distance, the time remaining after the last particle has entered the trap before the first 30 keV particle is reflected back to the entrance is 220 ns. Our 30 kV pulser has a 125 ns rise time, allowing 95 ns for jitter and uncertainty in the trigger timing under present conditions. A 60 kV pulser, capable of rise times of 90 ns has been constructed and tested at Los Alamos, and can be added to the system at a later date to capture a larger energy bite from the incident beam.

In testing the first part of the experiment, the PS200 catching trap, we have successfully captured up to approximately 1,000,000 antiprotons from a single, fast extracted pulse from LEAR. We have stored these particles for several minutes and have observed electron cooling to low energy with a time constant of approximately 10 - 20 seconds⁶⁾.

1.2 The free fall experiment

The principle of the proposed method to measure the gravitational acceleration of antiprotons has been pioneered by a challenging experiment performed at Stanford University in the 60's⁷⁾. The authors of this experiment reported a zero net force on free electrons traversing a vertical shield tube. This result was explained by the fact, that the free electron gas in the metal tube, used to shield external electric fields, sagged under the influence of gravity until an electric field was generated by the density gradient thus produced, which counteracted the force of gravity. To obtain this measurement, stray electric fields due to variations in the work function along the inside surface of the shield tube (the "Patch effect") had to be below 10^{-11} V/m.

In the experiment proposed here the requirements on reducing the field inside the tube can be relaxed by three orders of magnitude due to the inertial mass of the antiproton being 2000 times larger than the electron mass. Nevertheless, due to the controversy surrounding the interpretation of the electron experiment, a systematic study of the patch fields has been recognized as an important part of our experimental program. To minimize the electric fields on the central axis of a cylinder due to patches with differing work functions along the surface of the inside wall of the cylinder, one needs to minimize the variation in work function from patch to patch as well as the size of the individual patches. Patch sizes can either be dominated by the intrinsic structure of the material or by adsorbs on the surface. For these reasons it becomes apparent that an amorphous, non-reactive, single component (no alloys) material is the best choice. We have used a vibrating capacitive probe (Kelvin probe) to study small samples of different surfaces and have identified two candidates: graphite in form of an aerosol spray (AerodagTM) and ultra-thin layers of gold on a germanium sub-layer⁸⁾. Both these surfaces do not exhibit any work function variations at the level of the instrument resolution of the Kelvin probe of 1 mV, but the observed changes in overall work function when using graphite samples with different degrees of orientational disorder may point towards work function variations not too far below the resolution limit⁹⁾. Such an effect is not expected for the gold/germanium surfaces.

One obvious solution to increase the sensitivity of these studies is to use the time-of-flight method itself as a probe for the surface electric fields present in the

system. For this purpose we have constructed a test experiment consisting of an ion source, a drift tube, and a particle detector¹⁰⁾. Ions are being transferred from the ion source to the entrance of the drift tube at 1 - 2 keV energy to minimize the error in the definition of the start time. Here they are retarded by an electric potential to zero mean kinetic energy. Some ions are rejected, others are still traversing the drift tube at high energies, but with proper choice of the current density in the pulse from the ion source a small portion of ions (preferably less than one per pulse) is entering the drift tube near zero kinetic energy. These particles are sensitive to the fluctuations of the electric field along the tube axis and their time-of-flight will provide a measure of the rms fluctuations down to a level comparable to the force of gravity on the ion species used. We will start with relatively heavy ions (i.e. Xe) to probe electric field fluctuations at the 10^{-5} level and then successively increase the sensitivity of these tests by choosing lighter and lighter ions until ultimately arriving at protons.

Preliminary tests of this system have revealed a high neutral background density from the ion source which is saturating the particle detector during and shortly after the ion pulse is injected. This prompted us to install a small Penning trap as an intermediate storage unit between the ion source and the drift tube. This trap can dynamically capture more than 10^6 charged particles from the ion source and release them again with a delay sufficiently long to let the neutral background decay away. We anticipate first results from this set-up late 1994.

1.3. Alternative methods to measure “g” on charged particles

Prompted by the difficulties in shielding stray electric-fields to a level of 10^{-7} V/m and better, alternative methods to measure the gravitational acceleration of antimatter have been discussed. At first sight, an obvious solution would be to use a neutral particle. Before discussing this, I would like to briefly present an alternative idea to measure the gravitational acceleration of a charged elementary particle. This method has been discussed in the literature¹¹⁾ and recently a detailed technical analysis has been completed¹²⁾.

Imagine a charged particle moving horizontally along the field lines of a homogeneous magnetic field. This particle will experience a sideward drift due to the

crossed magnetic and gravitational field, similar to a particle moving in crossed electric and magnetic fields. The displacement by this drift motion is too small to be observed in a linear mode in an apparatus of sensible dimensions, but can be accumulated to macroscopic size by deflecting the particle back and forth between two electrostatic mirrors¹¹⁾. One detailed technical realization of this idea is to confine the particles in a horizontally oriented Penning trap. As shown by Lagomarsino et. al¹²⁾, the result of gravity is an upward displacement of the \mathbf{ExB} drift orbit (magnetron orbit) in the trap proportional to g . The detailed analysis given in reference 12 makes it clear that this is a technically very challenging experiment, but the main differences to the TOF method are the following: (1) The same statistical accuracy can be achieved with a much lower number of particles, and (2) the system is much less sensitive to the patch effect. As a matter of fact, reference 12 assumes a patch potential of 10^{-4} cm·Volts as given as upper limit in reference 8, a value which would completely inhibit the TOF method. This means, should the reduction of the patch field by surface treatments not reach the level necessary for the TOF method, this alternative method would still be able to provide a measurement of g , possibly even in an apparatus of smaller scale than assumed in¹²⁾, depending on the final out come of the patch effect studies. The reason for this is, that in the TOF method the electrical potential on the axis of the tube (which falls off like $1/R$, R being the radius of the tube) dominates the measurement, while in the \mathbf{gxB} method it is the electric field (proportional to $1/R^2$) which enters the problem.

2. ANTIHYDROGEN

While one certainly overcomes the problems with stray-electric fields, new complications enter the picture. First, the difficult step of producing and subsequently controlling, the antihydrogen atom has to be solved, and secondly, a method to actually measure g has to be devised and analyzed for its capabilities.

2.1 Antihydrogen production

Of the variety of schemes proposed in the literature¹³⁾ to produce antihydrogen only the reactions listed below (with the appropriate references for detailed discussions) will yield the ultra-low energy antihydrogen atoms needed.



In the first case, both constituents forming antihydrogen need to be trapped, while in the last two cases only antiprotons need to be confined before the recombination process, while the positron is delivered in form of a positronium beam. Both methods have distinct advantages and disadvantages and, depending on the final application, either one may be the better choice.

The rate constant for process (1) is strongly temperature dependent:

$$(4) \quad \Gamma = 6 \times 10^{-12} (4.2/T)^{9/2} n_e^2 \text{ s}^{-1},$$

and therefore benefits vastly from cooling the particles. Because of the mass difference between positrons and antiprotons, with positrons at 4.2 K the antiprotons could have energies as high as 1 eV before the recombination rate is significantly effected. This could be used to form an eV energy beam of antihydrogen atoms leaving the trap in axial direction, even though the beam quality would strongly depend on how well the axial energy of the antiprotons could be decoupled from the radial energy.

While the rate can be extremely high (with a $10^7/\text{cm}^3$ positron at 4.2 K one obtains $\Gamma = 6 \times 10^6 \text{ s}^{-1}$), a number of problems for practical applications have been identified by the original authors of this proposal: the antihydrogen atoms are being formed in highly excited Rydberg states ($n > 100$), and need to be quickly deexcited before electrostatic field gradients from the Penning trap fields will ionize these states, and before spectroscopic studies could be attempted. A mixture of collisional and spontaneous de-excitation may be used, provided the antihydrogen does not drift outside the positron plasma on the same time scale. Alternatively, laser induced de-excitation could be attempted. Additionally, magnetic field effects both on the recombination process and on the survival of the antihydrogen Rydberg atoms need to be studied in detail before a final assessment of advantages and disadvantages of this method for a specific application can be made.

Alternatively one may enhance the radiative antihydrogen formation rate by several orders of magnitude by coupling the recombination process to a third particle for energy and momentum conservation via collisions between positronium atoms and antiprotons¹⁹⁾. This process can be interpreted as Auger capture of the positron to the antiproton and cross sections have been estimated by Humberston et al.²⁰⁾, using charge conjugation and time reversal to link the cross section for positronium formation in collisions between positrons and hydrogen to the antihydrogen formation cross sections. Early calculations assumed both antihydrogen and positronium to be in the ground state, resulting in a production cross section of $3.2 \times 10^{-16} \text{ cm}^2$ with a broad maximum at a \bar{p} energy of 2.5 keV. Calculations of the total antihydrogen formation cross section using classical and semi-classical methods²¹⁾ have obtained values which are considerably larger than the ground-state results. Values for the formation of antihydrogen in excited states are given by Ermolaev et al.²²⁾ and indicate that there is a large cross section to low-lying excited states.

M. Charlton²³⁾ has discussed the formation of excited $\bar{\text{H}}^*$ atoms via collisions between antiprotons and excited positronium. The cross section follows a classical $n_{p_s}^4$ scaling, where n_{p_s} denotes the principal quantum number of the positronium atom, leading to large enhancements in the reaction rate. This process can also be utilized to preferentially populate specific low level excited states for spectroscopic purposes. An additional feature of the excited state process is the low recoil energy of the resulting antihydrogen atoms.

A final note on rates: much discussion has been put forward about the difference in rate constants for antihydrogen production using either one of the processes discussed here. In essence, this discussion may be viewed as non relevant if the product is to be trapped. At a limited number of antiprotons being available to work with, there is no fundamental difference between a rate of $10^6/\text{sec}$, converting all antiprotons to antihydrogen in a single second, or a rate of 1000/sec (or even 100/sec), requiring a production cycle of 16 minutes (or 3 hours), which still is much shorter than the storage time of antiprotons in traps of several months achieved²⁴⁾.

2.2. Measuring “g” on Antihydrogen

Even though it is repeatedly being mentioned, purely ballistic methods to measure gravitational acceleration on antihydrogen atoms may be ruled out. Even if the antihydrogen atoms can be laser cooled (requiring Lyman α light sources) to the photon recoil limit of $T = \frac{1}{2} \Gamma \hbar / K_B = 2.4$ mK, this temperature would still correspond to a distribution of height of approximately 1 m in the gravitational field. A precise determination of the centroid of a cloud of this dimension is not possible. Similarly, if one should be able to generate an atomic fountain with a mean energy of the photon recoil and a spread of half that value, the observed time-of-flight over a height of 10 cm will be $14 \text{ msec} \pm 7 \text{ msec}$, again not yielding a precision measurement of g . The only hope in the latter case would be to perform an end point measurement similar to the PS200 proposal, but this time with more particles near the end point.

A potentially much more powerful method could be developed based on the work of by S. Chu and collaborators²⁵⁾. In their experiment they have used velocity sensitive stimulated Raman transitions to measure the gravitational acceleration, g , of laser cooled sodium atoms in an atomic fountain geometry. An ultra cold beam from an atomic trap is launched upwards and is subjected to three subsequent pulses to drive a two-photon Raman transition between the $F = 1$ and 2 levels in the $3S_{1/2}$ state. A Raman transition is used to provide a large photon recoil velocity while satisfying the metastability of the states necessary for long interaction times. A first $(\pi/2)$ pulse will prepare the sample in a superposition of the $|1, \mathbf{p}\rangle$ and the $|2, \mathbf{p} + \hbar \mathbf{k}\rangle$ states; the second (π) pulse will reverse the populations, and the third $(\pi/2)$ pulse will bring the wave packets to interference. The interference can be detected by probing the number of atoms in state $|2\rangle$. In the absence of any external forces acting on the atoms the final state of an atom will depend on the phase of the driving Raman field. This result can be extended to an atom falling freely in the gravitational field. In the frame of reference falling with the atom the Raman light fields appear Doppler shifted linearly in time, which shows up as a phase shift varying as the square of the time: $\Delta\phi = -(\mathbf{k}_1 - \mathbf{k}_2) g T^2$. Since the Doppler shifts $\approx 2k_i g T$ are much larger than the Rabi frequency, an active frequency shift between the three subsequent pulses was used to compensate for the deceleration of the atoms in the fountain.

Using a 50 ms delay in-between the pulses distinct interference fringes were observed, and a least square fit to the data gave an uncertainty in the phase determination of 3×10^{-3} cycles. This represented a sensitivity to g of $\Delta g/g = 3 \times 10^{-8}$. A higher sensitivity is expected to be obtainable when cesium is used instead of sodium because of large reduction of the rms velocity spread. The current work was done with $30\mu\text{K}$ sodium atoms (representing a 30cm/s velocity spread). For the cesium sample one may expect an rms spread of 2cm/s only, therefore a much larger portion of the sample will be contributing to the fringes.

Hydrogen (and certainly antihydrogen) is notoriously unsuited for high precision measurements, notwithstanding the enormous advances in hydrogen spectroscopy over the last years²⁶⁾, and a translation of this method to the hydrogen/antihydrogen case will not be trivial and straight forward. A large problem will be imposed by the much higher photon recoil limit for laser cooling hydrogen atoms ($\approx 3\text{ mK}$) which gives an rms velocity spread of approximately 700 cm/sec . A much faster fountain beam, resulting in greatly increased experimental dimensions, will have to be used. Therefore a much larger fraction of the initial beam pulse will be lost due to ballistic spreading during the flight time of the sample and much less than 1% of the initial population can be expected to contribute to the fringes. This will also cause severe problems for the antihydrogen/hydrogen comparison, since the supply of antihydrogen atoms will be limited to small numbers and a re-trapping scheme needs to be incorporated into the experiment. Nevertheless, this method is the only one identifiable in the current literature which shows the potential of a high precision measurement of g on antihydrogen atoms.

3. CONCLUSIONS

A direct measurement of the gravitational acceleration of antimatter in the Earth's gravitational field is an important test of fundamental principles and may lead to new insight in the physics at the Planck mass scale. No entirely model independent conclusions can be drawn from normal mass experiments, despite the fact that these experiments have achieved an extreme high precision, and therefore a direct measurement should be performed. Currently a proposal to use low energy, charged, antiprotons in a TOF method exists and the experiment is under preparation. An

alternative approach using charged particles, exploiting the gxB drift in a crossed magnetic and gravitational field has been discussed, and potential advantages in terms of a lower sensitivity to stray electric fields have been found.

If the methods of atom interferometry, which have been highly developed for sodium and cesium atoms, can be adapted to the hydrogen/antihydrogen case and the measurement cycle can be made to accommodate the much smaller number of antihydrogen atoms available, a potentially very powerful method would become available to achieve a much higher precision on g than anticipated in the charged particle experiments. But before this could be attempted, the important step of producing and controlling the antihydrogen atom needs to be solved first. In the light of the difficulty of this endeavor and the limited life time of current low energy antiproton sources it may be a wise choice to push ahead with the charged particle experiments to obtain at least a preliminary result (at the level of a few tenth of a percent) while continuing development of the much more complicated antihydrogen system.

ACKNOWLEDGMENTS

The support of all members of the PS200 collaboration in the work presented here is acknowledged. Current members include R. E. Brown, T. W. Darling, P. L. Dyer, T. Goldman, D. Hajdukovic, M. H. Holzscheiter, K. Hosea, R. A. Kenefick, N. S. P. King, R. A. Lewis, M. M. Nieto, T. Otto, S. P. Parry, R. Ristinen, J. Rochet, M. M. Schauer, G. A. Smith, and F. C. Witteborn. On the issues of antihydrogen production I enjoyed numerous enlightening discussions with B. Deutch, G. Laricchia, and M. Charlton and other members of the UCL - Aarhus antihydrogen group. I am also indebted to V. Lagomarsino, G. Manuzio, and G. Testera for pointing out to me the distinct advantages of the gxB method over the TOF measurement. This work was supported by the US Department of Energy under contract no. W7405 ENG-36.

REFERENCES:

1. N. Beverini et al., CERN Proposal PSCC/86-2/P94;
2. M. H. Holzschneider; Phys. Scripta Vol 46 (1992) 277
3. H. G. Dehmelt; in "Advances in Atomic and Molecular Physics" eds. D. R. Bates and I. Esterman, Vol. 3 (1967) 53
4. S. Rolston and G. Gabrielse; Hyp. Int. 44 (1988) 233
5. G. Gabrielse, L. Haarsma, and S. L. Rolston; Int. J. Mass Spec. Ion Proc. 88 (1989) 319
6. M. H. Holzschneider, P. L. Dyer, S. Hoibraten, N. S. P. King, R. A. Lewis, D. C. Lizon, G. L. Morgan, J. Rochet, M. M. Schauer, J. A. Schecker, T. Otto; to be published in the 'Proceedings of the Third Biennial Conference on Nucleon-Antinucleon Physics' (NAN'93) in Moscow, Russia, Sept. 1993
7. F. C. Witteborn and W. M. Fairbank; PRL 19, (1967) 1049
8. J. B. Camp, T. W. Darling, and R. E. Brown; J. Appl. Phys. 69 (1991) 7126
9. J. B. Camp, T. W. Darling, and R. E. Brown; J. Appl. Phys. 71 (1992) 783
10. M. H. Holzschneider; Nuc. Phys. A558 (1993)
11. D. Hajdukovic; Phys. Lett. B 226 (1989) 352
12. V. Lagomarsino, V. Lia, G. Manuzio, G. Testera; INFN Report BE 93/06
13. M. H. Holzschneider; Hyp. Int. 81 (1993) 71
14. G. Gabrielse, S. L. Rolston, L. Haarsma, and W. Kells; PRL A129 (1988) 38
15. B. I. Deutch et al.; Phys. Scripta T22 (1988) 288
16. J. W. Humberston, M. Charlton, F. M. Jacobsen, and B. I. Deutch; J. Phys. B20 (1987) L25
17. B. I. Deutch, L. H. Andersen, P. Hvelplund, F. M. Jacobsen, H. Knudsen, M. H. Holzschneider, M. Charlton, and G. Laricchia; Hyp. Int. 44 (1988) 271
18. M. Charlton; Phys. Lett. A143 (1990)
19. B. I. Deutch, M. Charlton, M. H. Holzschneider, P. Hvelplund, L. V. Jorgensen, H. Knudsen, G. Laricchia, J. P. Merrison, and M. R. Poulsen; Hyp. Int. 76 (1993) 153
20. J. W. Humberston, M. Charlton, F. M. Jacobsen, and B. I. Deutch; J. Phys. B20, L25 (1987)
21. A. M. Ermolaev, B. H. Brandsen, and C. R. Mandel; Phys. Lett. A25, 44 (1987)
22. A. M. Ermolaev; Hyperfine Int. 44, 375 (1988)
23. M. Charlton; Phys. Lett. A143, 143 (1990)
24. M. Kasevich and S. Chu; Appl. Phys. B54 (1992) 321
25. T. W. Haensch and C. Zimmermann; Hyp. Int. 76 (1993) 47
26. G. Gabrielse, X. Fei, L. A. Orozco, R. L. Tjoelker, J. Haas, H. Kalinowsky, T. A. Trainor and W. Kells; Phys. Rev. Lett. 65 (1990) 1317

TESTING THE DISCRETENESS OF SPACETIME AT LOW ENERGIES

G. Chardin

*CEA, DAPNIA, Service de Physique des Particules
Centre d'Etudes de Saclay, F-91191 Gif-sur-Yvette Cedex, France***Abstract**

Expected at the Planck scale, the discreteness of spacetime can hardly be tested at such high energies in a foreseeable future. However, the Bekenstein relation suggests another approach, based on the study of time-asymmetry. After a brief introduction to discrete physics, we investigate the relation between gravitation and discreteness on the one hand, and between gravitation and time-asymmetry on the other. These relations have led to reconsider the possibility that CP violation in the neutral kaon system, the only known microscopic "arrow of time", could be attributed to gravitation. We summarize the arguments suggesting that "antigravity" could occur for antimatter and could even be expected in General Relativity itself. Experimental tests are proposed.

INTRODUCTION

The idea that Nature might be discrete rather than continuous may appear strange at first sight, although this hypothesis has recently received renewed attention¹⁻³). What are the motivations for this new atomism ? As we show in the following, this idea arises primarily, on the one hand, from the link between gravitation and time-asymmetry and, on the other hand, from the study of discrete logical systems — computers — and the logical and physical reversibility of calculation. This review provides a short introduction to the subject and literature of discrete physics to the reader who might be interested to pursue further the study of this fascinating subject, and proposes experimental tests, hopefully more realistic than a galactic collider at the Planck scale (10^{19} GeV), that may differentiate a discrete Universe from a continuous one.

DISCRETE PHYSICS

The idea that Nature is fundamentally discrete rather than continuous can be traced back to the Greek philosophers and their atoms — $\alpha\tau\omicron\mu\omicron\sigma$: indivisible — that we improperly identify with our atoms which do have a structure. More recently, the idea that the Universe could be discrete or even finite has appeared repeatedly : von Weizsäcker⁴), in his “urs” theory — basically bits of information — conjectured that the total number of “urs” in the Universe was of the order of 10^{120} . Kantor⁵) made a similar estimate (10^{122}) of the quantity of information that the Universe could encompass and found about the same number. Penrose⁶) also obtained the same number when trying to estimate the improbability of the presently observable Universe. Basically, it appears that, apart from philosophical prejudices, two main approaches towards discreteness have been identified : the first one arises from the study of the ultimate limits of dissipation in computers, while the second is based on the Bekenstein mass/entropy relation^{7,8}).

Similarly to the historical development of the Second Law of Thermodynamics, where Sadi Carnot discovered the concept of entropy by considering the ultimate limitations on the efficiency of heat engines, Rolf Landauer investigated at IBM, starting from the late fifties, the question of dissipation in computation. The most coveted patents for high-end computers were related to the cooling of hardware processors and the question was whether there existed a fundamental limit to the amount of dissipation in a computation. In other words, is the process of computation necessarily irreversible or can it be made reversible ? Landauer's principle⁹) represents the fact that irreversibility in computation can be traced back to the erasure of a register regardless of its previous state. Building on the work of Landauer, Bennet¹⁰) showed that the process of computation could be made reversible provided that intermediate results were not destroyed (keeping the “garbage bits”) and that no registers were

erased. The increase in the computing power has made possible numerical simulations which can adequately approximate the physical world and the study of discrete systems like cellular automata¹¹⁾ where some of the properties of Nature could be recognized ("speed of light", "particles", ...), the best known example of cellular automata being the Game of Life¹²⁾, devised by the mathematician John Conway.

Whereas the discretization was mainly used initially as a calculation technique, a number of major physicists, notably Feynman¹³⁾, Finkelstein¹⁴⁻¹⁶⁾, t'Hooft^{17,18)}, Landauer¹⁹⁾, Lee²⁰⁾, Penrose²¹⁾, Regge²²⁾ and Wheeler²³⁾, gradually considered the possibility that spacetime could be discrete. Reversing the problem, and following Laplace, Descartes and Maxwell who had represented the Universe as a vast machinery, the question then became : "Could it be that the Universe is a discrete system ?" or even a computer²⁴⁾. A Turing machine, the generic computer, can basically realize *any* algorithmic calculation and may be programmed to calculate, e.g., the evolution of a 3-dimensional system under the Schrödinger equation (in a discrete approximation).

Another question concerns the computation limits of the Universe. The observable Universe today contains a number of nucleons of the order of 10^{80} , give or take an order of magnitude. Is it then meaningful, asked Landauer, to speak about a real number like π if there is absolutely *no way* to calculate and print the first 10^{100} first digits of π , or to calculate its 10^{1000} th digit ? The use of real numbers and of the continuum then appears as an approximation if they cannot be given any physical implementation¹.

A related question, addressed by Borges in "The Babel library"²⁸⁾, concerns the storage capabilities of the Universe : what is the number of books that we could store in the Universe, or more precisely the maximum amount of bits, should we decide to devote all the physical resources of the Universe to this task ? Landauer argues that there is no limit to the amount of information that can be stored in a given volume with a given energy budget. However, there appears to exist some limits, studied by Bekenstein²⁹⁾ and by Unruh and Wald³⁰⁾ notably. To provide a more quantitative answer to the question of information content in the physical Universe, we now turn to the relation between gravitation and the Second Law of thermodynamics.

DISCRETENESS AND GRAVITATION

Planck noted that a relativistic quantum theory of gravitation would introduce a characteristic length scale

¹Weyl²⁵⁾ has warned us against the danger of considering irrational numbers and the continuum as more than a convenient approximation, whereas Gödel²⁶⁾ and Chaitin²⁷⁾, most notably, have shown that the natural numbers themselves contain their own logical doom and even complete randomness.

$$l_{\text{Planck}} \approx \sqrt{\hbar G/c^3} \approx 10^{-33} \text{ cm}$$

Wheeler insisted on the fact that the four dimensional character of spacetime will probably break down at the Planck scale and that we should only consider the continuum as an approximation. Numerically, the Planck scale corresponds to an energy of $\approx 10^{19}$ GeV. Even with acceleration performances of 1 GeV/m in a cavity — which we are far from achieving on macroscopic distances — a collider at the Planck scale would require an accelerator $\approx 10^{19}$ m long, about 1000 light-years, quite unlikely to be funded soon, even in Texas. In fact, at the present age of the Universe, collisions on the cosmic microwave background radiation at 2.7 K makes it an almost impossible feat to accelerate charged particles above $\approx 10^{20}$ eV or 10^{11} GeV. Does it mean that the Planck scale has no observable consequence in our low-energy world ?

In the early seventies, Bekenstein^{7,8)}, noting the striking analogy between the laws of black hole dynamics and the laws of thermodynamics, conjectured a relation between the area of a black hole and its entropy :

$$S = M^2,$$

in natural (Planck scale) units. But the notion of entropy, for a black hole of a given mass (energy), implied the existence of a non zero temperature, through the usual thermodynamic relation, a possibility which was severely criticized by Carter and Hawking, among others. Ironically, two years later, Hawking refined the Bekenstein formula and indeed proved the existence of the so-called “Hawking radiation” of black holes³¹⁾.

The physical significance of the Bekenstein entropy was further refined by Zurek³²⁾ and by Zurek and Thorne³³⁾ who showed that this entropy was fundamentally the physical entropy or, almost equivalently, the number of emitted photons of the radiation emitted by the black hole during its evaporation, providing the consistency of the Second Law of Thermodynamics. Coupled with the finding that a photon in a given energy level cannot encode more than one bit in its polarization state³⁴⁾, this relation effectively provides the information content in bits of the emitted radiation during the whole evaporation process (up to factors of order unity neglected here).

Wheeler³⁾ has stressed the simple geometrical interpretation of the Bekenstein formula : a black hole is a “hidden” region of spacetime beyond an horizon ; simply measure its boundary in natural units, or units of (Planck length)² and get its entropy (up to a factor 4π). This picture strongly suggests that there is some discreteness of spacetime at the Planck scale and that the information content of the underlying discrete structure is “revealed” in its evaporation.

Applying the Bekenstein formula to the observable Universe, we get an entropy

$$\approx (10^{80}/10^{19})^2 \approx 10^{122}.$$

This is an upper bound on the amount of information which can be stored in the Universe at the present time. If there exists no limiting factors, this quantity varies with time as t^2 and is essentially zero for $t = t_{\text{Planck}}$.

The lesson from this can then be summarized as follows : gravitation is the ultimate machinery to transform matter into photons ; stars begin the job and black holes finish it (and thermal photons represent the most extreme form of disorder).

DISCRETENESS AND CP-VIOLATION

Before the discovery of CP-violation in 1964, nobody would have bet a dime on a local failure of CP-invariance. The geometrical picture of an antiparticle as a particle going backward in time was so strong that when CP-violation was discovered³⁵⁾, Bell and Perring³⁶⁾ as well as Bernstein, Cabbibo and Lee³⁷⁾ proposed that CP-invariance could be restored by invoking a long-range interaction which would reveal the predominance of matter over antimatter in our environment. The vector interaction that these authors proposed is now experimentally clearly excluded.

In every theory where, following the expression of Wheeler, "everything is geometry", including the particles which could be pictured as knots in spacetime³⁸⁾, perfect local CP-symmetry is expected. The pregeometric models considered by Wheeler³⁸⁾ and others may then be inadequate to represent our Universe where we *do* observe CP-violation in the neutral kaon system. Could it be that "CP"-violation in the neutral kaon system is really due to gravitation, the only long-range interaction for which there is a notable matter/antimatter asymmetry in the environment ?

Some authors refer to the "strange idea that antimatter might be gravitationally repelled by matter"³⁹⁾. Bondi⁴⁰⁾ was the first to show that a solution with repulsive gravity — the so-called negative mass solutions — is perfectly compatible, if unstable, with General Relativity. But a more complete picture was given by Carter⁴¹⁾ who showed that the maximal extension of the Kerr and Schwarzschild geometries both incorporate attractive and repulsive gravity in the *same* solution. A test body initially at rest at the distance r of a center of attraction of mass M will "fall down" with an acceleration g , using the Newtonian language, or "fall up" with an acceleration $-g$, depending on which part of the solution the test particle is assigned. What is then the physical significance of this repulsive gravity that nobody has ever experienced ? Again, this repulsive part of the solution is expected to be unstable since it is interpreted locally as negative mass, but any undergraduate student, if not taught otherwise, would guess that this negative energy solution might bear some relation with antimatter and the "non physical" and unstable negative energy solutions of the Dirac equation.

The next surprise comes from the fact that antigravity would give just the right order of anomalous regeneration in the neutral kaon system that we have attributed to CP-violation⁴²⁻⁴⁴. It suggests a simple expression for the ϵ parameter

$$\epsilon = O(1) \hbar m_K g / \Delta m^2 c^3 = O(1) \times 0.88 \cdot 10^{-3}$$

predicts that the ϵ' parameter is zero, and that CP-violation in the B system should appear unobservably small.

The repulsive part of the Kerr or Schwarzschild solutions is usually ignored, considered as non physical, or attributed to "another universe". But could we find a solution to the coexistence of the two attractive and repulsive metrics by paying the price of a (hopefully) limited instability of the solution ? A first order estimate of the evaporation of the vacuum which would result from this kind of antigravity gives an evaporation "temperature" of

$$\hbar g/c$$

or $\approx 10^{-18}$ K at the surface of the Earth, inobservably small^{42,43}. This expression for the typical energy of the photons evaporated by a massive structure is the same as the Hawking temperature of a black hole. It is therefore far from obvious that this (limited) instability is a drawback. Nieto and Goldman⁴⁵ have reviewed critically the arguments against antigravity, although they propose a more conventional solution.

DISCRETENESS AND THE CARDINAL OF THE UNIVERSE

Finally, there exists a completely different approach of discreteness in Physics. We have seen that the information content of today's Universe is of the order of 10^{122} bits, if we apply to the observable Universe the Bekenstein relation between mass and entropy. Although this information content increases classically with time as t^2 , there are some quantum limitations (once all matter has been reprocessed in photons through the Hawking process, there is nothing which remains to be dissipated and therefore to be seen) which makes it probable that the information content of the entire Universe is really finite. It is therefore tempting to look at the possibility to build such huge numbers from the fundamental constants — another face of the "Eddington-Dirac large numbers" puzzle⁴⁶). Another possibility is to look for the various mathematical structures which possess such huge cardinal numbers : a few possibilities come to the mind, from iterated exponentiation⁴⁷) to large exceptional groups — the "Monster" group has a number of elements of the order of 10^{50} elements — and to the calculus of propositions³⁸) where such large numbers can be obtained from the theorems which can be built from a (small) set of axioms. Clearly, all these ideas are very vague — Wheeler calls the idea of pregeometry from the calculus of propositions "an idea for an idea"³⁸) — but it should be kept in mind that if it turns out that the very notion of the cardinal number of the Universe makes sense,

we will probably gain something from the knowledge of which particular mathematical structure may “generate” it.

If it turned out, for example, that the cardinal number of the Universe was in some way related to the number of elements of some exceptional group, then there is clearly only a very limited number of alternatives for the total information content of the Universe, whereas the “local” properties of the group could tell us something about, e.g., the mass hierarchy problem. Obviously, however, we will never be able to completely explore the particular mathematical structure which represents the Universe, a fundamental physical limitation in our mathematical capabilities.

CONCLUSIONS

In this paper, we have tried to provide a short introduction to the fascinating relation between discreteness and physics. The main two motivations for this hypothesis appear to be both related to time-asymmetry : in gravitation — through the Bekenstein relation and its geometric interpretation — and in the reversibility of discrete logical systems such as computers. Since the Planck scale of energy, where presumably this discreteness would become apparent, appears unreachable in a foreseeable future, we turned our attention to the dual aspect of time-asymmetry and reconsidered the possibility that CP-violation in the neutral kaon system, a feature which appears extremely difficult to incorporate in a pregeometric description of spacetime, and the only known example of the microscopic “arrow of time”, could be in fact explained by gravitation. Although we are unable to propose a complete solution, the attractive and repulsive gravity both present in even the simplest (e.g. Schwarzschild and Kerr) solutions of General Relativity itself suggests a simple expression for the ϵ -parameter of CP-violation in the neutral kaon system, allowing several predictions, notably on the ϵ'/ϵ parameter and on the CP-violation parameters in the B system. Direct measurements on the antiproton, and later on the antihydrogen, two experiments in preparation at LEAR (CERN), also appear fundamental, despite the present theoretical prejudices. The precision measurements of the CP-violating parameters under new experimental conditions using low-energy and isotropic kaon beams⁴⁸⁾ are yet other crucial tests of this idea.

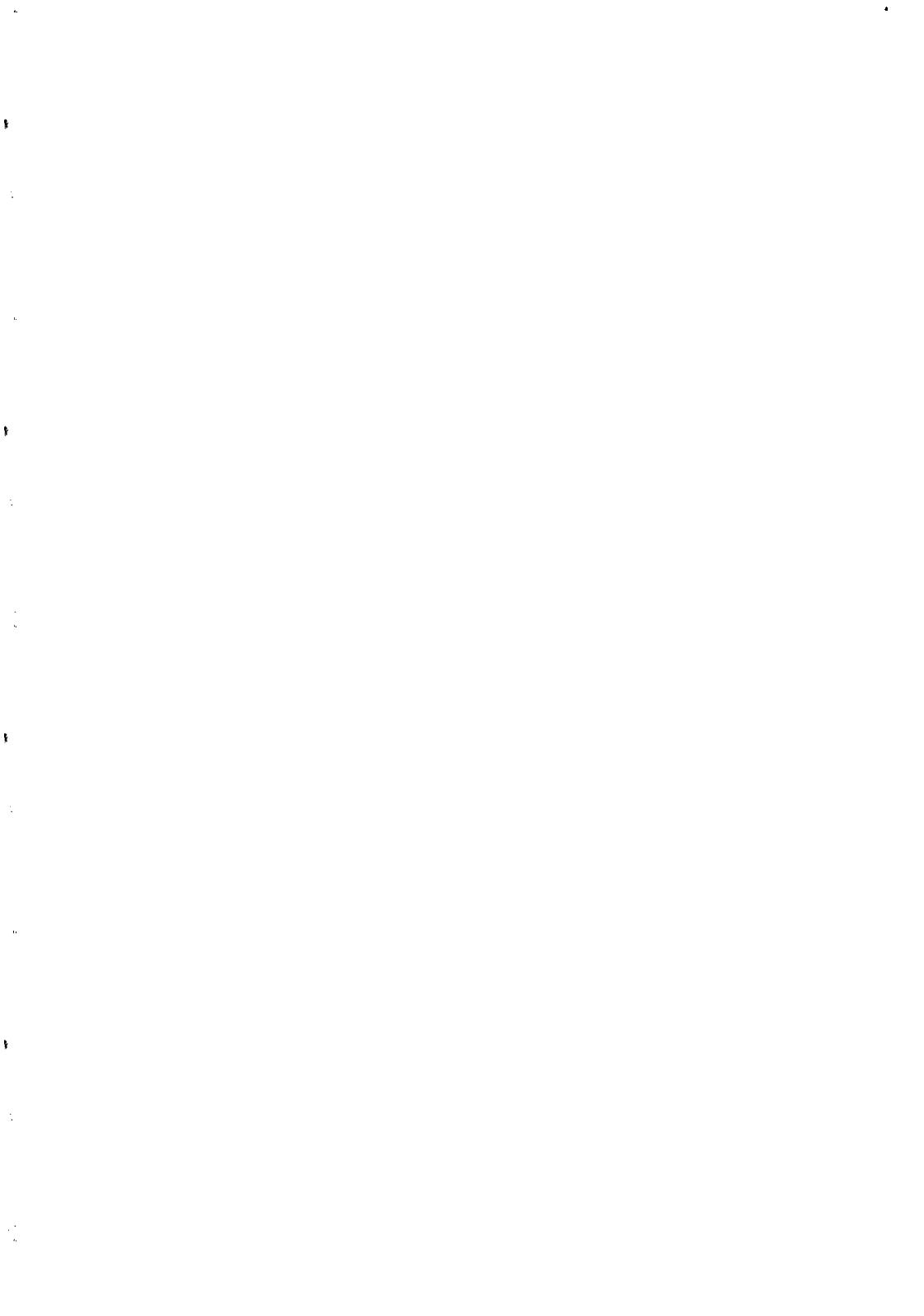
Acknowledgments : Discussions with G. Chaitin, E. Fredkin, R. Landauer, J-M. Rax and J. Rich are gratefully acknowledged.

REFERENCES

- [1] Int. J. Theor. Phys. 21 (1982) vol. 6-7, articles by Bennett, Feynman, Finkelstein, Fredkin and Toffoli, Kantor, Landauer, Minsky, Wheeler therein.
- [2] Found. Phys. 16 (1986) vol. 6, articles by Bennett, Deutsch, Feynman, Geroch and Hartle, Landauer therein.
- [3] J.A. Wheeler, in *Complexity, Entropy and the Physics of information*, ed. W.H. Zurek (Addison Wesley, Redwood City, 1990) and references therein.

- [4] C.F. von Weizsäcker, in *Quantum Theory and Beyond*, ed. E. Bastin (Cambridge Univ. Press, Cambridge, 1971)
- [5] F.W. Kantor, *Int. J. Theor. Phys.* 21 (1982) 525.
- [6] R. Penrose, in: *General Relativity, an Einstein Centenary Survey*, eds. S.W. Hawking and W. Israel, (Cambridge University Press, Cambridge, 1979) p. 281.
- [7] J.D. Bekenstein, *Nuovo Cimento Lett.* 4 (1972) 737.
- [8] J.D. Bekenstein, *Phys. Rev. D* 9 (1973) 3292; *Physics Today* 33 (1980) 24.
- [9] R. Landauer, *IBM J. Res. Dev.* 5 (1961) 183.
- [10] C.H. Bennett, *IBM J. Res. Dev.* 17 (1973) 525.
- [11] see e.g. S. Wolfram, *Rev. Mod. Phys.* 55 (1983) 601, and references therein.
- [12] W. Poundstone, *The recursive universe* (Oxford Univ. Press, Oxford, 1985)
- [13] R.P. Feynman, *Found. Phys.* 16 (1986) 507.
- [14] D. Finkelstein, *Phys. Rev.* 184 (1969) 1261 ; *Phys. Rev. D*5 (1972) 320 ; *Phys. Rev. D*5 (1972) 2922 ; *Phys. Rev. D*9 (1974) 2219.
- [15] D. Finkelstein, G. Frye and L. Susskind, *Phys. Rev. D*9 (1974) 2231.
- [16] D. Finkelstein, *Int. Journal of Theor. Phys.* 21 (1982) 489.
- [17] G. 'tHooft, in *Recent developments in gravitation*, Cargèse 1978, eds. M. Levy and S. Deser (Plenum Press, New York, 1978).
- [18] G. 'tHooft, *Nucl. Phys.* B342 (1990) 471.
- [19] R. Landauer, , *Found. Phys.* 16 (1986) 551.
- [20] T.D. Lee, *Phys. Lett.* B122 (1983) 217 ; R. Friedberg and T.D. Lee, *Nucl. Phys.* B225 (1983) 1.
- [21] R. Penrose, in *Quantum Theory and Beyond*, ed. E. Bastin (Cambridge Univ. Press, Cambridge, 1971)
- [22] T. Regge, *Nuovo Cimento* 19 (1961) 558.
- [23] J.A. Wheeler, *Int. J. Theor. Phys.* 21 (1982) 557.
- [24] E. Fredkin and T. Toffoli, *Int. J. Theor. Phys.* 21 (1982) 219.
- [25] H. Weyl, *Amer. Math. Monthly* 53 (1946) 2.
- [26] K. Gödel, *Monatshefte für Mathematik und Physik* 38 (1931) 173.
- [27] G. Chaitin, *Information, Randomness and Incompleteness*, 2nd Edition (World Scientific, Singapore, 1990)
- [28] J.L. Borges, "The Babel library", in *Ficciones* (Emecé, Buenos Aires, 1956).
- [29] J.D. Bekenstein, *Phys. Rev. D* 23 (1981) 287.
- [30] W.G. Unruh and R.M. Wald, *Phys. Rev. D* 26 (1982) 947.
- [31] S.W. Hawking, *Nature* 248 (1974) 30; *Comm. Math. Phys.* 43 (1975) 199; *Phys. Rev. D* 13 (1976) 191.
- [32] W.H. Zurek, *Phys. Rev. Lett.* 49 (1982) 1683.
- [33] W.H. Zurek and K.S. Thorne, *Phys. Rev. Lett.* 54 (1985) 2171.
- [34] W. Wothers and W.H. Zurek, *Nature* 304 (1983) 188.
- [35] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Phys. Rev. Lett.* 13, 138 (1964).
- [36] J. S. Bell and J. K. Perring, *Phys. Rev. Lett.* 13 (1964) 348
- [37] J. Bernstein, N. Cabibbo and T. D. Lee, *Phys. Lett.* 12 (1964) 146
- [38] see e.g. C.W. Misner, K.S. Thorne and J.A. Wheeler, in "Gravitation", pp. 1203-ff and references therein.
- [39] R. Hughes, *Contemp. Phys.* 34 (1993) 177.
- [40] H. Bondi, *Rev. Mod. Phys.* 29 (1957) 423.
- [41] B. Carter, *Phys. Rev.* D141 (1966) 1242 ; D174 (1968) 1559.
- [42] G. Chardin and J.M. Rax, (1992) *Phys. Lett.B* 282 (1992) 256.
- [43] G. Chardin, *Nucl. Phys. A* 558 (1993) 477c.
- [44] T. Goldman, M.M. Nieto and V.D. Sandberg, *Mod. Phys. Lett. A* 37 (1992) 3455.
- [45] M.M. Nieto and T. Goldman, *Phys. Rep.* 205 (1991) 221.
- [46] P.A.M. Dirac, *Nature* 139 (1937) 323 ; A.S. Eddington, *The Nature of the Physical World* (Cambridge Univ. Press, London, 1928).
- [47] E. Bastin, in *Quantum Theory and Beyond*, ed. E. Bastin (Cambridge Univ. Press, Cambridge, 1971).
- [48] R. Adler et al., *Phys. Lett. B* 286 (1992) 180.

SUMMARY



High, Very High, and Ultrahigh Energy Astrophysical Phenomena
A Summary of the Sessions

Hinrich Meyer
University of Wuppertal
D-42097 Wuppertal
GERMANY

The study of high energy phenomena in the cosmos is based on observations of x and γ rays in space and cosmic rays, TeV gammas and, hopefully, in the near future of neutrinos in experiments on earth. As of this moment the satellites in space are the winners of the game with beautiful new results from reports on results for Sigma, CGRO with Batse and Egret and also ROSAT although not directly present at this meeting. The earth bound experiments, WHIPPLE, THEMISTOCLE, HEGRA, DUMAND, AMANDA, NESTOR, and BAIKAL are on their way to catch up at higher energies (E. Paré, G. Fontaine, H. Meyer, P. Sokolsky, H. Dai, S. Katsanevas).

And now what's the matter? The processes looked at have large amount of matter (≥ 1 solar mass or more) falling very deep releasing high amount of energy that by means not too well understood yet is converted into very high energies of individual elementary particles. Supernovae are very prominent along this line leaving (supposedly) pulsars behind, neutron stars of ≈ 1.4 solar mass with 10^{12} G magnetic fields spinning very rapidly $>$ and \gg than 1 Hz. And one of them, Geminga, was long time observed only as a GeV γ ray star, and now very recently was proven to be a pulsar on the basis of a very feeble x ray flux by ROSAT (P. Ramanamurthy). Even weaker is the optical output (≈ 25 mag), however, the proper motion of it shows that GEMINGA is close and so was the Supernova at birth only ≈ 100.000 years ago at Neandertaler times (the Neandertal is only few km from Wuppertal university, for those who don't know). Furthermore, Egret observations of pulsars reveal on the basis of Crab, Vela, 1706-44, 1055-52 (and one more) that pulsars get more efficient with age (I wish it would happen for us

physicists too) in converting spin down energy into radiation at GeV, truly remarkable objects (I. Grenier). Beyond that it is supposed that Supernova remnants (SNR) are the source of all cosmic rays up to about 10^{16} eV in our galaxy and then unavoidably of cosmic rays in any galaxy. And indeed for the CRAB photons up to 17.000.000 MeV energy (Themistocle) have been observed to originate from the CRAB SNR. And what will be the next great observation surely to be expected in the near future in this fascinating field? (P. Duffy, P. Eschtruth, G. Fontaine, H. Meyer)

More mysterious, however, are the Gamma ray bursts (GBR) shown to be very isotropic on the sky. They never repeat (unfortunately) they coincide in position with nothing (or everything) and may happen everywhere in the universe, say with small probability in any ordinary galaxy. The most likely hypothesis, n-star-n-star mergers makes them unique and of course rather violent events. It is clear by now that this picture is hard to proof, many think that only a serendipitous observation will unveil the secret (M. Baring).

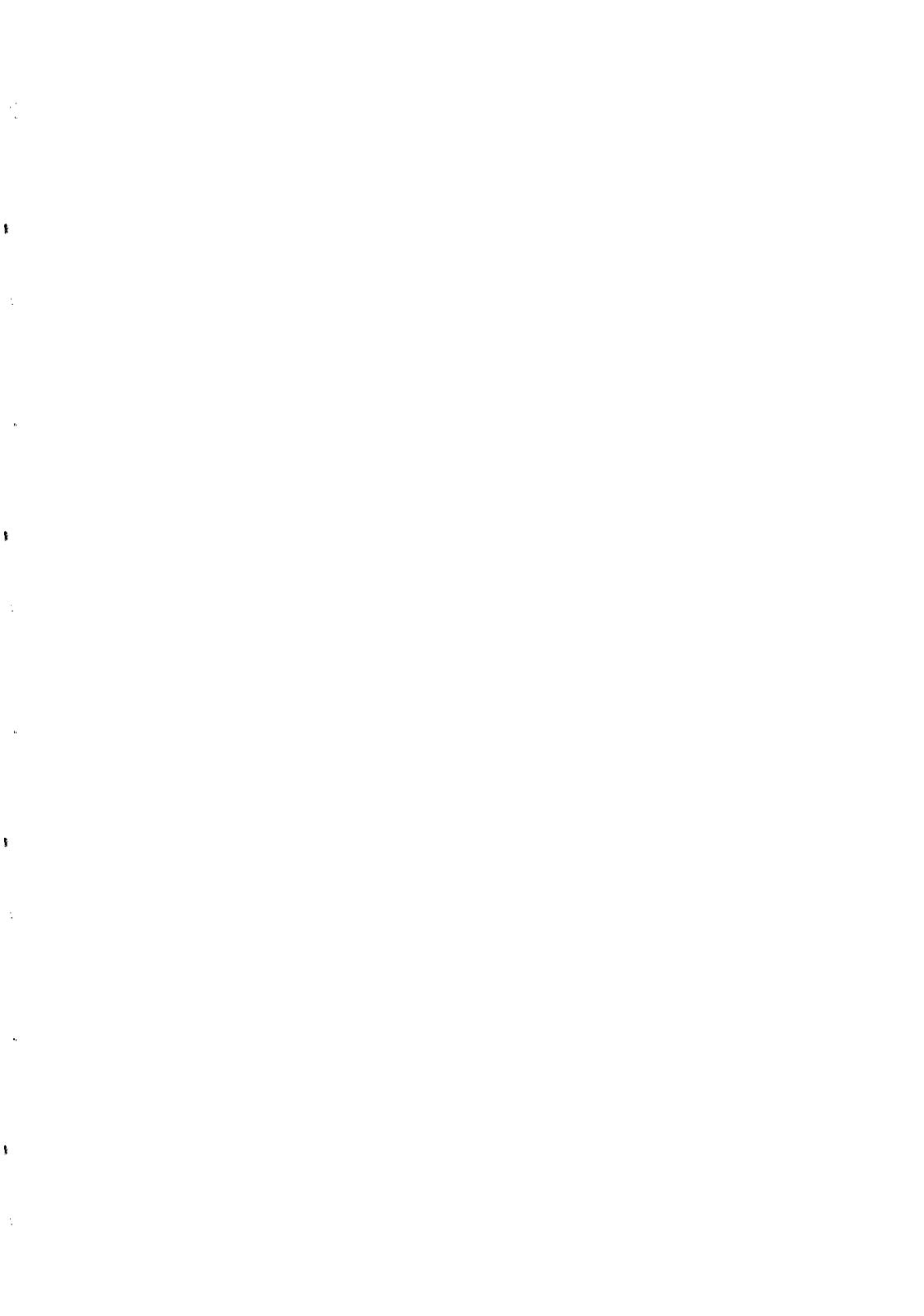
The deepest gravitational potential seems to be provided by black holes. They may exist - although there is no rigorous proof yet, not to forget (and contrary to what layman are sure about) - at vastly differing scales from 10 solar mass objects in binaries in our galaxy to 10^9 solar masses in the very center of active galactic nuclei (AGN) and then at very large distances from us ($z > 1$). X ray binaries very brilliant at times and then rather unspectacular again - they are repeating! - are prime candidates with black hole signature, however, the signature is necessary only and not sufficient, and nobody coming up with a full proof proposal, a challenge (J. Paul)! In AGNs the standard picture has two opposing jets emerging perpendicular to the rotating accretion disk of the black hole (C. Dermer). Shock wave acceleration pushes hadronic matter (protons) up to 10^{20} eV energy, but why not even further? It is the ambient photon field to stop them via violent energy crushing photoproduction of pions - an amusing picture for someone who has spent some years of his scientific life in measuring the rather unspectacular photoproduction of pions in the laboratory - and it is even conceivable that this "smashing the brickwall" of pion photoproduction with highest energy protons in AGNs has left over cosmological ages a very high energy ($\approx 10^{17}$ eV) neutrino gas in the universe, a truly cosmological signal, just below present experimental limits waiting to

be discovered by the next generation neutrino experiments, say Amanda at Antarctica? This same process leaves of course high energy photons (π^0 does it) that find their energy eating medium in the universal 3° K blackbody radiation background (Max Planck surely would have been amazed by that) until just below 100 TeV, where survivors would assemble and fill the universe.

In closing I like to remind you of a phrase in Gabriel Chardin's talk,

"Gravitation is a machine that converts energy into photons",

quite fitting for the sessions I tried to summarize.



LIST OF PARTICIPANTS

Kamal ABDULLAH
Harvard University
Physics Dept
Cambridge MA 02138
USA

Nicole BEL
Observatoire Meudon
5, Place Janssen
92190 Meudon
France

Eric ADELBERGER
Univ. of Washington
Nuclear Physics Lab. GL-10
Seattle WA 98195
USA

Pierluigi BELLI
INFN-Sezione di Roma 2
c/o Dipartimento di Fisica
Via Della Ricerca Scientifica
00133 Rome
Italy

Réza ANSARI
Univ. de Paris Sud
L A L
Bât. 200
91405 Orsay Cedex
France

David BENNETT
L L N L
L-413 IGPP
Livermore CA 94550
USA

Eric AUBOURG
CEN Saclay
DAPNIA - SPP
91191 Gif sur Yvette Cedex
France

Lars BERGSTROM
Stockholm University
Dept of Physics
Box 6730
11385 Stockholm
Sweden

Pierre BAREYRE
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Alain BLANCHARD
Observatoire Strasbourg
11 rue de l'Université
67000 Strasbourg
France

Matthew BARING
NASA
Goddard Space Flight Ctr
Lab. for H.E. Physics
Greenbelt MD 20771
USA

Roger BLANDFORD
CALTECH
130-33
Pasadena CA 91125
USA

Stephen BARR
Univ. of Delaware
Bartol Research Inst.
Newark DE 19716
USA

Anne BOUCHIAT
Ecole Normale Supérieure
Laboratoire de Physique
24 Rue Lhomond
75231 Paris Cedex 05
France

Saps BUCHMAN
Stanford Univ.
W.W. Hansen Lab. GP-B
Stanford CA 94305-4085
USA

Ralph CONTI
Univ. of Michigan
Physics Dept
Ann Arbor MI 48109
USA

Fabien CAVALIER
Univ. de Paris Sud
L.A.L.
Bât. 200
91405 Orsay Cedex
France

Adrien CORNAZ
Univ. of Zürich
Physics Institute
CH-8057 Zürich
Switzerland

Ludwik CELNIKIER
Observatoire Meudon
5, Place Janssen
92190 Meudon
France

Christian COUTURES
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Gabriel CHARDIN
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Hongyue DAI
Univ. of Utah
Physics Dept
Salt Lake City UT 084112
USA

Donghyun CHO
Univ. of Colorado
JILA
Boulder CO 80309-0440
USA

Thibault DAMOUR
IHES
35 Route de Chartres
91440 Bures sur Yvette
France

Timothy CHUPP
Univ. of Michigan
Randall Lab.
500 University Str.
Ann Arbor MI 48109
USA

Maryvonne DE JESUS
I. P. N. Lyon
43, bd de 11 Novembre 1918
69622 Villeurbanne
France

Ignazio CIUFOLINI
IFSI
CNR
Via G. Galilei C.P. 27
Frascati
Italy

David DE MILLE
Univ. of California
Physics Dept
Berkeley CA 94720
USA

Alvaro DE RUJULA
CERN
TH Division
CH-1211 Geneve 23
Switzerland

Jim FALLER
Univ. of Colorado
Campus Box 440
Boulder CO 80309
USA

Charles DERMER
Naval Research Lab.
Code 7653
Washington DC 20375-5352
USA

Ephraim FISHBACH
Purdue University
Physics Dpt 1396
W. Lafayette IN 47906-1396
USA

Peter DUFFY
Max Planck Inst/Kernphysik
Postfach 10 39 80
6917 Heidelberg
Germany

Ricardo FLORES
Univ. of Missouri
Physics Dept
St Louis MO 63121
USA

Paul ESCHSTRUTH
Univ. de Paris Sud
L A L
Bât. 200
91405 Orsay Cedex
France

Gérard FONTAINE
Ecole Polytechnique
I N2 P3
3, rue Michel Ange
75781 Paris Cedex 16
France

Carlos ESCOBAR
Univ. de Sao Paulo
Ist. de Fisica
Caixa Postal 20516
01498-970 Sao Paulo SP
Brazil

Georg FORSTER
Tech. Univ. München
Physik Department E15
James Franck Str.
D-85747 Garching
Germany

Francis EVERITT
Stanford Univ.
Gravity Probe B & STEP Progr.
HEPL
Stanford CA 94305-4085
USA

Yasunori FUJII
Nihon Fukushi University
Okuda Chita-Gun
Aichi 470-32
Japan

Orrin FACKLER
LLL Livermore
P.O.Box 808
L-403
Livermore CA 94550
USA

Denis GARRETA
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Ta GIRARD
Centro de Fisica Nucl.
av. Prof. Gama Pinto, 2
1699 Lisboa
Portugal

Yannick GIRAUD-HERAUD
Collège de France
Lab. Physique Corpusculaire
11, Pl. Marcelin Berthelot
75231 Paris Cedex 05
France

Fernand GRARD
Univ. Mons Hainaut
Av. Maistriau 19
B-7000 Mons
Belgium

Keith GREEN
Rutherford Appleton Lab.
Chilton
Didcot Oxon OX11 0QX
U.K.

Isabelle GRENIER
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex

Bruno GUIDERDONI
Inst. d'Astrophysique de Paris
98 Bis bd Arago
75014 Paris
France

Claude GUYOT
C E N Sackay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Tom HIJMANS
Univ. of Amsterdam
Van Del Waals Zeeman Lab.
Valckenierstraat 65
1018 X Amsterdam NL
Netherlands

Edward HINDS
Yale University
Physics Dept.
217 Prospect Street
New Haven CT 06511
USA

Michael HOLZSCHEITER
Los Alamos Nat. Lab.
MS K 480
Los Alamos NM 87545
USA

Eugen HOLZSCHUH
Univ. Zürich
Physik Institut
Schönberggasse 9,
CH-8001 Zurich
Switzerland

Bruno HUBLER
Univ. Zürich
Physik Institut
CH-8057 Zürich
Switzerland

Charles JOACHAIN
Univ. Libre de Bruxelles
Campus Plaine
Bd du Triomphe
B-1050 Bruxelles
Belgium

Lucile JULIEN
Lab. Spectr. Hertzienne ENS
Boite 74
4, Pl. Jussieu Tour 12 E1
75252 Paris Cedex 05
France

Marc KAMIONKOWSKI
School of Natural Sc.
Institute for Adv. Study
Princeton N.J. 08540
USA

Walter KUNDIG
Univ. of Zurich
Physics Inst.
8057 Zürich
Switzerland

Jean KAPLAN
Univ. de Paris VI et VII
Lab. Physique Théorique
2, Pl. Jussieu
75251 Paris Cedex
France

Steve LAMOREAUX
Univ. of Washington
Physics Dept
Seattle WA 98195
USA

Stavros KATSANEVAS
Univ. of Athens
Physics Lab.
104, Solonos Street
10680 Athens
Greece

Robert LANOU
Brown University
Dept of Physics
Providence R.I. 02912
USA

Maxim KHILOPOV
Center for Cosmoparticle Physics
Cosmion
Miusskaya Pl. 4
Moscow 125047
Russia

Bruno LAURENT
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Iosif KHRIPLOVICH
Inst. of Nuclear Physics
Novosibirsk 630090
Russia

Renaud LE GAC
CPPM
163, av. de Luminy
13288 Marseille Cedex 9
France

Jonny KLEINFELLER
Kernforschungszentrum
Karlsruhe
Postfach 3640
76021 Karlsruhe
Germany

Eric LESQUOY
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Andrzej KROLAK
Academy of Sciences
Inst. of Mathematics
Sniadeckich 8
00950 Warsaw
Poland

Adolph LU
Univ. of Santa Barbara
Inst. for Theoretical Physics
Santa Barbara CA 93106-4030
USA

Christophe MAGNEVILLE
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Bernhard MAIER
Max Planck Institut f.
Kernphysik
P.O.B. 103 980
69029 Heidelberg
Germany

G rard MAREL
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Sophie MAUROGORDATO
Observatoire de Meudon
5, Place Janssen
92195 Meudon
France

Yannick MELLIER
Observatoire de Midi-Pyr n es
14 av. Edouard Belin
31400 Toulouse
France

Hinrich MEYER
B U G H Wuppertal
FB 8
GaussStr. 20
42119 Wuppertal
Germany

Stephan MEYER
Enrico Fermi Inst.
Univ. of Chicago
5640 Ellis Avenue
Chicago IL. 60637
USA

Giulio MIGNOLA
Univ. di Torino
Dipart. di Fisica Teorica
Via P. Giuria, 1
10125 Torino
Italy

Alain MILSZTAJN
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

John MOFFAT
University of Toronto
Dept of Physics
Toronto Ont M5S 1A7
Canada

Marc MONIEZ
Univ. de Paris Sud
L A L
B t. 200
91405 Orsay Cedex
France

Luigi MOSCA
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Anna NOBILI
Univ. di Pisa
Diprt. di Matematica
Via Buonarroti 2
I-56127 Pisa
Italy

Hiroshi NUNOKAWA
K E K
Theory Group
1-1 Oho
Tsukuba-shi 305
Japan

Eric PARE
Ecole Polytechnique
LPNHE
91128 Palaiseau Cedex
France

Jacques PAUL
C E N Saclay
Service d'Astrophysique
91191 Gif sur Yvette Cedex
France

A.M. PENDRILL
Chalmers Univ. of Tech.
Dpt of Physics
41296 Göteborg
Sweden

M. C. PERILLO ISAAC
Collège de France
Lab. Physique Corpusculaire
11, Pl. Marcelin Berthelot
76231 Paris Cedex 05
France

Daniel PFENNIGER
Univ. of Geneva
Geneva Observatory
CH-1290 Sauverney
Switzerland

Marc PINSONNEAULT
Ohio State Univ.
174, W. 18th av.
Columbus OH 43210
USA

François QUEINNEC
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Poolla RAMANA MURTHY
N A S A
Goddard Space Flight Center
Code 662
Greenbelt MD 20771
USA

Norman RAMSEY
Harvard Univ.
Lyman Physics Lab.
Cambridge MA 02138
USA

Cécile RENAULT
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Jim RICH
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Rogers RITTER
Univ. of Virginia
Dept of Physics
Charlottesville VA 22901
USA

Stephen ROSS
Max Planck Inst. f.
Quantumoptics
D-85748 Garching
Germany

Remo RUFFINI
Univ. of Rome
Dept of Physics
Pl. Aldo Moro 5
00185 Rome
Italy

Jürgen SCHURR
Univ. Wuppertal
Fachbereich Physik
GaussStr. 20
42097 Wuppertal
Germany

Stefano SCOPEL
Univ. di Torino
Dipart. di Fisica Teorica
Via P. Giuria 1,
10125 Torino
Italy

Isaac SILVERA
Harvard Univ.
Dept of Physics
Cambridge MA 02138
USA

Gregory SMITH
Univ. of Washington
Nuclear Physics Lab.
GL-10
Seattle WA 98185
USA

Pierre SOKOLSKY
Univ. of Utah
Physics Dept
Salt Lake City UT 084112
USA

Michel SPIRO
C E N Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex
France

Christopher STUBBS
Univ. of California
Santa Barbara CA 93106-9530
USA

Joseph SUCHER
Univ. of Maryland
Dept of Physics
College Park MD 20742
USA

Timothy J. SUMNER
Imperial College
Physics Dept
Prince Consort Rd
London SW72BZ
U.K.

Michal SZYMANSKI
Warsaw Univ. Observatory
Al. Vjazdowskie 4
00-478 Warsaw
Poland

Catherine THIBAUT
C S N S M
Bât. 108
91405 Orsay Campus
France

Jean TRAN THANH VAN
Univ. de Paris Sud
L P T H E
Bât. 211
91405 Orsay Cedex
France

Sylvie VAUCLAIR
Obs. de Midi-Pyrénées
14 av. Edouard Belin
31400 Toulouse
France

Robert VESSOT
Smithsonian Astro. Obs.
60, Garden Street
Cambridge MA 02138
USA

Jirka VRANA
Ecole Polytechnique
LPNHE
91120 Palaiseau
France

Harald WALESCH
Univ. Wuppertal
GaussStr. 20
42097 Wuppertal
Germany

Antoine WEIS
Max Planck Inst.
für Quantum Optics
85748 Garching
Germany

Claus ZIMMERMANN
Max Planck Inst.
Postfach 1513
85740 Garching
Germany

