(2)

AD-A233 001

NPS-PH-91-005

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THE MEANING OF TIME
AN INTRODUCTION INTO PHILOSOPHICAL, BIOLOGICAL
AND PHYSICAL ASPECTS OF TIME

KARLHEINZ E. WOEHLER

JANUARY 1991

Technical Report

Approved for public release; distribution is unlimited.

Prepared for:
Naval Postgraduate School
Monterey, CA 93943

Naval Postgraduate School
Monterey, California

Rear Admiral R. W. West, Jr.
Superintendent

H. Shull
Provost

The work reported herein was not funded by the Naval Postgraduate School.

This report was prepared by:

KARLHEINZ E. WOEHLER
Professor of Physics

Released by:

PAUL J. MARTO
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPS-PH-91-005	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL (If applicable) PH	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Naval Postgraduate School Physics Department (Code PH) Monterey, CA 93943-5000	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION unfunded 8b. OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) THE MEANING OF TIME
AN INTRODUCTION INTO PHILOSOPHICAL, BIOLOGICAL AND PHYSICAL ASPECTS OF TIME

12. PERSONAL AUTHOR(S)
Karlheinz E. Woehler

13a. TYPE OF REPORT	13b TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM Jan 90 TO Dec 90	910115	312

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This report presents an introduction into philosophical, biological and physical aspects of time. Time in man's basic experience, symbolizations of time, the western view of historical time and the evolution of the concept of time in philosophy are outlined. A brief introduction to biological clocks, chemical oscillations and speculations about the nature of the human time sense follow. The major portion of the report deals with the search for the arrow of time in nature from physics. Absolute time in Newtonian physics, time in special relativity and the time inversion invariance of physical laws, appears to leave no room for an arrow of time in nature. Even the concept of entropy and the second law of thermodynamics are found not to be grounded in the laws of nature themselves but rather in the initial conditions of the time evolving systems. The search for the origin of the arrow of time leads to the big bang origin of the universe which was a very low entropy state. The proper description of the evolution of the universe in terms of general relativity (cont'd)

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Karlheinz E. Woehler	22b TELEPHONE (Include Area Code) (408) 646-2896 22c OFFICE SYMBOL PH/Wh

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603 UNCLASSIFIED

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

19. (cont'd from page 1) shows that time cannot be a dimension external to the universe but appears as an internal evolution parameter in recent attempts in the literature to give a quantum cosmological description of the origin of the universe.

DD Form 1473, JUN 86 (Reverse)

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

THE MEANING OF TIME

An Introduction into Philosophical, Biological and Physical Aspects of Time.

Karlheinz E. Woehler

January 1991

Abstract

This report presents an introduction into philosophical, biological and physical aspects of time. Time in man's basic experience, symbolizations of time, the western view of historical time and the evolution of the concept of time in philosophy are outlined. A brief introduction to biological clocks, chemical oscillations and speculations about the nature of the human time sense follow. The major portion of the report deals with the search for the arrow of time in nature from physics. Absolute time in Newtonian physics, time in special relativity and the time inversion invariance of physical laws, appears to leave no room for an arrow of time in nature. Even the concept of entropy and the second law of thermodynamics are found not to be grounded in the laws of nature themselves but rather in the initial conditions of the time evolving systems. The search for the origin of the arrow of time leads to the big bang origin of the universe which was a very low entropy state. The proper description of the evolution of the universe in terms of general relativity shows that time cannot be a dimension external to the universe but appears as an internal evolution parameter in recent attempts in the literature to give a quantum cosmological description of the origin of the universe.

AN INTRODUCTION INTO PHILOSOPHICAL, BIOLOGICAL AND PHYSICAL ASPECTS OF TIME

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

The final impulse to write this report on "The Meaning of Time" came during the month of February 1989 when I participated as a meteorological observer in the Coordinated Eastern Arctic Experiment (CEAREX) on board the "Polarbjörn" which operated around 77° N in the Barent Sea east of Spitzbergen (Svalbart). On many days the ship was made fast alongside a large ice plate and lay still during the night. At 5 o'clock in the morning, with just enough light to see the horizon and the cloud formations I made my first weather observation for the day. With the ship still quiet and only the hum of the generator audible I stood on deck, looking out on this incredible frozen seascape. All processing appeared to have ceased and I was struck by a sense of experiencing "Being" and "Time" in the purest form available in our world.

As in most cases of such profound experiences, this one too had a precursor. Forty years ago, as a young university student I was exposed to Heidegger's philosophy. The various flavors of Existentialist philosophies were hotly debated among young intellectuals then. Because Heidegger was so very difficult to read it became somewhat of an intellectual status symbol to have read in "Being and Time" (Sein und Zeit) [HEi 1]. But Heidegger's analysis of man's most fundamental experience, that of his own existence and its relation to time, the notion of "Time as the Horizon of Being", left a deep impression on me and a desire to see whether these philosophical analyses were somehow born out by what was to me the most trusted path to truth, physics.

No student of physics can totally escape some sense of awe when he studies Einstein's special theory of relativity and under difficulties learns to accept that, the clock on a fast spaceship will appear to run slower than the one here on earth and that he will be an old man when his twin brother returns from a space journey of merely a few years on his clock.

If nothing else, than this insight into the relativity of the measure of time depending on the relative state of motion tends to lead the student to ask a host of questions on the nature and meaning of time and soon to find himself in a multifaceted interdisciplinary field of intellectual pursuit. He begins with questions of philosophical nature (Part A). What is our most basic experience of time? How are we aware of the "arrow of time"? When did mankind become aware of time and how did man articulate his awareness? At least with regard to western culture, which forms our basic experiences, is there a culturally embedded sense of time from which we derive orientation? What has western philosophy to say about time? In part B we turn to biology. If we all seem to have a sense of time direction, what provides this sense? Is there an internal clock ticking away? Can we find biological clocks elsewhere in nature and if so, what kind of clocks are these? And so we are led to the physics of time, in part (C). We know that classical Newtonian physics was built on the concept of an absolute time in which all motion is embedded and can be described. The enormous success of this approach drives western technology and with it an ever more sophisticated technology for measuring time intervals. The properties of light propagation eventually led Einstein at the beginning of this century to analyze what we mean by simultaneity. This analysis led to Einstein's special theory of relativity and the relativity of time measures with regard to observers in a state of relative motion with respect to each other.

Even special relativity, incorporated into the laws of physics, preserves the most amazing fact about the laws of nature; that for each process described by these laws the process in opposite direction is also allowed by nature and should be expected to occur. There does not seem to be anything in the fundamental laws of nature that explains why we see one-way processes in nature, like organisms being born, aging and dying and never the reverse.

Thermodynamics captures this one-directedness of complex system in the concept of entropy and the second law of thermodynamics. The second law states that systems tend toward the state of maximal disorder. This seems to conflict with our every day experience that at least here on earth there appears to be a force at work toward building ever more complex, highly organized structures, culminating in man who in turn submits vast areas of nature to his will, erecting technological structures of great complexity and organization. That life here on earth is made possible by the energy from the sun, was recognized, described and celebrated by the earliest of ancient cultures. It finds a strong support from the thermodynamics of an open system in which the sun produces an enormous entropy increase which locally can be seen as responsible for our experience of the arrow of time. Still that leaves the question: What is the dynamic foundation for the second law of thermodynamics? How can the world behave according to the second law of thermodynamics when all the fundamental laws of physics on which thermodynamics rests, are time symmetric? This leads to search for the arrow of time in the overall evolution of the universe.

Modern cosmology has developed a "Standard Model" of the cosmological evolution for which there is now a great deal of observational support. Einstein's general theory of relativity is the formal apparatus to describe the evolution of the universe. The most striking feature of this model is the notion of a "Big-Bang" singularity at the beginning of the evolution of our universe and the possibility of Black-Hole singularities within our universe. At both singularities time ceases to be a meaningful concept in the classical sense. The question arises whether there are distinctions in the structure of the original Big-Bang singularity and the final Black-Hole singularities that offer an explanation for the thermodynamic arrow of time, leading from a state of very low entropy to one of very large entropy. Quantitative arguments lead to the notion that our universe would appear to be an event of most extraordinary improbability.

Again we must face the fact that entropy is a measure of the expected average behavior of a complex system, consisting of very large numbers of subsystems that collide and interact with each other. How is the thermodynamic arrow of time in the universe anchored in the fundamental dynamical laws that govern the evolution of the universe?

The internal structure of the equations of general relativity which describe the evolution of the space which constitutes our universe has been investigated intensely during the past seventy years. Over the last thirty years significant progress has been made to merge this theory with the other great foundation of modern physics, the quantum principle. It has become clear now that, the theory does not include external clocks with which one could measure the time evolution of the universe. Time information is intrinsic to the evolving 3-dimensional space geometry. Evolving geometry is so to speak its own clock.

Rudiments of a quantum cosmology can be now outlined. In quantum theories one takes seriously the message of the Heisenberg Uncertainty Principle according to which each physical system undergoes fluctuations in its microstructure which cannot be fully described by the laws of physics. The actual evolution of the states of real systems is one of many near-by evolutionary path ways. The laws of quantum physics are not deterministic in nature but rather probabilistic. They give probability distributions for possible outcomes of evolutions. They describe in one equation the evolution of a whole swarm of equal systems each going through its evolutionary track, slightly different from each other due to the individual micro fluctuations which can only be captured in a statistical sense but not for the individual system.

A quantum cosmology then is description of many equal parallel evolving universes and in that way assessing statistically what cannot be assessed in the individual one universe in which we live. Quantum cosmology looks at the universe from outside like God would, contemplating many possible evolutions, which are distinguished by the micro effects of the uncertainty principle but otherwise follow a common set of rules.

The most amazing discovery is that this quantum cosmology does not contain a time parameter. It only contains information about 3-dimensional geometries and gives probabilities for various states of that geometry to occur. There are however quantities, associated with the evolving 3-geometries that can be adopted as representing clocks.

During the last five years a number of proposals have appeared in the scientific literature that attempt to explain the appearance of a time like parameter in the solutions to these quantum cosmological equations. Some are more convincing than others. Many details remain to be worked out.

One of these proposals is particularly intriguing to me. The proposal was made by M. Castagnino and appeared in Physical Review in April 1989 [CAS 89]. Under some restricting assumptions the author shows that one can introduce a "probabilistic time". Probabilistic time intervals are taken to be proportional to the probability of finding the 3-geometry in a particular interval of possible geometries in the phase space of all geometries, regardless of the possible matter distributions for this geometry interval, i.e. the probability is integrated over all matter states. The probabilistic time since the Big-Bang singularity would then be given by the cumulative probability from the big bang singularity to the present expanded geometry averaged over all possible matter constellations. Castagnino shows that this concept of time becomes identical to the classical time in cosmological models in the classical approximation of quantum cosmology far enough from the singularity.

What is so intriguing about this concept of time is that, it is given by the summation of the probabilities of all probable geometry and matter states that the universe could have taken. This entire set of possible states of space and matter represents the possible states of Being. The probabilistic time measure incorporates in the one number "time" the entire spectrum of possible states of being. So, in some sense one can say that "Time represents the Horizon of Being".

Heidegger certainly knew little about general relativity and nothing about quantum cosmology. Yet his searching mind gave expression to a creative vision about the nature of being which we find now in resonance with certain scientific results. This situation is not so different from the insights provided in early mythologies when archetypal images foreshadowed the images which emerged later from rational exploration. It appears that the human mind does have an integrative power of comprehension which in a creative act can see basic patterns in our world and perhaps give expression to this insight in word or image. A mindful culture will listen to such signals and use them as guiding posts in the search for its path.

Although there is plenty of physics to be talked about regarding time, the philosophical and biological aspects are so deeply meshed with our own life that I wanted to give at least a brief overview over some of the connections. These comments can at best be a set of impressionistic reflections for the purpose of raising one's consciousness to the multifacetness of our time experience. Not even the physics part in this report is comprehensive. Many learned books have been written about Time from various aspects. The literature is vast. The reference list at the end of this report is not comprehensive but identifies that literature which guided me through these explorations of which this report is a record. What is original in this report is an attempt to make the latest results on time in quantum cosmology accessible to a wider audience than the technical experts. For these reasons I have chosen to break the treatment into a general presentation with some of the theoretical background delegated to a number of appendices. In this way it is hoped that also a non-physicist will be able to read the report even though it may be at places on a "I-take-your-word-for-it" basis.

2. TIME IN MAN'S BASIC EXPERIENCE

Man exhibits basic behavioral responses which are related to his time awareness, which are identified as memory, nostalgia, expectations, hope, anticipation, intention. These responses are also partially found in lower organisms as everyone knows who has lived with a dog for instance. But man has the unique ability to talk about his experience of time. Lewis Carrol took a close look at our world and asked questions about things which we usually take for granted. He describes what he sees through the experiences of his child and lets Alice view the wonders of the world in "Through the Looking Glass" [CAR 1]. In Alice's encounter with the Queen he explores the "arrow of time" in his own whimsical ways:

> Alice had just met the Queen who was much in need of having her hair set, which Alice kindly does for her. And here is what ensues:
>
> "Alice carefully released the brush, and did her best to get the hair in order. "Come, you look rather better now!" she said, after altering most of the pins. "But really you should have a lady's-maid!"
>
> "I'm sure I'll take *you* with pleasure!" the Queen said. "Twopence a week, and jam every other day."

Alice couldn't help laughing as she said "I don't want you to hire *me*-and I don't care for jam."

"It's very good jam." said the Queen.

"Well, I don't want any *to-day*, at any rate."

"You couldn't have it if you *did* want it," the Queen said. "The rule is, jam to-morrow and jam yesterday--but never jam *to-day*."

"It *must* come sometimes to 'jam to-day,'" Alice objected.

"No, it can't," said the Queen. "It's jam every *other* day: to-day isn't any other day, you know."

"I *don't understand you*," *said* Alice. "It'*s* dreadfully confusing!"

"That's the effect of living backwards," the Queen said kindly: "it always makes one a little giddy at first----"

"Living backwards!" Alice repeated in great astonishment. "I never heard of such a thing!"

"---but there's one great advantage in it, that one's memory works both ways."

"I'm sure *mine* only works one way," Alice remarked. "I can't remember things before they happen."

"It's a poor sort of memory that only works backwards," the Queen remarked.

"What sort of things do *you* remember best?" Alice ventured to ask.

"Oh, things that happened the week after next," the Queen replied in a careless tone.

And the story goes on becoming ever more confusing, as the Queen suddenly remembers that soon she will have pricked her finger on a needle, leaving Alice somewhere between bewilderment and amusement.

Time in our immediate experience is in terms of happenings, of chains of events that change the things around us conveying the notion of evolution of things on the one hand. On the other hand we perceive objects as enduring, giving us the notion of duration.

(3) Time exhibits three basic ordering characteristics of the events in our experience:

(i) A before-after relation holds between any two arbitrarily selected events

(ii) The totality of all events of our experience falls into a past-present-future ordering

(iii) There is a fundamental asymmetry in our relation between events that fall in the past and those that would be in the future.

(4) These ordering characteristics raise a number of questions that need

to be explored further.

- The Now:

  We have a sense that time flows. We recognize a one-dimensional continuum of instants at which events are temporally located. But in addition there seems to be a kind of gliding index-NOW-that gradually moves along this array in the direction from past to future. "It is as if we were floating on a river, carried by the current past the manifold of events which is spread out timelessly on the bank" (Plato).

- Asymmetry of Truth value:

  Statements regarding events in past, present and future seem assymetric in their truth value. Contingent statements about the future have no truth value, unlike statements concerning the past and present which are determinately either true or false. The events of the past for us are definite actuality and unchangeable. Events of the future are largely unknown and only exist as potentialities for us. They are to some extent undetermined. They can be influenced to some degree near the now. The image of a scroll may represent this, which is sliding by with the future side covered and invisible and the factual events coming into view near the narrow window of now and remaining in sight until they fade into the distant past. (Fig 1)

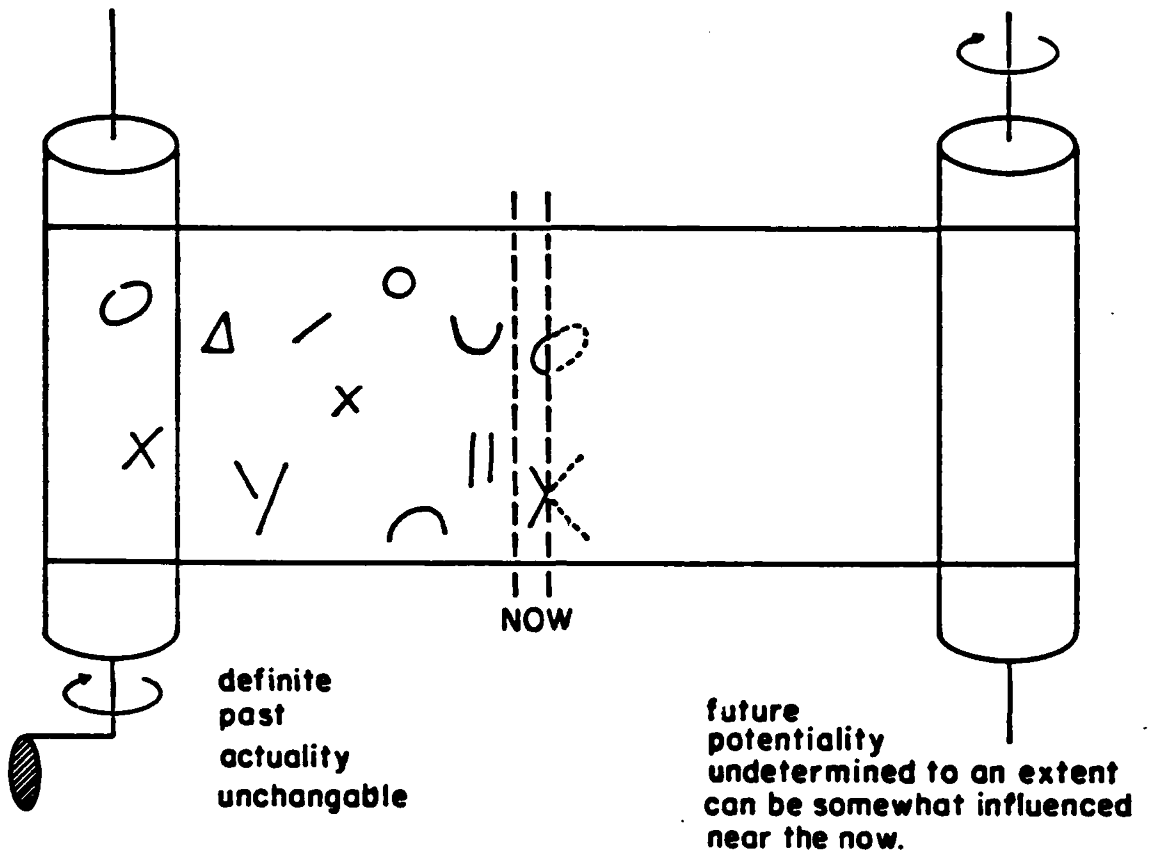


Fig. 1 TIME AS A MOVING TAPE IN MAN'S IMMEDIATE EXPERIENCE

- Defacto Irreversibility:

  There are many processes whose temporal inverse are possible but which in actuality never do occur. We will discuss later in detail, this profound asymmetry that reverse processes are allowed by the laws of nature but are overwhelmingly unlikely to occur which finds its expression in the second law of thermodynamics.

- Knowledge:

  We know more about the past than we know about the future. Knowledge involves an understanding of relations between events or arrays of events. Knowledge means that we have a coherent, transparent model of event fields, established through a comprehension of regularities and patterns in the event field. Such networks cannot be constructed as well for future events because the events are largely unknown to us.

- Causation:

  Effects seem never to precede their cause (see Alice and the Queen). We can influence the future but not the past. Backward causation does not appear to occur.

- Explanation:

  We seek explanations of phenomena in terms of antecedent, rather than subsequent, circumstances. Explanations account for the later event or state in terms of the prevailing conditions or circumstances at an earlier time together with knowledge about existing laws of regularities in event sequences related to the explanation. Explanations of a system's present state in terms of the attainment of some future goal are not accepted in the western scientific system.

- Decision:

  We act for the sake of the future, not the past. The future appears to be some what controllable. Our actions are designed to bring about with greater probability, events in the future which represent favorable outcomes with regard to our intentions. Past events cannot be affected that way. Most will agree that this behavior is based on our understanding of causal connections of events.

- Values:

  We care a great deal more what will happen to us than what has happened. We dread death - the future time at which we will not be alive, but we are much less concerned about the time we missed to be alive before our birth.

(5) Individual time awareness appears to be relative. The relational aspect of time that relates our sense of time to changes and duration of events or objects in our environment sometimes leads to the impression of speeding up or slowing down. The window of now seems to slide by the bank of events more or less fast. Or alternatively, the scroll appears to be pulled more or less fast past our now-window. We will investigate possible explanations for this

phenomenon in part B.

The connection of our time experience with notions of change as well as constancy implies that something that changes must in some sense retain its generic identity. All changes are relative to a constant background. Change is the replacement of at least one specific quality of the object by another.

(6) It appears that, although our individual time awareness may be somewhat relative, the fact that we are unable to change the basic ordering characteristic of past, present and future, the river of time seems to flow unalterably, leading to the notion of an objective time which proceeds independent of all events. Events are merely embedded, placed into this river of time. This content independent time is referred to as Absolute Time. We will return to this absolute time and the technology to measure time and devise time standards in part C.

3. SYMBOLIZATIONS OF THE TIME EXPERIENCE

From the moment word and gesture were born man has tried to condense his subjective experience of time and duration in manifold symbols. Symbols are the signs of communication transmitting in a condensed way experience from generation to generation. A word contains an evolution of various meanings attached to it in its etymological development which are unconsciously reverberating in the final semantic token [ME 66].

(1) The english word time comes from the root ti - to stretch. Thus it directly relates to time span, duration. Ti is also related to temps and tide indicating its relation to seasonal and atmospheric rhythms. A more subtle linguistic link is to tidy, tidiness. Psychoanalysis is very much aware of the relation between time cleanliness and orderliness.

(2) Time eats itself away: There is the age old image of the serpent or dragon swallowing its tail [Fig 2]. This token means both death and eternity. In chinese symbols it becomes the PI the round jade disk with a hole in the center, the symbol of heaven. As the aura that crowns the heads of the saints it appears in our culture. The mythical father Cronos the creator devours his own children. In dreams, time often represents the hostile father, the castrative old man with a scythe who cuts the cord and separates mother and child.

(3) Time as arrow: Symbolizes the western world with its emphasis on progress and growth and with its view of time as an entity going forward in a straight line from the infinite past to the unknown

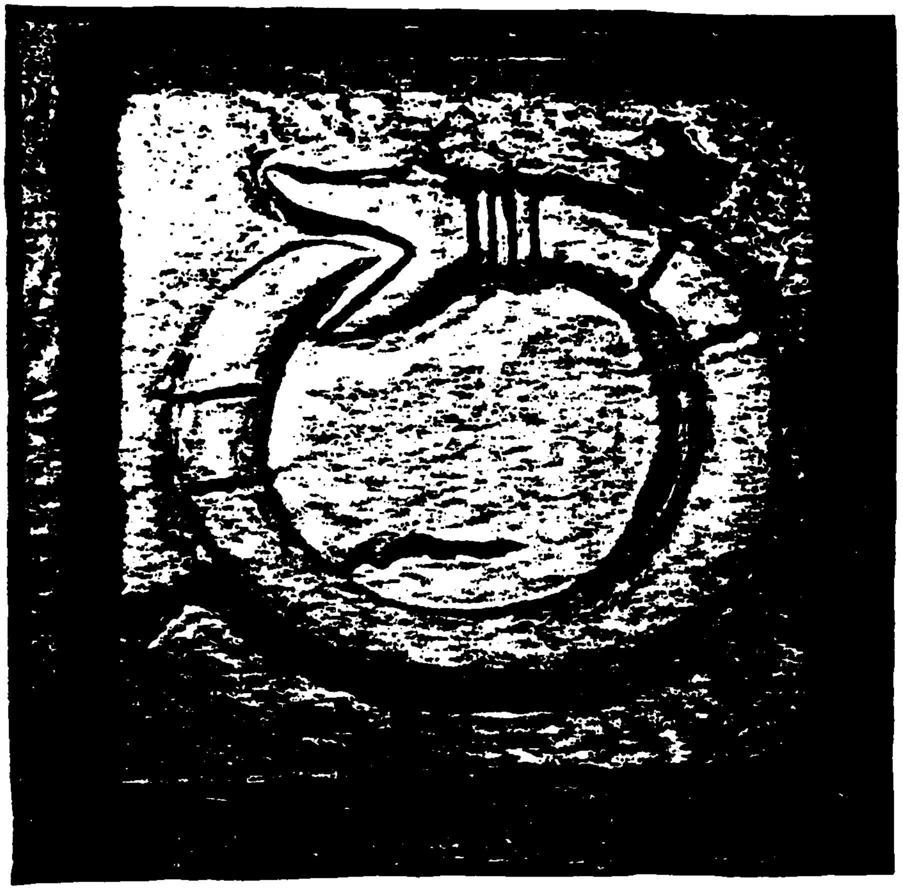


Fig. 2 TIME DEATH AND ETERNITY
Common African symbol of the cycle of life and death. Picture of a bas relief from the walls of the palace of King Ghezo of Dahomey.

Fig. 3 THE TRIUMPH OF TIME
Sixteenth century engraving by Pieter Brueghel
Metropolitan Museum of Art

future. It symbolizes mans feeling of the irreversibility of fate and the hasty agitation of life that devours man. Music traps us in the idea of passing time too. Song and symphony are gone before we can have a second look at them reminding us of the fleeting arrow of time.

(4) Time as the two faced god Janus: The roman god Janus, the god of gates and transition looks with one face into the past with the other into the future.

(5) A picture done in the 10th century by Pieter Brueghel, called the triumph of time, depicts many of the ideas and symbols of time through the ages [Fig 3]*. Central in the picture is anthropomorphized time, Father Time. He is shown in his creative and destructive aspect as Saturn or Cronos the creator devouring his child. In his left hand he holds the serpent biting its tail the symbol of the endless cycle of time in many cultures, sometimes representing the idea of cosmic wholeness. The chariot in endless motion carries the tree of life, the sustaining aspect of time. Sun and moon which govern our life rhythm are drawing the chariot. The globe which is nourishing the life tree is encircled by the signs of the zodiac representing the revolving firmament against which man has measured time since ancient history. In the background to the right, spring time, the time of mating. The church, the home of

---

* Metropolitan Museum of Art

scholasticism. Men and women dancing around the maypole in an ancient fertility ritual celebrating the new season of vegetation. In the left the storms of autumn left only barren trees, completing the suggestion of mans embeddedness into the seasonal rhythm. The wheels of time's chariot are mandala, archetypal Hindu symbols of completeness, in our day perhaps the idea of a finite but unbounded universe. Above Cronos is a weight driven clock ready to strike a bell and remind us of passing time. It suggests increased time consciousness. As time passes, the devices of man are crushed under the chariots wheel. The tools of industry instruments of art even books turn to litter, become outdated and must be replaced. Behind death a cherubim announces the last judgement and with it the end of time the end of life and death [FRA 66].

(6) Time as submission or revolt against paternal command: Man's time concept initially related to biological functions soon gets altered by the interaction between infantile rhythms and paternal schemes. Time and schedule become to symbolize the compelling and intruding external force that offend his infantile feeling of magic omnipotence. Being on time means being clean and obedient to the rule. External submission to paternal commands goes parallel to an internal rebellion preparing for revenge. The same characters may delay their submission, they may kill time, symbolically kill the coercive force, kill the parents in a magic way. The strategy of delay is an ambivalent attack on those that command us. Waiting has

become in our society of schedule the symbol of rejection or lack of power. Hurt pride, depression fill the interval of waiting with anticipation of revenge [ME 60].

(7) Time is money: Animals instinctively follow a magic strategy with their body products by encircling their field of action and possession with them. Homo sapiens chose the reversed way of adaptation to environment by shaping an environment that had to adapt to him. Yet with his more cultivated forms of creation he follows the same instinctual magic strategy as animals do with their body products. All that his hands and mind can shape is turned into one great defense against the emptiness of the world and of death. Hoarding and creating become for him an escape from death. Time is money means that time and money can buy defense against the final death. Time is something to be saved like money. Money genius and power are substitute symbols for time.

(8) Time as creation and the beloved: Man in the act of creation is unaware of time because he becomes part of history and continuity. He gives his time and attention, he gives himself to others. To lovers time symbolizes the moment of communion and shared emotion. Giving time is universally experienced as an act of love.

4. TIME AWARENESS IN THE WESTERN VIEW OF THE HISTORICAL PROCESS [BR 66]

The archeological record from the very appearance of homo sapiens clearly documents mans's consciousness of time and the exploitation of this. Making tools presupposes not only their future use but also the utilization of past experience for future benefit. The frescoes found in caves in the Pyrenees according to experts are designed as hunting magic. These pictures are found in the remotest parts of the caves that must have served as a sort of sanctum. The purpose was based on the belief that the painted representation on the wall of a sacred cave of some animal transfixed with lances will help to make a future hunt successful. In the cave of the "Trois Frères near Ariège there is depicted in the innermost cave the picture of a human clad in animals pelt and antlers mounted on his head: The dancing sorcerer. Such dances are known in all cultures as sympathetic magic, intending to gain power over the animal such represented. But to the primitive mind the magical action of the dance had to be preserved in some form to continue its magical efficacy. This practice of ritual perpetuation of the past is later found in many of man's religious faiths. In these activities of the earliest cultures we see therefore a indication of man's disposition for intelligent planning, involving anticipation of the future in the light of the past and the application of present efforts to future ends. These efforts are designed to gain security against often occurring temporal changes of his life conditions of which his time sense makes him aware. This ability gave him the evolutionary advantage over stronger animals and became the foundation of his complex technology. But the advantages which the time sense thus gives him are offset by a profound sense of spiritual insecurity. It also stems from consciousness of time in causing man to anticipate his own mortality and seek for some assurance of existence beyond death. The provision of food and other equipment found in graves of early civilization reasonably suggests that it was believed that the dead still continued in some form to exist and needed special preparation. This must mean that death for early man constituted a crisis. It is reasonable to assume that from this sense of insecurity given by the awareness of time and death evolved that complex reaction of hope and fear which we call religion.

In the 9th century B.C. Yahwist writers elaborated earlier prophetical appeals into a long narrative that traced the providential actions of their god Yahweh from the very creation of the world through the early history of mankind to the birth of Abraham, the progenitor of Israel and then on to the events of the exodus and the conquest of Canaan. This narrative became part of the sacred literature of the people. It caused the Jews to regard the passage of time as the field in which their god manifests his power and his providence on their behalf. History thus was interpreted as the revelation of the divine purpose. The whole cosmic process was regarded as the gradual unfolding of God's plan, the destiny of Israel being its central theme. Into this complex the destiny of the faithful Israelite was fitted. He had the assurance of Yahweh that as a member of his elected people he would finally be raised to life again, to participate in the glorious vindication of Israel which Yahweh in his own proper time would accomplish. The Christian movement inherited the outlook of Jewish eschatology in connection with the coming of the Messiah. A Christian world view was constructed which envisioned the purpose of God as unfolding majestically through the ages until the second coming of Christ which would mark the end of time and the final achievement of the divine purpose. The Christian lived out the days of his life in preparation for eternity, fortified by divine grace and ever mindful of the destiny of his soul. He was solemnly reminded of the constant passage of time by the church bells that sounded across the land recording the course of the hours. Time for the Christian was invested with the most profound significance. He was taught to see it in a two fold aspect as the gradual revelation of the mighty purpose of God in which his own personal destiny here and hereafter had its minute but essential place. As in earlier cultures the technique of ritual perpetuation of the past was used to conserve the efficacy of the saving death and resurrection of Christ. In baptism the Christian is ritually assimilated to Christ so that he can participate in the new life of resurrection. The efficacy of Christ's historical saving death is made available in the celebration of the Eucharist.

From the earliest days, Christianity was influenced by another view of time which derived from Greek philosophy. In Plato's philosophy God is eternal in the strict sense of transcending time altogether. There is no change and succession and therefore no before and after in God. he has his full and complete being in an eternal now. Plato saw in the perpetual uniform rotation of the planets an image of God's eternal activity. It was believed that the periods of the planets were commensurate with each other; so at constant intervals the same configuration would reappear. The interval between successive returns was the great year. Later philosophers extended this to the belief that this eternal cycle of the heavens influences history, that world history substantially repeats itself over and over in endless recurrent cycles. In this greek view, history was not fulfilling God's design progressively nor moving to any final irreversible consummation as in the Judeo-Christian view. For the human race there was no long term progress, only the endless repeated round of rise and decline, birth and death. Christianity has come under these conflicting ideas of time of the greek and hebrew philosophy in its early development. It was the task of the middle ages to work out a cosmology which should as far as possible combine both. The biblical teaching on history limited the extent to which greek theories of time could be acceptable, but within these limits it extended a powerful influence on medieval thought. It tended to inhibit any idea of systematic progress either on cosmic or human scale. As a result the idea that human history, taken as a whole has a significant structure, reflecting a divine plan disappeared largely from view. It was recognized that God is working out his purpose but we cannot hope to understand what these are. The medievals had no philosophy of history. Modern science has shown scientific evidence that the earth and its living organisms including man have had a finite duration only and that perhaps the whole universe has existed only for 15 billion years. In principle then, the Hebrew Christian finitist view of human history still stands as against the greek infinitist view. Acceptance of a systematically progressive view of history was slow in the beginning. It was slowed by the continuing influence of the greek thought, association of thoughts of progression with atheistic philosophies and most of all the Calvinistic interpretation of original sin which held that man's fall has forever barred him from consistent improvement of human society. Today the theory of evolution is widely accepted. Nevertheless, we are only beginning perhaps to fully assimilate the fact that the greek static view of time is not tenable. The way is open for a return to the more dynamic unidirectional view of world history of the bible but enriched and particularized by the results of modern science. Among those who have recently attempted to reinterpret the christian world picture in this way the most influential has probably been the Jesuit paleontologist Pièrre Teilhard de Chardin.

5. THE CONCEPT OF TIME IN WESTERN PHILOSOPHY

The following is only the very briefest overview of some of the principal philosophers of western culture who have contributed to the body of thoughts on the nature of time.

We mentioned already Plato (428-347BC) as the originator of the essential quasistatic view of time and history. He was preceded by Parmenides (540BC and Zeno (460BC) who taught that only the permanent and enduring are real. All time, flux, motion and change are apparent, unreal [BUR 08]. Aristoteles (384-322BC) was the great master of analytic definitions. To him motion was the central theme. We are only aware of time through changes. We perceive time and movement together. They correspond to each other in magnitude. The process of any change can be quantified by the rate of change. Time is the number of motion in respect to before and after [RO 42].

John Locke (1632-1714) was not a metaphysicist and was not concerned with the ontological status of time. He was an empiricist and tried to show how we build up the idea of time out of elements given to us in experience. Time for Locke is a complex idea build on the idea of succession of our experiences by abstraction to a hypothetical continuity of experiences [LOK 94].

Isaac Newton (1643-1727) saw space and time as absolute real entities outside the acts of our experiences into which the real objects of the world are embedded [NEW 47].

For Gottfried Leibniz (1646-1716) space and time are ideal actual principles, relations that the things of the real world follow. Time is the order of succession, space the order of coexistence [Lei 08].

Immanuel Kant (1724-1804) sees space and time as categories, forms of intuition build into the human mind. Time is not a property of the things but a property of the instrument by which we view things [KA 29].

The dichotomy of the philosophers into those who take a metaphysical viewpoint of time and those who follow Locke's and Kant's categorical interpretation continues into the modern philosophy of the 19th and 20th century. We will see that also science is permeated by this dichotomy.

Henri Bergson's thinking revolved about the phenomenon of life and so we find him more concerned with time as it presents itself in the unfolding of the states of our consciousness. Time is living time, a succession of creative acts [BE 10].

Samuel Alexander (1859-1941) on the other hand in one of the most complex analysis of the phenomena of space and time comes to the conclusion that space and time are the fundamental stuff of which the universe is constituted. Time itself is that creative agent, the origin from which matter, life, mind and ultimately deity itself emerge [AL 18].

6. BEING AND TIME IN HEIDEGGER'S PHILOSOPHY

(1) On July 27, 1915, the young Martin Heidegger, he was 26 then, delivered a lecture before the philosophical faculty of the University of Freiburg in Germany, entitled "The Concept of Time in the Science of History".

In this lecture, Heidegger contrasted the concept of time in modern physics - from Galileo's free fall experiments to Planck's quantum theory - to that underlying the study of history.

In the years following, the concept of time becomes a central theme in Heidegger's thinking, and it reaches a culmination in his seminal work: Sein und Zeit (Being and Time) in 1927 [HEi 1].

Heidegger's analysis of the concept of time developed in the years when major changes in the foundations of physics were taking place that affected the concept of time (Einstein's special theory of relativity 1905, general theory 1915). It appears that these changes that began then are reaching only now a level when the full consequences can be seen. For these reasons it is perhaps permissible and of interest to give some more space to a summary of Heidegger's ideas.

There are three parts to this introduction. In the introductory part (a) we must familiarize ourselves with Heidegger's terminology, with his method of analysis and with the fundamental question about the nature of being, which he poses. In part (b) we summarize Heidegger's analysis of Dasein, the particular form of being to which we have access in ourselves through our existential experience. In the third part then we look at the role of temporality in "Dasein", which for Heidegger becomes the point from which he penetrates into the deeper questions about the meaning of being in general.

(a) Introduction to the Theme in "Being and Time"

- The problem of being:

  Heidegger's concern is the meaning of "Being". The word being in english is used in three different ways

  (i) Being from the verb "to be" (sein in german) means being in the state or act of "Being".

  (ii) The "Being", a thing or object that is, that has being. The latin version is "Entity". In german it is "Seiendes".

  (iii) "Being", the beingness, that which makes things which are, that they are, that which makes for the essence of beings, the being of beings. (Das Sein, in german)

  It is this latter meaning that Heidegger is after.

- Being and Entities:

  Being is the most universal concept. It was

discovered by Parmenides ($\simeq$ 500 B.C.), who saw that of all the diversity of the objects in our world they have one property in common: They are, they are members of the class of beings, they are entities. The problem of how to reconcile this oneness of all entities with the multiplicity of their apparent forms and changes was the main problem in Plato's philosophy and remained one of the central problems of all western philosophizing. While Plato is seeking to understand the very nature of that "universality": "Being", his pupil Aristoteles, whose mind had a more pragmatic bent, was more concerned with these changes that these entities undergo, with all what he called "motions". So Aristoteles' thinking revolved more around attempts to explain the multiplicity of the entities and he focuses on an analysis of the possible modes of being. "An understanding of "Being" is always already contained in everything we apprehend in beings", says Aristoteles.
Due to the overwhelming influence of Aristotle in all of western philosophy the question about the essence of beingness became forgotten. It was assumed with Aristotle, that being cannot be

further defined.

Heidegger rejects this attitude and reopens the inquiry into the essence of beingness.

- <u>DASEIN</u>:

  Among all beings or entities, there is one subset whose being is of particular importance to us. These objects, beings, entities, Heidegger calls "DASEIN". This is the entity "Man". The particular beingness of man, of "Dasein", is "Existence". Existentiality is the particular mode of being of entities that exist, that is man, Dasein. The importance of the Dasein form of beings is that an understanding of the beingness of the entity man may lead to an understanding of beingness in general.

- <u>Dasein and World</u>:

  The essential element of Dasein seems to be its <u>being in world</u>. An understanding of being of Dasein can lead to an understanding of <u>world</u> in general. Any ontology (understanding of beingness) of entities which are not Dasein, is necessarily based on the ontic structure (structure of being) of Dasein. This is the only access we have, to explore beingness, namely in the beingness of man.

So, prior to all other investigation concerning being, there must be a "fundamental ontology", which is based on a existential analytic of Dasein. This analytic of Dasein is a preparatory task which can only reveal the being of the entity man, and is not expected to give directly an interpretation of the meaning of "Being".

- Heidegger's Method:

Heidegger's method of investigation is phenomenological. Phenomenology is the science of phenomena. Phenomenology does not describe the "what" of the objects of investigation by subscribing to particular points of view or direction. Racher it describes the "how" of the investigation itself. The word phenomenology comes from the greek work "phainomenon". It means, that which shows itself, the self showing, the manifest. The root of the word is "phaino" meaning: To bring to light. The other root is the word "logos" which means making manifest, discovering, placing in truth. Phenomenology then means, to let what shows itself be seen from itself, just as it shows itself from itself.

"Being" had been concealed, forgotten, undefinable.

The only access to the question of "Being" is to let it show itself. The only possible ontology is phenomenology. The phenomenological concept of phenomenon as "self-showing" means the "being of beings".

(b) The Analysis of Dasein

- The Essence of Dasein:

  The essence of Dasein, the being of the entity man consists in its existence. Existence is the particular mode of being of Dasein, of man. Dasein shows itself first in the "I am". In that assertion "I am", Dasein is essentially its own possibility. Dasein can "choose itself", it can "win itself" or "lose itself". So, we distinguish two modes of being of Dasein: Authentic existence and non-authentic existence.

- Being in the World:

  Being of Dasein is determined by "Being-in-the-World". The "in" here is not meant as spatial concept. It is meant as being involved in world, dealing with the world.

  Knowing the things of the world is a possible mode of dealing with them. So, knowledge is a mode of being in the world.

- The World at Hand:

  "World" is not the things which exist in the world, not nature or other entities. World here means a characteristic of Dasein itself. Man finds himself in a world that is not primarily "present" but "at hand".

  This observation leads Heidegger to a sideline of the investigation about the role of utensils, the things "at hand" and to the worldliness of world as res extensa, entities of extension, and from there to the spatiality of existence.

- Coexistence:

  Being in the world is characterized by being with others. The world of Dasein is a common world. Being of Dasein is coexistence. The coexisting are the "they", impersonal others. The "they" fulfills Dasein in its everyday life. The "they" is an existential and it belongs to Dasein as a primary phenomenon.

- Every day Life:

  Dasein is characterized by facticity and openness (Erschlossenheit). Openness means being essentially open to the things. Heidegger distinguishes two modes of being in the world.

(i) everydayness, trivial unauthentic existence

(ii) authentic existence

In everydayness, trivial unauthentic existence man exists as the undistinguishable impersonal "they". In this mode man sees himself as fallen, lost in the world. He finds himself in a state of "thrownness" (Geworfenheit). He is thrown into the world.

- Authentic Existence:

  The mode of authentic existence is found in anguish (german Angst). Anguish is not caused by this or that undesirable event that may occur and affect our life. That is fear. Anguish is caused by nothing. "Nothingness" reveals itself to us in anguish. Dasein is characterized by care (Sorge in german), that is concern, preoccupation.

- Death:

  Death is an essential characteristic of Dasein. Death is not the particular event. Death for Dasein is a "not yet", a matter of "coming to one's end". Dasein is always incomplete, because its conclusion implies at the same time a ceasing to be. Dasein is a being-towards-death and as such death is a constitutive part of Dasein. Dying is based on care (Sorge). Death is the most authentic possibility of existence. The "they", who live trivial unauthentic existence tend to hide this fact, the certainty of death. Death is the most proper possibility of Dasein. In authentic existence the illusions of the everyday life of the "they" are overcome. Dasein is free for death.

- Temporality:

  Dasein is characterized by care (Sorge). What is the meaning of care? Anguish in the face of death is a "not-yet". Concern is characterized by an awaiting. Care is a matter of something in the future. Temporality manifests itself as the sense of authentic care.

(c) The Role of Temporality in Decision:

In this final section of this exposition of Heidegger's Sein und Zeit [S & Z...] we will elaborate a little more on the ideas of the last paragraph of the previous section. For this we will stay closer to the central part of the original text by giving the corresponding page numbers.

- Resolve [S & Z. pp 296-298]:

  Dasein is found in the mode of anguish (german Angst). In authentic existence we understand our existence as acceptance of our obligations (german Schuldig-Sein). Obligation here means ethical obligation for the care for the other. These modes of Dasein constitute the openness of Dasein through conscience (german Gewissen). In our conscience we become aware of our anguish (german Gewissensangst). The quiet anguished acceptance of ones own obligations we call <u>resolve</u> (german Entschlossenheit). The german "Entschlossenheit" conveys more the openness and acceptance aspect. The english word resolve carries more of the sense of liberation in the acceptance of the obligation. In this resolve the self is thrown from the taking care of things into the caring being with others.

- <u>The Anticipating Resolve [S & Z, p 305]</u>:

  The existential acceptance of obligation is only really fulfilled when the resolve in its opening of Dasein has become so clear that it understands the obligations as continuous obligations.

  This understanding is only possible when Dasein infers its being to its end. Resolve, therefore, is actually understanding being as being-toward-the-end. It involves the anticipation of death.

- Care and Selfness [S & Z, p 322]:

  In saying "I", Dasein expresses itself as being-in-the-world. "I" means the being that is in the world. It means the being that is concerned with being. It means the being that cares. It means the "self" that has taken a stand in its caring. It means the self reliant (german selbst ständig) self. This self reliance is the anticipating resolve. Dasein is properly self in its original quiet anguished accepting resolve.

- Temporality as the Ontological Meaning of Care [S & Z, pp 325-3:

  Dasein becomes essential in authentic existence. This authentic existence shows itself in the anticipating resolve. The mode of authentic care reveals the original selfness and wholeness of Dasein. The anticipating resolve is being toward its own authentic possibilities of being.

  Dasein, then, is coming-toward-itself in its own possibilities.

  In this letting-itself-come-toward-itself it bears its possibility as possibility.

  This bearing the authentic possibility, letting itself-come-toward-itself is the original phenomenon of "coming-to", that is the "future". Future, in this view is not a later "now" an instant that has not yet become actual. It rather is the "coming" in which Dasein in its authentic possibilities of being comes toward itself.

The anticipating resolve opens each particular situation of its "Da", so that the existence takes care of the factual items at hand in the world through deliberate action.

The resolute being in situations at hand, letting one self encounter world in action is only possible when there is a "presencing" of the being.

Only as presence in the sense of presencing can resolve be what it is.

Coming toward itself in its future, resolve brings presencing into each situation. The "having been" emerges from the future (the coming-to) by letting pressure emerge from the prior (the past) coming-to. This unified phenomenon of having been-presence-coming to, we call temporality. Only through the determination of Dasein as temporality obtains Dasein the characteristic wholeness of the anticipating resolve.

Temporality emerges as the meaning of authentic care. Temporality is not a being. Temporality is not but "timeth" (german zeitigt).

- Finiteness and Temporality:

Care is being toward death. In death Dasein comes to the end of its possibilities. The anticipating resolve ends with death. The "coming-to" (the future) therefore is finite.

This seems to clash with our notion that "time goes on". The difficulty lies in the confusion between the original concept of temporality in its phenomenological characterization as given above and the "every-day" notion of time, which bases on the "former", "later", "not yet", "no more" of the things being in time. What does it mean: "Time goes on", "time is infinite", "things are in time"? It would have to be explained, how the notion of infinite time emerges from the finite original concept of temporality; not the other way around. Heidegger does that in paragraph 78 by showing that the phenomenon of presencing in the acts of taking care of worldly things leads to the notion of "datability". From there he proceeds to the notion of "duration", which eventually unravels the entire structure of the every-day notion of time.

- The Temporality of Being-in-the-World and the problem of the Transcendence of World [S & Z, pp350-384]

At this point Heidegger begins to slowly move into the temporality of the other beings that are not Dasein. And here the work becomes fragmentary. The "coming-to" (i.e. the future), "having been" and presence show themselves in the phenomena of the "toward-one-self", "back-onto" and "letting-be-encountered-by" of Dasein. The phenomena of the "toward", "onto", "by", show temporality as ecstasis par excellance. By ecstasis we mean a state in which man reaches outside one self. Future, past and present are called the ecstasies of temporality.

That form of being here called dasein is open to itself. It is "enlightened". What makes it open to itself is the care. This enlightenment makes possible the "seeing", the "having of world". It is the ecstatic temporality which opens the original "Da" (english there). Only through the rootedness of Dasein in temporality does the basic phenomenon of "being-in-the-world" become clear.

- <u>The Questions of the meaning of Being in General</u>: How is "world" possible? In what sense "is" world? What and how does world transcend? How does world manage to show itself in our being-in-the-world? The being of Dasein was determined as care. The meaning of care is temporality. In the openness of the "Da" ("there") "world" is opened up. Therefore the ontological structure of "world" must also be founded on temporality. The existential temporal condition of the possibility of world lies in that temporality as ecstatic unity has something like a <u>horizon</u>. The ecstasies of the temporality are not simply a "carried away". There is an associated "where to". That "where to" represents the horizontal scheme. In this horizonal unity of the ecstasies lies the ground that to being there belongs an open world. On the basis of the horizonal ecstatic structure Dasein is essential "in-a-world". World is not just there at hand. It "is" together with the "outside-one self" of the ecstasies. The problem of transcendence is the question: What makes it ontologically possible that beings can encounter in the world and that as encountering

beings they can be objectivised?
Something like "being" is opened up in the understanding which belongs to the existing Dasein. Dasein as existing being-in-the world relates to beings, those that it encounters in the world as well as to itself as existing being.
How is understanding of "being" from the perspective of Dasein at all possible?
Can this question at all be answered by going back to the original way of being of Dasein which comprehends the concept of "being"?
The existential ontological constitution of Dasein as a whole is founded on temporality, we saw. Therefore, Heidegger speculates, an original way of "effecting temporality" by the ecstatic temporality must somehow make possible the ecstatic design of being in general.
How is this "way of effecting temporality by temporality to be interpreted?
Is there a path that leads from the original temporality to the meaning of "being"? Does time itself reveal itself as the horizon of being?
This last question of the work rings out demanding, challenging, electrifying. It has that same

alarming quality of the last trill of the trumpet in Bach's second brandenburg concerto in F-major (BWV 1047), that leaves us in suspense and awe.

- We are going to leave Heidegger here and turn to time in biology and in physics. We will encounter at first the "every-day" notion of time which flows and the concept of time in the context of things being in time. Only toward the end of this report when we talk about time in Einstein's general theory of relativity will we return to aspects of temporality that appear similar to that of Heidegger's original phenomenological time.

B. BIOLOGICAL AND PSYCHOLOGICAL ASPECTS OF TIME

7. THE BIOLOGICAL CLOCK

- Phenomena pointing to biological clocks

On the biological level three lines of study have given scientific insight into the nature of our sense of time.

(a) Photoperiodism

(b) Celestial orientation of animals and their time sense

(c) Circadian rhythms in living organisms

Photoperiodism is the influence of the relative length of day and night on the flowering response of many plants. Plants seem to measure the length of the day by comparison with a built in rhythmic variation of their light sensitivity.

That the bees amazing ability of direction finding must be connected with an internal clock that was long known (Forel 1900). It was also known that birds fly great distances and arrive at specific destinations. Even young birds in their first migration if captured and held in captivity for some time would upon release head in the right direction.

Gustav Kramer proved in 1949 that the birds sense of direction was dependent upon the apparent position of the sun in the sky, and that in their direction finding they were compensating for the presumed movement of the sun across the sky during the day [KR 56].

The discovery of the circadian rhythm has provided most information concerning the possible nature of the biological clock which is probably in operation in all these phenomena. Already 1729 DeMarian found that many plants show diurnal leaf movements, extending their leaves in day time and folding them at night. This rhythm continues for several days even if the plants are placed in constant darkness [Bü 56]. We know now that circadian rhythms are exhibited by almost all organisms. Intensive study on cockroaches has finally shown, that their activities which persist in complete darkness are related to the rhythmic secretion of endocrines by nerve cells close to the brain. Even fungi show characteristic rhythmic growth and spore formation. Man shows circadian rhythms in the blood stain count, serum iron content, body temperature, heart rate, blood pressure, urine production, excretion of phosphate and potassium.

- Speculations on the nature of biological clocks

  Having mentioned the orientation ability of many animals, photoperiodism and the circadian rhythms one asks whether or not the same mechanism is used by all organisms to meter time. Most biologists are inclined to think that circadian rhythms are direct manifestations of the biological clock and that they are caused by rhythmic endogenous changes. In search for the nature of the mechanism one important finding is, that plants do not measure day length by measuring the absolute length either of the light period or the dark period but determine the length of day according to the time at which it receives light in relation to its endogenous circadian rhythm. The photoperiodic response seems to be dependent upon an endogenous circadian rhythm of light sensitivity. K.C. Hammer has shown, that cockroaches, fruit flies, hamsters and various plants maintain their circadian rhythm even at the southpole [HA 62]. This suggests that any factors associated with the earth's rotation cannot be responsible for the regulation of the internal clock. In fact the majority of biologists now are inclined to assume that not an external stimulus at all is governing the biological clocks but that the circa 24-hour clock is inherited. Evidence in favor of this hypothesis is mounting. Still it would remain to be explained then: If the clocks are inherent and inherited, why do the circadian rhythms tend do die away after long periods under constant conditions? A hypothesis put forward by Cloudsley Thompson [CLTH 61] is that the animals activity is controlled not by the one clock but a large number of cellular clocks which are kept in synchronization by an environmental factor like changes in light intensity. Deprived of this synchronization the clocks will gradually go wrong and get out of step. The animal loses its rhythm but a single exposure to light or the onset of rising temperature is sufficient to synchronize them.

The reason for the evolution of an inherited biological clock can be understood from the obvious advantage to a species. Advantages by synchronization of the activities of the members of a species are gained in their chances of mating, adaptation of feeding rhythms to the activity rhythms of the prey, avoidance of predators by adjustment to their activity times, food competition, migration orientation to find preferred breeding places etc.

- <u>The Chemical Pacemaker Model</u>

More information about the nature of the human time sense came from experiments done by H. Hoagland in 1933 [HO 33]. His experiments indicated that the human time sense is basically dependent upon the velocity of oxidative metabolism in some brain cells. A chemical pacemaker generally is the slowest reaction in a sequence of linked reactions. Control of rhythms in nerve and muscle cells is determined by continuous metabolic processes. The rhythms are relaxation oscillations in which some potential is built up to a critical value and then discharges through resistances that are reduced during the discharge.

A mechanical prototype of this type of relaxation oscillation is the simple water clock. A receptor is overbalanced by a weight. Water is slowly filling the reservoir. When it is filled, the balance tips to the other side spilling the water out of the reservoir upon which the balance tips back to repeat the cycle. The time interval is governed by the flow rate of the filling [Fig 4]. In the electrical analog the reservoir is the capacitor which is being charged by the battery. The gas diode becomes conductive above a certain voltage which leads to discharge of the capacitor. The gas diode acts like the tipping of the balance and spilling [Fig 5].

That chemical reaction systems can exhibit oscillatory behavior can be demonstrated in the behavior of two-species reaction systems in which one species N_2 "feeds" on the other N_1. The "prey" species N_1 has a constant birth rate and a loss rate due to being eaten by the "predator" species. Such models are useful in population models as well as in chemical systems. For suitable parameters of birthrates and feeding rates these populations can show definite oscillatory behavior [Fig 6]. These equations are so called Lotka-Volterra systems.

In all the examples the dynamical system has intrinsic nonlinear elements in the system. In the predator-prey model the nonlinearity lies in the characteristics of the predator-prey relationship which shows up in the nonlinear term in both differential equations. In the RC-circuit it is the non linear current-voltage characteristics of the gas diode that provides the non linearity. In the water clock it is the large

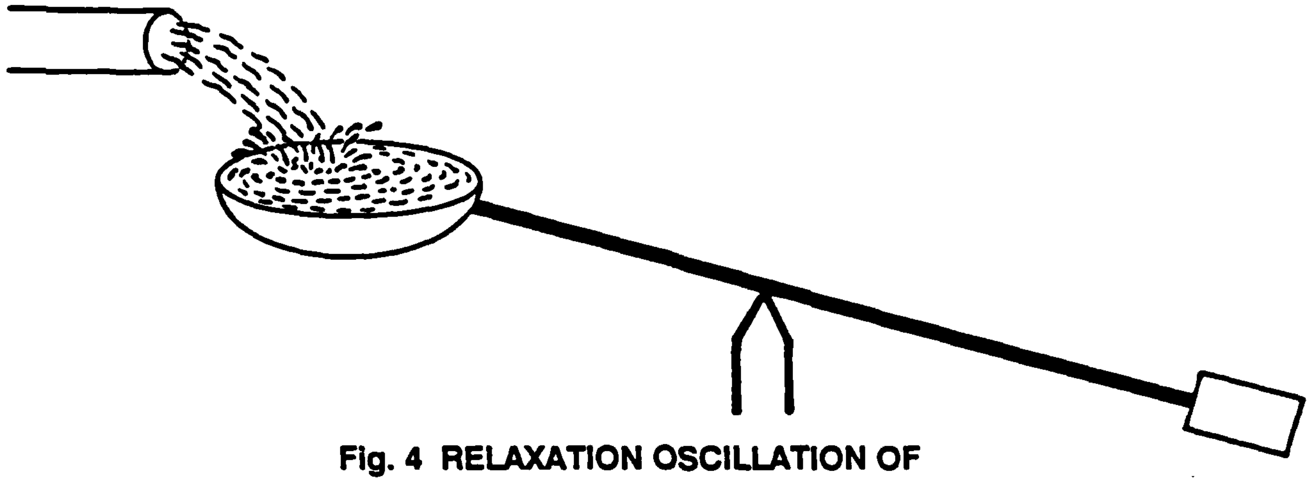


Fig. 4 RELAXATION OSCILLATION OF A SIMPLE WATER CLOCK.

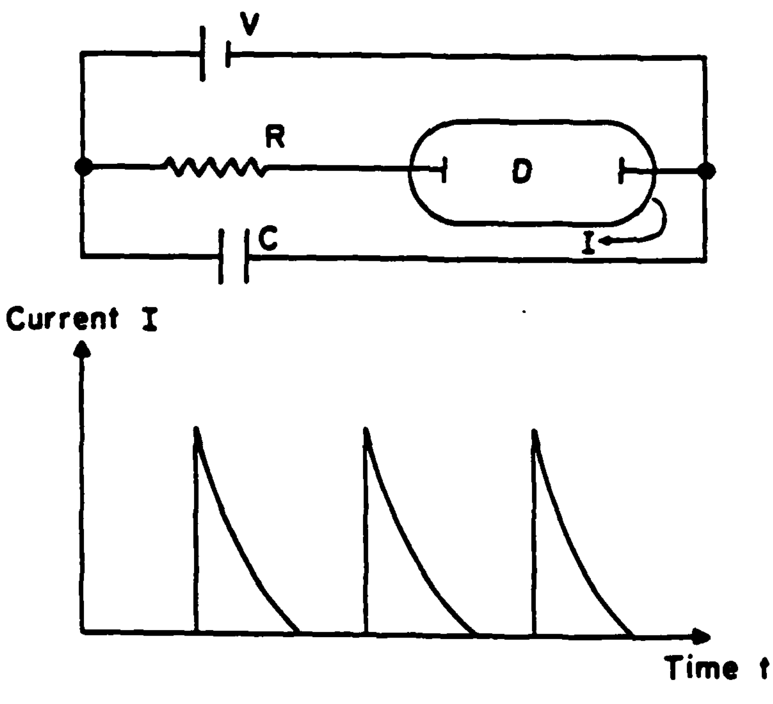


Fig. 5 RELAXATION OSCILLATION OF RC-CIRCUIT WITH GAS DIODE

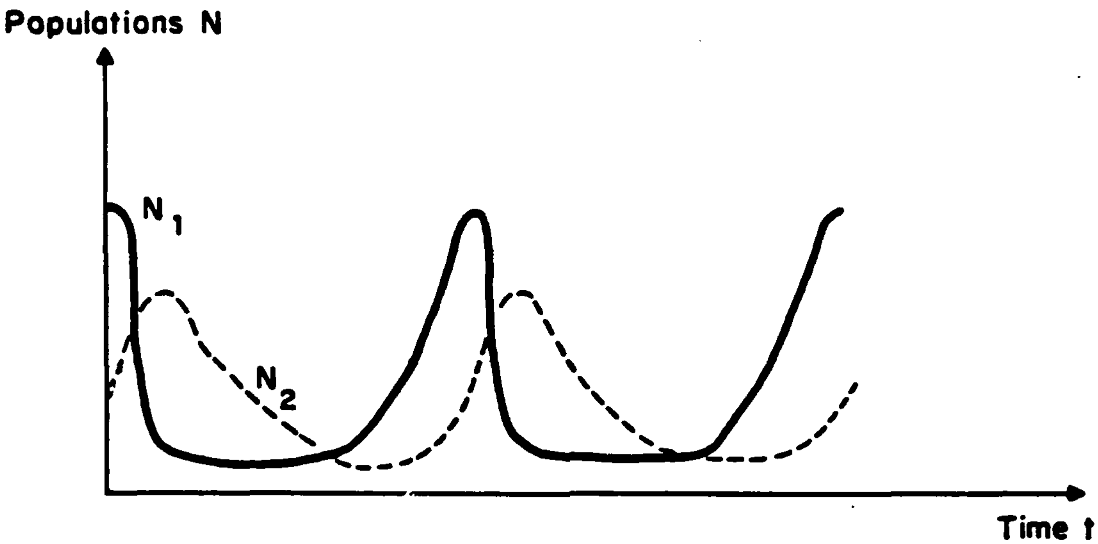


Fig. 6 OSCILLATORY BEHAVIOR OF A LOTKA-VOLTERA SYSTEM REPRESENTING PREDATOR - PREY RELATIONS FOR PREY POPULATION N_1 AND PREDATOR POPULATION N_2

$$\frac{dN_1}{dt} = aN_1 - k\,B_1\,N_1\,N_2$$

$$\frac{dN_2}{dt} = -cN_2 + k\,B_2\,N_1\,N_2$$

a = natural birthrate of prey; c = natural death rate of predator. $k\,B_2\,N_1$ = population growth rate of predator due to availability of prey; $k\,B_1\,N_2$ = decay rate of prey population due to predators.

amplitude excursion of the reservoir at the shift of balance and the abrupt spilling of the content which is intrinsically non linear. In modern theory of non linear dynamic systems such behavior is identified as limit cycle oscillations.
For chemical relaxation models to serve as a model of the time pacemaker in biological systems one would hope that this model would explain certain features of the time pace making.
All chemical reactions are temperature dependent. This dependence is described by Arrhenius' law:

$$\dot{N} = Z \exp[-\mu/RT]$$

$\dot{N}$ is the chemical reaction rate, Z is a constant, T the absolute temperature, R the gas constant and μ the energy of activation, i.e. the kinetic energy per mol which the molecules must acquire before they can react.
If the frequency ν of some physiological rhythm is directly proportional to the reaction rate of some underlying chemical pacemaker then it is:

$$\nu = K \dot{n} = a \exp[-\mu/RT]$$

or

$$\ln\nu = G - \mu/RT$$

In the case of the above Lotka-Volterra system one can solve the problem for small values of the oscillation by linearizing around the equilibrium point and one finds for the angular frequency.

$$2\pi\nu = (ac)^{1/2}$$

This shows that for small oscillation the linear birth and death rate are the important ones and that indeed the frequency of oscillation is proportional to the geometric mean of the birth and death rate of predator and prey population.

A logarithmic plot of the frequency versus 1/T should give a straight line with slope μ [Fig 7]. In this way μ-values for certain enzyme reactions have been measured by Hodidian and Hoagland [HO 39].

Enzymes react with substrate molecules oxidizing the substrate. In a number of cases the μ-values have been found to characterize the particular enzyme substrate reaction uniquely. By comparison with the μ-values characterizing certain physiological rhythms it is hoped to eventually identify the pacemaker reactions associated with the time sense of man. Indeed if heart beat rates, breathing movements, α-brain wave rhythms and other physiological rhythms are plotted on logarithmic scale vs. 1/T the Arrhenius relation is confirmed in all cases. Each process has a clearly defined μ-value. These fall into the range 8000-16000 cal. If private psychophysiological time is determined by chemical velocities

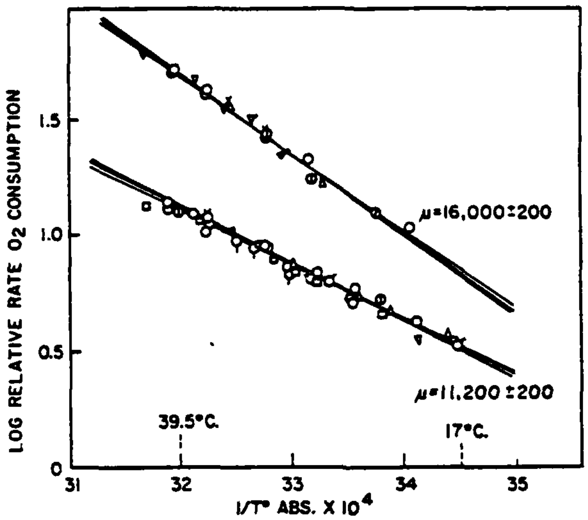


Fig. 7 USE OF ARRHENIUS' LAW OF CHEMICAL REACTION TEMPERATURE DEPENDENCE TO DETERMINE THE POSSIBLE PACE MAKING REACTIONS IN RHYTHMIC PHYSIOLOGICAL PROCESS

$$\dot{n} = Z \exp\left[-\frac{\mu}{RT}\right]$$

$\dot{n}$ = chemical reaction rate; Z = reaction constant
T = absolute temperature; R = universal gas constant
μ = energy of activation.
ν = frequency of physiological rhythm depending on pace making chemical reaction with activation energy μ.

$$\nu = K\dot{n}$$

$$\ln \nu = \text{const} - \frac{\mu}{RT}$$

then raising our internal body temperature should speed the reaction thus making more chemical change and hence more physiological time pass in a given interval of physical clock time. We would think that time was dragging on. Looking at the clock it would be slower than we think it should be. We would tend to arrive early at appointments. Lowering the internal body temperature decreases the metabolic rate, makes the physiological clock run slow, physical clock time seems to run fast, time flies by rapidly, the subject tends to arrive late at appointments. Experiments on persons whose body temperature was changed by diathermy were made. The persons subjective clock time was measured by letting them tap their subjectively counted second intervals. The data showed clearly the Arrhenius dependence and gave μ-value of

24000 cal.

This suggests that our sense of time depends on a chemical pacemaker in some group of brain cells with that μ-value. The chemical clock is not identical with the α-brain wave rhythm which has a μ-value of 8000 and whose enzyme system is identified. The relationship between time perception and metabolic rate of subjects can also be studied during states of excitation and tranquilization. The former is associated with increased metabolic rate and can be produced by hallucinogenic drugs like mescaline, LSD, etc. The latter state with decreased metabolic rate is produced by hyperthermia producing tranquilization. The change of subjective time under the influence of these drugs can be measured again by the tapping rate and the experiments show clearly the expected effect. But it is of great importance to note that LSD for instance not only raises the body temperature and thus produces over-estimation of clock time i.e. time contraction but also produces a simultaneous expansion of space. The latter phenomenon can be illustrated by the size of handwriting samples taken under standardized conditions before, during and after the excitatory drug produced experience. The increase in tapping rate as well as the increase in handwriting space suggests a deep interrelation of the subjective time and space concept. Metaphorically speaking we can compare ourselves to the spider excreting his web and thus creating his own space-time coordinates.

Efron conducted further experiments in connection with this possible interaction of our space and time sense [EF 63]. He found that when two stimuli differ in intensity they are relayed along the neural pathway with a different delay into the hemispheric space of the observer. This delay is not corrected for by the central nervous system but interpreted as space.

R. Fisher sees in these experiments evidence for a unifying origin of all of mans space-time concepts: [Fi 62] Its biological perception and conception through sensory transduction or instrumentation. The differences in the space-time concepts in physics, biology, psychology, etc. are only small differences suggesting how far our specialized approaches removed the concept from the original perceptual context. We are reminded of words by St. Augustinus in his confessions book 11: "It is in you O my mind that I measure time. I do not measure the things themselves whose passage produced the impress, it is the impress that I measure when I measure time. Thus either that is what time is, or I am not measuring time at all".

8. THE DATA STORAGE SIZE HYPOTHESIS OF THE HUMAN TIME SENSE

- Of the more recent investigation on the psychology of man's time sense a little book by R. Ornstein, Professor of Psychology at the University of California in San Francisco, "On the Experience of Time" (1969) is of particular interest in the context of this report [OR 69].
- Ornstein distinguishes between four basic kinds of time experience.

  (a) Short time experiences:

  Peterson & Peterson (1959) [PP 59] have demonstrated the existence of immediate memory processes. These memories are fleeting and rapidly decaying. They are quite distinct from long term memories. Miller showed (1956) that the information processing capacity of this immediate memory is fixed at a low amount and is extremely difficult to modify by training [Mi 56].
  Short time experiences are considered by most people as being events 3-4 seconds apart. Long term time, or duration begins somewhere above 10 seconds.
  Beside the aspect of estimate of short intervals, immediate perception and apprehension of short intervals, the short term time experience also include experience of rhythmic nature, motor aspects of timing, tapping of time intervals.

(b) Experience of long duration:

These are the experiences of time passing, lengthening or shortening of individual time duration experiences according to individual state of perception, aspects of permanence connected to remembering of things past and retrospection.

(c) General temporal perspective:

This is derived from the general cultural background. While western culture has evolved a very precise sense for timing, other, less technically oriented cultures have different basic units of time.

(d) Experience of simultaneity:

This experience appears to depend on how fine the "grain" of time experience is for the individual.

- Ornstein considers and critiques the sensory process metaphor. He observes that all theories on time experience appear to postulate some sort of "time base", a repetitive, cumulative pulse dispensing mechanism which delivers internal time signals, on "Organ of Time" so to speak. This time basis is identified either with a specific periodicity which is then usually called a time quantum or with a specific bodily process called a "biological clock".

  The postulated time base appears to have two characteristic scales.

(a) Perceptual moment scale ≈ 0.1 sec.

(b) Indifference interval ≈ 0.7 sec.

The indifference interval is the interval when observers who are asked to estimate time intervals appear to be most successful. The explanation of this is often given in terms of the sensory processing idea, according to which there is a real time existing independently of ourselves which we perceive more or less accurately. Below 0.7 seconds observers seem to over estimate and above 0.7 seconds observers tend to underestimate. It is not clear how this finding of a most accurately guessed time interval would resolve the confusion of our arbitrary clock intervals hours, minutes, seconds with anything like real time.

Norbert Wiener (1948) suggested that α-brain waves (8-12 cps) might be the "ticks" of the biological clock that breaks time into a sequence of "moments" of 0.1 seconds length [Wi 48]. The problem with this hypothesis is that the α-wave is not always present in the EEG. Does that mean that the clock stops occasionally? Does the clock change its rate when the frequency of the EEG changes out of the α-range?

It was already mentioned in the previous section that experiments on the influence of drugs on the time experience like those done with psilocybin by Fischer [Fi 67] show apparent lengthening of duration experience. The drug induces increase in the handwriting, decrease of taste differentiation, increase in the chosen rate of tapping time intervals and increase in rhythmic eye movement. Using such sensory, motor and psychomotor "clocks" Fischer found that these clocks do not run at the same rate.

So, one must ask why these processes are called chronometers or clocks. What are the criteria for judging a given physiological process a chronometer, an internal time keeper? Why would taste be a chronometer and not hair growth or heart rate or skin pigmentation? There are almost an infinity of physiological processes which might alter the rate in response to psilocybin or to some other manipulation. Also, if all these various processes are chronometers, then which is <u>the</u> biological clock? Is it a combination of all these? If they all run at different rates, which is the right one? From these considerations it seems very difficult to link the hypothetical time base to one particular process.

- These considerations lead Ornstein to his <u>storage size metaphor</u> of time. Perhaps Guyau (1890) [Gu 90] was the first to relate time experience to human information processing. According to Guyau, <u>time itself does not exist in the universe</u>, but rather time is produced by the events which occur. He considers time as a purely mental construction from the events which take place. Time experience is constructed on the intensity of the stimuli, the extent of the differences between the stimuli,the number of stimuli, the attention paid to the stimuli and the expectations called up by the stimuli.

The specific features of the storage size metaphor become clear if one considers the computer. If information is given into a computer and instructions are given to store that information in a certain way one can check the size of the array or number of spaces or number of words necessary to store the information given into the computer.

The storage size metaphor relates the experience of duration of a given interval to the size of the storage space for that interval in general information processing terms.

In the storage of a given interval either increasing number of stored events or the complexity of those events will increase the size of storage and as storage size increases the experience of duration lengthens.

This hypothesis leads to a working hypothesis according to which anything which might alter the size of storage of the information in a given interval, will also affect the experience of duration of that interval.

For instance, application of the storage size hypothesis to the drug experiments would lead to the following definite prediction: Drugs affect the cognitive processing. Drugs which increase awareness or alertness should result in more information from the stimulus array reaching consciousness. This should have the same result as would have an actual increase of the number of physical stimuli that are present. It should then increase the storage size and therefore lengthen the experience of duration of that interval. Our brain is a filter that sorts out the stimuli and reduces the number that reach consciousness to that size which is necessary to function and survive. It eliminates extraneous information. Stimulant drugs "open the gates". Psychedelic drugs open these gates so wide that a torrential flood of inner sensations reaches consciousness. There is then an increase of speed of conduction of nervous impulses and increasing rates of cortical firings [Fi 66] and increased information processing. The reports of users of psychedelic drugs on their experience of duration is unanimous: Duration is lengthened relative to ordinary experience.

- In his book, Ornstein reports on a number of experiments which were designed to contradict or confirm the storage size metaphor. Three of these experiments may serve as examples.

  (a) Observers listen to tapes of equal length but with different number of events recorded on it and are asked to estimate the tape length.

  The storage size hypothesis predicts that the tapes with more events will be judged longer. Results confirm the hypothesis.

(b) Observers look at figures of varying complexity for equal time intervals and are asked to judge the length of observation time given.

The storage size hypothesis predicts that the observation intervals will be judged longer in which the observers look at the more complex figures.

Results by and large confirm the hypothesis although there were some difficulties.

(c) Observers were listening to equal length tapes with a standard set of events. However the sound events were arranged on some tapes in an ordered sequence, on others in random sequence. The observers were to estimate the tape length.

The storage size hypothesis predicts that the disordered sequence of events are more difficult to code and process and therefore needing more storage size than the more ordered sequences and would therefore be judged longer.

The results confirm the hypothesis.

- The notion that time does not seem to exist in the human universe but is an experience associated with the required data storage size is a notion which we will encounter again when we explore the meaning of time in the large scale models of the universe.

C. TIME IN ITS PHYSICAL ASPECTS

9. ABSOLUTE TIME IN NEWTONIAN PHYSICS AND THE MEASUREMENT OF TIME

- We have already mentioned in earlier sections that in our immediate experience time is associated with the notion that the narrow window of the present seems to slide along like the continuing flow of a river, giving us the impression of "time flowing".

  This in turn eventually led to the representation of time as a string of numbers and the representation of our present as a point that is moving along this array of numbers.

  Events and processes are no longer only placed in the arena of a 3-dimensional space in which objects may or may not move about, but also into a further dimension of time. Even objects at rest spatially appear to move along the time dimension with constant rate. And all objects move in the same forward direction. Newton writes in his principia: "Absolute, true and mathematical time, of itself, and from its own nature, flows equably without relation to anything external".

  Space and time together are the fixed absolute arena into which all events and processes and objects of matter are embedded. The geometrical representation of this is a space-time coordinate system in which each point marks a localized event, point-objects perform paths that progresses in the positive time-direction [Fig 8].

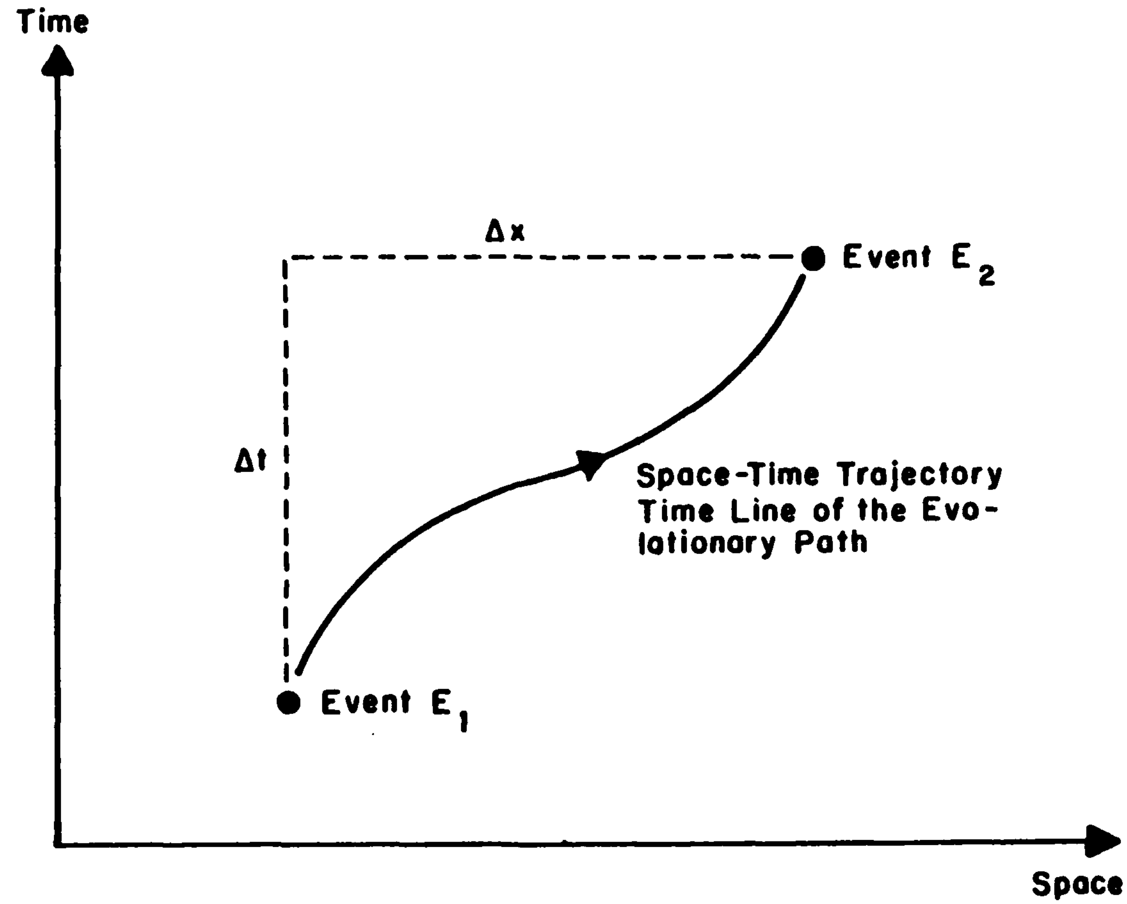


Fig. 8 THE NEWTONIAN WORLD VIEW
SPACE AND TIME ARE ABSOLUTE AND REPRESENT THE ARENA INTO WHICH PROCESSES AND EVENTS ARE EMBEDDED.

- MEASUREMENT OF TIME

The notion of absolute time and the understanding that the description of motion requires an assessment of the motion in time created the need for a way to measure accurately the "flow of time". This undertaking brings us immediately back to question this concept of absolute time. There does not seem an absolute "thing": Time, to which we could put our measuring stick to see how long it is. We are referred back to our immediate notion of time which is related for us to the duration of objects or the interval between recurring events. The daily recurrence of rising and setting of the sun, the phases of the moon, the shift of the culmination height of the sun and associated seasonal changes, the annual recurrence of certain stellar constellations on the night sky, all these are events which we perceive as occurring with regularity and they are therefore suitable to use as indicators of lapse of equal time intervals.

Until the middle of this century when the time keeping accuracy of atomic clocks began to exceed that of the Ephemeris time keeping the periodicity of the daily and annual earth motion was the principal means of timekeeping. The need to clock time intervals of shorter duration lead already in antiquity to construct "clocks". Mechanical devices based on a constant rate of flow of water (clepsydra) were perhaps used as early as

135 BC in Alexandria in order to keep time on cloudy days when the sun dial was of no use. Modern clock work technology began with Galilei's observation (1581) of the independence (in limits) of the period of a pendulum of given length on the amplitude, this providing a suitable means for timekeeping. Christian Huygens built the first pendulum clock (1651).

The pendulum provides a simple enough mechanism to demonstrate the basic principle of all accurate time keepers including the modern Caesium and Rubidium standards with accuracies of 10^{-13}. [see appendix A for some details]

The heart of any time keeping device is a high-Q resonator. Resonators like the pendulum can undergo oscillations. They can be forced to oscillate by a driving oscillator at any frequency ω of the driver. However the resonator responds best, i.e. gives large amplitudes of oscillations when the driver frequency ω is near the "resonance frequency ω_o of the resonator. The resonance frequency is the natural frequency at which the resonator would oscillate in the absence of the driver and all losses. The amplitude of the resonators shows the characteristic "resonance curve" [Fig 9] if the driver frequency ω is varied around ω_o. The resonance curve is the narrower around ω_o, the smaller the "dlamping" in the system is, the smaller the undesirable energy losses are, or the larger the Q-value is. If the resonance curve is narrow than the resonator provides the means for an accurate clock. The resonance curve behavior assures that the clock will only deviate from the accurate measure, given by ω_o, by a small amount. Any significant extraction of energy from the resonator by for instance making it drive a clock counter would constitute loss and spoil the resonance curve by making it wider, thus causing loss of accuracy. This is avoided by having as the driver oscillator an oscillator whose frequency is adjustable in the vicinity around ω_o [Fig 10]. In a feedback loop one measures the response amplitude of the resonator, which can usually be done with very little energy extraction, and feed that information back to the driver steering the driver frequency to stay near the maximum response of the resonator at ω_o. This control circuit can be arranged so that the system stays near ω_o with great stability. A signal from the driver oscillator is then used to drive the clock counter which basically counts the number of elapsed oscillations ω_o, which is taken as measure of elapsed time. Fig 11 shows the historical development of timekeeping accuracy. Modern Cesium clocks reach an accuracy of 10 nanoseconds per day which is about 10^{-13} seconds per second.

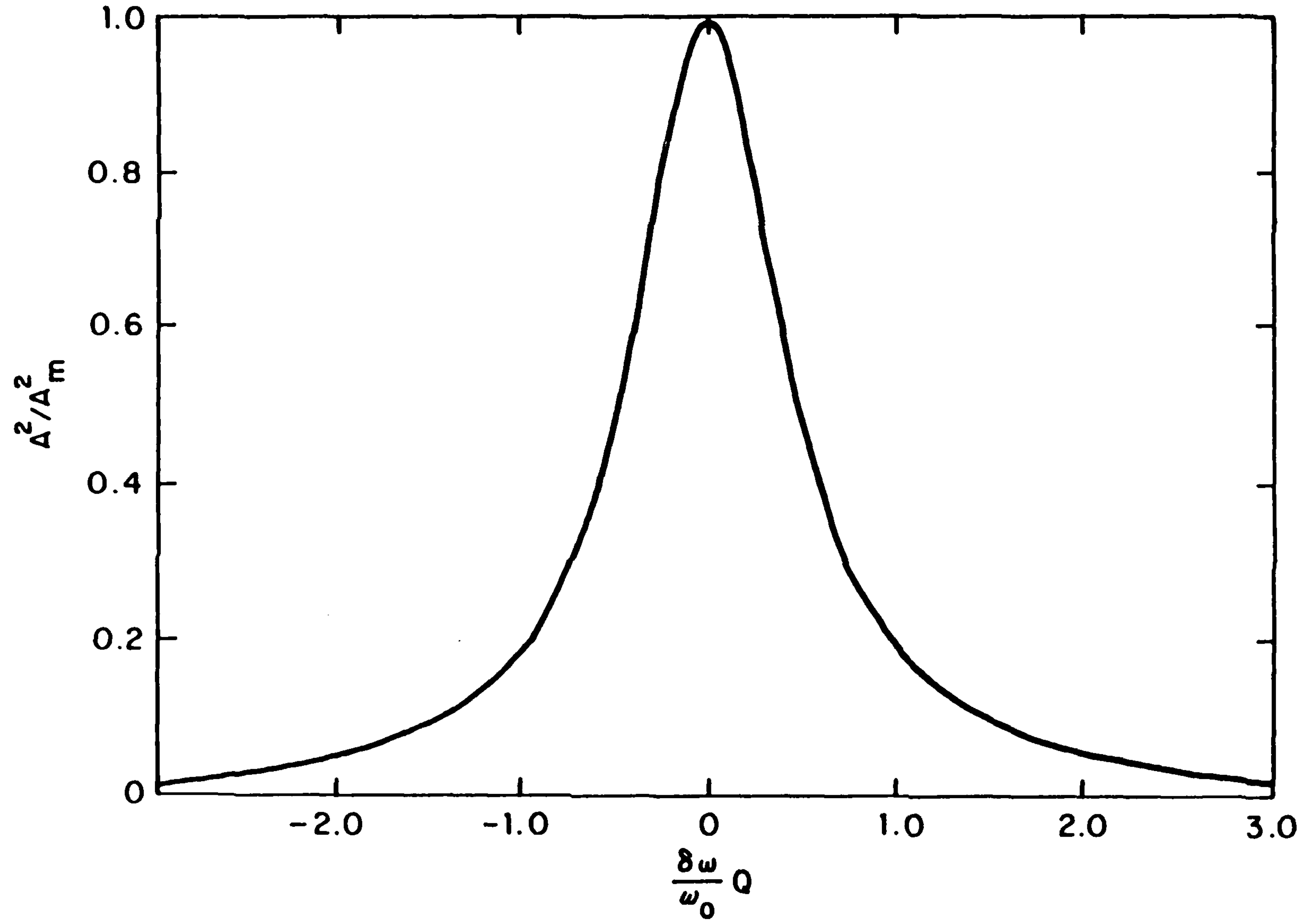


Fig. 9 UNIVERSAL RESONANCE CURVE OF FORCED HARMONIC OSCILLATOR

ω_0 = frequency of undamped oscillator; $Q = \frac{\omega_o}{\gamma}$ Quality factor; γ = damping constant; $F_o/m\gamma\omega_o$ maximum amplitude when driven at resonance. The approximation is valid for $\frac{\delta\omega}{\omega_o} < Q$

SCHEMATICS OF TIME KEEPER

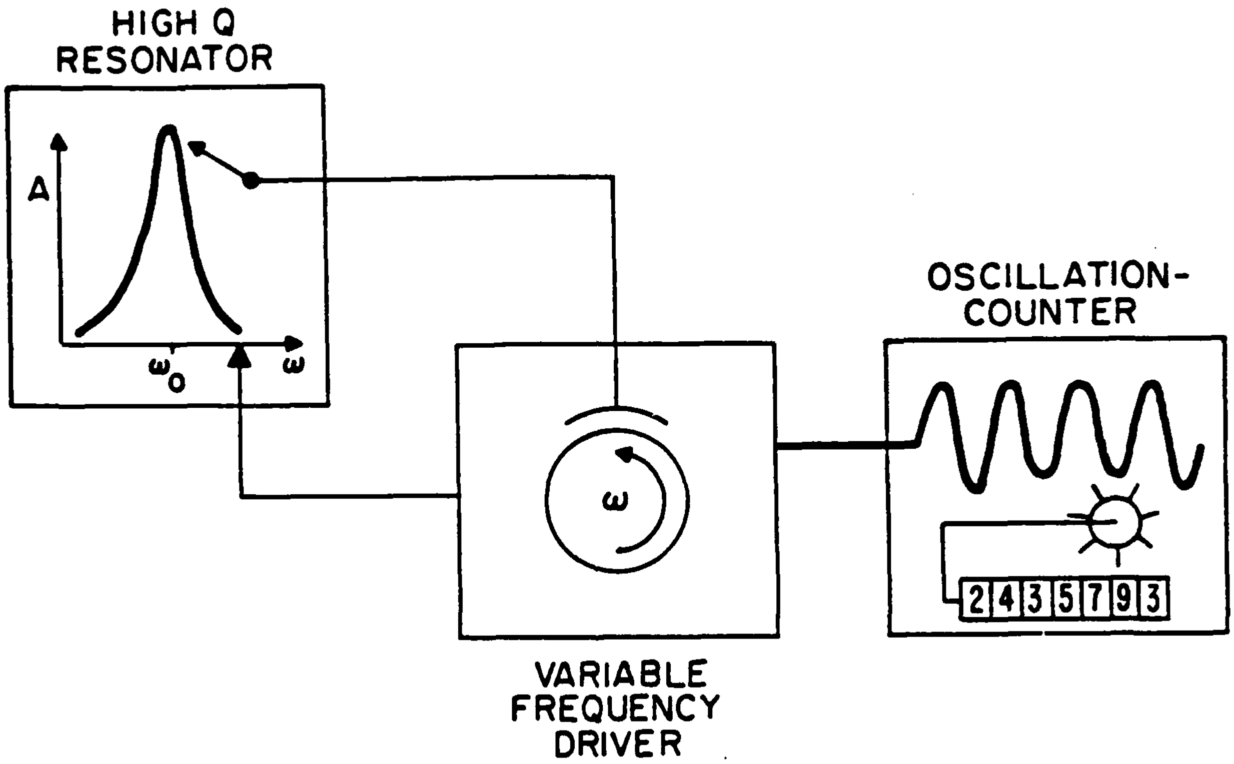


Fig. 10 Schematics of Time Keeping Devices

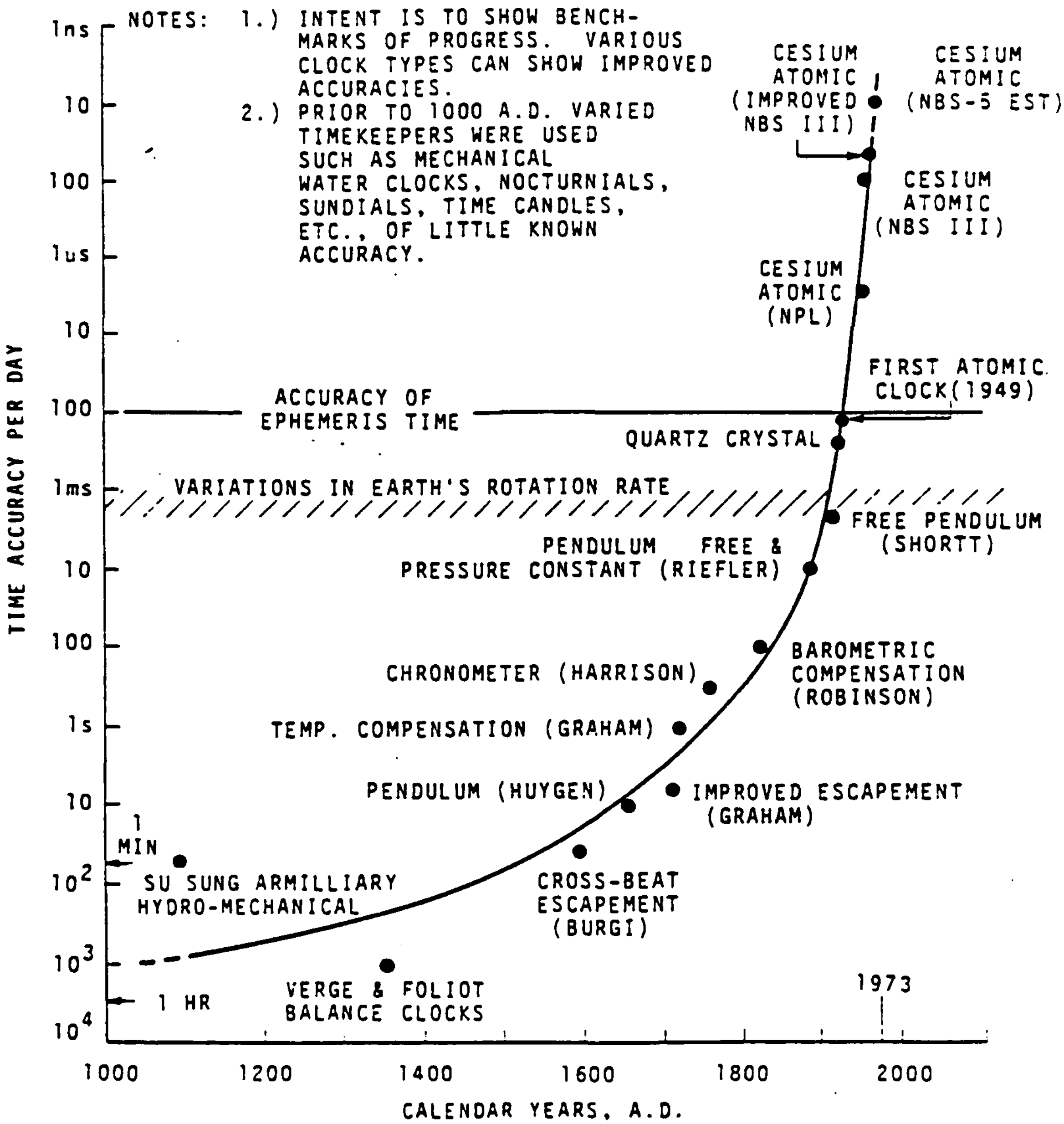


Fig. 11 Historical develoopment of timekeeping accuracy

10. RELATIVITY OF THE TIME CONCEPT IN EINSTEIN'S SPECIAL THEORY OF RELATIVITY

- Overwhelming experimental evidence about the nature of electromagnetic waves and the propagation of light gained around the turn of this century led Einstein in 1905 to develop his special theory of relativity. The principles of this theory pervade all physics with the recognition that signals require a finite length of time to propagate from the object to its observer. After the development of Einstein's theory a staggering amount of knowledge has been accumulated and built on this theory so that today almost the whole structure of physics is in some way affected by it. So far there has not been found any evidence against this theory. And yet some of the predictions made by the theory are most peculiar when compared with our every day experience. These peculiarities arise with respect to the space time coordinates of an event as seen by different observers who are in a state of relative motion with respect to each other. (Observer on the ground and the other on a fast train looking at the same event).

10.1 Inconsistencies in the Structure of Physical Theories Around 1900

- The laws of physics known around 1900, just prior to Einstein's publication of his special theory of relativity were characterized by a number of features which were inconsistent.

(a) Classical Newtonian laws of mechanics were known to be identical in coordinate reference frames which move with respect to each other with constant velocity.

This finding rests on the assertion that time as measured in two such reference frames O and O′ is identical: $t = t'$, and that a particle that is seen to move with velocity v in the stationary frame O will be seen to move with velocity $v' = v-u$ in the frame O′ that moves with velocity u with respect to O. This observation is referred to as the Galilei invariance of Newtonian physics [Fig 12].

(b) Optical and electromagnetic phenomena had found a triumphant unification of description in the famous set of Maxwell's equations of electromagnetism of 1864. These equations appeared to describe accurately all optical, electric and magnetic phenomena. Light, in this theory, appears as a wavemotion.

There were several difficulties with this state of affairs, however.

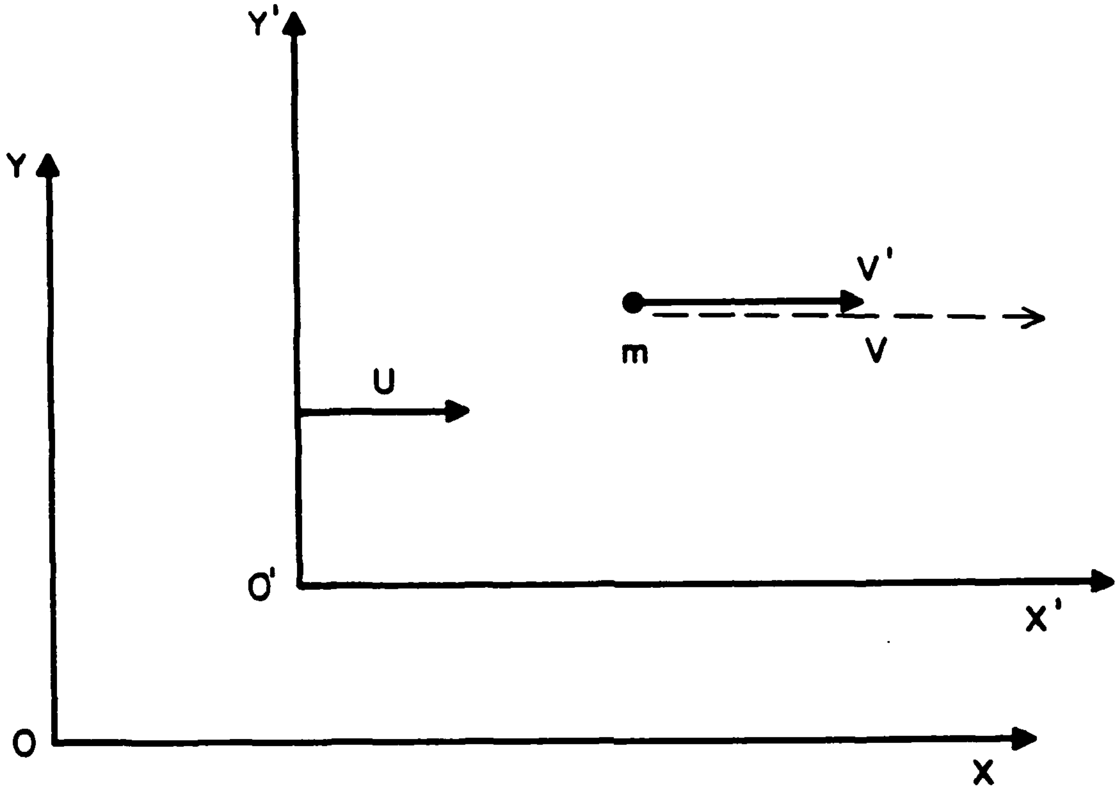


Fig. 12 CLASSICAL GALILEAN PRINCIPLE OF RELATIVITY
The motion of particle m is observed in reference frame O to have velocity V. In reference frame O' which is moving with velocity U with respect to O the particle is seen to have velocity V' = V - U.
The time in both referrence frames is identical.

(i) Maxwell equations are not invariant under Galilei transformations. They change their form, i.e. they describe different physics when one follows the prescription how to write the laws for a reference frame that is moving with a constant velocity.

(ii) In classical mechanics the description of wave motions involves the implicit assumption of a medium that is carrier of the waves. It was natural to assume the existence of some kind of substrate medium as mediator of the electromagnetic waves and light. This "ether", however had to be assumed to have rather strange mechanical properties, like zero mass density but non-zero elastic properties.

(iii) The ether provides a preferred reference frame. Light propagating in this preferred reference frame has the vacuum light velocity c. In every other frame, that is moving with some velocity u with respect to this ether the velocity is not c but c-u. These frames are not equivalent frames. This state of affairs sets the electromagnetic phenomena apart from all other physics which was found to be Galilei invariant.

(c) In order to avoid the chism between electromagnetic phenomena and all other physics a number of additional assumptions were proposed which were designed to put electromagnetic and mechanical phenomena on the same footing. All these assumptions were rather ad hoc and did not gain acceptance.

A number of important experiments eventually led to the abandonment of these hypotheses and paved the way for Einstein's radical revision in his special theory of relativity.

Among these experiments the most important were three.

(i) The aberration of star positions during the course of a year. This long known phenomenon of the slight variations of the position of stars during a year is due to the motion of the earth around the sun with a speed of $10^{-4}c$ which causes angular variations of 10^{-4} radians. This observation cannot be explained under the assumption that the light velocity is solely determined by the ether medium.

(ii) Fizeau's experiment on the velocity of light in moving fluids (1859) showed that the so called ether dragging hypothesis was not tenable.

(iii) The most famous is the Michelson-Morley experiment which was designed to detect the motion of the earth relative to the preferred reference frame "ether" which would be at rest and in which the velocity of light is c. The experiment gave a completely negative result. No such motion effect was detected. Today we see this experiment as the most definite statement that establishes the principle of relativity.

10.2 Einstein's Theory of 1905 - Relativity of Space and Time

- Einstein, in 1905 boldly proposed to abandon the concept of ether and accept the fact that light propagates through vacuum, that this vacuum is really empty, that light does not need a medium that "waves" and must be endowed with strange properties. With no ether frame, the only frame of reference of significance to an observer is his own fixed reference frame. The light velocity in any of these equivalent reference frames of observers at rest or constant velocity with respect to each other is c as found in the Michelson-Morley experiment.

Einstein postulates that:

Principle of Relativity

I. The laws of electromagnetic phenomena as well as the laws of mechanics are the same in all inertial reference frames, which can move with respect to each other with constant velocity.

II. The velocity of light is independent of the motion of the source, and is equal to c in vacuum.

These postulates required that either Maxwell equations would have to be modified to make them Galilei invariant if one wishes to retain the Gailelei principle of relativity or one would have to give up the Galilei invariance as formulation of the principle of relativity which would require that Newtonian mechanics had to be modified.

Einstein is led to the second of these alternatives by a careful analysis of the meaning of simultaneity of separate events.

The principle of Galilei invariance assumes that the time in different reference frames O and O′ is the same: $t = t'$. This assumption is based on the Newtonian notion of an absolute universal time. In the context of Galilei invariance this implies that it is always possible to have synchronized clocks established in all reference frames. Einstein shows that the process of actually synchronizing clocks between reference frames that move with constant speed with respect to each other is not trivial and actually leads to disagreements about the simultaneity of two events as seen by observers in the two reference frames.

Einstein shows that because of the finite propagation speed of light signals, the attempt to establish simultaneity leads to a mixing of time and spatial information. Using the fact that light propagates with velocity c regardless of the reference frame, as established in the Michelson-Morley experiment, Einstein derives the famous LORENTZ transformations which relate the space and time coordinates of events as observed by another observer who is moving with constant velocity relative to the first one and is observing the same event.

These transformations are different from the Galilei-transformations and they do not give the same time for an event as seen by both observers. Time and position coordinates become intermingled in a particular way. In Fig 13 this situation is illustrated. In a reference frame O distances are measured as Δx and time intervals as Δt. For simplicity only one space dimension is considered. Events E_1 and E_2 are located according to their position and time of occurrence as points in a space-time diagram. The intervals Δx and $c \cdot \Delta t$ are read off as the distances along the x-axis and ct-axis of the two event points E_1 and E_2 (it is convenient to use the distance ct that light travels during t rather than t). The same events can be observed by an observer O′ who is moving with velocity U along the x-axis of O. Observer O′ carries his own space-time diagram with axis x′ and ct′. In order to show both in the same diagram the x′-ct′ axis system must be drawn at an angle as in the figure, and the distance $\Delta x'$ between the two events E_1 and E_1 as seen by O′ and their time separation $c\Delta t'$ can be determined as shown by determining the distances parallel to the x′ and ct′ axis. The units in O and O′ in this diagram are not the same and so this representation is merely a schematic. If one uses proper units along x′ and ct′ axis,

$$x'\sqrt{\frac{1+v^2/c^2}{1-v^2/c^2}} \text{ and } ct'\sqrt{\frac{1+v^2/c^2}{1-v^2/c^2}} \text{ respectively,}$$

then one can read off the diagram two of the most surprising consequences, the time dilation and length contraction. (Fig 14). Signals from a clock at rest in the moving frame O′ which are given in time intervals $\Delta t'$ are observed by O to be longer according to

$$\Delta t = \frac{\Delta t'}{\sqrt{1-v^2/c^2}}$$

Similarly, length intervals $\Delta x'$ at rest in frame O′ appear to O shortened to

$$\Delta x = \Delta x'\sqrt{1-v^2/c^2}$$

For the time comparison the stationary clock in O′ appears to O as moving along the ct′ axis. Successive signals appear to come from different places and they can be observed directly and compared with synchronized clocks along the track of O′.

The yard stick comparison is somewhat more difficult. The yard stick $\Delta x'$ is at rest in O′. Observer O must

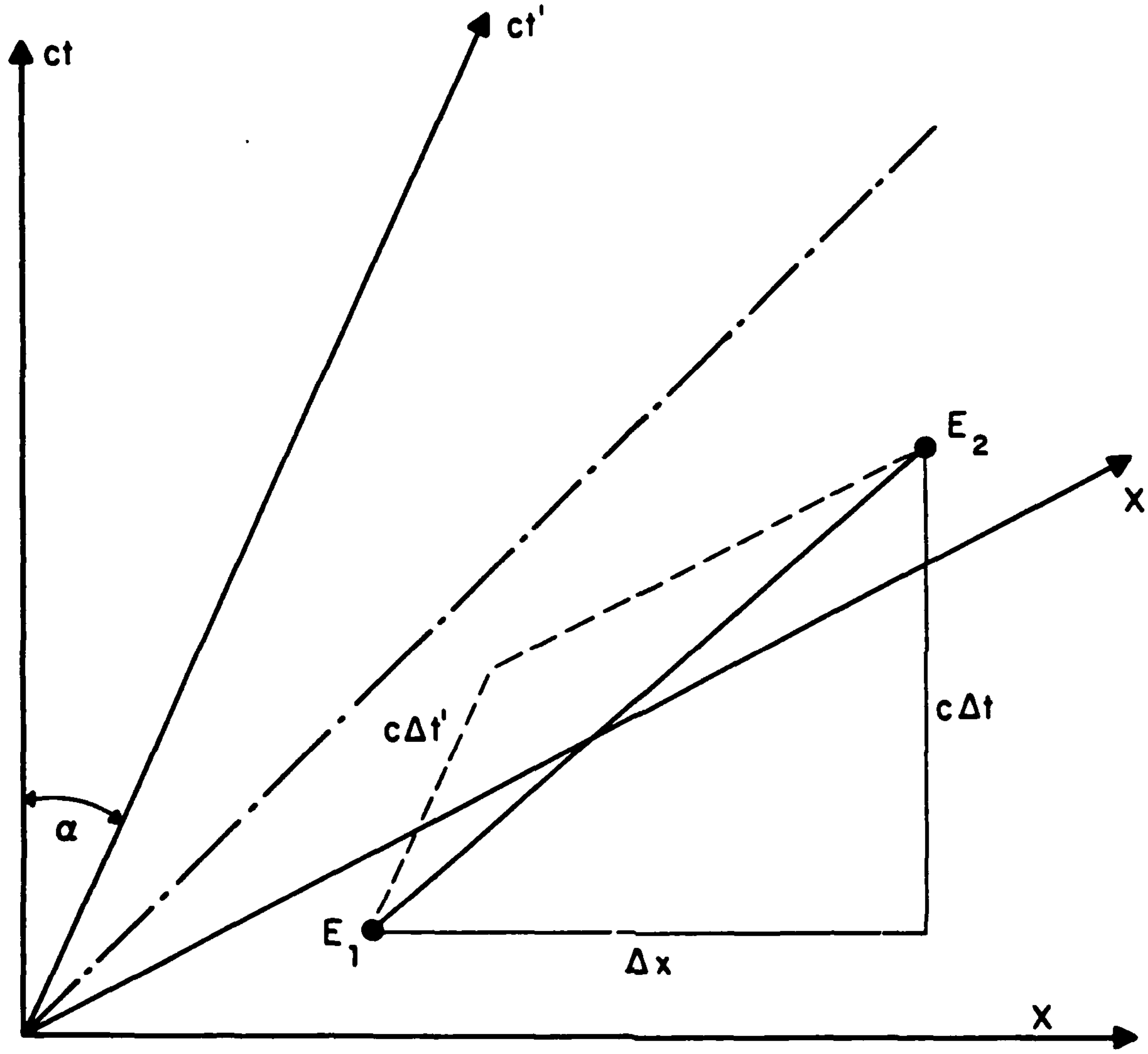


Fig. 13 LORENTZ TRANSFORMATION FOR SPACE AND TIME INTERVALS BETWEEN TWO EVENTS E_1 AND E_2 AS OBSERVED BY OBSERVER O (x ; ct) AND AN OBSERVER O′ (x′ ; ct′) WHO IS MOVING WITH VELOCITY U WITH RESPECT TO O IN x-DIRECTION THE TWO COORDINATE SYSTEMS CAN BE PLACED INTO SINGLE GRAPH BY EMPLOYING NON-RECTANGULAR COORDINATES FOR O′ SUCH THAT $\tan\alpha = U/c$

$$\Delta x' = \frac{\Delta x - U\Delta t}{\sqrt{1 - U^2/c^2}} \qquad c\Delta t' = \frac{c\Delta t - \Delta x\, U/c}{\sqrt{1 - U^2/c^2}}$$

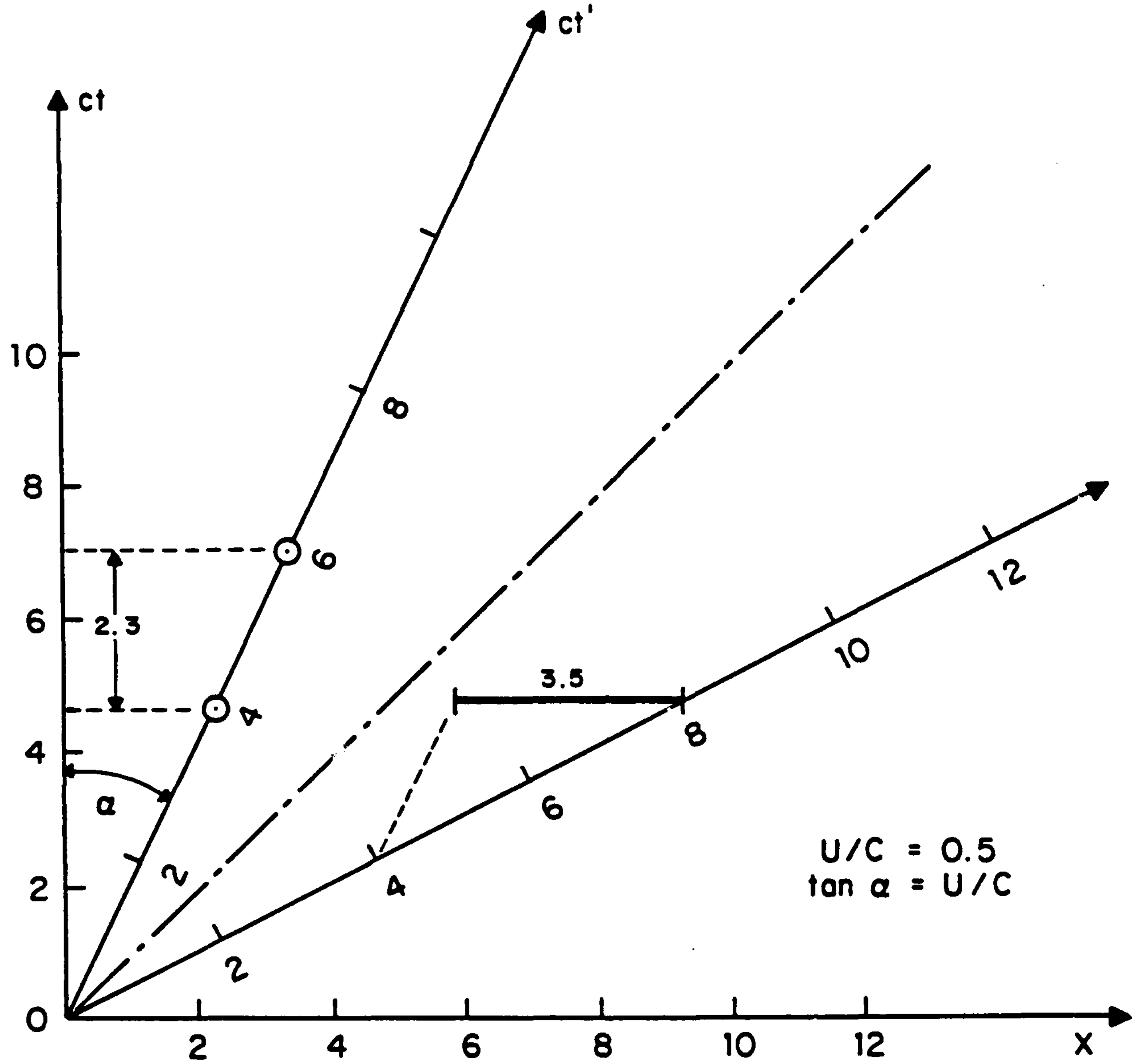


Fig. 14 TIME DILATION AND SPACE CONTRACTION IN SPECIAL RELATIVITY

(a) TIME DILATION: SIGNALS FROM A CLOCK AT REST IN THE MOVING FRAME O' WITH TIME INTERVALS Δt' ARE OBSERVED BY O TO HAVE INTERVALS

$$\Delta t = \frac{\Delta t'}{\sqrt{1 - U^2/c^2}}$$

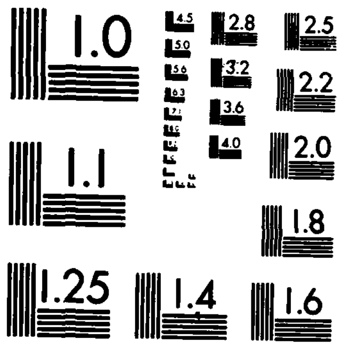
(b) SPACE CONTRACTION: A LENGTH Δx' AT REST IN O' IS OBSERVED BY O TO HAVE LENGTH

$$\Delta x = \Delta x' \sqrt{1 - U^2/c^2}$$

THE SCALES IN THE TWO COORDINATE SYSTEMS MUST BE CHOSEN APPROPRIATELY TO CONVEY THE PROPER RELATIONS

AD-A233 001 THE MEANING OF TIME AN INTRODUCTION INTO PHILOSOPHICAL BIOLOGICAL AND PHY..(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA K E WOEHLER 15 JAN 91 NPS-PH-91-005 XN-NPS

UNCLASSIFIED NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

some how manage to lay his own yardstick next to that of O′ and make a simultaneous matching up of the end-points. What is simultaneous to O is not simultaneous to O′. To O′ the coincidence of the two yardsticks happens with some time delay. At first, coincidence at the right edge is made and a little later the left edge is lined up.

10.3 Failure of the Notion of Certain Past and Uncertain Future in Special Relativity

- Mapping the reference frames of two observers O and O′, which are moving with respect to each other along the x-direction with a velocity v into one diagram as in the figures 13 and 14 shows another most surprising feature of special relativity. We can see that the "Now" (points on the x-axis) at all positions x > o in O lies in the past of O′ [Fig 17].

  Furthermore, some events in the future of O lie in the past of O′.

  For instance, event A lies in the "certain past" for O′ yet in the uncertain future for O.

  Similarly, event B lies in the certain past for O but in the uncertain future for O′.

  The events in the past of O′ are certain to O′. Observer O has no knowledge of them yet because signals from these events take time to reach O. These events are not really

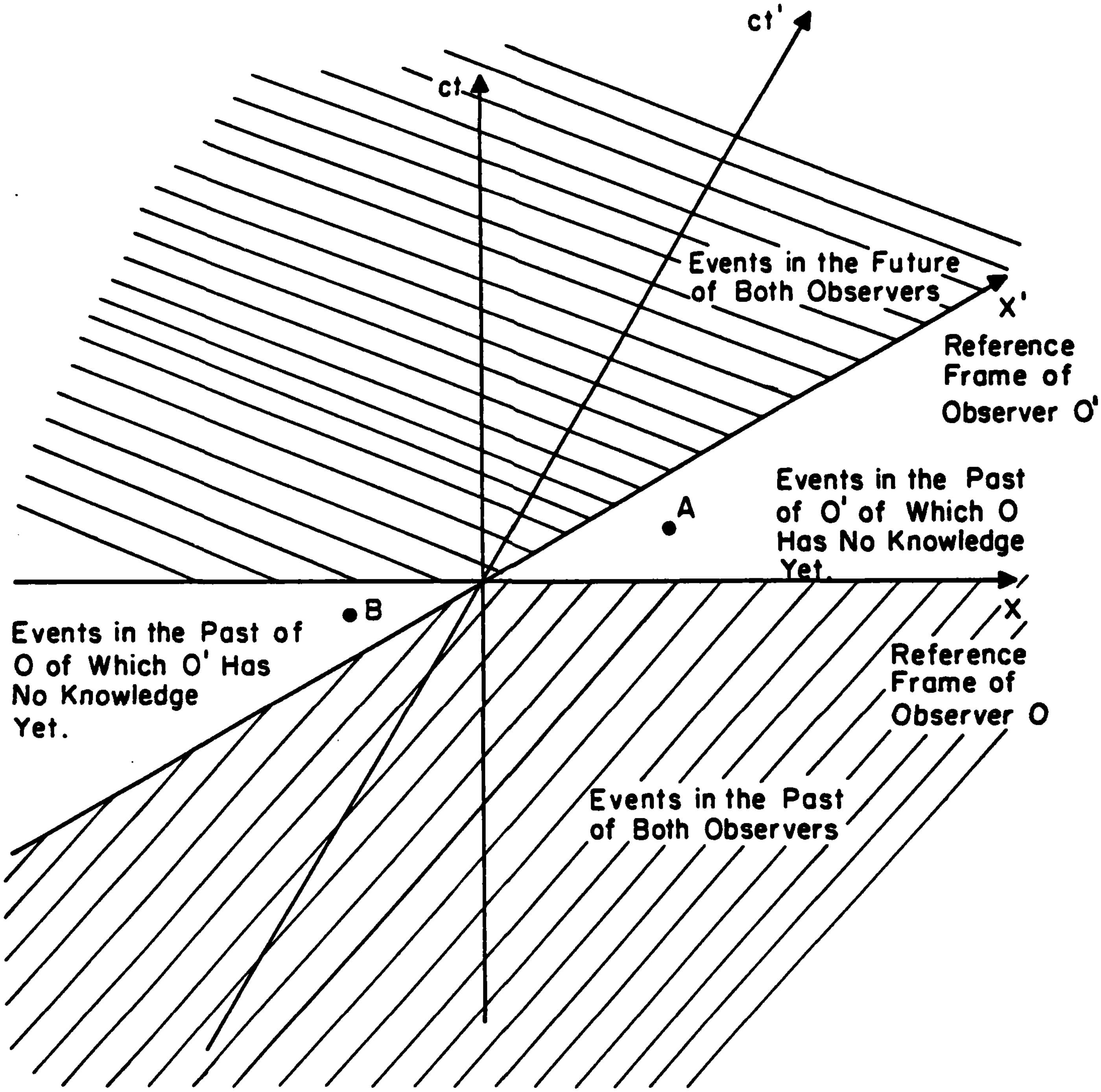


Fig. 15 RELATIVITY OF
PAST AND FUTURE IN
SPECIAL RELATIVITY

uncertain for O. They will happen for O eventually when the signals arrive. Observer O cannot actually influence what already happened in O′.

The diagram is for an observer O′ that moves with v = 0.5 c with respect to O. Events that lie in the "uncertain future for both observers O and O′ could never the less be in the past of an observer O″ who moves with ¾c with respect to O. In that case the angle between the axis ct″ and x″ is smaller than the one between ct′ and x′ and some of the events in the shaded region "above" the x′-axis with x′>o could lie "below" the x″-axis.

What does all this contribute to our understanding of the meaning and the nature of time?

The, for us so peculiar, properties of relative time, i.e. time relative to the relative state of motion in Einstein's theory has its origin in the acceptance of the fact that synchronization of clocks, moving relative to each other, involves the propagation of light signals, and that simultaneity has not a unique meaning for every observer.

This does seem to say that "absolute time" cannot be used in a practical operational way. The concept is not a useful concept. It appears, that time can practically only be defined in terms of a network of clocks with certain transformation rules given, how the readings of clocks, which move with respect to each other, must be related, in order to have consensus about spatio-temporal relations of events.

10.4 Other Consequences of the Postulates of Special Relativity

- More to complete the story for the reader who is not familiar with Einstein's theory than to say more about time in this theory, some of the important consequences shall be briefly addressed. As was mentioned, Einstein's postulates broke with the classical tradition of the Galilei prescription how to relate space, time and velocities for two observers who are moving with constant velocity with respect to each other. He set the Lorentz transformation in its place which follow directly from his postulates.

  Besides the time dilatation and space contraction mentioned there are a number of other consequences which follow from requiring the Lorentz transformation as the proper way of relating the observations of two observers. The most important consequences shall be briefly listed here.

  (a) Similar to distances and times in the two reference frames also velocities in O and O′ are related differently than as described by Galilei transformation.

One surprising result is that even if the two observers are moving toward each other, both with velocity near c, both will observe the other as moving at speed less than c.

(b) The law of conservation of momentum in a collision of two mass points moving at velocities near c toward each other, collide and then move apart, leads to the conclusion that the concept of mass must be given a relative meaning with values of the mass m(v) of a particle depending on the state of relative motion of the mass with respect to the observer.

$$m(v) = \frac{m_o}{\sqrt{1-v^2/c^2}}$$

Where m_o is the mass an observer would find when the mass is at rest in the laboratory, v is the speed of the mass.

The relation indicates that masses appear to increase their mass with velocity. This does not mean that material is added to the object. It will have the same number of atoms. It merely says that our attempt to determine the mass of a moving object does not give the same result as a measurement of the mass at rest. It means that our measurement of mass of a moving object requires that we let it collide with another mass and from the outcome of this collision together with the principal of momentum conservation, we determine the mass. But that observation requires observation of velocities i.e. determination of distances and time intervals. These, however must follow the Lorentz rules and therefore the mass determination becomes affected. Mass is not an absolute concept but is subject to the rules of the relativity postulates.

(c) The relativity of the mass automatically affects the equations of Newtonian mechanics into which the mass of the moving object enters prominently. The laws of mechanics become modified. After this modification these laws do not remain the same if one makes the transformation to another reference frame via the old Galilei prescription. Instead, the laws do remain the same if one makes this transformation using the new Lorentz transformation.

(d) An important principle of classical mechanics is that any work done on a particle of mass m by exerting accelerating forces on it, leads to an equivalent increase of the particles "energy of motion" (kinetic energy).

The fact that the mass of the particle would increase as its speed increases must now be taken into account, which leads to a revised form of the so called work-energy theorem. The most dramatic result of that revision is, that even a mass at rest must be attributed with having a "rest energy" $E = m_o c^2$ where m_o is the mass one would measure in the state of rest.

Even for very small masses m_o this amount of energy is of staggering magnitude, because c is such a large number.

This famous mass-energy equivalence relation of Einstein was a very important factor in the subsequent understanding of processes inside the tiny atomic nucleus and technically led to the atomic bomb and nuclear power reactors.

11. INVARIANCE OF THE LAWS OF PHYSICS UNDER TIME INVERSION AND ITS SIGNIFICANCE

11.1 The Laws of Physics

All laws of physics that describe dynamical processes of physical systems are deriveable from a variation principle which is called the principle of least action. All processes in nature proceeds in such fashion that when one considers all possible path ways along which the system could have proceeded, the one actually chosen by nature is the one for which the change of "action" is the least. For each physical system there is a characteristic function.

$$L\{\phi\} \equiv L\left\{\phi(x,t)\,;\,\frac{\partial\phi}{\partial x}\,;\,\frac{\partial\phi}{\partial t}\,;\,x;\,t\right\}$$

the so called Lagrangian for the system. This Lagrangian is a function of quantities $\phi(x\ t)$ that describe the state of the system at position x at time t, and the rates of change of these state variables.

The action for the system as it evolves beginning at time t_i until time t_f is

$$S = \int_{t_i}^{t_f} L\,dt$$

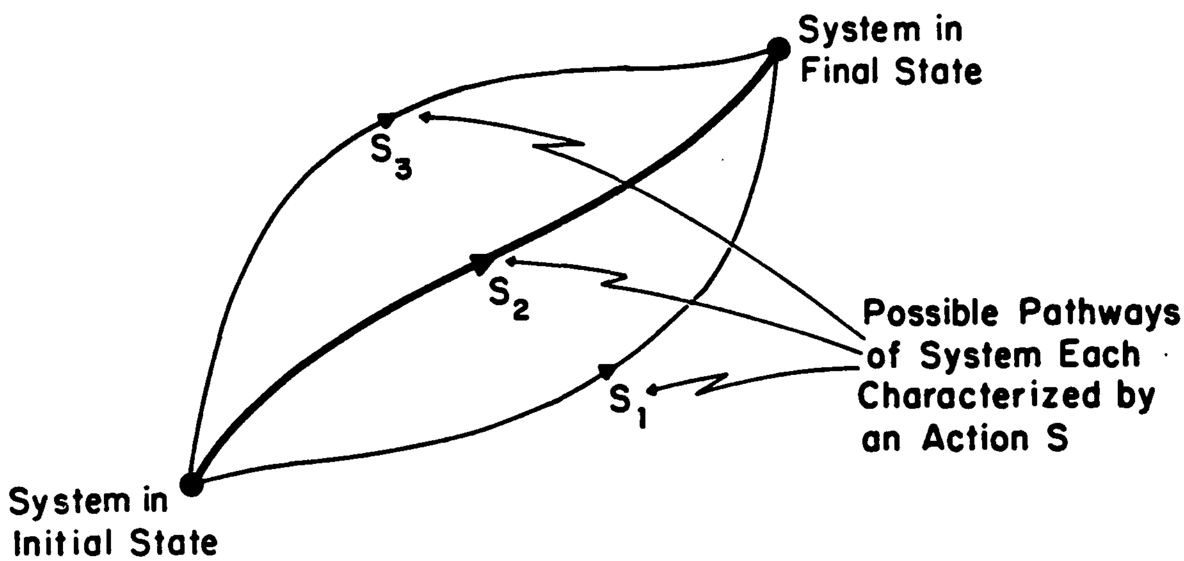


Fig. 16 PRINCIPLE OF LEAST ACTION
FOR EACH SYSTEM AN ACTION FUNCTION S IS DEFINED IN TERMS OF THE LAGRANGIAN FUNCTION L WHICH IS A FUNCTION OF THE SYSTEM STATE VARIABLES.

$$S = \int_{t_{final}}^{t_{initial}} L dt$$

THE CLASSICAL PATH WHICH THE SYSTEM TAKES FROM ITS INITIAL STATE AT TIME t_{init} TO ITS FINAL STATE AT t_{final} IS CHARACTERIZED BY BEING THE PATH OF LEAST ACTION.

The system evolves along that path for which S is a minimum. The physical law of motion of the system follows uniquely from this requirement. All information about the nature of the system is condensed in the one functional L $[\phi]$.

11.2 What is meant by Invariance of a Physical Law?

The physical laws derived from the respective functional $L[\phi]$ by applying the principle of least action have a structure which show some general characteristics of the system. One says: The Lagrangian and therefore the laws of motion that follow from it have certain "symmetries". These symmetries in the Lagrangian become apparent when one subjects the Lagrangian to a change of reference frame. Change of the coordinate system $x \to x'$ in which the system is being described, change of the reference clock used: $t \to t'$ and changes of the description of the state of the system: $\phi \to \phi'$ are such change of aspect under which the system is considered.

Concretely, the most common "transformations" of this sort are

(i) Shift of the zero of the clock used.

(ii) Shift of the origin of the coordinate system used.

(iii) Rotation of the axis of the coordinate system.

(iv) Change of state variables used to describe the state.

(v) Lorentz transformation to moving reference frame.

If one carries out such transformation, i.e. change of reference frame on the Lagrangian and on the equations of motion following from it then the so transformed Lagrangian may appear changed in its structure and the equations of motion may be changed. In that case these changed equations describe different motions, a different system. If however after such transformation the Lagrangian and the equations of motion look exactly like the original equations, then this transformed system has exactly the same behavior as the original one. The system and the equations describing it are "invariant" under these particular changes of reference. The system has a "symmetry" that lets it appear unchanged when viewed from the different reference frame.

One of the very important insights of modern theoretical physics is Noether's theorem (1901), which shows that there is a one-to-one correspondence between the invariance of laws of physics under certain groups of transformations of reference frame and associated conserved quantities which remain constant during the process described by the laws of motion.

Three of these correspondences are:

Invariance under:	Conservation of:
Change of time zero	Energy
Translation shift of coordinate system	Linear momentum
Rotation of coordinate system	Angular momentum

There are other kinds of transformation one can subject Lagrangian and equation of motion to, for which invariance exist but which have associated conserved quantities that are not as well known.

11.3 What is the implication of Time Reversal Invariance of all Known Physical Laws?

All laws of elementary dynamic processes as far as we know are invariant under timer reversal transformation (see appendix C for mathematical details).

If one replaces in the laws of motion in every place where the time t occurs:

$$t \rightarrow t' = -t$$

one finds that these laws appear unchanged in form. This means that the trajectory of the system that is time reversed is identical to the trajectory of the original system. The only effect of the time reversal is that the same motion runs in reverse sequence just like a movie run backwards. On the movie run backwards the events taking place are identically the same as in the forward movie, except the sequence of events proceeds in reverse order.

That the laws of motion in physics are invariant under time reversal means that to each process described by and possible under this law of motion, the reverse process is an equally possible process and must be expected to occur in nature.

That this should be so for all basic laws of dynamic processes is a most surprising statement and seems to clash with our immediate experience. We are aware of many processes in our life, the reverse of which does not occur in our experience. Never do we see an old man emerge from a grave and in the course of his subsequent life appear to look younger, get smaller and in the end as a baby disappear in a womans womb. The important point here is that, the time reversal invariance holds for all elementary dynamic processes. By this we mean Newton's laws of particle dynamics, the laws of electromagnetic phenomena, the laws of ideal fluid dynamics, and the laws of atomic physics.

However, whenever dissipative processes are included in these laws, i.e. processes that contribute to dispersion of the energy present in the system into forms that are not part of the same dynamic system or into forms of energy which are not incorporated fully into the level of description of that system, then one finds that the equations of motion are not time reversal invariant. These laws then have solutions and describe processes for which the reverse is not equally valid and possible process. The phenomena of the arrow of time make their appearance. In one of the next sections we will describe these problems in greater detail.

A more formal description of the time reversal invariance of some of the elementary dynamic laws of physics is given in appendix C.

11.4 P-Invariance, C-Invariance and the TCP Theorem

We need to look at a couple more transformations and nature's invariance under these.

Parity Transformation

If one looks at the motion of a particle in the x-y-plane with the particle moving counter-clock wise about the plane one can describe the motion in a so called right handed coordinate system as shown in Figure 17.

If one watches this coordinate system and the motion in a mirror one notices that the y-axis seems to have switched direction, turning the coordinate system into a left handed one and that the motion seems to run in opposite direction. The same left handed coordinate system is also obtained from the original coordinate system after a complete space reflection transformation, turning all axis into the opposite direction. One can see that this coordinate system (x″ y″ z″) can be obtained from the mirrored system (x′ y′ z′) by merely rotating the system by 180° around the y′ axis.

Whether one describes the motion of the particle in the right handed or left handed coordinate system makes not the slightest difference in Newtonian mechanics. Newtonian mechanics is invariant under parity transformation. The motion seen in the mirror is physically the same motion, follows the same law, as the one in the laboratory.

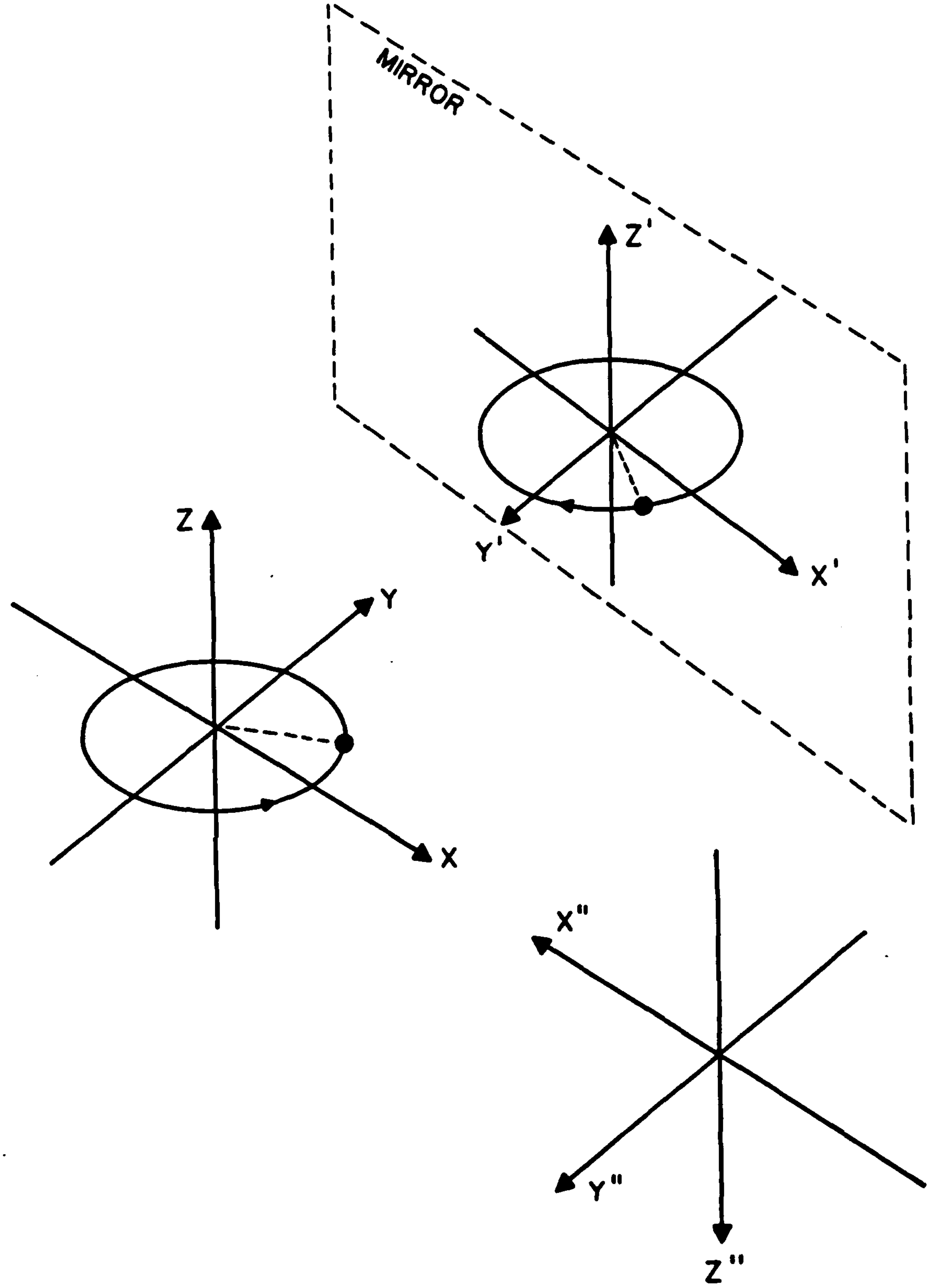


Fig. 17 MIRROR OR PARITY SYMMETRY OF PHYSICAL PROCESSES. THE MIRROR IMAGED MOTION IS THE SAME AS THE ORIGINAL MOTION DESCRIBED IN A LEFT HANDED SYSTEM.

Almost all other laws of physics have the same property.

There is an exception. The laws governing the weak interaction radioactive decays involving neutrinos do not follow the parity invariance.

This was shown by Mme. Wu and her collaborators in 1957 by measuring the direction of the emitted electrons relative to the orientation of the magnetic dipole moment of $_{27}Co^{60}$ nuclei when they undergo radioactive decay

$$_{27}Co^{60} \rightarrow {}_{28}Ni^{60} + e^{-} + \overline{\nu}$$

The cobalt nuclei represent little magnets as if a circular current were flowing in the nucleus. In our parity transformation picture one may think of the circular motion in the xy-plan as a current carrying wire. The magnetic field produced by that current points in the +z direction.

In radio active decay the cobalt nucleus emits an electron and antineutrino. Wu found that the electrons are <u>not</u> emitted symmetrically with respect to the plane of the current loop. Instead there is a preferred direction of emission that is related to the circulation of the current loop in the same way as the direction of advance of a left-hand screw is related to its rotation i.e. in the negative z-direction in the picture. In the mirror image the preferred direction of electron emission would be the same, the negative z-direction. But the circulation of the current loop appears to have reversed. The interpretation would be that the electron emission follows a right hand screw. Thus the description of this β-decay is not the same in the original and the mirror image.

This behavior is essentially due to the fact that the antineutrino has positive helicity. The antineutrino carries spin of 1/2 angular momentum units and its spin axis is always aligned with its direction of propagation like a right hand screw. The cobalt nucleus has its angular momentum aligned in positive z-direction in the experiment. The decay product nickel nucleus has angular momentum one full unit less than the cobalt nucleus. So the electron and antineutrino must together carry away one angular momentum unit. That requires that electron and antineutrino which both have spin 1/2 must have their spins aligned in the positive z-direction. But because the antineutrino has positive helicity its momentum must be directed in positive z-direction. In order to conserve linear momentum, the electron must move in negative z-direction, which is what is observed.

The mirror imaged process, if it occurred would have to have the electron moving in negative z-direction as in the previous case. But because the antineutrino has spin and momentum aligned it would also have to move downward and thus violating

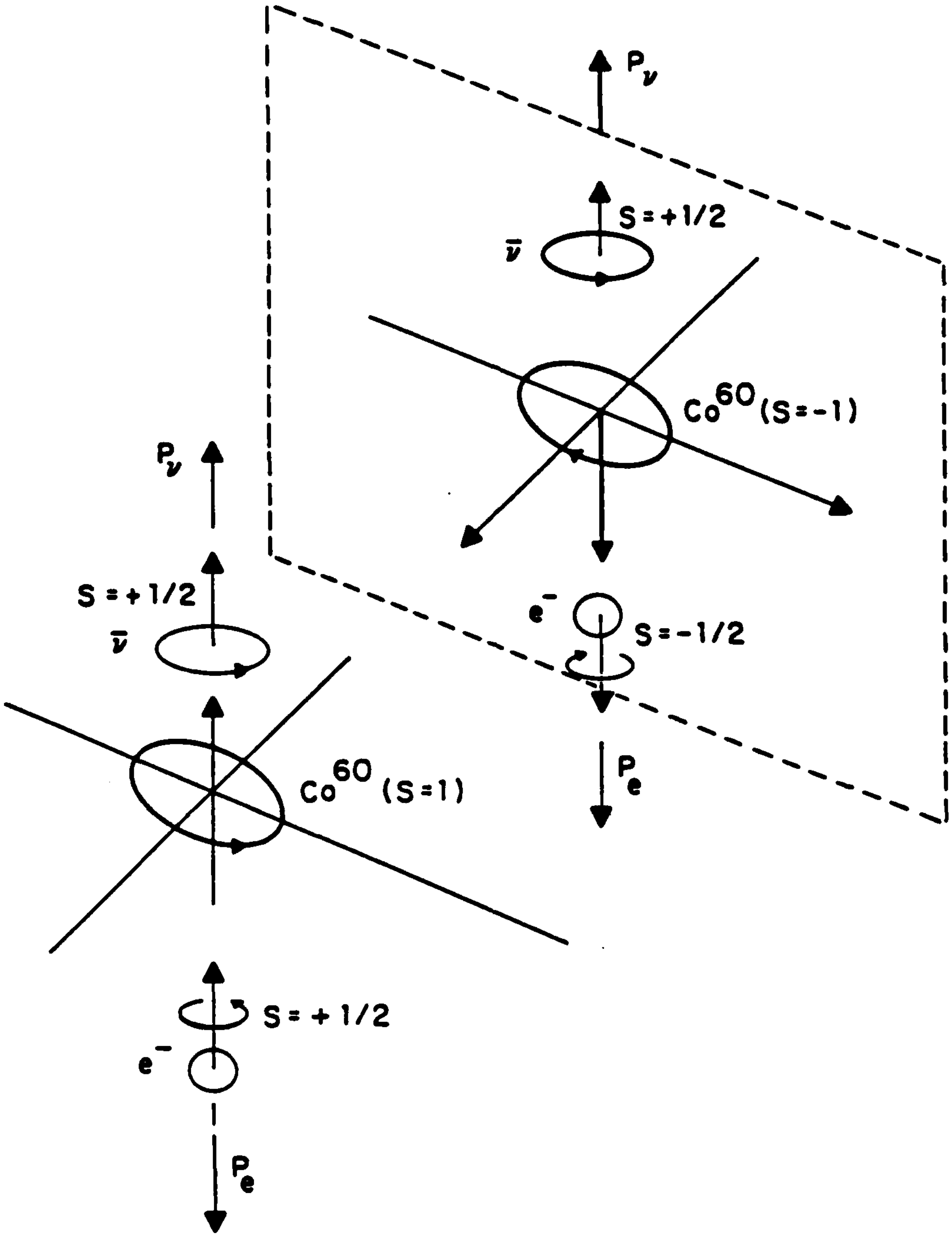


Fig. 18 PARITY VIOLATION IN β-DECAY OF POLARIZED Co^{60}

momentum conservation in this process [Fig 18].

The radioactive β-decay violates parity invariance.

No other process is known to do that, and it is one of the principle difficulties to incorporate this symmetry breaking into the full theory of all elementary particle processes.

Charge conjugation Invariance

One of the surprising consequences of the application of Einstein's special theory of relativity to nuclear physics and to the physics of elementary particles was the prediction that to every particle type found in nature there should be an "anti particle". An incling of this is seen when one looks at the relativistic energy equation:

$$E^2 = c^2p^2 = m_o^2c^4$$

or

$$E = \pm\sqrt{c^2p^2 + m_o^2c^4}$$

That would seem to say that for every particle of rest mass m_o and momentum p there are two possible versions, one with positive and one with negative energy. This notion of particles with negative energy led to great difficulties. With proper formulation of the quantum theory of relativistic particles it became clear that these negative energies are an artifact which disappears after proper quantum mechanical description, and that there is no mystery about the so called "anti matter". Proper relativistic quantum mechanical description of particles shows that this theory has a "symmetry". For each particle described by these theories another particle with equal mass, opposite electric charge, opposite helicity should exist that behaves in all respects similar to the original particle.

Soon after the discovery by Dirac that special relativity requires that such particle-anti particle pairs should exist, the first of these, the antiparticle to the electron, the positron was discovered. Presently some hundred particles are known, each having its anti particle.

The transformation that converts the equations that describe the behavior of a particle into the proper equation for the corresponding anti particle is called charge conjugation transformation. This name is because the dominant feature of the antiparticles is that they have the opposite charge.

Just like we represented the parity transformation by a mirror, we can think of a charge conjugation mirror in which each particle appears replaced by its charge and helicity opposite.

The charge conjugation invariance holds in most of the particle physics. Positrons behave in every respect like electrons except that the charge is reversed and so the electric and magnetic forces work in opposite direction on it. A atom built from an antinucleus and positrons would show the same optical and chemical behavior. The behavior in β-radioactive decays is the exception. These processes not only violate parity invariance but also charge conjugation invariance. In fact the same picture can be used to demonstrate both. The decay of the π-meson into a μ-meson and a μ-neutrino ν_μ is a suitable example (Figure 19).

If a π-meson is at rest i.e. P_z=o then the two emitted particles must be emitted in opposite directions to conserve linear momentum. The π-meson carries zero angular momentum. The μ-meson and neutrino are both spin 1/2 particles. In order to preserve angular momentum the spins of the emitted particles must be antialigned. Then there are two possibilities: (A) the spins are aligned with the momenta, or (B) the spins and momenta are anti aligned. As one can see, in the picture cases (A) and (B) are the mirror images of each other. In the mirror vectors pointing at the mirror appear inverted while circulation preserves its orientation.

The decay of the negative π-meson leads to

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$

The antineutrino $\bar{\nu}_\mu$ has positive helicity, with momentum and spin aligned. Like a right hand screw. One can see that only case (A) is possible. Case B cannot occur because helicity

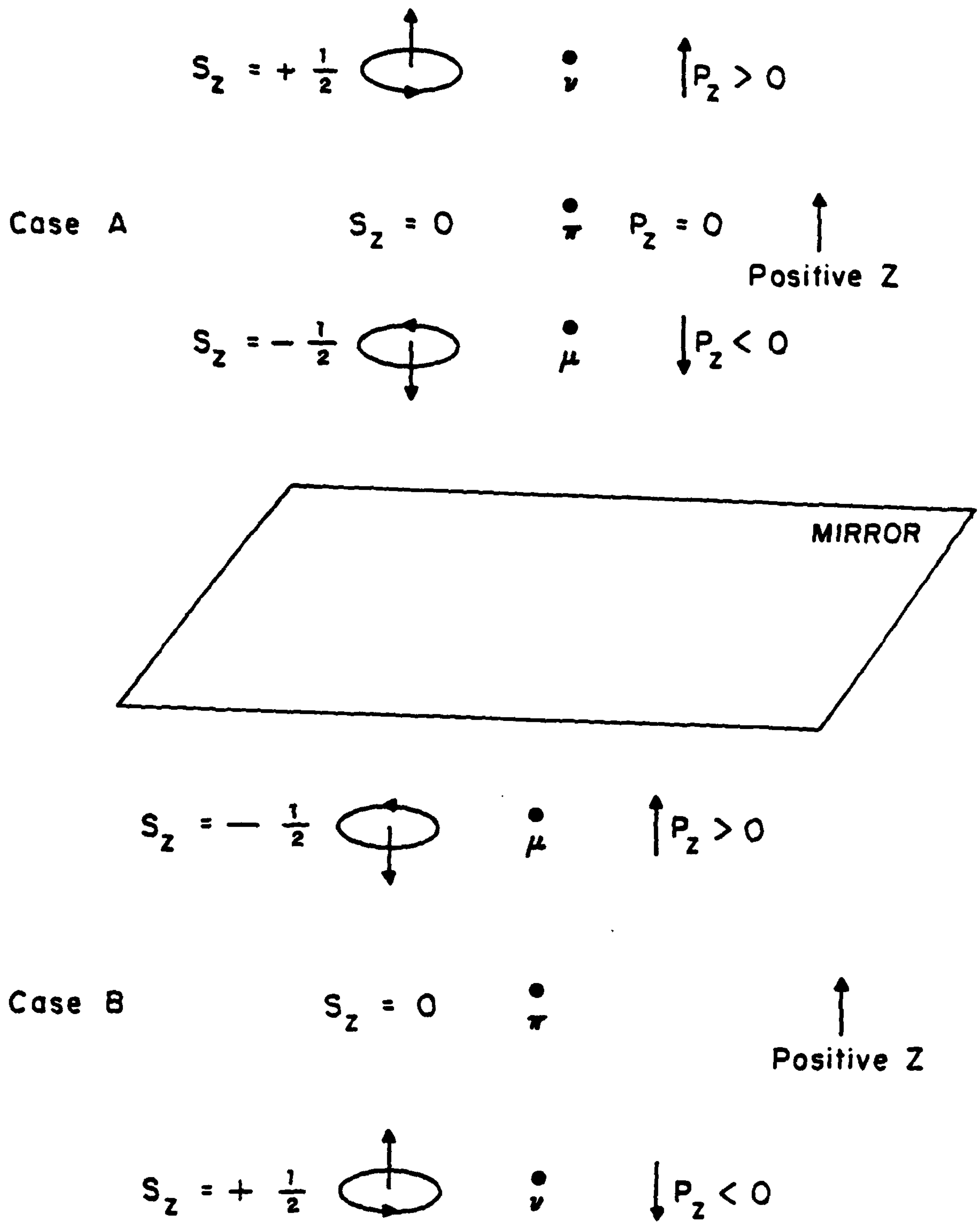


Fig. 19 CHARGE CONJUGATION AND PARITY VIOLATION IN π−MESON DECAY.

breaks the parity symmetry.

Now consider the charge conjugate decay

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$

This decay does not occur in the form (A) because the neutrino has negative helicity. So the handedness of the neutrino also causes non-invariance under charge conjugation. The charge conjugate process does occur in case (B).

When parity (P) and charge conjugation (C) were found not to be conserved in these so called weak interaction processes involving the neutrinos, the hope was that the combined operation CP would leave invariant the description of a system governed by this interaction. In the above case (A) which occurs for the π-meson one would first perform a charge conjugation $\pi^- \rightarrow \pi^+$ $\bar{\nu}_\mu \rightarrow \nu_\mu$ and $\mu^- \rightarrow \mu^+$. Then one would perform the mirror reflection which leads the process $\pi^+ \rightarrow \mu^+ + \nu_\mu$, in the case (B). This is indeed what is observed. this decay process violates parity conservation and charge conjugation conservation but obeys the invariance under CP transformation. Figure 20 shows another example of parity violation and CP conservation.

THE TCP THEOREM

The question whether there might be processes that do violate

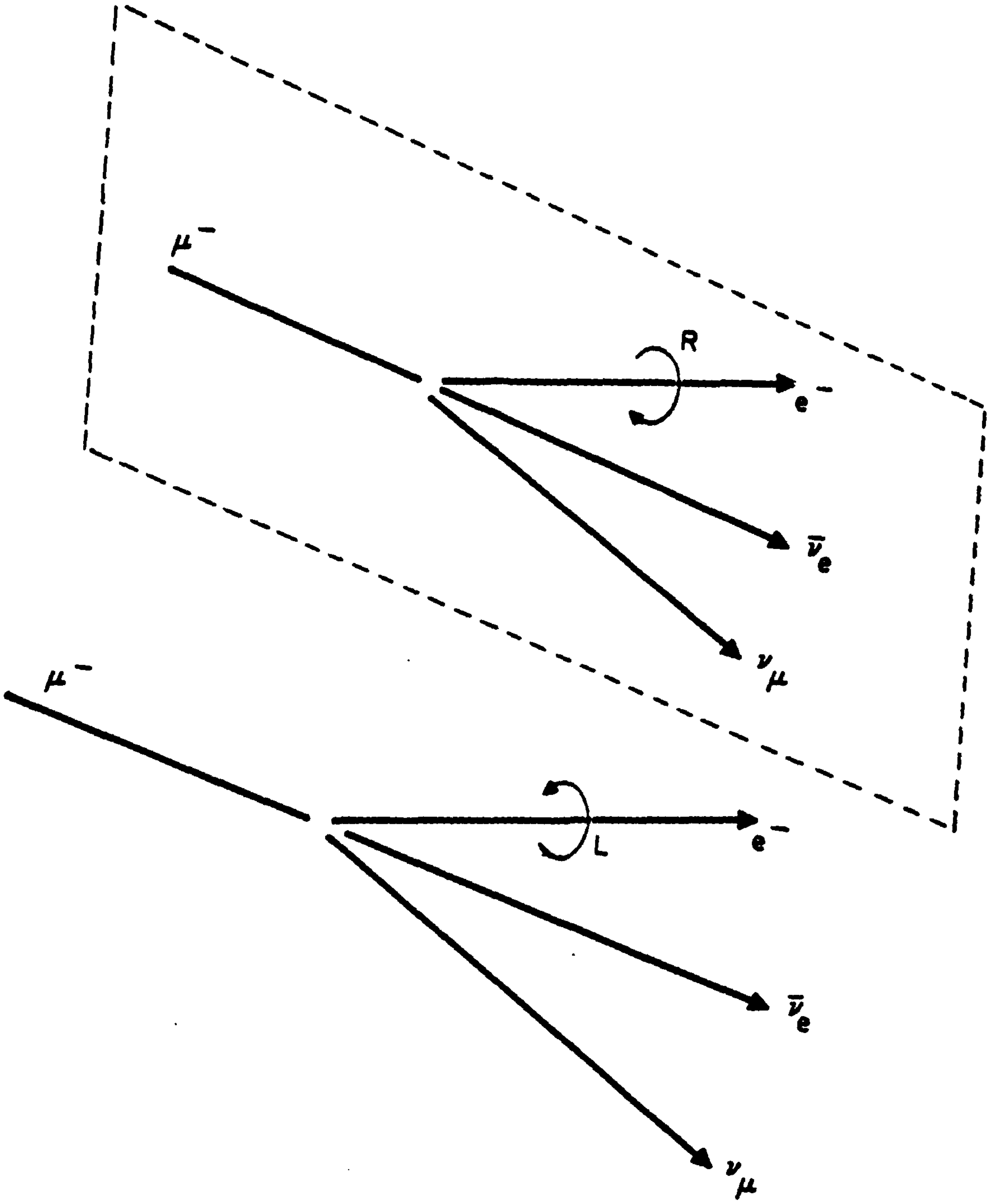


Fig. 20 PARITY VIOLATION IN μ-MESON DECAY
THE LEFT SPINNING ELECTRON APPEARS
1000 TIMES MORE OFTEN.

CP invariance becomes connected to the question whether there may be processes that violate time reversal invariance (T) through the TCP-theorem.

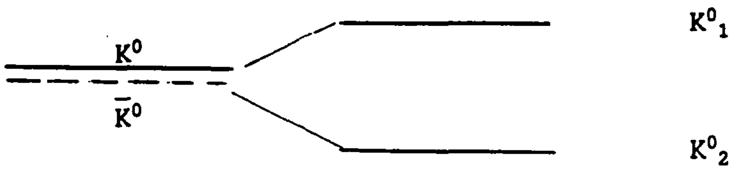
The TCP theorem states and can be proven in general that any theory that obeys Lorentz invariance and that does not contain instantaneous actions at a distance (so called local theories) will be invariant under the combined transformation of TCP. That means that for any process in nature the process that results by charge conjugating all participating particles, then mirror imaging the process and then running it reverse in time is also an equally possible process.

Particular implications of the theorem are (1) for stable particles a particle and its anti particle should have the same mass, (2) for unstable particles they should have equal decay rates or lifetime.

Because of the strong belief in the validity of Lorentz invariance we believe that TCP is conserved throughout physics. That means the observed violation of P in weak interaction must be compensated by another violation. We saw that C is also violated but that CP is conserved in the case we looked at. Should we find a process in which CP is violated then we would have to conclude that T is violated.

Such a violation was indeed found in 1964 by Christensen, Cronin, Fitch and Turley in connection with the neutral K^0-

meson and its antiparticle $\bar{K}^0$. These K^0 mesons have equal mass. In reality the existence of this particles is perturbed by the fact that the K^0 and $\bar{K}^0$ convert into each other at a very low yet finite rate involving two π-mesons as intermediate state. Effectively this breaks the doublet state of K^0 $\bar{K}$ into two states $K_1{}^0$ and $K_2{}^0$ with slightly different energies



The $K_1{}^0$ and $K_2{}^0$ states decay into π-mesons.

One can show that in order to preserve CP invariance the state $K_1{}^0$ can decay into $\pi^+ + \pi^-$ but $K_2{}^0$ cannot. Instead $K_2{}^0$ can decay into three π^0 but $K_1{}^0$ cannot. This makes $K_2{}^0$ a "slow" decaying state with lifetime $5x10^{-8}$ sec while $K_1{}^0$ is "fast" decaying in 10^{-10} sec.

$$K_1^0 \rightarrow \pi^+ + \pi^- \quad (t=10^{-10} sec)$$
$$K_2^0 \rightarrow \pi^0 + \pi^0 + \pi^0 \quad (t=5x10^{-8} sec)$$

In 1964 Christensen, Cronin, Fitch and Turley found in a beam of $K_1{}^0$ and $K_2{}^0$ at a distance at which all the $K_1{}^0$ have decayed and only the $K_2{}^0$ are left in the beam, that 0.1% of the $K_2{}^0$ decayed by the CP violating decay into two π-mesons.

After other experiments with the K^0-system had shown that TCP is not violated it was clear that along with CP also T must be violated at the level of 0.1% in this experiment.
This is the only known evidence that nature does recognize a direction of time at the microscopic level.
We will discuss later what the significance of this arrow of time might be.

11.5 The Time-Arrow in Radiation Processes

As was already mentioned all elementary dynamic laws of physics show the invariance under time reversal. The famous Maxwell equations that govern all electromagnetic phenomena including the generation and propagation of electromagnetic waves and light have the same invariance property.
There is however a curious assymmetry between possible phenomena based on Maxwell's equations with their time invariance and what actually occurs in nature.
All electromagnetic radiation is energy that is emitted in form of electromagnetic waves from a source region. The "cause" of this energy emission is the presence of time variable electric currents in that source region.
In all known practical situations the source region is a spatially finite region containing moving electric charges that constitute the currents. These currents, as they change create varying electromagnetic fields in the surrounding space.

Because of the finite propagation speed c of all electromagnetic influences, the field changes at distant field points arrive there at a time after the cause, which is equal to the propagation time [Fig 21].

The electromagnetic field at the observation point $\vec{x}$ at the time t is described by the relativistic vector field $A^{\mu}(\vec{x};t)$. This vector function must be a solution of the "wave equation" which is a direct consequence of Maxwell's equations:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right) A^{\mu}(\vec{x};t) = 4\pi\, j^{\mu}(\vec{x};t)$$

The source current at the point $\vec{x}'$ in the source region at a time t' is $j^{\mu}(\vec{x}';t')$.

All sources in the source region contribute to the field at $\vec{x}$ at the time t provided they are launched from such a source point at just the right time t' that they arrive at the observation point $\vec{x}$ at the same time t.

In the absence of any boundaries in space that could reflect waves the solution for $A^{\mu}(\vec{x};t)$ is written as

$$A^{\mu}_{ret}(\vec{x};t') = \int d^3x' \int dt'\, j^{\mu}(\vec{x}';t')\, \frac{\delta\left(t' - \left(t - \frac{|\vec{x}-\vec{x}'|}{c}\right)\right)}{|\vec{x}-\vec{x}'|}$$

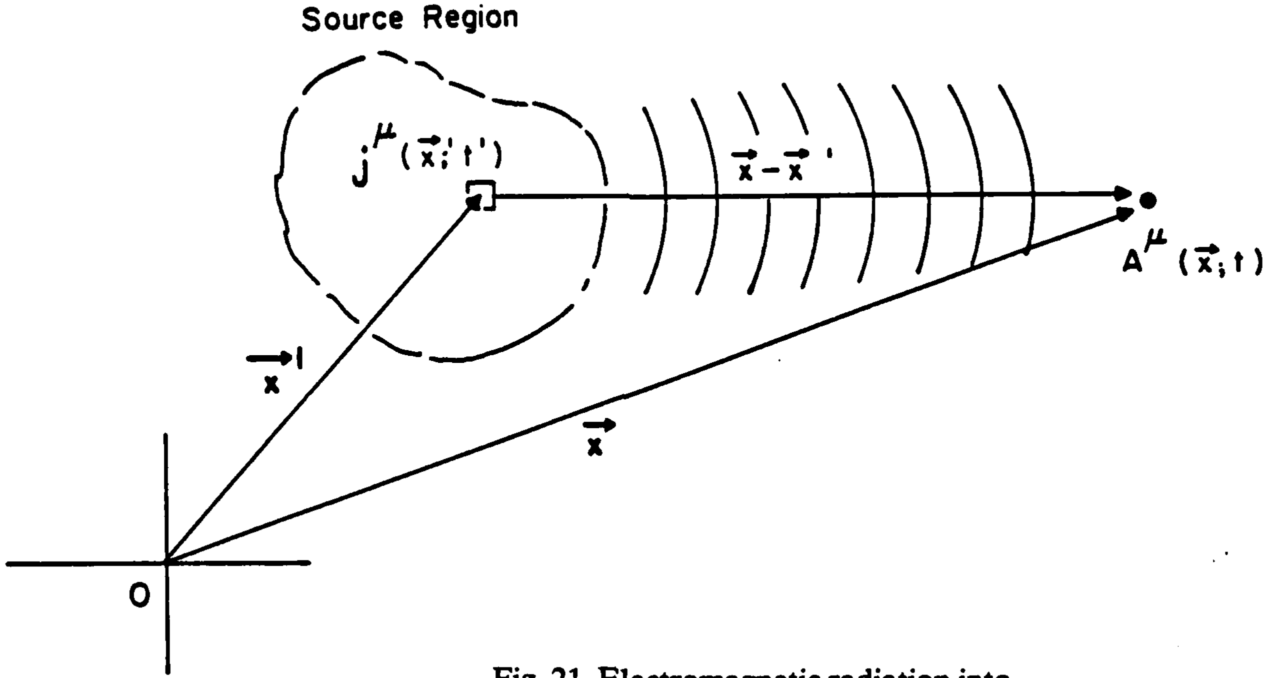


Fig. 21 Electromagnetic radiation into free space from a source region

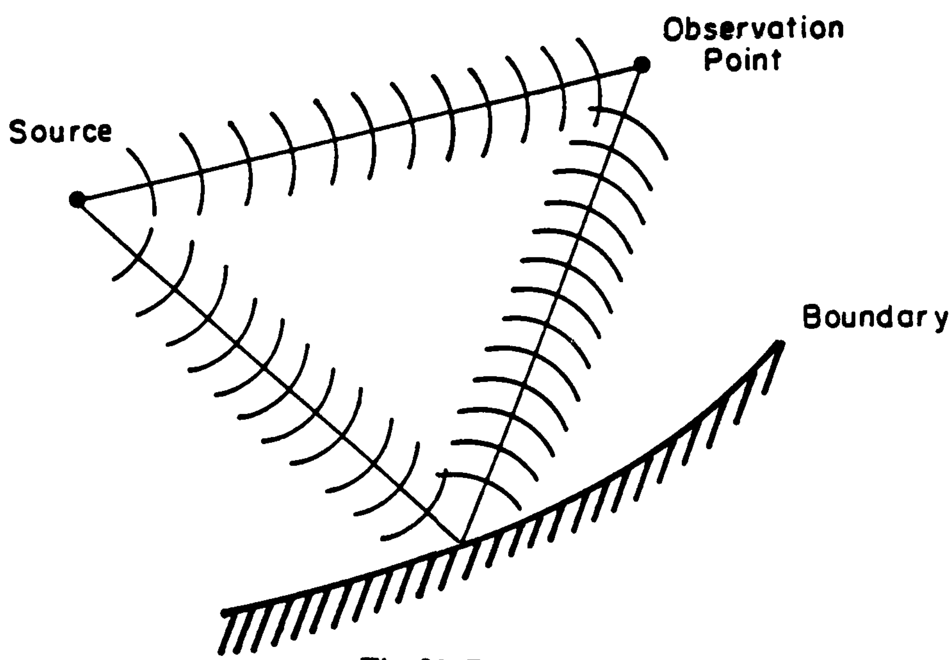


Fig. 22 Radiation influenced by presence of boundaries

The factor $\frac{1}{|\vec{x}-\vec{x}'|}$ comes from the fact that the signal decreases in strength as the distance traveled increases.

The so called δ-function is a device that sorts out the details which signals from what source point contribute at time t to the total signal.

$\frac{\vec{x}-\vec{x}'}{c}$ is the time the signal travels from source point $\vec{x}'$ to observation point $\vec{x}$. The time $t-\frac{|\vec{x}-\vec{x}'|}{c}$ would be the launch time of the signal at $\vec{x}'$. Unless t' is equal to that launch time the source $j^{\mu}(\vec{x}';t'$ cannot contribute to the signal. The δ-function $\delta(t'-t_{launch})$ is zero except for $t' = t_{launch}$ for which it is equal to one.

Implicit in this solution is that the signal leaves the source region at times t' and arrives at distant observation points at later times $t = t'+ \frac{|\vec{x}-\vec{x}'|}{c}$.

Another way of saying this is: In order to know the signal at a time t we must assess the source behavior at a <u>retarded time</u> $t' = t- \frac{|\vec{x}-\vec{x}'|}{c}$.

This language ar.d the so called retarded solution $A^{\mu}_{ret}(\vec{x};t)$ convey the picture of "outgoing" radiation, energy that is generated in the source region and propagates outward and is detected at distant points at corresponding later times. The very peculiar consequence of the time inversion invariance of Maxwell's equations is that, there also exists an advanced solution to the wave equation.

$$A^{\mu}_{adv}(\vec{x};t) = \int d^3x' \int dt' \, j^{\mu}(\vec{x}';t') \frac{\delta\left(t' - \left(t + \frac{|\vec{x}-\vec{x}'|}{c}\right)\right)}{|\vec{x}-\vec{x}'|}$$

Now, what happens at the source point $\vec{x}'$ at the time t' is related to the field at a time t a propagation time earlier. This conveys the picture of inward going waves that so to speak "turn on" the appropriate current distribution that matches the field configuration. This solution entails that waves come in concentrically from all directions into the source region producing a particular current distribution. We never see that happen in nature. Why is that? The picture of the retarded and advanced solution can be visualized in terms of water surface waves in a water pond. If one forces a cork swimming in the center of the pond into a rhythmic up and down motion then the cork acts like the "antenna" and represents the source region. One sees circular wave trains leaving the cork traveling outward toward the edge of the pond. If one stops the cork, then one can watch the waves hit the edge of the pond, reflect and begin to move inward concentrically. When they reach the cork they set it into an up and down motion. In this example the concentric incoming wave from the edge of the pond is clearly due to the presence of the pond edge that reflects waves.

The retarded and advanced solution for electromagnetic fields as given above are solutions to Maxwell equations for free space i.e. absence of all boundaries.

When boundaries are present then to the free space solutions there must be added contributions that arrive at $\vec{x}$ at time t due to reflections from boundaries.

$$A^{\mu}(\vec{x};t) = A^{\mu}_{ret}(\vec{x};t) + A^{\mu}_{in}(\vec{x};t)$$
$$A^{\mu}(\vec{x};t) = A^{\mu}_{adv}(\vec{x};t) + A^{\mu}_{out}(\vec{x};t)$$

In the first solution $A_{in}{}^{\mu}$ is the contribution from an incoming field due to reflections from boundaries at earlier times [Fig 22].

The contribution from the incoming wave is calculated from the values of the field at the boundaries at an initial time. The solution formulated in this way is a proper "initial value solution", determining the later outcome from initial conditions.

In the second form of the solution A_{out}^{μ} is an outgoing field from $\vec{x}$ to the boundary where it arrives in the future. Giving assumed values of the fields at the boundary at future time one can determine what A_{out}^{μ} must have been at the earlier time to match the final condition. This is the case of a final value problem where the solution is determined by the final outcome. Both complete solutions are correct solutions of Maxwell equations.

The question then is: Why is it that except in contrived artificial situations we find mostly radiation situations that are best described as an initial condition problem, rather than a final state condition *situation?* Is there something in radiation that provides an arrow of time ? If so, where does this arrow of time come in? It does not appear in Maxwell's equations. We will discover more arrows of time in the next section.

12. THE THERMODYNAMIC ARROW OF TIME

12.1 The Science of Systems of Very Many Particles.

The 19th century brought a complete understanding of the phenomena of heat and the transport of heat. All macroscopic material objects such as gases, liquids, solid bodies consist of very large numbers of elementary building blocks of the material, molecules.

These molecules are very small, having typical diameters of the order of a few Angstroms (10^{-10}m). Their masses are of the order of 10^{-26}kg. A thimble full of solid material contains of the order of 10^{23} molecules. These are very large numbers.

These molecules in the material are not totally at rest. They have energy and they move about with considerable velocities as in gases and liquids or they vibrate around their standard position in the crystal lattice of a solid object.

This motion of the molecules is their "thermal motion" and is perceived by us as heat. The larger the temperature of the object, the more heat is contained in it, the more agitated is the thermal motion of its molecules.

In order to describe in detail what happens in such a gas say, one would have to describe the motion of each of its molecules using Newton's laws of mechanics. One would describe how each molecule moves, how it collides with other molecules, is reflected taking a new direction, occasionally hit the wall of the container of the gas, is reflected from it and so on. And one would have to do this for everyone of the 10^{23} molecules. This is clearly impractical. The science of thermodynamics has developed a system of physical parameters that characterize the whole of the gas in the container sufficiently to make predictions about the behavior of the gas as a whole without having to know what each molecule is doing. Such quantities are the gas volume, the pressure, the temperature, the total internal thermal energy of all the molecules bouncing around in the container and some more quantities which we will meet shortly.

It is important to realize that the motion of the individual molecules follows Newton's laws of motion and is therefore time reversible. For each trajectory that a molecule follows in the container an equal but opposite motion is possible and with this many molecules may indeed occur.

12.2 The First Law of Thermodynamics

It is customary to use the behavior of gases as examples to demonstrate the laws of thermodynamics because of the relative ease with which the concepts can be shown.

We consider a gas in a well insulated cylinder with a piston on one side. The gas is at a temperature T and has a total internal energy U [Fig 23].

There are two ways in which the internal energy can be increased in the gas that is confined in the cylinder.

(a) One can inject heat energy ΔQ by an electric spark for instance.

(b) One can compress the gas by a volume change $-\Delta V$ by exerting a pressure p on the piston, doing compression work on the gas.

The total increase of internal energy in the gas is then

$$\Delta U = \Delta Q - p\Delta V$$

This is the formal expression for the first law of thermodynamics for a gas.

If no work is done on the gas by compression and if no heat is added to it then $\Delta U = o$. The internal energy stays constant when the system is isolated from its surrounding. Whatever energy U the gas has then is distributed into small portions over all the 10^{23} molecules, which bounce around in the cylinder. Typical velocities are of the order 500 m/s at room temperature.

During the many bounces each molecule makes, (for air at normal atmospheric pressure and temperature of the order 10^{10} collisions/sec. small portions of energy get transferred between molecules. At each instant one could imagine that the "microstate" of the gas system is characterized by how the

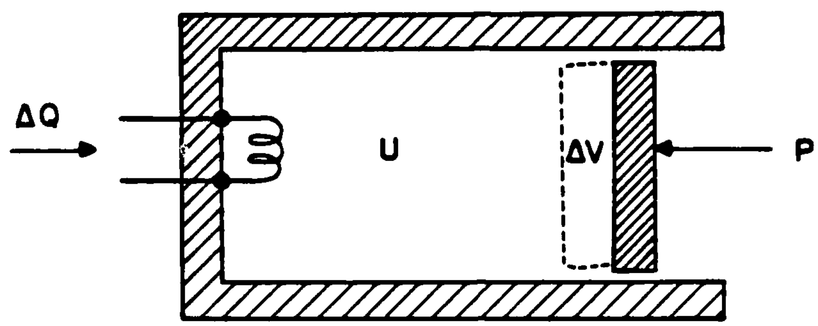


Fig. 23 THE FIRST LAW OF THERMODYNAMICS
THE INTERNAL ENERGY U CHANGES BY AN AMOUNT ΔU DUE TO THE HEAT ΔQ ADDED AND DECREASES BY THE WORK $P\Delta V$ DONE BY THE SYSTEM

$$\Delta U = \Delta Q - P\Delta V$$

total available internal energy U is distributed over the 10^{23} molecules.

The time evolution of the system is a sequence of microstates. If we look at a number of snap-shots of the system we catch it each time in one of its microstates (Figure 24).

In particular if the closed system of gas has been decoupled for awhile from all possible sources of heat exchange with the environment and compression work then we find the system in a sequence of microstates A, B, C that have no distinguishing features. Indeed it is impossible to determine what the temporal sequence of these states might be. There is no arrow of time visible in the sequence of microstates of an isolated thermodynamic system.

This indistinguishability of a temporal sequence in the microstates has its cause in the time reversibility of the microdynamic processes that the molecules undergo.

12.3 Entropy and the Second Law of Thermodynamics

(1) States of Increasing Disorder of Energy

Now we look at another sequence of a microstates of the same gas [Fig 25]. This sequence is quite different. From our general knowledge we are quite certain about the possible interpretation. We speculate that initially the gas was confine by some barrier to the left portion of the container as in B′. Then the barrier was suddenly

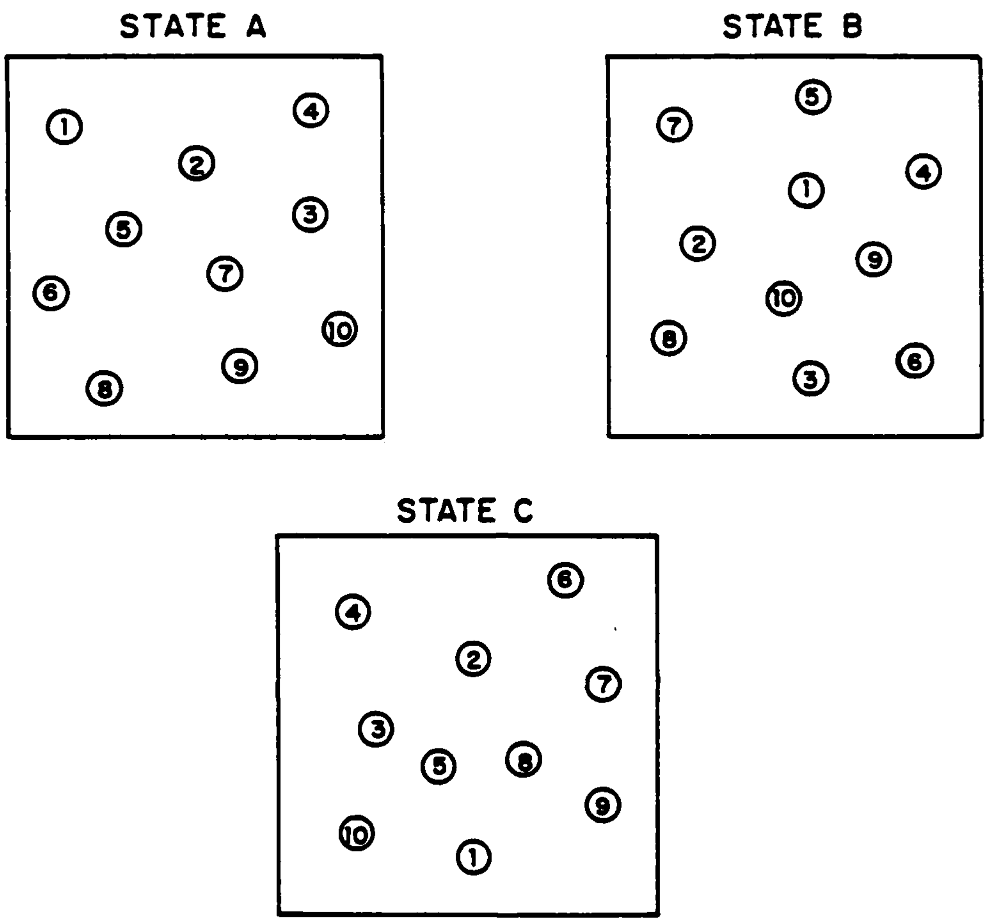


Fig. 24 CLOSED THERMODYNAMIC SYSTEM GOING THROUGH A SEQUENCE OF MICRO-STATES.
FOR SYSTEM NEAR EQUILIBRIUM NO TIME ORDERING OF THE STATES IS DISCERNIBLE.

removed and the molecules due to their energy begin to move to fill the newly available space (state A′) and that after awhile the system has reached an "equilibrium state" (C′) with the original energy spread rather evenly over the available space.

The system has evolved from an originally rather ordered state B′, with all energy concentrated in a small portion of the volume to a maximally disordered state C′ in which the energy and the molecules are distributed as evenly as possible. This process of going from a highly ordered state to a disordered state reveals the arrow of time. We are able to discern the temporal sequence of the successive states.

Thermodynamics has captured the importance of the concept of the degree of order of the microdistribution of the total energy of the system in the quantity "ENTROPY". Large disorder, dispersion of the available energy is equivalent to large entropy. High concentration of the energy is a low entropy state.

When a system goes from order to disorder, its entropy increases.

Entropy increase always involves transport of energy in such a way that the energy present in the system is dispersed.

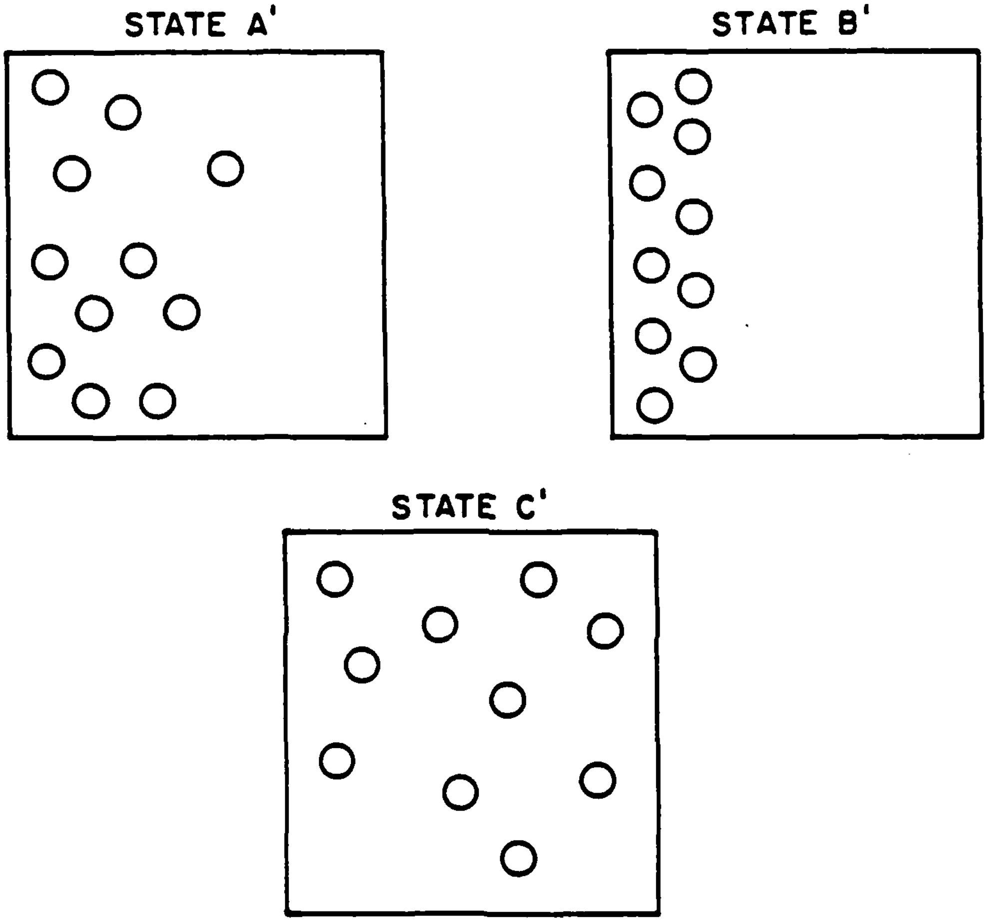


Fig. 25 THE THERMODYNAMIC ARROW OF TIME IN SYSTEMS GOING FROM STATE OF ORDER TO STATE OF LESS ORDER.

(2) The entropy change is given by the amount of heat transported divided by the temperature at which this heat was stored in the molecules of the gas.

$$\Delta S = \frac{\Delta Q}{T}$$

The heat exchange between two compartments of a system which are originally at different temperatures T_1 and T_2 with T_1 larger than T_2 may serve as example. If we keep the whole system isolated from all external influences and observe the two compartments we will see that gradually heat drifts from the hotter compartment (T_1) to the cooler one (Figure 26). After sufficiently long time the energy between the two compartments will have reached equilibrium. Both compartments reach the same temperature T_3 between T_1 and T_2.

When compartment (1) at temperature T_1 gives up an amount of energy ΔQ its entropy decreases by $\frac{\Delta Q}{T_1}$. Compartment (2) receives the same amount of heat ΔQ at temperature T_2. Its entropy increases by $\frac{\Delta Q}{T_2}$.

The total change of entropy for the two compartments together is positive. The entropy has increased. The original ordering of energy with more energy in compartment (1) has been disordered by some amount.

(3) The Second Law of Thermodynamics states:

Changes in an isolated system always occur in a way such that the entropy of the system increases irreversibly until it has reached the maximum value possible under the given constraints of the system.

This behavior of an irreversible change of state of a isolated system from a state of lower entropy to a state of largest possible entropy appears to provide an arrow of time.

(4) We call this process irreversible because in practice we never observe the reverse process in which the system beginning at equal temperature every where develops a significant temperature difference across the system.

We are struck by this observation because we know that the microscopic collision processes which lead to the dispersion of the energy to its maximal disorder are all time-reversible according to Newton's laws and the reverse processes should be occurring just as well. What is behind our experience that certain processes are more likely to occur than others?

A water glass standing at the edge of a table may drop due to a small vibration of the table. As it drops the water spills and splashes all over the floor. The glass

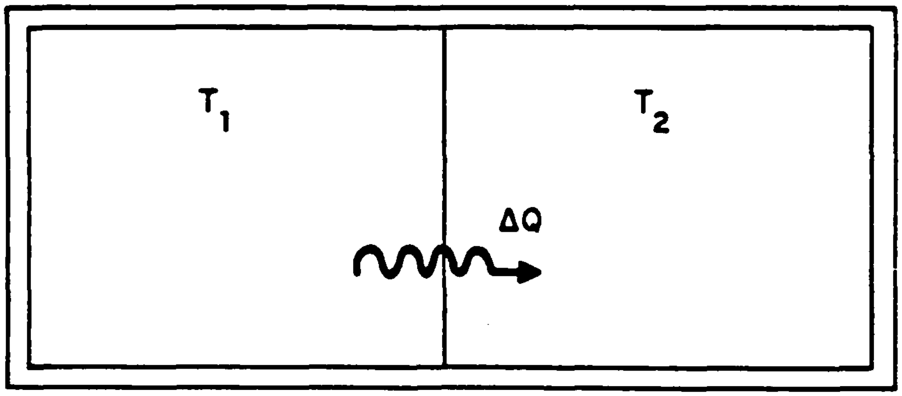


Fig. 26 THE SECOND LAW OF THERMODYNAMICS

"CHANGES IN AN ISOLATED SYSTEM ALWAYS OCCUR IN SUCH A WAY THAT THE ENTROPY OF THE SYSTEM INCREASES".

ISOLATED SYSTEM WITH TWO COMPARTMENTS IN THERMAL CONTACT WITH TEMPERATURES T_1 AND T_2. HEAT ΔQ TRANSFERS FROM (1) TO (2).

CHANGES OF ENTROPY IN THE TWO COMPARTMENTS:

$$\Delta S_1 = -\frac{\Delta Q}{T_1} \qquad \Delta S_2 = \frac{\Delta Q}{T_2}$$

TOTAL CHANGE OF ENTROPY:

$$\Delta S = \Delta S_1 + \Delta S_2 = \Delta Q \left[\frac{1}{T_2} - \frac{1}{T_1} \right] > 0$$

breaks into pieces and the whole thing ends in a mess with all the fragments of water and glass having peculiar coordinated motions due the laws of mechanics that govern the entire process.

We never see the inverse process of droplets and puddles of water and glass pieces gather up from the floor, moving upward and in the process assembling themselves into a glass filled with water standing at the edge of the table.

Yet this entire reverse process is energetically and dynamically quite possible in principle.

Even the fact that some of the energy goes into heat motion of the floor and evaporation of some of the water *does not matter*. The distribution of energy into many portions into a very complicated coordinated motion is reversible in principle. The first law of thermodynamics is indeed time reversal invariant.

The scrambling of an egg or the dissolving of a sugar cube in a cup of hot tea are similarly phenomena which we do not expect to occur in reverse direction although energetically they could.

We are used to accept complicated coordinated patterns of motion like the splashed water and glass pieces, or the scrambled egg mixture or the dissolved sugar molecules as the effect of an earlier cause. The cause is the placing of the highly ordered state of matter in form of the water glass, or the complete egg or the sugar cube at the beginning of the process. The process leads to a state of greater disorder in form of coordinated motion of many disjoined pieces. We do not think of such a state as the possible cause. We reason that it would take an absurdly precise coordination of the motion of all the pieces in order to assemble into the filled water glass.

If we were living in a world where reassembling water glasses, unscrambling of eggs and undisolving sugar cubes were common occurrences we would see as the "cause" not the extraordinarily improbable coordination of motion of the pieces but rather we would probably tend to attribute the process to a forward looking cause of a coordinated outcome. The pieces are coordinating their motion because of the intended outcome. The coordination is the effect. The cause would in time follow the effect [PE 89].

We do not observe this kind of world. We must ask why in our real world the cause comes before the effect. Why do precisely coordinated particle motions occur only after some large scale initial state has been set up that places the energy into an ordered confined state which then evolves into a state with the energy distributed into many coordinated fragments?

The second law of thermodynamics states that the entropy of an isolated system increases with time for irreversible changes and remains constant for a reversible system.

(5) <u>Measure of Disorder</u>

We have given a quantitative expression to the change of entropy in a compartmented system when a certain amount of energy is transferred from one compartment to the other.

But how can we give precise quantitative meaning to the concepts of "manifest order" and "manifest disorder"? How disordered is the state of the broken water glass and splashed water, the scrambled egg, the dissolved sugar cube?

The second law stated that the entropy increases for "irreversible" systems. What does this irreversibility mean? From an energy point of view all molecular motions are time reversible. Why is it that we consider the falling water glass, the egg scrambling, the dissolving sugar cube irreversible processes?

We do believe that to unscramble the egg, to undissolve the sugar cube, to unspill the water is in practice impossible. Does that mean the irreversibility is a concept of practicality? Is the entropy concept based on what is practical? Or can we give entropy as a measure of the degree of disorder a precise quantitative meaning that can help us understand why irreversible processes are so very unlikely to run backwards although from the energy point of view and the principles of dynamics of the motion of the molecules they can run in reverse.

12.4 The Phase Space Volume as Measure of Entropy

The total energy U of a volume of gas of 10^{23} molecules is broken up into tiny portions with each molecule sharing some of it in form of its kinetic energy of thermal motion.

In order to characterize the macrostate of the gas one can divide the energy range of the particles which must be somewhere between zero and U into M energy bins ΔE. One can then think of counting how many molecules have energies that falls into a particular bin and create a bar chart, with the total number of molecules $N = 10^{23}$ [Fig 27].

The total energy of the system is constrained to be U. One possible macrostate would be that one molecule has energy U. Then all the others must have energy zero. But this state can be realized in N different ways because each of the N molecules could be the one with energy U. Similarly all the other possible distribution of the N particles into the available

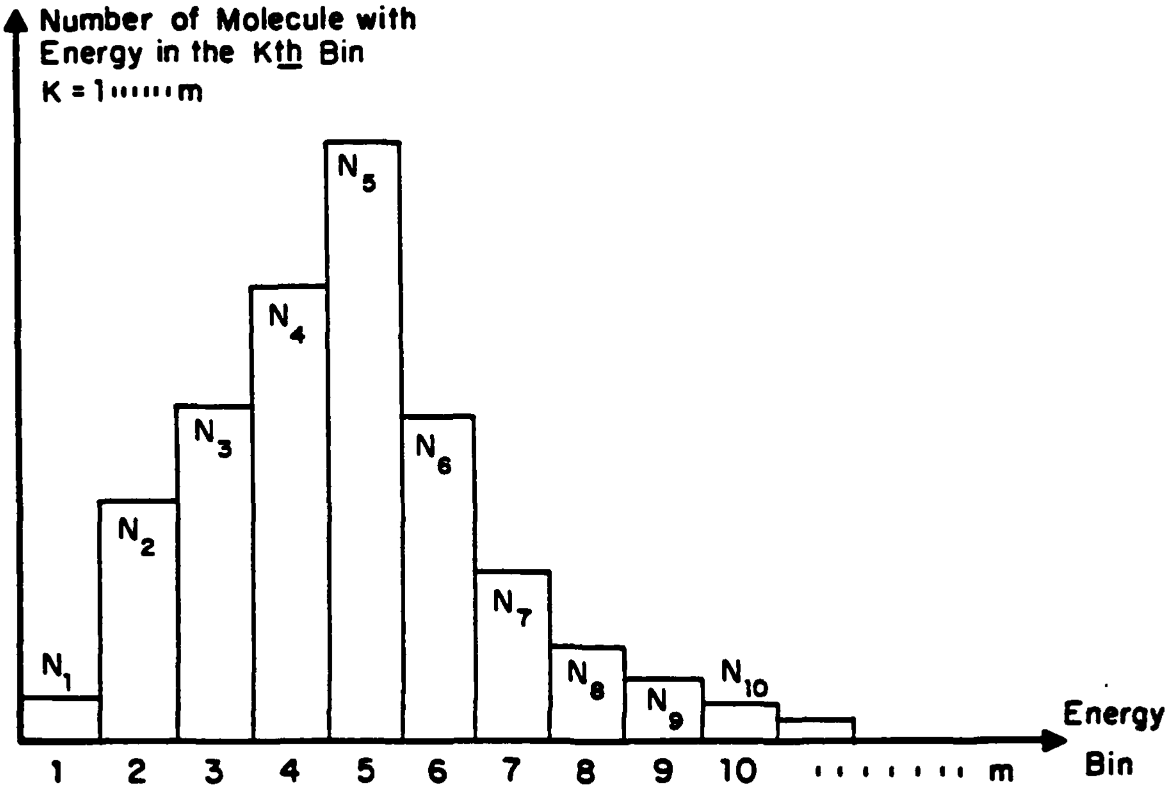


Fig. 27 ENERGY DISTRIBUTION OF MOLECULES DEFINING A MACRO STATE.
STATES WITH PARTICLES EXCHANGED BETWEEN ENERGY BINS BUT LEAVING THE DISTRIBUTION INTACT ARE DIFFERENT MICROSTATES.

energy bins can be done in very many different ways by exchanging particles from one bin to the other but leaving the number in each bin the same i.e. not changing the macroscopic energy distribution. Each realization of the same distribution be exchanging particles constitutes one microstate.

One can now ask which distribution N_1 N_2 N_3 ... N_m of the N particles into the available energy bins between E=o and E=U has the largest number of possible realizations.

This question has been answered by Maxwell and Boltzmann and is given in the famous Maxwell-Boltzmann energy distribution law:

$$N(E) = 4\pi N \left(\frac{m}{2\pi kT}\right)^{3/2} \frac{2E}{m} e^{-\frac{E}{kT}} \Delta E$$

where N(E) is the number of molecules with energy in a bin of width ΔE around the value E, m is the mass of the molecules, T is the temperature of the gas and k is a universal constant which will turn out to be an elementary unit of entropy. This distribution is the equilibrium distribution which an isolated volume of gas of N molecules at temperature T will assume after enough time. It is the state of maximal possible entropy for that gas at that temperature. This distribution has the largest number of possible realizations by exchanging molecules between bins. Any other distribution has fewer number of realizations, say $g(N_1\ N_2 \cdots N_m)$. The number of possible realizations of a given macrostate characterized by the string of numbers

$(N_1\ N_2 \cdots N_m)$ is a measure of probability for that macrostate to occur.

One could construct a "phase space" of all possible microstates in which each macrostate takes up a space volume proportional to the number of realizations of this state. In a two-dimensional model of this phase space the large area in the center corresponds to the equilibrium state which has the largest number of realizations. This area is surrounded by other smaller areas, representing macrostates in the neighborhood of the equilibrium state, that have fewer realizations. Outward direction is the direction of larger distance from equilibrium and the area is covered with smaller and smaller patches corresponding to the decreasing number of realizations of these states (Figure 28).

This figure does not give the proper proportion of the true situation. A quanitative estimate shows that the phase space area for the equilibrium state is extraordinarily large compared to neighboring areas associated with small departures from the equilibrium state. The probability for small departures from equilibrium in a isolated system is extremely small. The reason for this sharp peak of the probability near the equilibrium state is the large number of $N = 10^{23}$.

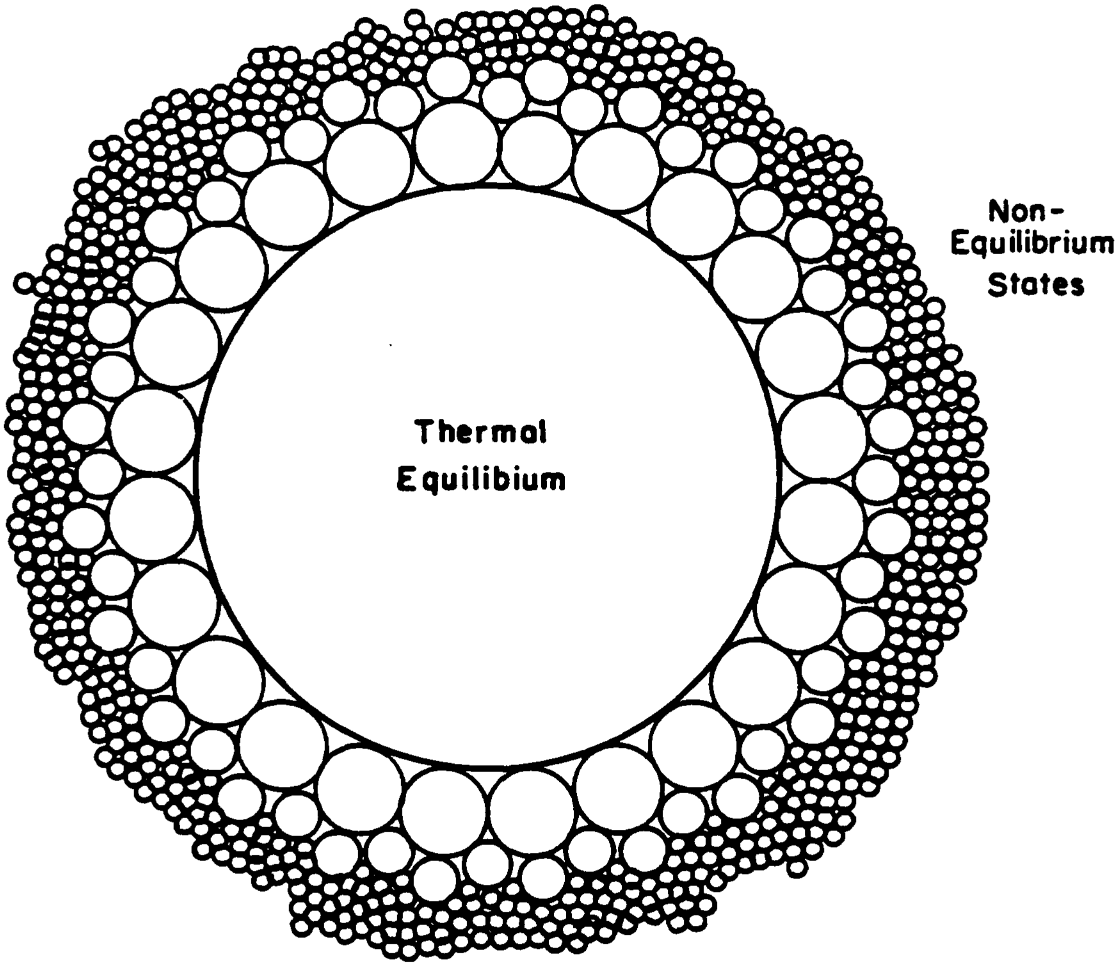


Fig. 28 PHASE SPACE VOLUME OF MICROSTATES

In appendix D the quantitative measure for the sharpness of this probability peak is given for the simple example of the distribution of N molecules into two equal sized compartments of a container.

In complete equilibrium each compartment would contain N/2 molecules. The number of realizations in which the N molecules can be distributed into the two compartments with N/2 in each is 2^N. The equilibrium state would be assigned an area of phase space equal to $A_o = 2^N$. A state in which one compartment contains $\frac{N}{2} + x$ molecules and the other $\frac{N}{2} - x$ has a number of realizations.

$$g(N,x) = 2^N e^{-\frac{2x^2}{N}}$$

The area associated to this state is a factor $e^{-\frac{2x^2}{N}}$ smaller than A_o.

If $N = 10^{23}$ and x is chosen 10^{13} it means that of 10^{10} molecules one is taken from one side and put in the other compartment, creating an imbalance of 1 in 10^{10}. The corresponding area in phase space and therefor the probability for that much imbalance is a factor 10^{-870} smaller. This is an incredible small number. There is simply no realistic comparison that can be offered from our world of experience that could make this smallness plausible. The only number that compares with this is the probability that of the dissolved sugar molecules in the cup of tea one half of the cup contains one sugar molecule more for each 10^{10} sugar molecules than the other half.

It was the great contribution of Ludwig Boltzmann that he made the connection between the probability of a macrostate to occur and the entropy. Boltzmann finds the famous relation

$$S = k \ln g(N_1 N_2 \cdots N_M)$$

Where $g(N_1, N_2 \cdots N_M)$ is the number of ways in which the given macrostate characterized by the distribution $N_1 \cdots N_M$ of the energy can be realized by microstates of the N molecules and k is the Boltzmann constant, a universal constant that is the entropy per particle and has a value of:

$$k = 1.38 \times 10^{-23} \text{ Joule K}^{-1}$$

In our example the entropy is

$$S = KN \left\{1 - \frac{2}{\ln 2}\left(\frac{x}{N}\right)^2\right\} \ln 2$$

In equilibrium with x=o and $N=10^{23}$ the entropy is S=0.95 Joule and goes toward 0.22 when $x = \frac{N}{2}$ i.e. all molecules assembled in one compartment. Because of the dependence of S on the logarithm of g(N, x) the numerical values in terms of S are not so incomprehensible. It is easier to compute the deviation of a state from equilibrium in terms of entropy than in terms of probabilities.

12.5 APPROACH TO THERMODYNAMIC EQUILIBRIUM

Following an argument given by Penrose [PE 89] we can see now that a system that starts in a state far from equilibrium in one of the small volume element at some distance from the center because of the thermal motion of the molecules will with high probability move into larger and larger volume elements and rapidly reach the very large volume of the equilibrium state. The system will continue to move about in the phase space in a random path through very many microstates. But because the thermal equilibrium volume is so incredibly large compared to all other states the system has extremely low chance to reenter one of the small volumes [Fig 29].

The motion of the system in phase space is governed by the laws of dynamics and the system runs in positive time direction. So we are finding that the entropy increasing path toward the equilibrium state seems associated with the positive time direction.

Because the molecular dynamics is time reversible we could imagine the system running in negative time direction. If the system is started out in a small compartment of the phase space then the probability argument would predict that in this negative time direction also the system would rapidly move into

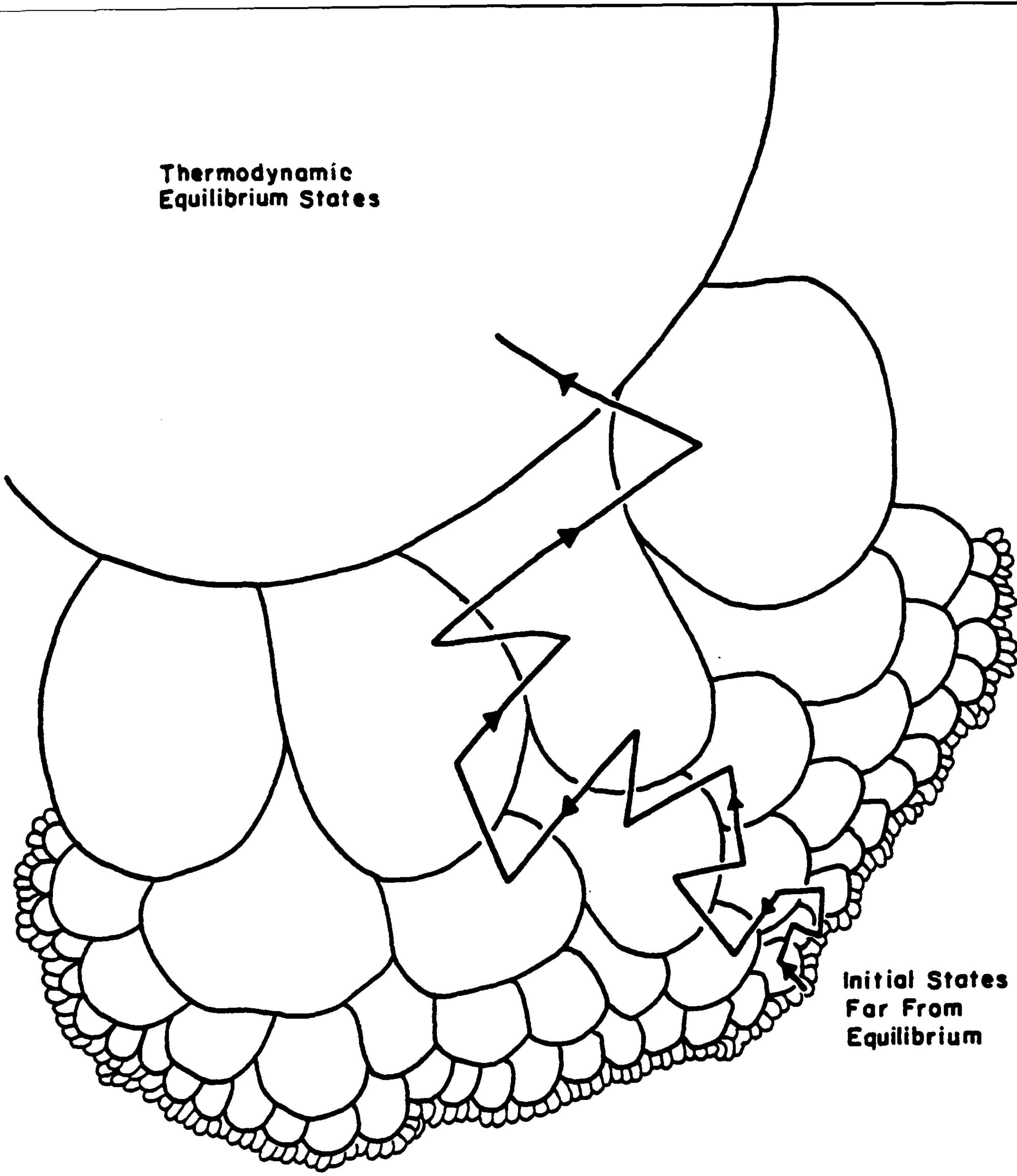


Fig. 29 AS TIME PROGRESSES THE PHASE SPACE POINT OF THE SYSTEM ENTERS COMPARTMENTS OF PHASE SPACE OF LARGER VOLUME, MACROSTATES OF LARGER NUMBER OF MICRO STATES. THE ENTROPY INCREASES.

larger and larger volumes and eventually end at large enough negative time in the equilibrium state.

Does that mean that the system has started at earlier time in a equilibrium state, moves out with decreasing entropy to the special initial state and then again with increasing entropy back to equilibrium? No, it does not mean that.

If for instance the initial state A is the state with all molecules in one compartment of a volume then this state was presumably initiated by filling this compartment with the gas and having it confined in that compartment by a separator. The path toward equilibrium begins when this constraint is lifted. Equally well we can imagine that the velocities of all molecules before lifting the constraint are inverted to the opposite direction. Then all molecules will run through their trajectories backward. They still will collide with the other molecules and the walls and there will be equally many molecules whose velocity points in the direction toward the empty compartment and eventually molecules will fill the empty compartment and entropy has increased [Fig 30].

If the system were started in thermal equilibrium we do not expect to see the system return to the constraint state even if we turn all velocities around, because once the system has entered thermal equilibrium the influence of the original constraint is no longer of importance. The phase space volume

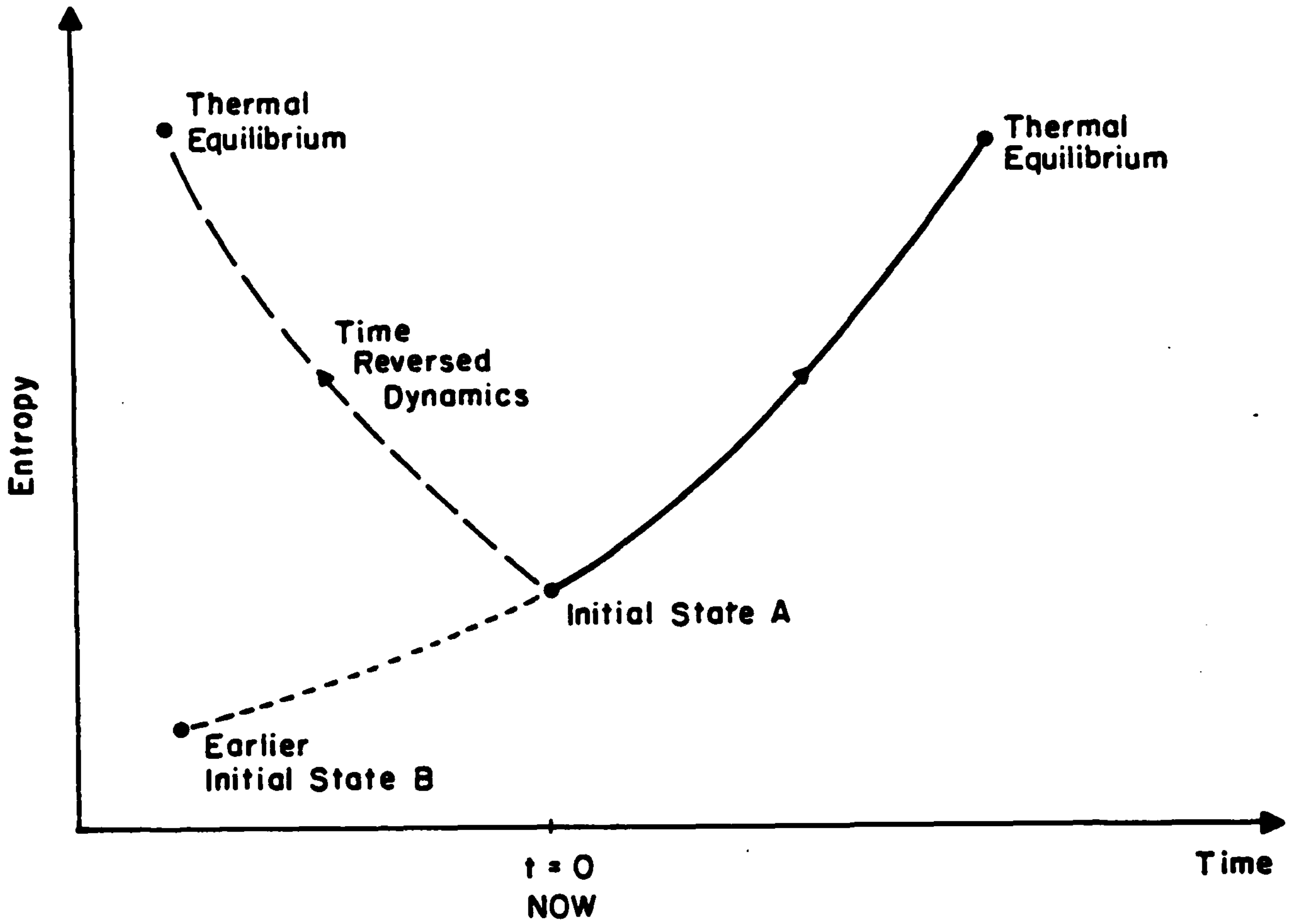


Fig. 30 ENTROPY INCREASE IN TIME REVERSED SYSTEMS

of thermal equilibrium i.e. the possible realizations of that state are so much larger that just reversing directions of all velocities does not mean that the system moves out of equilibrium.

Given an initial low entropy state and absence of any other factors constraining the system, the entropy would be expected to increase in both directions of time.

The predecessor to the low entropy state A in Fig 30 is not a less constrained higher entropy state but perhaps an even lower entropy state B.

So thermodynamics has not provided us with an arrow of time. The underlying molecular dynamics is time reversible. The second law tells us that the system has an overwhelming tendency to evolve toward thermal equilibrium.

That we see in our world entropy increasing in systems in forward dynamics in positive time direction is not because entropy increase is intrinsically coupled with positive time direction but because in our world all precesses of entropy increase appear to emerge from prior states of lower entropy due to certain constraints. We must investigate the reasons for the entropy states in the past.

Another issue is how the second law of thermodynamics is emerging from a consistent statistical mechanics model of a system. An overview is given in appendix D.

13. THE THERMODYNAMICS OF THE EARTH-SUN SYSTEM

THE QUESTION OF THE ORIGIN OF LIFE

13.1 The Origin of the Low Entropy States Around Us

The examples which we have given to exemplify entropy increasing processes were states of low initial entropy like the water glass on the tables edge, the sugar cube, the egg over the bowl. In all cases these low entropy initial states were prepared by man.

Man himself is a highly complex, organized system i.e. a low entropy system. How does man maintain such a low entropy state?

Man is a system of essentially constant energy. He ingests food which provides energy by chemical reaction. This energy leaves the body mostly as heat. Heat is a high entropy state. We are taking energy in a low entropy form as food and oxygen. In turn the energy is given up as heat and carbondioxide. So man interacts with his environment leaving the entropy of the environment increased [Fig 31]. The living body is a "dissipative structure" in the terminology of Prigogine [PR 80].

The entropy balance for the living organism is thus negative. Giving large entropy to the surrounding allows decrease of the organism entropy. This decrease manifests in internal self-organization of the system. New cells are grown, higher forms

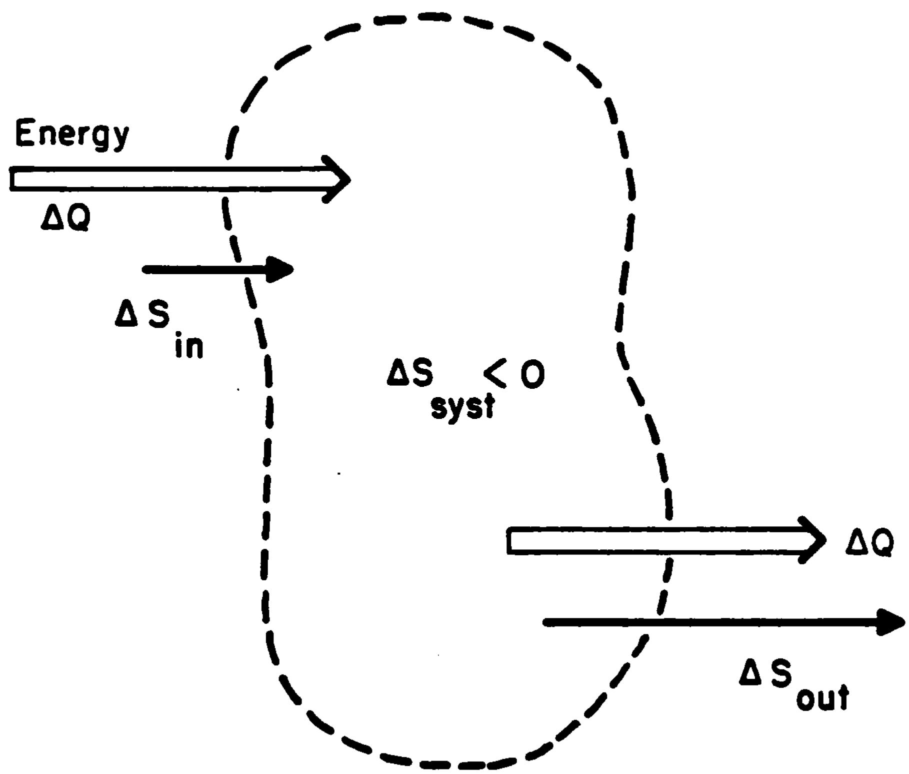


Fig. 31 MAN AS DISSIPATIVE SYSTEM
THE INTAKE OF ENERGY IN CONCENTRATED LOW ENTROPY FOOD IS DISSIPATED PREDOMINANTLY AS LOW TEMPERATURE HIGH ENTROPY BODY HEAT, ALLOWING TO DECREASE BODY ENTROPY AND CELL STRUCTURE FORMATION.

of cell organization are grown.

The source of the low entropy energy for living organism is the food. Where does the low entropy food come from?

In the biological chain the fast pace of evolution began after the occurrence of the low energy photo synthesis:

$$hv + Co_2 + \text{chlorophyl} \rightarrow C + o_2$$

The conversion of the earth atmosphere from a reducing atmosphere to an oxidizing one allowed the rapid evolution of organic structures of ever more complex forms. These low entropy structures became the elements of a food chain on earth.

13.2 THE SUN-EARTH SYSTEM

The sun radiates 3.9×10^{26} Joule/sec of energy into space. Most of that radiation is in the visible part of the electromagnetic spectrum. The earth intercepts in the average 1370 Joule per second per m^2 at the top of the atmosphere of the sunlit portion of the earth. Of this energy arriving, 34% is reflected back into space, 19% are absorbed in the atmosphere and 47% are absorbed by the earth surface. If this energy were continuously absorbed in the ocean with an average depth of 2000 m one would expect the ocean temperature to rise by about 2°K per year. That of course does not take place. Instead the earth temperature rises to a temperature of around 300°K (20°c) at which temperature the heat radiation form the earth equals the energy absorbed. The solar radiation is characterized by the surface temperature of the sun which is $T_{\odot} \approx 6000°K$. The radiation spectrum at that temperature is dominant in the visible part of the spectrum. The heat radiation from the earth is characterized by the temperature of the earth $T_E \approx 300°K$. At that temperature the spectrum is in the infrared region.

In a heat balanced earth then the amount of energy absorbed ΔQ_{in} arrives in form of high energy, low entropy photons, which bring to the earth the entropy $\Delta Q_{in}/T_{\odot}$.

The same amount of energy leaves the earth in the form of low energy high entropy infrared photons carrying away the entropy $\Delta Q_{out}/T_E$ where $\Delta Q_{out} = \Delta Q_{in}$. Therefore the earth is left with a decrease in entropy.

Again we can say that in principle this decrease in entropy can manifest in building complex dissipative structures [Fig 32].

The sun is an energy source of low entropy. Through the devices of forming dissipative structures on earth she produces low entropy energy sources on earth in form of wood and fossil fuels and the food chain for an evolutionary tree of living organisms.

So, we have traced the appearance of low entropy states back to the sun as a high temperature source. We must ask then what is

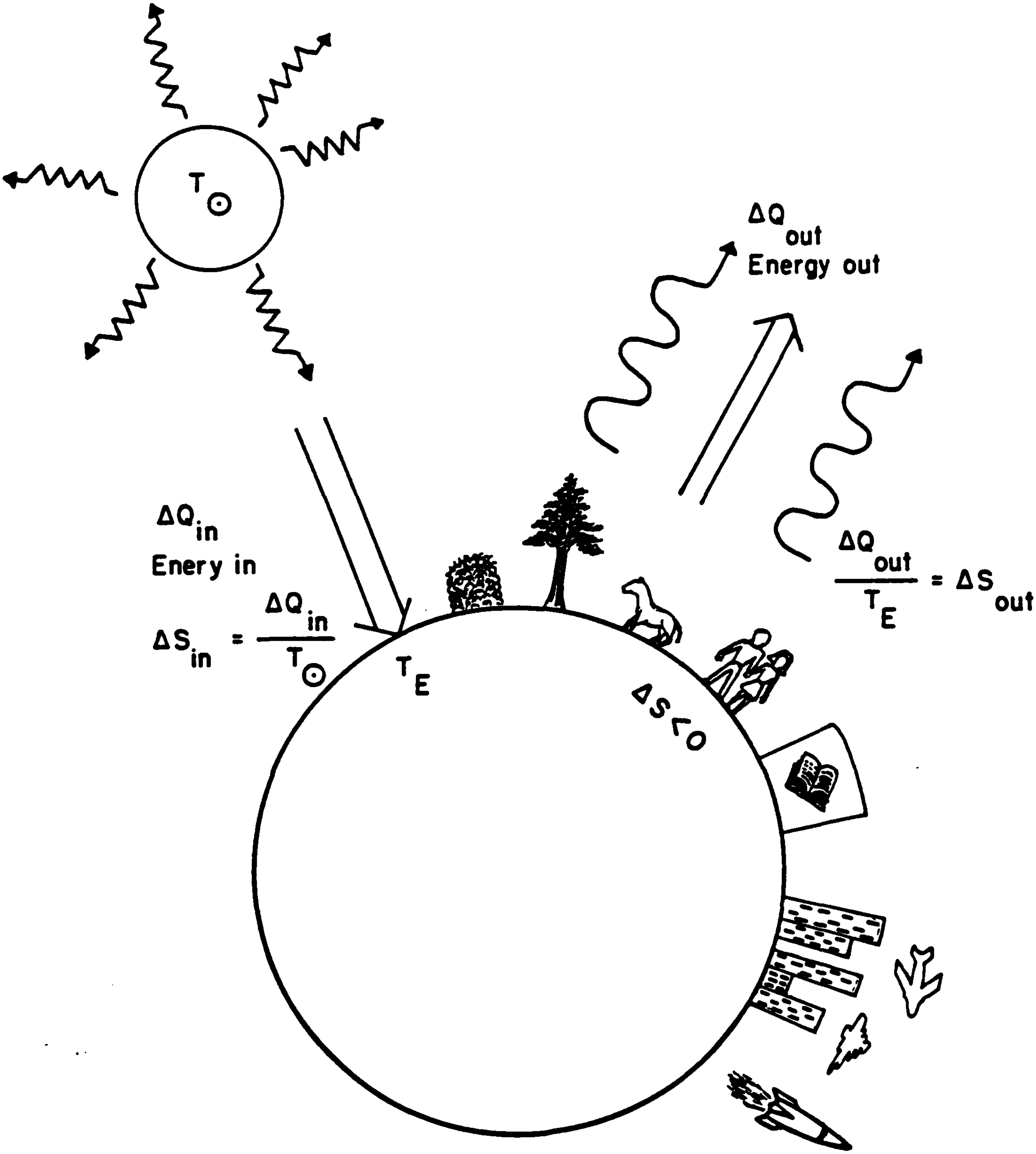


Fig. 32 THE SUN - EARTH SYSTEM

$\dot{\Delta Q} = 1368$ Watt/m²

$T_{\odot} \approx 6000°K \qquad T_E \approx 300°K$

$\dot{\Delta S} = \dot{\Delta S}_{\epsilon} - \dot{\Delta S}_{out} \approx -4.5$ WATT $m^{-2} k^{-1}$

the origin of this high temperature low entropy source?

13.3 The Sun as a High Temperature Low Entropy Source

The sun was formed by gravitational contraction of an interstellar gas cloud some 8-10 billion years ago. In such a gas cloud every atom finds itself in the gravitational attractive force field of all the other atoms. The attractive force is characterized by the gravitational potential energy Ω which the assembled gas cloud has. This potential energy is negative. It is the energy that would be required to remove all atoms to infinity against the gravitational force. For such a gas cloud to be balanced the inward gravitational force must be balanced by an outward directed pressure force. The pressure is due to thermal motion of the gas atoms. So the gas must have internal energy, U. The balance condition can be expressed in terms of the "Virial theorem": Twice the total internal energy plus the (negative) gravitational energy of the cloud is zero

$$2U + \Omega = 0$$

For a homogeneous density gas cloud the potential energy is

$$\Omega = -\frac{3}{5}\frac{GM^2}{R}$$

Where M is the mass and R the radius of the cloud. If the density is not constant then

$$\Omega \propto -\frac{M^2}{R}$$

If the cloud contracts to smaller radius then the potential energy becomes more negative. For every increment $\Delta\Omega$ in gravitational potential energy, the internal energy must grow by

$$\Delta U = -\frac{1}{2}\Delta\Omega$$

to keep the balance. Therefore, the contracting gas cloud will heat up. The gas cloud will contract to higher and higher densities and temperatures until the temperature has reached values of the order of 100 million degrees, in the interior. At these temperatures in the interior, fusion reactions of hydrogen nuclei occur in large numbers, forming the heavier hydrogen nuclei, deuterium and releasing large amounts of energy

$$H + H \rightarrow D + \text{energy}$$

This energy release causes the build-up of temperature and pressure in the interior to such values that the gravitational attractive forces can no longer contract the gas any further. The gas reaches an equilibrium state. The energy produced in the interior finds its way to the surface and is radiated as light. The gas cloud has become a star.

The hydrogen fusion process can go on and provide large amounts of energy for billions of years.

It is this fusion in the stellar interior that provides the long term stationary high temperature low entropy source from which stem all low entropy initial states in our immediate environment.

The establishment of this source in turn is due to the gravitational force and due to the initial presence of a gas cloud that can be contracted by the gravitational force.

There is another aspect that is surprising to many. The complex living structures here on earth depend crucially on the availability of heavier atoms like oxygen, nitrogen, carbon calcium phosphorus etc. These chemical elements were not produced in the early universe which consisted mostly of hydrogen and helium. The heavier element have their origin in the high temperature furnaces of certain types of massive stars which burn the available fuel in the relative short times of 100 million years and end their existence in a gigantic explosion in which the material produced in the interior becomes dispersed into space. This dispersed gas cloud gathers up again by gravitational contraction forming a new star and chunks of matter circling about the new star as planets.

The material we here on earth are made of, has once before gone through the furnace of a stellar interior where it was produced by fusion of lighter atoms.

The question about the origin of the low entropy energy sources has led to the two factors:

(a) The original presence of diffuse gas clouds

(b) The gravitational force's ability to contract gas to become a star with a thermonuclear fuel burning interior.

14. SEARCH FOR THE ORIGIN OF THE ARROW OF TIME IN THE EVOLUTION OF THE UNIVERSE

To this point we have concluded that the arrow of time cannot have its origin in the fundamental laws of physics, which are all time reversible. The small deviation observed in the CP violation is not yet understood in its significance. It does not appear to affect atomic and molecular physics and the behavior of large systems.

The thermodynamic arrow of time as expressed in the second law finds its root in statistical dynamics not in the principles of dynamics, which are time reversible but through an additional assumption about the distribution function which is equivalent to coarse graining and discarding of dynamical information. This coarse graining is a sufficient model element to account for the minute uncertainties of dynamical information in the individual collision processes with molecules which have finite size. These uncertainties lead to the H-theorem and the second law of thermodynamics, guaranteeing that a system in a low probability low entropy initial state will rapidly approach the high probability thermal equilibrium state (see appendix D).

That we observe second law behavior in our environment has its origin in the fact that by some mechanism low entropy initial states are continuously being set up. We saw that in our immediate environment man is responsible for setting up such low entropy states. He does so because of availability of low entropy energy which he harnesses and directs in engines in such a way as to create other low entropy states which are useful to him. In these processes a fraction of the used up low entropy energy becomes degraded to higher entropy energy. The efficiency of these processes is governed by the Carnot efficiency at the optimum.

We saw that man himself and with him all of organic life on earth can be seen as a complex low entropy product of the great engine represented by the sun-earth system, in which low entropy energy from the solar interior is degraded in entropy allowing in principle the construction of low entropy complex structures on earth. It should be clear that this only assures that the origin of life on earth is not forbidden by the most fundamental laws of physics. It says nothing about all the necessary conditions and constraints that must be met to make the evolution of life actually happen. Our knowledge about basic processes in large complex molecules and the origin of cell structures is much to fragmentary still. We do know that already 3.5 billion years ago, only a billion years after the earth had formed and shortly after the earth had cooled enough that the atmosphere rained out most of the water that fills the great ocean basins, the first single cell organisms, the blue green algae made their appearance.

We have traced the occurrence of starlike low entropy sources to the law of universal gravitational attraction and the availability of diffuse gases that can contract and heat up.

The understanding of stellar structure has grown to the point that it became clear during the last two decades that the existence of stars depends in a very sensitive way on the values of some fundamental constants of nature. A slight increase in the fundamental electromagnetic force, a slight decrease in the electronic's mass or a slight increase in the very weak gravitational force would result in stars contracting to the size of planets without ever getting hot enough to start fusion burning. Reversal of these small deviations in the other direction would lead to formation of blue super giant stars which would burn out in less than a million years or become black holes.

In either case, ordinary long enough living stars would not be possible. Many details of these fascinating insights are described in a book by Barrow and Tipler: "The Anthropic Cosmological Principle" [BA & Ti 86].

So the observation of the second law behavior that provides for us an arrow of time has led us to ask three questions:

(a) Is there anything in the nature of gravitation and its role in the evolution of the universe that provides an arrow of time?

(b) What mechanism provided the diffuse matter in the early universe that could contract to provide stars?

(c) Is there any indication that the universe at its birth might select a set of natural constants from a random set and we happen to be accidentally observers in a universe in which the natural constants were picked just right to make us possible in this universe with other not so fortunate universes equally possible?

The first two questions we will approach in the next sections, following closely a line of arguments given by Penrose [PE 89]. The third question is presently beyond reach of any theorizing about the origin of the universe in which we live.

15 COSMOLOGY - BIG BANG - THE STANDARD MODEL

15.1 Einstein's General Theory of Relativity - Cosmology

In 1915, Einstein published his famous papers on general relativity, a new theory of gravitation and cosmology. He requires that the laws of physics should not only be unchanged for observers moving with constant velocity with respect to each other, which was the subject of his special theory of relativity, but they should be the same even for observers that are accelerated with respect to each other. Being in a laboratory that is accelerated, an elevator moving up for instance is principly indistinguishable by any experiment from being in a gravitational force field. So, gravitational force could be interpreted as being in an accelerated reference frame. The cause of this accelerated motion is not a force in Einstein's theory but the inertial motion in a curved space. If one lets a marble move with large speed in a large fish bowl one will see the marble move along a curved trajectory prescribing some kind of orbit, constrained by the curvature of the surface on which the marble is moving. It is not a gravitational force that keeps the marble in an orbit but the constraining curvature. On this basic idea Einstein formulated a complete theory of gravitation. His theory describes planetary motion in the curved space around the sun as well as Newton's theory did with some minor modifications. The rotation of the perihelion of planets, the light bending when grazing by the sun and the red shift of the light due to the gravitational curvature affect were all brilliantly confirmed by observation in the years following the publication of his theory.

It should be noted that this changed view of gravitation was not merely by the whim of not liking Newton's theory but because there was a serious problem with Newton's law of gravitation. The way it stands, Newton's law of gravitation postulates an action at the distance. The force acts from the attracting mass as if by magic over large distances instantaneously on the attracted mass. According to the special theory of relativity no action can propagate over distance faster than with light velocity. Therefore a revision was needed.

With the formalism of general relativity, Einstein was able to construct a model of the universe as a whole. Einstein originally envisioned a static universe in which all masses are held in their place by an overall curvature of the universe space.

One can visualize this concept by assuming that space is 2-dimensional rather than 3-dimensional and is represented by the surface of a balloon [Fig 33]. All physical processes in this 2-dimensional universe are entirely confined to the

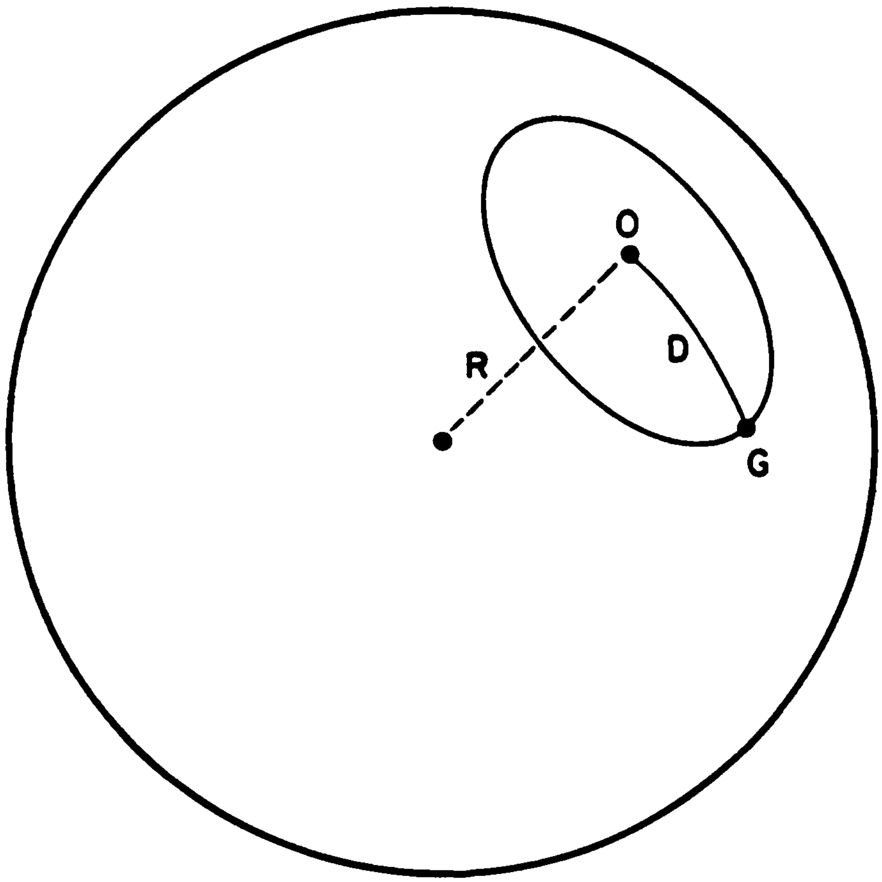


Fig. 33 Balloon model of Einstein's static universe. R is the "world radius", D the distance between observer O and galaxy G.

surface of the balloon. Masses like stars and galaxies are dots on this surface and light signals propagating from one galaxy to another are ripples that propagate with light velocity along the curved surface. The radius R of the balloon represents the "radius of the world". The distance D between the observer O and a distant galaxy G is a piece of a large circle on the balloon, the shortest connection between O and G. As far as our observations of the universe have reached we find this universe, and in our model the balloon surface, fairly homogeneously populated with galaxies. In Newtonian gravitation all these masses would tend to get closer to each other due to the gravitational forces. In Einstein's theory this force manifests in the tendency for the radius R to decrease. Einstein corrected this shortfall of his model by introducing the so called "cosmological term" into his equations, which amounts to a repulsive force in Newtonian language, which tends to increase the radius of the world. This allowed Einstein to construct a balanced static model of the universe. The repulsive cosmological force is like a negative surface tension that tries to expand the balloon. It is characterized by a "surface tension" constant Λ. The attractive force of the masses is characterized by the average mass density ρ of all the masses, which are assumed to be homogeneously smeared over the balloon surface. Einstein's theory gives between the world radius R, the "surface tension" Λ and the average mass density, ρ the relations

$$R = \Lambda^{-1/2} = \frac{c}{2\sqrt{\pi G \rho}}$$

From observations one estimates $\rho - 10^{-31}$ gr cm^{-3}. This gives

$$\Lambda \approx 10^{-58}\ cm^{-2} \qquad R \approx 10^{29}\ cm \approx 10 \text{ billion Ly}$$

15-2 The Hubble Effect - Expanding Universe

In 1929 Edwin Hubble, using the largest telescope then available, made the observation that the characteristic spectral lines of the light from distant galaxies all showed a shift to the red end of the spectrum form their usual laboratory position. Such red shift if normally associated with the source receding with some velocity from the observer with an amount of shift proportional to the velocity of recession.

The remarkable fact of Hubble's observation was that all galaxies show this red shift and it is the larger the further away the galaxy is from us. All galaxies seem to be receding from us the faster the further distant they are. This is the content of Hubble's law.

In our balloon model and in Einstein's theory the natural interpretation is that the balloon is not balanced but that it is expanding at some rate. In the balloon model one can see that by simple geometry it must be

$$\frac{\dot{D}}{D} = \frac{\dot{R}}{R}$$

where $\dot{R}$ is the rate of increase of the "world radius" and $\dot{D}$ is the rate of increase of the distance between observer O and the galaxy G, i.e. the apparent velocity of recession of the source.

The ratio $\dot{R}/R$ is called the "Hubble constant H", the relation then can be written

$\dot{D}$ = "velocity of recession = $H \cdot D$

which is exactly what Hubble observed.

There is an important conclusion to be drawn from this model. The observer O and galaxy G can be thought of as being glued to the surface of the balloon. They do not move in their space. As R increases, their distance is increasing not because one is actually moving away from the other but because the space in which they are located is stretching.

In general relativity space is no longer just the arena into which the events are embedded like in Newtonian physics, but rather space itself has become a dynamic object that undergoes changes, expands, contracts, warps.

That Einstein's theory does have solutions in which the "world radius" R changes in time, was discovered already seven years before Hubble's observations by Friedmann [FR 22]. These model

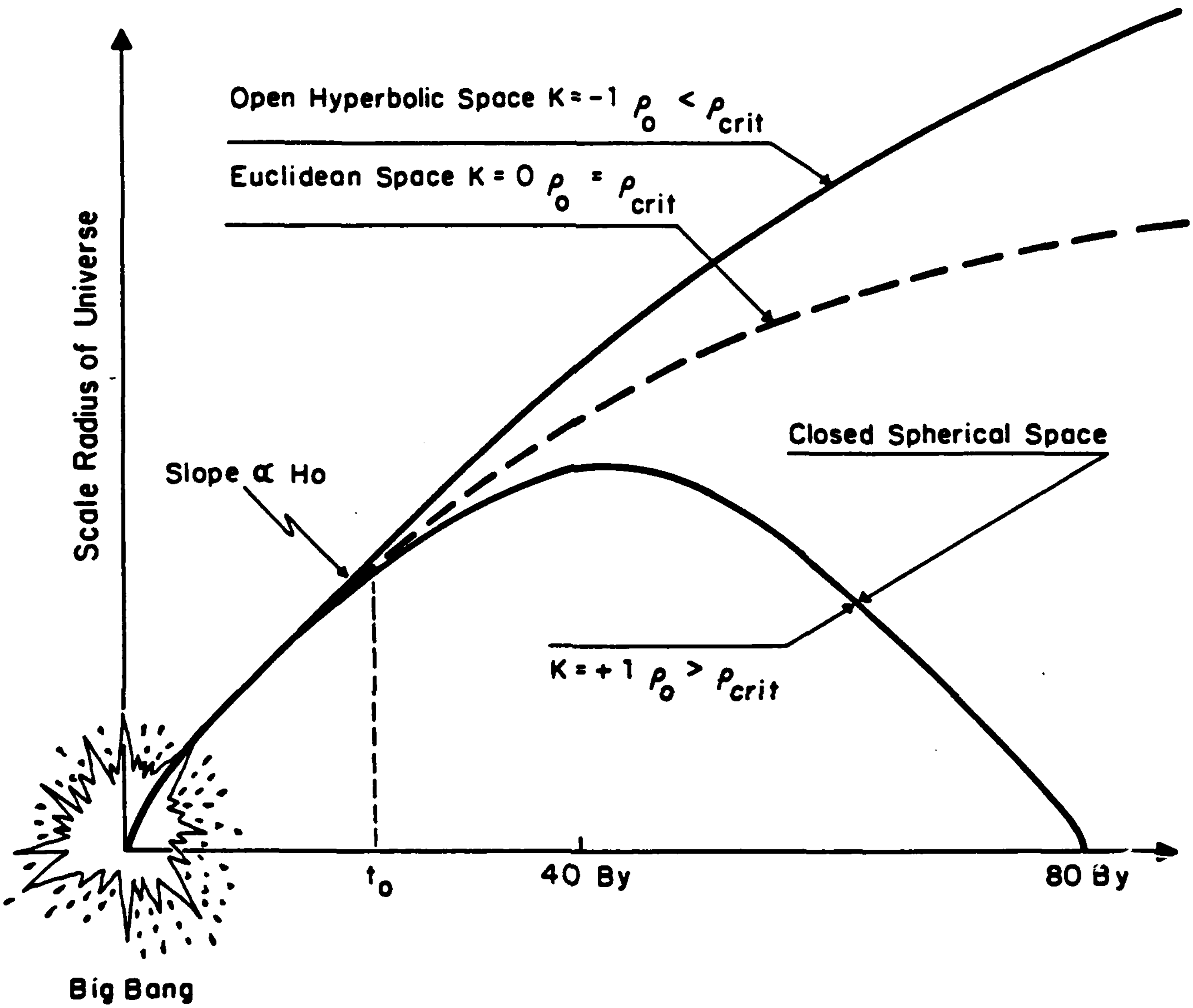


Fig. 34 Evolution of three types of Big Bang universes

H_O = Hubble constant of present expansion rate

ρ_O = Present average mass density

$\rho_{crit} = 3 H_O^2 / 8 \pi G$ critical mass density for closure of expanding space exist even when the cosmological term Λ is assumed zero. There are three types of models. They are distinguished by the type of geometry of the universe as spherical, euclidean and hyperbolic spaces. All have in common that they show an initial expansion beginning from $R = 0$ at $t = 0$. These are the Big Bang Models. After some time their evolutions diverge depending on the type of geometry [Fig 34].

(a) The spherical universe is very well modeled by our balloon model. It has positive curvature ($k=+1$). Like the balloon, it has a finite volume and is closed, i.e. large circle trajectories are closed orbits. The total mass in this universe is so large that the gravitational attraction eventually wins over the initial expansion which stops and is followed by a recollapse of the universe to zero radius. The mass density today ρ_0 would be larger than a certain oritical value.

(b) The hyperbolic universe cannot be realistically modeled similar to a balloon. It has negative curvature ($k = -1$). It is sparsely populated with mass. It is infinite in extent. Its initial expansion will slow down in time due to the gravitation but never cease. That universe will continue to expand. The mass density ρ_0 is smaller than the critical value.

(c) The Euclidean space ($k = 0$) is the limiting case between the spherical and hyperbolic case. It has zero curvature. It is infinite in extent and will continue to expand like the stretching of a flat rubber sheet. The observed density is equal to the critical density.

The present data base is the following:

(1) Although the observed mass distribution of galaxies shows accumulations and large voids one may consider the observed portion as having essentially homogeneous mass distribution

(2) The Hubble constant is

$$H_o = 100 \cdot h \quad Kms^{-1}\,Mpc^{-1}$$

where $0.4 \le h \le 1.0$

The critical density is related to the Hubble constant by

$$\rho_c = \frac{3H_o^2}{8\pi G}$$

For the above value

$$\rho_c = 1.88 \times 10^{-29}\, h^2 \quad gr\,cm^{-3}$$

(3) The best estimates of the luminous masses in the universe gives

$$\rho_{oLUM} \lesssim 0.01\, \rho_c$$

There has been growing evidence over the last decade that galaxies must have dark matter haloes. One estimates now

$$\rho_o \approx 0.1\rho_c$$

(4) The cosmological constant is very small

$$\Lambda \approx 10^{-58} \ \text{cm}^{-2}$$

On the basis of these values for the density one would conclude that the universe is open hyperbolic.

15-3 Cosmic Background Radiation - Hot Big Bang

In 1965 Penzias and Wilson [PER Wi 65] found that the universe must be filled with a black-body-spectrum microwave radiation of a corresponding temperature of $T = 3°K$. This radiation is highly isotropic (equally strong from all directions) with a variation of less than $\frac{1}{100}$%.

Such radiation had been predicted by Gamov in the forties. The argument for and interpretation of this radiation is that at very early epochs of the universe, when the mass density was nearly that of stellar interiors, there must have been nuclear reactions going on with copious production of radiation. As the universe expanded, this radiation began to cool. Eventually the temperature became too low for particle pair creations and the particle-anti particle annihilation into photons prevailed, leaving only a "small" amount of particles and many photons. The background radiation contains approximately 10^{11} photons for each heavy particle. Eventually the radiation became so cool that it could no longer ionize hydrogen atoms. At that point the photons could no longer interact with matter in a significant way and remained essentially decoupled. At present these photons have cooled down to a temperature of 3°K.

The thermal history of the universe prior to the decoupling of the photons is dominated by the photons. The temperature, density and radius are approximately related by these relations

$$T \approx \frac{1.2\times10^{10}}{\sqrt{t}} \quad ; \quad \frac{R}{R_0} \approx \frac{T_0}{T} \quad ; \quad \rho \approx \rho_0\left(\frac{R}{R_0}\right)^{-4}$$

where R_0 T_0 and ρ_0 are the world radius, the radiation temperature and the radiation equivalent mass density at present. Their values are

$$R_0 = 1\times10^{29}\mathrm{cm} \qquad T_0 = 3^{\circ K} \qquad \rho_0 = 6\times10^{-34}\ \mathrm{gr\,cm^{-3}}$$

With these data and the relations one can give some markers for the thermal history of the early universe.

t [sec]	$T^{\circ K}$	$\frac{R}{R_o}$	ρ[grcm^{-3}]	
10^{-4}	$1x10^{12}$	$2.5x10^{-12}$	$1.5x10^{13}$	large amounts of heavy particle antiparticle pairs annihilate producing high energy radiation
4	$6x10^{9}$	$5x10^{-10}$	10^{4}	Electrons, Positions annihilate neutron/proton ratio freezes to 1:5 neurtrinous are decoupled
180	$1x10^{9}$	$3x10^{-9}$	4.7	neutrons bind with protons and form light nuclei. He-abundance established at 27%, H^2 and He^3 abundances also established
400,000 yrs = $1.2x10^{13}$sec	4000	$8x10^{-4}$	$1x10^{-21}$	protons and electrons recombine to neutral hydrogen gas, photons decouple from matter

15-4 <u>The Thermodynamic and the Cosmological Arrow of Time</u>

We had previously concluded that the occurrence of the second law can be traced to the gravitational contraction of masses and the availability of diffuse gas clouds that can be contracting to form stars.

The matter in the universe becomes a neutral gas at $t \approx 400,000$ yrs. The temperature then was nearly that of the solar surface, the density was 10^{-21} grcm^{-3}, very small compared to solid densities. The world radius was the equivalent of 80 million light years.*

At that time the cosmic background radiation had its last scattering with matter. The isotropy of this radiation is a measure for the density fluctuations in the gas at that time. This relationship depends on the nature of the medium. For current models of structure formation the relationship is given as [KO-TU 90]

$$\left(\frac{\delta\rho}{\rho}\right) \approx C\left(\frac{\delta T}{T}\right)$$

where C = 10 - 100. Given the observed microwave anisotropy of $\delta T/T \approx 10^{-4}$ it follows that

$$\frac{\delta\rho}{\rho} < 10^{-2}$$

The universe at the decoupling of the radiation was an extremely smooth gas. The diffuse homogeneous gas was indeed a product of the Big Bang. How the formation of structure occurred is not fully understood yet.

This then would seem to be the answer to the question what brought the second law into the universe.

---

* The often asked question, how can the universe expand 80 million light years in 400,000 yrs if nothing can propagate faster than light velocity must be answered by saying that matter in the expanding space is stationary. What expands is space itself, which must not be compared with objects propagating <u>in</u> the space. Near the singularity the expansion rate becomes infinite.

The universe at the Big Bang started out in a state which soon led to a very homogeneous gas of about 70% hydrogen, 27% helium and traces of other light nuclei. This gas some time later began to clump into galaxies and stars by gravitational attraction.

Penrose has pointed out, that this cannot be the final answer. The argument goes in three steps.

(i) The primordial fire ball of the Big Bang was a state of thermal equilibrium of extremely high temperature and density. Because it was an equilibrium state it was a state of maximum entropy. This raises the question of how such a maximum entropy state can develop further increase of entropy.

(ii) The maximum entropy of a volume of gas increases with volume

$$S_{max} \propto \ln V$$

That means that as the universe expands the maximum attainable entropy would increase very rapidly and one might think that perhaps the processes in the gas that drive the system to maximum entropy could be too slow to keep up with the expansion of space and that there would be potential for subsequent increase of the entropy up to the new-increased maximum entropy.

(iii) Penrose argues that this cannot be the generally correct answer because it leads to difficulties in the case of a closed spherical geometry universe. In the collapse phase of the universe the maximum entropy would have to decrease. If the low entropy at the beginning is the cause for the second law behavior then this should apply to all phases of the universe.

However in recollapse the entropy maximum decreases in value.

Presumably sometime after the beginning of the contraction the entropy which had been trying to catch up with the ever increasing maximum value will finally catch up. From then on the gravitational contraction would continuously lower the entropy of the universe.

This conclusion was already reached by T. Gold in 1962 [Go 62]. Gold concluded further that in the recollapsing universe the direction of time would reverse. With this assumption the thermodynamic arrow of time defined to be the direction in which the entropy increases would continue to coincide with the cosmological arrow of time defined by the direction in which the universe is expanding.

Penrose finds the idea of reversal of the arrow of time at the point of largest expansion in a closed universe unacceptable.

S.W. Hawking [HA 85] in a paper "Arrow of Time in Cosmology" defends the notion of the reversal of time on the basis of his version of quantum-cosmology and the wave functions of the universe proposed by him.

Penrose's arguments employ the properties of collapsing masses within our space, black holes. He points out that if collapse of matter were to imply reversal of the time direction then an astronaut falling into a black hole would experience time reversal. There are no experiments to test that. However, Penrose argues that the properties of black holes do not support the notion of time reversal and he hypothesizes that the initial singularity of the big bang and the final singularities in black holes and the big crunch are fundamentally different.

We will describe the important properties of black holes and of space-time singularities in the next sections.

15-5 Cosmological Constant - De Sitter Space and Time Inflation Scenario

In the discussion of the standard cosmological model the cosmological term was neglected. The reasons for the neglect are two fold. First, we saw that according to best estimates the constant is very small at present. Secondly, the introduction of this term by Einstein was rather ad hoc, simply to be able to produce a solution of his equations which represents a static universe. Yet there was no observational evidence for a universal repulsive force. Also, the static solution he obtained turned out to be instable against small perturbations. With the discovery of the dynamic Friedmann solutions and the Hubble effect the cosmological term became less interesting.

The cosmological term, however, plays a crucial role in the "Early Inflation Scenario" which most cosmologists assume now as a necessary ingredient in the model of the initial expansion, in order to explain certain puzzling features of the standard model. Before we describe the inflation hypothesis we need to take a look at cosmological models in which a cosmological constant dominates.

It was already mentioned that the cosmological term acts like a general repulsive force. In the expanding balloon model it acts like a negative surface tension of the balloon skin. This negative surface tension represents an energy per unit area. With the force $F_{cosmol} \propto \Lambda R$, the corresponding potential energy is $U_{cosmol} \propto -\Lambda R^2$. The gravitational potential energy in Einstein's theory is $U_{grav} \propto -\rho R^2$. This comparison shows that energetically the cosmological constant acts like an additional mass density $\rho_{VAC} = \frac{\Lambda}{8\pi G}$. This equivalent mass density and corresponding energy density is present even if the universe contains no real mass. In that case the gravitational attraction on the balloon surface is zero and only the negative tension acts which leads to an accelerated inflation of the balloon. The radius of the universe grows exponentially:

$$R(t) = R_1 e^{\sqrt{\frac{1}{3}\Lambda}(t-t_1)}$$

This so called De Sitter space would have no real mass content but it contains through the cosmological constant a constant equivalent local energy density

$$\rho_{VAC} c^2 = \frac{\Lambda c^2}{8\pi G}$$

Today this energy density is interpreted as the non-zero vacuum energy density of certain quantum field theories of matter of the type which have proven highly successful in the attempts to come to a unified understanding of the basic building blocks of matter and the fundamental forces between them.

Associated with this vacuum density is a vacuum pressure which is equal to the vacuum energy density but negative. As the universe increases the cosmological term energy becomes more negative.

$$dU_{cosmol} = P_{VAC}\, dV < 0$$

But the volume increases, so the pressure must be negative. This peculiar sort of "ether" that is now thought to be possible, ascribes to the vacuum state of space, i.e. absence of any ordinary material particles a non-manifest form of energy, analogous to the latent heat that is present in water vapor. Upon condensation of the vapor, i.e. during a phase change of the vapor this latent heat becomes available as free energy. That same heat of vaporization must be supplied to the liquid at its boiling point to convert it into vapor. It is the energy needed to break the van-der Waal attraction between the water molecules. Similarly it is now believed that a latent form of energy can be present in space that can condense out to form material particles.

This interpretation of the cosmological constant and the De Sitter space became the principal ingredients for the so called inflationary scenario.

There were a number of problems plaguing the standard model. One was the observation that the observed mass-energy density ρ_0 in the universe is very close to the critical density, so that the total energy is close to zero, which would require an extremely accurate initial matching of kinetic energy and negative potential energy at the moment of the big bang. This is called the flatness problem. The second is the observation that the cosmic background radiation arriving today from opposite directions in the universe is isotropic to such a high degree. We had mentioned already earlier the high degree of uniformity of the whole universe at the time of the last radiation scattering. Yet one can show, when this primordial radiation was emitted the sources at opposite sides of the universe from us could not have been in thermodynamic equilibrium. They could not have communicated energy enough in the time available. In a rapidly expanding universe each observer has an absolute horizon such that radiation from sources beyond that horizon would take more time to reach the observer than is available since the Big Bang. This problem is referred to as the horizon problem.

The standard cosmological model gives for the entropy density in the early universe:

$$\frac{S}{L^3} \equiv s = \frac{2\pi^2}{45}\,\frac{k_B^4}{(\hbar c)^3}\,N(T)\,T^3 = 4\times10^{-15}N(T)\,T^3\,\mathrm{erg\,cm^{-3}K^{-1}}$$

Where N(T) is the number of particle degrees of freedom and is the order of 100 in the early universe. The size L is the size of any "patch" on the balloon we wish to consider and S is the entropy in it.

As a measure of deviation from a Euclidean space the standard model gives the relative mass density deviation from the critical density:

$$\left|\frac{\rho(t)-\rho_{crit}^{(t)}}{\rho(t)}\right| = 1.4\times10^{-56}\,N^{-1/3}(T)\left(\frac{E_{PL}}{k_B T}\right)^2$$

where $E_{PL} = 2x10^{16}$ erg $=10^{19}$ GeV is the "Planck energy".

At the time $t_r = 10^{-36}$ sec the temperature in the standard model is given by $k_B T_r = 10^{15}$ GeV which gives $|\ (\rho - \rho_{crit})/\rho\ | \approx 3\times10^{-49}$. The density is extremely close to critical density. Kinetic and potential energy of the universe are extremely finely tuned to nearly total energy equal to zero. This is an extraordinary result that raises the question, what mechanism could have initiated this remarkable fine tuning of the universe.

The horizon distance $D_H(t)$ is that distance from which a signal that left at t=o reaches the observer at time t. Let L(t) be the balloon patch size at time t that will evolve into the present universe size $L_p \approx 10^{28}$cm. Then at time t the ratio of the horizon volume to the patch size volume is given by the standard model as:

$$\left(\frac{D_H(t)}{L(t)}\right)^3 \approx 4\times10^{-86}\left(\frac{E_{PL}}{k_B T}\right)^3$$

At the time $t_r = 10^{-36}$ sec this ratio is 4×10^{-74}. That means that, the patch at 10^{-36} sec that evolves into our present universe consisted of 10^{73} causally disconnected patches and it is not possible to explain how our present universe can be so uniform, as evidenced by the uniformity of the cosmic background radiation.

We had seen earlier that in the early universe

$$R = 3\times10^{19}\ t^{1/2} \qquad T = 10^{10}\ t^{-1/2}$$

This shows that RT = const, and that therefore the entropy $S \propto (RT)^3$ is a constant during the early expansion. Under this assumption the flatness problem and the horizon problem persists for all times and leads to the difficulty of having to explain very special initial conditions for the universe. Comparison of the entropy density and the expressions for the degree of flatness and the horizon volume shows that:

$$|(\rho-\rho_{crit})/\rho| \propto S^{-2/3}$$

$$\left(\frac{D_H}{L}\right)^3 \propto S^{-1}$$

This observation led A. Guth [GU 81] to the proposal that these problems would disappear if the universe prior to a time t_r would undergo a De Sitter inflationary expansion during which the energy density in the expanding space is dominated by the vacuum energy. The space would exponentially grow by a factor $Z \approx 10^{28}$ during this expansion. The temperature would correspondingly drop exponentially by a factor 10^{28}, leading to a "super cooled" universe. At time t_r the vacuum energy would change its phase. The "latent heat" would become available and lead to a dramatic reheating of the universe to a temperature a factor 10^{28} larger. From then on the universe would evolve like an ordinary standard model. In the super cooled state prior to the reheating the entropy density is a factor 10^{56} smaller than after reheating and the flatness quantity would be

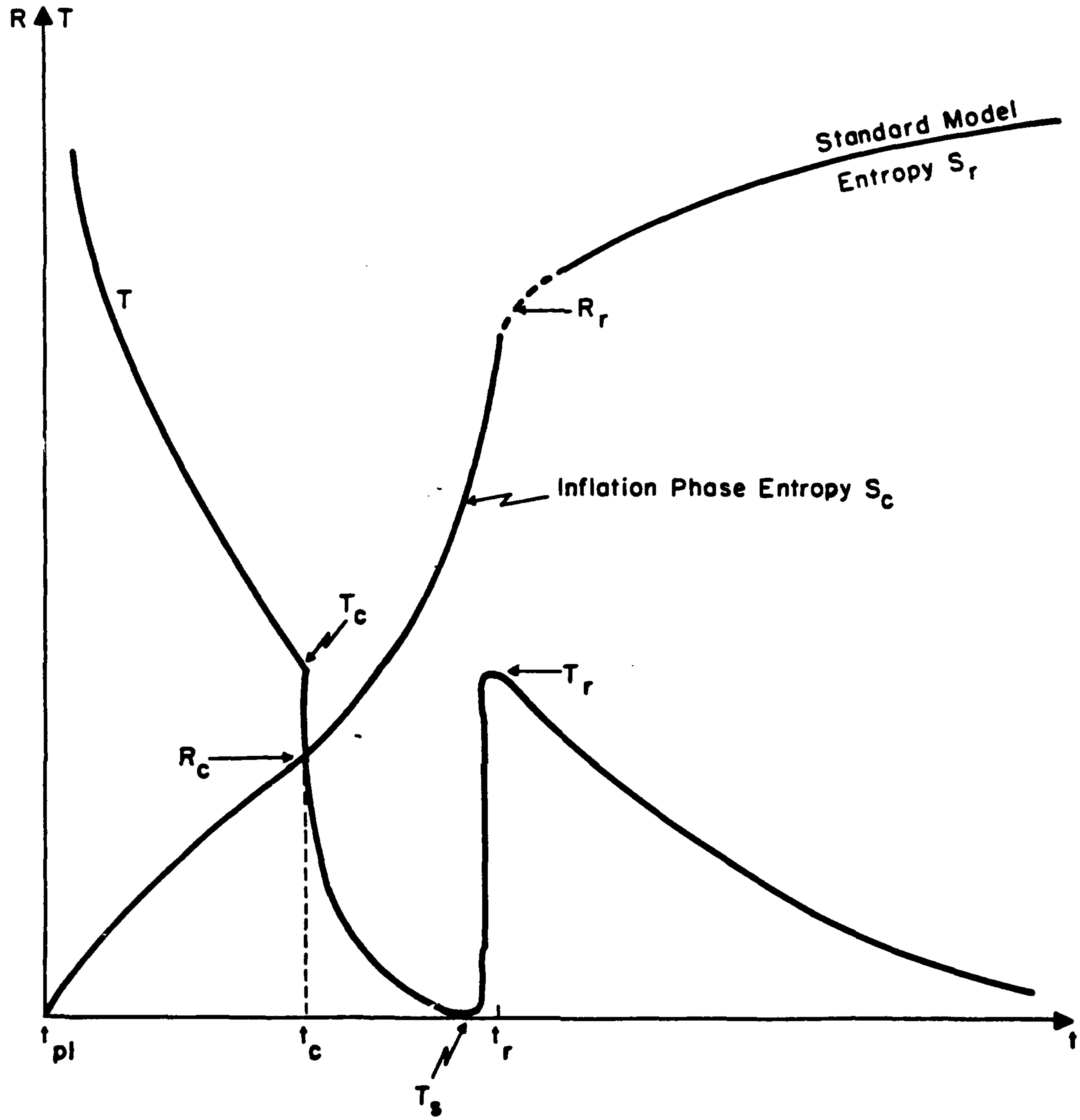


Fig. 35 THE EARLY INFLATION SCENARIO PHASE TRANSITION AT TIME t_c LEADS TO EXPONENTIALLY EXPANDING DeSITTER SPACE WITH SUPER COOLING. AT t_r PHASE TRANSITION LEADS TO RAPID REHEATING AND ENTROPY INCREASE.

of the order one, thereby eliminating that puzzling problem. The horizon volume ratio would be a factor 10^{84} larger which makes the horizon much larger than the patch which becomes our portion of the universe, thereby eliminating the horizon puzzle.

The inflation scenario so-to-speak decouples the present state of the universe from the conditions at the very moment of the origin of the universe. It is no longer necessary to worry about very special initial conditions. The early inflation creates a patch of very low entropy density.

The original version of an inflation by Guth had certain problems associated with the exact dynamics of the inflation, but the idea prevailed. Other better models were proposed. Although it is not possible today to point to the "right" model, because certain aspects of particle physics which are crucial to the story are not known yet, most cosmologists subscribe to the basic idea of an inflationary period sometime around 10^{-36} to 10^{-32} seconds.

At $t = 10^{-36}$ sec the standard model gives a radius of the universe of $R \approx 30$ cm. If this is the end point of the inflation, then inflation would start at 10^{-39} sec at $R \approx 10^{-27}$ cm.

16 BLACK HOLES - END OF TIME

16-1 The Concept of Black Holes

Newton's law of gravitation gives for the gravitational acceleration at the surface of a planet

$$g = \frac{GM}{R^2}$$

Where M is the mass of the planet, R the radius at its surface and G the universal gravitation constant. The weight of a mass m at the surface is

$$W = mg$$

From this one can see that the weight of a mass m will be larger when the same planetary mass M is squeezed into a smaller size sphere. If we think of the radius R shrinking down to smaller and smaller size the weight of the test mass m would continue to increase. Could this be done indefinitely? One consequence of having the surface gravity g increase is that it becomes more and more difficult to project a test mass m with a large enough velocity upward that it will escape from the planet. The characteristic escape velocity V_{esc} is given by the requirement that the kinetic energy of the projectile must exceed the total amount of work needed to lift the test mass against gravity from the surface of the planet to infinite distance. It must be kinetic energy $\frac{1}{2} m V_{esc}^2 = \frac{mMG}{R}$

This gives $$V_{esc} = \sqrt{\frac{2GM}{R}}$$

For the earth this value is 11.2 Km/s.

How small could one make R? The largest possible projectile velocity would be light velocity. So there is a limiting radius beyond which no object, light included, could escape from the surface of the compacted planet. That radius is called the Schwarzschild radius of the mass M.**

$$R_s = \frac{2GM}{c^2}$$

For the earth this radius is 0.8 cm, very small indeed. For the sun it is 3000 m. At the surface of the sun, collapsed to this radius, a body would have 1000 billion times the weight it has here on earth. Any body under this weight will crush under its own weight. In fact, one can show that this whole sun, if it had collapsed down to its Schwarzschild radius cannot support itself against its own gravity force. It will collapse further. There are no known physical principles that could prevent total collapse to a single point with infinite density. The mass will collapse to a singularity. Every notion of space and time loses its meaning at that singularity. One may

---

* The above derivation is done using Newtonian physics. At the Schwarzschild radius the gravitational force is so large that general relativity must be used instead. The result however turns out to be identical.

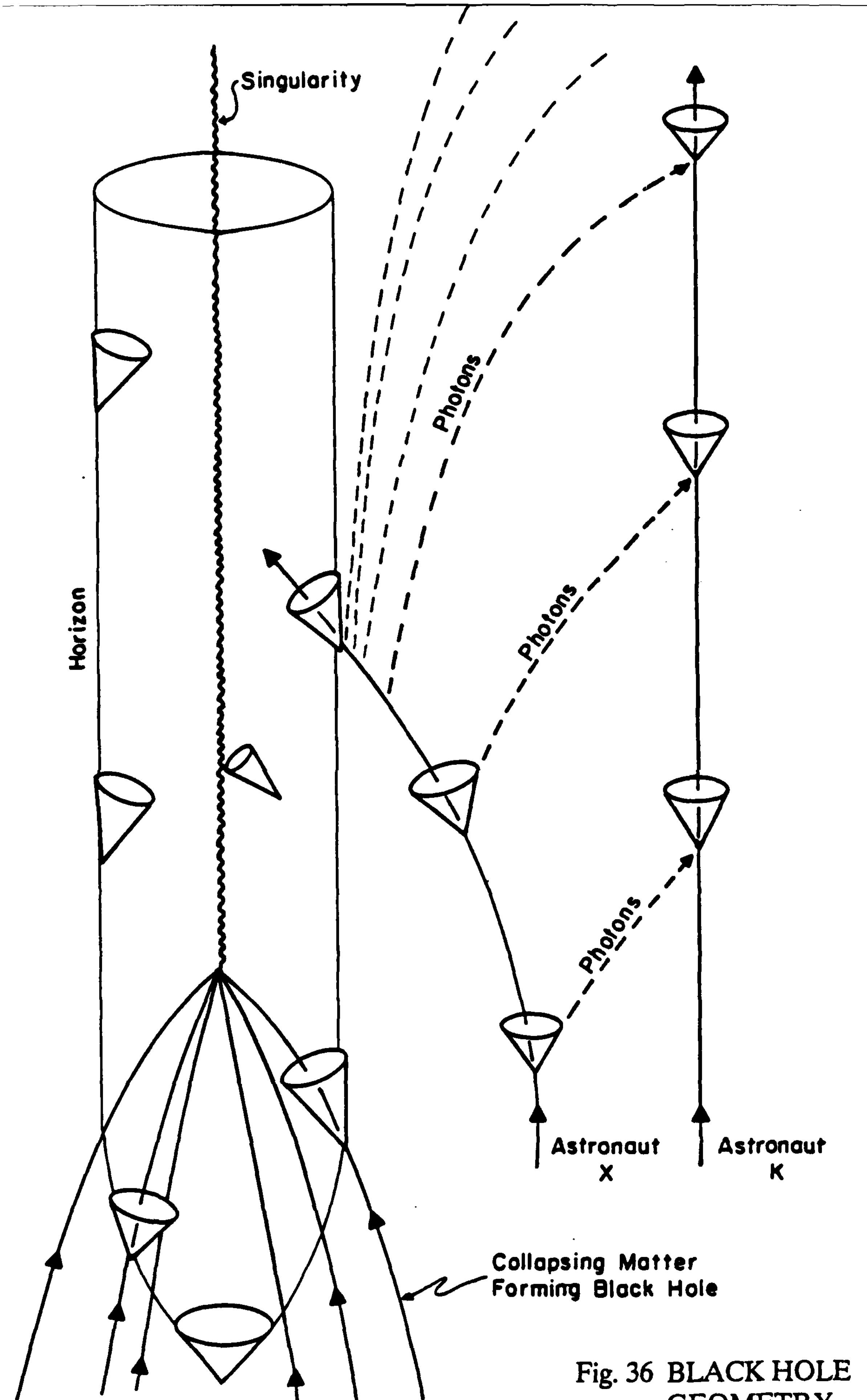


Fig. 36 BLACK HOLE GEOMETRY

describe the peculiar geometry in the vicinity of such a black hole singularity in terms of the experience of two astronauts in two space ships passing the black hole [PE 89]. Astronaut Kai keeps a respectful distance from the hole. In the diagram [Fig 36], the vertical direction is time and the horizontal directions represent the spatial dimensions. The little cones represent the space-time lines of propagating light signals. Far away from the horizon, which is given by the Schwarzschild radius of the collapsed total mass, the space is nearly euclidean. Closer to the horizon the light cones are tilted toward the center of the black hole, the more the closer one comes to the horizon because of the space curvature/gravity. Right at the horizon light can at most stay at the same place if directed outward. Slightly outside the horizon photons will very slowly climb out of the highly curved geometry near the horizon. Astronaut Xavier could be emitting signals as he moves closer to the horizon. These signals take longer and longer to reach tho observer K. As astronaut Xavier reaches and crosses the horizon the last signal takes an infinite time. Observer K actually never sees X crossing the horizon. If X had agreed to send signals every minute, K would receive these signals at longer and longer intervals and would have to conclude that X's clock has slowed down dramatically. The fool hardy astronaut X will notice nothing particularly strange when he crosses the horizon (provided the collapsed mass is very large and R_s is very large so that tidal effects are small. At the Schwarzschild distance from a 10^6 solar masses black hole the astronaut would experience a differential force along his body which is the weight of 10 grams). He is free falling toward the singularity and like the astronaut in a free falling satellite he will experience no gravity. He will look toward his companion K and will see him clearly all the time. Astronaut X finds himself like in a collapsing universe, together with all the mass falling into the black hole.

While stars represent a low entropy state because they are a source of energy at very high temperature, the opposite is true for black holes. They are no source of energy at all. No radiation appears to escape from the horizon. In the contrary, black holes are enormous sinks of energy into which all forms of mass/energy in their neighborhood are drawn and disappear. Black holes are states of very large entropy.

If black holes are viewed as thermodynamic systems then one can ask what the internal energy and the entropy of the system is. Then it would also follow that one would have to ascribe some kind of temperature to the system. The laws of black holes thermodynamics were given by Beckenstein [BE 73][BE 72] and Hawking.

The entropy of a black hole is given by

$$S_{BH} = 4\pi \frac{k_B G}{\hbar c} M^2 = 2.7\times10^{16} k_B M^2 = 1.0\times10^{77} k_B \left(\frac{M}{M_\odot}\right)^2$$

and the temperature is proportional to the surface gravity of the black hole horizon and can be expressed as [see appendix E]

$$T_s = \frac{\hbar c^3}{64\pi G k_B M} = 10^{-7} \frac{M_\odot}{M} \quad ,$$

where M is the mass of the black hole, $M_\odot$ the solar mass for comparison and k_B the Boltzmann constant. For a solar mass block hole the entropy is astronomically large compared to the entropy of our 10^{23} molecules in a box, which is 10^{23} k_B, or even the total entropy production of the sun over its entire life of 10 billion years which is approximately 10^{63} k_B.

On the other hand, the temperature at the Schwarzschild horizon is 10^{-7}°K, which is incredibly small even with respect to the black body cosmic microwave background radiation which has a temperature of 3°K. The black hole would continuously contribute to the entropy increase by energy flowing from the environment into it.

However, the fact that the horizon must be attributed with having a non-zero temperature at all must mean that this surface is radiating like any object with a temperature does. It was one of the great discoveries in theoretical physics of the last twenty years when Hawking gave the explanation for the mechanism that indeed produces such a radiation. This radiation is referred to as "Hawking Radiation" and is due to a quantum mechanical phenomenon which is called TUNNELING [HA 75]. The importance of this discovery lies in the fact that this is the first incident of a successful merging of gravitation with quantum mechanics.

Because the temperature is so low, the amount of radiation escaping is very small. In the present environment the higher temperature of the background would actually favor influx of energy over out flow. But if the universe were to expand indefinitely, the background radiation temperature would eventually fall below 10^{-7}°K and then the emission from the black hole would exceed the influx. It would take of the order of 10^{66} years for a solar mass black hole to "evaporate".

16-2 Do Black Holes Actually Exist?

So far all we said about black holes was purely speculative. If a solar mass were collapsed down to its Schwarzschild radius then we would have a black hole. Well, does that ever happen? The answer, we believe is YES! This belief has its support from three lines of reasoning and observations:

(a) The physics of the evolution of massive stars

(b) Inferential evidence from close X-ray binaries

(c) Increasing evidence for massive dark objects in the center of some galaxies

The patterns of stellar evolution are now well enough understood that one can say under what general conditions the collapse of a large mass to a black hole would occur.

Stars like our sun are expected to evolve "normally". The present age of the sun is approximately 5 billion years. At age of approximately 11 Byrs the inner core will be depleted of fusible hydrogen. The star will adjust to the loss of pressure in the core by contraction in the core and expansion of its shell and become a red giant star. However the contraction is not large enough and is not able to raise the core temperature enough to lead to substantial fusion of helium. Eventually the star will cool and become a very dense white dwarf star which cools out slowly. In this state the star has reached an equilibrium in which the atoms are packed so tightly that the repulsion between the charged electrons becomes significant. This repulsion, known as electron gas degeneracy which is caused by the action of the Pauli principle, is sufficient to prevent further contraction. For stars which are smaller that the so called "Chandrasekhar limit" of $M_{ch} \approx 1.4M_{\odot}$ it is estimated that this mechanism provides stability of the white dwarf state.

If the collapsing star after fuel depletion in the core is larger than 1.4 $M_{\odot}$ then during the collapse phase part of the outer shell of the star is shed. For very massive stars this mass shedding takes the form of a gigantic super nova explosion.* If the collapsing core is larger than 1.4 $M_{\odot}$ but smaller than the Oppenheimer-Landau-Volkor limit of $M_{OLV} \approx 2.5$ $M_{\odot}$ then the contraction of the core becomes so large that even the nuclei of the material become so close that there is no more space for the atomic electron shells between the nuclei. The electrons become so to speak squeezed into the nuclei and tend to combine with the positively charged protons to form neutrons. That makes the core material essentially consisting of neutrons. The star has become a "neutron star". The more massive the remaining neutron star is the closer these neutrons are squeezed together. Neutrons also follow the Pauli principle, which does prohibit particles of spin 1/2 like the electrons and the neutrons from overlapping. This principle again provides a stable equilibrium state for the neutron star. However when the mass of the neutron star exceeds $2.5M_{\odot}$ the contracting gravitational force is so large that the "neutron degeneracy" is no longer able to provide an equilibrium. The star continues to collapse and no new mechanism is known that would provide another equilibrium. The core would collapse to a black hole singularity.

---

* Pictures of the expanding gas shell from the 1987 supernova in the Magellanic cloud were recently obtained with the Hubble space telescope.

Observation of X-ray sources has revealed one, perhaps two sources which are associated with a binary system of two stars which are revolving around each other in a tight orbit, and one of which is a "dark" companion star. From the observations one infers that the dark collapsed object must have a mass larger than $2.5M_\odot$. Therefore it is not expected to be a neutron star but a black hole.

There is increasing evidence that many galaxies may have very massive dark central objects. The Andromeda galaxy, two million light years away appears to have a central dark core of 30-70 million $M_\odot$. The galaxy M87, 35 million Ly from earth has a core of three billion $M_\odot$ in its center. The "Sombrero galaxy" has a rapidly rotating core of one billion $M_\odot$. Our own "Milky Way galaxy" appears to have a compact core of several million $M_\odot$ with features that indicate the presence of a large mass black hole. These observations and considerations make it reasonable to assume that, the ultimate state of matter, after all possibilities of stable stellar energy production by fusion have been exhausted, is the black hole state. Even the many stars which end up as burned out cooling white dwarfs are expected to coalesce eventually to massive black holes.

So, we are finding that gravitational clumping of matter is the cause for a continuing process of entropy increase from the first gathering of the early homogeneous gas provided in the big bang forming stars until the final death of stars and coalescence to black holes.

We also found that these final black hole singularities seem to be fundamentally different from the initial big bang singularity of the universe. The final black hole singularities are high entropy states while the initial big bang singularity must have been a low entropy state. What is the difference in the structure of these singularities? And how does the final cosmological singularity in the case of a recollapsing, k = +1, universe compare to the initial and the black hole type singularity?

On the basis of classical general relativity Penrose has proposed a solution to these questions in terms of characteristics of the geometry near the singularity described by a uniform average curvature and a distortional part caused by non-uniformity. He concludes that initial and final singularities are fundamentally different geometrically and therefore provide the arrow of time. The arrow of time will not reverse in the collapsing phase [PE 89].

Hawking has taken the opposite position based on his version of quantum cosmology. His "wave function of the universe" is time symmetric. In a recollapsing universe the arrow of time will reverse.

We will consider Penrose's argument first, in the next section and Hawking's in a later section.

17. STRUCTURE OF SPACE-TIME SINGULARITIES

17-1 THE SINGULARITY THEOREMS

The question what the meaning of the black hole singularities and the big bang cosmological singularity are, which appear in the corresponding solutions of general relativity led in the 60's and 70's to an intensive investigation. Generally, the appearance of singularities in the solutions of a physical theory that is supposed to be proper description of reality signals the break-down of theory. No quantitative predictions about the trajectory of evolution of the system beyond the singularity can be made. Usually, in physical systems, additional processes influence the behavior near the singularity in such a way that the singularity in fact is avoided. The classical driven harmonic oscillator predicts a catastrophic infinite amplitude when the driver frequency approaches the resonance frequency. This singularity is avoided because damping effects limit the amplitude. Ocean waves approaching a shallow beach steepen up to an ever increasing front which would in the limit require infinite acceleration of the water particles. The non-linearities and surface tension properties modify the behavior near the singularity leading to the breaking of the wave.

Einstein's equations of general relativity are very non-linear equations. The trajectories of systems that are governed by non-linear equations are very difficult to calculate. General solutions are not known and cannot be found by a general strategy as in linear theories. Only special solutions for often highly symmetric situations are more easily obtainable. In general relativity the big bang cosmological solutions and the black hole solutions are very special and relatively simple because of the assumed spherical symmetry. So the question arose whether the singularities are perhaps an artifact due to the choice of this high degree of symmetry and whether near the singularities in a non-symmetric situation other non-linearities would come into play that would avoid the singularities.

Newly forged mathematical tools of differential geometry were used by Penrose (1965), Hawking and Penrose (1969) and Geroch to prove a number of general theorems for the solutions of Einstein's equations.

The Hawking-Penrose theorem states that inside a Schwarzschild-horizon there must exist a singularity of space-time at which the path or world line of an observer who hits it must terminate and physics as we know it must break down. Another theorem by Hawking (1972) shows that the horizon area can never decrease. The existence of singularities inside a horizon protects the outside world from being affected by the presence of the singularity, because no signals from inside the horizon can emerge to the outside. But are all singularities hidden by such a horizon or are there singularities that are "naked"? We have seen in the previous section that the collapse of matter through the Schwarzschild horizon into a singularity appears to the outside observer after an infinite time. The black hole singularity is for us a future singularity. Einstein's equations are time symmetric like all the other dynamic theories we know. So the inverse of the collapse toward a future singularity is a possible solution to Einstein's equations. Picturing the collapse movie run backward we would see at $t = -\infty$ at the horizon stellar material emerging turning into a brilliant star. This event has been called sometimes a white hole. Since the light and the material can reach us at a later time, this singularity has been called "naked". The black hole solution is the analog of the retarded solution in electromagnetic theory where from an initial state the future singularity is developed. The white hole solution is like the advanced solution where the observed final state is the determinant for the past constellation. Penrose has proposed to rule out these "advanced solutions" as not relevant to the real universe. This hypothesis is called the "cosmic censorship hypothesis". There is no known proof for this hypothesis. The hypothesis plays a similar role as the assumption of coarse graining in statistical mechanics.

Ruling out past singularities limits the solutions of Einstein equations to these in positive time direction, with the possibility to connect the thermodynamic arrow of time with the cosmological arrow of time.

The cosmic censorship hypothesis also rules out the big bang original singularity as having the structure of a past singularity.

Without going into the mathematical details we want to outline at least conceptually how Penrose expresses the corresponding geometric condition.

17-2 RIEMANN, RICCI AND WEYL

The geometry of a general curved space-time as is used in general relativity is described in terms of the "metric" $\tilde{g}$ which specifies for all space points and all times what the distances are between neighboring points. For 3+1 space time the metric has ten independent pieces of information.

In the context of Einstein's theory which relates gravitation to curvature of space-time the most important geometric quality is one that combines all that can be said about the intrinsic curvature. This quantity is the "Riemann tensor". It has 20 independent pieces of information about the curvature at each space-time point.

There are two more geometric quantities which are associated with the curvature information and which can be extracted from

"Riemann" in a particularly simple way. One of these is the "Ricci tensor", the other is the curvature scaler R. "Ricci" has 10 independent pieces of information which are a particular subset of the 20 components of "Riemann". The scaler R is just one number. The significance of these two quantities is that, a certain combination of them is the only "conserved curvature feature" that exists. Therefore it was suggestive to connect this "Einstein tensor" $\tilde{\tilde{G}}(\tilde{\tilde{R}},R)$ with the quantity in physics that is known to be always conserved: Energy. This in fact is the basis of the famous Einstein equations of general relativity.

Because of this connection of "Ricci" and R to the energy distribution in Einstein's theory it is useful to group the 20 components of "Riemann" into those that are related to "Ricci" and R, which are then directly attributable to the mass-energy distribution and a remainder. This remainder is called the "Weyl tensor". In a stylized way this decomposition would be written as

"Riemann" = "Ricci" + "Weyl"

The physical meaning of these two parts becomes somewhat clear if we consider a gravity detector consisting of a spherical mass shell on which we measure the changes of shape when we bring it into a gravitational field [Fig 37]. The shell freely falling shell outside the gravitation generating mass M becomes distorted into an ellipsoid due to the tidal effect just like the earth oceans bulge due to the tidal effect of the moon. The amount of tidal distortion is a measure for the non-uniformity of the gravitational field produced near M. This non-uniformity is mathematically captured in "Weyl".

When the detector shell envelopes the mass M then gravity pulls the shell inward by some amount. The change in volume of the detector shell is directly related to the amount of mass-energy M. This volume change is expressed by "Ricci".

In a region of space where there is no mass distribution present locally "Ricci" will be zero but "Weyl" may show an effect due to masses at some distance. On the other hand, inside a homogeneous mass distribution "Ricci" is governed by the mass density and is the larger, the larger the mass density is. But "Weyl" will be zero because the gravitational effect is uniform throughout the mass distribution.

17-3 "RICCI" AND "WEYL" NEAR SINGULARITIES - PENROSE HYPOTHESIS

Now we can see that near a black hole singularity the fool hardy astronaut X carrying his gravity detector shell would clearly find "Ricci" = 0 near the horizon because all mass is collapsed to a singularity at the center. He will however detect the non-uniformity due to the fact that the mass is all concentrated in the central singularity. "Weyl" will in fact become very large as he approaches the singularity, which will

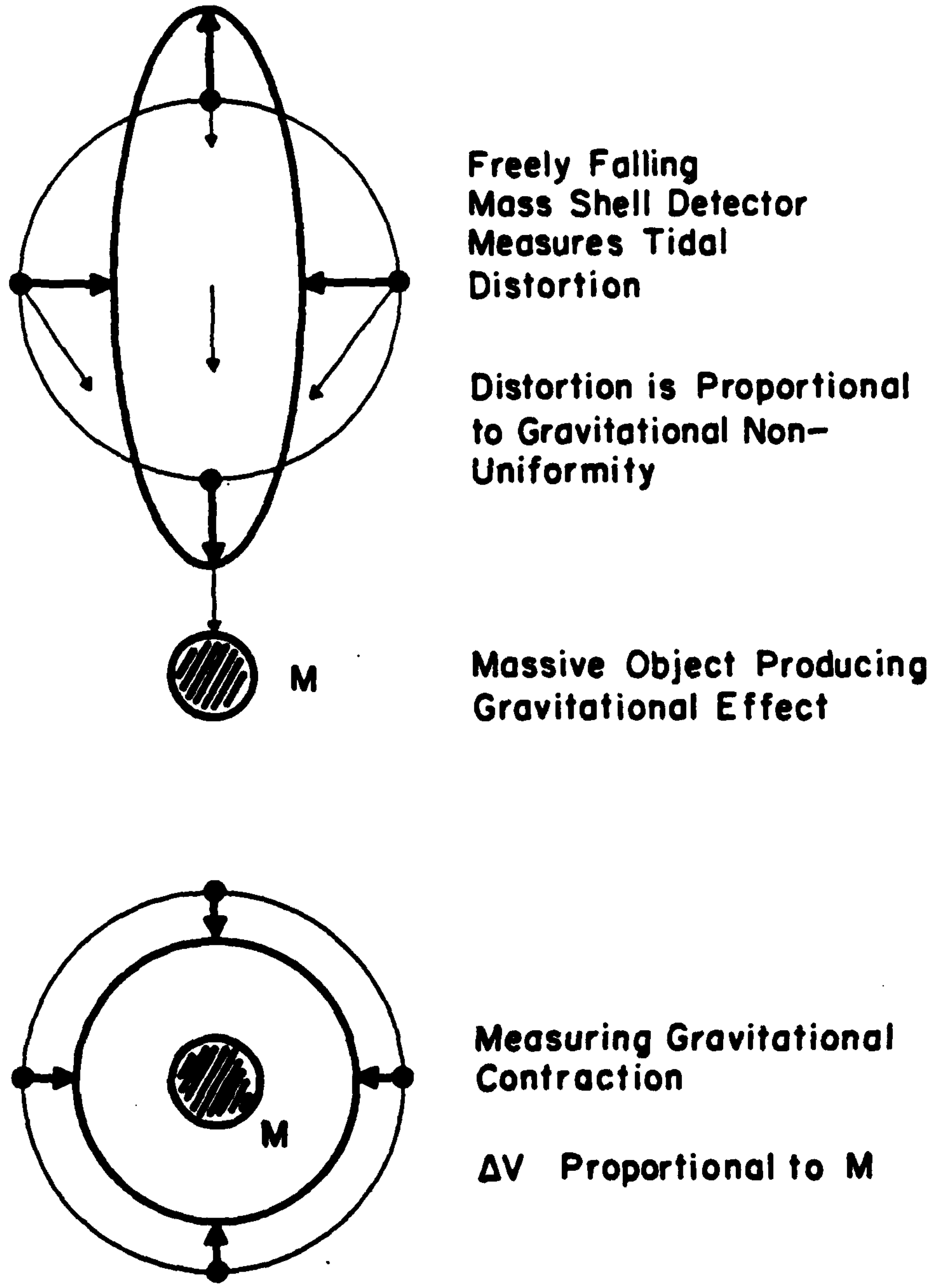


Fig. 37 MEASURING "RICCI" AND "WEYL"

affect his general health.

So one would conclude that the black hole singularities are characterized by

"Ricci" = 0

"Weyl" $\rightarrow \infty$

Entropy high

The initial big bang singularity, on the other hand would be characterized by the fact that near the singularity we have a highly uniform mass distribution at very high density and very high temperature. The gravitational curvature is very uniform. We have then:

"Ricci" $\propto \rho \rightarrow \infty$ near singularity

"Weyl" = 0

Entropy low

So, it would appear that Penrose's cosmic censorship hypothesis could be posed in the form: The cosmic initial singularity must have "Weyl" = 0. That excludes the possibility that the initial singularity is a "naked singularity", the time inverse of a black hole, emerging from a state of high entropy. It brings the thermodynamic arrow of time into conformity with the cosmological arrow of time.

There remains the question: What kind of singularity is the big crunch of the k = +1 spherical geometry universe? It would appear that as the density of radiation increases the universe returns to the state of homogeneous high density mass distribution with "Ricci" $\rightarrow \infty$ as the singularity is approached and "Weyl" = 0. That would indicate that the system has returned to a low entropy state and that for this future singularity we would have "Weyl" = 0 as for the past singularity.

Penrose argues that it is more likely that as density increases during recollapse large numbers of burned-out stars coalesce forming black holes, galaxies collapse to black holes. Rather than moving toward a state of uniform low entropy high density high temperature mass distribution, the system will move toward a very non-uniform distribution of very low temperature high entropy black hole singularities. In their neighborhood the space is characterized by "Ricci = 0 and "Weyl" $\rightarrow \infty$ as the singularity is approached.

Penrose's cosmic censorship hypothesis which prohibits past singularities or its equivalent the Weyl tensor hypothesis together with the assumption that in the case of a recollapsing universe that collapse will occur in non-uniform fashion giving rise to numerous black hole singularities which coalesce toward the big crunch, provides the time asymmetry for the universe. The initial low entropy state is not just due to the smallness of the initial space but due to "Weyl" = 0.

Hawking argued against Penrose's hypothesis [HA 85] saying that the hypothesis is rather ad hoc, amounting merely to putting in the arrow of time "by hand". It does not follow from anything more fundamental. Secondly, the whole line of argument is based on classical gravitation theory. It is generally agreed now, Hawkings argues that, some form of quantum gravity and quantum cosmology must be applied to these epochs of extreme gravitational fields.

Hawkings bases his view of time reversal and entropy decrease during recollapse on just such a quantum cosmology and we will consider this in a later section.

18. HOW IMPROBABLE WAS THE BIG BANG, OR HOW PRECISELY DID THE CREATOR PICK THE UNIVERSE?

In the chapter on the thermodynamic arrow of time we had shown in an example how enormously larger the phase space volume for the thermodynamic equilibrium state is than for any of the non-equilibrium states. The initial state of the universe with "Weyl" = 0 was such a non-equilibrium state. It is interesting to estimate by how much the entropy grows during the evolution of the universe.

We will consider the case of a closed universe near the critical case. The observed parameters are then:

$\rho_o = \quad = 1\text{x}10^{-30}$ grcm^{-3}	mass energy density
$T_o = 3°K$	radiation temperature
$R_o = 10^{29}$ cm	universe radius
$B = \frac{R_o^3 \rho_o}{m} = 6\times10^{60}$	number of heavy particles in the universe
$n_o = \frac{\rho_o}{m} = 6\times10^{-7}$	number of heavy particles per cm^3

The following natural constants will be needed:

$k_B = 1.38\text{x}10^{-16}$ erg K^{-1}	Boltzmann constant
$G = 6.6\text{x}10^{-8}$ dyn cm^2 g^{-2}	Newton's gravitation constant
$\hbar = 1.05\text{x}10^{-27}$ erg s	Planck quantum constant
$a = 7.5\text{x}10^{-15}$ erg cm^{-3} K^{-4}	Black body radiation constant

(1) Entropy of Black Hole

In a previous chapter we had given the expression for the entropy of a black hole. According to Beckenstein [BE 73] and Hawking [HA 75] it is

$$S_{BH} = \frac{A}{4}\left(\frac{k_B c^3}{G\hbar}\right)$$

Where A is the area of the Schwarzschild horizon

$$A = 4\pi R_s^2 = 4\pi\left(\frac{2GM}{c^2}\right)^2$$

This gives $$S_{BH} = 4\pi \frac{k_B G}{\hbar c} M^2$$

With all numbers inserted this gives

$$S_{BH} = 2\times10^{10} k_B M^2$$

(2) <u>Entropy of 3°K Background Radiation</u>

The entropy density of black body radiation of temperature T_o is given by

$$S_o = \frac{4}{3} a T_o^4 \quad \text{erg K}^{-1}\text{cm}^{-3}$$

The energy density is:

$$\rho_\gamma c^2 = a T_o^4 \ \text{erg cm}^{-3} = 2\times10^{-13} \text{erg cm}^{-3}$$

The number of photons per cm^3 is the energy density divided by the photon energy, which is approximately $k_B T_o$.

$$n_\gamma = \frac{aT_o^4}{k_B T_o} = \frac{a}{k_B} T_o^3 \approx 1500 \text{ photons cm}^{-3}$$

The number of photons per heavy particle in the universe is

$$\eta_\gamma = \frac{n_\gamma}{n_o} = \frac{1500}{6\times10^{-7}} \approx 2\times10^9 \text{ photons/particle}$$

The entropy per photon is

$$s_\gamma = \frac{\frac{4}{3} a T_o^3}{n_\gamma} = 1.7\times10^{-16} \text{ erg}K^{-1}\text{/photon} = 1.0 k_B\text{/photon}$$

The entropy of the entire black body radiation is then

$$S_{Rad} = \eta_\gamma \cdot B \cdot k_B = 2\times10^9 \times 6\times10^{80} \cdot k_B = 10^{90} k_B \text{ erg } K^{-1}$$

(3) Entropy of Solar Mass Black Hole

$S_{BH} = 2\text{x}10^{10} M^2 k_B$

Solar mass $M_\odot = 2\text{x}10^{33}$ gr

Number of heavy particles $\quad N_B(M_\odot) = \frac{M_\odot}{m} = \frac{2\times10^{33}}{1.6\times10^{-24}} = 10^{57}$

$S_{BH}(M_\odot) = 8\text{x}10^{76} k_B$

Black hole entropy per particle $\quad \tilde{S}_{Bh} = \frac{S_{BH}}{N_B} = 6\times10^{19} k_B\text{/particle}$

(4) Entropy of $10^6 M_\odot$ Black Hole in Galactic Center

$$S_{BH}(10^6 M_\odot) = 2\times10^{10}(10^6\cdot2\times10^{33})^2 k_B = 8\times10^{88} k_B$$
$$\tilde{S}_B(10^6 M_\odot) = \frac{8\times10^{88}\,k_B}{10^{11}\times1.0\times10^{57}} = 6\times10^{20} k_B/\text{particle in galaxy}$$

(5) Entropy of the Universe

This entropy depends on the kind of coalescence that takes place because the entropy depends on M^2

(a) all individual stars form black holes

$$S_u^{(1)} = \tilde{S}_{BH}\cdot6\times10^{80} = 3\times10^{100}\,k_B$$

(b) Galaxies form $10^6 M_\odot$ black hole cores

$$S_u^{(2)} = \tilde{S}_B(10^6 M_\odot)\,6\times10^{80} = 6\times10^{20}\cdot6\times10^{80} k_B = 4\times10^{100} k_B$$

(c) All galaxies collapse to galactic black holes

$$S_{Galax} = 2\times10^{10}(10^{11}\cdot2\times10^{33})^2 k_B = 8\times10^{98} k_B$$
$$\tilde{S}_{ga} = \frac{8\times10^{98}}{10^{11}\cdot10^{57}} = 8\times10^{30} k_B/\text{particle}$$
$$S_u^{(3)} = 6\times10^{80}\cdot8\times10^{30} k_B = 5\times10^{110} k_B$$

(d) Entire universal mass forms one black hole

$$M_{univ} = 6\times10^{80}\cdot1.6\times10^{-24} = 1\times10^{57}\,\text{gr}.$$
$$S_{univ}^{(4)} = 2\times10^{10}(1.0\times10^{57})^2 k_B = 2\times10^{124} k_B$$
$$\tilde{S}_{univ} = 3\times10^{43} k_B/\text{particle}$$

(6) Phase space volume of the largest entropy state

$$S_f - S_i = k_B \ln\frac{\Delta V_f}{\Delta V_i}$$

From this we find

$$\frac{\Delta V_f}{\Delta V_i} \approx e^{\frac{S_f - S_i}{k_B}} \approx e^{2\times 10^{124} - \frac{S_i}{k_B}}$$

The entropy of the initial state can be estimated

$$S_i = \frac{4}{3} a T_i^3 R_i^3$$

for T_i at the Planck time $T_i \approx 10^{32}$ $R_i \leq 10^{-3}$ cm

one obtains:

$$S_i \approx 10^{-15} \cdot 10^{96} \cdot 10^{-9} \approx 10^{72} \frac{erg}{{}^\circ K} = 10^{88} k_B$$

Then $$\frac{\Delta V_i}{\Delta V_f} \approx e^{10^{88} - 10^{123}} \approx e^{-10^{123}}$$

One might say: The creator aimed at a phase space volume of one part in $e^{10^{124}}$, when he set up the universe that allowed the second law. This is an incredibly small number. If one believes in the active participation of a creator who aims at initiating a universe that allows the unfolding of a multitude of structures and forms one might not be so surprised about such a number. The creator presumably knew what he was doing.

If one believes that the occurrence of our universe is a natural phenomenon that has a natural explanation, then one would seek to explain this occurrence as the accidental highly correlated state in a fundamental system of random processes that underlies the space-time structure. Such considerations have become possible in recent years within the frame work of quantum cosmology. We will discuss these developments in the following section.

19. TIME AS INTERNAL EVOLUTION PARAMETER IN EINSTEIN'S GENERAL THEORY OF RELATIVITY

19-1 The Road Toward Quantum Gravity

In the expression for the entropy of a Black hole for the first time in gravitation theory the Planck constant $\hbar$ appeared which signals the importance of quantum effects in the system described such.

The problem how to merge Einstein's theory of gravitation with the other pillar of modern physics, quantum theory, has been persued for over sixty years, but no satisfactory solution has been found. For the other three known forces in nature, the strong force, the electromagnetic force and the weak force formulations have been found that incorporate the quantum theoretical principles and these models have been extensively tested with high energy collisions in large accelerator facilities. The proper quantum theoretical descriptions of gravitation theory has eluded this unification, partly because gravitation theory as generally covariant theory of geometries is fundamentally different from the other theories of interactions. This difference makes unworkable the usual path of transiting from the unquantized classical theory to the proper quantum theory in a way that real observable results can be calculated from this theory. The other reason for the difficulty is the lack of experimental observations to study quantum gravity behavior. The gravitational force is fundamentally much weaker than the other three forces and is drowned out by these in all experiments that we can perform in the laboratory. Quantum gravity behavior makes itself only felt in very extreme laboratories, like the physics near the black hole singularity or very near the origin of the universe. The criterion for the point when quantum gravitation becomes the dominating phenomenon is that the gravitational energy of a system becomes comparable to the total system energy. The gravitational energy of a massive system confined to a size ΔL is given by $GM^2/\Delta L$. Only when the system of mass M is confined to small enough size ΔL, will its gravitational energy become significant. Quantum-mechanically a mass cannot be localized mere precisely than given by its Compton wave length $\lambda_c = \frac{\hbar c}{Mc^2} = \frac{\hbar c}{E}$. Thus one finds the ratio of gravitational energy to system energy for a system energy confined to its Compton wave length, when quantum effects become important as

$$\frac{W_{grav}}{E} \approx \left(\frac{L_{PL}}{\Delta L}\right)^2 \quad ,$$

where $L_{PL} = (G\hbar/c^3)^{1/2} = 1.6\times10^{-33}$ cm is a universal length that can be formed from the three fundamental constants; gravitational constant G, Planck constant $\hbar$, and light velocity

c. This so called Planck length becomes the fundamental quantum length of geometry where space quantum graininess makes its appearance. Just as $\hbar$ limits the accuracy with which one can hope to measure the action of a process, so L_{PL} limits the accuracy with which one can determine the nature of the geometry by local measurement. Beyond that, quantum fluctuation make results uncertain.

The above relation tells us that, when a system of mass is confined to size of the order of L_{PL}, then quantum gravity effects need to be considered.

In considering the inflationary scenario we gave as radius of the universe at the beginning of the inflationary phase $R = 10^{-27}$ cm. This is still six orders of magnitude larger than the Planck length. So one expects that quantum effects play a role only near the very origin at $t \approx 10^{-33}$ cm and temperatures corresponding to 10^{19} GeV. So, if one wishes to theorize about the very origin of the universe one needs a "Quantum cosmology".

In order to arrive at a quantum cosmology, one needs to subject the classical theory of gravitation, Einstein's theory of general relativity to the procedure that has proven in the rest of physics to give the proper transition from the classical theory to the corresponding quantum theory, for instance from classical mechanics of point particles to quantum mechanics of point particles, from classical electromagnetic field theory to quantum electrodynamics, etc.

It is this process that the peculiarly different role of time in Einstein's theory makes its appearance compared to Newtonian theory with the role of time as an absolute arena in it.

To motivate the results to be presented we will take a brief look at the so called "canonical quantization procedure" in quantum mechanics as a model for all other quantization procedures.

19-2 Hamilton Equations and Canonical Quantization in Mechanics

In classical mechanics the dynamics of a conservative motion of a point particle of mass m is based on the "Hamiltonian" for this system. The Hamiltonian is the total energy of the system consisting of kinetic energy and potential energy. The kinetic energy is expressed in terms of the particle momentum P (we consider a 1-dimensional system for simplicity):

$$K = \frac{P^2}{2m}$$

The potential energy is a function of position x only.

$$U = U(x)$$

The Hamiltonian of the system is

$$H(x,P) = \frac{P^2}{2m} + U(x)$$

The complete dynamics of the system is incorporated in the set of Hamilton canonical equations:

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} \quad ; \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x}$$

For our simple particle system these become

$$\frac{dx}{dt} = V = \frac{P}{m} \qquad \frac{dp}{dt} = -\frac{dU}{dx} = F(x)$$

We find the velocity in terms of the momentum and we find Newton's second law of dynamics.

The Lagrangian for the system is the kinetic energy minus the potential energy where the kinetic energy is expressed in terms of the velocity:

$$L(\dot{x}\ x) = \frac{1}{2}m\dot{x}^2 - U(x)$$

The momentum p is given by

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

and it is

$$H = p\dot{x} - L$$

These are the fundamental relations of the Hamilton-Lagrangian dynamics.

The formulation in terms of Hamiltonian and canonical Hamilton equations provides a "canonical" way of making the transition to quantum mechanics.

Quantum physics rests on the uncertainty principle which expresses the fact that microscopic particles behave such that position x and momentum p cannot be simultaneously determined with arbitrary accuracy. Both can only be determined to some uncertainty Δx and Δp such that

$$\Delta x \, \Delta p \geq \hbar$$

That renders inoperative the basic tenant of Newtonian physics that requires that only precise knowledge of the initial values of x and p allows to determine the later trajectory from the equation of motion. If the canonical equations of motion are to be retained for an electron, say, then x and p can no longer be treated as regular numerical quantities if the limitations imposed by the uncertainty principle are to be included. Instead x and p must be treated as higher order arrays of possible values to represent the quantum fuzziness. They become operators, that operate on a probability like function

$\Psi(x,t)$ that describes the probability to find the particle at position x at time t.

In this so called Schrödinger formulation of quantum mechanics the prescription for transitioning from the classical Hamiltonian to the quantum mechanics is:

$$H \rightarrow i\hbar\frac{\partial}{\partial t} \qquad p \rightarrow -i\hbar\frac{\partial}{\partial x} \qquad X \rightarrow x$$

such that the differential operators operate on the function $\Psi(x,t)$.

$$i\hbar\,\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\,\frac{\partial^2\Psi}{\partial x^2} + U(x)\,\Psi$$

This is the Schrödinger equation for a point particle of mass m. The justification for this process comes from for this process comes from the theoretical foundation of the ideas of quantum mechanics which has proven enormously successful in all areas of modern physics. All of atomic and molecular physics rests on this equation and its extension to more complicated systems. Quantized electromagnetic theory, the theory of the weak and strong interactions, all are built on the same fundamental principles.

It was then natural to attempt to build a quantized gravitation theory on the same principles. The first obstacle is that, Einstein's theory of gravitation has an appearance totally different from Hamiltonian mechanics. It is a theory that expresses gravitation in terms of space metric and curvature of space due to the presence of mass. What in this theory plays the role of position, momentum, potential energy, kinetic energy, Hamiltonian? Only in the sixties were these questions properly answered. They had to be answered before a quantization procedure could be made to work out, because no other road to quantization that was mathematically feasible was known.

In the next section we will show the result of the "Hamiltoniazation" of general relativity.

19-3 The Hamilton Form of General Relativity

General relativity is a theory about dynamics of space-time itself.

The distances between two space-time events are described as

$$ds^2 = g_{00}dt^2 + g_{11}(dx^1)^2 + g_{22}(dx^2)^2 + g_{33}(dx^3)^2$$

Where x^1 x^2 x^3 are three spatial coordinates. This is the Pythagoras law of geometry of distances. In its general form for general curved spaces this "metric" becomes

$$ds^2 = \sum_{\mu,\nu} g_{\mu\nu}dx^\mu dx^\nu \qquad \mu,\nu = 0,1,2,3$$

In general the components $g_{\mu\nu}$ of the metric tensor are general functions of space and time coordinates and form a 4x4 matrix which is symmetric

$(g_{\mu\nu} = g_{\nu\mu})$

Einstein's theory describes the evolution of this metric in terms of a set of ten differential equations which are given by the ten components of a curvature measure for which a conservation law holds.

$$R_{\mu\nu}[g_{\lambda\sigma}] - \frac{1}{2} g_{\mu\nu} R[g_{\lambda\sigma}] = 0$$

$R_{\mu\nu}$ is the "Ricci" tensor, R is the curvature scaler which can be derived from $R_{\mu\nu}$. The right hand side is zero when there is no mass present, otherwise an appropriate measure of the mass distribution will be on the right hand side.

$R_{\mu\nu}[g_{\lambda\sigma}]$ and R are functions of all metric components $g_{\lambda\sigma}$ and their first and second derivatives.

Clearly these ten equations do not have the form of Hamilton canonical equations. They are second order differential equations whereas Hamilton's canonical equations contain only first derivatives. Also, Hamilton's equations are solved for the time derivatives explicitly

$$\frac{d}{dt} \text{(canonical variable)} = \text{something.}$$

In Einstein's theory the time coordinate is treated on

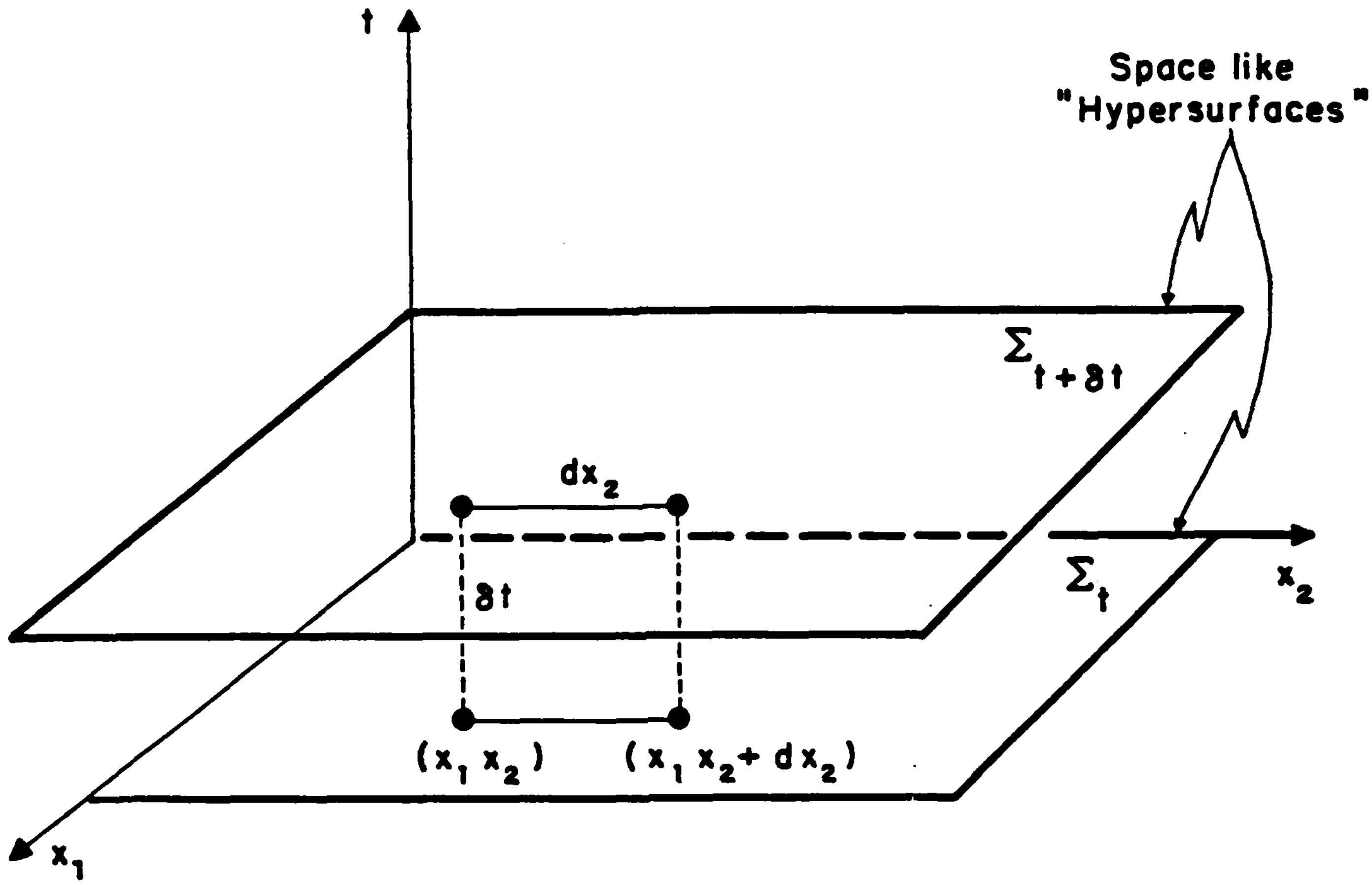


Fig. 38 FOLIATION OF EUCLIDEAN 2-DIM. SURFACES FORMING 3-DIM. EUCLIDEAN SPACE-TIME

completely equal footing with the other coordinates. The important feature of Einstein's theory is that, it is "generally covariant".

That means, one can subject the whole set of ten equations to a general coordinate *transformation*.

$$x^{\mu} \rightarrow x^{\mu'} = x^{\mu'}(x^0 x^1 x^2 x^3) \qquad \mu = 0,1,2,3$$

This leads to corresponding transformations of the metric and the curvature tensor

$$g_{\mu\nu} \rightarrow g'_{\mu\nu} \quad R_{\mu\nu} \rightarrow R'_{\mu\nu} \quad R \rightarrow R'.$$

After all this the Einstein equations for the new quantities in the new coordinates look exactly the same

$$R'_{\mu\nu}\left[g'_{\lambda\sigma}(x')\right] - \frac{1}{2} g'_{\mu\nu}(x')\, R' = 0$$

This is what is meant by general covariance.

In order to give the time coordinate a special position like in the canonical Hamilton dynamics Arnowitt, Deser and Misner [ADM 62] introduced the so called 3+1-formulation of Einstein's theory. They considered a foliation of space time into sandwiches between space like hypersurfaces Σ_t.

In a Euclidean 2+1 geometry which is a foliation of planes in a Euclidean space time the idea is straight forward [Fig 38].

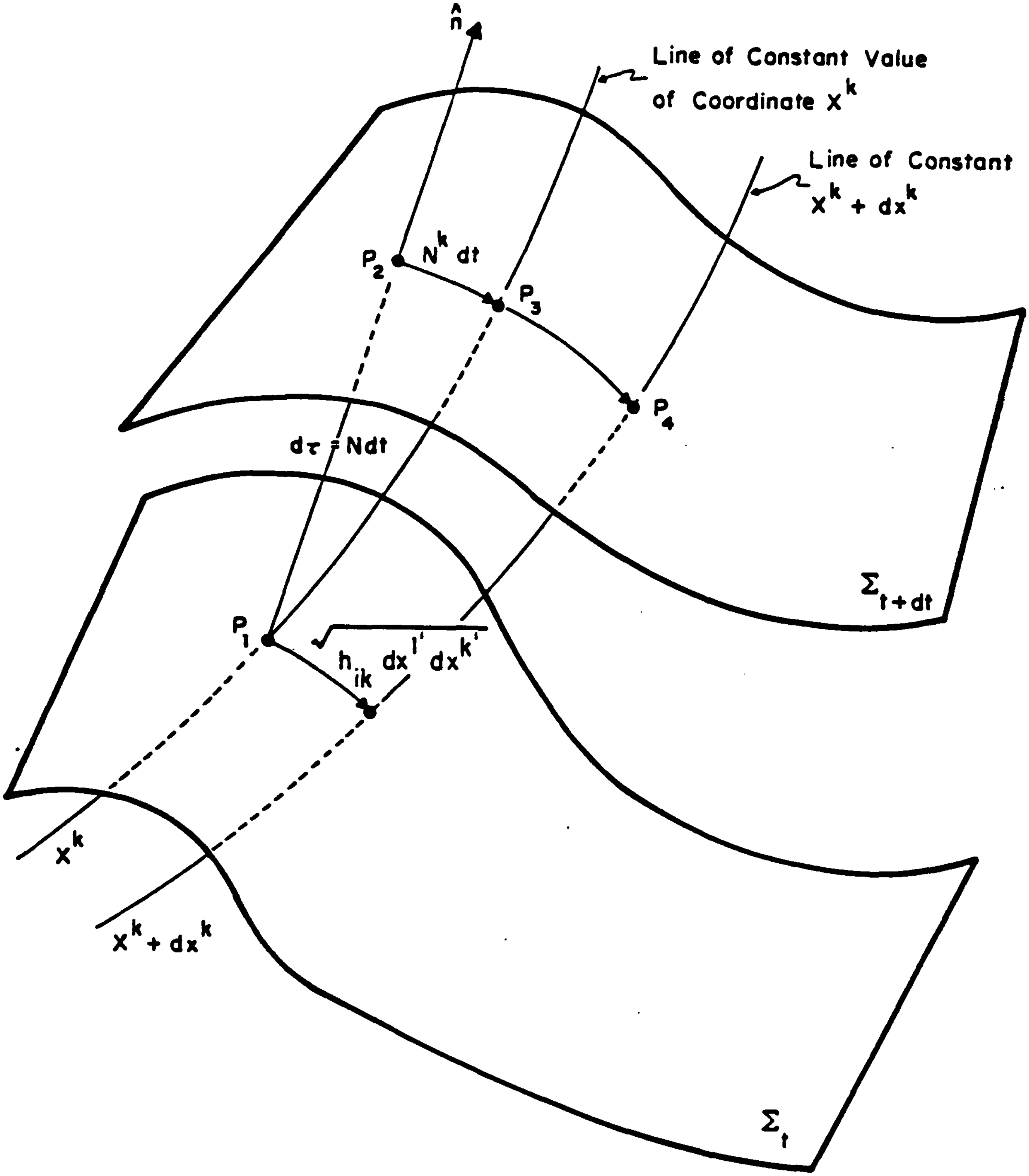


Fig. 39 FOLIATION OF CURVED 2-DIM.
SURFACES FORMING 3-DIM. GEOMETRY
"N = LAPSE FUNCTION N^K = SHIFT FUNCTION"

In a general curved 3+1 geometry the situation is more complicated. It becomes advantageous to decompose the full 4-dimensional metric $g_{\lambda\sigma}$ into a 3-dimensional metric h_{ik} $i,k = 1,2,3$ intrinsic to the space like hyperplanes Σ_t and the temporal part. The time axis can be taken to be the normal to the hyperplane. We see immediately that for a warped sandwich [Fig 39], the time axis may not intersect $\Sigma_{t+\delta t}$ at the same coordinate point as on Σ_t. A coordinate line x^k that connects points with same values for x^k on different hyperplanes may be shifted from the time line. also the proper time distances between the curved hypersurfaces Σ_t and $\Sigma_{t+\delta t}$ may be different at different coordinate points. The measure of time depends on the curvature.

Therefore, besides the intrinsic metric h_{ik} there are "lapse-and-shift-functions", N and N^k $k = 1,2,3$.

The metric decomposition is given by:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = (Ndt)^2 - h_{ik}(N^i dt + dx^i)(N^k dt + dx^K)$$

where $$g_{\mu\nu} = \begin{pmatrix} N^2 - N_i N_k h^{ik} & -N_k \\ -N_i & -h_{ik} \end{pmatrix}$$

With this decomposition it is possible to write the lagrangian density of the system

$$\mathcal{L} = \mathcal{L}\left[N; N_k; h_{ik}; \frac{\partial N_k}{\partial x^i}; \frac{\partial h_{ik}}{\partial x^i}\ \dot{h}_{ik}\right]$$

In this Lagrangian density the N_i N_k h_{ik} act like the coordinate variables in a mechanical system. The $\dot{h}_{ik}$ represent the velocities.

The momenta canonically corresponding to the coordinates according to the Hamiltonian prescription follow from the derivatives of the Lagrangian with respect to the velocities.

$$\frac{\partial \mathcal{L}}{\partial \dot{h}_{ik}} \equiv \pi^{ik} \qquad \frac{\partial \mathcal{L}}{\partial \dot{N}} = \pi \equiv 0 \qquad \frac{\partial \mathcal{L}}{\partial \dot{N}_k} = \pi^k = 0$$

The momenta π and π^k turn out to be zero because the Lagrangian does not contain the velocities $\dot{N}$ and $\dot{N}_k$. These are called <u>primary constraints</u> of the system.

The Lagrangian can be written into the form

$$\mathcal{L} = \pi^{ij}\dot{h}_{ij} - N\mathcal{H}^0 - N_i\mathcal{H}^i$$

Where $\mathcal{H}^0$ and $\mathcal{H}^i$ are known functions of the coordinates and momenta and their spatial derivatives. With this Lagrangian the dynamics of the system can be derived by variation of the action integral with respect to the "coordinates" and their "momenta"

$$\delta S = \delta\int \mathcal{L}\, d^4x = 0$$

The variation with respect to the "coordinates" h_{ik} and the "momenta" π^{ik} leads to the dynamic equations of general relativity in the Hamiltonian canonical form where the velocities and accelerations are given as functions of the coordinates and momenta and their spatial derivatives.

$$\dot{h}_{ij} = f\{h_{ij}\ pu^{ij}\ N_k\} \qquad \dot{\pi}^{ij} = P\{h_{ij}\ \pi^{ij}\ N\ N_k\}$$

The variation with respect to the shift and laps function N_k and N leads to the <u>secondary constraint</u> relations:

$$\mathcal{H}^0\{h_{ij}\ \pi^{ij}\ N\ N_k\} = 0 \qquad \mathcal{H}^k\{h_{ij}\ \pi^{ij}\ N\ N_k\} = 0$$

These relations are constraint relations on the h_{ij} and π_{ij} because they contain no time derivatives of these "coordinates" and "momenta". They are conditions that are preserved throughout the dynamic process.

Because the lapse function N and the shift functions N_k never appear differentiated with respect to time, their role in this system is not that of a dynamical variable. Instead they appear in the Lagrangian as factors to the constraints. This is the characteristic form of a variation principle with constraints imposed. Such problem is successfully treated by including the constraints into the variation integral with arbitrary factors, so called Lagrangian multipliers.

The Hamilton density according to the mechanical analog is given as

$$\mathcal{H} = \pi^{ij}\dot{h}_{ij} + \pi^i\dot{N}_i + \pi\dot{N} - \mathcal{L}[h_{ij}\cdots]$$

Inserting the above expression for the Lagrangian one obtains

$$\mathcal{H} = \dot{N}_i \pi^i + \dot{N}\pi + N\mathcal{H}^0 + N_i \mathcal{H}^i$$

Formally one find the canonical equations belonging to the "coordinate-momentum" pairs $(N\ ;\ \pi)$ and $(N_i\ ;\ \pi^i)$:

$$\dot{\pi} = -\frac{\partial H}{\partial N} = -\mathcal{H}^0 \qquad \dot{\pi}^i = -\frac{\partial H}{\partial N_i} = -\mathcal{H}^i$$

$$\dot{N} = -\frac{\partial H}{\partial \pi} = \dot{N} \qquad \dot{N}_i = \frac{\partial H}{\partial \pi^i} = \dot{N}_i$$

The lower equations simply tell that, the N and N_i are totally arbitrary. The canonical equations for them are trivially satisfied, no matter what is chosen for N and N_i. This freedom represents the arbitrariness of the four coordinates $(x^0\ ;\ x^k)$ which are used.

The upper equations by themselves do not imply the constraint relations $\mathcal{H}^0 = \mathcal{H}^i = 0$. Rather, these constraints must be imposed separately to insure consistency with the primary constraint relations which were that, $\dot{\pi} = \dot{\pi}^i = 0$. This is why they are called the secondary constraints. Dirac showed that the appearance of secondary constraints in a theory which must be imposed on the solutions of the Hamilton equations to assure consistency, is a common feature of all generally covariant theories.

The most surprising result is that, the primary and secondary constraints together render the Hamiltonian zero

$$\mathcal{H} = 0$$

If $\mathcal{H}$ is identified with the total energy density and if one assumes that the energy eigen states of a eventually properly quantized system evolve according to exp (iEt) in the Schrödinger picture, then one is led to the conclusion that <u>quantum ware functions in quantum gravity can have no time dependence</u>. <u>Time is at most an arbitrary parameter</u>. <u>The intrinsic geometry h_{ik} itself carries all the dynamically relevant information about time</u>.

19-4 <u>The Initial Value Problem in Hamiltonian Formulation</u>

The properly posed initial value problem is posed by giving the initial geometry ${}^{(3)}\hat{G}$. *This is best* done in the so called "Thin-sandwich formulation". For the given geometry ${}^{(3)}\hat{G}(o)$ one picks a definite but arbitrary set of coordinates $x^k = (x , y , z)$. In terms of these coordinates one finds a unique intrinsic metric h_{ij} (x , y , z) that describes ${}^{(3)}\hat{G}$ as a hypersurface Σ_0 in 4-dimensional space. Then one constructs coordinates $x^k + \dot{x}^k dt$, infinitesimally displaced in terms of a parameter t and finds the metric $h_{ij} + \dot{h}_{ij}dt$ which defines the geometry ${}^{(3)}\hat{G}(0) + {}^{(3)}\dot{\hat{G}}dt$ for the infinitesimally neighboring hypersurface Σ_{dt}. These two geometries are equivalent to giving ${}^{(3)}\hat{G}(o)$ and the rate of change of it on an initial hypersurface.

For this initial geometry the secondary constraints $\mathcal{H}^0 = o$ and $\mathcal{H}^i = o$ are relations to determine the lapse and shifts N and N_i. With this everything is uniquely determined on the initial hypersurface.

Now, to continue the integration one chooses new lapse and shift functions N(x y z) and N_i(x y z) and inserts them together with the initial h_{ij} and π^{ij} into the dynamic equations

$$\dot{h}_{ij} = f\{h_{ij}\ \pi^{ij}\ N_k\} \qquad \dot{\pi}^{ij} = p\{h_{ij}\ \pi^{ij}\ N\ N_k\}$$

and one determines the new geometry on a hypersurface that is shifted at each point by an amount N(x y z)dt. The initial $^{(3)}\hat{G}$ is embedded in a 4-dimensional geometry $^{(4)}\hat{G}$. It is a slice through $^{(4)}\hat{G}$. The choice of N(x , y , z) at each point pushes the integration forward in $^{(4)}\hat{G}$ by a chosen step size creating another slice with a completely determined new geometry $^{(3)}\hat{G}'$. The choice of N determines where the new slice $^{(3)}\hat{G}'$ occurs in $^{(4)}\hat{G}$. In other words a particular geometry $^{(3)}\hat{G}$ has a particular location in $^{(4)}\hat{G}$ relative to the initial geometry. The slice $^{(3)}\hat{G}$ carries within it information about its location in $^{(4)}\hat{G}$.

The ${}^{(3)}\hat{G}$ carries in it information about "time".

We see now that, there is no room in this theory for a time coordinate in the usual sense which continues to increase somehow autonomously forming one direction of ${}^{(4)}\hat{G}$. What forms the ${}^{(4)}\hat{G}$ are slices of ${}^{(3)}\hat{G}$. How closely these are packed in ${}^{(4)}\hat{G}$ is a geometric matter inherent to the ${}^{(3)}\hat{G}$. The parameter t is an arbitrary parameter that merely labels hypersurfaces of ${}^{(3)}\hat{G}$ embedded in the full ${}^{(4)}\hat{G}$ [Fig 40].

The Hamiltonian is zero. We expect that in the quantum version of the theory this means that there is no time dependence in the wave functions Ψ and the energy eigen states should be zero.

In the next section we will take a look at some aspects of quantum cosmology.

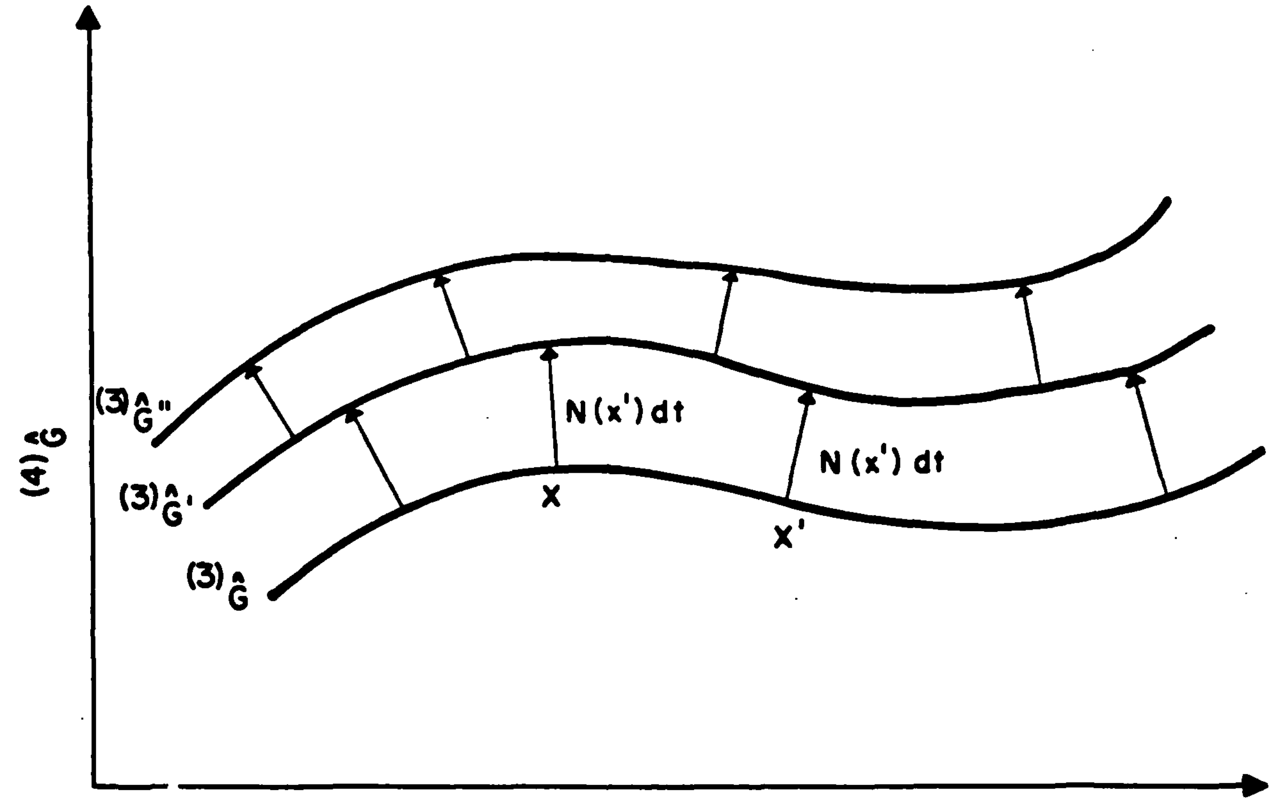


Fig. 40 STACK OF $^{(3)}\hat{G}$ FORMING $^{(4)}\hat{G}$
The choice of the lapse function N(x) determines a new slice $^{(3)}\hat{G}$ and where it is positioned in $^{(4)}\hat{G}$.

20. TIME IN QUANTUM COSMOLOGY

20-1 The Transition to Quantum Cosmology: Wheeler-DeWitt Equation

The Hamiltonian formulation of dynamics is the form in which the transition to the corresponding quantum-dynamics is easily made by a simple prescription: Convert all momenta in the Hamiltonian into differential operators with respect to the corresponding coordinate variables which act on a "wave function" $\Psi[h_{ij}]$ which is a function of the coordinate variables.

The "coordinate variables" in the gravitational Hamiltonain are the h_{ij}. The corresponding momenta are the π^{ij}.

Because of the secondary constraints it is sufficient to use the constraints $\mathcal{H}^0 = \mathcal{H}^i = 0$ instead of the Hamiltonian, where $\mathcal{H}^0 = \mathcal{H}^0[h_{ij}\ \pi^{ij}]$.

For the transition to the equivalent of the Schrödinger equation in atomic physics one makes the replacement

$$\pi^{ij} \rightarrow -i\hbar\left(\frac{c^3}{16\pi G\hbar}\right)^{3/2} \frac{\delta}{\delta h_{ij}}$$

The so called WHEELER-DeWITT equation then becomes

$$\mathcal{H}^0\left[h_{ij}\ ;\ \frac{\delta}{\delta h_{ij}}\right]\Psi[h_{ij}] = 0$$

The interpretation of this only known quantum gravity model has been difficult.

It is a second order differential equation for the "coordinate variable" $h_{ik}(x\ y\ z)$. These six variables are not really coordinates. They are functions of the coordinates (x y z) of the space they are describing. They are the metric of a 3-dimensional geometry $^{(3)}\hat{G}$. To describe the metric completely one needs the six h_{ik} at every space point (x , y , z). So one needs six times infinite many peices of information to describe one geometry $^{(3)}\hat{G}$ completely. So this Wheeler-De Witt equation describes events in a six-times-infinite-dimensional "phase space". Each geometry $^{(3)}\hat{G}$ is represented by one point in this six times infinite dimensional space.

In the spirit of quantum dynamics one takes into account the uncertainty principle by not considering one single system but many equivalent systems, represented by a swarm of points in this "phase space". The "Schrödinger wave function" $\Psi[h_{ij}]$ is expected to represent some sort of probability distribution for this swarm of points in this six-times-infinite-dimensional SUPER SPACE. If the distribution has a fairly defined maximum which traces out a fairly well defined trajectory in superspace then this trajectory is the equivalent of the expected classical evolution smeared out somewhat by the quantum fuzziness due to the uncertainty principle which enters through the presence of the Planck quantum constant $\hbar$.

Nobody knows how to deal with differential equations in infinite-dimensional space. Besides there are difficulties with the probability interpretation.

But what is most surprising is that, the Wheeler-DeWitt equation of quantum cosmology <u>does not contain time derivatives of anything</u>. There does not appear an absolute time or clock external to the universe that measures evolution of the universe. What measure there might be, it must be contained in the geometries of the universe.

The difficulties in dealing with the full Wheeler-DeWitt equation in superspace led to extensive investigation of the quantum cosmology of homogenious isotropic metrics like the Friedmann-Robinson-Walker metric of the simple big bang models in which the only metric component that is dynamic is the scale factor R. For more realistic models matter fields ϕ may be included which are also assumed to be homogeneous. The assumption of homogenity removes the infinite degrees of freedom of variations with respect to the coordinates (x y z). The superspace reduces to a "Minisuperspace" of very few dimensions, R plus however many matter fields one considers.

20-2 The Boundary Condition Problem

Still remaining is another difficulty associated with the Wheeler-DeWitt equation. The equation has infinite many solutions $\Psi(R , \phi)$. As always with differential equations a single solution is selected from that infinity of solutions by posing suitable boundary conditions or initial conditions. But what are we to assume as the appropriate boundary conditions for the wave function of the universe?

Two approaches have been eplored most extensively. The first is the Hartle-Hawking boundary condition, [HH 83] which led to their famous paper on the wave function of the universe. The second is by Vilenkin [Vi 86].

Hartle and Hawking confine their considerations to closed universe geometries. The "ground state" of quantum cosmology is the state of lowest excitation. It is that solution of the Wheeler-DeWitt equation which is at the beginning of the expansion.

They employ the so called "path integral" technique to quantization. A dynamical state of system in classical dynamics is reached by a unique path, that of least action. In quantum systems other pathways are possible in the vicinity of the least action path due to the uncertainty principle. Excursions from the least action path by action differences of the order $\hbar$ occur with finite probability. The probability is smaller for larger excursions. In order to assess the probability to find the system in a given state one averages over all pathways which lead to the particular state, each weighted by the probability for this path. The probability for the state, i.e. the Schrödinger function for the state is expressed in terms of a sum of possible histories leading to this state.

In our six-times-infinite dimensional phase space of states of 3-dimensional geometries ${}^{(3)}\hat{G}$, where each geometry is one point there are infinite many histories that could lead to one particular state, i.e. one point in superspace. Hartle and Hawking then need to specify a method by which to confine the set of histories in a reasonable way such that the ground state of the system is defined. To specify the ground state wave function $\Psi_o[h_{ij}]$ they propose to sum over all "compact 4 geometries" leading to a given 3-dimensional h_{ij}. This means that the universe allowed in competition do not have any boundaries in space or time This eliminates the need for posing any boundary conditions. One can interpret the resulting function $\Psi_o[h_{ij}]$ as giving the probability amplitude for the 3-geometry h_{ij} to arise from a zero 3-geometry, i.e. a single point. In other words the ground state is the amplitude for the universe to appear from nothing.

The ground state wave function $\Psi_o[h_{ij}]$ of the universe so constructed is symmetric with respect to application to the original big bang singularity or the final big crunch of the closed universe. There is no measure of time in the wave function and no directionality.

In a famous paper "Arrow of Time in Cosmology" [HA 85] Hawking shows that the proposed wave function of the universe is in a sense CPT invariant.

Furthermore Hawking showed that, the wave function, when represented as a harmonic superposition of perturbations from homogeneity, these perturbations would grow during the expansion phase, making the universe more irregular, thus increasing entropy.

Thus in the universe described by the Hartle Hawking wave function the thermodynamic arrow of time, defined by the direction of increasing entropy is related to the cosmological arrow of time defined by the direction of time in which the universe expands. Hawking argues that because of the proven CPT invariance of the wave function, the thermodynamic arrow would reverse during a contracting phase of the universe or inside the event horizon of black holes. The wave function whose harmonic content had grown during expansion would become smoother again during recollapse, the entropy would decrease.

In a companion paper in the same issue of Physical Review D.N Page [PA 85] questions Hawkings conclusion. He points out that while the superposition of harmonics that make up the wave function may be CPT invariant as Hawking showed but not all component waves may be CPT invariant individually. A CPT transformed component Ψ_n, CPT•Ψ_n is not necessarily identical to Ψ_n, but the sum Ψ_n + (CPT)•Ψ_n is CPT invariant still. So the universe could have component wave functions which give monotomically increasing irregularities in the universe, throughout the expansion and contraction phase. In that case the thermodynamic arrow of time would not reverse at recontraction and the cosmological arrow would be opposite to the thermodynamic arrow during contraction.

In a little footnote to his paper Hawking conceeds that, "Page may well be right in his suggestion".

In the other approach proposed by Vilenkin [Vi 86] the boundary condition is based on the picture that the universe spontaneously nucleates in a de Sitter space and then evolves along the lines of an inflationary scenario. The mathematical description of this cosmic nucleation is analogous to that of quantum tunneling through a potential barrier. It is referred to as "quantum tunneling from nothing". The boundary condition selects only "outgoing waves" from the set of solutions of the Wheeler- DeWitt equation. Vilenkin shows that the tunneling wave functions predict initial states that lead to inflation while the Hartle Hawking wave function does not.

There is no discussion of the connection of cosmological arrow of time and thermodynamic arrow of time in Vilenkin's paper. But it seems that the chosen boundary which selects only outgoing waves already provides an evolutionary assymmetry in that it gives inflationary scenarios. The "tunneling" boundary condition can be thought of as a quantum version of the Penrose condition (Weyl condition) in classical general relativity.

20-3 The Concept of Time in Quantum Cosmology

The wave functions of the universe discussed in the frame work of minisuperspace where Ψ depends only on the universe scale variable R and on one or more matter fields ϕ are completely independent of time. It seems to give a completely static picture of the universe.

Already in his seminal paper DeWitt [DW 67] pointed out that the time coordinate is an arbitrary label with which to label 3-dimensional slices through the 4-dimensional space of general reliability. A physically meaningful time could be defined internally by using for instance the geometrical or matter variables. One could use R or ϕ itself as internal measure of time. Because it is only an interal measure it is not necessary that the same measure provide a global time scale for the entire trajectory in superspace. In order for a variable to serve as good time variable it should satisfy these conditions: (i) it should be semiclassical following the WKB approximation, for instance, so that the connection to classical time can be made. (ii) it should be monotonic.

The field ϕ is a good time variable in some models and not in others. One can however introduce a separate "clock variable" ζ into the theory which enters into the Wheeler-DeWitt equation as a potential function $V(\zeta)$ such that the evolution of ζ is monotonic and the potential constitutes neglegible influence on the cosmological dynamics.

In a more recent paper M. Castagnino [CA 89] introduces a "probabilistic time". He considers wave functions $\Psi(R , \phi)$ in minisuperspace where R is the universe scale variable and ϕ the matter field variable. The "phase point" representing the state of the universe is moving about in this minisuperspace. Minisuperspace itself is endowed with a metric $G_{R\phi}$ which has a determinant G(R). The quantity $|\Psi(R,\phi)|^2$ can be interpreted as the probability density to find *the metric and the field ϕ in the volume element* $\sqrt{-G(R)}\, dRd\phi$ containing the "phase point" (R,ϕ). The probability of finding the state (R,ϕ) is then

$$d^{(2)}p = |\Psi(R,\phi)|^2 \sqrt{-G(R)}\, dRd\phi$$

The probability of finding the metric in the interval R, R+dR regardless of the field ϕ is

$$dp = dR \int |\Psi(R,\phi)|^2 \sqrt{-G(R)}\, d\phi$$

where the integration is over all possible field configurations. The probability dp symbolizes a quantity proportional to the number of possible metrics in the intervall R, R+dR.

One can now imagine that the universe stays in a metric given by R for a duration proportional to dp. That means one can give to each metric a time unit of occupancy. The metric of the universe must lay between R and R+dR a period proportional to dp x time unit.

The element of probabilistic time is then

$$d\theta = c\,dR\int|\Psi(R,\phi)|^2\sqrt{-G(R)}\,d\phi$$

The probabilistic time of the universe when it has metric R is:

$$\theta = c\int_o^R dR\int|\Psi(R,\phi)|^2\sqrt{-G(R)}\,d\phi$$

where R = o corresponds to the big bang moment. The constant c is chosen such that in the transition from semiclassical time θ to classical time these two time scales are equal.

Castaguino's construction of probabilistic time coincides with Vilenkin's in semiclassical approximation.

The intriguing fact is that in this definition of time which is a purely internal measure in the universe, the entire spectrum of possibilities of the matter field serves as the clock mechanism that pushes the clock hand through some interval when the universe expands by dR. That clock intervall incorporates not only the actual matter constellation but all quantum dynamic alternatives each weighted by their probability of occurance.

When the universe enters from the quantum epoche to the classical behavior the trajectory becomes a definite one and the probabilistic time sharpens to the classical time.

21. **SUMMARY**

In spite of all the avenues of inquiry around the phenomenon of time that have been explored and some of which have been described in this report, it appears that this phenomenon retains its mysterious quality in our lives.

The knowledge of the elaborate theoretical construct does not seem to change our basic internal experience of the flowing of our time. This experience, though somewhat modulated, depending on our psycho-physical state, seems to be a fundamental constant. Its mechanism is still not well understood.

But when we change our focus and search for the manifestations of the time phenomenon in the objective external world we do seem to be led to a rather definite conclusion.

Searching through our attempts to measure the lapse rate of time we discover that we are measuring time very dissimilar to measuring geometric qualities. In the latter case we can actually hold the yard stick next to an external object and compare its length with that of our yard stick. Measuring time we are not making such comparison of our clock with some absolute time. We are watching our clock intervals and compare it with other clock like mechanisms.

In the special theory of relativity time intervals appear relative to the state of motion of the observer with respect to the clock. The relativity of time is our first indication that time appears intrinsic to the material world and its processes rather than absolute. The material universe appears to present us with a number of arrows of time.

We saw in the slight variation of the CP invariance in basic interactions a time assymetry in the forces of nature. The origin of this assymetry must reside in those processes that establish the fundamental patterns of matter and forces at the very birth of the universe. It must be associated with the fundamental constants of nature in our universe. There are presently only very few rudimentary attempts at "fundamental theory" of theories that aim at explaining the appearance of the fundamental constants.

The thermodynamic arrow of time can be seen as a fundamental pattern in a universe of distinct individual entities in motion. The states of assemblies of distinct objects show the general characterization of order and disorder. We saw that there is in our universe a large scale transformation from ordered states to disordered states due to the fact that the universe was apparently created in a highly ordered state. This continuing process of entropy increases in the universe for the observer becomes a metaphor for the flowing of time. The universe was born in a very low entropy state. In addition gravitation acts to produce low entropy matter lumps from distributed matter.

The initial low entropy state of the universe and the cosmological expansion seem to be connected in the origin of the universe. The thermodynamic and the cosmological arrow of time appear to be clearly connected for the expansion phase of the universe. How these two are related during recollapse, if that occurs, is controversial.

The clearest indication that time is a phenomenon that is a purely intrinsic feature of the evolving universal geometry comes from general relativity in its classical Einsteinian form as well as in the rudimentary models of a quantum cosmology. The Hamiltonian version of general relativity is independent of time. The analysis of the initial data structure of solutions of the theory shows that time must be information that is carried by the 3-dimensional geometries themselves intrinsically. The canonical quantum theoretical version of theory, embodied in the Wheeler-DeWitt equation shows the same feature. The time parameter is absent from the theory. It must be especially introduced by defining clock like mechanisms associated with the evolving 3-geometries, that measure time intervals.

The probabilistic time is such a measure. It is special because in it the universe in all its possibilities is its own clock. The evolution of a matter-space universe defines time.

TIME IS THE HORIZON OF BEING.

APPENDIX A: FORCED HARMONIC OSCILLATOR

Newton's law of motion for the forced harmonic oscillator is given as

$$m\ddot{x} + m\gamma\dot{x} + \kappa x = F_o \sin \omega t$$

where x as a function of time t represents the instantaneous displacement of the oscillator from equilibrium position, ($\dot{}$) refers to time derivative, γ is the damping constant, κ the restoring force constant, ω the frequency of the driving force, F_o its amplitude. In the case of a rotational oscillation like the balance wheel in wrist watches. X is the angular displacement, m its movement of inertia and F_o the torque of the driver. The steady state solution of this equation of motion is

$$x(t) = A \sin(\omega t + \Phi)$$

The amplitude A of the oscillation is frequency dependent and given by

$$A = \frac{F_o/m}{((\omega_o^2-\omega^2)^2 + \gamma^2\omega^2)^{1/2}}$$

where $\omega_o^2 \equiv \frac{\kappa}{m}$.

At a driver frequency $\omega = \omega_o$ the amplitude is maximal and given by

$$A_m \equiv \frac{F_o}{m\gamma\omega_o}$$

Written in terms of A_m and expressing ω in terms of its separation from ω_o

$$\omega = \omega_o + \delta\omega = \omega_o\left(1 + \frac{\delta\omega}{\omega_o}\right) \equiv \omega_o(1+\alpha)$$

one has
$$A = \frac{Am}{\left(\frac{\omega^2_o}{\gamma^2}(1-(1+\alpha)^2)^2 + (1+\alpha)^2\right)^{1/2}}$$

The factor $\frac{\omega_o}{\gamma}$ is called the quality factor Q.

If one expresses the denominator in terms of a normalized quantity $\Omega \equiv Q\alpha$ one has

$$A = A_m\left(1+4\Omega^2+2\frac{\Omega}{Q}+\frac{\Omega^2}{Q^2}+\frac{4\Omega^3}{Q}+\frac{\Omega^4}{Q^2}\right)^{-1/2}$$

For large Q all but the first two terms can be neglected and

$$A \approx A_m\,(1+4\Omega^2)^{-1/2}$$

For Q > 10 the Q dependence of A is very little and the above approximation of a "universal resonance curve" which is independent of Q is very good. The maximum amplitude occurs at $\Omega = 0$ and is A_m. The width of the resonance curve can be given as that value of Ω for which the amplitude square is down by a factor two

$$A^2 = \frac{A_m^2}{2} = \frac{A_m^2}{1+4\Omega_{1/2}^2}$$

i.e. $$\Omega_{1/2} = \pm \frac{1}{2}$$

This corresponds to

$$\frac{\delta\omega_{1/2}}{\omega_o} = \pm \frac{1}{2Q}$$

For a sharp resonance curve a large Q, i.e. $\gamma << \omega_o$ is required.

APPENDIX B SPECIAL THEORY OF RELATIVITY

B-1 The Lorentz Transformation

Reference frame O′ moves with velocity v in x-direction in reference frame O [Fig B-1].

At time t=0 the two frames coincide and O′ sets his clock to t′ = 0. At that time t=t′=0. O′ flashes a light. The electromagnetic signals expands with vacuum light velocity c in all directions. The position of the light front is observed by both observers O and O′.

For O′ the wave has spread after a time t′ on his clock to the surface of the sphere given by

$$(x')^2 + (y')^2 + (z')^2 = c^2(t')^2 \tag{1}$$

Observer O observes the wavefront to be the surface of the sphere

$$x^2 + y^2 + z^2 = c^2t^2 \tag{2}$$

Guided by the Galilei transformation in classical Newtonian physics one proposes the following transformation law

$$\begin{aligned} x^1 &= \gamma(x-vt) \\ y^1 &= y \\ z^1 &= z \\ t^1 &= \gamma(t+\delta) \end{aligned} \tag{3}$$

$$\lim_{v<c} \gamma = 1 \quad \text{and} \quad \lim_{v<c} \delta = 0$$

We expect that

in order to recover the classical Galilei transformation at low velocities.

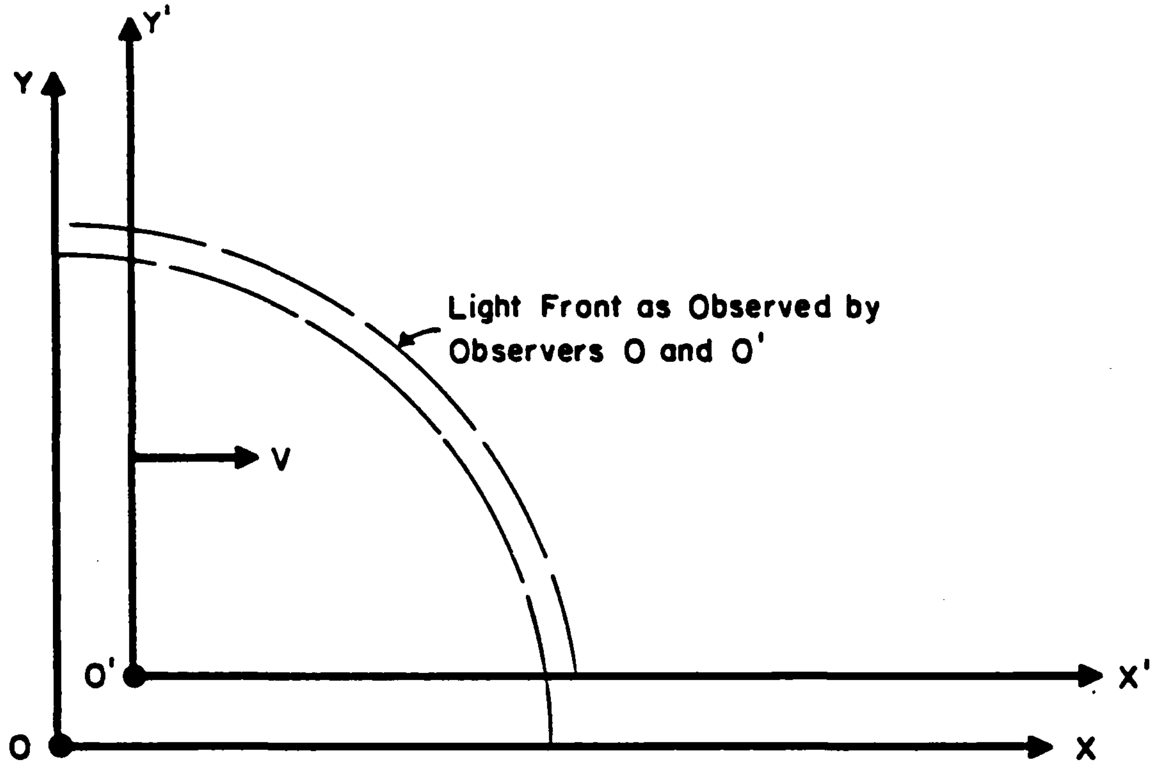


Fig. B-1 LIGHT PROPAGATION AND LORENTZ TRANSFORMATION

$$x^2 + y^2 + z^2 = c^2t^2$$

$$(x')^2 + (y')^2 + (z')^2 = c^2(t')^2$$

Inserting the rules (3) into (1) one obtains

$$\gamma^2(x^2-2vxt+v^2t^2)+y^2+z^2 = c^2\gamma^2(t^2+2\delta t+\delta^2)$$

If this is to be identical to equation (2) we can subtract (2). The remaining expression must be zero for arbitrary x, and t.

$$(\gamma^2-1)x^2-2vxt\gamma^2 = (c^2\gamma^2-v^2\gamma^2-c^2)t^2+2\delta c^2\gamma^2 t+c^2\gamma^2\delta^2$$

To get rid of the mixed term for all x and t it must be

$$-2vx = 2\delta c^2$$

That is $\delta = -\frac{v}{c^2}x$

For the remainder we have

$$\gamma^2-1-\frac{v^2}{c^2}\gamma^2)x^2 = (\gamma^2-1-\frac{v^2}{c^2})c^2t^2$$

So, it must be

$$\gamma^2-1-\frac{v^2}{c^2}\gamma^2 = 0 \qquad \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

The Lorentz Transformation laws are then:

$$x' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(x - vt)$$
$$y' = y$$
$$z' = z$$
$$t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t-\frac{v}{c^2}x\right)$$

B-2 Time Dilation and Lorentz Contraction

These two effects can be derived from the transformations by applying the transformations to space and time intervals

$$\Delta x' = \gamma(\Delta x - v\delta t)$$
$$\Delta t' = \gamma\left(\Delta t-\frac{v}{c^2}\Delta x\right)$$

(a) For a clock at rest in O′ it is $\Delta x' = o$ O

sees the intervals Δt at points separated by Δx such that

$$\Delta x = v\Delta t$$
$$\Delta t' = \gamma\left(\Delta t-\frac{v^2}{c^2}\Delta t\right)$$
$$\Delta t = \frac{\Delta t'}{\sqrt{1-\frac{v^2}{c^2}}}$$

Time appears dilated

(b) A yardstick in O′ is measured by coincidence comparison with a yardstick Δx in O. the coincidence in O requires $\Delta t = 0$.

Therefore

$$\Delta x' = \gamma \Delta x$$

or
$$\Delta x = \Delta x' \sqrt{1-\frac{v^2}{c^2}}$$

Length appears contracted.

B-3 Relativistic Velocity Transformation

$$u'_x = \frac{dx'}{dt'} \quad u'_y = \frac{dy'}{dt'} \quad u'_z = \frac{dz'}{dt'}$$
$$dx' = \gamma(dx-vdt)$$
$$dy' = dy$$
$$dz' = dz$$
$$dt' = \gamma(dt-\frac{v}{c^2}dx)$$

$$u'_x = \frac{dx'}{dt'} = \frac{\gamma(dx-vdt)}{\gamma(dt-\frac{v}{c^2}dx)} = \frac{\frac{dx}{dt}-v}{1-\frac{v}{c^2}\frac{dx}{dt}} = \frac{u_x-v}{1-\frac{vu_x}{c^2}}$$
$$u'_y = \frac{dy}{\gamma(dt-\frac{v}{c^2}dx)} = \frac{1}{\gamma}\frac{u_y}{1-\frac{vu_x}{c^2}}$$
$$u'_z = \frac{1}{\gamma}\frac{u_z}{1-\frac{vu_x}{c^2}}$$

B-4 Relativistic Mass

Collision of two particles of equal rest mass in laboratory frame x-y with $\theta_{1i} = \theta_{2i}$. Because of conservation of momentum for elastic collisions it is $\theta_{1f} = \theta_{2f}$.

The impact is chosen to be such that $\theta_{1f} = \theta_{1i}$ [Fig B-2].

Two observers O_1 and O_2 are moving with large velocities with respect to each other toward each other. Each sees the other moving with velocity v.

Each of the observers tosses a billiard ball with mass m_o in his rest frame along his y direction in such a way that they collide with a velocity u which is small compared to c.

To observer O_1 the collision event looks like in the figure B-3. Ball B_1 in his frame moves in $+y_1$-direction, with speed u, elastically collides with B_2 and returns with speed -u. Ball B_2 in observer O_1's frame retains a large horizontal velocity v. The y-component of B_2 in O_2's frame would be u. Transformed into O_1-frame using the velocity transformation rules gives B_2 a velocity

$$u\sqrt{1-\frac{v^2}{c^2}}$$

in y-direction which is very small. The <u>conservation</u> of <u>y-momentum</u> from observer O_1's point of view will be

$$m(u)u - m(v)u\sqrt{1-\frac{v^2}{c^2}} = -m(u)u + m(v)u\sqrt{1-\frac{v^{2\,1}}{c^2}}$$

$$\text{or} \qquad m(u)u - m(v)u\sqrt{1-\frac{v^{2\,1}}{c^2}} = 0$$

It is clear that this equation could not be satisfied if one inserted for the rest mass m_o in both cases.

The law of conservation of momentum and the rules of Lorentz transformation can be applied consistently only when one allows the

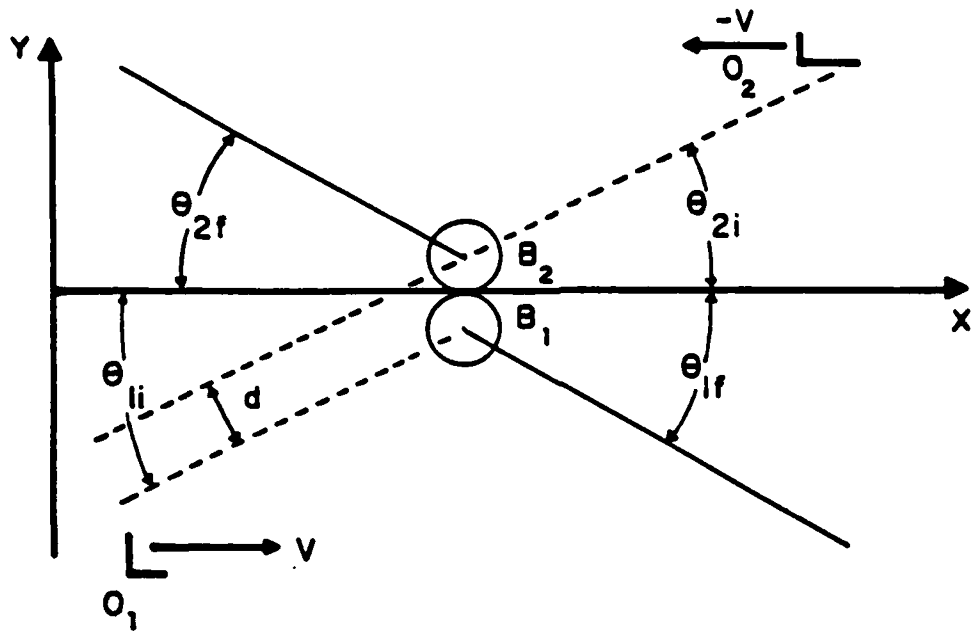


Fig. B-2 COLLISION OF TWO MASSES IN THE LABORATORY FRAME. OBSERVERS O_1 AND O_2 MOVE WITH VELOCITIES SUCH THAT EACH SEES THE OTHER MOVING WITH V TOWARD HIM. EACH TOSSES HIS MASS WITH u IN Y-DIRECTION.

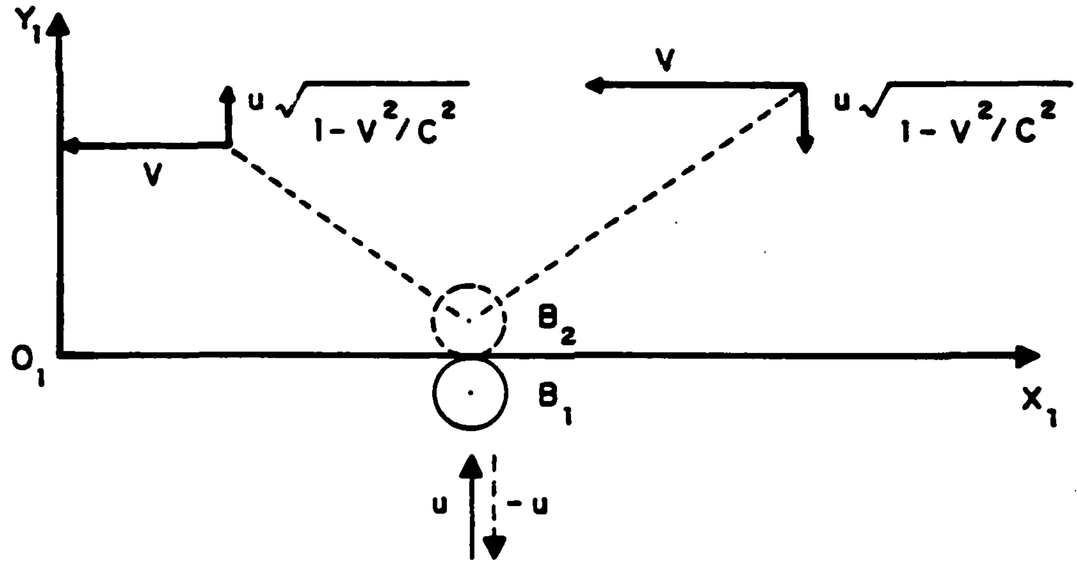


Fig. B-3 COLLISION AS OBSERVED BY OBSERVER O_1 Y-COMPONENT OF u IN O_1 FRAME IS $u\sqrt{1 - V^2/C^2}$

mass to be dependent on the speed. The velocity u in observer O_1's frame was assumed small compared to c. So we can set $m(u) \approx m_o$. With this the momentum conservation requires

$$m(v) = \frac{m_o}{\sqrt{1-\frac{v^2}{c^2}}}$$

B-5 Relativistic Energy

The work energy theorem states that a particles kinetic energy increases according to the work done by the forces acting on the particle to accelerate it.

$$\Delta K = \int_{x_i}^{x_f} F\, dx$$

The force is equal to rate of change of momentum

$$F = \frac{d(m(v)v)}{dt} = m\frac{dv}{dt} + v\frac{dm}{dt}$$

so

$$\Delta K = \int_{x_i}^{x_f} \left(m\frac{dv}{dt} + v\frac{dm}{dt}\right)dx$$

To carry out the integration one makes the following substitutions:

From the mass formula

$$m^2\left(1-\frac{v^2}{c^2}\right) = m_o^2$$

or

$$m^2c^2 - m^2v^2 = m_o^2c^2$$

Differentiating with respect to time gives:

$$2mc^2\frac{dm}{dt} - 2m^2v\frac{dv}{dt} - 2v^2m\frac{dm}{dt} = 0$$

From this follows after dividing by 2mv

$$m\frac{dv}{dt} + v\frac{dm}{dt} = c^2\frac{dm}{dt}\frac{1}{v} = c^2\frac{dm}{dt}\frac{dt}{dx} = c^2\frac{dm}{dx}$$

This gives for the integral

$$\Delta K = \int_{x_i}^{x_f} c^2\frac{dm}{dx} = c^2(m(v_f) - m(v_i))$$

Let $m(v_i) = m_o$ $\quad m(v_f) = \frac{m_o}{\sqrt{1-\frac{v^2}{c^2}}}$

then we have for the kinetic energy:

$$\boxed{K = \frac{m_oc^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_oc^2}$$

For $v \ll c$ one can approximate

$$K = m_oc^2\left[\left(1-\frac{v^2}{c^2}\right)^{-1/2} - 1\right] \approx m_oc^2\frac{1}{2}\frac{v^2}{c^2} = \frac{1}{2}m_ov^2$$

This suggests to identify the first term on the right with the total energy of the particle

$$E(v) = \frac{m_o c^2}{\sqrt{1-\frac{v^2}{c^2}}} \equiv m(v)c^2 = K + m_o c^2$$

The total energy of the particle is its kinetic energy plus an energy that must be associated with every mass at rest, the "Rest Energy"

$$\boxed{E(o) = m_o c^2}$$

For practical calculations one likes to have the energy of the particle in terms of its momentum. In classical Newtonian physics

$$E = \frac{1}{2} m_o v^2 = \frac{p^2}{2m_o}$$

The corresponding relativistic expression is obtained by evaluating

$$m^2c^4 - m_o^2c^4 = m_o^2c^4\left(\frac{1}{1-v^2/c^2} - 1\right) = m_o^2c^4\frac{\frac{v^2}{c^2}}{1-v^2/c^2}$$

$$= \frac{m_o^{\circ}c^2v^2}{1-\frac{v^2}{c^2}} - c^2\, m^2v^2 \equiv c^2p^2$$

Thus $$m^2c^4 = c^2p^2 + m_o^2c^4$$

$$\boxed{E^2 = c^2p^2 + m_o^2c^4}$$

APPENDIX C INVARIANCE OF LAWS OF PHYSICS UNDER TIME REVERSAL

C-1 Invariance of Newton's Law of Motion Under Time Reversal

In classical Newtonian physics for the motion in a conservative force field one has

$$m\frac{d^2\vec{x}(t)}{dt^2} = \vec{F}(\vec{x})$$
$$\vec{P}(t) = m\frac{d\vec{x}(t)}{dt}$$

Making transformation $t \rightarrow t' = -t$ gives

$$\frac{d^2\vec{x}(t')}{dt'} = \frac{d^2\vec{x}(-t)}{dt^2} = \frac{\vec{F}(\vec{x})}{m}$$
$$\text{and} \quad P(t') = m\frac{d\vec{x}(t')}{dt'} = -m\frac{d\vec{x}(-t)}{dt}$$

The equation of motion has the same form in the time reversed system and so the trajectory $\vec{x}(-t)$ is identical in shape as $\vec{x}(t)$. However the momentum vector is reversed and the motion runs reverse.

C-2 Invariance of Electromagnetism Under Time Reversal

Maxwell's equations relating electric and magnetic fields with charge and current densities are

$$\nabla o \vec{E} = \rho$$
$$\nabla \times \vec{B} - \frac{1}{c}\frac{\delta \vec{E}}{\delta t} = \frac{\vec{J}}{c}$$
$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\delta \vec{B}}{\delta t}$$

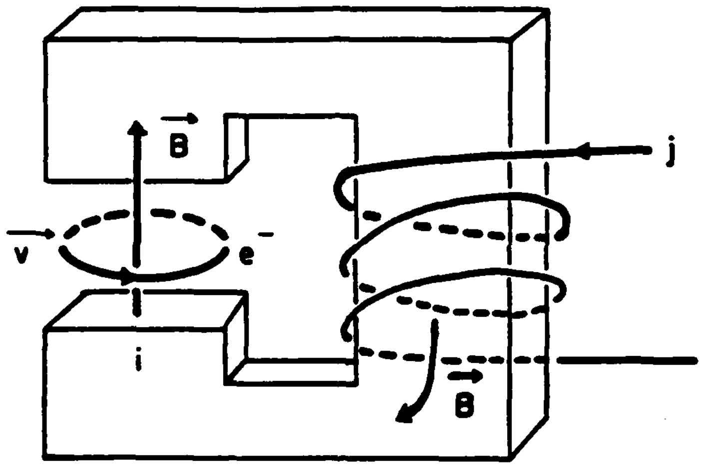


Fig. C-1 INVARIANCE OF ELECTRODYNAMICS UNDER TIME REVERSAL. EXAMPLE OF CHARGED PARTICLE MOTION IN MAGNETIC FIELD.

$$\vec{F} = q\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right)$$

$$m\frac{d\vec{v}}{dt} = \vec{F}$$

TRANSFORMATIONS UNDER TIME REVERSAL:

$t \to -t \quad \vec{v} \to -\vec{v} \quad \vec{F} \to \vec{F}$

$q \to q \quad \vec{E} \to \vec{E} \quad \vec{B} \to \vec{B}$

$j \to -j$

For particle motion in electromagnetic fields to this must be added the Lorentz force law and Newton's equation of motion (we consider small particle velocities only here)

$$\vec{F} = q(\vec{E} + \frac{1}{c}\vec{V} \times \vec{B})$$
$$m\frac{d\vec{V}}{dt} = \vec{F}$$

Consider the example of an electron moving in the gap of an electromagnet (cyclotron) [Fig C-1]

Under time reversal we expect from Newton's law

$$t \to -t$$
$$\vec{V} \to -\vec{V}$$
$$\vec{F} \to \vec{F}$$

Because the charge of the particle stays unaltered, the quantities q, ρ, E, B, and $\vec{j}$ must transform under time reversal as follows

$$q \to -q$$
$$\rho \to -\rho$$
$$\vec{J} \to -\vec{J}$$

This inverts B: $\vec{B} \to -\vec{B}$

While the E-field retains its direction according to Faraday's law.

C-3 Invariance of Eulerian Fluid Motion Under Time Reversal

The Euler equations of motion are

$$\rho\left(\frac{\partial\vec{v}(t)}{\partial t} + (\vec{v}\cdot\vec{\nabla})\,\vec{v}\right) = -\text{grad } p$$

Under transformation $t \to -t$ $\quad \vec{v} \to -\vec{v}$ these become

$$\rho\left(\frac{\partial\vec{v}(-t)}{\partial t} + (\vec{v}(-t)\cdot\nabla)\,\vec{v}(-t)\right) = -\text{grad } p(-t)$$

If the fluid flow is viscous and contains a viscous force term of the form

$$\frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\nabla)\,\vec{v} = -\frac{1}{\rho}\nabla p + \frac{\mu}{\rho}\;\nabla^2\vec{v}$$

then we see that the viscous term changes sign under the transformation and $\vec{v}(-t)$ is not a solution of the same equation as $\vec{v}(t)$.

C-4 Time Reversal Invariance in Non-Relativistic Quantum Mechanics

Non relativistic quantum mechanics of spinless charged particles is well described by the Schrödinger equation with conservative potential.

$$\frac{-\hbar^2}{2m}\;\nabla^2\Psi(\vec{x};t) + V(\vec{x})\,\Psi(\vec{x};t) = i\hbar\frac{\partial\Psi(\vec{x};t)}{\partial t}$$

If $\Psi(\vec{x};t)$ is a solution to this equation we can immediately see that $\Psi(\vec{x};t')$ with $t' = -t$ is not a solution of the same equation. If we insert $\Psi(\vec{x};t')$ we must set

$$\frac{\partial}{\partial t}\Psi(\vec{x};t') = \frac{\partial}{\partial t'}\Psi(\vec{x};t')\cdot\frac{\partial t'}{\partial t} = \frac{\partial}{\partial t'}\Psi(\vec{x};t')\,(-1)$$

So $\Psi(\vec{x};-t)$ solves an equation that has the opposite sign on the right side.

But we can show that if $\Psi(\vec{x};t)$ is a solution, then $\Psi^*(\vec{x};-t)$ is also a solution.

To show this we take the conjugate complex of Schrödinger equation

$$\frac{-\hbar^2}{2m}\nabla^2\Psi^*(\vec{x};t) + V(\vec{x})\,\Psi^*(\vec{x};t) = -i\hbar\frac{\partial\Psi^*(\vec{x};t)}{\partial t}$$

Now we make the time reversal $t \rightarrow t'(t) = -t$

$$\frac{-\hbar^2}{2m}\nabla^2\Psi^*(\vec{x};-t) + V(\vec{x})\,\Psi^*(\vec{x};-t) = i\hbar\frac{\partial\Psi^*(\vec{x};t)}{\partial t}$$

which is indeed the same Schrödinger equation

That the function $\Psi^*(\vec{x};-t)$ is the proper state function to describe the time reversed process can be demonstrated in the simple case of the one-dimensional motion of a free particle. The Schrödinger equation reduces then to

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi(x;t)}{dx^2} = i\hbar\frac{\partial\Psi(x;t)}{\partial t}$$

Solutions to this are plane waves

$$\Psi = ae^{i(kx-\omega t)}$$

which solve the equation provided the dispersion relation

$$\frac{\hbar^2 k^2}{2m} = \hbar\omega$$

holds. With the identification $\hbar k \to p$ and $\hbar\omega \to E$ this is the energy momentum relation for the free particle.

Taking the function

$$\Psi^*(x_1; -t) = ae^{-i(kx+\omega t)} = ae^{i(-kx-\omega t)}$$

we can read off that it belongs to a particle of the same energy and opposite momentum.

APPENDIX D ENTROPY IN THERMODYNAMICS AND STATISTICAL MECHANICS

D-1 First Law of Thermodynamics

There exists for every thermodynamic system a unique state function "Total Internal Energy" U(T,V) which is a function of volume V and Temperature T.

The first law of thermodynamics states:

The internal energy increases by an amount dU due to heat $đQ$ added to the system and decreases by the amount of work $đW = pdV$ done by the system

$$dU = đQ - pdV$$

D-2 The Carnot Cycle

In a cyclic process in which the working gas operates between two heat reservoirs of temperatures $T_1 > T_2$ the efficiency of the process can be given [Fig D-1].

For a reversible cycle like the Carnot cycle it is

$$\text{Efficiency} \quad \eta = \frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1}$$

and also:

$$\frac{Q_1}{T_1} - \frac{Q_2}{T_2} = 0$$

D-3 The Entropy

For reversible cycles it is $\oint đQ/T = 0$

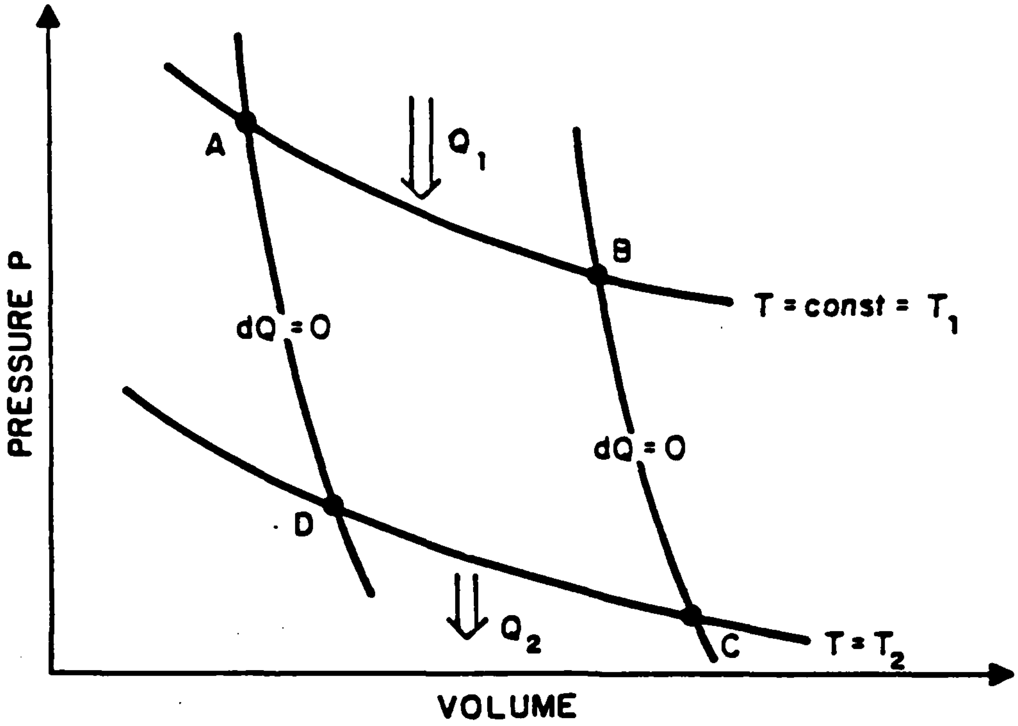


Fig. D-1 CARNOT CYCLE

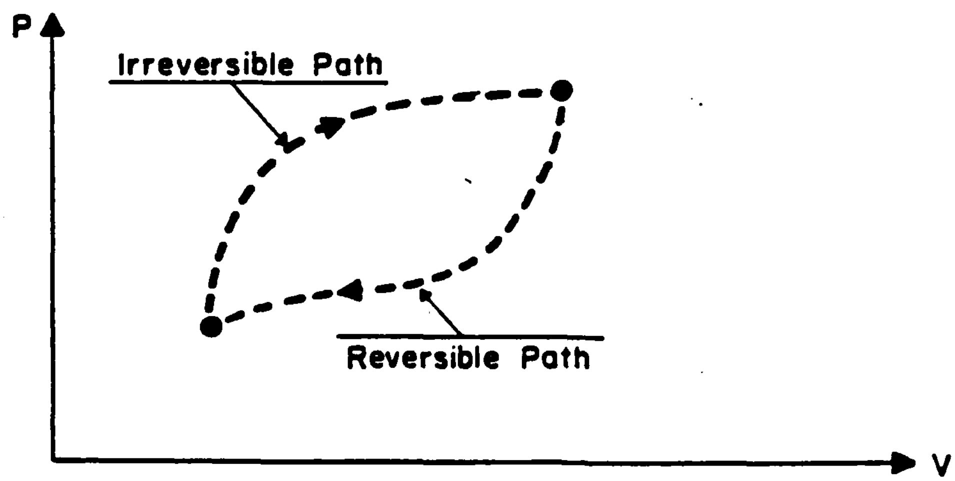


Fig. D-2 CHANGE OF ENTROPY IN IRREVERSIBLE PROCESS

That legitimates the quantity đQ/T as a state variable in the Carnot process:

$$\int_A^C \frac{đQ}{T} - \int_C^A \frac{đQ}{T} = 0 \Rightarrow \int_{A \text{ path } 1}^{C} \frac{đQ}{T} = \int_{A \text{ path } 2}^{C} = \frac{đQ}{T} = S_c - S_A$$

D-4 Clausius Theorem for Non Reversible Cycles Between T_1 and T_2

$\eta_{irrev} \leq \eta_{rev}$

This implies

$$\frac{Q_1 - Q_2}{Q_1} \leq \frac{Q_{1\,rev} - Q_{2\,rev}}{Q_{1\,rev}}$$

$$1 - \frac{Q_2}{Q_1} \leq 1 - \frac{Q_{2\,rev}}{Q_{1\,rev}} \Rightarrow \frac{Q_2}{Q_1} \geq \frac{Q_{2\,rev}}{Q_{1\,rev}} = \frac{T_2}{T_1}$$

It follows

$$\frac{Q_2}{T_2} \geq \frac{Q_1}{T_1} \Rightarrow \frac{Q_1}{T_1} - \frac{Q_2}{T_2} = \oint \frac{đQ}{T} \leq 0$$

D-5 Change of Entropy in Irreversible Process

$$\oint \frac{đQ}{T} \leq 0$$

$$\int_{A \, irrev}^{B} \frac{đQ}{T} + \int_{B \, rev}^{A} \frac{đQ}{T} \leq 0$$

$$\int_{A \, rev}^{B} \frac{đQ}{T} = S_B - S_A \Rightarrow \int_{A \, irrev}^{B} \frac{đQ}{T} \leq S_B - S_A$$

For irreversible processes [Fig D-2]

$$ds \geq \frac{đQ}{T}$$

D-6 Second Law of Thermodynamics

For isolated systems $đQ = 0$

This implies $\boxed{dS \geq 0}$

Entropy of an isolated systems cannot decrease. Entropy of an isolated system must tend toward a maximal value. The final equilibrium state will be a state of maximal entropy.

D-7 Entropy of System Plus its Surrounding

(1) System undergoing complete irreversible cycle initial and final state are identical.

Because entropy is state function $\Delta S_{syst} = 0$

From Clausius theorem it follows:

$$\oint \left(\frac{đQ}{T}\right)_{syst} \leq \Delta S_{syst} = 0$$

Heat is exchanged with surrounding. The best possible exchange is a reversible exchange [Fig D-3]

$$đQ_{sourr} = -đQ_{syst}$$

It follows:

$$0 = \Delta S_{syst} \geq \oint \left(\frac{đQ}{T}\right)_{syst} = -\oint \left(\frac{đQ}{T}\right)_{sourr}$$

Only if the heat transfer is reversible can we equate the last

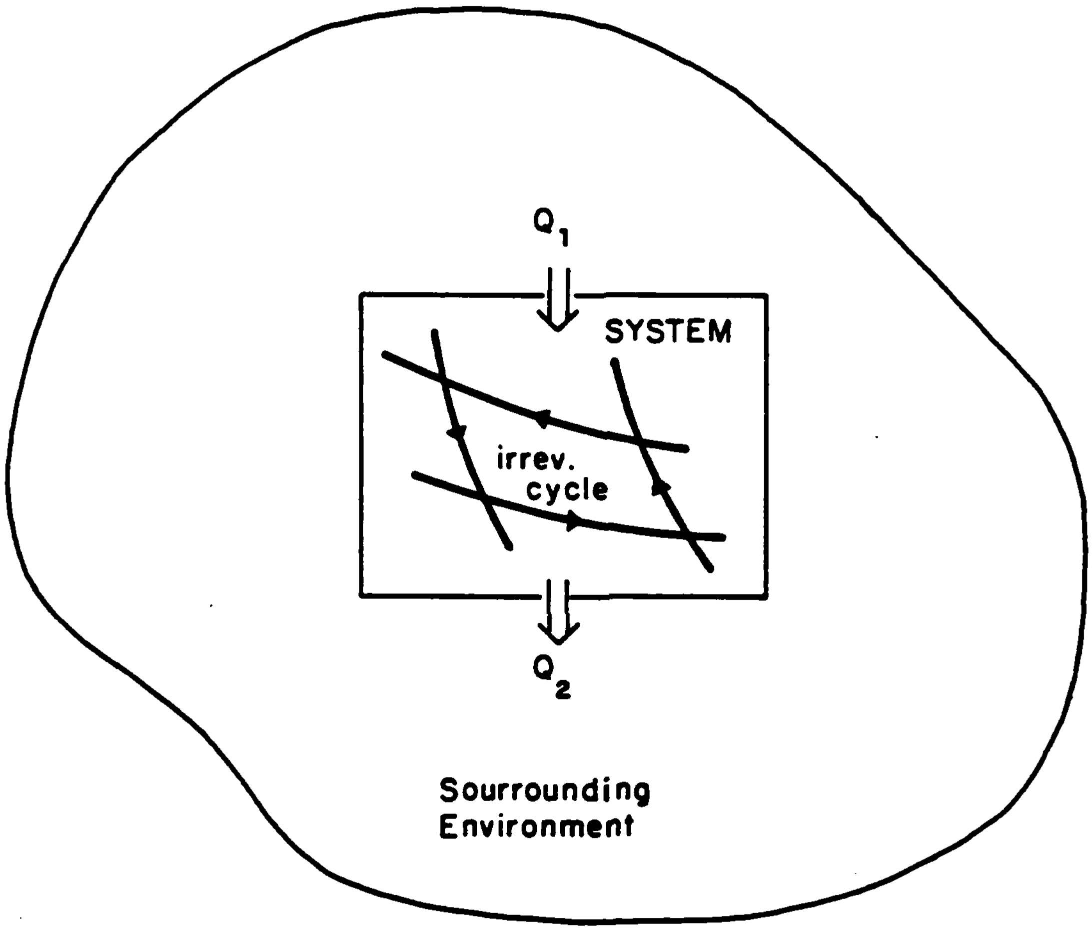


Fig. D-3 ENTROPY OF SYSTEM AND ITS ENVIRONMENT quantity to ΔS_{sourr}.

Thus $\quad 0 \geq -\Delta S_{sourr}$

or $\quad \Delta S_{sourr} \geq 0$

(2) Irreversible System Change From State 1 → state 2

$$\Delta S_{syst} \geq \int_1^2 \frac{dQ_{irr}}{T} = -\int\left(\frac{dQ}{T}\right)_{sourr} = -\Delta S_{sourr}$$

Therefore $\quad \Delta S_{syst} + \Delta S_{sourr} \geq 0$

In the presence of irreversible processes the total entropy of the universe increases.

D-8 The Entropy in Statistical Interpretation

(1) Entropy as measure of probability of state

System 1	System 2
State a	State b
Entropy S_a	Entropy S_b

Probability of finding system 1 in state a $P_a^{(1)}$

Probability of finding system 2 in state b $P_b^{(2)}$

Probability of finding system 1 in state a and system 2 in state b

$$P(a,b) = P_a^{(1)} \cdot P_b^{(2)}$$

Entropies: $\quad S_a = f(P_a^{(1)}) \qquad S_b = f(P_b^{(2)})$

In order that entropy be a proper state function it must be additive:

$$S_{total} = S_a + S_b$$

f must be a universal function

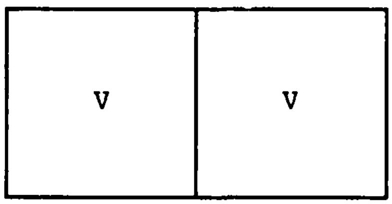
$$f(P(a,b)) = f\left(P_a^{(1)} \cdot P_b^{(2)}\right) = f\left(P_a^{(1)}\right) + f\left(P_b^{(2)}\right)$$

The function f that has this property is the logarithmic function

$$S = C \ln P$$

(2) <u>Determination of Constant C</u>

Consider Joule expansion of 1 mol of gas from volume V to 2V



Probability for one molecule to be in left compartment V: $P_1 = 1/2$.

Probability for all N molecules to be in V $P_1 = (\frac{1}{2})^N$

Probability for all molecules to be in 2V $P_2 = 1$

The entropy difference is then:

$$S_2 - S_1 = C \ln \frac{P_2}{P_1} = C \ln 2^N = CN \ln 2$$

Comparison with thermodynamics:

$$S_2 - S_1 = R \ln V_2/V_1 = R \ln 2$$

Therefore:

$$R = C\,N$$

$$C = \frac{R}{N} = k_{Boltzmann} = 1.38 \times 10^{-23} \text{ Joule/K}$$

D-9 Statistics of Distribution of States - An Example

As a simple example for which the distribution function can be easily derived we consider the case of the statistics of distributing N molecules into two equal compartments of a container

(1)	(2)
m	(n-m)
molecules	molecules
$m = \frac{N}{2} + x$	$N - m = \frac{N}{2} - x$

The number of ways in which the N molecules can be distributed with m in compartment (1) and (N-m) molecules in compartment (2) is given by the rules of combinatorics as

$$g(N,m) = \frac{N!}{m!\ (N-m)!}$$

The probability of the state (N,m) is taken to be proportional to the number of ways in which this state can be realized by interchanging particles.

$$P(N,m) = P(N,x) \propto g(N,x) = \frac{N!}{\left(\frac{N}{2}+x\right)!\left(\frac{N}{2}-x\right)!}$$

The entropy associated with the state (N,x) is given by

$$S(N,o) - S(N,x) = k \ln \frac{g(N,o)}{g(N,x)} = k \ln \frac{P(N,o)}{P(N,x)}$$

It is: $\ln g(N,x) = \ln N! - \ln\left(\frac{N}{2}+x\right)! - \ln\left(\frac{N}{2}-x\right)!$

For large N one can use the stirling formula to deal with the large factorials

$$\ln N! \approx N \ln N - N$$

With this we obtain:

$$\begin{aligned}
\ln g(N,x) &= \left[N \ln N - N - \left(\frac{N}{2}+x\right) \ln\left(\frac{N}{2}+x\right) + \left(\frac{N}{2}+x\right) + \left(\frac{N}{2}-x\right) \ln\left(\frac{N}{2}-x\right) + \left(\frac{N}{2}-x\right) \right] \\
&= \left[N \ln N - \left(\frac{N}{2}+x\right) \ln \frac{N}{2}\left(1+\frac{2x}{N}\right) - \left(\frac{N}{2}-x\right) \ln \frac{N}{2}\left(1-\frac{2x}{N}\right) \right] \\
&= \left[N \ln N - N \ln \frac{N}{2} - \left(\frac{N}{2}+x\right) \ln\left(1+\frac{2x}{N}\right) - \left(\frac{N}{2}-x\right) \ln\left(1-\frac{2x}{N}\right) \right] \\
&= N \ln 2 - \left(\frac{N}{2}+x\right) \ln\left(1+\frac{2x}{N}\right) - \left(\frac{N}{2}-x\right) \ln\left(1-\frac{2x}{N}\right)
\end{aligned}$$

for $2x \ll N$ we can approximate further

$$\ln\left(1 \pm \frac{2x}{N}\right) \approx \pm \frac{2x}{N} - \frac{1}{2}\frac{4x^2}{N^2} = \frac{2x}{N}\left(\mp 1 - \frac{x}{N}\right)$$

This gives:

$$\ln g(N,x) \approx N \ln 2 - \left(\frac{N}{2}+x\right)\frac{2x}{N}\left(1-\frac{x}{N}\right) - \left(\frac{N}{2}-x\right)\frac{2x}{N}\left(-1-\frac{x}{N}\right)$$

$$\ln g(N,x) = N \ln 2 - \frac{2x^2}{N}$$

Therefore

$$\boxed{g(N,x) = 2^N e^{\frac{-2x^2}{N}}}$$

The entropy difference is:

$$S(N,o) - S(N,x) = k \ln \frac{2^N}{2^N e^{\frac{-2x^2}{N}}} = k \ln \frac{1}{e^{\frac{-2x^2}{N}}}$$

The probability density distribution is given by:

$$P(N,x)dx = C\, g(n,x)dx = \frac{4}{\sqrt{2\pi N}} e^{\frac{-2x^2}{N}} dx$$

The probability distribution has Gaussian shape [Fig D-4].

For $N = 10^{23}$ the half width is $x = 1.8 \times 10^{11}$ which corresponds to a non equilibrium of one in 10^{12}. For a non equilibrium of one in 10^{10}, i.e. $x = 10^{13}$ the probability becomes $e^{-2000} \simeq 10^{-868}$. With $k = 1.38 \times 10^{-23}$ the entropy difference is a mere 2000k while $S(N,o) \approx kN \ell n\, 2$.

D-10 <u>The Foundation of the Second Law in Statistical Mechanics</u>

(1) <u>Distribution Function in μ-Phase Space</u>

In a recently published book "The Physical Basis of the Direction of Time" H.D Zeh [ZE 89] gives a detailed analysis of the ingeniousness and subtleness of Boltzmann's H-theorem, which forms the basis of the second law.

Boltzmann's derivation is based on the statistics in the 6-dimensional phase space of the space-and-momentum coordinates $\vec{q},\vec{P}$ of the molecules.

A system of N molecules, with $N \approx 10^{23}$ is represented by its N points in this 6-dimensional "μ-space". These phase points move about in this phase space due to the thermal motion of the molecules.

Strictly speaking this set of points should be described by a

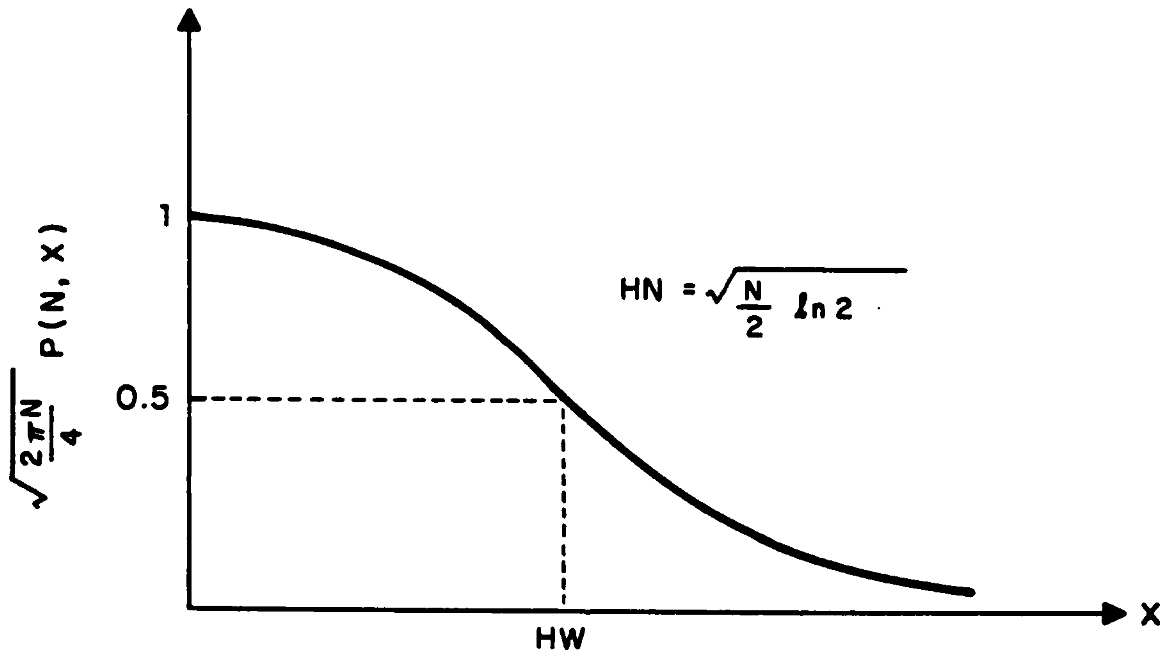


Fig. D-4 PROBABILITY DISTRIBUTION FUNCTION phase space distribution function that consists of N δ-functions, placing the N objects at their respctive places into the phase space

$$\rho_\mu = \sum_{K=1}^{N} \delta(\vec{q}-\vec{q}_K^{(t)})\, \delta(\vec{P}-\vec{P}_K^{(t)})$$

The integral over the entire phase space is N

$$\int \rho_\mu \, d_q^3 \, d_P^3 = N$$

Any physical variable that is a function of the molecules is gi 'en as an average

$$\bar{f} \equiv \frac{1}{N} \int f(\vec{q}, \vec{P})\, \rho_\mu(\vec{P}, \vec{q})\, d_q^3 \, d_P^3$$

Because N is such a large number and the points in phase space lie so densely it is suggestive to replace the discrete distribution function by a continuous function $\rho_\mu(\vec{q}, \vec{P}\ t)$. To capture principal features like the width of the distribution function Boltzmann introduced the famous H-functional:

$$H[\rho_\mu] \equiv \int \rho_\mu(\vec{q}\,\vec{P}\,t)\, \ln\, \rho_\mu(\vec{q}\,\vec{P}\,t)\, dq^3\, dP^3 = N\, \overline{\ln \rho_\mu}$$

The average value of $\ln \rho$ is obviously large for narrow peaked distributions and small for flat spread out one's.

(2) <u>The Boltzmann H-Theorem</u>

This famous theorem is derived under the following assumptions:

(a) the system is isolated

(b) the system is homogeneous $\rho_\mu = \rho_\mu(\vec{P}, t)$

(c) Collision number assumption: ρ_μ changes in time as molecules through collisions disappear from one phase space volume element and appear in another. Energy and momentum in these collisions are conserved. The collision probabilities are space and time symmetric.

$$\frac{\partial \rho_\mu(\vec{P}t)}{\partial t} = \text{Gain} - \text{Loss}$$

Under these assumptions one can prove the statements

(I) $$\frac{dH[\rho_\mu]}{dt} \leq 0$$

(II) ρ_μ (t) in time approaches a Maxwellian distribution ρ_μ and for this H is minimal

(III) The minimal H function is proportional to the entropy

$$S = -k\, H[\rho_\mu] \equiv S_\mu[\rho_\mu]$$

The so defined entropy is a measure of the width of the μ-phase space distribution function ρ_μ.

The collision model provides the mechanism that lets the distribution evolve toward the Maxwell distribution.

The fact that this model gives a finite entropy comes from the assumption of a smooth continuous distribution function ρ_μ $(\vec{P},t)$ instead of a sum of δ-functions. The discrete, δ-functions form of the distribution function gives an infinite H functional. The proof of the theorem cannot be carried out with it.

The justification for replacing the discrete distribution by a continuous one is that coordinates and momenta are assumed to be "uncertain" or incompletely observable.

(3) The Instability of particle trajectories

Suppose a molecule collides with a number of fixed hard spheres [Fig D-5].

The velocity of the particle remain unchanged and only the direction of the velocity is considered. Starting from the origin if the direction makes a small change. $\delta\theta(o)$ the trajectory will change for every collision. Assuming that the density of the hard spheres is low and that the mean free path λ of the particle is large compared to the radius a of the hard spheres we can see from the figure that

$$\delta\theta' = 2\delta\alpha \doteq \frac{2\lambda\,\delta\theta}{a\cos\alpha}$$

Therefore $$|\delta\theta'| \geq (2\lambda/a)\ |\delta\theta|$$

After n collisions

$$|\delta\theta(t)| \geq |\delta\theta(o)|\ e^{t/\tau}$$

where $$\tau = \frac{\lambda}{V \ln \frac{2\lambda}{a}}$$

With V being the speed of the particle

If $t \geq \tau\ \ln\left(1/|\delta\theta(o)|\right)$ then $|\delta\theta(t)| \sim 1$ and the direction is

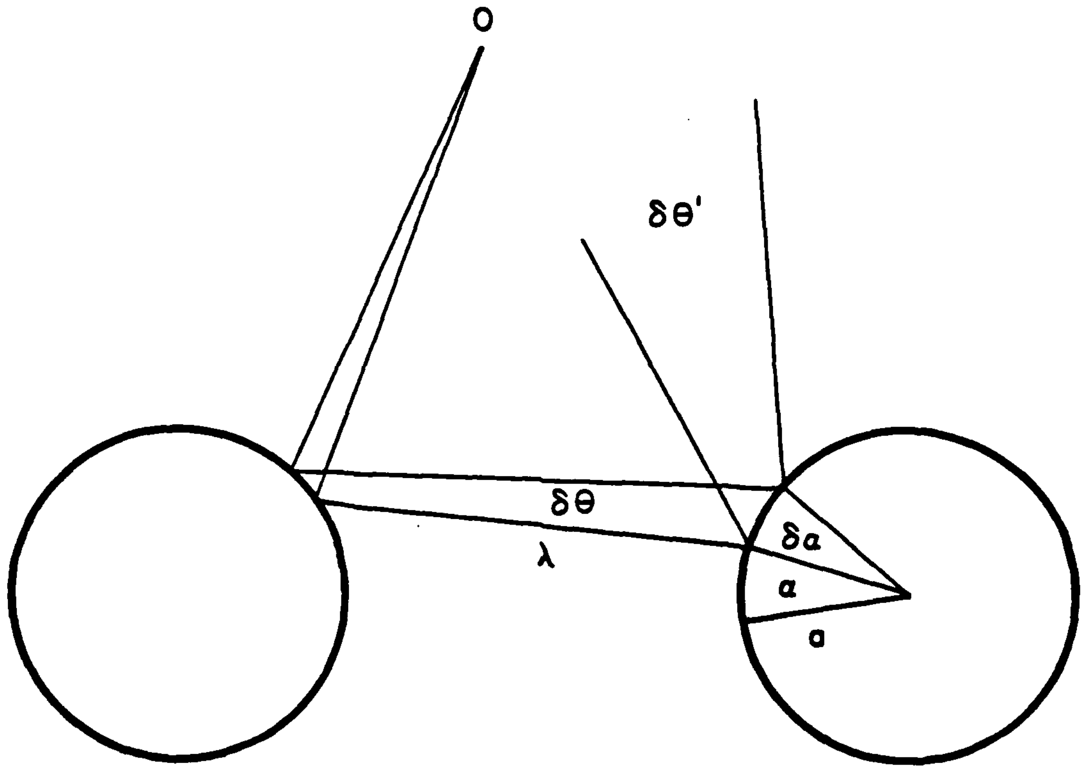


Fig. D-5 UNCERTAINTY IN COLLISION PROCESSES

$$\delta\theta' = 2\,\delta\alpha = \frac{2\lambda\delta\theta}{a\cos\alpha} \qquad \lambda > a$$

$$|\delta\theta'| \geq (2\lambda/a)\ |\delta\theta|$$

After n collisions: $\delta\theta(t) = \left(\frac{2\lambda}{a}\right)^n \delta\theta(o) \qquad n = \frac{t}{\lambda/v}$

$$\ln\frac{\delta\theta(t)}{\delta\theta(o)} = \frac{t}{\lambda/v}\ln\left(\frac{2\lambda}{a}\right)$$

$$|\delta\theta(t)| \geq |\delta\theta(o)|\,e^{t/\tau}$$

$$\tau = \frac{\lambda}{v\ln\frac{2\lambda}{a}}$$

randomized.

One can estimate that the displacement of grams of mass by centimeters at the distance of 4 light years will have a gravitational effect that cause a deflection $\delta\theta(o)$ during one free path which would have grown after 180 collisions in a few nanoseconds to a magnitude of $|\delta\theta| \approx 1$.

This consideration does two things:

(a) it gives justification to the notion that the positions and moments are uncertain to some extent and therefore justify a continuous distribution function. This leads to a loss of information about correlations that are created by the collision. Which must obey conservation and symmetry laws.

(b) No system is ever totally isolated from these minute. gravitational influences of the surrounding mass distribution in the universe. That is, why the Boltzmann model of a continuous distribution function and the H-theorem is such a successful model, for ideal gases.

4. <u>Gibbs' Γ-Phase Space Dynamics</u>

For the statistical mechanical description of real gases, and solids and liquids correlations between particles are an essential feature of the dynamics of such systems. The Boltzmann μ-space statistics is clearly insufficient because it neglects those correlations that are associated with the collisional interactions. To deal with these kinds of problems Gibbs introduced the ensemble concept and the 6N-dimensional Γ-phase space.

The Γ-space has one dimension for each of the six degrees of freedom for each of the N molecules. The phase space coordinates are $q = q_1 \cdots q_{3N}$ $P = P_1 \cdots P_{3N}$ $dq\ dp = d^{3N}q\ d^{3N}p$.

A complete system of N molecules is represented by a single point in Γ-space. This single point in 6N-dimensional space carries all dynamic information about the system. As the system evolves in time the phase point moves about through a trajectory in Γ-space. All information about correlations that are associated with the changing state are preserved in the phase-point.

To do statistics Gibbs considers very many systems with identical conditions, an "ENSEMBLE" of states.

The statistics of this ensemble of points in Γ-space can be described by a probability density function $\rho_\Gamma(qP)dq\ dp$ which is the probability to find the state of an arbitrary member of the ensemble to be given by the values P,q in a volume element dp dq.

The probability is normalized such that

$$\int \rho_\Gamma(P,q)\,dp\,dq = 1$$

The analog to Boltzmann's H-functional is Gibbs' "Extension in phase" η

$$\eta[\rho_\Gamma] = \overline{\ln \rho_\Gamma} = \int \rho_\Gamma(Pq) \ln \rho_\Gamma(P,q)\, dp\, dq$$

With it one can define a analog to the entropy, the statistical entropy S_Γ

$$S_\Gamma = -k\,\eta[\rho_\Gamma]$$

The meaning of these functionals becomes clear in the simple example when ρ_Γ is constant in a small volume element ΔV_Γ of phase space and zero otherwise.

Then one simply has: $\rho_\Gamma \Delta V_\Gamma = 1$

$$\eta[\rho_\Gamma] = \int \frac{1}{\Delta V_\Gamma} \ln \frac{1}{\Delta V_\Gamma}\, dq\, dp = -\int \frac{1}{\Delta V_\Gamma} \ln \Delta V_\Gamma\, dp\, dq = -\ln \Delta V_\Gamma$$

and $S_\Gamma = k \ln \Delta V_\Gamma$

The statistical entropy is proportional to the logarithm of the size of the phase space volume element occupied by the ensemble.

With these principles a number of important results and theorems can be derived.

(a) Applied to a distribution of statistically independent particles the distribution functions breaks into a product of N single particle distribution functions ρ_μ. The extension in phase η then becomes

$$\eta[\rho_\Gamma] = H[\rho_\Gamma] - N \ln N$$

The statistical entropy is equal to the Boltzmann entropy except for the self mixing entropy kN ℓn N $\simeq$ kN ℓn N!

(b) When the system is constrained to have fixed mean energy

$$\bar{E} = \int H(p,q)\,\rho(p,q)\,dp\,dq$$

Then S_Γ can be maximized by the canonical distribution

$$\rho_{can} \equiv Z^{-1}\exp\{-H(p,q)/kT\}$$

Where the "partition function" is

$$Z \equiv \int e^{-H(p,q)/kT}\,dp\,dq$$

ρ_{CAN} represents an absolute maximum for S_Γ

(c) Liouville Theorem.

The motion of the ensemble points in phase space can be treated like a fluid flow with a local velocity vector

$$\vec{V}_\Gamma \equiv \{\dot{p}_1 \cdots \dot{p}_{3N}\ \dot{q}_1 \cdots \dot{q}_{3N}\}$$

These derivatives must satisfy Hamilton equations

$$\dot{p}_K = -\frac{\partial H}{\partial q_K} \qquad \dot{q}_K = \frac{\partial H}{\partial p_K}$$

The ensemble point flow must conserve the probability. So, Gauss' law of continuity must hold:

$$\frac{\partial \rho_\Gamma}{\partial t} + \mathrm{div}_\Gamma(\rho_\Gamma \vec{V}_\Gamma) = 0$$

From this follows that the flow must be incompressible

$$\mathrm{div}_\Gamma \vec{V}_\Gamma = 0$$

That means if we pick a volume element ΔV_Γ and watch the phase points in it move we find that the density ρ_Γ remains constant. That means that the volume occupied by these constant density points remains constant.

(d) The ensemble entropy S_Γ remains constant throughout the time

$$\frac{dS_\Gamma}{dt} = 0$$

This can be proven directly by taking the time derivative of the defining relation for S_Γ, and making use of Liouville theorem.

That the statistical entropy does not show the arrow of time has its reason in the fact that in this statistical treatment all information of the N-particle dynamics is preserved through the explicit use of the Hamilton equations, which are time symmetric.

In order to introduce the second law behavior one has to in some way discard dynamical information. That procedure of discarding information must be such that it preserves the important aspects about the width of the distribution but ignores dynamically unimportant details.

Gibbs has introduced such a device in analogy to the behavior of an ink drop in water, which behaves as nearly incompressible fluid. When the water is stirred the ink will rearrange into a network of tiny tubes. The totality of these tubes still has the same volume as the original drop on account of Liouville's theorem. But from a coarse grained view the ink seems to have spread over a much larger volume. We have discarded the information in the details of the tiny tubes.
Analogously Gibbs defines a coarse grained entropy S_{Gibbs} which is a functional of a coarse grained distribution function ρ^{cg}, which is defined to be a constant within very small but otherwise arbitrarily chosen fixed volume elements ΔV_m in Γ-space. This device loses just enough of the details of the Hamiltonian dynamics of the motion of the phase points that it is perceived as a general dispersion of the phase points on a coarse grained level just like the ink appears to spread and turn a larger volume slightly blue.
Using this definition of coarse grained entropy one can indeed prove

$$\frac{dS_{Gibbs}}{dt} \geq 0$$

(5) In concluding this overview of the statistical mechanics

foundation of the second law *) we see that classical statistical mechanics in its virgin state has no place for an arrow of time. The laws of mechanics are time symmetric. The arrow of time and the second law behavior is brought into the theory by an additional element. This additional assumption amounts to a coarse graining whereby certain information is discarded which is considered however dynamically irrelevant for the behavior of the macroscopic system. With this coarse graining device, classical statistical mechanics gives a proper account of the observed behavior of systems following the second law.

*) This presentation has closely followed that given by Zeh [ZE 89]

APPENDIX E BLACK HOLE PHYSICS

E-1 The Schwarzschild Metric

In the theory of general relativity the line element for a spherical symmetric mass distribution outside the mass is given by the Schwarzschild metric

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right)c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\Psi^2).$$

The spatial portion of the metric is in spherical polar coordinates with the mass M centered at $r = 0$. One can see that the metric becomes singular at the Schwarzschild radius:

$$R_s = \frac{2GM}{c^2}$$

There is nothing singular about the physics at that radius. The singularity is merely an artifact of the choice of coordinates, just as planar polar coordinates are unsuitable to describe the surface of a sphere beyond $r = R$ when R is the radius of the sphere. By transformation to other suitable coordinates the interior of the Schwarzschild sphere can be properly described.

Propagation of light is along null-geodesics characterized by $ds = 0$. For a radially propagating light signal the propagation speed appears to an observer as

$$\frac{dr}{dt} = c\left(1 - \frac{2GM}{c^2 r}\right)$$

At a distance large compared to the Schwarzschild radius the propagation velocity is c. The metric for large distances assumes Euclidean form.

At the Schwarzschild radius the propagation speed goes to zero. That means that no light from the interior will ever be observable at a distant observer. This led to the notion of Black Hole.

E-2 Black Hole Entropy and Temperature

The sphere of radius R_s is called the Schwarzschild horizon. The area of this horizon sphere is

$$A_s = 4\pi R_s^2 = 16\pi G^2 M^2/c^4$$

If the mass of the black hole is increased by an increment dM of mass falling into the black hole then the horizon area increases by the amount

$$dA_s = \frac{32\pi G^2}{c^4} M dM = \frac{32 G\pi}{c^2} R_s dM$$

Since nothing that has entered the black hole through the Schwarzschild horizon can escape it again, the horizon area can only increase

$$dA_s > 0$$

This behavior is analogous to the increase of entropy of a closed system. This idea has led to a quantitative comparison of the black hole behavior and thermodynamics which turned out to be extremely fruitful.

In thermodynamics the first law in the absence of pressure for a closed system states

$$dU = T ds \qquad ds > 0$$

For black holes we have the analog:

$$d(Mc^2) = \frac{c^4}{32\pi G R_s} dA_s \qquad dA_s > 0$$

The analogy suggests a relation between change of horizon area and change of entropy

$$dS = f\, dA$$

From the black hole relation we state:

$$\frac{1}{32\,\pi\, R_s}\, dA_s = \frac{G}{c^4}\, d(Mc^2) = \frac{G}{c^4}\, k_B T\, \frac{dS}{k_B}$$

It must be $\frac{dS}{k_B} = \frac{f\, dA_s}{k_B}$

This ratio is dimensionless. Therefore

$$\frac{k_B}{f} = (\text{Length})^2$$

The length must be a universal length involving G in some way. The only universal length known that involves the constants G and c is the Planck length

$$Lp = \left(\frac{G\hbar}{c^3}\right)^{1/2} = 10^{-33}\,\text{cm}$$

Therefore $\qquad f = k_B/L_p^2$

The black hole relation then gives

$$\frac{1}{32\pi R_s}\, dA_s = \frac{G}{c^4}\, T_s\, f\, dA_s = \frac{G}{c^4}\, T_s\, \frac{k_B}{L_p^2}\, dA_s$$

This gives for the black hole the temperature

$$T_s = \frac{c^4 L_p^2}{32\,\pi\, k_B G R_s} = \frac{\hbar c}{32\,\pi\, k_B R_s}$$

Using the expression for L_p this can be rewritten as

$$k_B T_s = \frac{(\hbar c/G)\, c^2}{64 \pi M} = \frac{\hbar}{16 \pi c} \left(\frac{GM}{R_s^2} \right)$$

This shows that the black hole temperature is proportional to the surface gravitational acceleration at the horizon.

Using the solar mass of $M_\odot = 2 \times 10^{33}$ gr one can write

$$T_s = 10^{-7} \left(\frac{M_\odot}{M} \right) \ ^\circ K$$

The entropy is:

$$S = f \cdot A = 16 \pi \frac{k_B G}{\hbar c} M^2$$

A more sophisticated approach to determine the magnitude f that goes beyond mere dimensional analysis gives another factor 1/4. This factor comes from some statistical argument.

The accepted expression for the entropy is

$$\boxed{S = 4 \pi \frac{k_B G}{\hbar c} M^2}$$

E-3 Thermal Emission From the Surface of Black Holes

If the horizon has a temperature assigned to it, it should radiate according to classical statistical thermodynamics. The emission from the surface A_s should have the flux per unit area

$$F_s = \sigma T_s^4$$

Where σ is the Stefon Boltzmann constant

$$\sigma = \frac{2\pi^5 k_B^4}{15 c^2 h^3}$$

The rate of energy loss of the black hole is then

$$\frac{d(Mc^2)}{dt} = -F_s A_s = -\alpha \frac{c}{L_p} \frac{M_p^3}{M^2} c^2$$

Where $M_p = (\hbar c/G)^{1/2} = 10^{-5}$ gr is the "Planck mass", α is some numerical constant of order one and $L_p/c = t_p = 10^{-44}$ sec is the "Planck time". Integrating the equation from M to M = 0 gives as total evaporation time

$$T_H \sim \frac{1}{\alpha}\left(\frac{M}{M_p}\right)^3 t_p \approx 10^{66}\left(\frac{M}{M_\odot}\right)^3 \text{ years.}$$

The radiation that leaves the horizon with a power level of

$$P_H \approx \frac{\alpha c^2}{t_p}\left(\frac{M_p^3}{M^2}\right)$$

is referred to as "Hawking Radiation".

The appearance of the Planck constant $\hbar$ in the expressions for the temperature and entropy signals that quantum physics has entered the picture. The entry of $\hbar$ was here made through the universal Planck length. It suggests that the Hawking radiation is due to some sort of quantum tunneling effect by which energy leaks out of the system that classically could not overcome the gravitational potential barrier.

E-4 Quantum Tunneling From Vacuum Fluctuations Near the Black Hole Horizon

A simple picture of the mechanism of Hawking radiation is obtained by assuming that in the strong gravitational field near the horizon virtual pair creation is occurring with one particle falling into the hole the other escaping to large distance to be observed.

The quantitative argument for this radiation can be made approximately by assuming that after virtual pair creation the newly created particle may separate by one Compton length, which is the accuracy with which the position of a particle can be determined according to the uncertainty principle. If in this distance the work done by the gravitational tidal forces is of the order of the rest energy of the particle pair, then pair production is energetically favorable. Using classical Newtonian quantities, the differential of gravitational force over a radial distance Δr is

$$\Delta F = \frac{\partial F}{\partial r}\,\Delta r \approx \frac{2GmM}{r^3}\,\Delta r$$

The work done by the tidal forces in separating the particles by Δr is

$$\Delta W = \Delta F\cdot\Delta r = \frac{2GmM}{r^3}\,\Delta r^2$$

If the separation is taken to be the *Compton length of* the particle of mass m

$$\Delta r = \lambda_c = \frac{\hbar}{mc}$$

and $r = R_s = \dfrac{2GM}{c^2}$ then the condition for pair creation becomes

$$\frac{2GMm}{R_s^3}\,\lambda_c^2 \geq 2mc^2$$

This gives $\lambda_c \geq R_s\sqrt{2}$

For photons it is

$$\lambda_c = \frac{h}{m_{ph}\, c} = \frac{hc}{h\nu} = \lambda,$$

the wavelength of the photons that can be emitted.

The black body temperature corresponding to this radiation is

$$k_B T_s \sim h\nu = \frac{hc}{\lambda_c} \sim \frac{hc}{R_s} \qquad T_s \sim \frac{\hbar c}{k_B R_s}$$

This is up to some factor 32π identical to the previous expression for T_s. Another way of showing that the Hawking radiation must exist as consequence of the uncertainty principle is to say that the energy of a system confined to a sphere of radius R_s must be

$$\Delta E \approx \frac{\hbar}{\Delta t}$$

The amount ΔE can escape through the Schwarzschild surface if $\Delta t \approx R_s/c$

Thus, $\Delta E \approx \frac{\hbar c}{R_s} = \frac{\hbar c^3}{2GM}$

The power radiated is then

$$P = \frac{\Delta E}{\Delta t} = \frac{\hbar}{\Delta t^2} = \frac{\hbar c^2}{R_s^2} = \frac{\hbar c^6}{4G^2M^2}$$

$$= \left(\frac{\hbar c}{G}\right)\left(\frac{\hbar c}{G}\right)^{1/2}\left(\frac{c^3}{\hbar G}\right)^{1/2} \frac{c^3}{4M^2}$$

$$= \frac{1}{4}\,\frac{c^2}{t_p}\,\frac{M_p^3}{M^2}$$

Which is the same as derived above.

REFERENCES

ADM 62 Arnowi, H.R., Deser, S., and Misner C.W., The Dynamics of General Relativity in: Gravitation, Introduction to Current Research, L. Witten (Ed.), Wiley, (1962)

Al 18 Alexander, Samuel "Space, Time and Deity."

BA & Ti 86 Barrow, J.D. and Tipler F.J. "The Anthropic Cosmological Principle." Clarendon, Oxford, (1986)

BE 10 Bergson, H. "Time and Free Will." McMillan, New York, (1910)

BE 73 Beckenstein, J.D. "Black Holes and Entropy." Phys. Rev D7 2333, (1973)

BR 66 Brandon, S. G.F. Time and the Density of Man, in "The Voices of Time." J.T. Fraser (Ed.), Brazillier, New York, (1966)

BUE 56 Buenning E. "Endogenous Rhythms in Plants." Ann. Rev. of Plant Physiology, Vol 7, P. 71, (1956)

BUR 08 Burnet, J. "Early Greek Philosophy." London, (1908)

CAR 1 Carrol, Lewis "Through the Looking Glass." St. Martin's Press, New York, (1977)

CAS 89 Castagnino, M. "Probabilistic Time in Quantum Gravity." Phys. Rev. D, Vol. 39, No. 8, 15 April 1989

CL-TH 61 Cloudsley-Thompson, J.L. "Rhythmic Activity in Animal Physiology and Behavior." Acad. Press, New York, (1961)

DW 67 DeWitt, B.S. "Quantum Theory of Gravity." Phys. Rev. 160, p. 1113, (1967)

EF 63 Efron, R. "An Extension of the Pulfrich Stereoscopic Effect." Brain, Vol. 86, p. 295, (1963)

Fi 62 Fischer, R. et al, "Biological Aspects of Time," Ann, N.Y. Acad. Sci., Vol. 96, p. 44, (1962)

Fi 66 Fischer, R., Biological Time in "The Voices of Time." J.T. Fraser (Ed.), Brazillier, New York, (1966)

Fi 67 Fischer, R. "The Biological Fabric of Time." Ann. N.Y. Acad. Sci., Vol. 138, (1967)

FR 22 Friedmann, A., Z. F. Phys 10 317, (1922), ibid. 21 326, (1924) FRA 66 Fraser, J.T. "The Voices of Time." Brazillier, New York, (1966)

GO 62 Gold T. in "Recent Developments in General Relativity." p. 225-234, Oxford Univ. Press, New York, (1962)

GU 90 Guyav, M. "La genese de l'idee de temps." Alcan, (1990)

HA 62 Hamner, K.C. et al. "The Biological Clock at the South Pole." Nature, Vol. 195, p. 476, (1962)

HA 75 Hawking, S.W. "Particle Creation by Black Holes." Comm. Math. Phys. Vol. 43, p. 199, (1975)

HA 85 Hawking, S.W. "Arrow of Time in Cosmology." Phys. Rev. D Vol. 32, p. 2489, (1985)

HH 83 Hartle, J.B. and Hawking, S.W., Phys. Rev. D, Vol. 28, p. 2960, (1983)

HO 33 Hoagland, H. "The Physiological Control of Judgements of Duration. Evidence for a Chemical Clock." J. Gen. Physch., Vol. 9, p. 267, (1933)

HO 39 Hadidian, C. and Hoagland, H. "Chemical Pacemakers." J. Gen Physch., vol. 23, p. 31, (1939)

HEI 1 Heidegger, M. "Sein und Zeit." Jahrbuch f. Phaenemonologie v. phaenemenologische Forschong Bd. VIII E. Husserl (Ed.) (1927)

KA 29 Kant, Jm. "Critique of Pure Reason." N.K. Smith (transl.), McMillan, London, (1929)

KO-TU 90 Kolb, E.W. and Turner, M.S. "The Early Universe." Ch. 9, Adison-Wesley Publ. (1990)

KR 56 Kramer, G. Orientierte Zugaktivität gekäfigter Vögel Naturwiss. 7 (1956)

LEi 08 Leibniz "The Philosophical Works." G.M. Duncan (transl.), New Haven, (1908)

LOK 94 Locke, J. "An Essay Concerning Human Understanding." A.C. Fraser (Ed.), Clarendon, Oxford, (1984)

ME 66 Meerloo, J.M. The Time Sense in Psychiatry in "The Voices of Time." J.T. Fraser (Ed.), New York, (1966)

Mi 56 Miller, G. "The Magical Number 7 ± 2, Some Limits for our Capacity for Processing Information." Physcol. Rev., Vol. 63, p. 81, (1956)

NEW 47 Newton, Isaac "Mathematical Principles of Natural Philosophy." A. Mott (transl.), Berkeley Univ. Of Calif. Press, (1947) OR 69 Ornstein, R. "On the Experience of Time." Penguin Books, Inc. New York, (1969)

PA 85 Page, D. "Will the Entropy Decrease if the Universe Recollapses?" Phys. Rev. D, Vol. 32, p. 2496, (1985)

PE 89 Penrose, R. "The Emperors New Mind." Oxford Univ. Press, (1989)

PE-Wi 65 Ap. Journ. 142 419, (1965)

PE 79 Penrose, R., Singularities and Time Asymmetry in: General Relativity: an Einstein Centenary, Hawking, S.W., Israel W. (Ed.), Cambridge Univ. Press, (1979)

PP 59 Peterson, L.R. and Peterson, M.Y. "Short Term Retention of Individual Verbal Items." J. Exp. Phsychol., Vol. 58, p. 193, (1959)

PR 80 Prigagine, Ilya "From Being to Becoming, Time and Complexity in the Physical Sciences." Freedman, San Francisco, (1980)

RO 42 Ross, W.D. "The Students Oxford Aristotle." Oxford Univ. Press, (1942)

Vi 86 Vilenkin, A. "Boundary Conditions in Quantum Cosmology." Phys. Rev. D, Vol. 33, p. 3560, (1986)

Vi 88 Vilenkin, A., Phys. Rev. D, Vol. 37, p. 888, (1988)

Wi 48 Wiener, Norbert, "Cybernetics." Wiley, New York, (1948)

ZE 89 Zeh, H.D. " The Physical Basis of the Direction of Time." Springer, (1989)

BOOKS ON THE SUBJECT OF TIME:

(1) Schlegel, R. Time and the Physical World
Michigan Univ. Press (1961)

(2) Whitrow, G.T. The natural Philosophy of Time
London (1961)

(3) Fraser, J.T. (ed.) The Voices of Time
New York (1966)

(4) Gold, T. (ed.) The Nature of Time
Ithaca, New York (1967)

(5) Grünbaum, A. Geometry and Chronometry
Minneapolis (1968)

(6) Zeman, J. (ed.) Time in Science and Philosophy
Elsevier Amsterdam (1972)

(7) Heidegger, M. On Time and Being
Harper & Row Publ. (1972)

(8) Fraser, J.T. et al. (ed.) The Study of Time
Springer Verl. New York (1972)

(9) Eranus Yearbooks 3, Man and Time
Bollinger Series XXX · 3 Princeton Univ. Press (1973)

(10) Ornstein, R.E. On the Experience of Time
Pelican Books (1975)

(11) von Müller, A.C. Zeit und Logik
Wolfgang Bauer Verl, Bamberg (1983)

(12) Horwich, P. Asymmetries in Time
MIT Press (1987)

(13) Davies, P.C.W. The Physics of Time Asymmetry
Surrey Univ. Press (1974)

(14) Zeh, H.D. The Physical Basis of the Direction of Time
Springer Verl. (1989)

(15) Hawking, S. W. A Brief History of Time: From the Big Bang to Black Holes
Bantom Books (1988)

DISTRIBUTION LIST

1.	Library (Code 52) Naval Postgraduate School Monterey, CA 93943-5000	2
2.	Dean of Research (Code 08) Naval Postgraduate School Monterey, CA 93943-5000	1
3.	Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
4.	Superintendent (Code 00) Naval Postgraduate School Monterey, CA 93943-5000	1
5.	Provost (Code 01) Naval Postgraduate School Monterey, CA 93943-5000	1
6.	Dean of Faculty and Graduate Studies (Code 07) Naval Postgraduate School Monterey, CA 93943-5000	1
7.	Chairman, Department of Physics (Code Ph/Wh) Naval Postgraduate School Monterey, CA 93943-5000	10