THE MYSTERIES OF THE ATOM

By

H. A. WILSON, M.A., M.Sc., F.R.S.
Professor of Physics, Rice Institute.
Fellow of the American Philosophical Society,
formerly Fellow of Trinity College, Cambridge

LONDON
CHAPMAN & HALL, LTD.
11 HENRIETTA STREET, W.C. 2

1934

Printed in U. S. A.

Press of
GEO. S. FERGUSON CO.
Philadelphia, Pa.

CONTENTS

APPENDICES

CHAPTER I

Matter and Electricity in the Nineteenth Century

The modern chemical theory of atoms and molecules dates from early in the nineteenth century, when it was proposed by John Dalton of Manchester, England. According to this theory the different sorts of matter may be divided into two classes, namely, compounds and elements. Hydrogen, oxygen, copper, and iron are familiar examples of elements, while water, a compound of hydrogen and oxygen, and rust, a compound of iron and oxygen, are well known compounds.

A quantity of any particular element, iron for example, according to the atomic theory consists of an immense number of minute particles or atoms which are all exactly equal. Atoms may be compared with standard machine screws or other parts which are all made so nearly alike that they are interchangeable, but the atoms of any particular element are supposed to be exactly, not merely approximately, alike. If two or more atoms stick together so as to form a more or less stable particle, it is called a molecule. For example, hydrogen gas at ordinary temperatures is found to consist of molecules each consisting of two hydrogen atoms. At very high temperatures these hydrogen molecules split up into single atoms, and the hydrogen is then called atomic hydrogen, to distinguish it from ordinary or molecular hydrogen.

If two atoms of different elements combine or stick together, we get a molecule of a compound. For example, two hydrogen atoms and one oxygen atom combine together to form a molecule of water. Thus water consists of exactly equal molecules, each consisting of one oxygen atom and two hydrogen atoms firmly stuck together. It is found that water contains

11.19 per cent by weight of hydrogen and 88.81 per cent of oxygen. This shows that an oxygen atom is $88.81 \times \frac{2}{11.19}$ or 15.88 times as heavy as an atom of hydrogen. In the same way two atoms of oxygen combine with one of sulphur forming a molecule of sulphur dioxide. Sulphur dioxide is a bad-smelling gas quite different from either oxygen or sulphur, just as water is quite different from oxygen or hydrogen. Sulphur dioxide is found to contain 50 per cent by weight of sulphur and 50 per cent of oxygen. Since it contains twice as many atoms of oxygen as of sulphur, it follows that an atom of sulphur is just twice as heavy as one of oxygen. These two examples show how chemists have been able to determine the relative weights of the atoms of all the elements by finding the percentage composition of their compounds with each other.

The following table gives the chemical atomic weights of several important elements, found in this way. It is customary to take the chemical atomic weight of oxygen to be sixteen, so that for example, that of sulphur is 32, since an atom of sulphur is twice as heavy as one of oxygen, and the atomic weight of hydrogen is $\frac{16}{15.88} = 1.0078$, since an atom of oxygen is 15.88 times as heavy as one of hydrogen.

In the nineteenth century atoms were commonly thought of as minute hard particles which were supposed to excite a field of force in the space around them. Just as, for example, a magnet attracts a piece of iron so that the iron sticks to the magnet, so two atoms were supposed to attract each other and stick together forming a molecule. A magnet attracts a piece of iron even when they are some distance apart and it is supposed that the magnet produces or excites a field of force, or a magnetic field, around it. The nature of such fields of force is unknown but the effects they produce can be observed. We see that the magnet attracts the iron and we say that the mag-

CHEMICAL ATOMIC WEIGHTS
TAKING THAT OF OXYGEN TO BE 16

Name of Element	*Physical Properties*	*Atomic Weight*
Hydrogen	Lightest known gas	1.0078
Helium	Very light gas	4.0022
Beryllium	Light metal	9.02
Carbon	Solid found in various forms, *e.g.*, graphite and diamond	12.0036
Nitrogen	Gas. Air contains about 80% nitrogen	14.008
Neon	Rare gas found in air	20.18
Sodium	Metal; very easily tarnished	23.00
Aluminum	Light metal	26.97
Chlorine	Yellow gas	35.457
Sulphur	Yellow solid	32.065
Iron	Metal	55.84
Copper	Metal	63.55
Zinc	Metal	65.38
Silver	Metal	107.88
Tungsten	Metal used for lamp filaments	184.1
Platinum	Metal	195.2
Gold	Metal	197.21
Lead	Metal	207.22
Radium	Radioactive metal	225.97
Uranium	Radioactive metal	238.18

net produces a field of force which we suppose is there whether the iron is put in it or not. The different properties of the atoms of different elements were thought of as due to differences between the fields of force and differences between the shapes of the atoms. For example, an atom of oxygen readily combines with two atoms of hydrogen forming a water molecule, but an atom of carbon combines with four atoms of hydrogen forming a molecule of the gas methane. An atom of carbon was therefore sometimes thought of as having four sides like a tetrahedron, and each side was supposed to be able to hold an atom of hydrogen, while an atom of oxygen was thought of as a flat disk with two sides. It was supposed that in solid and liquid bodies the atoms were packed closely together so that most of the space in a solid was occupied by the hard impenetrable particles or atoms. In gases the atoms were supposed to be separated and moving about freely.

Rough estimates of the size of atoms were made during the nineteenth century. For example, Lord Rayleigh found that a cubic inch of oil, when put on water, spread out into a film which covered about one hundred thousand square yards of surface. The film was therefore only about one one hundred millionth of an inch thick. It was supposed that the film consisted of a single layer of oil molecules so that the diameter of the molecules was supposed to be about one one hundred millionth of an inch. This rough estimate has turned out to be about right. According to this, one hundred million molecules placed side by side would form a row one inch long. In the same way one hundred million shot, each one tenth of an inch in diameter, would form a row about sixteen miles long.

In a gas like air the molecules occupy only about one thousandth of the volume so that when one thousand cubic feet of air are condensed, about one cubic foot of liquid air is obtained.

The atoms were thought of as hard and solid, and it was supposed that they were the smallest possible particles. They were indestructible and permanent so that the total amount of each element in the universe remained constant. The atoms were supposed to be the ultimate units out of which the material universe was built up. As we shall see, many of these nineteenth century ideas about atoms have proved to be erroneous.

During the first half of the nineteenth century the researches of Faraday and others placed the science of electricity on a firm basis. The development of electrical theory was largely independent of that of chemical theory and it was not until near the end of the century that it was realized that matter and electricity are practically identical.

An electric current in a copper wire was thought of as a flow of the electric fluid through the wire. A conductor like copper was supposed to contain equal quantities of positive and negative electric fluids which flowed in opposite directions when a current was passing through the wire. The electricity was not regarded as a constituent of the copper atoms, but as present in the copper along with the atoms. The atoms were not supposed to move along with the current. It was known that when a current flows from one metal into another, say from copper into iron, that no copper is carried into the iron and no iron into the copper, so that it was clear that the atoms did not move along with the current. Very different results were obtained when a current was passed through a solution of a salt in water.

Suppose, for example, that two pieces of sheet copper *A* and *B* are dipped into a solution of copper sulphate in a glass jar. Suppose an electric current is passed through the solution from *A* to *B*. Then it is found that the plate *A* slowly dissolves in the solution while copper is deposited on the plate *B*. The amount of copper which dissolves at *A* is equal to the amount deposited on *B*. Thus it is clear that the copper atoms flow through the solution with the current. The amount of copper deposited is found to be exactly proportional to the amount of electricity passed through from *A* to *B*. Other metallic salts behave in the same way as copper sulphate. For example, silver can be deposited from a solution of silver nitrate.

Faraday passed the same current through solutions of different salts and compared the weights of the different bodies deposited. He found that for a number of elements the amounts deposited were accurately proportional to the atomic weights. For example, the atomic weight of silver is one hundred and seven times that of hydrogen and Faraday found that the same current deposited one hundred and seven times as much silver as hydrogen. This shows that equal numbers of silver and hydrogen atoms are deposited and therefore that the amount of electricity required per atom is the same for silver as for hydrogen.

It is supposed that in a salt solution each salt molecule splits up into a metallic atom, which is charged with positive electricity, and the rest of the salt molecule which is negatively charged. When a current is passed through the solution the current is carried by the charged metallic atoms, which move with the current, and the negatively charged particles which move in the opposite direction. Such charged atoms or molecules are called *ions*. Faraday's results show that the ions of hydrogen, silver, sodium, and many other elements, all have exactly equal positive electric charges. He also found many different sorts of ions which carry negative charges which are equal though of opposite sign to the charges on hydrogen ions. Faraday also found that some elements, for example, oxygen, copper, and calcium, form ions which carry just twice as much electricity as hydrogen ions. Oxygen carries negative electricity, and copper and calcium positive. Other elements give ions which carry just three times as much electricity as hydrogen ions. Thus if e stands for the amount of electricity carried by one hydrogen atom or ion in a solution, then the charge carried by any other ion is e or $2e$ or $3e$ or $-e$ or $-2e$ or $-3e$. It appears therefore that the charge e is a sort of natural unit of electricity and that the charge carried by an atom in a solution is always an exact multiple of this unit. This naturally suggests that electricity is made up of equal particles or atoms of electricity which are of two different sorts, namely particles with the charge $+e$, and particles with the charge $-e$.

This idea of an atomic constitution of the two kinds of electricity, though clearly suggested by Faraday's results, was not adopted until many years later. Electricity was regarded as some sort of imponderable and non-material entity and it was thought that the charges carried by ions in solutions were determined in some unknown way by the nature of the atoms. Thus it was supposed that hydrogen and silver ions carry equal positive charge because the hydrogen and silver atoms have equal powers of attracting positive electricity.

A metal like copper was supposed to consist of copper atoms which were thought of as hard particles packed closely together. These particles were supposed to contain equal quantities of positive and negative electric fluids which could flow about freely in the copper from one atom to another. The copper was therefore a good conductor of electricity. An element like sulphur, which does not conduct electricity was also supposed to consist of hard particles or atoms which contained equal quantities of the two electric fluids but the electric fluids were supposed not to be able to flow from one atom to another. Chlorine and copper combine to form copper chloride, and it was supposed that when a chlorine atom and a copper atom came in contact then some electricity flowed from the chlorine to the copper so that the copper atom became positively charged and the chlorine atom negatively charged. Positive electricity attracts negative electricity so that the oppositely charged copper and chlorine atoms became firmly stuck together forming a molecule of copper chloride. When the copper chloride is dissolved in water it was supposed that the two atoms separated although retaining their charges so that in the solution there are positively charged copper ions and negatively charged chlorine ions. Thus chemical combination was supposed to be largely due to electrical charges on the atoms but the atoms were not always charged and the charges were not regarded as an essential part of the atoms.

The lightest atom is the hydrogen atom of atomic weight 1.0078 and it was supposed that, since atoms were indivisible, no lighter particle than a hydrogen atom could exist. Electricity was regarded more as a quality than as a material entity. A charged particle excited an electric field in the space around it and an uncharged particle did not. The difference between a charged particle and an uncharged particle was thought of as due to the presence of the field around the charged particle and not as due to the presence of something material on it which excited the field.

The above is a brief outline of the ideas about electricity and matter current about 1890. The universe was a collection of hard particles, the atoms, which could be electrified. These particles were indestructible and indivisible. They excited fields of force in the space around them and were supposed to move under the action of forces in the same way as large masses of matter were observed to move. The laws of motion of large masses were known and the atoms were supposed to obey the same laws.

Some scientists about 1890 thought that all the important facts about matter and electricity had been discovered and that nothing remained for physicists to do but to make more and more accurate measurements of important quantities, the values of which were already approximately known. As a matter of fact, however, in 1890 the stage was set for a series of scientific performances more surprising and revolutionary than could be imagined. The atoms were to be dissected into many parts and shown to be neither permanent nor indestructible. Electricity was to be shown to consist of particles out of which the atoms of all the supposed elements are built up. The most firmly established theories, like Newton's law of gravitation and the wave theory of light were to prove inadequate and new theories quite different from anything known before were to appear. The very conception of a particle in a definite position at a definite time was to be tried and found wanting.

CHAPTER II

ELECTRONS

The discovery of the electron first showed that the physics and chemistry of the nineteenth century were inadequate. The electron is a particle nearly two thousand times lighter than a hydrogen atom and electrons are contained in every kind of matter. Atoms contain electrons and so are not merely indivisible hard particles. The existence of electrons was established about 1897 by J. J. Thomson in the Cavendish Laboratory at Cambridge, England. Wiechert and Kaufmann in Germany at the same time also carried out experiments which indicated the existence of electrons. The theory of electrons was rapidly developed by J. J. Thomson, H. A. Lorentz, and many others.

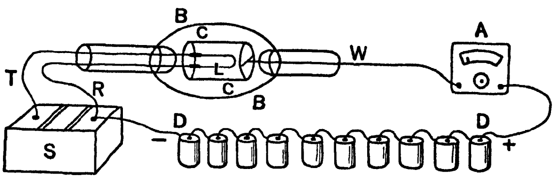

FIG. 1

We shall not consider here all the experiments which first showed the existence of electrons as a universal constituent of matter but shall describe more recent experiments which show the same thing more plainly.

Negative electricity is emitted by metals such as platinum or tungsten, when heated in a vacuum. This phenomenon, first investigated thoroughly by O. W. Richardson, may be studied with the apparatus shown in Fig. 1.

BB is a glass bulb from which practically all the air has been pumped out so that there is an almost perfect vacuum inside it. A wire W is sealed through the glass as shown and supports a thin hollow metal cylinder CC. Two wires T and R are sealed into the other end of the bulb and support a tungsten wire loop L inside the metal cylinder so as not to touch it. The wires T and R are connected to a battery S which sends a current through the wire loop L and heats it to a high temperature so that it becomes incandescent like the filament of an ordinary electric light bulb. One of the wires R is also connected to the negative end of another battery DD of small dry cells. The positive end of this battery is connected to an ammeter A and through the ammeter to the wire W and cylinder CC. It is then found that the ammeter indicates a current flowing from the cylinder CC to the wire loop L and round the circuit containing the battery DD and ammeter A. If the heating current is stopped by disconnecting the wire T from the battery S so that the loop cools down then the current through the ammeter stops. Also if the battery DD is reversed so that its negative end is connected to the cylinder then the ammeter shows no current even when the loop is hot. This experiment shows that the wire loop, when it is hot, emits negative electricity, but not positive electricity.

When the cylinder is connected to the positive end of the battery DD it is positively charged and attracts the negative electricity emitted by the loop, but when the cylinder is negatively charged it repels the negative electricity which is therefore prevented from getting across from the loop to the cylinder. The current obtained with the cylinder positive increases rapidly as the temperature of the loop is raised. It appears that negative electricity escapes from a hot metal in much the same sort of way that vapor escapes from a hot liquid. The negative electricity evaporates from the hot metal just as a hot liquid evaporates. The negative electricity is not accompanied by any material from the hot metal; it is just pure negative electricity.

Fig. 2 shows an apparatus by means of which the negative electricity emitted by a hot metal may be studied.

It consists of a conical shaped glass bulb from which the air is pumped out. L is a small tungsten wire loop connected to two wires R and T sealed through the glass. The loop can be heated by passing a current through it as in the previous apparatus. This loop is enclosed in a small glass tube into which a narrow metal tube T is sealed. This tube T is connected to a wire W sealed through the bulb wall. If the loop is heated and the tube T charged positively, by connecting it to a battery of about three hundred dry cells, then some of the negative electricity which is emitted by the hot loop passes through the tube T, and forms a narrow jet or stream which hits the glass bulb at E. The inside of the glass near E is coated with a thin layer of finely powdered zinc sulphide which emits a greenish light when the negative electricity strikes it.

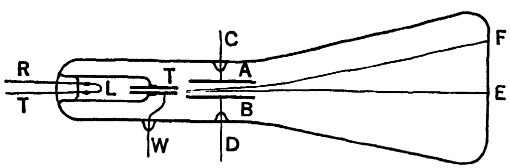

FIG. 2

A small bright spot of greenish light at E is therefore produced by the narrow jet of electricity from the loop. The stream of negative electricity passes between two parallel metal plates A and B connected to two wires C and D. If the plates are connected to a battery so that A is positively charged and B negatively charged, the stream of negative electricity is deflected towards A and the spot of light moves from E to F. The positive charge on A attracts the stream and the negative charge on B repels it as was to be expected.

The stream of negative electricity can also be deflected by bringing a magnet near the bulb. If we use a horse shoe shaped magnet and place it so that one pole is in front of the bulb and opposite the middle of the plates A and B and the other behind the bulb, then the spot at E is deflected up when the South pole is in front, and down when the North pole is in front.

J. J. Thomson assumed that the negative electricity consisted of minute particles all equal and each having a definite charge and also a definite weight, and he was able to calculate the ratio of the charge to the weight or the charge per unit weight from the observed electric and magnetic deflections of the spot of light.

The details of this calculation are given in Appendix 1. It was found that the charge per unit weight was very much larger than any previously known value.

In experiments like those of Faraday on the conductivity of solutions the charge carried by a known weight of hydrogen ions in a solution had been determined, so that the charge per unit weight of hydrogen atoms was known. J. J. Thomson found that for the negative electricity emitted by the hot loop the charge per unit weight was nearly two thousand times greater than for the hydrogen atoms in solutions. The fact that the spot of light was all deflected to the same extent and so not spread out when deflected shows that all the particles have the same charge per unit weight. Faraday's results which showed that ions in solutions always have charges which are exact multiples of a definite charge e suggested that electricity is made up of equal particles. This idea was adopted by J. J. Thomson and he therefore supposed that the particles of negative electricity emitted by the hot loop have each the charge $-e$, equal but of opposite sign to that on a hydrogen atom in a solution. But since the charge per unit weight of the particles of negative electricity is two thousand times as great as for hydrogen atoms it follows that if both have numerically equal charges then the weight of a particle of negative electricity must be two thousand times smaller than that of a hydrogen atom. These particles of negative electricity are called electrons. The assumption made by J. J. Thomson that the stream of negative electricity consists of particles each having the electric charge $-e$ has proved to be correct. These electrons can be obtained from any kind of matter, for example, as we shall see later on, they can be knocked out of the atoms of any gas by means of X-rays. It is clear therefore that atoms are not indivisible hard particles, for they contain very much smaller and lighter particles, the electrons.

If electricity is really made up of equal particles then any quantity of electricity must be an exact multiple of the charge on one particle, or $+ ne$ where n denotes a whole number. If then a method of comparing extremely small electric charges could be devised it ought to be possible to test the theory in a very direct way. Suppose for example, that a large number, say ten thousand, small charges were examined and that it was found that about five thousand were all equal and that about two thousand five hundred of them were all equal but double those in the set of five thousand and that another two thousand five hundred were all equal but three times greater than those in the first set. This would clearly indicate that the electricity was made up of equal units, each equal to the charge of those in the set of five thousand.

An experiment which showed such relations among a large number of very small electric charges was performed by the writer in 1903. An ordinary fog or mist consists of minute drops of water which fall only very slowly through the air. Such a cloud of droplets can be easily obtained artificially in a closed vessel containing moist air by a sudden increase of volume or expansion. The expansion cools the air and the moisture present condenses into droplets. The writer produced a cloud of water droplets which were very slightly charged with negative electricity by passing X-rays through the air. These droplets could be seen to fall slowly and they all fell at the same rate, showing that they were all equal in weight because heavy drops fall faster than lighter ones. But if the charged cloud was produced between two metal plates one above the other, and these plates were connected to a battery so that the lower plate was charged positively and the upper one negatively, then the droplets did not all fall at the same rate. The negatively charged drops were attracted by the lower positively charged plate and repelled by the upper plate so that they fell faster than before and the drops with bigger charges fell faster than those with smaller charges. Many thousands of the charged droplets could be seen falling and it was found that charging the plates caused them to divide into a number of sets. The most numerous set fell slowest but all in it fell at the same rate, showing that they all had equal charges. The next most numerous set all fell at the same rate but faster than the first set, showing that they all had equal charges larger than those of the first set. A third set was also observed which fell still faster. By measuring the rates of fall the charges on the different sets could be compared and it was found that they were nearly proportional to the numbers one, two, and three. This experiment showed clearly the atomic nature of electric charges. Faraday's results on ions in solutions showed that the average charges on the ions were $\pm$ e or $\pm$ 2e or $\pm$ 3e but did not prove that the individual charges on all the ions of one kind were equal. It was natural to assume that they were equal, but the above experiment with the charged water droplets showed that the charges on the drops in each set were all equal since all the drops in each set fell at the same rate.

The actual average value of the atomic charge was measured rather roughly by Townsend and by J. J. Thomson in 1897 and 1898 and by the writer in 1903. Millikan several years later succeeded in measuring very small individual charges accurately. He measured the charges on small droplets of oil slowly falling in air between two metal plates as in the writer's experiment just described. The use of oil drops instead of water drops enabled accurate measurements to be made, because the oil drops do not evaporate and so remain of constant size for a long time. Millikan charged the upper plate positively and so made the drops fall less rapidly and was able to actually balance a drop so that it did not fall at all. He found that the charges were always exact multiples of the smallest charge observed and found this smallest charge to be 4.77×10^{-10} electrostatic units of either positive or negative electricity. The theory of the measurement is given in Appendix 2.

That the charge on the electron is equal to this smallest possible negative charge was shown by J. J. Thomson. When ultraviolet light falls on a metal like zinc it causes electrons to escape from the zinc. In the apparatus shown in Fig. 2 the hot wire may be replaced by a small zinc plate. If this plate is illuminated by ultraviolet light from an arc lamp it emits a stream of electrons which behave exactly like those emitted by the hot wire. J. J. Thomson measured the charges on a cloud of droplets which were charged by means of electrons emitted from a zinc plate and found it equal to the smallest negative charge. This smallest negative charge is therefore called the electronic charge.

The electronic charge or atom of negative electricity is very small. The amount of electricity which goes through a 40 Watt 110 volt lamp in one second is equal to 2,300,000,000,000,000,000 or 2.3×10^{18} times the electronic charge.

The charges carried by charged atoms or molecules in gases, that is by gaseous ions, have been shown to be equal to the charge on hydrogen ions in solutions. For example, air at a very high temperature can be made conducting by the presence of a small amount of sodium or other alkali metal salt vapor. The salt vapor dissociates into gaseous ions just as it does in a solution. The writer found that the amount of electricity carried by the sodium vapor was equal to that which the same amount of sodium carries in a solution. This showed that gaseous sodium ions at a very high temperature carry the same charge as sodium ions in a solution as was to be expected.

The ratio of the charge to the weight for the ions in solutions has been determined by measuring the weight deposited by passing a known quantity of electricity through the solution. Knowing this ratio and also the charge on one ion we can calculate the weight of one ion. The ratio of the charge e to the weight in grams m, of a hydrogen ion is about 300,000,000,000,000. This may be conveniently written 3×10^{14}, 10^{14}

meaning the number one followed by fourteen zeros. We have then $e/m = 3 \times 10^{14}$ and therefore $m = e/3 \times 10^{14}$. The charge e is about $\frac{5}{10,000,000,000}$ or $5/10^{10}$ so that the weight m is equal to $\frac{5}{10^{10} \times 3 \times 10^{14}}$ or $1.66/10^{24}$ gram. The weight of an electron is therefore about $1/10^{27}$ gram.

A silver atom weighs one hundred and seven times more than a hydrogen atom or $1.78/10^{22}$ gram. A cubic inch of silver weighs one hundred and sixty-five grams so that there are about 10^{24} atoms of silver in one cubic inch. This shows that one hundred million silver atoms placed side by side would form a row one inch long. This result agrees with that of Lord Rayleigh on the size of oil molecules, mentioned in the previous chapter.

Very recently Anderson of the California Institute of Technology has discovered a new kind of particle which appears to be quite similar to an electron but to have a positive instead of a negative electric charge. These new particles are called positrons. They are emitted by lead when cosmic rays pass through the lead and are also produced by the action of gamma rays on aluminum. The positron is believed to have a weight equal to that of an electron. Apparently the high energy photons of the cosmic or gamma rays collide with an atom and are converted into two particles, an electron and a positron. These new results are evidently of fundamental importance but their precise significance has not yet been determined.

CHAPTER III

PROTONS

All atoms contain electrons so that an electrically neutral or uncharged atom must contain an amount of positive electricity exactly equal to the negative electricity on the electrons in it. If a neutral atom loses an electron it is left with a charge $+ e$ and if it gains an electron then its charge is $- e$.

Since negative electricity consists of equal particles or electrons it is natural to expect a similar constitution for positive electricity. It is found in fact that positive electricity consists of equal particles with charge $+ e$ exactly equal to the electronic charge. These positive particles are called protons; they are found to be very much heavier than electrons. One proton is equal in weight to one thousand eight hundred and forty-five electrons. A hydrogen atom is found to be just one proton and one electron and all other atoms consist of nothing but protons and electrons. According to this matter is nothing but electricity. Electrically neutral or uncharged matter is just a mixture of equal numbers of protons and electrons.

An ordinary gas like air consists of molecules moving about freely and colliding with each other and with the walls of any vessel in which the gas may be contained. Such gases do not conduct electricity; they are almost perfect insulators. This shows that the gas molecules do not carry electric charges. They must be electrically neutral for if they were charged the gas would be a conductor just as a salt solution conducts because it contains charged ions. If two parallel metal plates are put in a gas and connected to a battery of dry cells so that one plate is charged positively and the other negatively no current flows through the gas from one plate to the other. If the gas molecules were charged then those with negative charges would be attracted to the positive plate and those with positive charges to the negative plate, so that we should get a flow of positive electricity one way and of negative electricity the other way, or a current through the gas. It is clear, therefore, that ordinary gas molecules are exactly electrically neutral or uncharged.

But these molecules contain electrons which each have a charge $-e$. If then a gas molecule contains, say n electrons and so a total negative charge $-ne$, it must, when electrically neutral, also contain a positive charge exactly equal to $+ne$ so that its total charge may be exactly zero. The mere fact that exactly neutral molecules exist is therefore sufficient to show that positive electricity is made up of equal units, each of amount $+e$.

Gases, however, can be made to conduct electricity in several ways. The gases in a flame conduct. At the high temperature, about 3600° F., of a flame, some of the molecules lose electrons so that in the flame we have free electrons and molecules with charge $+e$ besides neutral molecules, and the flame conducts electricity. If X-rays are passed through a gas electrons are knocked out of some of the molecules with great velocity. These electrons in turn knock out many more electrons from other molecules and so the gas becomes a conductor while the X-rays are passing through it. Gases at low pressures conduct, provided a battery of a sufficient number of cells is used to produce the current. The gas insulates perfectly for a small number of cells, but breaks down and conducts with a larger number.

Fig. 3 shows a so-called vacuum tube with which the conductivity of a gas at a low pressure may be studied.

The tube contains a metal disc A, supported by a wire C, sealed through the glass at one end of the tube, and another disc B, with a small hole at its center, supported by a wire D, sealed through the glass as shown. The air is nearly all pumped out of the tube, but about one part in one hundred thousand is left in, so that there is a fairly good, but not perfect, vacuum in the tube. If the wires C and D are connected to a very large battery of say, twenty thousand dry cells, or to an induction coil so that C is strongly charged with positive electricity and D with negative, then a current passes through the tube. It is found that electrons are emitted from the disc B and pass along the tube towards A. When they strike the glass they cause it to emit light so that the tube glows. The electrons can be deflected by bringing a magnet near the tube between A and B.

Also it is found that a streak of faint glow appears extending from the hole in B to the end of the tube at E. Where this streak strikes the glass at E it causes the glass to emit light so that there is a bright spot on the glass at E. It is found that this streak is due to positively charged molecules moving

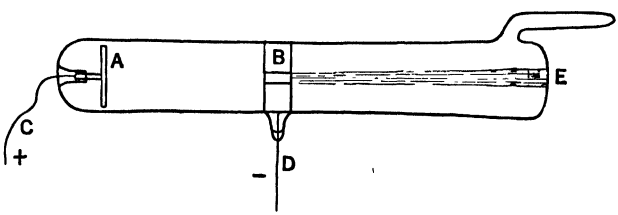

FIG. 3

through the hole in B with great velocity. The electrons emitted by B and moving towards the disc A which attracts them strongly, knock electrons out of the gas molecules with which they collide so that a great many positively charged gas molecules are formed in the tube between A and B. These positively charged molecules are strongly attracted by the negatively charged disc B and rush towards it. Some of them pass through the small hole and travel to E along straight lines. This stream of positively charged molecules is called a stream of positive rays.

Such positive rays can be obtained from any kind of gas and they have been investigated by J. J. Thomson and F. W. Aston with very interesting and important results. They can be deflected by electric and magnetic fields just as electrons can be deflected, but they are much less deflected than electrons and in the opposite direction. J. J. Thomson measured the deflections and so determined the charge per unit weight for these rays, just as he had previously done for electrons. It was found that the charge per unit weight for positive rays in hydrogen gas was equal to the charge per unit weight for hydrogen ions in a solution. This shows that hydrogen positive rays are hydrogen atoms with charge $+e$ or hydrogen atoms which have lost one electron.

J. J. Thomson also got positive rays in hydrogen for which the charge per unit weight was just one-half that for hydrogen ions in a solution. These rays were evidently hydrogen molecules carrying the charge e and consisting of two atoms.

With mercury vapor in the vacuum tube seven different sorts of positive rays could be detected. The atomic weight of mercury is 200 so that a mercury atom with charge e would have charge per unit weight two hundred times smaller than a hydrogen atom. J. J. Thomson found mercury positive rays with charge per unit weight 200, 100, 66.6, 50, 40, 33.3 and 28.6 times smaller than for hydrogen atoms with charge e. These were clearly mercury atoms with positive charges e, $2e$, $3e$, $4e$, $5e$, $6e$ and $7e$. Similar results were obtained with other elements. With neon gas the atomic weight of which is 20.2, J. J. Thomson obtained two sorts of positive rays having charges per unit weight twenty and twenty-two times smaller than for hydrogen. This was interpreted as indicating that neon is a mixture of two sorts of atoms having atomic weights 20 and 22 respectively. The atomic weight 20.2 is therefore merely the average value for all the atoms present. This was the first case in which it was found that all the atoms of an element are not of equal weight. As we shall see later the same thing is true for many other elements.

Following J. J. Thomson's researches, F. W. Aston devised an improved method of studying positive rays with which he has obtained results of far-reaching importance. In Aston's apparatus the rays are first deflected downwards by an electric field and then upwards by a magnetic field. The rays then fall on a long narrow photographic plate. The electric and magnetic fields are so designed that the position on the plate where the rays strike depends only on the weight per unit charge of the rays. On developing the plate a row of small spots is found on it, one spot for each sort of rays present.

The weight of an atom is proportional to its atomic weight and so may be taken equal to its atomic weight if we take as our unit weight that of an atom of atomic weight unity. In the same way, we may express the charge on the atom in terms of the charge e as unit. The weight per unit charge for positive rays which have the charge e is then just equal to the atomic weight. For rays with charge $2e$ it is just one-half the atomic weight and for rays with $3e$, one-third of the atomic weight.

With oxygen gas in his apparatus Aston got three spots on the photographic plate which he attributed to oxygen atoms with charge e, and so weight per unit charge 16, and oxygen atoms with charge 2e, and so weight per unit charge 8, and to oxygen molecules each consisting of two atoms with total charge e, and so weight per unit charge 32. In the same way, with nitrogen, the atomic weight of which is 14, he got spots corresponding to weights per unit charge of 7, 14, and 28. The distances of these spots from one end of the plate could be exactly measured, and was found to increase in a regular way with the weight per unit charge. By measuring the distance of the spot due to any positive rays from the end of the plate it was therefore possible to determine the weight per unit charge accurately, and so get the atomic weight. In this way the atomic weights can be found to within one part in ten thousand. A more detailed description of Aston's apparatus will be found in Appendix 3.

Aston found, for example, that chlorine gas, the atomic weight of which is 35.45, gave positive rays with weights per unit charge 35, and 37, exactly. This indicates that chlorine is a mixture of two sorts of atoms, having atomic weights 35 and 37, in such proportions that the average atomic weight as found by the chemists is 35.45. Atoms which have exactly similar chemical properties so that they are regarded as atoms of the same element, but nevertheless have different atomic weights, are called isotopes. Thus chlorine is said to have two isotopes with atomic weights 35 and 37. Aston and others have found that nearly all the so-called elements are mixtures of isotopes. For example, xenon, the chemical or average atomic weight of which is 131.3 is found to be a mixture of atoms with atomic weights 128, 129, 130, 131, 132, 134, and 136. In all cases the atomic weights found for positive rays are very nearly whole numbers. This clearly indicates that the atoms of all the elements are built up out of particles of atomic weight unity.

Since hydrogen atoms are nearly of atomic weight unity and give positive rays carrying the charge $+e$, and since, as we have seen, the positive electricity in atoms must be made up of units of amount $+e$, it has been concluded that the particles out of which atoms are built up are hydrogen atoms with charge $+e$ and electrons. The positively charged hydrogen atom is called a proton. For example, an electrically neutral or uncharged atom of mercury of atomic weight 200, is believed to consists of two hundred protons and two hundred electrons. The weight of the electrons is very small compared to that of the protons. In very recent years it has been shown that protons can be knocked out of many different atoms. This will be considered more fully later on.

Throughout the nineteenth century the atoms of the ninety-two known elements were regarded as the ultimate indivisible particles of the material universe. Also, all the atoms of any one element were supposed to be exactly alike, and the atomic weights determined by the chemists were believed to be the actual relative weights of the atoms. The epoch making discoveries of J. J. Thomson and Aston have shown that the atoms are not the ultimate particles, but that all the different sorts of atoms are built up out of only two kinds of particles,* protons and electrons, which are nothing but positive and negative electricity. Also all the atoms of any one element are not exactly alike, but have different atomic weights which are always nearly whole numbers, so that chemical atomic weights are merely average values of no fundamental significance. Matter therefore is just electricity and nothing else, since all negative electricity is electrons, and all positive, protons.

It may be asked if protons and electrons are to be regarded as material particles charged with electricity. The answer is that this idea is not justified by the facts. The operation of charging a body with negative electricity consists of adding electrons to it, and a body is charged positively by removing electrons from it so as to leave an excess of protons in it. Thus we cannot suppose that an electron is charged negatively because adding an electron to an electron would give two electrons. Electrons and protons are just atoms of electricity, and so far as is known at present, are indivisible. We only know of electricity in the form of electrons and protons, so that it is meaningless to speak of these indivisible particles as if they consisted of two parts, electricity and matter. The properties of electricity are the properties of electrons and protons. Two electrons or two protons repel each other and an electron and a proton attract each other, but there is no reason to suppose that these properties would remain if the electrons and protons could be divided into still smaller particles, or that they can be attributed to the parts of electrons and protons. There is as much sense in supposing that a part of a proton has the properties of a proton as in supposing that a part of a mercury atom has the properties of mercury. The properties of matter are the properties of electrons and protons arranged in small groups called atoms. (The different properties of different atoms are due to differences in the number and arrangement of the electrons and protons in them.)

* The recently discovered neutron is probably a very close combination of a proton and an electron, and the positive electron or positron, also recently discovered, has only a very transitory existence and so is not a constituent of matter.

CHAPTER IV

The Structure of Atoms

Radium and other radioactive elements emit rays called alpha-rays which are just like positive rays. These rays are found to have a weight per unit charge of value 2, and to carry a positive charge $2e$, so that they have atomic weight 4. The gas helium, which is used in the United States for filling balloons, has atomic weight 4, and it is found that the rays from radioactive elements are atoms of helium which have lost two electrons. An electrically neutral helium atom consists of four protons and four electrons, so an alpha-ray consists of four protons and two electrons.

The alpha-rays are shot out from radium with the enormous velocity of about ten thousand miles a second. They travel through air in nearly straight lines and knock electrons out of the air molecules with which they collide. They therefore leave a track of electrons and positively charged molecules behind them.

If a microscopic particle of radium is supported in a closed vessel containing moist air, and if the air is cooled by a sudden expansion, the moisture condenses on the electrons and charged molecules forming small drops which can be easily seen in a bright light. In this way the track of an alpha-ray can be made visible. This method of rendering such tracks visible was discovered by C. T. R. Wilson in the Cavendish Laboratory at Cambridge in England. A description of his apparatus is given in Appendix 4.

Fig. 4 shows a sketch of the alpha-ray tracks from a particle of radium obtained in this way. The alpha-rays are gradually slowed down by their collisions with the air molecules so that they go only a few inches before they are stopped.

It will be seen that they all go about the same distance, which shows that they are all shot out from the radium at the same speed. These photographs seem to show very clearly that alpha-rays are small particles.

Very important information about atoms can be deduced from these alpha-ray tracks. A proton is nearly two thousand times heavier than an electron so an alpha-ray is about eight thousand times heavier. A collision between an alpha-ray and

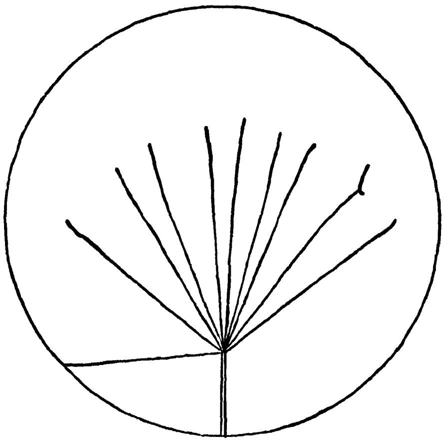

FIG. 4

an electron is therefore like one between a motor car weighing four thousand pounds and a stone weighing half a pound. The stone is knocked out of the way but the car goes straight on. A collision between an alpha-ray and a proton is like one between the car and a rock weighing one thousand pounds. In this case the car will be very seriously affected and will not go straight on.

The alpha-rays give practically straight tracks in air so that it is clear that they do not collide with protons, but only with electrons. They knock electrons but not protons out of the air atoms. This shows that practically all the space inside the atoms contains no protons, but only electrons, for if the electrons and protons were uniformly distributed throughout the atoms, then the alpha-rays would be as likely to knock out protons as electrons, since there are as many protons in the atoms as electrons. When a very large number of alpha-ray tracks are examined a very few are found which suddenly turn through a large angle. There is one such track in Fig. 4. Such large deflections of the alpha-rays indicate a collision with a heavy particle.

When alpha-rays are passed through very thin sheets of metal like gold leaf, most of them go nearly straight through, but a very few are deflected or scattered through large angles. This scattering was carefully studied by Lord Rutherford, and he showed that it could be explained by supposing that the gold atoms each contain a very small heavy particle strongly charged with positive electricity. This particle is called the nucleus of the atom and it contains all the protons. The charge on the nucleus can be estimated from the way in which the alpha-rays are scattered. The positively charged nucleus repels the positively charged rays so that a ray is deflected as it passes by a nucleus, and the nearer the ray goes to the nucleus, the more it is deflected.

Consider an alpha-ray going through a thin plate of thickness t inches in a direction perpendicular to the plate. Let the number of atoms per cubic inch in the plate be n. Then if we imagine a small sphere of radius r inches drawn around the alpha-ray, it is easy to see that as the ray goes through the plate it will come within a distance r of all those atomic nuclei which are inside a cylinder of radius r and length t. The volume of this cylinder is $\pi r^2 t$ so that the alpha-ray will come within a distance r of $n\pi r^2 t$ nuclei as it goes through the plate.

There are about 10^{24} atoms, and therefore nuclei, in a cubic inch of gold so that if $t = \frac{1}{100,000}$ inch we have

$n\pi r^2 t = \pi r^2 \times 10^{19}$. Suppose it were found that one alpha-ray in ten million is deflected through an angle greater than a right angle in passing through the plate then we could say that $10^7 \times \pi r^2 \times 10^{19} = 1$ so that $r^2 = \frac{1}{\pi \times 10^{26}}$ which gives $r = 5.6/10^{14}$ inch. Therefore in order for the ray to be deflected through a right angle it has to get a distance of $5.6/10^{14}$ inch from the nucleus. Thus by finding the fraction of the rays which are deflected through different angles it is possible to find the angles through which the rays are deflected when they get to different distances from a nucleus. This enables the force which the nucleus exerts on an alpha-ray at different distances to be calculated and so gives the electric charge on the nucleus.

In this way Rutherford found that the charge on the nucleus of a gold atom is equal to seventy-nine times the charge of one proton. The atomic weight of gold is 197, so that an atom of gold contains one hundred and ninety-seven protons and one hundred and ninety-seven electrons. It appears therefore that the gold nucleus contains one hundred and eighteen electrons so that its total charge is 197 — 118 or seventy-nine times the charge *e*. There are therefore seventy-nine electrons outside the nucleus. The electrons outside the nucleus are attracted by it and are supposed to form a sort of little cloud of electrons moving about around it. The fact that alpha-rays are so seldom deflected through appreciable angles shows that the nucleus is very small. Rutherford estimated its diameter to be about one million times smaller than that of the atom. Since the nucleus contains one hundred and ninety-seven protons and one hundred and eighteen electrons it is clear that electrons and protons are exceedingly small particles. They are so small that practically all the space in a heavy solid body like gold is empty, or rather contains nothing but the electric and magnetic fields which the electrons and protons excite in the space around them.

It will be seen that we have now left the nineteenth century idea of an atom very far behind. In a solid like gold, instead of hard indivisible particles packed closely together, we have mostly empty space with excessively minute, heavy, positively charged nuclei arranged at intervals of one one hundred millionth of an inch. Between the nuclei minute electrons move about. When alpha-rays are shot through a very thin sheet of gold they nearly all miss the nuclei and are not appreciably deflected by the electrons.

It is found that all atoms are similar to gold atoms. The chemical and physical properties of atoms are found to depend on the number of electrons outside the nucleus. All atoms with the same number outside have the same properties. We should therefore expect to find atoms with one, two, three, four, etc., electrons outside the nucleus. It is natural to expect the number outside to increase with the total number. The total number of electrons in an atom is equal to its atomic weight so if we make a list of all the elements in the order of their atomic weights and number them one, two, three, four, etc., in order, we might expect the number assigned to each atom to be equal to the number of its electrons outside the nucleus, and this is found to be the case. For example, gold is number seventy-nine in the list and has seventy-nine outside electrons. The number of an element in the list is called its atomic number. Hydrogen is number one. Its atomic weight is one, and its nucleus is just one proton, and it has just one electron outside. Helium is number two. Its atomic weight is 4, so that it has four protons in the nucleus and two outside electrons. Oxygen is number eight. It atomic weight is sixteen, so it has a nucleus of sixteen protons and eight electrons with eight electrons outside. The element with the largest atomic number is uranium. Its atomic weight is 238 and its number 92, so its nucleus contains 238 protons and 238 — 92 or 146 electrons and it has 92 electrons outside its nucleus.

The chemical properties of the elements are related in a remarkable way to their atomic numbers. The elements fluorine, chlorine, bromine, and iodine, have very similar properties. They all give negatively charged ions in solutions, and very readily combine with metals like sodium. Their atomic numbers are 9, 17, 35, and 53.

The elements neon, argon, krypton, and xenon, the gases which are now used so much in electrical signs, all have no chemical properties. They do not form compounds with other elements. Their atomic numbers are 10, 18, 36, and 54, which are greater by unity than the numbers of flourine, chlorine, bromine, and iodine.

The elements sodium, potassium, rubidium, and caesium, have atomic numbers 11, 19, 37, and 55. They are all very similar metals which give positively charged ions in solutions, and very readily combine with elements like chlorine.

The explanation of such facts was first suggested by J. J. Thomson. He supposes that the atoms with 10, 18, 36, and 54 electrons outside the nucleus have very stable arrangements of their electrons. The electrons form a complete group around the nucleus, and these atoms do not easily either gain or lose an electron. The elements with one outside electron less, easily take on one more electron, so forming a stable group, and the elements with one more outside electron easily lose an electron. Thus an atom with one more and another with one less readily combine, because the atom with one more loses its extra electron to the atom with one less. Both atoms then have stable groups, but the one which has lost an electron is positively charged, and the other which has gained an electron is negatively charged, so that the two atoms attract each other and form a stable molecule. For example, consider potassium 19, argon 18, and chlorine 17. The eighteen electrons of argon form a stable group and argon does not form compounds. A potassium atom with 19 and a chlorine atom with 17, form two stable groups of 18, the potassium atom losing an electron to the chlorine atom. The potassium then has a charge $+e$ and the chlorine atom $-e$, so they attract each other and stick together as a molecule of potassium chloride. J. J. Thomson showed that the properties of many other elements could be explained in this way.

It was mentioned in the first chapter that an oxygen atom combines with two hydrogen atoms, whereas a carbon atom combines with four. The atomic number of oxygen is 8, so that it requires the addition of two electrons to an oxygen atom to form the stable group of ten electrons. Two hydrogen atoms each give up one electron to an oxygen atom, so forming the stable group. The oxygen atom then has a charge $-2e$ and the hydrogen atoms have each $+e$, so that the three stick together forming a molecule of water.

The atomic number of carbon is 6, so that a carbon atom has six electrons outside its nucleus and it requires four more electrons to make a stable group of ten. Four hydrogen atoms therefore combine with one carbon atom, just as two combine with oxygen.

It appears that the chemical properties of atoms depend on the number of electrons outside the nucleus, which is equal to the atomic number. Atoms may have the same atomic number, but different atomic weights. For example chlorine, as we have seen, consists of a mixture of atoms with atomic weights 35 and 37, but all having the same atomic number 17. The atoms of weight 35 have a nucleus containing eighteen electrons and thirty-five protons, and those of weight 37 have one containing thirty-seven protons and twenty electrons. Both sorts have seventeen electrons outside the nucleus, and so have identical chemical properties and are regarded as atoms of the same element. Thus Rutherford's nucleus theory of atoms enables us to understand the existence of isotopes or atoms with different weights but identical properties.

When alpha-rays are passed through a thin plate of an element of rather small atomic weight, for example aluminum. it is found that the plate emits rays which travel much farther through air than the alpha-rays. These rays are found to have a weight per unit charge of value 1, and are therefore believed to be hydrogen atoms with charge $+e$, that is to say, just protons. Protons can be knocked out of many elements in this way. This result due to Lord Rutherford confirms the theory that atoms are built up of protons and electrons.

P. M. S. Blackett has taken photographs of about a million alpha-ray tracks in nitrogen and other gases. Among these million photographs he found about a dozen which show a proton knocked out of an atom by an alpha-ray. Tracings from some of these remarkable photographs are shown in Fig. 5. The proton makes a much thinner track than the alpha-ray.

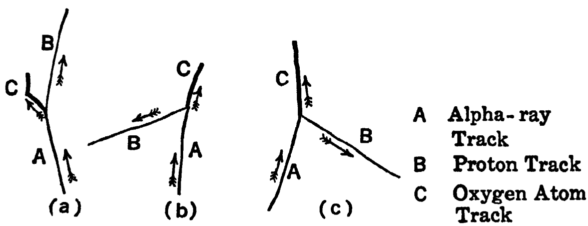

FIG. 5

The alpha-ray hits the nitrogen atom so hard that the atom gets a high velocity and makes a short, thick track. The unexpected result that appears is that the alpha-ray is absorbed by the nitrogen atom and disappears. Nitrogen has atomic weight of 14 and atomic number 7, so that its nucleus contains fourteen protons and seven electrons. The addition of an alpha-ray and the removal of a proton must change the nucleus to one containing seventeen protons and nine electrons. The new nucleus therefore belongs to an atom of atomic weight 17 and atomic number 17 — 9 or 8, so it is an isotope of oxygen. The alpha-ray therefore changes nitrogen into oxygen. This is the first known artificial transformation of one element into another.

When an atom of a radioactive element emits an alpha-ray it changes into an atom of another element. For example, radium has atomic weight 226 and atomic number 88. It is a metal rather like silver in appearance. The nuclei of radium atoms are unstable and about one in two thousand of them emit alpha-rays in a year. An alpha-ray consists of four protons and two electrons, so that when a radium atom emits an alpha-ray its atomic weight is changed to 222, and its atomic number to 86. The new element formed is a gas having no chemical properties, like the gases helium, neon, argon, krypton, and xenon. It is called radon. The alpha-ray is a helium atom, so that radium may be said to spontaneously decompose into helium and radon. The rate of decomposition of radium is in no way affected by high or low temperatures, or any other known agencies. Pure radium decomposes at the same rate as any compound of radium, for example radium chloride or sulphate.

This spontaneous decomposition of radioactive elements is remarkable, because the chance of a particular atom decomposing is independent of its age. The atoms disappear just as the population of a country would disappear if there were no births and a death rate independent of age. The number of deaths on any day is proportional to the population on that day. For example, if half the atoms remain after one hour, then one-quarter will be left after two hours, one-eighth after three hours, one-sixteenth after four hours, and so on. We do not know what decides which particular atoms shall decompose, so we usually say that it is merely a matter of chance. All the atoms appear to be exactly alike and do not change in any way until they suddenly decompose. Some decompose at once and some last for thousands of years, for no obvious reason so far as we can see.

CHAPTER V

LIGHT AND PHOTONS

Up to about the beginning of the nineteenth century light was thought of as minute particles shot out of luminous bodies with enormous velocities. In the nineteenth century this particle theory was replaced by a wave theory according to which light consists of waves traveling through a medium of some sort, just as waves travel over the surface of the sea. In recent years, however, the wave theory has run into difficulties and it now appears that there are particles in light as well as waves, and moreover, the particles seem to be much more real than the waves.

We shall begin with the wave theory, and then take up the new facts which require a return to the particles. When waves travel through a medium the medium does not move along with the waves; it moves backwards and forwards as the waves are passing, but is left in its original position when they have gone by. For example, when a stone is dropped into a pond waves move out over the surface of the water in concentric circles. If a small body floating on the water is carefully watched as the waves pass by it, the body will be seen to move up and down and backwards and forwards but it will be left where it was after the waves have gone by.

If a hundred men were standing in a row, and if the first man knocked down the second man, who got up and knocked down the third man, who got up and knocked down the fourth man, and so on along the row, a wave of falling down and getting up could be said to have run along the line of men.

A wave motion requires a practically continuous medium so that when any part of the medium is disturbed the disturbance is propagated through the medium by the action between adjacent parts.

A vibrating tuning fork sets up a series of waves in the air around it, which produces the sensation of sound in the ear. In the same way the electrons in a luminous body are thought of as vibrating and sending out light waves which travel through the surrounding space. A tuning fork or a bell in a vacuum cannot produce any sound, but light waves travel through a vacuum. For example, the light from the filament of an ordinary electric light bulb gets out through the vacuum in the bulb.

During the nineteenth century it was generally supposed that a so-called vacuum was not really empty, but that it contained a medium called the ether, through which the light waves were propagated. This ether was thought of as filling up all space, including that between the atoms in solid bodies, and was supposed to be some sort of elastic fluid. This idea of an ether filling all space was not really justified. Empty space was conceived as having only geometrical properties, which depended merely on the relations between distances measured in it. When it was realized that light waves travel through empty space and also that electric and magnetic forces act across it, it was assumed that there must be a medium present having the observed physical properties as well as geometrical properties. It is now realized that this imaginary separation of empty space into two parts, one having the geometrical, and the other the physical properties, is meaningless and without justification. There is no imaginable way of separating the geometrical and physical properties; in fact, the geometrical properties are really physical properties. If the physical properties could be removed, the geometrical properties would probably also disappear. Space as we know it has these geometrical and physical properties which cannot be separated, and so we now merely regard them as properties of empty space and do not introduce the idea of an ether. Light waves then, may be thought of as waves traveling through empty space. As we have seen, the electrons and protons in material bodies occupy only an infinitesimal fraction of the total volume, so that when light travels through matter it is still going through practically empty space. According to this view, so-called empty space is not nothing. It has properties, and so must be something. Descartes said that if everything inside a closed vessel were removed, the sides of the vessels would be in contact; there would not even be space left in it. But space is not material. Matter consists of electrons and protons, and there are none of these in empty space.

The idea that light is of the nature of waves of some sort is suggested by experimental results which can be explained on

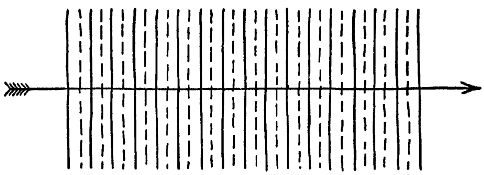

FIG. 6

the wave theory. The waves themselves are not visible. We observe effects produced by light but we do not observe the waves. We can see water waves moving over the sea, but we cannot see light waves.

A series of waves may be represented by the parallel lines in Fig. 6. The continuous lines represent crests, and the dotted lines, troughs, of the waves. The waves are supposed to be moving along in the direction of the arrow. The distance from one crest to the next is called the wave length. The number of waves which go past any point in one second is called frequency of the waves. The distance a wave goes in one second is called the wave velocity. It is easy to see that the wave length multiplied by the frequency is equal to the wave velocity, for a length of the series of waves equal to the velocity passes any point in one second and this length contains a number of waves equal to the frequency.

If light from a small source such as an automobile headlight bulb is passed through a small hole in a metal plate PP we get a narrow beam or ray of light as shown in Fig. 7.

Such a ray of light in air is straight, so that light is said to

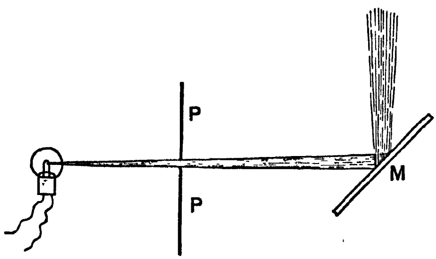

FIG. 7

travel in straight lines. If the ray falls on a plane mirror it is reflected so that the reflected and incident parts are equally inclined to the mirror. The properties of such rays of light

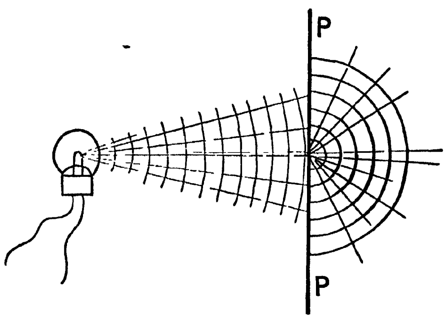

FIG. 8

suggested the particle theory. The particles were supposed to bounce off the mirror.

Newton pointed out that if light consisted of waves, we should expect the waves to spread out in all directions from the hole in the plate PP as shown in Fig. 8, and not to form a straight ray. It is found that the light passing through the hole does spread out slightly and that this spreading increases as the hole is made smaller, so that with an exceedingly small hole no ray is obtained, but the light spreads out in all directions.

The formation of the straight ray when the hole is not too small may be explained on the wave theory by regarding the large hole as made up of a large number of very small holes and supposing that the waves from each small hole spread out in all directions, just as from a single very small hole. Along the straight ray the wavelets from all the small holes have traveled equal distances, so that all the crests and troughs coincide, or as we may say, all the wavelets are in step so that they combine together and form an intense ray of light. But along any other direction the wavelets from the small holes, that is, from different parts of the large hole, have traveled different distances and so are not in step. The wavelets therefore do not combine to give strong light but neutralize each other so that there is no appreciable light except along the straight ray. Thus the propagation of rays along straight lines can be explained on the wave theory, and also the fact that a very narrow ray spreads out in all directions, which would not be expected on the particle theory.

We see that the idea that light travels in straight lines is only an approximation. It is near enough to the truth for large scale phenomena where all the distances involved are large compared with the wave length of the light, but it fails completely when this is not the case.

Fig. 9 shows two equal sets of waves moving in slightly different directions, indicated by the arrows. Points where two crests are superposed are marked +, and points where two troughs come together, —. Points where a trough and a crest are superposed are marked 0. We see that the + and — points lie on parallel lines, and that the 0 points lie in between.

Where two crests come together there will be a crest twice as high as the crests in either of the two sets, and where two troughs come together, a trough twice as deep. Thus along the + and — lines the two sets of waves will unite to give waves twice as great as the waves in either set. When a crest and a trough come together they will cancel each other and there will be no wave motion. Thus along the lines marked 0 there will be no waves. The two sets of waves are said to interfere and destroy each other along these lines. Such interference between two sets of waves can be seen when two stones

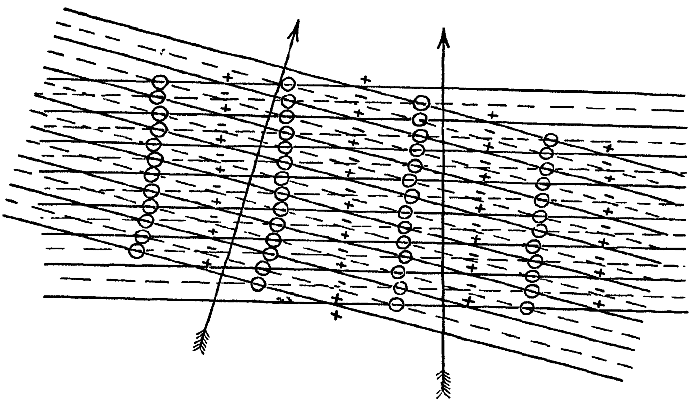

Fig. 9

are dropped into a pond at the same time and rather near together, so that the two sets of waves overlap.

Such interference can be obtained with light. That is to say, it is possible to illuminate a white screen with two beans of light and get bright bands on the screen, with dark lines in between. On the dark bands the two beams destroy each other. An experiment of this kind was first done by Thomas Young near the beginning of the nineteenth century, and this experiment and other similar ones became the foundation of the wave theory of light.

Young's experiment is shown in Fig. 10. AB is a thin metal plate with a short, very narrow slit S in it. CD is another similar plate with two narrow slits TT quite close together. EF is a white screen. The first slit S is strongly illuminated by focusing sunlight on it with a lens L. Also a piece of red glass is put over the slit S so that only red light gets through. Some of the red light from the first slit falls on the two slits TT, and goes through them onto the white screen EF. This screen is thus illuminated by two beams of red light. If either of the two slits TT is covered up there is a uniform patch of red light on the screen. With both slits open the patch on the screen is not uniform, but consists of narrow bright bands with dark bands in between. This remarkable experimental result can be explained by supposing that each

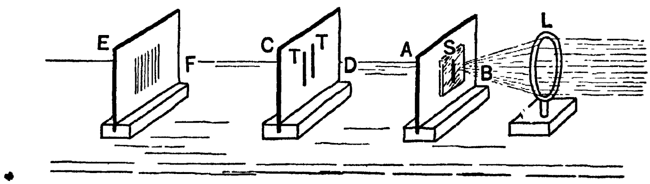

FIG. 10

of the two slits TT sends out a train of waves which interfere at the screen in the way explained above. Similar results can be obtained with light of any color. The distance between the bright bands on the screen is different with light of different colors. It is nearly twice as great with red light as with violet light. If the distance from the pair of slits to the white screen is six feet and if the two slits TT are one-twentieth of an inch apart, then the distance between the bright bands on the screen with red light is about one twenty-fifth of an inch. If we assume that the bright and dark bands are due to interference between two sets of waves we can calculate the wave length. This calculation is explained in Appendix 5. It is found in this way that the wave length of red light is about three one-hundred-thousandths of an inch, and that of violet light about half as great. The wave lengths of yellow, green, and blue light come in between those of red and violet light. The frequency of light is obtained by dividing the velocity by the wave length. It is very large because the velocity is big and the wave length small.

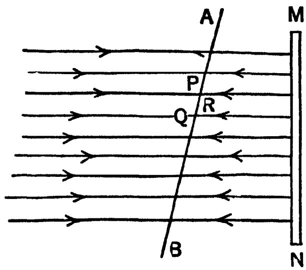

FIG. 11

During the nineteenth century a great variety of experiments on interference of light waves was devised, and the wave lengths of all sorts of light were very exactly measured. It was found that light always behaves exactly in accordance with the wave theory in such experiments, and this theory came to

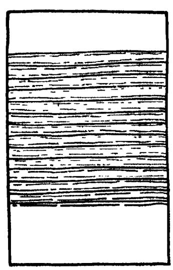

FIG. 12

be regarded as so well established that it was believed to be really true.

Only one more experiment on interference will be described here, a very interesting one due to Wiener. His experiment is shown in Fig. 11.

A beam of light was reflected straight back along its path by a flat mirror MN. A thin photographic film AB was placed in the light so as to be nearly, but not quite, parallel to the mirror. On developing the film he found it covered with equidistant parallel bands as shown in Fig. 12. This can be explained by interference between the two beams of light going through the film.

The curved line in Fig. 13 may be supposed to represent a series of waves moving along in the direction of the arrow. As the waves move along past any fixed point P, there is an up and down motion at P. If we consider another point Q at a distance from P equal to the wave length, it is easy to see that the up and down motion at Q will be exactly the same as

FIG. 13

at P. At a point R half way between P and Q, the up and down motion will be opposite to that at P or Q. Thus we see that the wave motions at any two points half a wave length apart are in opposite directions.

Consider now a point Q on the film in Wiener's experiment and suppose it is where one of the bands appears when the film is developed. At such a point the two beams of light must reinforce each other and not interfere. Now consider a point P on the film and suppose it is half a wave length nearer the mirror than Q. Moving half a wave length along a train of waves in either direction reverses the motion in the waves, so at P the motions in both beams of light will be opposite to those at Q, and so will still be together and therefore there will be another band at P. At R, half way between P and Q the motions in the two beams will be opposite, so that there will be interference and therefore no light, and the plate will not be affected.

Thus we see that the bands on the film must be arranged so that each one is half a wave length farther from the mirror than the next one. If, for example, the inclination of the film to the mirror is one in one hundred, then the distance between the bands will be equal to fifty times the wave length. Wiener found that the wave lengths found in this way agreed with those measured in other ways.

The wave lengths of ordinary or visible light are between about three one-hundred-thousandths of an inch, and one-and-a-half one-hundred-thousandths. Light of longer or shorter wave length does not affect the human eye, and so is invisible. Radiation of the same nature as visible light can be obtained in various ways with wave lengths greater or smaller than those of visible light. For example, the electric radiation used

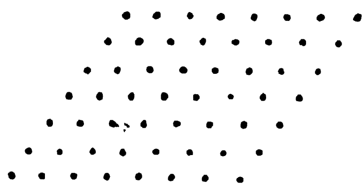

FIG. 14

in radio broadcasting is of the same nature as visible light and travels with the same velocity, but has wave lengths enormously greater. Ultra-violet light and X-rays have much smaller wave lengths than visible light. The wave length of ordinary X-rays is about ten thousand times smaller than that of visible light.

The wave lengths of X-rays are too small to be measured in the same way as ordinary light waves. They have been measured by means of crystals. The atoms in a crystal like Iceland Spar or Diamond are arranged in regular layers at equal distances apart, as shown in Fig. 14.

If a beam of X-rays is allowed to fall on a crystal, each layer of atoms reflects a small fraction of the rays. We thus get a very large number of very weak reflected beams, all going in the same direction, as shown in Fig. 15, in which the first four only of the layers of atoms are represented by the four equally spaced parallel lines AB.

When a large number of equal trains of waves are all moving along the same path they will destroy each other by interference unless they are all in step, that is to say, unless the crests and troughs in all the trains coincide. All the waves in a train are alike so that moving a train of waves back a distance equal to the wave length makes no difference. The waves reflected from any layer have to go a slightly greater distance than those reflected from the next layer above, so that if this extra distance is exactly one wave length all the reflected trains

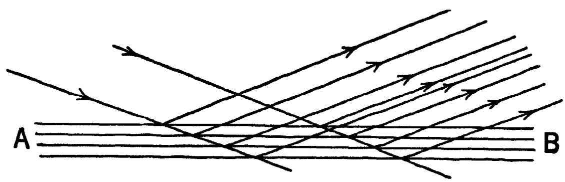

FIG. 15

of waves will be in step, and there will be a strong reflected beam due to the superposition of all the weak beams, one from each layer. There will also be a strong reflected beam if the extra distance is equal to exactly two, three, or more wave lengths.

The extra distance changes as the angle between the incident beam and the surface of the crystal is changed. It is found that if the crystal is slowly turned so as to gradually increase the angle between the incident rays and the surface of the crystal, then no X-rays are reflected except at a series of definite angles which can be easily measured. For the smallest angle the extra distance is one wave length, and for the next smallest, two wave lengths, and so on. The wave length can then be easily calculated from the observed angles and the distance between the layers of atoms in the crystals. The method of calculation is given in Appendix 6. For example, it is found that the strongest X-rays emitted by copper have a wave length of about one-half of one-hundred-millionth of an inch.

The velocity of light in empty space is one hundred and eighty-six thousand miles per second. No greater velocity has ever been observed, and it is believed that the velocity of light is the greatest possible velocity.

It was mentioned in the chapter on electrons that when ultra-violet light falls on a metal like zinc, it causes electrons to be emitted by the metal. This effect is called the photo-electric effect, and it is found that it cannot be explained on

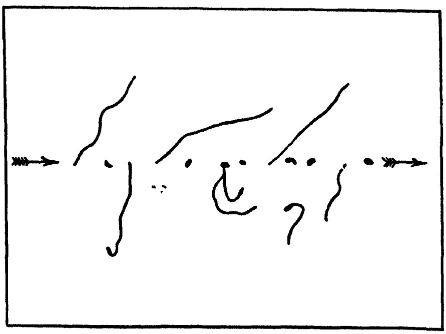

FIG. 16

the wave theory of light. X-rays cause the emission of electrons from atoms of any sort, and the electrons are shot out with very great velocities, so that they can knock other electrons out of atoms and produce tracks in a gas, which can be made visible like the tracks of alpha-rays.

Fig. 16 is a sketch of a photograph of electron tracks produced by passing a narrow beam of X-rays through a moist gas and cooling the gas by a sudden expansion so as to condense moisture on the electrons and positively charged molecules. The direction of the narrow beam of X-rays is indicated by the arrows. It will be seen that the tracks nearly all start from the beam and are all about the same length. The tracks are not straight like alpha-ray tracks because electrons are easily deflected by collisions with the electrons in the air molecules.

It is found that the length of these tracks depends only on the wave length of the X-rays used. X-rays of shorter wave length give longer tracks so that, for example, waves half as long or with twice the frequency give tracks just four times as long. Increasing the intensity of the X-rays increases the number of tracks obtained, but does not change the length of the tracks.

The length of an electron track is found to be proportional to the square of the initial energy of motion of the electron. The greater this energy, the more electrons it can knock out of atoms before it is stopped.

It appears therefore that when X-rays knock electrons out of atoms the number knocked out is proportional to the intensity of the X-rays, but the energy with which they are knocked out is independent of the intensity of the rays and directly proportional to their frequency.

Ultra-violet light does not give electrons enough energy to make visible tracks, but it has been shown by other methods, which need not be considered here, that the energy which ultra-violet light gives to electrons in atoms is proportional to the frequency of the light and independent of its intensity, just as with X-rays.

These results are not what might be expected on the theory that X-rays are waves. Consider waves in the sea acting on a ship with a number of small boats tied to it. The ship and boats may be supposed to be like the nucleus of an atom with its electrons round it. Unless the waves were high enough we should expect that they would not cause the boats to break away from the ship. Very low waves would never break a boat away and the energy given to a boat when broken away would increase with the height of the waves. It would be extremely surprising to find that the energy with which the boats were knocked away was the same with very low waves as with very high ones of the same frequency.

X-rays are produced by allowing a stream of electrons to strike against a metal plate. Fig. 17 shows an X-ray tube used for the production of X-rays.

It consists of a glass bulb from which practically all the air has been pumped out. At S a small spiral of tungsten wire is supported by two wires C and D sealed through the glass. A current is passed through the spiral, from a small battery connected to the wires C and D, so as to heat it to a bright white heat. The spiral then emits electrons. K is a metal block supported by a copper rod sealed through glass at T. Several copper plates are welded to the rod between T and P as shown, so as to expose a large surface to the air with the object of keeping the rod from getting too hot. The

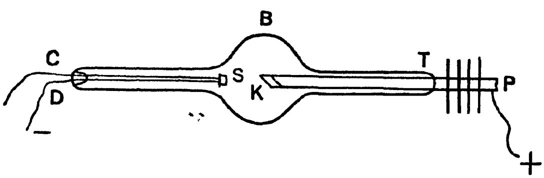

Fig. 17

copper rod and metal block K are strongly charged positively so that the electrons from the spiral are attracted by the block and rush into it with great velocity. The block then emits X-rays.

It is supposed that the electrons collide with the atoms in the block and give them energy, part of which is then emitted as X-rays. In 1881, long before X-rays were discovered by Roentgen, J. J. Thomson predicted, on theoretical grounds, that collisions between charged particles and a solid body should produce radiation.

The electrons flow along the copper rod and out through the wire attached to it so that a current of electricity flows through the tube. Such an X-ray tube emits X-rays with a long range of wave lengths, but it is found that no X-rays are produced with wave lengths shorter than a definite value.

The frequency of these shortest waves produced is found to be exactly proportional to the energy per electron of the electrons striking the block. It is independent of the number of electrons striking the block. The intensity of the X-rays, as might be expected, is proportional to the number of electrons striking the block.

X-rays are more penetrating, the shorter their wave length, so that the waves of shortest wave length from an X-ray tube can be easily separated from the rest by passing the rays through suitable absorbing screens.

If the X-rays of shortest wave length are allowed to fall on a metal plate they cause it to emit electrons, and it is found that these electrons have energy of motion or kinetic energy just equal to that of the electrons striking the block in the X-ray tube. This very remarkable fact cannot be explained on the wave theory.

On the wave theory it is supposed that an electron in the tube hits an atom in the block and causes it to emit waves. The waves which escape from the block spread out in all directions so that it is clear that only a very minute fraction of the energy in the waves can possibly fall on any atom in the metal plate. Yet atoms in the plate emit electrons having as much energy as the original electron in the tube. We might suppose that the atom in the plate absorbed energy from the X-rays until it had got enough to emit the electron, but a simple calculation (See Appendix 7.) shows that this would take many hours, while it is found that the plate begins to emit electrons at once when the X-rays fall on it.

It is just as though the waves produced by throwing a stone into a pond could spread out in all directions and travel a long way to the shore, and there throw out another stone with as much energy as was used in throwing the first one in.

It is perfectly clear that the facts of photoelectricity are incompatible with the wave theory of light.

In order to explain these facts, Einstein therefore proposed a particle theory of light and X-rays. According to this theory light consists of minute particles called photons, which always move with the velocity of light, and which have kinetic energy proportional to the "frequency" of the light as determined by interference experiments on the wave theory. Of course, if light is not waves, but particles, then the idea of the frequency of light loses its meaning, but Einstein supposed that the frequency given by the interference experiments must have some significance and that the energy of the photons can be supposed proportional to it whatever it really is. Photons are very different from electrons and protons. They are not electricity, and they always move with the enormous velocity of light, while electrons and protons can be supposed to come to rest. Also photons can appear and disappear, or can be created and destroyed. The facts of photoelectricity are quite easily explained on Einstein's photon theory.

A narrow beam of X-rays passing through air, on this theory, is thought of as a stream of photons. If the X-rays are all of the same wave length, or frequency, then the photons all have the same energy. If a photon collides with an electron in an atom it may give all its energy to the electron, and so disappear. Increasing the intensity of the rays merely increases the number of photons, without altering the energy of each one. Thus it is clear that the number of electrons knocked out will be proportional to the intensity of the rays, but the energy with which they are knocked out will be independent of the intensity. The energy per electron will be proportional to the frequency because that of the photons is supposed to be.

In an X-ray tube, when an electron hits the block it gives its energy to an atom which then emits a photon. The energy of the photon may be equal to or less than that of the electron. The photon goes off in some definite direction and if it strikes a metal plate it may knock an electron out, which may get all the energy of the electron in the tube. The difficulty that waves spread out so that all the energy cannot be concentrated on to one electron at a distance, is avoided by the particle theory.

As we have seen, when X-rays are passed through a gas, only a very few of the electrons in the atoms are knocked out. This is difficult to explain on the wave theory, but follows at once from the particle theory. On the wave theory we should expect all the electrons to be affected by the waves because the waves are continuous and pass over all the electrons present. On the particle theory a photon traveling through the gas goes on between the electrons until it hits one. The small number hit can be attributed to the extremely small size of the electrons and photons.

The way in which light acts on a photographic plate is also much more easily explained on the particle theory than on the wave theory. A photographic plate is coated with a thin layer of gelatin containing minute grains of silver bromide. If such a plate is exposed equally all over to very weak light for a short time, and then developed, it is found that only a few of the silver bromide grains are affected. On the particle theory this obviously can be taken to mean that only those grains hit by a photon are affected, but on the wave theory we should expect all the grains to be equally affected, which is not the case.

There are many other important branches of optics in which it is found that the particle theory is much more successful than the wave theory, for example, the theory of the radiation from hot bodies and the theory of the light emitted by gases when excited by an electric discharge, but it is not proposed to discuss these questions here. We shall, however, consider one more important effect which seems to require very definitely the particle theory for its explanation.

When X-rays are passed through elements of small atomic weight, like carbon or sodium, it is found that a small fraction of the rays is scattered in all directions. On the wave theory this was explained by supposing that the electrons in the scatterer were set oscillating by the waves, and so emitted waves in all directions. On this theory the scattered X-rays should have exactly the same frequency as the incident X-rays, because the electrons must oscillate with the frequency of the waves falling on them, and must emit waves of the same frequency. However, it is found that the frequency of the scattered waves is not exactly equal to that of the incident waves, but is slightly less. This effect was carefully studied by A. H. Compton, who showed that it can be explained on the particle theory, and it is known as the Compton effect.

In the photoelectric effect a photon collides with an electron in an atom and gives all its energy to the electron, so that the photon disappears. There is no scattering of the X-rays in the photoelectric effect.

To explain the scattering of X-rays on the particle theory, Compton supposed that it is possible for a photon to collide with an electron and bounce off from it, just as in a collision between two billiard balls. When the photon hits the electron it gives some of its energy to the electron, and so comes off with less than its original energy. But the frequency of a photon is proportional to its energy, so that the frequency of the photon is reduced by the collision. Compton showed that the change of frequency observed agrees exactly with that calculated on this theory. This calculation is given in Appendix 8. The electron gets a small fraction of the energy of the photon, and so is set in motion. In the sketch shown in Fig. 16 of the electron tracks due to a narrow beam of X-rays passing through gas, there are several very short tracks besides the long tracks due to the photoelectric effect. These short tracks are supposed to be due to electrons from which photons have bounced off, so giving only a small fraction of their energy to the electron. The length of these short tracks and their direction, agrees very well with the values calculated on Compton's theory.

It appears, therefore, that there are a large number of phenomena which can only be explained by supposing that light consists of particles or photons, and also that there are a large number of phenomena of the nature of interference, which require a wave theory for their explanation. Thus we have two radically different, and apparently incompatible theories of light.

Whenever light produces an observable effect, for example, when it acts on a photographic plate or knocks an electron out of an atom, it appears to act like particles. In interference experiments it is not the waves which are observed, but the distribution of light intensity. This is done by means of a photographic plate, or in some other equivalent way. The observed distribution of intensity is thus not a distribution of waves but a distribution of photons, or rather a distribution of effects attributed to photons. The photons themselves are not observed any more than the waves are.

Consider Wiener's interference experiment shown in Fig. 11 on the particle theory. On this theory we have photons moving towards the mirror, which bounce off from it and come back. These photons pass through the photographic film as they move towards the mirror, and as they come back. If a photon hits a grain of silver bromide in the film, it is absorbed by the grain and so disappears. It is clear that the plate should be uniformly affected all over and not show the bands which are actually observed. In the bands, however, only some of the grains are affected, which cannot be explained on the wave theory. It is clear that what is required is some sort of combination of the two theories. This will be considered in the next chapter.

CHAPTER VI

Waves and Particles

The state of optical science at which we have arrived in the last chapter was described by an eminent physicist when he said that we think of light as particles on Mondays, Wednesdays, and Fridays, and as waves on Tuesdays, Thursdays, and Saturdays. To this we may add that on Sundays we admit that only God knows what it really is.

So far we have considered three kinds of particles, electrons, protons, and photons. The photons are supposed to be particles, but they have a wave frequency associated with them in some mysterious way.

In 1926 Louis de Broglie suggested that just as light appears to be a combination of waves and particles, so electrons and protons may also have a wave aspect. That is to say, electrons may be the particles in a combination of waves and particles analogous to light.

According to this idea it ought to be possible to get interference with electrons, just as with light. This has proved to be the case, but the wave lengths of electrons are much shorter than those of visible light; they are, in fact, about the same as the wave lengths of X-rays.

The wave character of electrons was discovered experimentally by Davisson and Germer in 1927, and they found that the wave lengths were just equal to those given by de Broglie's theory.

The wave lengths of electrons are so short that it is not possible to perform experiments like Young's or Wiener's experiments with them. The method by which Davisson established the wave aspect of electrons was similar to that used to measure the wave lengths of X-rays, which was described in the previous chapter.

A narrow stream of electrons in a vacuum was allowed to fall on the surface of a crystal of nickel. It was found that the electrons were reflected strongly only when the inclination of the stream of electrons to the surface of the crystal had certain definite values, just as when X-rays are reflected from a crystal.

It is easy to see that if we regard electrons as particles then no possible explanation of this result of Davisson's can be imagined. Obviously a particle could bounce off the crystal at any angle.

By measuring the angles at which the electrons were strongly reflected, and knowing the distance between the layers of atoms in his crystal, Davisson was able to calculate the wave length of the electrons. He found that it depended on the velocity of the electrons, so that doubling the velocity just halved the wave length. That is to say, the wave length was inversely proportional to the velocity.

For electrons moving ten thousand miles per second, which is not very fast for electrons, he found the wave length to be about three-fifths of one one-thousand-millionth of an inch. According to this, the wave length for electrons moving only one foot per second would be about one-tenth of an inch, but there is no known way of experimenting with such slowly moving electrons.

Davission's results on the wave lengths of electrons have been confirmed in an interesting way by G. P. Thomson, the son of J. J. Thomson, who showed that electrons are particles in 1897.

Any metal, for example gold, consists of a great many very small crystals packed closely together in a random manner so that the layers of atoms in the crystals lie in all directions. If a narrow beam of X-rays is passed through a plate of gold the rays fall on the small crystals, and when the angle between the rays and the layers of atoms in a crystal has one of the values for which there is strong reflection, the rays are reflected. There are so many small crystals arranged in all directions, that there are always plenty for which the angles are right for reflection. If the rays passing through the plate are allowed to fall on a photographic plate, the result is that we get a series of concentric rings on the plate when it is developed, as shown in Fig. 18.

The wave length of the X-rays can be calculated from the distances between the layers of atoms in the gold crystals, and the angles through which the rays are deviated by being reflected so as to form the rings, just as when a single large crystal is used.

G. P. Thomson tried this experiment with electrons instead

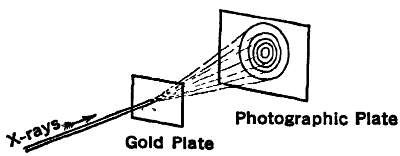

FIG. 18

of X-rays. He had to use high-velocity electrons and an extremely thin gold plate, in order to get the electrons to go through. He got concentric rings on the plate, just as with X-rays, and was able to calculate the wave lengths of the electrons. His results agreed with de Broglie's theory, like those of Davisson.

It has been shown rather recently by Dempster of Chicago that protons can be reflected from a single crystal and their wave length found just as with electrons. The wave lengths of protons are much smaller than those of electrons.

In the apparatus shown in Fig. 4 a narrow stream of electrons travels along the tube until it hits the glass. This stream of particles can be deflected by means of electric or magnetic forces as we have seen. On the wave theory the stream of electrons is to be thought of as a narrow beam or ray of waves, so that we ought to be able to show that the beam of waves will go along the same path as the particles, and be deflected in the same way.

If the width of a narrow beam or ray of waves is not large compared with the wave length of the waves, the ray diverges appreciably. For electrons moving one thousand miles per hour, if the width of the ray was one-hundredth of an inch, it would be more than a million times the wave length, and so the ray would not diverge appreciably. The wave length of alpha-rays is nearly eight thousand times smaller than for electrons, so that a ray only about one millionth of an inch across would not diverge. Thus the alpha-ray tracks obtained in the cloud expansion apparatus do not indicate any less divergence than might be expected for a narrow beam of alpha-ray waves.

When a particle such as an electron is moving in a fixed field of force like an electric field, the velocity of the particle changes in magnitude and also in direction, as it moves along. The field of force may either accelerate or retard the motion. It is supposed that the energy of the particle is made up of two parts, the energy of motion or kinetic energy, and the energy of position, or potential energy. The total energy remains constant. If the kinetic energy increases, then the potential energy decreases by an equal amount.

The potential energy depends only on the position of the particle in the field, so that, for a given total energy, the kinetic energy, and therefore the velocity of the particle, will depend only on the position.

Any curved path can be regarded as made up of a large number of very short straight paths, so if we like we can regard the path of the particle in a field of force in this way. We can also suppose that the velocity of the particle, and so its kinetic energy remains constant along each of the very short straight paths.

At a junction between two of the short paths the direction and magnitude of the velocity change suddenly, and therefore the kinetic and potential energies also change suddenly though their sum remains constant. The curved path of a narrow beam or ray of waves can be regarded as made up of short straight parts in the same way. To show that the ray of waves and the particle will go along the same path, it is therefore sufficient to consider any two of the short straight paths.

In Fig. 19 let EFAB represent a narrow beam or ray of waves traveling with velocity U_1 in the space above a boundary PQ. Let the wave velocity below the boundary be

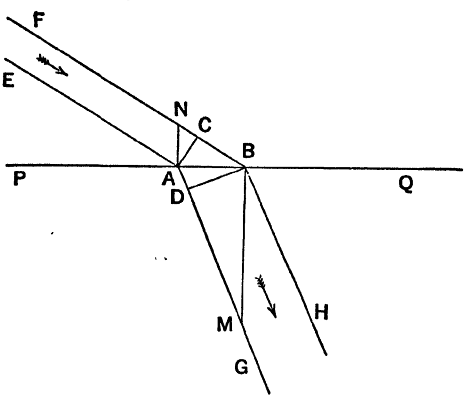

FIG. 19

U_2 so that at the surface PQ, the wave velocity changes suddenly from U_1 to U_2. Let ABGH be the ray below the boundary. Let AC be a wave which has just arrived at the boundary at A, and let DB be the position of this wave when it is just leaving the boundary at B. While the wave goes from C to B along FB, it goes from A to D along AG so that CB is to AD as U_1 is to U_2 or $CB/AD = U_1/U_2$. This equation expresses the way in which the direction of the ray is changed at the boundary.

Now suppose that instead of waves we have a stream of particles moving with velocity V_1 parallel to EA and FB, and suppose that at the boundary PQ their velocity is changed to

V_2, and that they then move parallel to BH and AG. We may suppose that the potential energy of the particles above the surface PQ is greater than that below it, so that V_2 is greater than V_1.

As a particle passes through the boundary its velocity parallel to the boundary will not be changed because the potential energy does not change as we go along the boundary. Draw AN and BM perpendicular to PQ. Then if we suppose that NB represents the velocity V_1, NA will represent the part of it perpendicular to PQ, and AB the part parallel to PQ. In the same way, if AM represents V_2, then BM will represent the part of it perpendicular to PQ, and AB that parallel to PQ. Thus if we take NB to represent V_1 and AM to represent V_2, this makes the two parts parallel to the boundary equal, as they should be, so that we have $V_1/V_2 = NB/AM$. Now $AM/AB = AB/AD$ and $NB/AB = AB/CB$. Hence $AM = AB^2/AD$ and $NB = AB^2/CB$ so that $NB/AM = AD/CB$. But $NB/AM = V_1/V_2$ so that $V_1/V_2 = AD/CB$ or $V_2/V_1 = CB/AD$. This equation expresses the way in which direction of motion of the particle is changed at the boundary.

For the ray of waves we found $U_1/U_2 = CB/AD$. In order that the waves and the particles may go along the same path it is necessary to have CB/AD the same in each case, so that we must have $U_1/U_2 = V_2/V_1$ or $U_1 V_1 = U_2 V_2$. Thus it appears that the product of the wave and particle velocities must remain unchanged when these velocities change.

In the case of photons, as we have seen, the wave and particle velocities are both equal to the velocity of light which we will denote by C. The product of the wave and particle velocities is therefore equal to C^2 for photons. According to de Broglie's theory this product is equal to C^2 for any kind of particle, so that $U V = C^2$ or $U/C = C/V$.

If the wave velocity U is say, one-quarter that of light, this gives $1/4 = C/V$ or $V = 4C$ so that V, the particle velocity, is four times that of light.

In Davisson's experiments the velocity of the electrons was about ten thousand miles per second or about one-twentieth of that of light, so that the wave velocity was about twenty times that of light, or four million miles per second.

At first sight it seems absurd to suppose that the waves associated with a particle move with a much greater velocity than the particle, but de Broglie showed that it is quite possible for this to be the case.

When a stone is dropped into a pond, waves spread out in concentric circles. If the ring of waves is watched closely, it can be seen that the waves in the ring do not move with the same velocity as the ring. They appear at one side of the ring, move across it, and disappear at the other side. The

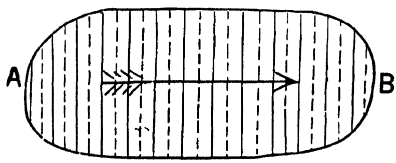

FIG. 20

ring of waves is called a group of waves, and it appears that the velocity of the group is not the same as that of the individual waves. De Broglie suggested that the waves associated with an electron form a group of waves which moves with the electron although the wave velocity is much greater than that of the electron.

Such a group is shown in Fig. 20. The continuous lines are supposed to represent the wave crests, and the dotted lines the troughs. The whole group moves along with the particle velocity v while the waves appear at A, rush across the group with wave velocity u, and disappear at B.

As we have seen, the energy of a photon is proportional to its wave frequency *n*, which is obtained by dividing the velocity of light *C*, by the wave length *l*.

The energy divided by the frequency is therefore equal to a constant which has the same value for all photons. This constant is called Planck's constant because it was first used by him in his theory of heat radiation, and it is usually denoted by h. Thus if E denotes the energy of a photon we have $E/n = h$ or $E = hn$.

De Broglie supposed that the same thing would be true for electrons and other particles. Now, according to Einstein's theory of relativity, the energy of a particle of weight m is equal to m multiplied by the square of the velocity of light, or to mC^2. This is explained in Chapter X. De Broglie therefore supposed that the frequency n of electron waves will be given by the equation $mC^2 = hn$ where m is the weight of the electron.

The wave velocity u is equal to nl, where l is the wave length, so that we have $l = u/n = uh/mC^2$. But $C^2 = uv$ so that $l = h/mv$. The product of the weight and the velocity of a particle is called its momentum, so that it appears that the wave length of the electron waves should be equal to Planck's constant divided by the momentum of the electron. h/m is equal to 1.13 for an electron, so that $l = 1.13/v$. This gives the wave length l in inches when v is expressed in inches per second. If v is ten thousand miles per second or one billion nine hundred million inches per second, we get the wave length to be $\frac{1.13}{1{,}900{,}000{,}000}$ inch or $\frac{0.6}{1{,}000{,}000{,}000}$ inch. This value calculated by means of de Broglie's theory agrees with that found experimentally by Davisson.

The wave particle theory of electrons and protons is analagous to the wave particle theory of light. The distribution of the effects observed can be calculated correctly by a wave theory, but the effects observed are such as might be expected to be produced by particles and not by waves. Neither the waves nor the particles are directly observed.

It appears that we are in exactly the same sort of fix as regards electrons and protons as in the case of light and photons. In a way this improves the situation because it is now clear that this mysterious blend of two apparently incompatible sets of qualities, that, of waves and particles, is universal and not merely peculiar to light.

It was pointed out by Einstein many years ago that physical theories ought to deal with quantities which can actually be observed. Quantities which cannot be observed are always of doubtful character.

Let us consider Young's interference experiment again from this point of view. It is found that if the distance between the two slits TT (Fig. 10) is doubled, then the distance between the bands of light is halved, or the product of the distance between the slits, and that between the bands, is a constant. Also this product is found to be proportional to the distance from the slits to the screen or photographic plate. If, then, we denote the distance between the pair of slits by d, and that between the bands by b, and the distance from the slits to the screen by s, then we may express the results obtained by saying that db/s is a constant. Now d, b, and s, are all directly observed, so there is no doubt about them. If we assume that the light consists of waves which can interfere, then we find that db/s is equal to the wave length of the light. Denoting this supposed wave length by l we have the equation $db/s = l$. Now l is not observed, it is merely calculated, and we do not really know that it is a wave length. The same length l always appears when any kind of so-called interference experiment is performed with the same sort of light, so that it is clear that this length expresses some property of the light. It is, however, quite possible that some other quite different theory of light might be discovered which would give the same equation $db/s = l$ as the wave theory, and on this other theory l would not be a wave length, but something quite different. The wave theory enables the distribution of the light to be calculated correctly; that is, it gives formulae which agree with the facts, but we do not know whether the quantities in the formulae which cannot be observed are really of the nature assumed. That is, we do not know whether the theory is true or not. When we consider the effects produced by the light which can be observed, such as its action on the grains of silver bromide in a photographic plate, we find they are not effects which we should expect waves to produce. This shows that the wave theory is probably not true. The effects observed are such as can be attributed to particles or photons, but the theory that light consists of photons does not offer any explanation of the observed distribution of intensity. The photons, like the waves, are not observed. It seems probable therefore, that light is neither waves nor particles, but something else which somehow has some of the properties of both.

We can say, however, that whatever the nature of light may be, that of electrons and protons is similar.

If we consider two points A and B on a photographic plate which has been exposed, then the number per unit area of grains affected at A is to the number at B, as the light intensity at A is to that at B. On the particle theory this means that the number of photons striking the plate is proportional to the intensity or energy of the light. If the intensity at A is, say, ten times that at B, then ten times as many photons fall on equal areas in equal times at A as at B.

Suppose, however, that the plate is only exposed to very weak light for a very short time such as one millionth of a second. In this case we can imagine that very few photons, only one for example, arrive at the two equal areas at A and B. If only one photon arrives, it must fall either on A or B, so that it is not possible for the number falling at A to be ten times that at B. Thus we see that the number of photons is only proportional to the intensity when large numbers are considered. If we consider a single photon in Young's experiment, then it may fall anywhere on the screen, but we suppose that the chance of its falling on any small area is proportional to the intensity of the waves at that area.

In the same way, if the death rate in a city of population one million is ten per thousand, then we can say that ten thousand people will die in a year in that city, but if we consider a particular person we cannot say whether he will die or not in any year, but we can say that his chance of dying is one in a hundred.

The view as to photons, electrons, and protons at which we have now arrived may be summed up as follows. In any experiment with one of these three sorts of particles, the distribution of the particles may be calculated by assuming a wave theory, and calculating the intensity of the waves. The number of particles which arrive at any place will be proportional to the intensity of the waves at that place, provided the number of particles is large. If the number of particles considered is small, the chance of a particle arriving at the place will be proportional to the intensity. Strictly speaking, we ought not to say the number of particles arriving, but only the number of effects produced which can be attributed to a particle. The particles, like the waves, are hypothetical, and only the effects produced are observed.

CHAPTER VII

The Uncertainty Principle

The corner stone of nineteenth century physics was the idea of the uniformity of nature; that identical causes always produce identical effects. This idea was expressed by Clerk-Maxwell in the following words:

"The difference between one event and another does not depend on the mere difference of the times or the places at which they occur, but only on differences in the nature, configuration, or motion of the bodies concerned."

All matter was supposed to consists of particles which moved according to definite laws so that if the positions and motions of all the particles concerned were known at any time, then the positions and motions at any future time were exactly fixed and could, theoretically at least, be predicted.

We are not now in a position to support any such dogmatic statement as that of Clerk-Maxwell. The idea of the uniformity of nature was based on superficial observations of large scale phenomena, and it now appears that with given initial conditions many future events are possible. At best all we can hope to do is to calculate the relative probabilities or chances of the different possible events. Also Clerk-Maxwell's statement involves the assumption that it is possible to fix the initial configuration and motion of the bodies concerned exactly, and as we shall see this is not the case.

Consider, for example, the decomposition of radium into helium and radon. It is found that about one radium atom in two thousand decomposes in a year. Suppose we are asked to predict the state of a particular radium atom at some future time, say after ten years. All we can say is that it will either be unchanged or will have decomposed, and that the chance of its having decomposed is about one in two hundred. The same sort of uncertainty is involved in all other phenomena involving only small numbers of particles. Moreover there is no absolute certainty even when large numbers of particles are involved. Any particular radium atom may last a thousand years, so if we consider, say a million radium atoms, it is possible, though very improbable, that they will all last a thousand years. That is to say, exceptional events apparently contrary to the laws of nature are not impossible, and so presumably do occasionally happen. Miracles are not impossible.

Let us consider a very simple experiment which will serve

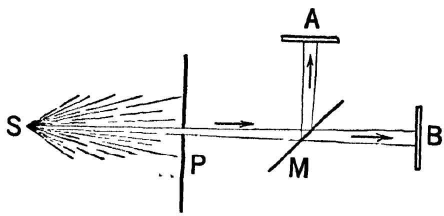

FIG. 21

to illustrate these ideas. In Fig. 21 let S be a small source emitting either photons, electrons, or protons. Let the radiation from it fall on a metal screen P with a small hole in it, so that a narrow beam gets through. Let this beam fall on a very thin metal mirror M, so thin that it reflects only one-half of the radiation.

Let the reflected rays be received on a photographic plate A and the transmitted rays on another plate B. Let us suppose that the hole in the screen can be opened and closed by a shutter, and that the source is extremely weak so that it emits very few particles. Then we may imagine the shutter opened and closed so as to let just one particle through. What can we say as to the fate of this particle?

According to the ideas explained in the last chapter, we are to imagine that the source emits waves, and we are to calculate the intensity of these waves. If the hole is not too small the waves will form a straight narrow ray or beam as shown in the figure, which will be divided by the thin mirror into two beams of equal intensity, which will fall on the photographic plates.

The chance of the particle appearing anywhere is proportional to the intensity of the waves, and so we can say that there is an even chance that it will appear at A or B. The intensity of the waves will be uniform over the area on each plate on which the beam falls, so that the particle will be equally likely to appear at any point on these areas. A grain of silver bromide somewhere on one of these areas will be affected by the particle and will show up when the plates are developed. We cannot predict which plate the particle will appear at, or at what point it will appear.

The fact that a particular grain in one of the plates is affected may seem to justify the assumption of a particle, but the waves seem to be purely imaginary; they merely serve for purposes of calculation.

If, instead of one hole in the plate P, there are two narrow slits close together, we have a Young's interference experiment, and the intensity of the waves will not be uniform on the plates, but will be distributed in narrow bands as we have seen. In this case the particle will be certain to fall in one of the bands, and not in between two bands.

An experiment of this sort was tried by G. I. Taylor. He photographed an interference pattern due to light, and used such weak light that not more than one photon at a time could be supposed to be falling on the plate. After a very long exposure he found on the plate a pattern identical with that obtained with a short exposure and strong light. This showed that the distribution of the photons is given by the wave theory even when they go in one at a time as we have supposed.

We may say then that we have a theory which enables us to calculate correctly the chance that any one of many possible effects will occur. For example in the above experiment the chance, that any one of the grains on the plates will be affected, can be calculated. When a very large number of such effects are produced the theory gives correctly the way they are distributed.

The theory does not give the path along which the particle may be thought of as moving from the source to the silver bromide grain which is affected. For instance, in Young's experiment it does not indicate which of the two slits the particle goes through. It is impossible to make an observation which would locate the particle on its path to the grain without stopping it or diverting it on to an entirely different path. Thus if a photographic plate were put immediately behind the two slits, a grain would be affected opposite one of them, but the photon would be absorbed by the grain. Also, since different photons go to different points on the plates, the path cannot be located by stopping different photons at different places along the narrow beam of light.

The assumption of waves and particles helps us to understand what happens, to some extent, although we have no idea why or how the particles are guided by the waves. The waves and particles are merely imaginary. They are crude models of the underlying reality which they perhaps resemble slightly in some way which we do not understand.

We can calculate what will happen when a very large number of effects are produced, but the symbols, representing wave lengths and other quantities not observed, in our mathematical formulae, probably really represent quantities quite different from those suggested by our crude model.

When we are trying to explain any phenomena the best we can do is to imagine some combination of things with which we are familiar, such as waves and particles, which seems to work in accordance with the facts, and if we are successful, then we have a more or less satisfactory theory. But the underlying reality probably consists of things quite different from anything we are familiar with, and so we cannot hope to get anything better than a crude model. Such models are probably about as much like reality as shadows are like the objects producing them.

The imaginary character of the waves in the wave particle model becomes clearer when the theory of two particles which repel each other, like two electrons, is considered. It is then found necessary to consider a set of waves in an imaginary space of six dimensions. Such waves are certainly imaginary, so that there is little doubt that if there is anything at all like reality in the wave particle model, it is the particles, and not the waves.

In what follows we shall suppose the particles to be thought of as real, and the waves as merely serving as a basis for the calculation of the probable distribution of the particles.

A particle moving along in empty space may be thought of as accompanied by a group of waves. If the group is say six inches long, then since the chance of the particle being anywhere is proportional to the intensity of the waves, we can say that the particle must be somewhere in its group, so that its position is uncertain to the extent of six inches. If the position of the particle were exactly known, the group would have to be supposed to have no appreciable size.

Now a group of waves cannot be supposed to contain only waves of one definite length. In order to form a group it is necessary to suppose that there are a large number of trains of waves present, all moving with the group and having slightly different wave lengths, so that they are all in step at the middle group and so give strong waves there, but are out of step and so interfere and destroy each other at points outside the group. This is explained in Appendix 9. Since there must be a certain range of wave lengths in a group of waves, it follows that the wave length is uncertain to the extent of this range. This uncertainty in the wave length involves a corresponding uncertainty in the wave and particle velocities. It is found that the uncertainty in the position of the particle, which is equal to the group length, multiplied by the uncertainty in the particle or group velocity, is equal to a constant, the value of which can be calculated.

According to this, if we know the position of the particle accurately, which requires a very short group of waves, then the uncertainty in the velocity becomes very large, and if we know the velocity accurately, the uncertainty in the position becomes very large.

Thus it is impossible to know both the position and the velocity accurately. This result which follows from the wave particle theory is called the uncertainty principle. It was discovered by Heisenberg.

This surprising result can be checked by considering any possible experimental way of determining the position and velocity of a particle. Suppose, for example, that we observe it with a microscope. The smallest length which can be seen with a good microscope is about equal to the wave length of the light used, so the uncertainty in the position of the particle will be equal to the wave length of the light used to illuminate it.

When the particle is illuminated, the light photons bounce off it and some of them go into the microscope. But when a photon bounces off the particle, it will set the particle moving or change its velocity. The velocity observed will therefore be uncertain by an amount equal to the greatest change which the photons can produce. If we use light of short wave length, and so reduce the uncertainty in the position, we increase the uncertainty in the velocity, because the energy of a photon is proportional to its frequency, and so inversely as its wave length.

The product of the two uncertainties obtained in this way agrees with the uncertainty principle. This is shown in Appendix 10.

Suppose a beam of electrons is passed through a small hole in a screen. Then the position of the electrons as they go through is uncertain by the diameter of the hole. If we make the hole very small, the waves diverge from it so that the electrons also diverge. The sideways velocity of an electron is therefore uncertain to an extent depending on this divergence.

The smaller the hole, the greater the divergence becomes, and in this case also the product of the two uncertainties agrees with the uncertainty principle. This is worked out in Appendix 11.

It appears, therefore, that the position and velocity of a particle cannot both be accurately determined. It is not therefore possible to have a system of particles for which the positions and velocities are exactly known at a given time. So even if the particles did obey definite laws of motion, the positions and velocities could not be calculated at any future time.

The uncertainty of the wave particle theory is present in all actual observations, so that this theory is superior to the classical theories of the nineteenth century in that it does not assume the possibility of results more accurate than can be really obtained. In the classical theories particles were thought of as in definite positions, and moving with definite velocities, whereas in the new theory they are only thought of as being somewhere in a group of waves.

The uncertainty in the position and velocity may be thought of as due to the disturbance which an observation of the particle necessarily makes in its position or velocity, as in the observation with a microscope considered above.

To know the position it must first be observed. Thus if we like we can imagine that the particle really has a definite position and velocity but that it is impossible to determine them both accurately. Quantities which cannot be observed, however, are really meaningless and the particles themselves are never observed, but only the effects which they are imagined to produce.

CHAPTER VIII

The New Theory of Atoms

De Broglie's wave particle theory has been developed by Heisenberg, Schroedinger, Dirac, and others, with the object of explaining the optical and other properties of atoms.

The new theory of atomic phenomena is known as wave mechanics or quantum mechanics. Heisenberg's theory and that of Schroedinger were based on different ideas, but they have turned out to be mathematically equivalent. This is a remarkable example of how two entirely different theories, based on quite different assumptions, may lead to identical results, and really be essentially the same. Dirac's theory is a sort of combination and generalization of those of Heisenberg and Schroedinger.

The simplest atom is that of hydrogen, which, as we have seen, is just a proton and an electron, and we shall only consider here the theory of this atom which will serve to illustrate the principles involved. When an electric current is passed through hydrogen gas, the atoms are excited and emit light. This light is not all of the same wave length, but a large number of definite wave lengths are found in it which are always exactly the same, and have been accurately measured. Instead of the wave lengths it is more convenient to use the number of waves in an inch or a centimeter, which is called the wave number. One inch is equal to 2.54 centimeters.

The table on page 71 gives some of the wave numbers or waves per centimeter of the light emitted by hydrogen atoms. It is found that these wave numbers are related to each other in a simple way.

The three numbers in the first column are proportional to $(1 - 1/4)$, $(1 - 1/9)$ and $(1 - 1/16)$. The numbers in the second column are proportional to (1/4 — 1/9), (1/4 — 1/16), (1/4 — 1/25), (1/4 — 1/36) and so on, and those in the third column are proportional to (1/9 — 1/16), (1/9 — 1/25) and (1/9 — 1/36). Thus it appears that all these

82258	15233	5332
97481	20565	7799
102823	23032	
	24373	
	25181	
	25706	
	26066	
	26323	
	26513	

wave numbers are equal to a constant multiplied by $(1/n^2 - 1/m^2)$ where n and m stand for two whole numbers 1, 2, 3, 4, etc. For example, in (1/9 — 1/36) we have $n = 3$ and $m = 6$. The series of wave numbers in the middle column was discovered by Balmer and is called Balmer's series.

The constant which when multiplied by $(1/n^2 - 1/m^2)$ gives the observed wave numbers, is equal to 109678. For example, multiplying this by 1/4 — 1/25 or 21/100 we get 23032, which is equal to the third number in the second column.

A theory of these wave numbers was worked out by Bohr before the discovery of wave mechanics or the wave particle theory. Bohr's theory was based on Einstein's idea that light consists of photons and on Rutherford's nucleus theory of the atom.

Bohr supposed that the electron in the hydrogen atom revolves around the proton in an orbit just like a planet revolving around the sun. He assumed that only orbits of certain definite sizes are possible. The possible orbits were determined by the assumption that the product of the momentum of the electron and the radius of the orbit must be an exact multiple of Planck's constant divided by 2π. He calculated the energy of the atom for each of these supposed possible orbits, and so got a series of possible energy values. The possible energy values came out equal to the expression $A - B/n^2$ where A and B are constants and $n = 1, 2, 3, 4$, etc.

Bohr then assumed that an atom could suddenly change from one of its possible energy values to any other smaller value, and that when such a change took place the energy difference was emitted as a photon. The frequency and wave length of a photon are determined by its energy, so Bohr was able to calculate the wave numbers of the light emitted according to his theory, and he found that they agreed exactly with the observed values. The wave numbers are proportional to the energy difference or to $(A - B/n_2^2) - (A - B/n_1^2)$ which is equal to $B(1/n_1^2 - 1/n_2^2)$. This surprising result was considered to justify the peculiar assumptions which he had made. Bohr's calculations are given in Appendix 12.

Bohr's assumptions were quite contrary to classical nineteenth century ideas. According to these, an electron could revolve around a proton in an orbit of any size, just as the orbit of a planet around the sun may be of any size. Also, an electron revolving in an orbit was supposed to produce electric waves, or light of frequency equal to the frequency of revolution, but Bohr supposed the atom did not radiate at all except when the electron jumped from one of its possible orbits to another one. The frequency of the light emitted on Bohr's theory was not equal to the frequency of revolution, but was determined solely by the energy difference emitted. But Bohr's theory gave the correct wave numbers, and it was found that it could be applied to other atoms, and was successful in explaining a great variety of atomic optical phenomena. In spite of its unsatisfactory features it was a great advance over all previous theories of atoms, and contributed more than anything else to the rapid progress which has been made in this fundamental branch of physics.

Heisenberg, Schroedinger, and Dirac have shown that Bohr's assumptions can be justified by means of the wave particle or wave mechanics theory, and they have developed a theory of atoms which has already gone a long way towards explaining their optical and other properties. Their theory may be regarded as a development of Bohr's theory.

We regard an electron as moving in a group of waves, the intensity of which is a measure of the chance of the electron being found at any point. If the electron is moving round a small orbit in an atom, and the distance round the orbit is only a few wave lengths, we must suppose that the group of waves may extend several times round the orbit. In this case, at any point on the orbit we shall have several trains of waves superposed. Unless the crests and troughs coincide, or the waves are in step, the several trains will interfere and destroy each other. It is easy to see that the distance round the orbit must be equal to some whole number of wave lengths if this is not to happen. Thus only those orbits will be possible the distances round which are one, two, three, or more wave lengths. Moreover, it is found that this rule gives exactly the same possible orbits as Bohr's assumption. See Appendix 13.

When the orbit is only a few wave lengths long, the group of waves fills up all the space near the proton, so that we have a distribution of wave motion around the proton, and not merely a train of waves going round an orbit. Such a distribution is analagous to the vibrations of an elastic solid such as a solid rubber ball. The theory of the wave vibrations around the proton was worked out by Schroedinger, and he showed that the possible frequencies of vibration gave possible energies which agreed with those calculated by Bohr. The possible energies are obtained by multiplying the possible frequencies by Planck's constant as for a photon. Schroedinger's wave mechanics theory of the hydrogen atom also explains why the atom does not radiate except while it is changing from one possible energy to another. It cannot yet be said, however, that any very satisfactory explanation of how the atom does radiate has been given.

CHAPTER IX

Cosmic Rays

It has been discovered rather recently that there is a very penetrating kind of radiation in the atmosphere which seems to be coming in from outside. This radiation has been called cosmic rays, and since it has new and interesting properties a brief account of them will be given here.

This radiation can be detected and its intensity measured by means of the electrical conductivity which it produces in air and other gases. An electroscope suitable for measuring such conductivity is shown in Fig. 22.

It consists of a strong steel cylinder, AB, closed at the top by a steel plate, CD, bolted on. EF is a light metal frame attached to the plate. This frame carries two very fine quartz fibers, GH, which are attached at G to a small rod of insulating material and at H to a thicker quartz fiber, MN. The two fibers, GH, are coated with a very thin film of gold to make them conduct electricity. They can be charged with electricity and then repel each other and spread out as shown, pulling up the fiber, MN, slightly. There are two thick glass windows, WW, in the steel cylinder, one in front of, and the other behind, the two fibers. The two fibers are illuminated from behind through the back window and can be observed through the front window. The distance between the fibers can be exactly measured with a microscope. If the gas in the cylinder becomes slightly conducting, the electricity on the fibers slowly escapes to the cylinder, and the distance between the fibers slowly diminishes. The rate at which the distance decreases is proportional to the conductivity of the gas.

If some radium is brought near the cylinder, the penetrating radiation from the radium makes the gas conduct, and the fibers move together. The penetrating radiation from radium is called gamma rays, and is found to be similar to X-rays, but of shorter wave length and more penetrating.

The gamma ray photons passing through the gas in the cylinder knock electrons out of some of the atoms and these electrons have enough energy to knock electrons out of a great many more atoms.

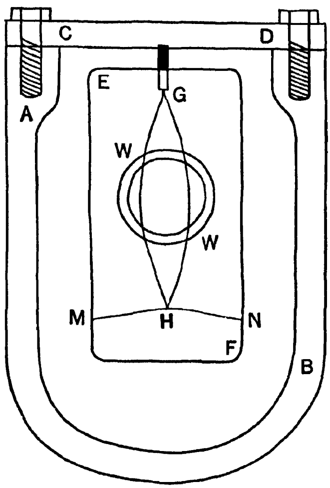

FIG. 22

If the fibers are positively charged they attract the free electrons which move on to the fibers and neutralize the positive charge. The gas is then said to be conducting.

The conductivity obtained can be increased by pumping more gas into the cylinder so in using it to detect the cosmic rays it is best to pump the gas in to a pressure of about one thousand pounds on the square inch. Also it is found that argon gas is better than air.

It is found that there is a small amount of conductivity in the electroscope when it is anywhere at about sea level. Part of this can be got rid of by taking the electroscope in a boat on a deep lake some distance from the shore. This part is believed to be due to a small amount of gamma rays emitted by traces of radium in the earth.

Part of the remaining conductivity can be eliminated by sinking the electroscope deep in the lake so that it is surrounded on all sides by a large thickness of water. The remaining conductivity is believed to be due to radiation emitted by the electroscope itself.

Instead of sinking the electroscope in deep water it may be surrounded by thick walls of lead or iron.

The second part of the conductivity which can be removed by sinking in deep water or by thick metal walls, is supposed to be due to a small amount of penetrating radiation present at the earth's surface, but not coming from the earth.

Experiments like these were first done by McLennan and Burton about 1903. Hess in 1911 and 1912 took such an electroscope up in a balloon. He found that the conductivity decreased at first up to a height of about 3300 feet, but then began to increase, and at 16,500 feet was two or three times that on the ground.

Similar but more accurate measurements were made in 1913 by Kolhorster. He went up to 30,000 feet and found that the radiation increased rapidly up to about 23,000 feet, but increased much less rapidly between 23,000 and 30,000 feet.

These results clearly indicated that the penetrating radiation comes down through the atmosphere and is partly absorbed, so that it gets weaker and weaker as it comes down.

Many experiments on cosmic rays have been made by Millikan from 1926 to the present time. He has measured the conductivity due to the radiation at different heights on the mountains of California and South America, and also at different depths in mountain lakes at high levels.

It is found that the decrease in intensity of the radiation due to sinking the electroscope in water is very nearly the same as that due to sinking it through an equal weight of air. Water is about seven hundred and fifty times as heavy as air at sea level, so seven hundred and fifty feet of air near sea level reduces the radiation as much as one foot of water. Of course, as one goes up in the air it gets lighter, so that at great heights one foot of water is equivalent to much more than seven hundred and fifty feet of air.

In 1929 Regener reported very exact measurements with an electroscope in Lake Constance in Switzerland. He was able to detect the radiation even at a depth of seven hundred and fifty feet below the surface of the water.

The following table gives the relative intensities of the radiation at different depths below the top of the atmosphere, reckoned in feet of water. The whole atmosphere is equivalent to about thirty-four feet of water. The depths given in the table are the thickness of water above the electroscope, plus the water equivalent of the air above it.

Depths in Feet	Differences	Relative Intensities
12		100
	28	
40		50
	46	
86		25
	49	
135		12.5
	65	
200		6.3
	120	
320		3.1
	120	
440		1.6
	100	
540		0.8
	100	
640		0.4

The second column gives the differences between the depths in the first column. Each intensity given is one-half the preceding one. We see that below two hundred feet the radiation is reduced to about one-half by one hundred feet of water. At about one hundred feet it is reduced one-half by about fifty feet of water, and at small depths by only thirty feet.

It is supposed that the radiation coming in at the top of the atmosphere is a mixture of radiations having different penetrating powers. The more easily absorbed parts are removed first, as the radiation goes through matter, so that after going through two hundred feet of water only the most penetrating radiation is left. The thickness of water required to reduce the radiation to one-half is a measure of its penetrating power. The most penetrating part of the radiation is reduced to one-half by about one hundred feet of water, which is equivalent to about nine feet of lead.

The most penetrating radiation previously known, the gamma rays from radium, is reduced to one-half by about six inches of water. The most penetrating cosmic rays are therefore about two hundred times as penetrating as any previously known radiation.

Recent experiments by A. H. Compton have shown that the intensity of the cosmic rays is slightly greater during the day than at night which suggests that a small fraction of the rays may come from the sun or the region round it. Compton also finds that the cosmic rays get stronger as the magnetic poles of the earth are approached. The intensity is least on the magnetic equator, which is half way between the magnetic poles, and is greater by about thirty per cent at points 70 degrees north or south of the magnetic equator. Earlier experiments by Millikan were supposed to show that the intensity is the same all over the earth but his results are probably not as reliable as Compton's. If the cosmic rays are deflected by the earth's magnetic field, as Compton's measurements indicate, it seems probable that they must be electrons or protons since photons are not affected by a magnetic field.

There is another way of detecting the cosmic rays which enables the direction along which they are moving to be determined. Instead of an electroscope, a Gieger counter is used. A Geiger counter is shown in Fig. 23. It consists of a glass tube, AB, about two feet long and two inches in diameter, closed at both ends. A fine wire, CD, is stretched along the axis of the tube, and sealed through the glass at each end. Inside the tube surrounding the wire there is a copper cylinder, EF, which fits into the glass tube. This cylinder is connected to a wire, G, sealed through the glass.

The tube is filled with gas at a pressure of about one-thirtieth of an atmosphere. It is found that argon is the best gas to use. A large battery of about fifteen hundred dry cells is connected to the copper cylinder and to the wire CD. The number of cells in the battery is adjusted so that it is nearly, but not quite, enough to produce an electrical discharge through the gas in the tube. It is found that if a high speed electron or proton is shot through such a tube a momentary discharge or current through the gas is produced. The elec-

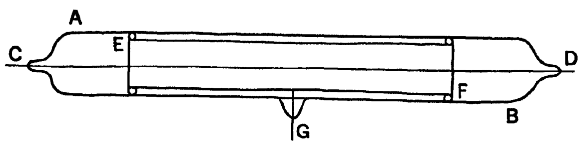

FIG. 23

tron or proton knocks electrons out of a number of atoms in the tube, and so starts the current.

The momentary current can be amplified and made to work a mechanical counter so that every time an electron from outside goes through the tube it is registered by the counter.

When a Geiger counter is set up anywhere it is found that it counts a large number of particles passing through it. If the counter is sunk in a lake the number counted diminishes as the depth below the surface of the water is increased in just the same way as the intensity of the cosmic rays measured with an electroscope diminishes. It is clear, therefore, that the cosmic rays are high speed electrons or protons or else that they produce high speed particles when they pass through matter. The number of particles counted by a Geiger counter like the one described above is, at sea level, about one hundred and fifty per minute.

Bothe and Kolhorster in 1929 tried some very interesting experiments on cosmic rays with Geiger counters. They used two similar Geiger tubes which were connected to the mechanical counter in such a way that it only registered a count when there was a momentary discharge in both tubes at almost exactly the same time. The counter therefore was supposed not to count electrons or protons which passed through only one of the two tubes, but only those which passed through both tubes. The velocity of the high speed particles is so large, over one hundred thousand miles a second, that they go through both tubes at almost exactly the same time.

When the two Geiger tubes are arranged so that one is directly above the other and they are about one foot apart, the counter registers a number of particles, about ten per minute, which seem to go in a nearly vertical direction through both tubes. If the tubes are put horizontally side by side about one foot apart, very few particles are counted. This shows clearly that most of the particles are moving in a nearly vertical direction.

Bothe and Kolhorster put a block of gold about one and a half inches thick between the two counters when one was directly above the other, and found that the gold reduced the number of particles counted about twenty-five per cent. Now one and one-half inches of gold also reduces the cosmic rays as measured in an electroscope by about twenty-five per cent, so that this experiment seems to show that the penetrating power of the particles is equal to that of the cosmic rays. This experiment has been repeated by Mott-Smith using lead instead of gold, and he also finds that the penetrating power of the particles is equal to that of the cosmic rays. He measured the number of particles getting through different thicknesses of lead up to about two feet.

Now the penetrating power of the electrons knocked out of atoms by X-rays and gamma rays is very much smaller than that of the rays so that if cosmic rays were of the same nature as gamma rays—that is, if they were high energy photons—

we should expect them to produce electrons of much less penetrating power.

Bothe and Kolhorster therefore suggested that possibly cosmic rays are not photons, like X-rays and gamma rays, but may be electrons or protons with the enormous energy necessary to give them the great penetrating power observed. A. H. Compton's results which show that cosmic rays are deflected by the earth's magnetic field support Bothe and Kolhorster's suggestion, since photons are not affected by a magnetic field.

Mott-Smith tried an experiment in which he had three Geiger tubes one above the other, and a large block of iron just below the middle tube. The three tubes were connected to a counter so that it only worked when there was a momentary discharge at the same time in all three Geiger tubes. It was thought that if the counter registers electrons or protons passing through all three tubes, then magnetizing the block of iron ought to deflect the electrons or protons off the lower tube and so reduce the number counted, but it was found that magnetizing the iron block did not make any difference to the number counted. It seems probable that the magnetic field in this experiment was not strong enough to produce an observable effect.

As we have seen, C. T. R. Wilson discovered that the tracks made by electrons and alpha-rays when they go through moist air can be made visible and photographed by cooling the air by a sudden expansion.

We should, therefore, expect to find tracks on such photographs due to the cosmic ray particles. The expansion only lasts a small fraction of a second, so that the chance of getting tracks due to the cosmic rays is very small. Since the cosmic ray electrons have great energy and are moving nearly vertically, we should expect them to produce straight tracks in a nearly vertical direction.

Skobelzyn, Mott-Smith and G. L. Locher, and Anderson have all taken a very large number of such photographs and have found a number of nearly vertical straight tracks, presumably due to cosmic ray particles, on them.

Locher and Mott-Smith took a large number of photographs with a Geiger counter just above the expansion chamber, and found that when the counter was excited at the moment of the expansion, then a track was found on the photograph. The direction of the track in the chamber was such as to pass through the counter. These experiments show clearly that the tracks obtained are really cosmic ray tracks.

Skobelzyn and Anderson placed the expansion chamber in a magnetic field and found that the tracks obtained were curved, showing the particles were deflected by the field. Anderson found that some tracks are deflected like electrons, and some the opposite way, like protons.

The experimental results which have been obtained with cosmic rays seem to suggest that they are very high energy electrons or protons. These particles come down through the atmosphere and so probably come in from outside.

It has been suggested that cosmic rays may be uncharged particles consisting of a proton very closely combined with an electron. Such particles are called neutrons and appear to have been obtained recently, but they do not have the properties of cosmic rays. A short account of neutrons is given in Appendix 15.

There is little doubt that cosmic rays are very high energy electrons or protons but we do not know where they come from or how they are produced. They may be produced by high energy photons coming into the earth's atmosphere from outside.

CHAPTER X

Space, Time, and Relativity

Einstein's theory of relativity deals mainly with large scale phenomena but some of the results which follow from it are of fundamental importance for the theory of electrons and other particles. We shall therefore consider this theory briefly here.

The principle conclusion that follows from the relativity theory is that the motion of the earth through space makes no difference, so that it is perfectly proper to regard the earth as at rest. The average man has been in the habit of regarding the earth as at rest for several thousand years and so now has the satisfaction of knowing that he has been conducting his affairs in strict accordance with Einstein's epoch making discoveries.

The fact that position and motion can only be specified relatively to some arbitrarily selected body is of course not new. It has been known for hundreds of years. The body selected is usually the earth, and the position of a point on the earth can be specified by giving its distance north and west of some known point and its height above sea level. For example, if we know that a place is one hundred miles north of the Washington monument, thirty miles west, and one thousand feet above sea level, we can say that we know where it is. This information does not fix the position in any absolute sense because the earth is moving rapidly through space, but it does fix the position relative to the earth.

If a body is moving about on the earth, then its velocity is expressed in miles per hour or feet per second, measured relative to the earth, so of course the velocity is relative to the earth and is in no way affected by the very large velocity of the earth through space.

When it was thought that all space was filled with an ether or some sort of elastic fluid which transmitted light waves, it was naturally supposed that the motion of the earth through this fluid would affect optical phenomena. All sorts of optical and electrical experiments were tried with the object of detecting such effects, but not the slightest sign of any could be detected. The most celebrated of these experiments was one due to Michelson of Chicago. He compared the time taken by light to go a certain distance and back again with the time to go an equal distance and back in a perpendicular direction. The apparatus could be rotated so as to change the two directions relative to the supposed motion of the ether, but there was no effect, although a time difference of one in many hundreds of millions could have been detected.

The absence of observable effects in such experiments was the experimental foundation of Einstein's theory. He assumed that all phenomena are such that uniform motion, in a straight line, of the rigid body on which or relative to which they are observed makes no observable difference. It follows that an observer on a practically rigid body like the earth has no means of determining with what velocity the rigid body is moving through space. Of course, he can observe its velocity relative to other bodies, but he does not know their velocities through space. We can observe the velocity of the earth relative to the sun or the stars, but for anything we know they may all have an enormous velocity about which we know nothing.

The laws of nature have been deduced from the results of experiments done on the earth or observed from the earth. According to Einstein's theory these experiments and therefore the laws deduced from them, are in no way affected by a uniform motion of the earth through space.

This means that if we imagine two earths, moving relative to each other with any uniform velocity, the results of all experiments will be the same on both. In particular the velocity of light will be the same in any direction to an observer on either of the two earths.

But the relative motion will cause the experiments to appear different when seen on one earth by observers on the other.

To see this let us consider a very simple optical experiment. Let a flash of light be sent out from a point S half way between two points A and B on one of the two earths. The observer who does this experiment will consider that the light from S arrives at A and B at the same time because, according to Einstein's theory, any motion of his earth through spaces makes no difference so that the results he obtains will be the same as if his earth was at rest.

It is important to note that this is quite contrary to what might have been expected. We should expect that if the earth on which the experiment is done is moving in the direction from A to B with a great velocity, then as the light moves from S towards A, A will be moving towards S so that the earth's velocity will reduce the time taken by the light to get to A. In the same way, B will be moving in the direction away from S, which will increase the time the light takes to get to B. But according to Einstein's assumption, which is supported by all the facts, the motion of the earth does not affect the times, which are taken to be equal just as if the earth were at rest.

Einstein's assumption is equivalent to supposing that the addition of any velocity to the velocity of light gives a velocity equal to that of light.

Now suppose this experiment is watched by an observer on the other earth. He will regard his earth as at rest and will see the other earth to be moving. He will therefore consider that while the light from S is moving towards A and B, these points are moving. If, for example, they are moving along the direction from A to B, then the light will get to A before it gets to B. Thus in this case two events which occur at the same time for one observer, occur at different times for another. We are forced to conclude that time measurements are relative to the observer and have no absolute significance.

In a similar way it may be shown that equal distances on one earth will not appear equal when seen by an observer on the other.

For example, suppose that on one earth a flash of light is sent out from S in Fig. 24 in two perpendicular directions SA and SB, and that mirrors are put up at A and B so as to reflect

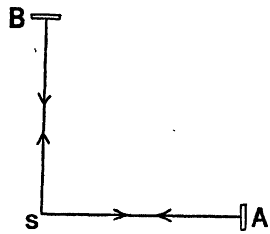

FIG. 24

the light back to S. Also let the distances SA and SB be adjusted so that the light reflected at A gets back to S at exactly the same time as the light reflected at B.

The observer who does this experiment will consider that SA and SB are equal because he will consider that S, A and B are

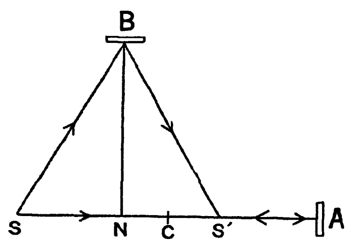

FIG. 25

at rest on his earth, and that the light moves along SA with the same velocity as along SB.

An observer on the other earth will consider that S, A, and B are moving during the experiment and so will not consider that SA and SB are equal.

Let us suppose that the second observer considers that the earth on which the experiment is done is moving with one-half the velocity of light in the direction from S to A in Fig. 25.

It will appear to this observer that S moves to S′ while the light goes from S to B and back to S′. Also SS′ will be equal to either SB or S′B since we are supposing that S moves with half the velocity of light.

While the light goes along SB + BS′ the other ray will go along SA + AS′ so that

$$SA + S'A = SB + BS' = 2SS'$$

But $SA = SS' + S'A$ so that $S'A = \frac{1}{2}SS'$. The mirror A is shown in its position at the moment when the light gets to it. SA is three-quarters of SA + AS′, so that when the mirror is at A, S will have moved along three-quarters of SS′ and so will be at C. Thus we see that the distance between the source and the mirror A is three-quarters of SS′. The distance between the source and the mirror B is equal to NB, which is $\sqrt{\frac{3}{4}}$ of SS′. The distance from the source to the mirror B therefore appears to the second observer to be greater than that from the source to the mirror A. The ratio of the two lengths is $1/\sqrt{3/4}$ or $\sqrt{4/3}$ which is equal to 1.15.

We are therefore forced to conclude that measurements of distances as well as times are relative to the observer and have no absolute significance.

It would actually be extremely difficult, if not absolutely impossible, for the observer on one earth to make observations of what happened on the other earth when the two earths were rushing past each other with an enormous velocity, but for purposes of theoretical calculation we imagine that it could be done.

The difference between what the two observers see is very small unless the relative velocity of the two earths is very large. So long as it is less than say one hundred miles a second, there is practically no difference. Now the velocities with which we have to deal in practice are always less than this, so that practically speaking observations of times and distances made by different moving observers do not differ appreciably.

In the first of the above experiments if the distances from S to A and S to B were one hundred and eighty-six miles, the observer who did the experiment would consider that the light would get to either A or B one-thousandth of a second after it left S since the velocity of light is one hundred and eighty-six thousand miles per second. If the relative velocity of the two earths was half that of light, in the direction from A to B, the observer on the other earth would consider that the light from S would get to A in 2/3000 second and to B in 2/1000 second. So the light would get to A 4/3000 second before it got to B.

The relation between the times and distances observed from the two earths may be obtained from the assumption that the velocity of light is the same for an observer on either earth.

Suppose two events are observed from one of the two earths and that the distance between them is found to be d and the time between them to be t. Also suppose that the same two events are observed from the other earth and that the distance between them is found to be d′ and the time between them to be t′. We suppose the two earths to be moving relatively to each other with any uniform velocity in a straight line so that t will not be equal to t′ nor d to d′. If we suppose that a body was present at both the events, so that it moved the distance between them in the time between them, then its velocity would be d/t to the first observer and d′/t′ to the second.

If this velocity was equal to the velocity of light it would have the same value for both observers, so that denoting the velocity of light by c, we should have $d = ct$ and $d' = ct'$. For light going the opposite way we should have $d = -ct$ and $d' = -ct'$. The equations $d = ct$ and $d = -ct$ may be combined into the equation $d^2 = c^2t^2$ and in the same way we get $d'^2 = c^2t'^2$. These equations may be written $d^2 - c^2t^2 = 0$ and $d'^2 - c^2t'^2 = 0$.

If, then, we suppose that for *any* two events, for which $d^2 - c^2t^2$ may have any value, the equation $d^2 - c^2t^2 = d'^2 - c^2t'^2$ is true, then this relation between d, t, and d′, t′ will make the velocity of light the same to observers on either earth because if $d/t = c$, then $d^2 - c^2t^2 = 0$, so that since $d^2 - c^2t^2 = d'^2 - c^2t'^2$ we get $d'^2 - c^2t'^2 = 0$, and therefore $d'/t' = c$. Therefore a body which appears to be moving with the velocity of light to one observer will also appear to be moving with the velocity of light to the other if for any two events $d^2 - c^2t^2 = d'^2 - c^2t'^2$. The quantity $d^2 - c^2t^2$ is therefore the same for all observers and so has an absolute significance. Let us denote it by s^2 so that $s^2 = d^2 - c^2t^2$. It is sometimes more convenient to have s^2 equal to the sum of two parts instead of the difference, so we introduce a new quantity T such that $T^2 = - t^2$ and so get $s^2 = d^2 + c^2T^2$. The new quantity T may be called the imaginary time, since quantities having negative squares do not really exist. The equation $T^2 = - t^2$

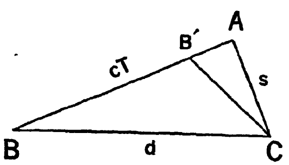

Fig. 26

gives $T = it$ where i stands for the square root of minus one, and so is an imaginary quantity which does not really exist.

The equation $s^2 = d^2 - c^2t^2$ or $d^2 = s^2 + c^2t^2$ may be represented geometrically as in Fig. 26.

ABC is a rightangled triangle such that $BC = d$, $AB = ct$, and $AC = s$. For such a triangle $BC^2 = AC^2 + AB^2$ so that the triangle represents the same relation as $d^2 = s^2 + c^2t^2$. If another observer gets d' and t' instead of d and t then these values can be represented by the triangle AB'C. We see that if t' is smaller than t, then d' is also smaller than d, so as to keep s the same.

The equation $s^2 = d^2 + c^2T^2$ may be represented geometrically in the same way. In Fig. 27 let ABC be a rightangled triangle such that $BC = s$, $AB = cT$, and $AC = d$. Then $BC^2 = AB^2 + AC^2$ or $s^2 = d^2 + c^2T^2$. In this representation, however, cT is an imaginary distance since $cT = ict$. The triangle as drawn represents real values of CT, d, and s. It is impossible to draw the triangle so that the lengths of its sides represent the actual imaginary values. For example, if $d = 5$ and $ct = 3$ or $cT = 3i$, we get $s^2 = 25 - 9 = 16$, so that $s = 4$. But a triangle with $BC = 4$, $AC = 5$, and $AB = 3i$, cannot be drawn to scale. It is an imaginary triangle.

The absolute quantity s is a combination of distance and time. Its distance and time components are different for different observers, but its value is the same for them all.

The quantity s may be called the absolute interval between the two events.

As an example suppose an observer finds the time between

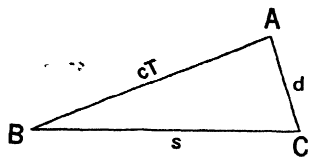

FIG. 27

two events to be 1/1000 second, and the distance between them to be two hundred miles. Then the absolute interval between these two events is given by $s^2 = 200^2 - (186,000/1000)^2 = 7300$, so that $s = 85.5$. If another observer found the time between the same two events to be one hundred years or about three thousand million seconds, his value of the distance d between them would be given by the equation $7300 = d^2 - (186,000)^2 (3,150,000,000)^2$ which gives $d = 600,000,000,000,000$ miles. An observer who found that the same two events occurred at exactly the same time would get for the distance d between them the value given by $7300 = d^2$ or $d = 85.5$ miles.

An interesting way of thinking of the events which happen in the universe was proposed by Minkowski, based on this idea of the absolute interval. We imagine a diagram, which may be called the s-diagram, constructed, in which *each event* is represented by a point and the points are so arranged that the distances between them are equal to the absolute intervals s. It is possible to imagine this when s is taken equal to $\sqrt{d^2 + c^2T^2}$.

The events as seen by any particular observer happen one after another in different positions. In the s-diagram they do not happen, they merely exist. The s-diagram is drawn in a region which is not ordinary space, but is a sort of blend of space and imaginary time. Space and time are combined in it into a single absolute quantity s equal to $\sqrt{d^2 + c^2T^2}$. The separation of s into two parts d and cT is a purely relative operation of no real significance.

Minkowski therefore considered that the s-diagram represents the universe as it really is. He suggested that the separation of events which exist in the s-diagram into a series of happenings in space and time is due to the one-sided view which any particular observer necessarily gets. This is discussed more fully in Appendix 16.

Let us now try to find an absolute velocity which is the same for all observers just as the absolute interval s is.

The velocity of light is the same for all observers, so it satisfies this necessary condition for an absolute velocity.

Consider a particle moving along a straight line with uniform velocity v. Its arrival at two points A and B on the line may be regarded as two events, so if t denotes the time it takes to go from A to B, and d the distance from A to B, then we have

$$s^2 = d^2 - c^2t^2 \text{ or } s^2 = d^2 + c^2T^2.$$

The quantities d and ct may be said to be the space and time components of s. cT is the imaginary time component.

If we suppose that any particle has an absolute velocity along s equal to the velocity of light c, then this velocity will have a space component parallel to d equal to $c \times (d/s)$ and a time component equal to $c \times (ct/s)$.

For in Fig. 28 if we suppose that BC represents the velocity c along s, then BA and AC will represent its components parallel to cT and d. The component along AC is therefore equal to c $\times$ (AC/BC) or c $\times$ (d/s) and the component along BA to c $\times$ (BA/BC) or c $\times$ (cT/s). This is the imaginary time component so that the time component is c $\times$ (ct/s).

If the ordinary velocity v of the particle is small compared with that of light, d^2 will be very small compared with c^2t^2 so that in the equation $s^2 = d^2 - c^2t^2$ the term d^2 may be neglected, and we have approximately $s^2 = -c^2t^2$, which gives $s = ict$ where i stands for the square root of minus one or $\sqrt{-1}$. The space component c $\times$ (d/s) is therefore equal to

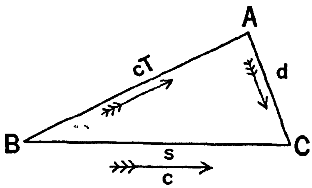

FIG. 28

c $\times$ (d/ict) or d/it which is equal to v/i since v is equal to d/t. Now the space component of the absolute velocity should be equal to v, not v/i. It is therefore necessary to take the absolute velocity along s equal to ic instead of to c. The space component is then ic $\times$ (d/s) which is equal to ic $\times$ (d/ict) and so to d/t or v.

It appears therefore that if we suppose every particle to have a velocity ic along s, then its ordinary or space velocity will be d/t or v. The time component of the velocity ic is ic $\times$ (ct/s) which is equal to ic $\times$ (ct/ict) or just c. Different observers will get different values for v, but the absolute velocity is the same for all observers.

A very important result was obtained by Einstein by means of this idea of absolute velocity. This result depends on force and work or energy, which we must therefore consider briefly.

The momentum of a body is simply the product of its weight or mass and its velocity. When a force acts on a body it changes the momentum of the body. The force is taken to be equal to the rate of change of the momentum. Thus if a force F acts on a body of weight m and changes its velocity from v_1 to v_2 in a time t, we have $F = (mv_2 - mv_1)/t$. The work done by the force is taken to be equal to the force multiplied by the distance through which it acts.

Let us apply these ideas about force and work to a body moving with the absolute velocity ic of which the time component is c. Suppose a force F acts on the body so as to tend to increase the momentum of the body. The velocity c is a constant and cannot be changed so the only way in which the force can be supposed to increase the momentum is by increasing the weight m. If then, we suppose that the force F increases m from m_1 to m_2 in a time t, we have $F = (m_2c - m_1c)/t$. mc is the time component of the momentum, so that F must be the time component of the force. The time component of s is ct and this is the distance through which the force acts, so that the work done by the force is Fct and

$$Fct = (m_2c - m_1c)\ c = (m_2 - m_1)c^2.$$

Thus the work done by the force or the energy which it gives to the body is equal to the increase in the weight of the body multiplied by the square of the velocity of light. We may conclude from this that the energy in a body of weight m is equal to mc^2. Work or energy is usually expressed in terms of the foot-pound or the work required to lift a pound weight up one foot.

To get the energy mc^2 in foot pounds it is necessary to express m in pounds, c in feet per second, and divide by 32. This is explained in Appendix 15.

For a body weighing one pound, the energy is equal to 30,000,000,000,000,000 or 3×10^{16} foot pounds. According to this there is enough energy in a one-pound weight to drive an engine of one hundred thousand horse power for nearly twenty years.

Einstein's theory therefore leads to the very interesting result that there is an enormous store of energy in matter. So far no way of getting this energy out and using it has been discovered.

Since the energy of a body of weight m is mc^2 we may conclude that energy has weight so that the principle of the conservation of energy is really the same thing as the principle of the conservation of matter.

The energy radiated away by the sun in one second according to this weighs about four million tons. The amount of this which is intercepted by the earth is about twelve pounds per second.

There is an interesting application of these ideas about energy and matter to atomic weights. The atomic weight of hydrogen is 1.0078 and that of helium is 4. But a helium atom is formed out of four hydrogen atoms and so should have an atomic weight of 4.0312. It is supposed that during the formation of a helium atom out of hydrogen, enough energy is radiated away to account for the decrease of weight from 4.0312 to 4.

The atomic weight of uranium is about 238, and that of radio-lead 206, so that the difference is approximately 32. During the radio-active transformation of uranium into radio-lead eight alpha-particles are emitted so that since the atomic weight of an alpha-particle is four we should expect the change in the atomic weight to be 32.

But the alpha-rays are emitted with very high velocities and so have considerable kinetic energy. It is found that the weight of this energy, assuming the weight to be equal to the energy divided by c^2, corresponds to an atomic weight of 0.05. We should therefore expect the difference between the atomic weights of uranium and radio-lead to be 32.05 instead of 32. The best estimates of these atomic weights indicate that the difference is a little greater than 32, and so support the theory that an amount of energy E has weight equal to E/c^2.

In getting the space component of the velocity we supposed that d^2 in the equation $s^2 = d^2 - c^2t^2$ could be neglected. If we do not make this approximation we can get the exact value of the space component. The velocity v is equal to d/t so that $d = vt$ and $s^2 = v^2t^2 - c^2t^2$ so that $s = t\sqrt{v^2 - c^2}$ or $s = it\sqrt{c^2 - v^2}$ where $i = \sqrt{-1}$. The space component of the velocity is equal to $ic \times (d/s)$ or $icvt/s$ which is equal to $icvt/it\sqrt{c^2 - v^2}$ or to $cv/\sqrt{c^2 - v^2}$ and this may be written $v/\sqrt{1 - v^2/c^2}$. The exact value of the space component of the velocity is therefore equal to $v/\sqrt{1 - v^2/c^2}$ which is equal to v when v/c is small.

Another very interesting result can be obtained by considering the momentum of a body or the product of its weight or mass m and its velocity.

Corresponding to the absolute velocity ic and its space component $v/\sqrt{1 - v^2/c^2}$ we have absolute momentum mic with space component $mv/\sqrt{1 - v^2/c^2}$. We may regard this as the product of a weight $m/\sqrt{1 - v^2/c^2}$ and the velocity v. The weight $m/\sqrt{1 - v^2/c^2}$ is equal to m when v/c is very small, but as v/c increases it becomes greater and gets very large when v is nearly equal to c.

The following table gives values of the weight of a body when moving with different velocities calculated by means of the expression $m/\sqrt{1 - v^2/c^2}$.

Velocity in Miles Per Second	Weight
0	1
1,000	1.00006
10,000	1.006
20,000	1.023
50,000	1.04
100,000	1.18
150,000	1.69
180,000	3.96
185,000	9.7

Such an increase of weight with velocity was first predicted by J. J. Thomson in 1881, long before the theory of relativity was discovered by Einstein.

This variation of weight or mass with velocity has been verified experimentally by experiments on electrons moving with great velocities. Radium emits electrons which have velocities ranging up to nearly the velocity of light. Kaufmann and Bucherer measured the weight per unit charge for these electrons by finding the way in which they are deflected in electric and magnetic fields as was explained in Chapter II. They both found that the weight varied with the velocity in accordance with the expression $m/\sqrt{1 - v^2/c^2}$ in which m stands for the weight when the velocity v is very small.

This increase of weight with velocity is due to the weight of the energy of motion or kinetic energy. The energy when the particle is at rest is mc^2 and when it is moving so that its weight is $m/\sqrt{1 - v^2/c^2}$ instead of m its energy is $mc^2/\sqrt{1 - v^2/c^2}$. When v/c is small $\sqrt{1 - v^2/c^2}$ is equal to $1 - v^2/2c^2$ so that $1/\sqrt{1 - v^2/c^2}$ is equal to $1 + v^2/2c^2$ and therefore $mc^2/\sqrt{1 - v^2/c^2}$ is equal to $mc^2(1 + v^2/2c^2)$ or $mc^2 + \frac{1}{2}mv^2$. The increase in the energy when v/c is small is therefore equal to $\frac{1}{2}mv^2$ which, as we have seen, is equal to the kinetic energy.

As we have seen, the energy of a photon is equal to Planck's constant multiplied by the frequency of the light. Denoting Planck's constant by h, the frequency by n, and the energy of the photon by E, we have $E = hn$. The weight of a photon is equal to E/c^2 or hn/c^2. The velocity of a photon is always equal to c, so that its momentum is equal to $(hn/c^2) \times c$ or hn/c. The wave length l of the light is equal to c/n so that the momentum is equal to h/l.

In the same way for an electron or other particle of weight m moving with velocity v, the momentum mv is supposed to be equal to h/l where l is the wave length of the electron waves. Hence $mv = h/l$ or $l = h/mv$. Also the frequency n is determined by the energy as for a photon. The energy is equal to

mc^2 so that $mc^2 = hn$. The wave velocity which we will denote by u is equal to nl. Hence we have

$$u = nl = hn/mv = mc^2/mv = c^2/v.$$

so that $uv = c^2$. Thus the product of the wave and particle velocities for any particle is equal to the square of the velocity of light. For a photon u and v are both equal to c, but for electrons and protons v is less and u greater than c.

Thus the theory of relativity leads to very interesting and important results about energy and matter. These results appear to be in accordance with the facts so we may conclude that the theory of relativity is a reliable theory.

CHAPTER XI

Relativity and Gravitation

According to the theory of relativity uniform motion of an earth along a straight line makes no difference to phenomena observed on the earth. But non-uniform motion certainly does make a difference. For example, on a liner moving over a calm sea at constant speed, passengers in the cabins cannot tell that the boat is moving, but when the sea is rough so that the motion is not uniform, it produces easily noticeable effects. Suppose a passenger in one of the cabins did not know he was on a ship and knew nothing at all about what was outside his cabin; would it be possible for him to prove that the cabin was moving in a non-uniform manner by making observations inside the cabin? According to Einstein he could not find out anything about the motion of his cabin whether it was uniform or not. To see how this surprising conclusion is reached it is necessary to consider gravitation. By taking gravitation into account Einstein extended his theory of relativity so as to include any kind of motion as well as uniform motion.

By gravitation we mean the attraction between heavy bodies which causes them to move toward each other. For example, all bodies when not supported fall towards the earth. Gravitation was shown by Newton to be a universal property of matter. The sun, earth, moon, and all the planets attract each other in just the same way that the earth and any body near it do.

It is found that all bodies of whatever kind fall towards the earth at the same rate. For example, a ball of lead and a ball of cork fall at the same rate. As a body falls its velocity increases. After falling freely for one second its velocity is

32 feet per second and after two seconds, 64 feet per second, and so on. The velocity increases by 32 feet per second in each second while the body is falling freely. The body is said to fall with an acceleration of 32 feet per second per second. Anyone who drives a car is familiar with acceleration. If a car gets up to 60 miles an hour in half a minute its acceleration is 120 miles an hour per minute. One hundred and twenty miles an hour is equal to 176 feet per second, so an acceleration of 120 miles an hour per minute is equal to 176 feet per second per minute or 176/60 feet per second per second.

A region in which gravitational forces are present is called a gravitational field. The fundamental property of a gravi-

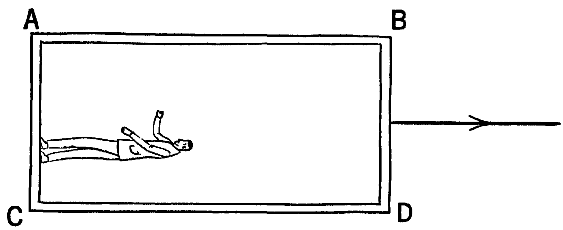

FIG. 29

tational field is that it produces the same acceleration of all bodies, large or small, and of whatever kind. A uniform gravitational field is one in which the acceleration is the same everywhere in the field. The gravitational field of the earth is nearly uniform near to the earth's surface.

Let us now consider a simple imaginary experiment, due to Einstein, which enables us to obtain important information about gravitation. We imagine a man inside a large box in a region where there is no gravitational field. The man is supposed not to be able to see outside the box and not to know anything about what is done outside. Suppose that a rope is attached to the box and that the rope is pulled so as to make the box move with a constant acceleration as in Fig. 29.

The acceleration of the box will cause the man to fall on to the end of the box AC and he will be pressed against this end with a force sufficient to make him move with the same acceleration as the box. He will be able to stand up on the end AC and walk about just as if the box was on the earth with the end AC on the ground.

If he picks up any small body in the box and lets it drop, it will fall towards the end AC with an acceleration relative to the box, equal but opposite to that of the box. For a force is required to make a body move with an acceleration so that a body in the box which is unsupported and so has no force acting on it will have no acceleration. It will therefore appear to have an acceleration relative to the box equal and opposite to that of the box.

The man in the box will therefore find that all bodies of whatever kind and size fall in the box on to the end AC with the same acceleration.

If the man knows about gravitational fields he will very likely come to the conclusion that his box is supported in a gravitational field which produces the acceleration which he observes. But it will be impossible for him to tell whether the acceleration is due to a gravitational field or to an acceleration of the box, or to a combination of the two because either could produce exactly the effects which he observes.

We might suppose that although it is impossible for the man to distinguish between a gravitational field and an acceleration of the box by experiments with falling bodies, yet he might be able to do it by optical or electrical experiments. Einstein, however, makes the assumption that there will be no observable difference between the effects due to a gravitational field and an acceleration of the box in optical, electrical, or any other kind of experiments. This assumption is called Einstein's principle of equivalence. The results of his theory are found to agree with the facts so that the principle of equivalence is probably correct.

A very interesting result can be immediately deduced from the principle of equivalence. Suppose that the man in the box observes the path of a ray of light going across the box as in Fig. 30.

If the box was at rest, the ray would go along the straight line ABC. But if the box has an acceleration then as the light goes from A to B, the box moves so that B′ gets to B. The man will therefore consider that the light goes from A to B′. Also as the light goes from A to C the box will move so that C′ gets to C, so that it will appear to the man that the light goes along AB′C′. Now the velocity of the box is increasing all the time, so that C′C will be more than double B′B, and therefore the apparent path AB′C′ will not be straight but curved.

The acceleration of the box will therefore cause light to

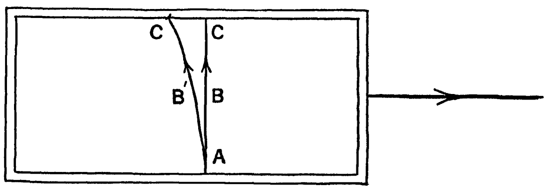

FIG. 30

appear to move along curved instead of straight lines. According to the principle of equivalence a gravitational field will have exactly the same effects as an acceleration of the box, so that we are forced to conclude that light must be deviated by a gravitational field. This effect is very small because the velocity of light is so large, but Einstein calculated that there should be an appreciable deviation of a ray of light going by the sun, due to the very strong gravitational field near the sun.

This effect can be detected by observing the apparent positions of stars when they are nearly behind the sun. The bright light of the sun makes it impossible to see stars when they appear to be near to it, but during a total eclipse of the sun by the moon they can be seen and photographed.

Astronomers have made photographs of the stars near the sun during several total eclipses, and have also photographed the same stars with the same cameras when not near the sun. On comparing the two sets of photographs they have found that the presence of the sun does slightly alter the apparent positions of the stars, and that the effect found is just equal to that which Einstein calculated. These results therefore strongly support the principle of equivalence and Einstein's theory, which is based on it.

There is another important result which can be deduced from the experiment with the man in a box. Let us suppose that the box is not pulled with a rope, but is put in a uniform gravitational field. This field will give to the box and to everything in it the same acceleration. A body in the box will not then fall if unsupported because it will move at the same rate as the box. It will have the same acceleration as the box, and so no acceleration relative to the box. The man in the box will therefore not be able to detect any effects due to the gravitational field. Everything in the box will behave exactly as if the box was at rest or moving uniformly. We see therefore that on a system moving freely in a uniform gravitational field there are no observable effects due to the field. In effect, there is no field to observers on the system.

For example, if the rope of an elevator breaks and the elevator then falls freely, things in the elevator behave as if there were no gravitational field. If a person in the falling elevator drops a parcel, it will not fall to the floor of the elevator because the elevator itself is falling freely and so moves at the same rate as the parcel.

In the same way, the earth is falling freely in the gravitational field of the sun, so that this field does not produce observable effects on the earth, except in so far as it is not a uniform field.

According to Newton's first law of motion, a body which is not acted on by any force moves with constant velocity in a straight line. This is really nothing more than a definition of what is meant by the absence of any force acting on the body.

If the body does not move with constant velocity along a straight line, then we say that a force is acting on it.

Now the motion of a body can only be observed relatively to some body like the earth on which the observer is. Let us suppose an observer on a large body observes a small body moving along somewhere out in space, and finds that it appears to him to move with constant velocity in a straight line. He will consider that it is moving freely with no force acting on it.

Now suppose another observer on another large body also observes the small body. If the second large body is moving relatively to the first one with uniform velocity in a straight line, he also will find that the small body is moving with constant velocity along a straight line. His value of the velocity will not be the same as that found by the first observer, but it will be constant.

Suppose, however, that the second observer is on a body which is not moving uniformly in a straight line relative to the first large body. For example, suppose the second body is rotating or moving along a curved path with a varying velocity. In this case the second observer will not find that the small body is moving relatively to the second large body with a uniform velocity in a straight line, and so may consider that forces are acting on it. In this way we see that Newton's first law of motion is only true for observations made from bodies moving uniformly with respect to a body relatively to which the law is true. It is not true in any absolute sense.

If the large body from which the observations are made has an acceleration, the small body will appear to have an equal and opposite acceleration, and so will appear to be acted on by a force. Such apparent forces are exactly like gravitational forces, so that a body moving with an acceleration seems to produce a gravitational field. For example, the acceleration with which all bodies fall towards the earth is only partly due to the gravitational field of the earth. Part of it is due to the acceleration of the earth due to its rotation. It is easy to see that all gravitational fields are not merely apparent ones due to accelerations of the body from which the observations are made. The acceleration with which bodies fall to the earth cannot be attributed to an upward acceleration of the ground because this would require the earth to be getting rapidly larger, which is certainly not the case.

If an observer on any large body observes forces on, or motions of, bodies, apparently due to a non-uniform motion of the large body he is on, he cannot be sure that the forces or motions he observes are not due to a gravitational field or partly to such a field and partly to motion of the body he is on. For example, a passenger in a cabin on a ship may observe effects apparently due to an irregular motion of the ship, but without further information, he cannot be sure that the ship is not at rest and the effects due to a varying gravitational field.

We are now in a position to consider Einstein's theory of gravitation. A body acted on by no forces, apparent or gravitational, moves with uniform speed along a straight line. Let A and B be two points on this line, and let d denote the distance from A to B and t the time the body takes to go from A to B. Then as we saw in the previous chapter, different observers on different uniformly moving bodies will observe different values of d and t, but they will all get the same value for $s^2 = d^2 - c^2t^2$. The absolute interval s is a blend of space and time. Putting $-t^2 = T^2$ as before, so that $s^2 = d^2 + c^2T^2$ we may regard s as a distance in a region which is not ordinary space, but involves both space and imaginary time. This supposed imaginary time, space region is of course purely imaginary. It is the same thing as the s-diagram of Minkowski considered in the previous chapter.

In ordinary space the three sides of a rightangled triangle ABC are related by the equation $BC^2 = AB^2 + AC^2$ and in the space time region in the same way $s^2 = d^2 + c^2T^2$. The time quantity cT plays the part of a distance in this imaginary region. Ordinary space in which the distance d is measured is said to be of three dimensions because three mutually perpendicular directions exist in it or because it has height, length, and width, but the space time region in which the absolute interval s may be imagined to exist, has four dimensions, namely, the three dimensions of ordinary space and also a fourth dimension corresponding to cT. The interval s is the distance between two points in the space time region. The straight path of the particle as it moves from A to B in ordinary space corresponds to a straight line of length s in the space time region.

If the body is observed from a large body which is not moving uniformly it will appear not to move uniformly, so that its line in the space time region will appear to be curved. This apparent curvature is not due to any real change in the line, but merely to the distortion of the time and space measurements by the non-uniform motion of the body from which the observations are made. The line appears to be curved but it is still the straightest possible line, since it is not really changed.

If, however, the moving body is moving in a gravitational field, then it will really describe a curved path. Einstein supposes that a gravitational field is a distortion or curvature of the space time region so that the straightest possible line between two points in it is not straight, but curved. The moving body goes along the straightest possible path, and so describes a curve as if a force was acting on it. According to this idea gravitation is not a force in the usual sense, at all. The line of a moving body is supposed always to be the straightest possible line in the space time region. When this region is not distorted or curved the straightest line is straight, but in a gravitational field the region is distorted or curved so that the straightest line possible is curved.

It is very difficult to imagine a curved four dimensional region, but we can get an idea of what it means by considering a two dimensional one. A surface which has length and breadth only, and no thickness is a two dimensional region. If such a surface is flat, then a straight line can be drawn in it between any two points, but if the surface is curved then it is not in general possible to draw a straight line on it between any two points. The shortest line between the two points is the straightest line. For example, the shortest line between two points on a spherical surface is a part of a great circle; that is, a circle which divides the sphere into two equal hemispheres.

The surface of a sphere is a curved two dimensional space which is in a three dimensional region. At any point in it two perpendicular lines can be drawn which are both perpendicular to a radius of the sphere. The radius, of course, lies outside the two dimensional surface.

We can imagine a curved three dimensional region located in a four dimensional space. At any point in such a region it will be possible to draw three mutually perpendicular lines which will all three be perpendicular to the radius of the three dimensional region. In the same way, we may try to imagine a curved four dimensional space located in a five dimensional region. At any point in the space four mutually perpendicular lines can be drawn which are all perpendicular to the radius of the space. The radius, of course, is outside the four dimensional space.

According to Einstein, then, a gravitational field is just a curvature of the imaginary four dimensional region in which the absolute intervals s can be imagined to be. It is a distortion of space and time and not a field of force. Just why the presence of a heavy body should distort the surrounding space and time has not been, and very likely never will be, explained. Also we cannot offer any reason why the line belonging to a body should be the straightest possible line in the curved space-time.

Einstein worked out the path of a planet round the sun on his new theory, and found that it came out almost exactly the same as on Newton's theory, which of course agrees very exactly with the observed motion.

However, in the case of the planet Mercury, which is nearest to the sun, the two theories differed appreciably. Einstein's new theory showed that the elliptical orbit of Mercury should slowly rotate round the sun so as to go once round in three million years. This very slow rotation of the orbit was not indicated by Newton's theory, but it had been observed by astronomers who were unable to find any satisfactory explanation of it. Einstein's new theory gave exactly the observed rotation, which is a strong point in its favor.

There is one other effect indicated by Einstein's theory which can be observed. When light waves and photons are emitted by a star, the photons should be attracted by the star, and so should lose some energy in overcoming the attraction.

But the frequency of a photon is proportional to its energy, so that the frequency of the light should be slightly decreased as it escapes from the star. This effect can be detected by comparing the light emitted by the atoms of any element in the star with the light emitted by the same element on the earth. It is found that there is a very small difference which is just equal to that indicated by Einstein's theory.

The above outline of Einstein's theories about space, time, and gravitation is necessarily very incomplete. A full account would be too mathematical for inclusion in this book. The writer hopes, however, that what has been written is sufficient to enable anyone to get some idea of the principles on which these epoch-making theories are based.

It will be observed that Einstein's theories are largely based on ideas derived directly from known facts. They are not philosophical discussions based on vague speculations, but are examples of supremely logical argument from reliable premises.

CHAPTER XII

Conclusion

According to Dirac, who is one of the leading authorities on the new mechanics of atoms, "the only object of theoretical physics is to calculate results that can be compared with experiment, and it is quite unnecessary that any satisfying description of the whole course of the phenomena should be given."

To this it may be replied that if ever some genius does discover a satisfying description of the course of phenomena, it will be welcomed with enthusiasm by a good many people. Meanwhile we have to admit that we cannot offer any such description, and so have to get along without one.

The tentative conclusion to which we have come is that the universe seems to consist of several sorts of particles moving about in so-called empty space. These particles, however, are never directly observed, but effects are observed which are such as might be expected from particles. The distribution of these effects, however, is not such as might be expected for particles, but is what might be expected for a distribution of wave intensities. The chance of an effect, apparently due to a particle, appearing, is proportional to the wave intensity. But the waves for two or more particles which interact are not waves in actual space, but are waves in imaginary space of many dimensions. The wave theory therefore seems to be merely auxiliary mathematics which enables the distribution of the effects to be calculated but does not represent directly any physical reality. We have got a theory which is very successful in calculating results which agree with experiments, but we do not really understand it. The mathematical formulae contain symbols of unknown meaning or are assigned meanings which will very likely prove to be quite wrong. The underlying reality is probably not either waves or particles, but something so different from any familiar large scale phenomena that it is going to be extremely difficult ever to form a conception of it.

Waves and particles are familiar things and so we try to explain phenomena in terms of them. So long as we remember that they are only models probably quite unlike the underlying reality they do no harm, and may even serve a useful purpose.

It is generally realized now that it is important not to form crude conceptions about the physical meaning of quantities but to preserve an open mind and be prepared to change any conceptions one may have got into the habit of using.

In the nineteenth century the universe was thought of as a collection of particles, the atoms, which obeyed Newton's laws of motion. The future position of every particle was therefore determined by the positions and motions of the particles at any given time. The course of events was therefore fixed by natural law. Free will was impossible. A man seemed to decide what he would do, but the process by which he decided was controlled by natural laws, and the result was determined beforehand. According to this idea it was difficult to believe in any supernatural powers controlling the evolution of the universe, and in particular that of life on the earth. It was, however, extremely difficult not to believe in free will. The conclusion that there is no free will seems to be contrary to the facts and so requires the theory to be modified. Moreover no one really believed that art, literature, religion, and all the other human activities of a more or less spiritual character could be regarded as the results of a purely mechanical process based on Newton's laws of motion. The theory was obviously quite inadequate to explain these facts.

These serious difficulties are avoided by the new mechanics of waves and particles. The course of events is not determined by the laws of nature. A given state of things may be followed by a variety of events. The laws merely enable the probability of each possible event to be calculated. Moreover, most events are not repeated a great many times, but only happen once, so that it is not usually possible to find out if the probabilities agree with the theory.

The laws of nature appear to have been designed so as to allow the course of events to be guided from outside without any violation of the laws.

It is true that in large scale phenomena involving very many particles the most probable course of events is so much more probable than any other that appreciable deviations from it practically never happen. Large scale phenomena therefore are, practically speaking, controlled by exact natural laws. But the actions of human beings are controlled by processes in the brain which are not necessarily large scale phenomena. Unfortunately we do not yet know very much about such brain processes, but there seems to be no reason to doubt that they are small scale phenomena involving relatively few particles, so that there is always a choice of many possible events. There is a possibility of free will without any violation of natural law.

What then does decide what shall happen if natural laws do not? The only answer to this question is that we do not know unless it is that the brain is controlled by spiritual forces or qualities, not usually included in the physicist's scheme. If we do not believe that human activities can be explained on a purely mechanical basis, then we must suppose they are controlled by so-called supernatural qualities. That is to say, a man is not merely a machine.

When we consider the beginning and evolution of life on the earth it is difficult to believe that it is a purely mechanical process. It seems necessary to suppose that this evolution is controlled or guided from outside. The new ideas about waves and particles make this possible without any violation of natural laws.

The conclusion that large scale phenomena involving very large numbers of atoms are practically speaking controlled by nearly exact natural laws, while atomic phenomena are not so controlled, cannot be regarded as very satisfactory. A theory which merely gives the relative probabilities of various possible events and leaves what actually happens either to pure chance, whatever that may mean, or to supernatural guidance can hardly be the physicist's final solution of his problems.

APPENDICES

APPENDIX 1

Charge per Unit Weight of Electrons

The weight of a body is usually expressed in pounds and is found by weighing the body on a balance with pound weights. The weight so found is used as a measure of the amount of matter in the body which in scientific books is called the mass of the body. In this book we shall use the word weight, in the conventional way, to mean the number of pounds in a body as found with a balance.

In scientific books the word weight is sometimes taken to mean the force with which the earth attracts the body and the weight is then expressed in terms of some unit of force. The unit of force most commonly used in scientific work is called a *dyne*. It may be defined as a force which gives one gram a velocity of one centimeter per second in one second. One pound is equal to 453.593 grams and one foot is equal to 30.48 centimeters. The force with which the earth attracts a pound is often used as a unit of force and is said to be a practical or engineering unit of force.

When a force acts on a body and the body moves, the force is said to do work. The work is taken to be equal to the product of the force into the distance through which it acts.

If a force F acts on a body of weight or mass m and causes its velocity to increase from v_1 to v_2 in a time t, then

$$F = \frac{mv_2 - mv_1}{t}.$$

The product of the weight m and the velocity v is called the momentum of the body so that the force is equal to the increase in the momentum divided by the time or to the rate of increase of the momentum.

The distance the body moves in the time t is equal to its average velocity $(v_1 + v_2)/2$ multiplied by the time t. So if s denotes this distance, then

$$s = \frac{v_1 + v_2}{2} \times t.$$

The work done by the force F is equal to Fs and

$$Fs = \frac{m(v_2 - v_1)}{t} \times \frac{(v_1 + v_2)\,t,}{2}$$

so that

$$Fs = \frac{m}{2}(v_2^2 - v_1^2).$$

The work done is equal to the increase in the energy of motion or the kinetic energy of the body. We see therefore that the kinetic energy is equal to $\frac{1}{2}mv^2$. Work is often expressed in foot-pounds. A foot-pound is the work done in lifting a pound up one foot. The work done by a force of one dyne when it acts through one centimeter is called an *erg*.

Quantities of electricity are often expressed in terms of a unit called

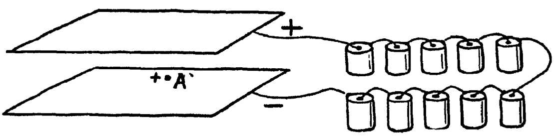

FIG. 31

the electrostatic unit of electricity which may be defined in the following way. If two equal charges of electricity, placed one centimeter apart, in a vacuum, repel each other with a force of one dyne then each charge is one electrostatic unit. This unit is very small and the units used in practical work are much larger. For example, an ampere is a current of 3,000,000,000 electrostatic units per second.

If two parallel metal plates are connected to a battery of dry cells as in Fig. 31 so that one is charged positively and the other negatively there will be an electric field in the space between the plates.

If a small body A charged with positive electricity is put between the plates it will be attracted by the lower negative plate and repelled by the upper positive plate so that there will be a force on it vertically downwards.

The strength of the electric field between the plates is taken to be equal to the force, expressed in dynes, on the small body when the charge on it is one unit of charge. If the distance between the plates is d centimeters and the force on the small body is F dynes, then the work required to move the small body, with its unit charge, from the lower plate to the upper plate will be Fd ergs. This work is called the *potential difference* between the two plates. It is usually denoted by P.D. If it is just one erg then the P.D. is one electrostatic unit of P.D. The practical or engineering unit of P.D. is called a *volt* and is such that 300 volts are equal to one electrostatic unit of P.D. One dry cell gives 1.5 volts, so that the 10 cells, in Fig. 31, will give 15 volts or 1/20 of an electrostatic unit of potential difference between the plates.

If the two plates are connected by a wire there will be a current through the wire. Let us suppose that there are 200 cells in the battery so that the P.D. between the plates is 300 volts or one electrostatic unit of P.D. If the current in the wire is one ampere this means that 3,000,000,000 electrostatic units of electricity flow from one plate to the other in one second. The work done on the electricity in one second will then be 3,000,000,000 ergs because when the P.D. is one electrostatic unit of P.D. the work required to take one electrostatic unit of charge across is one erg.

If the P.D. was only one volt instead of 300 volts the work done on the electricity when a current of one ampere was flowing through the wire would be 300 times smaller or 10,000,000 ergs per second. This work or energy appears as heat in the wire, which gets hot.

Thus a current of one ampere flowing across a P.D. of one volt gives ten million ergs of work or energy per second. This rate of doing work is called a *watt*. One horsepower is equal to 33,000 foot-pounds per minute which is equal to 746 watts.

A current of C amperes flowing across a P.D. of P volts gives CP watts or CP/746 horsepower. Thus if the P.D. is 110 volts then a current of 746/110 or 6.78 amperes will give one horsepower theoretically. In practice about 7 amperes at 110 volts is required for a one-horsepower motor because some of the energy is wasted in the motor.

A kilowatt is 1000 watts and a kilowatt-hour is a kilowatt for one hour. Electrical power is usually sold in kilowatt hours. The electricity comes into our houses through one wire and goes out through another wire, so that we don't get any, but we pay for the power.

Magnetic quantities are defined in much the same way as electric quantities. The unit magnetic pole is such that two of them, one centimeter apart in a vacuum, repel each other with a force of one dyne. The strength of a magnetic field is taken to be equal to the force in dynes on a unit pole put in the field.

It is found that if an electric charge e is moving along with a velocity v in a magnetic field of strength H which is perpendicular to the velocity v, then there is a force on the charge equal to Hev/c where c denotes the velocity of light. This force is in a direction perpendicular to both v and H. There is no force on a charge at rest in a magnetic field.

We are now in a position to consider how the charge per unit weight of electrons can be determined with the apparatus shown in Fig. 2 on page 12.

If the P.D. in electrostatic units between the tube T and the filament L is denoted by P, then the work done on an electron with charge e as it goes from the filament to the tube will be Pe ergs. If m denotes the weight or mass of the electron in grams and v its velocity, then its energy of motion or kinetic energy is equal to $\frac{1}{2}mv^2$. We have therefore $Pe = \frac{1}{2}mv^2$. Also if there is an electric field of strength F between the two plates A and B then there will be a force Fe on the electron due to this field. If a magnetic field of strength H is produced between the plates it will give a force on the electron equal to Hev/c. By adjusting the electric and magnetic fields the force Fe may be made equal and opposite to Hev/c so that the electrons are not deflected as they go between the plates. We have then $Fe = Hev/c$ so that $v = cF/H$. The equation $Pe = \frac{1}{2}mv^2$ then gives $e/m = v^2/2P$. By measuring F, H, and P the value of e/m or the charge per unit weight can therefore be obtained.

For example, if P is equal to 3000 volts or 10 electrostatic units, then it is found that the velocity v given by $v = cF/H$ is equal to 3.26×10^9 centimeters per second or about one-ninth that of light which is 3×10^{10} centimeters per second. e/m is then equal to $\frac{1}{2}\frac{(3.26 \times 10^9)^2}{10}$ or 5.3×10^{17} electrostatic units of charge per gram.

APPENDIX 2

Determination of the Charge on Droplets of Oil or Water

If a very small spherical droplet of oil or water falls through the air with a constant velocity v then its weight w, *expressed in dynes,* is given by the equation

$$w = 18\pi \sqrt{\frac{(sv)^3}{2Dg}}$$

In this equation D is the difference between the density of the droplet and that of the air, g is the acceleration of gravity, that is to say the velocity acquired by a freely falling body in one second, and s is the viscosity of the air.

This equation was first obtained by Sir G. G. Stokes in 1851.

Now suppose such a droplet is slowly falling through the air between two horizontal metal plates one above the other. Let there be a charge of electricity E on the droplet and suppose that the plates are connected to a large battery so that there is a vertical electric field of strength F between them.

This field will produce a force on the droplet equal to FE. The velocity with which the droplet moves through the air is proportional to the force acting on it.

When there is no electric field the force driving it is just its weight w and in the field it is w — FE assuming that the force FE is directed upwards. If v′ is the velocity with which the droplet falls in the field then

$$\frac{v'}{v} = \frac{w - FE}{w}.$$

This equation gives $E = (w/F)\ (1 - v'/v)$ so that E can be found by measuring v and v′.

The theory of this method of finding the charge on a small droplet was given by the writer in 1903.

It is easy in this way to measure a charge of one thousand millionth part of an electrostatic unit, which is a much smaller charge than can be measured in any other way.

APPENDIX 3

Aston's Positive Ray Apparatus

Aston's apparatus for measuring the weight per unit charge of positive rays is shown in Fig. 32. B is a large glass bulb with a tube connected to it at D. In this tube there is a small quartz bulb supported by a quartz rod and a wire sealed through the glass at A. Opposite D at C the bulb is connected to another tube which is almost closed by an aluminum plug with a narrow slit S in it. This tube also contains a second plug and slit at S′ and two parallel metal plates supported by wires E and F. The other end of this tube is connected to a box K through a stop-cock G. The part M of the box is put between the circular pole pieces of a large magnet so that a strong magnetic field perpendicular to the plane of the paper can be produced inside the circle M. A long narrow photographic plate can be put in the box

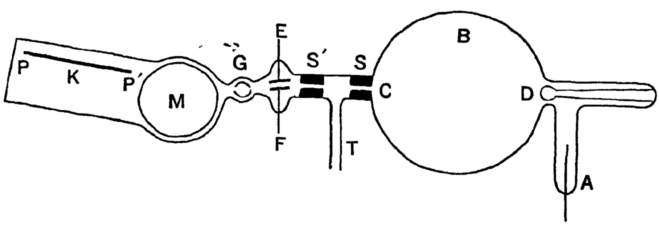

FIG. 32

at PP′. There is a side tube T between the slits S and S′ which is connected to a pump with which a good vacuum is produced in the tube and box. A very slow stream of the gas to be examined is allowed to flow into the bulb B through a small side tube not shown in the figure. The gas escapes through the slit S and is removed by the pump. In this way a very small gas pressure can be maintained in the bulb B and a good vacuum in the rest of the apparatus. The plug C is charged negatively and the wire A positively, to a potential difference of about 50,000 volts.

A discharge then passes through the gas in the bulb and the plug C emits electrons which strike the quartz bulb at D. The object of this bulb is to prevent the electrons striking the glass and melting it.

The electrons from C knock electrons out of atoms of the gas and the positively charged atoms so formed are attracted by the plug and some of them go through the two slits S and S′ as a narrow stream of positive rays. This stream goes between the plates E and F which are charged so as to deflect it downwards. Part of the deflected stream is admitted into the box K through the stop-cock G. The width of the stream admitted can be adjusted by means of this stop-cock.

The stream then goes through the magnetic field which deflects it upwards so that it falls on the photographic plate PP′.

We can show that by properly arranging the electric and magnetic fields and the photographic plate, the point on the plate at which the rays strike it can be made to be independent of the velocity of the rays and so to depend only on the weight per unit charge.

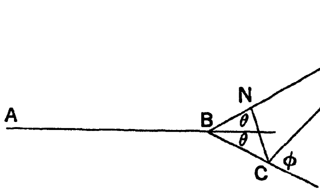

FIG. 33

In Fig. 33 let AB be the direction of the positive rays before they enter the electric field. Let this field be at B and let it deflect them so that they go along BC and are deflected down through an angle θ. Let the magnetic field be at C and let it deflect the rays along CD through the angle ϕ.

As the rays go through the electric field they are acted on by a sideways force FE where F is the field strength and E the charge on each particle. This force gives them a sideways velocity proportional to the time during which it acts, which is inversely as v the velocity of the rays. The distance they are deflected sideways is proportional to the sideways velocity and to the time they move with it so the sideways deflection is inversely as the square of the velocity v. Hence we have $\theta = A/v^2$ where A is a constant. The sideways force due to the magnetic field is proportional to v, so that $\phi = B/v$ where B is another constant.

Now consider rays with a slightly different velocity v′ and let θ' and ϕ' be the angles through which they are deflected, then we have $\theta' = A/v'^2$ and $\phi' = B/v'$. If b is the distance BC, the separation of the two rays at C will be $(\theta' - \theta)b$. Also at a distance r from C along CD the separation will be

$$(\theta' - \theta)b + (\theta' - \theta)r - (\phi' - \phi)r$$

Now $\theta' - \theta = A/v'^2 - A/v^2$ so that $(\theta' - \theta)/\theta = v^2/v'^2 - 1$. If $v = v' + s$ where s is small, then v^2/v'^2 is equal to $(v'^2 + 2v's)/v'^2$ or

$$1 + 2s/v'$$

so that

$$(\theta' - \theta)/\theta = 2s/v'.$$

In the same way

$$(\phi' - \phi)/\phi = s/v'.$$

Hence

$$(\phi' - \phi)/\phi = 2(\theta' - \theta)/\theta$$

or

$$\phi' - \phi = 2(\phi/\theta)(\theta' - \theta)$$

so that the separation is

$$(\theta' - \theta)(b + r - 2r\phi/\theta).$$

In order to have no separation it is therefore only necessary to arrange the apparatus so that $b + r - 2r\phi/\theta = 0$ or so that $r = b/(\phi/2\theta - 1)$.

Draw a line BD making an angle equal to θ with AB produced and let it cut CD at D. Also draw CN perpendicular to BD.

Let the angle CDB be denoted by α and let $CD = q$. We have then, since α is a small angle, $CN = q\alpha$. Also $CN = 2b\theta$ since θ is really also small although it is drawn rather large to make the figure clearer.

Hence $q\alpha = 2b\theta$.

But we have also $\phi = \alpha + 2\theta$, so that $q(\phi - 2\theta) = 2b\theta$ which gives $q = 2b\theta/(\phi - 2\theta)$ or $q = b/(\phi/2\theta - 1)$.

But this is just the value of r which makes the separation of the two rays zero. Thus it appears that if the photographic plate is put along BD there will be no separation of the two rays, with velocities v and v', at it. All the positive rays which have the same weight per unit charge will then strike the plate at the same point.

APPENDIX 4

C. T. R. Wilson's Cloud Apparatus

Fig. 34 shows a form of C. T. R. Wilson's cloud chamber designed by G. L. Locher. AB is a glass cylinder about 6 inches in diameter and 2 inches high. It is closed at the top by a glass plate C. DEFG is a metal box open at the top on which the glass cylinder rests. The box is made in two halves bolted together with a sheet of reinforced rubber

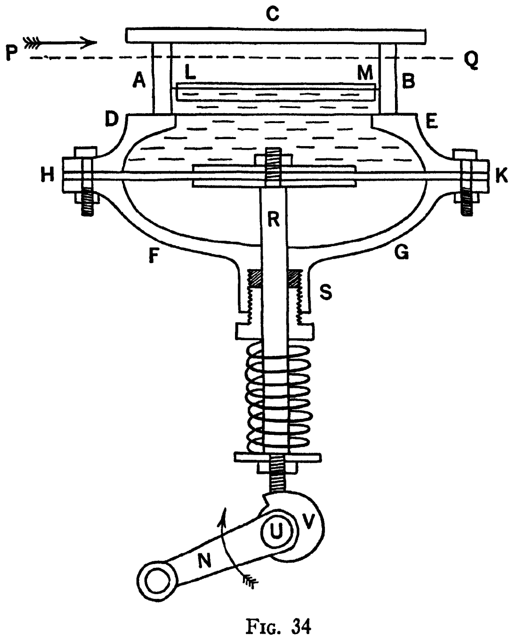

FIG. 34

HK between them. The rubber sheet is clamped between two metal disks on to a steel rod R which goes out of the box through a stuffing box S. The lower end of the rod rests on a cam V which can be rotated by turning the handle N attached to the shaft U. The box above the rubber sheet contains water on which a wooden disk LM floats. The rod is pressed against the cam by a strong spring T. The space between the float LM and the glass plate C contains moist air or some other gas. When the handle is turned the cam allows the rod to drop suddenly so that the rubber sheet is suddenly pulled down about 1/4 inch. This suddenly lowers the level of the water and float, so that the volume of the moist air above the float is suddenly increased. The sudden expansion cools the air and causes the water vapor to condense on any ions or dust particles which are present in it.

If the handle is turned slowly round and round the cam slowly raises the rod and lets it drop once in each complete turn.

The air in the cylinder can be strongly illuminated by a flash of light from a mercury arc and the tracks formed can be photographed in a camera placed above the glass plate C.

The first time the air is expanded a dense fog is produced due to condensation of the water vapor on dust particles in the air. This fog is allowed to settle, so removing the dust. After a few expansions the air becomes dust free, and then only a few large water drops which fall rapidly are produced at each expansion.

If a narrow ray of X-rays is passed through the chamber along the dotted line PQ, just before an expansion, then the water vapor condenses on the ions produced by the electrons knocked out of atoms by the X-ray photons. Each ion gives a small droplet so that the electron tracks are made visible and can be photographed. The tracks only last for a very short time and so must be photographed immediately after the expansion.

APPENDIX 5

Theory of Young's Experiment

In Fig. 35 AB is a metal screen with a narrow slit in it at Q. This slit is illuminated by a source of light S. CD is another screen with two slits in it at R and T. EF is a white screen on which the light from the two slits falls. At the point O the distances RO and TO are equal so that the two trains of light waves from R and T are in step and the crests and troughs coincide, so producing a strong light at O. Now consider a point P at which RP and TP are unequal. If the difference TP — RP is any whole number of wave lengths the two trains of waves at P will be in step and give strong light, but if TP — RP is any whole number of wave lengths plus one-half of a wave length, then the two

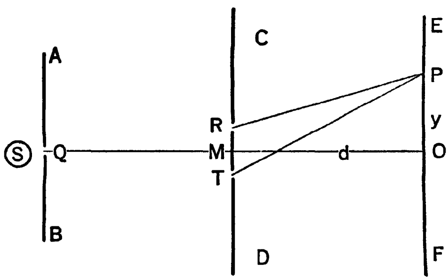

FIG. 35

trains will be out of step and will destroy each other so that there will be no light at P.

Let $OP = y$, $MO = d$, and $MR = MT = s$.

Then we have $RP^2 = d^2 + (y - s)^2$ and $TP^2 = d^2 + (y + s)^2$.

Hence $TP^2 - RP^2 = (y + s)^2 - (y - s)^2$.

Therefore $(TP - RP)(TP + RP) = 4ys$.

Now TP and RP are both nearly equal to $MO = d$, since y is small, so that $TP - RP = 4ys/2d = 2ys/d$.

If l denotes the wave length of the light and if P is at a bright band on the screen, then $TP - RP = nl$ where $n = 0$ or 1 or 2 or 3 or 4 and so on. Hence we have $nl = 2ys/d$.

The distance between any two adjacent bright bands is equal to y/n

so that if this distance is put equal to D, then $l = 2sD/d$. 2s is the distance between the two slits so that l can be found by measuring 2s, D and d.

For example, if $d = 72$ inches, $2s = 1/20$ inch, then it is found that $D = 1/25$ inch with red light. This gives $l = 1/20 \times 1/25 \times 1/72$ or 1/36000 inch.

APPENDIX 6

Reflection of X-rays from a Crystal

In Fig. 36 the horizontal lines are supposed to represent the equidistant layers of atoms in a crystal. RA, SB, and the other parallel lines represent the path of the incident X-rays which are reflected as shown along AT, BU, and the other parallel lines. The X-ray waves are perpendicular to the paths and so all the reflected beams must be imagined to overlap each other. Draw AN and AM perpendicular to SB and BU. The path SBU is longer than RAT by the distance NB + BM.

In order for all the reflected beams to be in step so as not to destroy each other it is necessary that this distance NB + BM and all the

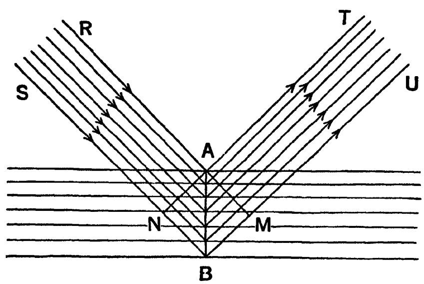

FIG. 36

similar path differences should be equal to whole numbers of wave lengths. Let the angle between the incident X-rays and the surface of the crystal be denoted by θ. Then the angle between the reflected rays and the surface is also equal to θ.

The angle θ is equal to the angles NAB and MAB so that if the distance between adjacent layers of atoms is d, then AB = 6d and NB + BM is equal to 2NB or to 2AB sin θ which is equal to 2 $\times$ 6d sin θ. The increase of path from one layer of atoms to the next is therefore 2d sin θ. If l denotes the wave length and n a whole number, then if nl = 2d sin θ all the reflected beams of X-rays will be in step and there will be a strong reflected beam.

For example, when palladium X-rays are reflected from a crystal of potassium chloride it is found that there is strong reflection when the angle θ is equal to 5° 23′ or 10° 49′ or 16° 20′. The sines of these angles are equal to 0.0938, 0.1877, and 0.2813, which are very nearly as 1 : 2 : 3.

This shows that $n = 1$ for $\theta = 5° 23'$, $n = 2$ for $\theta = 10° 49'$, and $n = 3$ for $\theta = 16° 20'$.

The wave length l of the X-rays is therefore equal to $0.0938 \times 2d$ or $\frac{1}{2} \times 0.1877 \times 2d$ or $\frac{1}{3} \times 0.2813 \times 2d$ which are very nearly equal.

The value of d in this case is 3.16×10^{-8} centimeter so that $l = 0.592 \times 10^{-8}$ centimeter or 2.33×10^{-9} inch.

APPENDIX 7

Emission of Electrons in the Photo-electric Effect

The energy of X-rays emitted by an X-ray tube is found to be about one per cent of the electrical energy supplied to the tube. If we suppose that the potential difference applied to a tube is 100,000 volts and the current through it one-tenth of an ampere, then the power expended is 10 kilowatts or 10^{11} ergs per second, since a watt is ten million ergs per second. One per cent of this is one thousand million ergs per second. Suppose the X-rays from the tube are allowed to fall on a metal plate at a distance of 100 centimeters from the tube. The X-ray waves at 100 centimeters from the tube are hemispheres of area $2\pi\,100^2$ or about 60,000 square centimeters. The energy of the X-rays which can be supposed to be absorbed by an atom or electron is the amount falling on an area of about $1/10^{16}$ of a square centimeter.

The energy absorbed by one electron in a second is therefore not more than $10^9/6 \times 10^4 \times 10^{16}$ or $1/6 \times 10^{11}$ ergs. The energy of the electrons which these X-rays cause to be emitted is equal to the P.D. on the X-ray tube multiplied by the charge e on one electron or 100,000 e/300. The charge e is equal to $5/10^{10}$. The energy of the electrons is therefore $100{,}000 \times 5/300 \times 10^{10}$ or $1/6 \times 10^6$ erg. It will therefore take $6 \times 10^{11}/6 \times 10^6$ or 100,000 seconds, which is about 28 hours, for the electron to get the energy with which it is emitted. But it is found that the electrons are emitted immediately when the X-rays are started, so that it is clear that the wave theory is quite inadequate.

APPENDIX 8

The Theory of the Compton Effect

In Fig. 37 let AB represent the path of a photon which collides with an electron at B and bounces off along the direction BC making an angle θ with AB produced. The electron is set moving along the direction BD with a velocity v.

Produce AB to E and make BE equal to BC. Draw BN perpendicular to CE so that BN bisects the angle θ.

The photon only loses a very small fraction of its energy and momentum in the collision, so that we may take BE to represent its momentum before the collision and BC that after the collision.

To change BE into BC it is necessary to go from E to C, so that EC will represent the momentum lost by the photon.

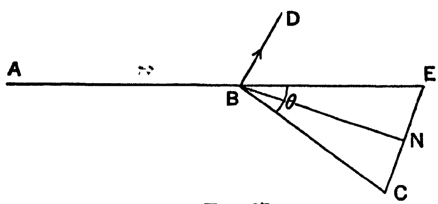

FIG. 37

The momentum of a photon is equal to h/l where h is Planck's constant and l the wave length of the photon. Hence if BE represents the momentum h/l before the collision, and EC that lost, then the momentum lost is equal to $(h/l) \times EC/BE$. But $EC = 2EB \sin \theta/2$ so that the momentum lost is equal to $(2h/l) \sin \theta/2$. If m denotes the weight or mass of the electron then its momentum is mv so that, since the momentum lost by the photon is equal to that gained by the electron we have

$$mv = (2h/l) \sin \theta/2.$$

The kinetic energy of the electron is equal to $\frac{1}{2} mv^2$ or $(mv)^2/2m$ which is therefore equal to $((2h/l) \sin \theta/2)^2/2m$.

The energy lost by the photon will be equal to that gained by the electron. The energy of a photon of frequency n is equal to hn or to hc/l where c is the velocity of light. If the wave length before the collision is l, and after the collision l′, then the energy lost is hc/l — hc/l′ so that $hc/l - hc/l' = ((2h/l) \sin \theta/2)^2/2m$. This gives $(l' - l)/ll' = (2h/mcl^2) \sin^2 \theta/2$ but l and l′ are nearly equal and $2 \sin^2 \theta/2 = 1 - \cos \theta$ so that finally $l' - l = (h/mc)(1 - \cos \theta)$. Putting in the known values of h, m, and c we get $l' - l = 0.0242 \times 10^{-8} (1 - \cos \theta)$. If $\theta = 90°$ then $\cos \theta = 0$ so that $l' - l = 0.0242 \times 10^{-8}$ centimeter. Thus, for example, X-rays of wave length 2×10^{-8} would be changed to rays of wave length 2.0242×10^{-8} when scattered through a right angle.

APPENDIX 9

Theory of Wave Groups

If a large number N of trains of waves all of the same height but with slightly different wave lengths are all moving along the same direction, they will form a group of waves. At the center of the group all the trains are in step so that all the crests coincide and the waves are N times higher than those of one of the trains.

As we go away from the center of the group, owing to the wave lengths being different, the trains of waves get more and more out of step, so that at some distance from the center they interfere and destroy each other. The group therefore has a more or less definite length L. This group length L is determined by the range of wave lengths in the trains of waves. Let us suppose that the wave lengths are all between l and l′ and that l′ is slightly greater than l.

Two equal trains destroy each other when one is half a wave length ahead of the other. The N trains can destroy each other in pairs if the train with the shortest waves is a whole wave length ahead of that with the longest.

Let the mean wave length $\frac{1}{2}(l + l')$ be denoted by $\bar{l}$. Then if $\frac{1}{2}L = n\bar{l}$ we must have $\frac{1}{2}L = (n + \frac{1}{2})l$ and $\frac{1}{2}L = (n - \frac{1}{2})\,l'$. These equations give $\bar{l}/L = 1/2n$ and $n\ (l' - l) = \frac{1}{2}(l + l') = \bar{l}$ or $(l' - l)/\bar{l} = 1/n$ so that $\bar{l}/L = (l' - l)/2l$. But $\frac{1}{2}(l' - l) = l' - \bar{l}$ so that $\bar{l}/L = (l' - \bar{l})/\bar{l} = (\bar{l} - l)/\bar{l}$.

Now as we have seen, the wave length is equal to Planck's constant h divided by the momentum. The momentum is the product of the weight m and the particle velocity v. We have therefore $l' = h/mv'$ and $\bar{l} = h/m\bar{v}$ so that $\bar{v} - v' = (h/m)\,(1/\bar{l} - 1/l') = (h/m)\,(l' - \bar{l})/\bar{l}\,l'$. We may put $\bar{l}\,l' = \bar{l}^2$ since l and l′ are nearly equal. We have therefore $\bar{v} - v' = (h/m)\,(l' - \bar{l})/\bar{l}^2$. But $(\bar{l} - l)/\bar{l}^2 = 1/L$ so that $\bar{v} - v' = h/mL$.

$\bar{v} - v'$ may be called the uncertainty in the particle velocity and L is the uncertainty in the position, so that the product of the two uncertainties is given by $(\bar{v} - v')\ L = h/m$.

APPENDIX 10

Uncertainty of Observations with a Microscope

The uncertainty in the position of the electron is equal to the wave length l of the light used. The uncertainty in the momentum of the electron may be put equal to the momentum of the photons, which is equal to h/l. Thus the product of these two uncertainties is equal to l $\times$ h/l or just h. The uncertainty in the velocity of the electron is equal to that in the momentum divided by the weight or mass m. The product of the uncertainty in the velocity and that in the position is therefore equal to h/m.

APPENDIX 11

Uncertainty for Narrow Beam of Electrons

In Fig. 38 let AB be a metal screen with a small hole of diameter d in it at O. Also suppose a stream of electrons falls on the plate in a direction perpendicular to its surface, with velocity v. Consider an electron which has just got through the hole.

The uncertainty in its position in the vertical direction is equal to the diameter of the hole. After passing through the hole the electron waves diverge to an extent depending on the wave length l and the diameter of the hole. The wave intensity at a point P some distance

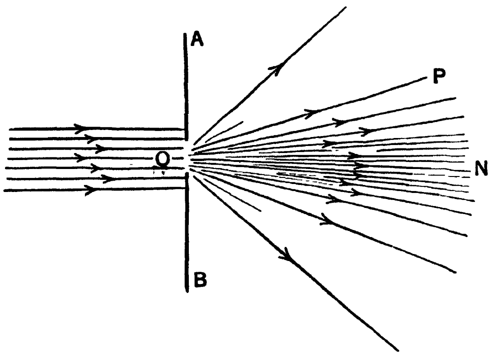

FIG. 38

from the hole will be zero if the wavelets from all the different parts of the hole interfere and destroy each other. For this to happen it is necessary to have a path difference of about one wave length between the greatest and shortest distances from points on the hole to P. This means that the inclination of OP to the normal ON must be not less than about l/d. The electrons may therefore have velocity components parallel to the screen anywhere between + vl/d and — vl/d.

The uncertainty in the velocity parallel to the screen may therefore be said to be equal to vl/d.

The product of the uncertainties in the position and velocity is therefore d $\times$ vl/d or vl. But l is equal to h/mv so that the product of the uncertainties is equal to h/m.

APPENDIX 12

Bohr's Theory of the Hydrogen Atom

In Fig. 39 let O be a proton and P an electron moving around O in a circular orbit of radius r with a velocity v. The electron and proton attract each other with a force equal to e^2/r^2 so that if m is the weight or mass of the electron we have $mv^2/r = e^2/r^2$. The kinetic energy of the electron is $\frac{1}{2}mv^2$ and its potential energy is $A - e^2/r$ where A is a constant so that its total energy E is given by $A - e^2/r + \frac{1}{2}mv^2$. The equation $mv^2/r = e^2/r^2$ gives $mv^2 = e^2/r$ so that $E = A - e^2/2r$. The angular momentum of the electron is equal to mvr and Bohr assumed this to be equal to $Nh/2\pi$ where N = 1, 2, 3, 4, etc. The equations $mv^2 =$

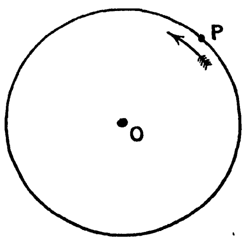

Fig. 39

e^2/r and $mvr = Nh/2\pi$ give $e^2/2r = 2\pi^2me^4/N^2h^2$ so that $E = A - 2\pi^2me^4/N^2h^2$. The different possible energies are got by putting N equal to different whole numbers 1, 2, 3, 4, etc. The energy of a photon of frequency n is hn, so that the frequencies of the light emitted are given by $hn = A - 2\pi^2me^4/N_2^2h^2 - A + 2\pi^2me^4/N_1^2h^2$ or $n = (2\pi^2me^4/h^3)(1/N_1^2 - 1/N_2^2)$.

The velocity of light c is equal to nl where l is the wave length so that the wave numbers are given by $1/l = (2\pi^2me^4/ch^3)(1/N_1^2 - 1/N_2^2)$. The experimentally found values of 1/l are given by the equation $1/l = 109678\ (1/N_1^2 - 1/N_2^2)$.

The following are the numerical values of m, e, c, and h.

$e = 4.77 \times 10^{-10}$ electrostatic units.
$m = 9.035 \times 10^{-28}$ gram.
$c = 2.998 \times 10^{10}$ centimeters per second.
$h = 6.55 \times 10^{-27}$ erg second.

Using these values we find $2\pi^2me^4/ch^3 = 109500$ which agrees very well with the experimentally found value 109678.

APPENDIX 13

Wave Mechanics Explanation of Bohr's Assumption

The wave length of electron waves is equal to h/mv where mv is the momentum of the electron. The length of a circular orbit of radius r is $2\pi r$ so that if this is a whole number N of wave lengths we have $Nl = 2\pi r$ but $l = h/mv$ so that $Nh/mv = 2\pi r$ or $Nh/2\pi = mvr$. But mvr is the angular momentum so that this equation shows that the angular momentum is a multiple of $h/2\pi$ as Bohr assumed.

APPENDIX 14

Calculation of Energy in Foot-Pounds

According to Einstein's theory the energy of a particle of weight or mass m is equal to mc^2 where c denotes the velocity of light. If m is expressed in grams and c in centimeters per second, then the energy mc^2 will be expressed in ergs. The erg is the work done by a force of one dyne when it acts through one centimeter and a dyne is a force which gives a gram a velocity of one centimeter per second in one second.

If m is expressed in pounds and c in feet per second, then the energy will be expressed in terms of a unit equal to the work done by a force, which gives a pound a velocity of one foot per second in one second, acting through one foot.

Now a force equal to the weight of a pound, that is, the force with which the earth attracts a pound gives the pound a velocity of 32 feet per second in one second. The foot-pound is therefore equal to 32 of the units of work mentioned in the preceding paragraph.

The energy mc^2 can therefore be expressed in foot-pounds by expressing m in pounds, c in feet per second, and dividing by 32.

APPENDIX 15

Neutrons

In 1930 Bothe and Becker discovered that the metal beryllium when bombarded by alpha-rays emits a very penetrating radiation. This radiation has been studied by several physicists, including Chadwick and others at the Cavendish laboratory. Chadwick has concluded that this radiation is of a new type consisting of particles having the same weight as protons but having no electric charge. It is suggested that these particles may consist of a proton and an electron very closely stuck together so that their charges neutralize each other. It is proposed to call these particles neutrons.

The intensity of this radiation is found to be reduced to one-half by passing through about 1/10 of a foot of lead so that it is much less penetrating than the cosmic rays.

The beryllium radiation when passed through any substance containing hydrogen knocks out high velocity protons.

When passed through moist air in a C.T.R. Wilson expansion chamber it produces no visible tracks itself and no electron tracks, but it does produce some short thick tracks believed to be due to nitrogen atoms with which the neutrons have collided.

It thus differs from X-rays and gamma rays which produce no tracks themselves, but knock electrons out of atoms, which make long thin tracks.

High velocity protons and alpha-rays also collide with atoms, giving them enough energy to make short thick tracks, but the protons and alpha-rays make easily visible tracks themselves.

A particle with no charge has no electric field around it, and so cannot influence another particle unless it makes a direct hit on it. In this respect it is like a photon. Photons with the great penetrating power of the beryllium radiation would be expected to produce very high velocity electrons just as gamma-rays do. Since the beryllium rays do not produce high velocity electrons they cannot be photons, but must be particles moving with velocities which are not large enough to give electrons enough energy to make tracks. Protons are such particles, but they make easily visible tracks themselves. A particle of the same mass as a proton but no charge, moving with about the same velocity as an alpha-ray would have enough energy to make a nitrogen atom give a short thick track, and to make protons give long tracks, but would not have enough velocity to make electrons give tracks. It is therefore suggested that the rays from beryllium may be neutrons.

APPENDIX 16

The S-diagram

The position of a point can be conveniently specified by giving its distances from three mutually perpendicular planes which are regarded as fixed. For example, we can fix the position of a point in a room by giving its height above the floor and its distances from two walls which enclose a corner. This is illustrated in Fig. 40. The position of the point P is fixed if we know PL, PM, and PN. The distance OP between the point P and the corner O is given by $OP^2 = PL^2 + PM^2 + PN^2$.

The lengths PL, PM, and PN are called the coordinates of the

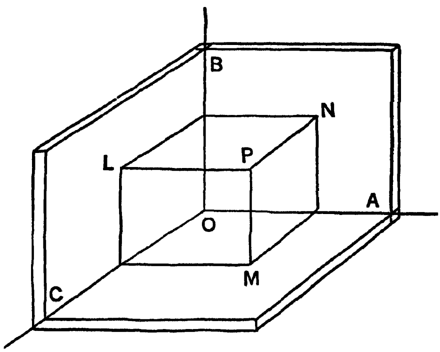

FIG. 40

point P and the three mutually perpendicular lines OA, OB, and OC are called the coordinate axes. The point O is called the origin of the coordinate axes.

If we imagine that the walls of the room are turned about O into any new position without moving O or P, we get a new set of axes with the same origin O. Let the coordinates of P with the new axes be PL′, PM′, and PN′. Then OP is given by $OP^2 = PL'^2 + PM'^2 + PN'^2$. With any set of axes the square of the distance of a point from the origin is equal to the sum of the squares of its three coordinates, and this sum has the same value for all sets of axes which have the same origin.

Let $OP = d$, $PL = x_1$, $PM = x_2$ and $PN = x_3$ so that x_1, x_2, x_3, are the coordinates of the point P and $d^2 = x_1^2 + x_2^2 + x_3^2$. If x_1', x_2' and x_3' are the coordinates of P in another set of axes having the same origin, then $d^2 = x_1'^2 + x_2'^2 + x_3'^2$.

Thus the position of a point relative to a set of axes can be specified by giving the values of its three coordinates x_1, x_2, x_3. Space is therefore said to be of three dimensions; it has height, length, and breadth.

If an event happens at a point then to specify the event we require its position, and also the time at which it happens. The event may therefore be fixed by giving the coordinates x_1, x_2, x_3 of the point at which it happened and the time t when it happened. To specify an event therefore requires four quantities.

Consider two events, one happening at the origin O at the time zero and the other at a point with coordinates x_1, x_2, x_3 at the time t. The absolute interval s between these events is given by the equation $s^2 = d^2 - c^2t^2$ or $s^2 = d^2 + c^2T^2$ where $T = it$. But $d^2 = x_1^2 + x_2^2 + x_3^2$ so that if we put $cT = x_4$ then $s^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$.

If the two events are observed by another observer who gets d′ and t′ instead of d and t, then $s^2 = d'^2 - c^2t'^2$ or $s^2 = d'^2 + c^2T'^2$ so that if $d'^2 = x_1'^2 + x_2'^2 + x_3'^2$ and $CT' = x_4'$ we get $s^2 = x_1'^2 + x_2'^2 + x_3'^2 + x_4'^2$.

We can imagine that x_1, x_2, x_3, and x_4 are the distances of the event from the origin measured along *four* mutually perpendicular lines as coordinate axes. This is impossible in actual space, but we can imagine a space in which four mutually perpendicular lines can be drawn through a point. Such a space is said to be of four dimensions. In this imaginary space of four dimensions we suppose we have a set of four perpendicular coordinate axes, and that the two events are represented by two points in this space. The first event is represented by the point at the origin O and the other by a point P with coordinates x_1, x_2, x_3, and x_4.

The distance OP between the two points is given by $OP^2 = s^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$. Just as in actual space, if the coordinate axes are turned, in any way, into a new position keeping the origin O fixed, the distance of the point P from the origin remains unchanged. If the coordinates of P with the new axes are x_1', x_2', x_3', and x_4', then $OP^2 = s^2 = x_1'^2 + x_2'^2 + x_3'^2 + x_4'^2$. Thus we see that the coordinates x_1, x_2, x_3, x_4, of the point P, or of the event which it represents, as found by one observer, and the coordinates x_1', x_2', x_3', x_4' found by another observer may be regarded as the coordinates of P relative to two different sets of axes having the same origin in the imaginary space of four dimensions. Rotating the axes about O does not change $OP = s$, and different observers all get the same value of s.

If we imagine all events represented by points in the four dimensional space, each point having four coordinates x_1, x_2, x_3, x_4, of which x_1, x_2, x_3, fix its position in actual space and x_4 fixes the time of the event, as seen by one particular observer, then the distances between the points will represent the absolute intervals between the events.

If x_1, x_2, x_3, x_4, are the coordinates of the one event and y_1, y_3, y_2, y_4, those of another, then the absolute interval s between them is given by $s^2 = (y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2 + (y_4 - x_4)^2$. The distance d between them is given by $d^2 = (y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2$ and the time t between them by $(y_4 - x_4) = cT = ict$. Thus $s^2 = d^2 + c^2T^2 = d^2 - c^2t^2$ which, as we have seen, has the same value for all observers. The four dimensional space with the points in it representing all events is Minkowski's s-diagram. The absolute intervals s between the events are the same for all observers. The coordinates x_1, x_2, x_3, x_4, of the events are different for different observers because they use different sets of axes relative to which they measure the positions and times of the events.

Consider a particle moving along a line or path. The arrival of the particle at the different points on its path may be regarded as a series of events which will be represented in the s-diagram by points which will lie on a line. Such a line in the s-diagram may be called an s-line. There will be an s-line for every particle. The s-lines of the particles in a solid body, like a stone for example, will all be close together and may be thought of as forming a sort of bundle of many lines in the four dimensional space.

Events which appear to an observer to occur at the same time all have the same value of $x_4 = ict$. x_4 increases with the time at the uniform rate ic. The events for which x_4 has the same value may be thought of as represented by points lying in a plane perpendicular to the x_4 axis. This plane moves along with the uniform velocity ic; it may be called the observer's time plane.

The points at which the s-lines of the particles are cut by this moving plane represent simultaneous events. Thus the s-diagram suggests that an observer is merely watching the points at which his time plane cuts the s-lines as it moves along.

Different observers have different time planes moving in different directions, and so get entirely different views of the succession of events.

The time planes are not really planes, they are three dimensional spaces containing all the lines perpendicular to the time axis through the point where it meets the time plane. That is to say, containing all actual space.

The events merely exist in the s-diagram, they do not happen. There is no space or time in the s-diagram, only the absolute intervals s. A particular observer introduces his particular space and time axes and so separates the events which exist in the s-diagram into a series of events happening one after another in different positions. This separation is purely relative and of no absolute significance. It is entirely different for different observers.

Minkowski therefore suggested that his s-diagram represents the universe as it really is, and that the separation of the absolute intervals into space and time by a particular observer is an illusion.

To this it may be replied that the s-diagram is purely imaginary. Actual space is of three dimensions and so far as we know a space of four dimensions is impossible. Also the s-diagram involves imaginary time so that it could not be drawn to scale even if there were a space of four dimensions to draw it in. The s-diagram is interesting and it shows up very clearly the relative character of space and time, but it is difficult to see how such a purely imaginary mathematical conception can be supposed to represent reality.

INDEX OF SUBJECTS

INDEX OF NAMES