MARCH 1 AND 15, 1942 PHYSICAL REVIEW VOLUME 61

Theory of Electrical Contact Between Solids

H. Y. FAN
Radio Research Laboratory, National Tsing Hua University, Kunming, China

(Received November 6, 1941)

The phenomena of electrical contact are examined in detail. The difference between the behaviors of metals and semi-conductors is pointed out. It is shown that the density of conduction electrons in a semi-conductor can be expected to change appreciably from its normal value which may be important for the explanation of certain phenomena.

INTRODUCTION

THE theory of electrical contact between two metals gives the well-known relation between the contact potential and the work functions of the metals. A theory has also been developed for the contact between a metal and a semi-conductor.[1] The theories are necessarily very crude as the subject concerns the surface phenomena. Since the surface is a region of discontinuity, the phenomena taking place there are much more difficult to handle than those inside the body and rigorous treatment would be very difficult. However, by examining the problem more closely we can get a clearer picture, bring out explicitly the assumptions and limitations of the existing theories and derive some additional information about the problem. This is the purpose of the present paper.

To simplify the problem we shall assume that the electron energy states are not affected by the surface. Various investigators[2] have shown that the surface gives rise to energy states within the forbidden band and other changes. To take such effects into consideration would greatly complicate the problem. As the number of surface states is small compared with the number of normal states we shall make the approximation of neglecting them.

SURFACE POTENTIAL BARRIER

The potential of an electron inside a solid is periodic with the period of the lattice, going through deep depressions in the neighborhood of each atomic nucleus. It rises sharply at the surface forming a potential barrier which keeps the electrons confined to the body. The potential energy is lowered inside a solid owing to the actions of the positive charges of the atomic nuclei and of the negative charges of the rest of the electrons. (Determining factors in the latter action are the exchange forces between electrons of the same spin and the electrostatic repulsive forces between pairs of electrons.) These actions of the charges continue outside the body. In the case of metals these actions reduce to the image force at sufficient distances from the surface. The above cause of the lowering of the potential energy inside a solid is a volume effect. In addition the potential barrier is affected by surface phenomena. The electron distribution around the surface ions is not symmetrical because of the fact that the charge about these ions on the inner side of the surface is different from that on the outside. This unsymmetrical distribution of negative electrons around positive ions results in a double layer which produces a potential difference across the surface. It has been shown by Bardeen[3] that for pure metals the surface double layer plays only a minor part.

Let us consider the relation between the potentials inside two solids in contact. When the bodies are far apart the relation is indicated by Fig. 1a. As they approach each other closer and closer the two parts of Fig. 1a simply shift toward each other (Fig. 1b) until at sufficiently

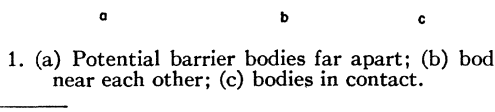

FIG. 1. (a) Potential barrier bodies far apart; (b) bodies near each other; (c) bodies in contact.

---

[1] A. H. Wilson, Proc. Roy. Soc. **A136**, 487 (1932).

[2] E. T. Goodwin, Proc. Camb. Phil. Soc. **35**, 205 (1939).

[3] E. Wigner and J. Bardeen, Phys. Rev. **48**, 84 (1935); J. Bardeen, Phys. Rev. **49**, 653 (1936).

close distances the distributions of electrons near the surfaces of the two bodies begin to change under mutual interaction of the charges of the two bodies. The potential barriers will then be changed. As a result we cannot simply shift the two parts closer to each other at such close ranges. If we neglect, however, the change of potential barriers owing to changes of charge distribution near the surface then the relation will be given by Fig. 1c. Actually such a picture does not correspond to equilibrium, which requires that the number of electrons entering one body from the other must be just balanced by the flow in the opposite direction. (Of course the energies carried by the electrons flowing in opposite directions must also balance each other.) This condition has not yet been taken into consideration.

CONTACT BETWEEN TWO METALS

The energy levels of solids are grouped into bands separated by gaps of forbidden energy. The lower bands are filled with electrons, each level having its full quota of two electrons. In metals there is a band only partially filled above the full bands. This is the conduction band, only electrons of this band participating in conduction. Upon coming into contact with another metal the full bands having their full quota of electrons cannot receive any more from the latter, and under equilibrium conditions no electrons should go over to the other metal from these bands. Therefore these bands can be disregarded in considering the exchange of electrons between the two metals and only the conduction bands need be considered. The electrons are distributed among the energy levels according to Fermi-Dirac statistics:

$$n=\frac{2}{e^{(E-\zeta)/KT}+1}. \tag{1}$$

The parameter ζ is the thermodynamic potential per electron and is determined by the total number of electrons which should be equal to the summation of n over all energy states. The choice of energy zero is arbitrary and it is convenient to choose the bottom of the conduction band. Figure 2a shows the potential barriers and conduction bands for two metals and Fig. 2b shows the two metals in contact. We have disregarded the condition for equilibrium. The bottom of the conduction band should actually not be at the same level clear up to the surface as indicated by horizontal line O_1 or O_2, since the potential near the surface is not the same as that inside the body. But this takes place near the very surface and we neglect it to simplify the picture and bring out the points we want to emphasize.

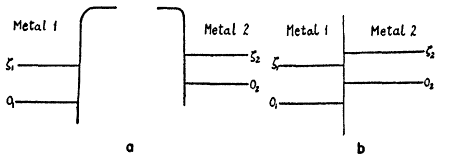

FIG. 2. (a) Potential barriers and conduction bands; (b) metals in contact.

Let us now consider the equilibrium condition. Frenkel[4] gave a good discussion using the free electron model. Bethe[5] gave substantially the same treatment, but he used a wave-mechanical treatment. We shall give Bethe's derivation with a few additions for clearer understanding. The wave function of an electron in a periodic field inside a solid is of the form:

$$\psi=e^{ikr}u, \tag{2}$$

where u is a periodic function with period of the lattice. The velocity of the electron in the direction x is

$$v_x=(2\pi/h)(\partial E/\partial k_x), \tag{3}$$

the electron energy being a function of $k(k_x, k_y, k_z)$. The number of electrons in body 1 arriving at the boundary surface (YZ plane) with a velocity component toward body 2 is

$$\frac{2}{(2\pi)^3}\int dk_x^1\int_{-\infty}^{\infty}dk_y^1\int_{-\infty}^{\infty}dk_z^1 \times\frac{2\pi}{h}\frac{\partial E_1}{\partial k_x^1}\frac{1}{e^{(E_1-\zeta_1)/KT}+1}. \tag{4}$$

The integration with respect to k_x^1 is over all values for positive v_x. This integral can be re-

---

[4] J. Frenkel, *Wave Mechanics, Elementary Theory.*

[5] A. Sommerfeld and H. Bethe, *Handbuch der Physik* Vol. 24/2.

written by changing the variable from k_x^1 to E_1

$$\frac{2}{(2\pi)^3}\int_{-\infty}^{\infty} dk_y^1 \int_{-\infty}^{\infty} dk_z^1 \int_{(E_1)}^{\infty} \frac{dE_1}{e^{(E_1-\zeta_1)/KT}+1}, \tag{5}$$

where (E_1) is the minimum value of E_1 for the set values of k_y^1 and k_z^1. Similarly the number of electrons arriving at the contact surface from body 2 heading for body 1 is

$$\frac{2}{(2\pi)^3}\int_{-\infty}^{\infty} dk_y^2 \int_{-\infty}^{\infty} dk_z^2 \int_{(E_2)}^{\infty} \frac{dE_2}{e^{(E_2-\zeta_2)/KT}+1}. \tag{5a}$$

Let the wave number (k_x^1, k_y^1, k_z^1) of an electron in body 1 change to (k_x^2, k_y^2, k_z^2) upon crossing the contact surface. To insure the same relation along the contact surface between the wave functions on both sides we must have $k_y^1 = k_y^2$ and $k_z^1 = k_z^2$. Since the electron energy should remain the same we have the following relation:

$$E_2(k_x^2, k_y^2, k_z^2) = E_1(k_x^1, k_y^1, k_z^1) - P, \tag{6}$$

where P is the difference between the energy zeros adopted for the two bodies. Whether an electron may enter a body or not depends in the first place upon the occupation of the particular energy level by electrons. The probability can be written in the form:

$$1 - \frac{\alpha}{e^{(E_2+P-\zeta_1)/KT}+1} \tag{7}$$

for electrons of energy E_2 in body 2 to enter body 1, and correspondingly

$$1 - \frac{\alpha}{e^{(E_1-P-\zeta_2)/KT}+1} \tag{7a}$$

for electrons of energy E_1 in body 1 to enter body 2. Furthermore electrons may suffer reflection at the boundary surface where two different periodic fields join together. We have to introduce a transmission coefficient which is a function of the wave numbers: $D(k_x, k_y, k_z)$ or $D(E, k_y, k_z)$. The number of electrons entering body 2 is

$$\frac{2}{h(2\pi)^2}\int_{-\infty}^{\infty} dk_y^1 \int_{-\infty}^{\infty} dk_z^1 \int_{(E_1)}^{\infty} D_{12}(E_1, k_y^1, k_z^1) \times \frac{1}{e^{(E_1-\zeta_1)/KT}+1}\left(1 - \frac{1}{e^{(E_1-P-\zeta_2)/KT}+1}\right) dE_1. \tag{8}$$

The number of electrons entering body 1 is

$$\begin{aligned}\frac{2}{h(2\pi)^2}&\int_{-\infty}^{\infty} dk_y^2 \int_{-\infty}^{\infty} dk_z^2 \int_{(E_2)}^{\infty} D_{21}(E_2, k_y^2, k_z^2) \times \frac{dE_2}{e^{(E_2-\zeta_2)/KT}+1}\left(1 - \frac{1}{e^{(E_2+P-\zeta_1)/KT}+1}\right)\\ &= \frac{2}{h(2\pi)^2}\int_{-\infty}^{\infty} dk_y^1 \int_{-\infty}^{\infty} dk_z^1 \times \int_{(E_2)+P}^{\infty} D_{21}(E_1-P, k_y^1, k_z^1) \times \frac{dE_1}{e^{(E_1-P-\zeta_2)/KT}+1}\left(1 - \frac{1}{e^{(E_1-\zeta_1)/KT}+1}\right).\end{aligned} \tag{8a}$$

The lower limit of integration for E_1 in (8) and (8a) should be the same. Let $(E_1) < (E_2) + P$. The electrons with energies $E_1 < (E_2) + P$ cannot, however, enter body 2 in which no electrons can have energy less than (E_2). These electrons must be reflected. The transmission coefficient obeys a general theorem which states in terms of our notation: $D_{12}(k_x^1, k_y^1, k_z^1) = D_{21}(k_x^2, k_y^2, k_z^2)$; it follows: $D_{12}(E_1, k_y^1, k_z^1) = D_{21}(E_1 - P, k_y^2, k_z^2)$. Thus we have

$$\begin{aligned}&\int_{-\infty}^{\infty} dk_y^1 \int_{-\infty}^{\infty} dk_z^1 \int_{E_1^0}^{\infty} D_{12}(E_1, k_y^1, k_z^1) \times \frac{dE}{e^{(E_1-\zeta_1)/KT}+1}\left(1 - \frac{1}{e^{(E_1-P-\zeta_2)/KT}+1}\right)\\ &= \int_{-\infty}^{\infty} dk_y^2 \int_{-\infty}^{\infty} dk_z^2 \int_{E_1^0}^{\infty} D_{12}(E_1, k_y^1, k_z^1) \times \frac{dE_1}{e^{(E_1-P-\zeta_2)/KT}+1}\left(1 - \frac{1}{e^{(E_1-\zeta_1)/KT}+1}\right),\end{aligned} \tag{9}$$

where E_1^0 is the common lower limit of integration. This reduces to

$$\int_{-\infty}^{\infty} dk_y^1 \int_{-\infty}^{\infty} dk_z^1 \int_{E_1^0}^{\infty} D_{12}(E_1, k_y^1, k_z^1) \times \left(\frac{1}{e^{(E_1-\zeta_1)/KT}+1} - \frac{1}{e^{(E_1-P-\zeta_2)/KT}+1}\right) dE_1 = 0. \tag{10}$$

This equation can be satisfied only with

$$E_1-\zeta_1=E_1-P-\zeta_2$$

or

$$P=\zeta_1-\zeta_2. \tag{11}$$

This means that the energy zero for body 2 (chosen at the bottom of its conduction band) should be higher than that chosen for body 1 by the amount $(\zeta_1-\zeta_2)$, therefore the levels ζ_1 and ζ_2 must be the same energy level when a common energy scale is taken for the two bodies. This is the condition for equilibrium.

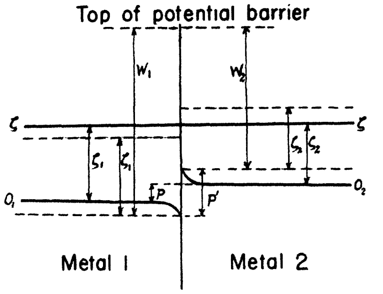

FIG. 3. Energy relations for metals in contact.

The same result is obtained directly from a general consideration of statistical equilibrium. When groups of electrons are in equilibrium their thermodynamic potentials must be equal. For the conduction electrons of the two bodies to be in equilibrium we must have

$$\zeta_1=\zeta_2=\zeta,$$

when a common energy zero is used for the two bodies. Difference between ζ_1 and ζ_2 must be due to the difference in the chosen energy zeros:

$$P=\zeta_1-\zeta_2.$$

Thus the same result as given by (11) is obtained. The first derivation is given because it shows the physical process more clearly.

Such condition as shown by Fig. 2b does not therefore correspond to equilibrium. More electrons will flow from body 2 to body 1. The former will be charged positively and the latter—negatively. We can look into the problem more closely by using the free electron model and consider the space distribution of the excess or deficiency of electrons and the average potential inside the two bodies. According to Fermi-Dirac statistics the density of electrons in space is given approximately by[6]

$$n=\frac{8\pi}{3h^3}[2m(\zeta-V)]^{\frac{3}{2}}, \tag{12}$$

where V is the potential energy of the electron at the point under consideration. The force driving the electrons from a point of higher to a point of lower potential is compensated by the force resulting from the difference in the electron densities at the two points. In metals the density of conduction electrons is very high and a space charge (excess or deficiency of electrons) enough to produce an appreciable potential field makes little change in the electron density. The action due to the difference of electron densities is, therefore, unimportant in metals. The space distribution of electrons is predominantly governed by electrostatics a fact which shows that charges on metals must be all on the surface (actually—near the surface). At the contact a sort of double layer will be formed. The electron potential inside metal 1 which is negatively charged will be raised and that of metal 2 will be lowered. The energy difference between the level ζ and the bottom of the band, being a measure of the electron density, remains the same inside each metal. At the contact surface we have $\zeta_1'>\zeta_1$ and $\zeta_2'<\zeta_2$ due to the excess of electrons in metal 1 and deficiency in metal 2. The energy difference at the contact between the bottoms of the conduction bands is approximately given by (see Fig. 3):

$$P'=W_1-W_2,$$

in which we have neglected the effects of the proximity of the two metals and of the charges resulting from electron flow on W_1 and W_2.[7] ζ_1' and ζ_2' adjust themselves by the metals' losing or gaining electrons near the surface until the condition for equilibrium (11) is satisfied. The electron potential and consequently the bottom of the conduction band of metal 1 have a convex form near the surface due to the negative charge of excess electrons and those of metal 2, concave. The detailed examination of

---

[6] J. Frenkel, *Wave Mechanics, Elementary Theory.*

[7] Actually these effects may be very important.

electron exchange at the surface gives in fact only

$$P' = \zeta_1' - \zeta_2', \tag{13}$$

whereas from the general consideration of statistical equilibrium it follows that we should also have

$$P = \zeta_1 - \zeta_2.$$

The work function is defined as the energy difference between an electron at rest outside of the surface and an electron at energy level ζ. The space charge near the surface due to excess or deficiency of electrons causes an additional potential difference between the interior of the metal and the outside. The work function will thus be changed by the contact. Actually such effect should be very small everywhere except at the contact. This can be made clear by the following consideration. The charges at the contact surface are responsible for producing the shift of potential

$$\begin{aligned} P' - P &= W_1 - W_2 - (\zeta_1 - \zeta_2) \\ &= (W_1 - \zeta_1) - (W_2 - \zeta_2) = \varphi_1 - \varphi_2. \end{aligned} \tag{14}$$

These charges are distributed over a very narrow region near the contact surface, this shift having to be produced within a very short distance, therefore the density of charges must be very high. A uniform layer of positive charge density ρ_1 of thickness a joining a uniform layer of negative charge density $-\rho_2$ of thickness b produces a potential difference between the two ends

$$\Delta V = \rho_1 a^2/2 + \rho_2 b^2/2. \tag{15}$$

To make $dv/dx = 0$ at the two ends the following relation has to hold

$$\rho_1 a = \rho_2 b.$$

Therefore

$$\Delta V/4\pi = \tfrac{1}{2}\rho_1 a(a+b) = \tfrac{1}{2}\sigma(a+b). \tag{16}$$

σ is the charge of each sign per unit area. Such a rough picture shows that the charge near the contact per unit area is

$$\sigma_c = \frac{P' - P}{\frac{1}{2}(a+b)} = \frac{\varphi_1 - \varphi_2}{\frac{1}{2}(a+b)}. \tag{17}$$

Consider on the other hand the charges at the surfaces A and B of Fig. 4. The potential difference between two points each immediately outside the surface of one metal is defined as contact potential and is given by the difference between the work functions of the two metals:

$$V_C = \varphi_1 - \varphi_2. \tag{18}$$

The potential difference between the surfaces A and B is

$$4\pi\sigma d = \varphi_1 - \varphi_2,$$

therefore

$$\sigma_A = \sigma_B = (\varphi_1 - \varphi_2)/4\pi d. \tag{19}$$

Since for metals charges must be nearly on the surface, the values of a and b of (17) must be very small. σ_C and σ_A will be of different orders of magnitude. The effect of σ_A on the potential between a point inside and a point immediately outside the metal must be negligible compared with such effect of σ_C which is measured by $(\varphi_1 - \varphi_2)$. It follows that the work functions of the metals are not changed appreciably by their coming into contact. The contact potential is quite accurately given by the difference between the original work functions of the metals.

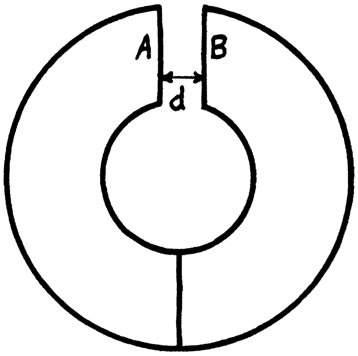

FIG. 4. Arrangement of two metals.

CONTACT BETWEEN METALS AND SEMI-CONDUCTORS (OR INSULATORS)

In an insulator the highest occupied band is completely full of electrons at 0°K and the next higher band is empty. At higher temperatures a small number of electrons will go over to the higher band. Depending upon whether the number of electrons going over to the higher band is large or small the substance is classified as insulator or semi-conductor. A semi-conductor has usually another source of electrons. Such

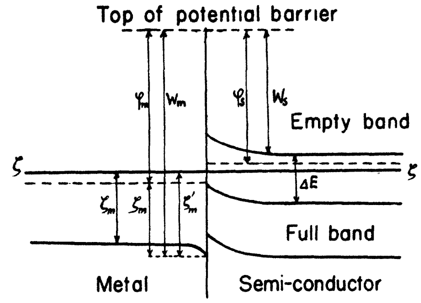

FIG. 5. Energy relations for a metal and semi-conductor in contact.

substance usually has impurity atoms with localized electron energy levels. These atoms may supply electrons to the empty band by being ionized or absorb electrons from a full band in case of electronegative atoms. Besides, their energy levels may provide a sort of stepping stone for the electrons of a full band to go over to the higher empty band.[8] For each kind of substance we have a definite relation between the numbers of electrons in the different bands and impurity levels. The numbers of electrons are all functions of the thermodynamic potential ζ. The latter can therefore be determined from this relation. Fowler[9] gives expressions of ζ for different types of semi-conductors in terms of the density of impurity atoms and the energy gaps between the different bands and the impurity levels. It is shown that as first approximation ζ lies half way between the two sets of energy levels, one full and one empty. The contact between a semi-conductor and a metal is usually treated by simply putting ζ's of the two substances on the same level.

As we have seen, in order to bring about equilibrium an initial unbalanced flow of electrons is required, resulting in the bodies becoming charged. The electron density in the semi-conductor will therefore be changed upon coming into contact with the metal. The normal relation between the number of electrons in the different bands will be upset and ζ will not be given by the usual expression. But it can be shown by a rough estimation that such effect is ordinarily small. For an intrinsic semi-conductor (no impurity) ζ obtained by equating the number of electrons in the conduction band and the number of holes in the next lower band is given by:[9]

$$\zeta = \tfrac{1}{2}\Delta E + KT \log C, \qquad (20)$$

where C is a constant and ΔE is the energy gap between the two bands. If the number of electrons in the conduction band should be increased to N times the number of holes in the lower band the increase in ζ will be:

$$\Delta\zeta = KT \log \sqrt{N}. \qquad (21)$$

With $N = 10$ and at $T = 300°\mathrm{K}$ this amounts to only 0.03 electron volt. This is small compared with the order of magnitude of ΔE (0.72 ev for Cu_2O).[10]

Although the effect of the contact on ζ itself is small the electron density in the conduction band which is normally small can be changed appreciably. Therefore the change in ζ should not be neglected when the electron density is being considered and it is not proper simply to equate ζ of the metal with the original ζ of the semi-conductor. Since there are only a few electrons in the conduction band of a semi-conductor their distribution is Maxwellian. For simplicity of treatment it is usually assumed that these electrons behave like free electrons with an effective mass. While this is not strictly correct[11] rigorous treatment would be too complicated. With this approximation the density of electrons is given by the usual formula for an electron gas:

$$n = 2\frac{(2\pi mkT)^{\frac{3}{2}}}{h^3} e^{\zeta/KT}, \qquad (22)$$

the potential energy of the electrons or the bottom of the conduction band being taken as zero. When the semi-conductor is in contact with a metal ζ is changed due to the gain or loss of electrons. By entering the expression for n in the exponential factor it is found a small change in ζ will cause considerable change in n, in agreement with the statement made above. The energy difference between the bottoms of the conduction bands is given by $(W_m - W_s)$, there-

---

[8] B. Gudden, Ergebnisse d. exakten Naturwiss., Vol. 13.

[9] R. H. Fowler, *Statistical Mechanics*, second edition.

[10] W. Juse and B. W. Kurtschatow, Physik. Zeits. de Sowjetumon **2**, 453 (1932).

[11] N. F. Mott and H. Jones, *Theory of Properties of Metals and Alloys*, p. 83.

fore as shown by Fig. 5:

$$\zeta = \zeta_m' - W_m + W_s. \tag{23}$$

The semi-conductor is charged positively or negatively depending upon whether $\varphi_m > \varphi_s$ or $\varphi_m < \varphi_s$. (Figure 5 shows the semi-conductor positively charged.) In the first case the electron density in the conduction band will be less than normal and the reverse is true for the second case. The excess of deficiency in electrons will not be concentrated near the surface as in metals. Because the numbers of electrons in the conduction band and holes in the lower band are small a semi-conductor can support a space charge and appreciable potential difference. The density of conduction electrons at the contact surface is given by (22). The space charge (conduction electrons and holes in the lower band) inside the body is determined jointly by Poisson's equation and Boltzmann's distribution law. With the additional condition of equilibrium between the electrons in the two bands the density of conduction electrons and the density of the holes in the lower band can be determined separately.

The space variation of the density of conduction electrons to be expected in semi-conductors may be important in the explanation of certain phenomena. Upon this is based Schottky's theory[12] of semi-conductor rectifiers which assumes the semi-conductor to be charged positively with region of low electron density at the contact.

---

[12] W. Schottky, Zeits. f. Physik **113**, 367 (1939).

MARCH 1 AND 15, 1942 PHYSICAL REVIEW VOLUME 61

Distribution in Angle of Protons from the D-D Reaction

H. P. MANNING, R. D. HUNTOON,* F. E. MYERS,† AND V. J. YOUNG
Department of Physics, New York University, New York, New York

(Received December 5, 1941)

The angular distribution of the disintegration protons from the D–D reaction has been investigated as a function of energy from 60 kev to 390 kev. The distribution was found to be well represented by $I(\theta) = 1 + A\cos^2\theta$ at any one energy. The value of A increases smoothly with bombarding energy over the range investigated.

INTRODUCTION

MEASUREMENTS of the relative yield of protons from the reaction

$$H^2 + H^2 = H^3 + H^1$$

as a function of the angle of the emitted proton with the line of approach of the deuterons have shown an anisotropy of the form

$$I(\theta) = 1 + A\cos^2\theta$$

in the center of mass system of coordinates. Previous investigators[1,2] have disagreed markedly on the value of A and its dependence upon energy.

The desired condition for angular distribution measurements where the anisotropy is energy dependent is that all observed disintegrations be due to impinging particles having the same energy at the time of impact. This condition is most nearly realized in a gas target. Without using foils thick enough to invalidate the use of a gas target, permanent gases require the use of a very small capillary and high speed pumping at high vacuum and at an intermediate pressure. Vapors can be handled with a larger capillary and a cold trap, refrigerated with liquid air. This allows higher counting rates and a great reduction in statistical error without loss in angular resolution. Thus, since the deuteron reaction with oxygen is highly improbable below

---

* Now at National Bureau of Standards, Washington, D. C.

† Now at Frankford Arsenal, Philadelphia, Pennsylvania.

[1] Huntoon, Ellett, Bayley and Van Allen, Phys. Rev. **58**, 97 (1940).

[2] Kempton, Browne, and Maasdorp, Proc. Roy. Soc. **A157**, 386 (1936); Haxby, Allen, and Williams, Phys. Rev. **55**, 1940 (1939); H. Neuert, Ann. d. Physik **36**, 437 (1939).