Les Houches 5 - 1955

Messiah, A. : Reactions nucleaires sans formation de noyau compose, 1955.....1-36

Ramsey, N.F. : Interaction of a nucleus with atomic and molecular fields, 1955.....37-95

Schwinger, J. : Quantum mechanics, 1955.....96-302

Uhlenbeck, G.E. : Notes on the theory of phase transitions, 1955 303-366

Uhlenbeck, G.E. : The statistical mechanics of non-equillibrium phenomena 367-444

REACTIONS NUCLEAIRES
SANS FORMATION DE NOYAU COMPOSE

Albert MESSIAH
Centre d'Etudes Nucléaires, Saclay

ECOLE d'ÉTÉ de PHYSIQUE THÉORIQUE

Ces notes ont été rédigées par P. LEHMANN, Centre d'Etudes Nucléaires, Saclay.

Université de Grenoble

Conférences données à
l'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie), France

Août 1955

E R R A T A

Page	Ligne	
6	2	$V = \sum_{i=1}^{Z} e^2 \left[\frac{1}{\lvert \vec{r} - \vec{r}_i \rvert} - \frac{1}{r} \right]$ (3)
7	9	$P_{\alpha \to \beta} = \frac{1}{\hbar^2} \left\lvert \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}(\Delta E)t} \langle \beta \lvert V(t) \rvert \alpha \rangle \, dt \right\rvert^2$ (10)
9	11	Lire ... $(\sin \frac{\theta}{2})^{-4}$...
10	2, 7	Lire Ψ au lieu de ψ
11	4	Lire $e^{i\vec{k}_\alpha \cdot \vec{r}_\alpha}$ au lieu de $e^{i\vec{k}_\alpha \cdot r_\alpha}$
13	10	Lire $e^{i\vec{k}_\alpha \cdot \vec{r}_\alpha}$ au lieu de $e^{i\vec{r}_\alpha \cdot r_\alpha}$ (Sous "Voieα")
	16	Lire $e^{-i\vec{k}_\beta \cdot \vec{r}_\beta}$ au lieu de $e^{-i\vec{k}_\beta \cdot r_\beta}$
15	3	Lire $\hbar\vec{k}_i$ et $\hbar\vec{k}_f$ l'impulsion initiale...
19	dernière	Lire $\varphi_c(\vec{r}_\mu)$ au lieu de $\varphi_c(r_\mu)$
20	3	Lire ... l'effet de la non-coïncidence
22	1	Lire ... de vecteur $\vec{K}$
23	dernière	Ajouter : De même : $\left(-\frac{\hbar^2 \Delta_n}{2M} - \frac{\hbar^2 q^2}{2M} \right) e^{i\vec{q} \cdot \vec{r}_n} = 0$
24	2	Lire $- e^{i\vec{q} \cdot \vec{r}_n}$ au lieu de $- e^{iq \cdot r_n}$
	15	Lire $j = \lvert j_A \pm 1/2 \rvert$ au lieu de $j = j_A \pm 1/2$
	16	Lire : La fonction d'onde $\psi_{jB}^{\mu_B}$
31	14	Lire : ... en partant de la formule de la page 28 au lieu de : en partant de la formule (8)
32	4	Lire ... traitées au lieu de ... traités.

Première conférence.-

REACTIONS NUCLEAIRES SANS FORMATION DE NOYAU COMPOSE

INTRODUCTION .

Le formalisme général des réactions nucléaires consiste à diviser l'espace en une région interne ou zone du noyau composé et une zone externe ou zone de voies. L'hypersurface (S) qui sépare ces deux zones est choisie de telle sorte que, dans la zone des voies, le système nucléaire considéré se scinde en deux noyaux (ou plus) dont la distance relative est suffisemment grande pour que leur interaction nucléaire puisse être négligé

Les sections efficaces se déduisant de la forme asymptotique de l'état stationnaire de collision, il suffit de résoudre l'équation de SCHRÖDINGER dans la zone des voies; ceci se fait sans difficulté lorsqu'on connaît les conditions aux limites sur la surface (S), c'est à dire la relation satisfaite par la fonction et sa dérivée normale sur cette surface On sait que cette relation s'exprime à l'aide de la matrice dérivée ou matrice R de WIGNER. Il n'est donc pas nécessaire de résoudre l'équation de SCHRÖDINGER dans la zone du noyau composé, du moment que l'on connaît la matrice R qui contient la totalité des renseignements internes indispensables à la solution du problème externe.

A priori, la position de la surface (S) est largement arbitrair et le formalisme général s'applique quelque soit la valeur du rayon de chaque voie, du moment qu'il est supérieur à un rayon minimum en deça duquel l'interaction entre les fragments nucléaires cesse d'être négligeab En fait, tout ce formalisme n'a d'utilité que dans la mesure où l'on peut, à une bonne approximation, représenter R par une expression simple et maniable, comme on le fait par exemple dans l'approximation à un niveau. La position exacte de la surface (S) joue alors un rôle important.

En pratique, on attribue aux rayons des voies leur valeur minimum. On imagine les noyaux comme des boules de matière nucléaire de rayon bien délimité et dont l'interaction devient négligeable (mise à part l'interaction Coulombienne pure) dès qu'elles cessent d'être en contact. Le rayon dans chaque voie est ainsi défini sans ambiguité. Soit (S_o) la surface (S) correspondante. Avec ce choix (S_o) que nous adopterons dans toute la suite, la zone du noyau composé est une zone où les nucléons évoluent en contact étroit les uns avec les autres et l'on peut espérer décrire cette évolution par des méthodes statistiques avec quelque chance de succès, justifiant ainsi les approximations faites habituellement sur la matrice R telles que l'approximation à un niveau. On dit souvent que la réaction nucléaire procède par formation du "noyau composé" (1) en

(1) Ne pas confondre le sens donné ici au terme "noyau composé" avec le sens plus général qui lui a été attribué jusqu'ici.

entendant par là que la particule incidente fusionne avec la cible au cours de la collision pour former dans une étape intermédiaire un "noyau composé" dans lequel les nucléons évoluent en contact étroit les uns avec les autres comme les nucléons d'un noyau ordinaire.

Si les noyaux étaient vraiment des boules de matière de rayon bien délimité, ce mécanisme par formation du "noyau composé" serait le seul possible. Mais certaines réactions nucléaires peuvent aussi se produire sans que la particule incidente ait pénétré en totalité ou en partie dans le noyau cible. Elles peuvent se classer en deux catégories :

(1) Excitation Coulombienne, dûe au fait que même à l'extérieur de (S_o), l'interaction entre les fragments nucléaires n'est pas une interaction Coulombienne pure dépendant uniquement de la coordonnée relative de leur centre de masse. Ainsi, un proton ou une particule α peut exciter un noyau sans y pénétrer. C'est le mécanisme principal d'excitation lorsque le noyau cible est lourd et l'effet de barrière important.

(2) Interaction nucléaire directe où quelques nucléons seulement participent à l'interaction à l'exclusion des autres. Les réactions de ce type les plus importantes sont les réactions de stripping du deuton, réactions dp (ou dn), où le neutron (proton) incident pénètre dans le noyau tandis que le proton (neutron) qui lui est lié dans le deuton continue son chemin sans subir d'interaction nucléaire notable. Dans cette catégorie, il faut

aussi ranger certaines réactions (np) qui s'expliquent qualitativement par l'interaction directe entre le neutron incident et le proton du noyau cible, alors que ce dernier se trouve à l'extérieur de la surface nucléaire proprement dite; un tel mécanisme est mis en évidence par le fait qu les sections efficaces sont plus importantes que l'effet de barrière ne le laisserait prévoir et par l'importance de l'émission de particules vers l'avant, caractéristique d'une contribution cohérente des grands moments cinétiques orbitaux, contrairement à ce qui se passe dans la formation de "noyau composé" où la distribution angulaire est nettement plus isotrope (et en tous cas symétrique par rapport à 90°).

Toutes les théories jusqu'ici proposées de ces phénomènes "d'interaction externe"consistent essentiellement à traiter l'interaction externe comme une perturbation. Une telle approximation est certainement justifiée dans le cas de l'excitation Coulombienne. Elle est plus douteuse dans les autres cas; elle conduit néanmoins à des distributions angulaires de stripping en surprenant accord avec l'expérience.

THEORIE SEMI-CLASSIQUE DE L'EXCITATION COULOMBIENNE.

(Cf. Ter Martirosyan Th. Eksper Fz. USSR 22, 284 (1952)

Bohr Mottelson Dan. Mat. Fys. Medd. 27, n° 16 (1953) p.162)

Le traitement de perturbation de l'excitation Coulombienne se simplifie grandement si l'on traite classiquement le mouvement du proton (ou de la particule chargée) incident(e).

Considérons un proton d'énergie $E = 1/2\ Mv^2$ se dirigeant vers un noyau N de charge Z. Soit ΔE l'énergie d'excitation du noyau après collision. Nous supposerons dans ce qui suit la distance classique d'approche minimum b du proton grande devant le rayon R du noyau

$$b = \frac{Ze^2}{E} \ll R \qquad (1)$$

ce qui nous autorise à négliger toute interaction nucléaire entre proton et noyau. Cette condition est bien réalisée dans le choc de protons de quelques MeV sur des noyaux moyens ou lourds.

L'Hamiltonien du système s'écrit alors :

$$H = H_N + \frac{p^2}{2M} + \frac{Ze^2}{r} + V \qquad (2)$$

où H_N est l'Hamiltonien du noyau, $(\frac{p^2}{2M} + \frac{Ze^2}{r})$ est celui du proton en interaction Coulombienne avec le noyau supposé ponctuel, V la différence entre l'interaction Coulombienne exacte et l'interaction ponctuelle $\frac{Ze^2}{r}$,

à savoir :

$$V = \sum_{i=1}^{Z} e^2 \left[\frac{1}{|\vec{r} - \vec{r}_i|} + \frac{1}{r} \right] \qquad (3)$$

En l'absence de V, l'Hamiltonien se sépare; la seule réaction possible est la diffusion Coulombienne pure dont la section efficace est donnée par la formule de RUTHERFORD :

$$\frac{d\sigma_0}{d\Omega} = \left(\frac{b}{4}\right)^2 \left(\sin \frac{\theta}{2}\right)^{-4} \qquad (4)$$

Introduisons les paramètres

$$\eta = \frac{Ze^2}{\hbar v} \qquad \xi = \frac{1}{2} \eta \frac{\Delta E}{E} \qquad (5)$$

Nous allons supposer que :

$$\eta \gg 1$$

$$\frac{\Delta E}{E} \ll 1 \qquad (7)$$

La condition (6) permet de traiter la diffusion Coulombienne dans l'approximation classique; le mouvement du proton est alors décrit par un paquet d'onde de dimensions négligeables dont le centre $\vec{r}(t)$ suit les lois du mouvement classique. Ce mouvement étant donné, le noyau est soumis à l'Hamiltonien

$$\mathcal{H} = H_N + V(t) \qquad (8)$$

où

$$V(t) = e^2 \sum_{i=1}^{Z} \left[\frac{1}{|\vec{r}(t) - \vec{r}_i|} * \frac{1}{r(t)} \right] \qquad (9)$$

est une perturbation dépendant du temps agissant sur les variables dynamiques du noyau et susceptible d'y induire des transitions dont la probabilité peut se calculer par la méthode des perturbations dépendant du temps.

Soit $|\alpha\rangle$ l'état initial du noyau, $|\beta\rangle$ son état final. La probabilité de transition $P_{\alpha \to \beta}$ est :

$$P_{\alpha \to \beta} = \frac{1}{\hbar^2} \left| \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}(\Delta E)t} \langle \beta | V(t) | \alpha \rangle \, dt \right|^2 \qquad (10)$$

d'où la section efficace d'excitation Coulombienne

$$\frac{d\sigma_{\alpha \to \beta}}{d\Omega} = \frac{d\sigma_0}{d\Omega} \times P_{\alpha \to \beta} \qquad (11)$$

Bien noter qu'en considérant le mouvement du proton comme donné, on néglige la réaction de l'excitation du noyau sur ce mouvement, approximation justifiée par la condition (7).

Pour calculer l'expression (10), on fait choix d'un axe polaire (qui sera pris comme axe de quantification des moments cinétiques des états $|\alpha\rangle$ et $|\beta\rangle$)et l'on développe V(t) en harmoniques sphériques.

$$V = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{+\ell} (-)^m Q_\ell^m T_\ell^{*m}(t) \qquad (12)$$

$$T_\ell^m(t) = \frac{4\pi Ze}{(2\ell+1)} \frac{Y_\ell^m(\Omega)}{r^{\ell+1}} \qquad (13)$$

$$Q_\ell^m = e \sum_{i=1}^{Z} r_i^\ell \, Y_\ell^m(\Omega_i) \qquad (14)$$

d'où :

$$P_{\alpha \to \beta} = \frac{1}{\hbar^2} \left| \sum_{\ell m} (-)^m S_\ell^{-m} \langle \beta J_\beta M_\beta | Q_\ell^m | \alpha J_\alpha M_\alpha \rangle \right|^2 \qquad (15)$$

avec

$$S_\ell^m = \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}(\Delta E)t} T_\ell^m(t) \, dt \qquad (16)$$

Les $(2\ell + 1)$ opérateurs Q_ℓ^m ($m = -\ell, \ldots +\ell$) sont les composantes de l'opérateur tensoriel $\underline{Q}^{(\ell)}$, moment 2^ℓ-polaire électrique. Les seuls éléments de matrice non nuls sont ceux qui satisfont aux règles de sélection de spin et de parité :

$$|J_\alpha - J_\beta| \leq \ell \leq J_\alpha + J_\beta \qquad m = M_\beta - M_\alpha \qquad \Pi_\alpha \Pi_\beta = (-)^\ell \qquad (17)$$

De plus :

$$\langle \beta J_\beta M_\beta | Q_\ell^m | \alpha J_\alpha M_\alpha \rangle = \frac{1}{\sqrt{2J_\beta + 1}} \langle J_\alpha \ell M_\alpha m | J_\alpha \ell J_\beta M_\beta \rangle \langle \beta \| \underline{Q}^{(\ell)} \| \alpha \rangle \qquad (18)$$

L'élément de matrice réduit $\langle \beta \| \underline{Q}^{(\ell)} \| \alpha \rangle$ intervient également dans l'expression de la vie moyenne de l'état $|\beta\rangle$, plus précisément dans sa

probabilité de transition à l'état $|\alpha>$ par émission γ 2^{ℓ}- polaire électrique. Une évaluation grossière des ordres de grandeurs des termes intervenant dans (15) montre que, toutes choses égales, la contribution des termes $2^{\ell+2}$-polaires est environ $(\frac{R}{b})^4$ fois plus petite que celle des termes 2^{ℓ}-polaire. On peut donc ne retenir que le terme de ℓ le plus bas compatible avec les règles de sélection, soit ℓ_o = soit $|J_\alpha - J_\beta|$, soit $|J_\alpha - J_\beta| + 1$.

Si en outre, les orientations de spins initiaux et finaux ne sont pas spécifiées, il faut de plus moyenner sur le premier et sommer sur le dernier, ce qui donne, tous calculs faits :

$$\frac{d\sigma_{\alpha\to\beta}}{d\Omega} = \left(\frac{b}{4}\right)^2 \left(\sin\frac{\theta}{2}\right)^{-4} \frac{|<\beta\| Q^{(\ell_o)} \|\alpha>|^2}{(2J_\alpha+1)(2\ell_o+1)} \frac{1}{\hbar^2} \left(\sum_m |S^m_{\ell_o}|^2\right)$$

Les intégrales $S^m_{\ell_o}$ dépendant de la trajectoire classique suivie par le proton, c'est à dire de la direction $(\Omega, \Omega + d\Omega)$ d'émission du proton. Elles doivent être calculées numériquement.

Deuxième Conférence.-

GENERALISATION DE L'APPROXIMATION DE BORN - APPLICATION A L'EXCITATION COULOMBIENNE.

Soit H l'hamiltonien d'un système de nucléons. Un état stationnaire de collision $\Psi^{(+)}(\alpha, \vec{k}_\alpha)$ d'énergie E est entièrement défini par la donnée de la voie d'entrée α et de l'impulsion relative $\hbar \vec{k}_\alpha$ dans cette voie. Nous nous limiterons aux voies à deux particules et, afin de simplifier l'exposé, nous supposerons ces particules neutres et sans spin; ces deux dernières restrictions peuvent être facilement levées. La forme asymptotique de $\Psi^{(+)}(\alpha, \vec{k}_\alpha)$ est la suivante :

dans la voie α $$\Psi^{(+)}(\alpha, \vec{k}_\alpha) \underset{r_\alpha \to \infty}{\sim} \psi_\alpha \left[e^{i \vec{k}_{\alpha 0} \vec{r}_\alpha} + f^{(+)}_{\alpha\alpha}(\Omega_\alpha) \frac{e^{i k_\alpha r_\alpha}}{r_\alpha} \right.$$

(onde plane + onde sortante) (I)

dans toute autre voie β $$\Psi^{(+)}(\alpha, \vec{k}_\alpha) \underset{r_\beta \to \infty}{\sim} \psi_\beta f^{(+)}_{\alpha\beta}(\Omega_\beta) \frac{e^{i k_\beta r_\beta}}{r_\beta}$$

(onde sortante pure)

$$H \Psi^{(+)}(\alpha, \vec{k}_\alpha) = E \Psi^{(+)}(\alpha, \vec{k}_\alpha) \qquad (1)$$

Si H est invariant par renversement du sens du temps, il existe

une autre solution stationnaire $\Psi^{(-)}(\alpha, \vec{k}_\alpha)$ de même énergie E (déduite de $\Psi^{(-)}(-\alpha, -\vec{k}_\alpha)$ par renversement du sens du temps) dont le comportement asymptotique est

dans la voie α $$\Psi^{(-)}(\alpha, \vec{k}_\alpha) \sim \psi_\alpha \left[e^{i\vec{k}_\alpha r_\alpha} + f^{(-)}_{\alpha\alpha}(\Omega_\alpha) \frac{e^{-ik_\alpha r_\alpha}}{r_\alpha} \right.$$

(onde plane + onde entrante) (II

dans $\beta \neq \alpha$ $$\Psi^{(-)}(\alpha, \vec{k}_\alpha) \sim \psi_\beta f^{(-)}_{\alpha\beta}(\Omega_\beta) \frac{e^{-ik_\beta r_\beta}}{r_\beta}$$

(onde entrante pure)

On sait que la section efficace

$$\frac{d\sigma(\alpha \vec{k}_\alpha \to \beta \vec{k}_\beta)}{d\Omega_\beta} = \frac{v_\beta}{v_\alpha} \left| f^{+}_{\alpha\beta}(\Omega_\beta) \right|^2 \qquad (2)$$

où v_β , v_α sont les vitesses relatives dans l'état final et initial respectivement. Il est commode de définir la matrice T de transition

$$\langle \beta \vec{k}_\beta | T | \alpha \vec{k}_\alpha \rangle = - \frac{2\pi\hbar^2}{M_\beta} f^{(+)}_{\alpha\beta}(\Omega_\beta) \qquad (3)$$

On vérifie que

$$\frac{d\sigma}{d\Omega_\beta} = \frac{2\pi}{\hbar v_\alpha} \left| \langle \beta \vec{k}_\beta | T | \alpha \vec{k}_\alpha \rangle \right|^2 \rho_\beta(E) \qquad (4)$$

où $\rho_\beta(E)$ est la densité des états finaux $(\beta, \vec{k}_\beta)$.

Supposons que H se mette sous la forme :

$$H = \tilde{\tilde{H}} + V$$

où V est un terme d'énergie potentielle que nous serons amené dans la suite à traiter comme une perturbation. Soit $\tilde{\tilde{T}}$ la matrice de transition relative à l'hamiltonien $\tilde{\tilde{H}}$. Nous allons démontrer que :

$$\boxed{\langle \beta \vec{k}_\beta | T | \alpha \vec{k}_\alpha \rangle = \langle \beta \vec{k}_\beta | \tilde{\tilde{T}} | \alpha \vec{k}_\alpha \rangle + \langle \tilde{\tilde{\Psi}}^{(-)}(\beta \vec{k}_\beta) | V | \Psi^{(+)}(\alpha \vec{k}_\alpha) \rangle} \quad (6)$$

Démonstration :

Nous supposerons les voies α et β distinctes. La démonstration s'applique avec une légère variante lorsque $\alpha = \beta$. Nous allons calculer le deuxième terme du second membre de (6) en utilisant le fait que $\Psi^{(+)}$ est solution de l'équation (1) et que :

$$(H - E)\, \tilde{\tilde{\Psi}}^{(-)}(\beta, \vec{k}_\beta) = V\, \tilde{\tilde{\Psi}}^{(-)}(\beta \vec{k}_\beta) \quad (7)$$

Désignant par H^+ l'opérateur hermitique conjugué de H, il vient

$$\langle \tilde{\tilde{\Psi}}^{(-)}(\beta \vec{k}_\beta) | V | \Psi^{(+)}(\alpha \vec{k}_\alpha) \rangle = \langle \tilde{\tilde{\Psi}}^{(-)} | (H^+ - E) - (H - E) | \Psi^{(+)} \rangle$$

$$= \langle \tilde{\tilde{\Psi}}^{(-)} | (H^+ - H) | \Psi^{(+)} \rangle \quad (8)$$

Pour évaluer cette dernière intégrale, on note que dans $H^+ - H$ la seule contribution ne peut provenir que de termes en $-\frac{h^2}{2M}\Delta$, car les

autres constituants de H sont hermitiques. Quant aux termes en $-\frac{h^2}{2M}\Delta$, on les calcule/dans chacune des voies ouvertes communes aux deux fonctions d'ondes $\bar{\bar{\Psi}}^{(-)}$ et $\bar{\Psi}^{(+)}$:

../en appliquant le théorème de GREEN

$$\langle \bar{\bar{\Psi}}^{(-)} | (H^+ - H) | \bar{\Psi}^{(+)} \rangle = \sum_\gamma \frac{\hbar^2}{2M_\gamma} \lim_{r_\gamma \to \infty} \left\langle \int \left[\bar{\bar{\Psi}}^{(-)*} \frac{d\bar{\Psi}^{(+)}}{dn} - \bar{\Psi}^{(+)} \frac{d\bar{\bar{\Psi}}^{(-)*}}{dn} \right] dS_{r_\gamma} \right\rangle_\gamma \quad (9)$$

la sommation étant étendue à toutes les voies ouvertes γ communes aux deux fonctions et le symbole $\langle \ldots\ldots \rangle_\gamma$ signifiant intégration sur les variables internes de la voie γ . Pour calculer chaque terme de la somme, rappelons les formes asymptotiques des deux fonctions dans les voies ouvert

	Voie α	Voie β	Voie $\nu = \alpha, \beta$
$\bar{\Psi}^{(+)}$	$e^{i\vec{k}_\alpha \cdot \vec{r}_\alpha} + f^{(+)}_{\alpha\alpha}(\Omega_\alpha)\frac{e^{ik_\alpha r_\alpha}}{r_\alpha}$	$f^{(+)}_{\alpha\beta}(\Omega_\beta)\frac{e^{ik_\beta r_\beta}}{r_\beta}$	$f^{(+)}_{\alpha\nu}(\Omega_\nu)\frac{e^{ik_\nu r_\nu}}{r_\nu}$
$\bar{\bar{\Psi}}^{(-)}$	$\bar{\bar{f}}^{(-)}_{\beta\alpha}(\Omega_\alpha)\frac{e^{-ik_\alpha r_\alpha}}{r_\alpha}$	$e^{i\vec{k}_\beta \cdot \vec{r}_\beta} + \bar{\bar{f}}^{(-)}_{\beta\beta}\frac{e^{-ik_\beta r_\beta}}{r_\beta}$	$\bar{\bar{f}}^{(-)}_{\beta\nu}(\Omega_\nu)\frac{e^{-ik_\nu r_\nu}}{r_\nu}$

Il est facile de voir que seules les contributions des voies α et β ne s'annulent pas et que ces dernières proviennent respectivement de la présence de l'onde plane dans $\Psi^{(+)}$ et de l'onde plane dans $\bar{\bar{\Psi}}^{(-)}$. Ains dans la voie β , il vient

$$\lim_{r_\beta \to \infty} r_\beta^2 \int \left[e^{-i\vec{k}_\beta \vec{r}_\beta} \frac{d}{dr} \frac{e^{ik_\beta r_\beta}}{r_\beta} - \frac{e^{ik_\beta r_\beta}}{r_\beta} \frac{d}{dr} e^{-i\vec{k}_\beta \cdot r_\beta} \right] f^{(+)}_{\alpha\beta}(\Omega_\beta) d\Omega_\beta = 4\pi f^{(+)}_{\alpha\beta}(\hat{k}_\beta)$$

où $\hat{k}_\beta$ désigne la direction de $\vec{k}_\beta$. D'où :

$$\langle \bar{\bar{\Psi}}^{(-)} | V | \Psi^{(+)} \rangle = \frac{2\pi\hbar^2}{M_\alpha} \bar{\bar{f}}^{(-)*}_{\beta\alpha}(-\hat{k}_\alpha) - \frac{2\pi\hbar^2}{M_\beta} f^{(+)}_{\alpha\beta}(\hat{k}_\beta) \quad (10)$$

Cette dernière équation reste valable lorsque $V = 0$ et se réduit alors à :

$$0 = \frac{2\pi\hbar^2}{M_\alpha} \tilde{f}^{(-)*}_{\beta\alpha}(-\hat{k}_\alpha) - \frac{2\pi\hbar^2}{M_\beta} \tilde{f}^{(+)}_{\alpha\beta}(\hat{k}_\beta) \qquad (11)$$

Par conséquent :

$$\langle \bar{\Psi}^{(-)} | V | \Psi^{(+)} \rangle = \frac{2\pi\hbar^2}{M_\beta} \left[\tilde{f}^{(+)}_{\alpha\beta}(\hat{k}_\beta) - f^{(+)}_{\alpha\beta}(\hat{k}_\beta) \right]$$

qui est bien la relation (6) compte tenu des définitions (3).

Q.E.D.

Notons que la relation (11) n'est rien d'autre que l'expression du principe de microréversibilité dont nous avons ainsi obtenu une démonstration basée uniquement sur l'invariance de l'Hamiltonien par renversement du sens du temps :

L'approximation de BORN consiste à traiter V comme une perturbation et, par suite, à remplacer dans l'élément de matrice du second membre de l'équation (6) la fonction d'onde exacte $\Psi^{(+)}(\alpha, \vec{k}_\alpha)$ par la fonction d'onde correspondante $\bar{\Psi}^{(+)}(\alpha \vec{k}_\alpha)$ de l'Hamiltonien non perturbée.

APPLICATION A L'EXCITATION COULOMBIENNE $p + N \longrightarrow p' + N'$

$$H = \bar{H} + V$$

$$\bar{H} = H_N + V_{nucl.} + \frac{p^2}{2M} + \frac{Ze^2}{r}$$

$$V = e^2 \sum_{i=1}^{Z} \left[\frac{1}{|\vec{r} - \vec{r}_i|} - \frac{1}{r} \right] \qquad (12)$$

$V_{nucl.}$ est l'interaction nucléaire du proton avec le noyau N. Soient $\hbar \vec{k}_i$ et $\hbar \vec{k}_f$ l'impulsion initiale et finale du proton respectivement. L'expression (6) s'écrit :

$$\langle \vec{k}_f | T | \vec{k}_i \rangle = \langle \vec{k}_f | \tilde{T} | \vec{k}_i \rangle + \langle \tilde{\Psi}^{(-)}(\vec{k}_f) | V | \Psi^{(+)}(\vec{k}_i) \rangle$$

et l'approximation de BORN :

$$\langle \tilde{\Psi}_f^{(-)} | V | \Psi_i^{(+)} \rangle \simeq \langle \tilde{\Psi}_f^{(-)} | V | \tilde{\Psi}_i^{(+)} \rangle \qquad (13)$$

$\langle \vec{k}_f | \tilde{T} | \vec{k}_i \rangle$ est l'amplitude de transition dûe aux interactions nucléaires seules. Lorsque $\frac{b}{R} \gg 1$, on peut négliger la présence du terme $V_{nucl.}$ et cette amplitude est nulle; dans le calcul du second membre de (13), les fonctions d'ondes se factorisent

$$\tilde{\Psi}^{(-)}(\vec{k}_f) = \Psi_{N'} \Phi^{(-)}_{Coul.}(\vec{k}_f)$$

$$\tilde{\Psi}^{(+)}(\vec{k}_i) = \Psi_N \Phi^{(+)}_{Coul.}(\vec{k}_i)$$

où Ψ_N, $\Psi_{N'}$ sont les fonctions d'ondes du noyau initial et final respectivement; $\Phi^{(+)}_{coul.}(\vec{k})$ est la fonction décrivant la diffusion Coulombienne pure d'un proton d'impulsion $\hbar \vec{k}$ par une charge ponctuelle Ze, $\Phi^{(-)}_{coul.}(\vec{k})$ la fonction solution du même problème avec ondes rentrantes. Dans l'approximation $\frac{b}{R} \gg 1$, l'amplitude de transition est une fonction entièrement calculable des moments électriques 2^l - polaire de transition de N à N', résultat en tous points analogue à celui de la théorie semi-classique. Les cal-

culs numériques sont beaucoup plus longs que dans cette dernière théorie et nécessitent pratiquement l'emploi de machines électroniques. (Cf. BIEDENHARM , Mac HALE et THALER à paraître dans Physical Review (1).) Si la condition $\frac{b}{R} \gg 1$ n'est pas remplie, $V_{nucl.}$ ne peut être négligé mais la théorie générale des réactions nucléaires permet d'écrire l'amplitude de transition $\langle \vec{k}_f | \bar{T} | \vec{k}_i \rangle$ ainsi que les fonctions d'onde $\bar{\Psi}^{(+)}$ et $\bar{\Psi}^{(-)}$ qui rentrent dans le calcul de l'élément de matrice de V.

(1) Cf. également les formules plus simples obtenues par BIEDENHARM et CLASS (Phys. Rev. 98 691 (1955)) à la limité où $\frac{\Delta E}{E} \ll 1$ en faisant l'approximation $k_i = k_f$.

Troisième Conférence.-

THEORIE DU STRIPPING.

La première tentative d'explication des réactions (dp) ou (dn) est dûe à OPPENHEIMER et PHILLIPS (1935). Il s'agissait d'expliquer les sections efficaces "anormalement" grandes de réaction (dp) pour des deutons de très faible énergie (quelques centaines de Kev). Dans l'hypothèse d'une pénétration complète du deuton dans la cible, un calcul grossier tenant compte du facteur de pénétration conduit à une section efficace considérablement plus faible. En fait, le deuton est un noyau "anormalement" grand, la distance internucléaire moyenne est de l'ordre de $2,5 \times 10^{-13}$ cm Le neutron et le proton qui le constituent se trouvent environ la moitié du temps hors de portée de leur force d'attraction mutuelle. Dans le mécanisme imaginé par OPPENHEIMER et PHILLIPS, le neutron pénètre dans le noyau sans que le proton franchisse la barrière Coulombienne; la liaison neutron-proton est alors rompue, le neutron étant capturé et le proton continuant son chemin.

Le mécanisme du stripping devait être invoqué plus tard pour expliquer des expériences faites au grand cyclotron de BERKELEY dans lesquelles le bombardement de cibles nucléaires par des deutons de grande énergie donnait lieu à l'émission vers l'avant de protons ou de neutrons d'énergie cinétique environ moitié de l'énergie incidente. SERBER (Phys. Rev. 72

1008 (1947)) rendit compte de la variation de la section efficace en fonction de l'énergie, et de la distribution angulaire des nucléons émis à l'aide d'un modèle semi-classique de ce mécanisme de stripping. De nombreuses expériences faites depuis à toutes énergies mais notamment dans le domaine 1-8 Mev, donnent invariablement cette distribution angulaire vers l'avant caractéristique du stripping ou, tout au moins, de l'absence de formation du noyau composé. En fait, on peut imaginer 3 mécanismes pour expliquer les réactions (dp) ou (dn) :

(1) Formation du "noyau composé" - pourrait devenir important autour des résonances du noyau composé formé par le deuton et le noyau cible.

(2) Excitation Coulombienne provoquant la rupture du deuton. DANCOFF (Phys.Rev. 72, 1012 (1947)) a montré que ce phénomène ne devient notable qu'aux énergies de l'ordre de 50 Mev.

(3) Mécanisme de stripping - celui que nous envisageons dans la suite.

Nous exposerons d'abord la théorie semi-classique de PEASLEE dans laquelle les états quantiques du noyau résiduel ne sont pas analysés, puis la théorie de BUTLER qui donne les sections efficaces de réaction A(dp)B donnant un noyau B dans un état quantique bien spécifié. Les distributions angulaires prévues par BUTLER sont en surprenant accord avec l'expérience; elles sont sujettes en outre à des règles de sélection remarquables qui font des réactions (dp) et (dn) une méthode importante de

détermination des spins et des parités des noyaux. Enfin nous discuterons les améliorations qu'il est possible d'apporter à la théorie de BUTLER. On notera que toutes ces théories procèdent implicitement ou explicitement d'un traitement de perturbation.

THEORIE DE PEASLEE (Phys.Rev. 74, 1001 (1948))

Le modèle de PEASLEE traite le noyau comme une surface sphérique de rayon R sur laquelle le proton (ou le neutron) a une probabilité ξ_p (ou ξ_n) d'être capturé. La section efficace totale de réaction (dn) est :

$$\sigma_{dn} = \pi R^2 \rho_p(R) \xi_p \qquad \text{de même : } \sigma_{dp} = \pi R^2 \rho_n(R) \xi_n$$

Dans cette formule $\rho_p(R)$, $\rho_n(R)$ sont les densités de probabilité de trouver le proton, le neutron respectivement à la surface du noyau. PEASLEE les calcule à l'aide de la fonction d'onde $\varphi(\vec{r}_n, \vec{r}_p)$ décrivant le deuton incident. Ainsi :

$$\rho_n(R) = \int_{r_p > R} d\vec{r}_p \int_{r_n = R} \frac{d\Omega_n}{4\pi} \left| \varphi(\vec{r}_n, \vec{r}_p) \right|^2$$

la fonction d'onde choisie est

$$\varphi(\vec{r}_n, \vec{r}_p) = \psi_d(\vec{r}_n - \vec{r}_p)\, \varphi_c(\vec{r}_p)$$

où $\varphi_c(r_p)$ représente l'onde Coulombienne d'une particule de masse 2 M,

ayant l'impulsion initiale du deuton et $\Psi_d(\vec{r})$ la fonction d'onde interne du deuton. On notera le rôle disymétrique joué par le neutron et le proton dans cette fonction d'onde à seule fin de simuler l'effet non-coïncidence du centre de masse et du centre de charge dans le deuton. ξ_p et ξ_n sont des quantités en principe inférieures à 1 mais de l'ordre de grandeur de l'unité.

Afin de comparer cette théorie grossière avec les mesures expérimentales de sections efficaces totales faites par activation, il faut soustraire la contribution des réactions (d, pn). Nous n'insisterons pas ici sur ces détails qui sont exposés dans l'article de PEASLEE. Il y a meilleur accord qu'on ne s'y attendait entre cette théorie et des expériences faites sur divers corps à des énergies allant de 2 à 15 Mev. Ci-dessous quelques valeurs des paramètres $r_o = R \times A^{-1/3}$ et ξ_n ou ξ_p, utilisés pour obtenir l'accord avec les courbes expérimentales :

Cible	$r_o(10^{-13}cm)$	ξ_n (réaction dp)	ξ_p (réaction dn)
Na^{24}	1,5	1,8	
Co^{59}	1,65	0,8	
Cu^{63}	1,5	0,9	
Br^{81}	1,65	1,1	
Bi^{209}	1,8	0,23	0,25

THEORIE DE BUTLER-BORN

(S.T. BUTLER, Proc. Royal Soc. A 208, 559 (1951)
P.B. DAITCH et J.B. FRENCH, Phys. Rev. 85, 695 (1952)
BHATIA et al., Phil. Mag. 43, 485 (1952))

La méthode que nous suivons pour dériver la formule de BUTLER mettra en évidence la relation entre la théorie de BUTLER et l'approximation de BORN. La section efficace différentielle de la réaction A(dp)B est rigoureusement donnée par la formule :

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar v} |T|^2 \rho_f \qquad (1)$$

où v est la vitesse du deuton incident, ρ_f la densité des états finaux. D'après le théorème démontré au cours de la deuxième conférence, l'élément de matrice de transition T est égal à :

$$T = \langle \Phi_f | (P+V) | \Psi_d^{(+)} \rangle \qquad (2)$$

où $\Psi_d^{(+)}$ représente l'état stationnaire de collision décrivant la réaction considérée , $\Phi_p = \Psi_B e^{i\vec{k}.\vec{r}_p}$ est le produit de la fonction d'onde du noyau B par une onde plane de proton, P et V l'interaction du proton-noyau A et proton-neutron respectivement. L'approximation de BORN consiste à remplacer $\Psi_d^{(+)}$ par le produit

$$\Phi_d \equiv \Psi_A \Psi_d(\vec{r}_n - \vec{r}_p) \exp\left[i\vec{K}.\frac{\vec{r}_n + \vec{r}_p}{2}\right] \qquad (3)$$

représentant une onde plane de vecteur $\vec{K}$, vecteur d'onde du deuton incident. Ψ_d est la fonction d'onde du deuton, Ψ_A celle du noyau A. Pour obtenir la formule de BUTLER, il faut faire en outre "l'approximation du stripping" qui exprime que le proton ne pénètre pas dans le noyau ; à cet effet, on écrit $P = 0$ (négligeant du même coup tout effet Coulombien) et l'on suppose l'interaction neutron-proton effective seulement lorsque le neutron est extérieur au noyau A, ce qui revient à remplacer V par $V_{ext.}$:

$V_{ext.} = V$ si $r_n > R$

$V_{ext.} = 0$ si $r_n < R$ R rayon du noyau A

Nous écrivons donc :

$$T \simeq \langle \Phi_p | V_{ext.} | \Phi_d \rangle \qquad (4)$$

A cette approximation, le calcul de T peut être poussé jusqu'au bout. Posons $\vec{r} = \vec{r_p} - \vec{r_n}$, $\vec{q} = \vec{K} - \vec{k}$, et soit $f^*(\vec{r_n})$ l'intégrale de $\Psi_B^* \Psi_A$ sur les variables dynamiques du noyau A .
Il vient :

$$T \simeq \int_{r_n > R} d\vec{r}_n \int d\vec{r}_p \, e^{-i\vec{k}.\vec{r}_p} f^*(\vec{r}_n) V(\vec{r}) \Psi_d(\vec{r}) \, e^{i\vec{K}.\frac{\vec{r}_n + \vec{r}_p}{2}}$$

$$= \left(\int_{r_n > R} e^{i\vec{q}.\vec{r}_n} f^*(\vec{r}_n) d\vec{r}_n \right) \times \left(\int e^{i\left(\frac{\vec{K}}{2} - \vec{k}\right).\vec{r}} V(\vec{r}) \psi_d(\vec{r}) d\vec{r} \right)$$

Le calcul de la deuxième intégrale peut se faire en notant que

$$V(\vec{r}) \psi_d(\vec{r}) = \left(\frac{\hbar^2 \Delta}{M} - E_d \right) \psi_d(\vec{r})$$

où E_d est l'énergie de liaison du deuton et en applicant le théorème de GREEN ce qui donne :

$$-\left[\frac{\hbar^2}{2M} \left(\frac{\vec{K}}{2} - \vec{k} \right)^2 + E_d \right] \times P\left(\frac{\vec{K}}{2} - \vec{k} \right)$$

où

$$P\left(\frac{\vec{K}}{2} - \vec{k} \right) = \int e^{i\left(\frac{\vec{K}}{2} - \vec{k}\right).\vec{r}} \psi_d(\vec{r}) d\vec{r} \qquad (5)$$

est l'amplitude de probabilité de trouver un proton d'impulsion $\hbar\vec{k}$ dans l'onde plane de deuton initiale. La première intégrale se calcule aussi à l'aide du théorème de GREEN. Dans la région $r_n > R$, le neutron est hors de portée des forces qui le lient aux autres nucléons de B; $f(\vec{r}_n)$ est donc solution de l'équation de SCHRÖDINGER

$$\left(-\frac{\hbar^2 \Delta_n}{2M} + \frac{\hbar^2 t^2}{2M} \right) f(\vec{r}_n) = 0 \qquad r_n > R$$

où $\frac{\hbar^2 t^2}{2M}$ est l'énergie de liaison du neutron dans B. De même :

d'où

$$F\left[\vec{q}, f(\vec{r})\right] \equiv \frac{\hbar^2}{2M}(q^2+t^2)\int_{r_n=R} e^{i\vec{q}\cdot\vec{r}_n} f^*(\vec{r}_n)\, d\vec{r}_n = \int_{r_n=R}\left[f^*(\vec{r}_n)\frac{d}{dn}e^{i\vec{q}\cdot\vec{r}_n} - e^{i\vec{q}\cdot\vec{r}_n}\frac{d}{dn}f^*(\vec{r}_n)\right] dS$$

intégrale sur la surface $r_n = R$ où $\frac{d}{dn}$ signifie dérivée normale extérieure. Comme la conservation de l'énergie exige que :

$$\frac{\hbar^2}{2M}(q^2+t^2) = \frac{\hbar^2}{2M}\left(\frac{\vec{K}}{2} - \vec{k}\right)^2 + E_d \ .$$

$$T \simeq P\left(\frac{\vec{K}}{2} - \vec{k}\right) \times F\left[\vec{q}, f(\vec{r})\right] \qquad (7)$$

REGLES DE SELECTION.

Dans tout ce qui précède, nucléons et noyaux sont supposés sans spins. La formule (7) reste valable en toute généralité. Soient j_A, j_B les spins des noyaux A et B, μ_A, μ_B, σ et π, les composantes respectives suivant l'axe oz des spins de A, B, du deuton et du proton. Désignons par $\chi^{\nu}_{1/2}$ la fonction de spin du neutron ayant composante ν suivant oz et introduisons les fonctions U^{μ}_{j} formés par combinaison des fonctions $\chi^{\nu}_{1/2}$ et $\psi^{\mu_A}_{j_A}$

$$U^{\mu}_{j} = \sum_{\nu\mu_A} \langle j_A \tfrac{1}{2}\, j\mu \,|\, j_A \tfrac{1}{2}\, \mu_A \nu \rangle\, \psi^{\mu_A}_{j_A}\, \chi^{\nu}_{1/2} \qquad j = j_A \pm 1/2$$

La fonction d'onde s'écrit :

$$\cdot \psi_{j_B}^{\mu_B} = \sum_{\substack{j\ell \\ \mu m}} \alpha_{j\ell} \langle j\ell\, j_B \mu_B \mid j\ell\mu m \rangle U_j^{\mu} g_\ell^m(\vec{r}_n)$$

+ termes orthogonaux à $\psi_{j_A}^{\mu_A}$

Dans la région $r_n \geq R$, $g_m^\ell(\vec{r}_n)$ s'écrit simplement

$$g_\ell^m(\vec{r}_n) = Y_\ell^m(\Omega_n)\, h_\ell^{(1)}(ikr_n)$$

où $h_\ell^{(1)}$ est la fonction de HANKEL sphérique de première espèce.

$$F\left[\vec{q}, g_\ell^m(\vec{r})\right] = 4\pi R^2 (-i)^\ell Y_\ell^m(\hat{q}) \left[h_\ell^{(1)}(ikr) \frac{d}{dr} j_\ell(qr) - j_\ell(qr) \frac{d}{dr} h_\ell^{(1)}(ikr) \right]_{r=R} \quad (8)$$

$$= 4\pi R \times Y_\ell^m(\hat{q}) \times A_\ell(q)$$

Ceci posé, on trouve sans peine que l'élément de matrice de transition $T(\delta\mu_A \to \pi\mu_B)$ est donné par la formule (7) si l'on prend pour fonction $f(\vec{r}_n)$ la combinaison linéaire :

$$\sum_{j\ell} \alpha_{j\ell} \sum_{m\mu\nu} \langle j\ell\, j_B \mu_B \mid j\ell\mu m \rangle \langle j_A \tfrac{1}{2} j\mu \mid j_A \tfrac{1}{2} \mu_A \nu \rangle \langle \tfrac{1}{2} \tfrac{1}{2} 1\delta \mid \tfrac{1}{2} \tfrac{1}{2} \nu\pi \rangle g_\ell^m(\vec{r}_n)$$

Lorsqu'on effectue la sommation de $|T|^2$ sur les spins finaux et la moyenne sur les spins initiaux, on obtient, après simplications au moyen des relations d'orthogonalité des coefficients de CLEBSCH-GORDON

$$\frac{1}{3(2j_A+1)} \sum_{spin} |T|^2 = \frac{2\pi R^4 (2j_B+1)}{(2j_A+1)} \times P^2\left(\frac{\vec{k}}{2} - \vec{k}\right) \times \sum_\ell \left[A_\ell^2(q) \times \left(\sum_j |\alpha_{j\ell}|^2 \right) \right]$$

Chaque valeur de ℓ correspond à un moment cinétique orbital du neutron capturé. Elle doit satisfaire aux règles de sélection de spin et de parité

$$\pi_A \, \pi_B = (-)^{\ell} \qquad \text{minimum} \left| j_A \pm j_B \pm 1/2 \right| \leq \ell \leq j_A + j_B + 1/2$$

A chaque valeur de ℓ correspond une distribution angulaire caractéristique. La fonction $P^2 \left(\frac{\vec{K}}{2} - \vec{k}\right)$ possède un pic très prononcé à 0°, tandis que la fonction

$$A_\ell^2(q) = R^2 \left[h_\ell^{(1)}(ir) \frac{d}{dr} j_\ell(qr) - j_\ell(qr) \frac{d}{dr} h_\ell^{(1)}(ir) \right]^2_{r=R}$$

a un comportement oscillatoire dépendant de la valeur de ℓ considérée. La distribution angulaire résultante présente une succession de pics très prononcés à des angles d'autant plus proches de 0°, en règle générale, que ℓ est plus petit ou que l'énergie incidente est plus grande. On trouvera des exemples caractéristiques de distributions angulaires dans l'article original de BUTLER (loc.cit.). Le résultat important est que les distributions angulaires correspondant à différentes valeurs de ℓ, notamment les positions des pics, sont très faciles à distinguer les unes des autres. Dans l'hypothèse où la théorie de BUTLER prévoit correctement les distributions angulaires, l'analyse des distributions angulaires expérimentales permet de déterminer la ou les valeurs de ℓ du neutron capturé et, par le jeu des règles de sélection d'une part, la parité relative des noyaux A et B d'autre part, une condition restrictive concernant les valeurs possibles

des spins j_A et j_B . C'est ce qui fait le grand intérêt des réactions (dp) et (dn) en spectroscopie nucléaire.

Quatrième Conférence.-

DISCUSSION ET AMELIORATION DE LA THEORIE DE BUTLER.

L'approximation du stripping faite dans la théorie de BUTLER des réactions (dp) doit être réexaminée. Il n'est pas justifié de négliger complètement l'interaction entre la cible et le proton ($P = 0$) sous prétexte que ce dernier passe trop loin de la cible pour en sentir l'effet. En fait, le rayon du deuton n'est pas tellement grand que le neutron puisse entrer en interaction avec le noyau sans que le mouvement du proton en soit affecté; le proton subit de toutes façons l'interaction Coulombienne, qui est une interaction à longue portée et il subit également de façon non négligeable l'interaction nucléaire. Tous les travaux visant à l'amélioration de la théorie de BUTLER sont basés implicitement sur l'approximation de BORN généralisée. Reprenant les notations des paragraphes précédents, désignons par $\Psi_p^{(-)}$ l'état stationnaire de collision (avec onde rentrante) du proton subissant l'interaction avec la "partie interne" du noyau B, c'est à dire le potentiel P d'interaction avec les nucléons de A et la "partie interne" $(V - V_{ext.})$ de l'interaction neutron-proton. L'amplitude de transition est donnée rigoureusement par

$$T = \langle \Psi_p^{(-)} \mid V_{ext.} \mid \Psi_d^{(+)} \rangle$$

Cette formule ne fait intervenir que les parties externes des fonctions $\Psi_d^{(+)}$ et $\widetilde{\Psi}_p^{(-)}$. Toutes les théories proposées jusqu'ici tendent à remplacer ces fonctions par des fonctions d'ondes approchées. La théorie de BUTLER consiste à les remplacer toutes deux par des ondes planes. HOROWITZ et MESSIAH (Journal Phys. 14, 695 (1953) et Phys.Rev. 92, 1326 (1953)) substituent une onde plane à $\Psi_d^{(+)}$ mais remplacent $\widetilde{\Psi}_p^{(-)}$ par l'état stationnaire de collisions d'un proton par une sphère impénétrable de rayon égal à celui du noyau B. Leurs calculs montrent que, dans la plupart des cas, la distribution angulaire n'est guère affectée (la théorie de BUTLER surestime un peu les rayons nucléaires) mais que la valeur absolue de la section efficace est beaucoup plus faible (de 3 à 5 fois plus faible suivant le cas) que dans la théorie de BUTLER, de sorte que les paramètres nucléaires α_{jl} obtenus en cherchant à ajuster les courbes de BUTLER avec les courbes expérimentales sont sans doute sous-estimées d'autant. On explique ce résultat en notant que, parmi les ondes partielles de proton participant à la réaction, seules celles de moment cinétique plus faible ($\ell < R / \lambda\!\!\!^{-}_p$) sont modifiées par la présence du potentiel nucléaire d'où un effet spectaculaire sur les valeurs absolues sans que la distribution angulaire soit notablement affectée.

Les calculs les plus complets du type exposé ici ont été faits par TOBOCMAN et KALOS utilisant un calculateur électronique (cf. Rapport de TOBOCMAN - Conference on Statistical Aspects of the Nucleus, BNL 24-26/1/1955) pour un certain nombre de noyaux et de valeurs de l'é-

nergie incidente. $V_{ext.}$ est considéré comme un potentiel de portée nulle; cette approximation, certainement justifiée, simplifie beaucoup les calculs. Dans chaque cas, ils calculent T :

1) Dans l'approximation de BUTLER : $\widetilde{\Psi}_p^{(-)}$ et $\Psi_d^{(+)}$ remplacés par ondes planes.

2) BUTLER + effet Coulombien : $\widetilde{\Psi}_p^{(-)}$ et $\Psi_d^{(+)}$ remplacés par des ondes Coulombiennes pures de proton et de deuton.

3) Effet Coulombien + diffusion potentielle : $\Psi_d^{(+)}$ approché comme dans (2) ; $\widetilde{\Psi}_p^{(-)}$ remplacé par une onde de proton diffusée élastiquement par une sphère impénétrable.

4) Effet Coulombien + sphère totalement absorbante : idem sauf que la sphère impénétrable est remplacée par une sphère totalement absorbante (onde de proton se prolongeant en e^{-ikr} à l'intérieur du noyau B).

La variante (4) semble donner le meilleur accord avec l'expérience pour les distributions angulaires.

Outre les valeurs absolues des sections efficaces, ces améliorations de la théorie du stripping diffèrent notablement de la théorie du stripping sur deux points importants :

1) le proton émis peut être notablement polarisé (jusqu'à 30 %);

2) si le noyau résiduel est laissé dans un état excité, la distribution angulaire des rayons γ de désexcitation qui, dans la théorie de BUTLER, est symétrique de révolution autour de l'axe du vecteur $\vec{q} = \vec{K} - \vec{k}$, cesse d'avoir cette propriété.

Des expériences de polarisation ou de corrélation ($dp\gamma$) permettraient donc de vérifier les diverses théories proposées.

On s'attend à des déviations spectaculaires de la distribution angulaire de BUTLER lorsqu'on fait varier l'énergie du deuton au voisinage d'une résonance du noyau composé (d + A). Pour qu'un effet clair se manifeste, il est nécessaire de se trouver dans le domaine des résonances étroites ; c'est le cas pour les noyaux C_{12} et O_{16} avec des deutons de 1 à 4 Mev. Il semble que l'on puisse expliquer les résultats expérimentaux en partant de la formule (8) et en y remplaçant $\Psi_d^{(+)}$ et $\tilde{\Psi}_p^{(-)}$ par des ondes de diffusion élastique résonante avec des paramètres de résonance appropriés. Cependant aucun calcul n'a été publié jusqu'ici à l'appui de ces prévisions.

En conclusion, la théorie de BUTLER convenablement améliorée semble en bon accord avec l'expérience dans pratiquement tous les cas.

Ceci confirme l'image proposée au début d'une division de l'espace de configuration en une zone restreinte du "noyau composé" où les nucléons sont en contact étroit les uns avec les autres et une zone externe où toutes les interactions peuvent être traités par perturbation.

INTERACTION OF A NUCLEUS
with
ATOMIC AND MOLECULAR FIELDS

by

Norman F. RAMSEY
Harvard University

Chapters of a book on Molecular Beams to be published.

ECOLE d'ÉTÉ de PHYSIQUE THÉORIQUE

Université de Grenoble

Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) - France

Juillet 1955.

CONTENTS

pages

III_I. INTERACTION OF A NUCLEUS WITH ATOMIC AND MOLECULAR FIELDS.

- Introduction. I

2. Electrostatic Interaction. 3

- 2.I. Multipole expansion. 3

- 2.2. Theoretical restrictions on electric multipole orders II

- 2.3. Nuclear electric quadrupole interactions . . . I3

- 2.4. Multipole moments of arbitrary order. . . . 20

3. Magnetic Interaction. 2I

- 3.I. Magnetic multipoles. 2I

- 3.2. Magnetic Dipoles. 26

- 3.3. Magnetic dipole interactions with internal and external magnetic fields. 30

- 3.4. Magnetic interactions in molecules 32

4. Intermediate Coupling. 36

- 4.I. Introduction 36

- 4.2. Solution of secular equation 37

- 4.3. Perturbation theory 40

IX.9. Magnetic Octupole Interactions 42

INTERACTION OF A NUCLEUS WITH ATOMIC AND MOLECULAR FIELDS

III. 1. Introduction.

Many of the most important molecular beam measurements are determinations of the interactions between the nuclei of the molecule and the molecular fields at the nuclei from the remainder of the molecule. For this reason, detailed considerations of molecular beam measurements are best preceded by a theoretical discussion of the interaction of a nucleus with atomic and molecular fields.

The interactions between a nucleus and the atom or molecule containing it may be studied in various ways including the methods of optical spectroscopy, molecular beams, nuclear paramagnetic resonance, microwave spectroscopy, and paramagnetic resonance as described in the author's book on Nuclear Moments (RAM 53e). All of these measurements are consistent with the following assumptions concerning atomic nuclei :

a. A nucleus has a charge Ze confined to a small region of the order of 10^{-12} cm diameter. It is the interaction of this charge with the electrons of an atom that gives rise to the gross features of optical atomic spectra.

b. A nucleus whose mass number A is odd obeys Fermi-Dirac statistics (the sign of the wave function is reversed if two such identical nuclei are interchanged), and a nucleus whose mass number A is even obeys Bose-

Einstein statistics (wave function unaltered on interchange).

c. A nucleus possesses a spin angular momentum capable of being represented by a quantum mechanical angular momentum vector $\underline{a}$ with all the properties[1] usually associated with such vectors. In nuclear moment work it is usually most convenient to use a dimensionless quantity I to measure the angular momentum in units of h where I is therefore defined by :

$$a = h I \qquad \text{(III.1)}$$

The spin $\underline{I}$ of a nucleus is defined as the maximum possible component of I in any given direction.

d. The nuclear spin $\underline{I}$ is half integral if the mass number A is odd and integral if the mass number is even.

e. A nucleus has a magnetic moment (PAU24) μ_I which can be represented as

$$\mu_I = \gamma_I h_I = g_I \mu_N I \qquad \text{(III.2)}$$

where $\underline{\gamma_I}$ and $\underline{g_I}$ are defined by the above equation and are called the nuclear gyromagnetic ratio and the nuclear g factor, respectively. μ_N is the nuclear magneton defined as e h/2 Mc where M is the proton mass[2]. μ_N has the numerical value of $(5.05038 \pm 0.00036) \times 10^{-24}$ erg gauss^{-1}. The scalar quantity which is a measure of the magnitude of μ_I and which is called the magnetic moment μ_I is

---

[1] Good discussions of the properties of angular momentum vectors are given in books by Condon and Shortley (CON35), Wigner (WIG31), Feenberg and Pake (FEE53), Racah (RAC42), Pauli (PAU32), and Schwinger (SCH46, SCH52).

[2] In the literature on the subject of nuclear moments there is considerable confusion as to notation. In different papers μ_o, μ_n, μ_{NM} represent the nuclear magneton whereas μ_o, μ_1 and β designate the Bohr magneton. Also in many papers g or g_I is written for $\mu_I / I \mu_N$, but in others g represents $\mu_I / I \mu_o$.

(III.3)

With this, eq. (III.2) can be written :

$$\mu_I = (\mu_x / I)\, I \qquad \text{(III.4)}$$

f. Many nuclei with spin equal or greater than 1 possess an electrical quadrupole moment, i.e., have electrical charge distributions which depart from spherical symmetry in a manner appropriate to an electrical quadrupole moment. (A precise definition of an electrical quadrupole moment is given in the next section).

g. A nucleus with a spin 3/2 or greater may possess a magnetic octupole moment (JAC54). (The definition of a magnetic octupole interaction is given in the next section.)

h. The nucleus is not infinitely heavy and of negligible size but instead with certain experiments allowance must be made for the finite mass and appreciable size of the nucleus (BRE32, RAM53e).

i. The nucleus has a finite polarizability and can be polarized by a strong electric field (KOP49, BRE50, GUN51, RAM53f).

III.2. Electrostatic Interaction.

III.2.1. Multipole expansion.

The general electrostatic interaction $\mathcal{H}_E$ between a charged nucleus and the charged electrons and nuclei of the remainder of the molecule is as follows if the finite extension of the nucleus is taken into account :

$$\mathcal{H}_E = \int_{\tau_e} \int_{\tau_n} \frac{\rho_e \rho_n \, d\tau_e \, d\tau_n}{r} \qquad \text{(III.5)}$$

ρ_e is the charge density of the electrons and of the other nuclei in the volume element $d\tau_e$ at position $\underset{\sim}{r}_e$ relative to the centroid of the nucleus

concerned so that ρ_e is a function of $\underset{\sim}{r}_e$. ρ_n is the nuclear charge density of the nucleus concerned in the volume element $d\tau_n$ at position $\underset{\sim}{r}_n$ relative to the centroid of the nucleus whence ρ_n is a function of $\underset{\sim}{r}_n$. In addition $\underset{\sim}{r}$ is the magnitude of the radius vector $\underset{\sim}{r}$ down from $d\tau_n$ to $d\tau_e$ as shown in Fig. III.1. The definition of ρ_e is such that it is negative for electrons and positive for positive charges.

Let θ_{en} be the angle between $\underset{\sim}{r}_e$ and $\underset{\sim}{r}_n$ and let r_e be greater than r_n, as may be assured by limiting consideration to electronic charges more distant than the radius R_N of the nucleus. This limitation will be explicitly indicated by the use of a superscript e with ρ_e^e to indicate charges external to a small sphere of radius R_N surrounding the nucleus and centered on the nuclear centroid. Then $1/r$ may be expressed as follows from the cosine law of trigonometry and expanded in the well-known fashion as a power series in r_n/r_e (cf. (MAR43, p.96)):

$$1/r = (r_e^2 + r_n^2 - 2\, r_e r_n \cos\theta_{en})^{-1/2} = 1/r_e + \frac{r_n}{r_e^2} P_1 + \frac{r_n^2}{r_e^3} P_2 + \frac{r_n^3}{r_e^4} P_3 + \cdots \qquad \text{(III.6)}$$

where P_k is the k'th Legendre polynomial of $\cos\theta_{en}$ so that

$$\begin{aligned} P_o &= 1 \\ P_1 &= \cos\theta_{en} \\ P_2 &= 1/2\,(3\cos^2\theta_{en} - 1) \\ P_3 &= 1/2\,(5\cos^3\theta_{en} - 3\cos\theta_{en}) \end{aligned} \qquad \text{(III.7)}$$

$$P_k = \frac{1}{2^k k!} \frac{d^k}{(d\cos\theta_{en})^k} (\cos^2\theta_{en} - 1)^k$$

From Eqs. (III.5) and (III.6), one can write

$$\mathcal{H}_E = \sum_k \mathcal{H}_{Ek}$$

where

$$\mathcal{H}_{Ek} = \int_{\tau_e} \int_{\tau_n} \frac{\rho_e^e \rho_n}{r_e} \left(\frac{r_n}{r_e}\right)^k P_k(\cos\theta_{en})\, d\tau_e\, \tau_n. \qquad \text{(III.9)}$$

The $\mathcal{H}_k$ term which involves P_k is said to be the interaction energy from the multipole moment of order 2^k. Thus the first term corresponds to a monopole or single charge, the second to the electric dipole moment, the third the electric quadrupole moment, the fourth to the electric octupole moment, etc.

The direct evaluation of Eq. (III.9) would necessarily be difficult since θ_{en} depends upon both the nuclear and electronic coordinates whence the double integration is not separated into two integrations each of which involves only the electron or the nuclear coordinates. However, such a separation can be achieved by using the spherical harmonic addition theorem (see (CON35, p. 53)) according to which

$$P_k(\cos\theta_{en} = = \frac{4}{2k+1} \sum_{q=-k}^{k} (-1)^q Y_k^{-q}(\cos\theta_n, \phi_n)\, Y_k^{+q}(\cos\theta_e, \phi_e) \qquad \text{(III.10)}$$

where θ_n is the angle between r_n and an arbitrary fixed z axis, ϕ_n the azimuthal angle of $d\tau_n$ from the xz plane while θ_e and ϕ_e are the corresponding quantities for $d\tau_e$ as shown in Fig. III.1. Y_k^q is the normalized tesseral harmonic

$$Y_k^q(\cos\theta, \phi) = \Theta_k^q(\cos\theta) \frac{1}{\sqrt{2}}\, e^{iq\phi} \qquad \text{(III.11)}$$

where Θ_k^q except for normalization is the associated Legendre polynomial and is defined for $q > 0$ as

$$\Theta_k^q = (-1)^q \sqrt{\frac{(2k+1)(k-q)!}{2(k+q)!}} \sin^q \theta \frac{d^q}{(d\cos\theta)^q} P_k(\cos\theta) \quad \text{(III.12)}$$

and for $q < 0$ as

$$\Theta_k^q = (-1)^q \Theta_k^{-q} \quad \text{(III.13)}$$

With these expressions, the 2^k interaction can be written as

$$\mathcal{H}_{Ek} = \underset{\sim}{Q}_k \cdot \underset{\sim}{F}_k , \quad \text{(III.14)}$$

where $\underset{\sim}{Q}_k \cdot \underset{\sim}{F}_k$ for the irreducible (RAC42) tensors $\underset{\sim}{Q}_k$ and $\underset{\sim}{F}_k$ of degree k indicates the tensor scalar product defined by

$$\mathcal{H}_{Ek} = Q_k \cdot F_k = \sum_q (-1)^q Q_{k,q} F_{k,-q} \quad \text{(III.15)}$$

where

$$Q_{k,q} = \sqrt{\frac{4\pi}{2k+1}} \int_{\tau_n} \rho_n r_n^k Y_\kappa^q(\cos\theta_n, \phi_n) \, d\tau_n \quad \text{(III.16)}$$

and

$$F_{k,q} = \sqrt{\frac{4\pi}{2k+1}} \int_{\tau_e} \rho_e^e r_e^{-(k+1)} Y_k^q(\cos\theta_e, \phi_e) \, d\tau_e \quad \text{(III.17)}$$

The integrals in Eqs. (III.16) and (III.17) have the great advantage that they separately concern either the electron or the nuclear coordinates, but not both simultaneously (RAC42, POU50).

Although eqs. (III.15), (III.16) and (III.17) give the general 2^k- pole interaction energy, they are inconvenient to work with in their most general form. However, the specific expression applicable for the three lowest multipoles may easily be derived as in the following paragraphs.

The monopole interaction is :

$$\mathcal{H}_{EO} = Ze\,\phi^e \qquad \text{(III.18)}$$

where

$$Ze = Q_{0,0} = \int_{\tau_n} \rho_n \, d\tau_n \qquad \text{(III.19)}$$

and

$$\phi^e = F_{0,0} = \int_{\tau_e} \frac{\rho_e^e}{r_e} \, d\tau_e \qquad \text{(III.20)}$$

ϕ^e is the lecetrostatic potential at the centroid of the nucleus from the electrical charges external to a small sphere of radius R centered on the nuclear centroid and surrounding all the nuclear charge.

In a similar fashion the electric dipole interaction is

$$\mathcal{H}_{E1} = -p_z E_z^e - \tfrac{1}{2} p_+ E_-^e - \tfrac{1}{2} p_- E_+^e \qquad \text{(III.21)}$$

where p_z is the z component of nuclear electrical dipole moment in a coordinate system whose origin is at the nuclear centroid so that

$$p_z = Q_{1,0} = \int_{\tau_n} \rho_n r_n \cos\theta_n \, d\tau_n = \int_{\tau_n} \rho_n z_n \, d\tau_n \qquad \text{(III.22)}$$

and $p_\pm$ is defined by

$$p_\pm = p_x \pm i p_y = \pm \sqrt{2}\, Q_{1,\pm 1}$$

$$= \int_{\tau_n} \rho_n r_n \sin\theta_n \cos\phi_n \, d\tau_n \pm i \int \rho_n r_n \sin\theta_n \sin\phi_n \, d\tau_n$$

$$= \int_{\tau_n} \rho_n x_n \, d\tau_n \pm i \int \rho_n y_n \, d\tau_n \qquad \text{(III.23)}$$

It is then apparent that p_x is the x component of the nuclear electrical dipole moment. Similarly, E^e_z is the lecrtic field at the nuclear centroid from the electrical charges external to the small sphere of radius R surrounding the nucleus so that

$$E^e_z = -F_{1,0} = -\int_{\tau_e} \frac{\rho^e_e}{r^2_e} \cos\theta_e \, d\tau_e = -\int_{\tau_e} \frac{\rho^e_e}{r^3_e} z^e \, d\tau_e \qquad \text{(III.24)}$$

$$= -\left[\frac{\partial \phi^e}{\partial z}\right]$$

and

$$E^e_{\pm} \equiv E^e_x \pm i E^e_y = \mp F_{1,\pm 1} = -\int_{\tau_e} \frac{\rho^e_e}{r^2_e} \sin\theta_e \cos\phi_e \, d\tau_e$$

$$\pm \int_{\tau_e} \frac{\rho^e_e}{r^2_e} \sin\theta_e \sin\phi_e \, d\tau_e = -\int_{\tau_e} \frac{\rho^e_e}{r^3_e} x^e \, d\tau_e \mp i \int_{\tau_e} \frac{\rho^e_e}{r^3_e} y_e \, d\tau_e$$

$$= -\left[\frac{\partial \phi}{\partial x}\right] \mp \left[\frac{\partial \phi^e}{\partial y}\right] \qquad \text{(III.25)}$$

It should be noted in the derivation of the preceding equations that such differentiations as ∂/∂ z with respect to the field point can be replaced by $-\partial/\partial$ z_e under the integral sign.

Likewise the electric quadrupole interaction can be written as

$$\mathcal{H}_{E2} = Q_0 (\nabla E^e)_0 - Q_{+1} (\nabla E^e)_{-1} + Q_{+2} (\nabla E^e)_{-2} - Q_{-1} (\nabla E^e)_{+1}$$

$$+ Q_{-2} (\nabla E^e)_{+2} \qquad \text{(III.26)}$$

In the above Q_0 is the zz component of the nuclear electrical quadrupole moment tensor. The Q_q' s in general are defined by

$$Q_0 = Q_{2,0} = \frac{1}{2}\int_{\tau_n} \rho_n r_n^2 (3\cos^2\theta_n - 1)\, d\tau_n = \frac{1}{2}\int_{\tau_n} \rho_n (3z_n^2 - r_n^2)\, d\tau_n$$

$$Q_{\pm 1} = Q_{\pm,\pm 1} = \mp\sqrt{\frac{3}{2}}\int_{\tau_n} \rho_n z_n x_{n\pm}\, d\tau_n$$

$$Q_{\pm 2} = Q_{2,\pm 2} = \sqrt{\frac{3}{8}}\int_{\tau_n} \rho_n x_{n\pm}^2\, d\tau_n \qquad \text{(III.27)}$$

where $x_{n\pm}$ is an abbreviation for $x_n \pm i\, y_n$. The (∇E^e)'s are components of the gradient of electric field at the nuclear centroid arising from the electrical charges external to the sphere of radius R. They are explicitly given by

$$(\nabla E^e)_0 = F_{2,0} = \frac{1}{2}\int_{\tau_e} \frac{\rho_e^e}{r_e^3}(3\cos^2\theta_e - 1)\, d\tau_e = -\frac{1}{6}\left[3\frac{\partial E_z^e}{\partial z} - \nabla\cdot E\right]$$

$$= -\frac{1}{2}\left[\frac{\partial E_z^e}{\partial z}\right]$$

$$(\nabla E^e)_{\pm 1} = F_{2,\pm 1} = \mp\sqrt{\frac{3}{2}}\int_{\tau_e} \frac{\rho_e^e}{r_e^3}\cos\theta_e \sin\theta_e (\cos\phi_e \pm i\sin\phi_e)\, d\tau_e$$

$$= \pm\frac{\sqrt{6}}{6}\frac{\partial E_1^e}{\partial z} = \pm\frac{\sqrt{6}}{6}\partial_\pm E_z^e \qquad \text{(III.28)}$$

$$(\nabla E^e)_{\pm 2} = F_{2,\pm 2} = \sqrt{\frac{3}{8}}\int_{\tau_e} \frac{\rho_e^e}{r_e^3}\sin^2\theta_e (\cos\phi_e \pm i\sin\phi_e)^2 d\tau_e$$

where

$$\partial_\pm = \frac{\partial}{\partial x} \pm i\frac{\partial}{\partial y} \qquad \text{(III.29)}$$

In the expansions of $\mathcal{H}_{Ek}$ in Eqs. (III.21) and (III.26), each of the indicated terms in general is itself a sum of separate terms. As a result, there is no uniqueness concerning the expansion in indicated

terms; the individual terms can be recombined into other quite different indicated terms to yield alternative expansions which superficially appear to be quite different but which on full expansion are completely equivalent. One such alternative can be obtained by combining the expressions +q and -q in Eq. (III.11) so as to make the expansion in terms of cos q ϕ and sin q ϕ instead of $e^{\pm iq\phi}$. A still different alternative is to make the expansion in terms of the Cartesian components of the tensor. In this case

$$\mathcal{H}_{E1} = - \underset{\sim}{\rho} \cdot \underset{\sim}{E}^e = + \rho_x E_x^e - \rho_y E_y^e - \rho_z E_z^e \qquad \text{(III.30)}$$

Also

$$\mathcal{H}_{E2} = - \frac{1}{6} \sum_{i=x}^{z} \sum_{j=x}^{z} Q_{ij} \left(\nabla E^e \right)_{ij} \qquad \text{(III.31)}$$

where

$$Q_{ij} = \int_{\tau_n} \rho_n \left(3 x_{ni} x_{nj} - \delta_{ij} r_n^2 \right) d\tau_n$$

$$\left(\nabla E^e \right)_{ij} = - \int_{\tau_e} \frac{\rho_e^e}{r_e^5} \left(3 x_{ei} x_{ej} - \delta_{ij} r_e^2 \right) d\tau_e \qquad \text{(III.32)}$$

and where x_{nz} denotes z_n, etc. The equivalence of Eqs. (III.30) and (III.31) to Eqs. (III.21) and (III.26) may be shown by expansion in Cartesian coordinates. In the author's book on Nuclear Moments (RAM53e) the direct calculation of Eqs. (III.30) and (III.31) from Eq. (III.9) is given.

In quantum mechanical calculations of the expectation values for the above expressions, ρ_n should be replaced by its expectation value (SCH46)

$$\rho_n = \int \psi^*(\tau_1, \cdots \tau_A)\, \rho_{op}\, \psi(\tau_1, \cdots, \tau_A)\, d\tau_1 \cdots d\tau_A \qquad \text{(III-32a)}$$

where $$\rho_{op} = \sum_i e_i\, \delta(\underset{\sim}{r}_n - \underset{\sim}{r}_i) \qquad \text{(III-32b)}$$

With this, for example from Eq. (III-16)

$$\langle Q_{k,q} \rangle = \sqrt{\frac{4\pi}{2k+1}} \int_{\tau_n} \int \psi^*(\underset{\sim}{r}_1, \cdots, \underset{\sim}{r}_A)\, r_n^k\, Y_k^q(\cos\theta_n, \phi_n)$$

$$\sum_i e_i\, \delta(\tau_n - \tau_i)\, \psi(\underset{\sim}{r}_1, \cdots, \underset{\sim}{r}_A)\, d\tau_1 \cdots d\tau_A\, d\tau_n$$

$$= \sum_{i=1}^{A} e_i \int \psi^*(\underset{\sim}{r}_1, \cdots \underset{\sim}{r}_A)\, r_i^k\, Y_k^q(\cos\theta_i, \phi_i)$$

$$\psi(\tau_1, \cdots, \tau_A)\, d\tau_1 \cdots d\tau_A \qquad \text{(III-32c)}$$

III-2.2. Theoretical restrictions on electric multipole orders. If a nuclear multipole moment is associated with a nucleus of definite spin I whose orientation properties are fully determined by the orientation of the spin angular momentum I, the possible varieties of nuclear electric multipole moments are restricted by two general theoretical theorems.

The first theorems arises from a parity consideration and is as follows : If all nuclear electrical effects arise from electrical charges, if there is no degeneracy of nuclear states with different parity, and if the nuclear Hamiltonian is unaltered by an inversion of coordinates (replacement of $\underset{\sim}{r}$ by $-\underset{\sim}{r}$), no odd (k odd) electrical multipole can exist.

The proof of this theorem is given in detail in the author's book on Nuclear Moments (RAM53e, P.23). The essentials of the argument are that if the above assumptions apply, then the wave function must be of either odd or even parity, i.e.,

$$\psi(x_n, \cdot\ y_n, z_n) = \pm\, \psi(-x_n, -y_n, -z_n) \qquad \text{(III-33)}$$

However, with either odd or even parity

$$\rho_n(x_n, y_n, z_n) = |\psi(x_n y_n z_n)|^2 = \rho_n(-x_n, -y_n, -z_n) \qquad \text{(III-34)}$$

On the other hand, for even k, the y_k^q of Eq. (III-II) is unaltered for the same coordinate inversion whereas for odd k the sign is reversed upon inversion. Therefore with the latter case in Eq.(III-I6) the integrand is of opposite sign at $(-x_n, -y_n, -z_n)$ than at (x_n, y_n, z_n). Consequently when the completed integration is carried out over both these positions the net result is zero, which proves the theorem.

In atomic problems, exceptions to the above theorem sometimes arise due to degeneracy of states with different parity. On the other hand, the energy separations of different nuclear energy levels are so large compared to normal interaction energies with nuclear electric moments that such degeneracies are very unlikely. Purcell and Ramsey (PUR50) have pointed out that the above proof depends on the assumption that the electrical effects of a nucleus arise only from electrical charges or on a related parity assumption and that these assumptions are not necessarily self obvious in the case of little-understood particles like nucleons an nuclei. Therefore, in an experiment with J.Smith (SMI5Ia, SMI5Ib), they searched with high precision for a possible electric dipole moment of the neutron. They find, however, that if such a dipole moment exists its magnitude must be less than the charge of the electron multiplied by a distance D of 5×10^{-2I} cm.

The second theorem is the following: For a nuclear spin I it is impossible to observe a nuclear electrical multipole moment of order 2^k greater than corresponding to $k = 2I$.

This theorem is also derived in detail in the author's Nuclear Moments (RAM53e, p.24). The essentials of the argument are that if ρ_n in Eq.(III-I6) is written as $\psi_n^* \psi_n$, the equation can be expressed as

$$Q_{k,q} = \int_{\tau_n} \psi_n^* r_n^k y_k^q \psi_n \, d\tau_n \qquad \text{(III-35)}$$

However, y_k^q is a possible orbital wave function for an orbital angular momentum of k (CON35, p.52) . As ψ_n is a wave function angular momentum I and by the vector addition of angular momenta (CON35, p.56), the product $y_k^q \psi_n$ can be considered a wave function of a system whose angular momentum is between $k + I$ and $|k - I|$. However, since ψ_n^* and hence $\psi_n^* r_n^k$ corresponds to angular

momentum I, they will be orthogonal to $y_k^q \psi_n$ unless they can correspond to the same angular momentum eigenvalue . Therefore, to prevent Eq.(III-35) from vanishing, I must lie between k + I and $|k - I|$, or $k \leq 2I$, which is the desired theorem.

Analogous theorems to the above are applicable to the field quantities $F_{k',q}$ of Eq.(III- I7 when they arise from an atom or molecule whose angular momentum is J. Thus, $F_{k',q}$ vanishes unless $k \leq 2J$. Even a nucleus with a large I and a nuclear quadrupole moment can have no electric quadrupole interaction energy with an atom whose $J = \frac{1}{2}$.

III-2.3. Nuclear electric quadrupole interactions. The most important orientation dependent electrical interaction of a nucleus is that arising from its electric quadrupole moment . It is therefore desirable to study in further detail the effect of a nuclear quadrupole interaction upon the energy levels of an atom or molecule . In general, the calculation of the energy levels is performed either by a perturbation theory calculation or by the solution of the secular equation which results from the matrix of the Hamiltonian when the quadrupole contributions of Eq. (III-26) are included. In either case, the matrix elements of the quadrupole terms must be calculated for the representation that is used.

The matrix elements may be calculated in a number of different ways. One procedure is the direct use of a general formula given in the next section for any order of multipole . Another procedure, the one to be followed in this section, depends on the observations that one usually uses a representation based on the angular momentum operators of the atom, molecule, or nucleus. Matrix elements of the functions of the angular momentum operators can usually be calculated easily in such representations from the standard forms (CON35) for $(I, m_I | \underline{I} | I' m_I')$ that all such matrix elements vanish except those corresponding to

$$(I, m_I | I_z | I\ m_I) = m_I \qquad \text{(III-35a)}$$

$$(I, m_I \pm 1 | I_\pm | I, m_I) = \sqrt{(I \mp m_I)(I \pm m_I + 1)}$$

where

$$\underline{I} = I_z \underline{k} + \tfrac{1}{2} I_+ (\underline{i} - i\underline{k}) + \tfrac{1}{2} I_- (\underline{i} + i\underline{k}) \qquad \text{(III-35b)}$$

where $\underline{i}, \underline{j}, \underline{k}$ are the Cartesian unit vectors. If, therefore, the quadrupole moment tensor can be expressed as simply proportional to a tensor in the

angular momentum variables the calculation of the matrix elements can be simplified .

The theorem necessary to make the above transformation is most simply taken from group theory or from the tensor commutation rules (WIG3I,RAC42, WEY39,CAS35,CAS36,NIE48). However,it can also be shown directly from simple multiplications and from the matrix elements for vectors.Such a proof is straightforward,though numerically tedious,and is given in the author's book on Nuclear Moments (RAM53 f.p.2I). By whatever the means of proof,the theorem applies to all tensors which transform under space rotations in the same way as does Eq. (III-II). This will be true of tensors that are similar in form (including symmetry with respect to commutation) to Eqs.(III-II),(III-I6), (III-I7) but which may be based on some vector $\underset{\sim}{A}$ which is other than the space vector $\underset{\sim}{r}$ upon which the above equations depend provided $\underset{\sim}{A}$ satisfies the same commutation rules with respect to the total angular momentum operator $\underset{\sim}{I}$ as does $\underset{\sim}{r}$ (the dependence of the above equation on the components of $\underset{\sim}{r}$ is most clearly illustrated when they are expressed in forms similar to Eq.(III-27).The significance of the phrase in the preceding sentence " including symmetry with respect to commutation" is,for example,that one component of the tensor in Eq(III-27) is proportional to $z_n x_{n\pm1}$ which is also equal to $x_{n\pm1} z_n$ since the components of r commute among themselves. Consenquently,the appropriate form for the corresponding component of a tensor based on $\underset{\sim}{I}$ is :

$$\tfrac{1}{2}\left(I_{\pm} I_z + I_z I_{\pm} \right)$$

since this expression is symmetric with respect to commutation . The theorem which applies to tensors satisfying the above conditions is that the quantum mechanical matrix elements diagonal in $\underset{\sim}{I}$ of all such tensors have the same orientation dependence,e.g.,the same dependence on the magnetic quantum number m_I . As discussed and proved in the author's Nuclear Moments (RAM53e) a similar theorem applies to tensors,such as those of Eq.(III-32),which are symmetric with respect to the interchanges of any two indices, and for which all spurs vanish (FAL50). It also applies to tensors based on more than one vector,as can be shown by the mathematical process of polarization (FAL50,WEY39).

As a result of the above theorem,the calculation of matrix elements diagonal in $\underset{\sim}{I}$ can be simplified by writing the tensor components as proportional to the corresponding expressions in components of I,which by the above theorem have the same dependence on m_I. Thus,from Eq (III-27) one can write :

$$Q_0 = \tfrac{1}{2}\int \rho_n (3Z_n^2 - r_n^2)\, d\tau_n = C\,\tfrac{1}{2}(3I_z^2 - I^2)$$

$$Q_{\pm 1} = \sqrt{\tfrac{3}{2}}\, C\,\tfrac{1}{2}(I_z I_\pm + I_\pm I_z) \qquad \text{(III-36)}$$

$$Q_{\pm 2} = \sqrt{\tfrac{3}{2}}\, C\, I_\pm^2$$

The value of C in the above equation can be related to the scalar quantity Q which is conventionally taken as the measure of a nuclear quadrupole moment . By convention the quantity Q, often called simply the quadrupole moment, is defined by :

$$Q \equiv \frac{1}{e}\int_{\tau_n} \rho_{n\, m_I = I}\,(3Z_n^2 - r_n^2)\, d\tau_n \qquad \text{(III-37)}$$

where $\rho_{n\, mI = I}$ is the charge density when the nucleus is in the orientation state with $m_I = I$. If $(I\, m_I | Q_0 | I\, m_I')$ designates a matrix element of Q_0, the value of C can be related to Q so follows with the aid of Eqs. (III-36) and (III-37)

$$Q = \frac{2}{e}(II|Q_0|II) = \frac{C}{e}\left[3I^2 - I(I+1)\right]$$
$$= \frac{C}{e}\, I(2I-1) \qquad \text{(III-38)}$$

Therefore, Eq. (III-36) can be written as

$$Q_0 = \frac{eQ}{2I(2I-1)}\left[3I_z^2 - I(I+1)\right]$$

$$Q_{\pm 1} = \frac{\sqrt{6}}{2}\,\frac{eQ}{2I(2I-1)}\left[I_z I_\pm + I_\pm I_z\right] \qquad \text{(III-39)}$$

$$Q_{\pm 2} = \frac{\sqrt{6}}{2}\,\frac{eQ}{2I(2I-1)}\, I_\pm^2$$

Since the orientation of the atom or molecule is specified by the orientation of its angular momentum J, a similar procedure to the above is applicable in the calculation for matrix elements diagonal in J of the tensor components of the gradients of the molecular electric field . Thus, for such calculations, Eq. (III-28) can be written as

$$(\nabla E^e)_0 = \frac{e q_J}{2J(2J-1)}\left(3J_z^2 - J(J+1)\right)$$

$$(\nabla E^e)_{\pm 1} = \mp\frac{\sqrt{6}}{2}\,\frac{e q_J}{2J(2J-1)}\left[J_z J_\pm + J_\pm J_z\right] \qquad \text{(III-40)}$$

$$(\nabla E^e)_{\pm 2} = \frac{\sqrt{6}}{2}\,\frac{e q_J}{2J(2J-1)}\, J_\pm^2$$

where q_J is defined by

$$q_j \equiv \frac{1}{e}\int_{\tau_e} \frac{3\cos^2\theta_e - 1}{r_e^3}\,\rho^e_{e,m_j=j}\,d\tau_e$$

$$= \frac{1}{e}\int_{\tau_e} \frac{3z_e^2 - r_e^2}{r_e^5}\,\rho^e_{e,m_j=j}\,d\tau_e = \frac{1}{e}\left\langle \frac{\partial^2 V^e}{\partial z^2}\right\rangle_{JJ} \qquad \text{(III-4I)}$$

The symbol $\langle\ \rangle_{JJ}$ indicates the average potential when the molecule is in the state with $m_j = J$.

In a similar fashion, when the tensors are written in the form of Eq (III-32) they may be reexpressed in calculations of matrix elements diagonal in I and J as

$$Q_{ij} = \frac{eQ}{I(2I-1)}\left[3\,\frac{I_iI_j + I_jI_i}{2} - \delta_{ij}I(I+1)\right] \qquad \text{(III-42)}$$

$$(\nabla E^e)_{ij} = -\frac{eq_J}{J(2J-1)}\left[3\,\frac{J_iJ_j + J_jJ_i}{2} - \delta_{ij}J(J+1)\right]$$

The exact forms for the coefficients are obtained analogously to the above by considering $(I\,I|Q_{zz}|I\,I)$. Details of this simple calculation are given in the author's Nuclear Moments (RAM53e, p. I7).

A still different and often useful form for the quadrupole interaction is

$$\mathcal{H}_{EQ} = \frac{e^2qQ}{2I(2I-1)J(2J-1)}\left[3(\underline{I}\cdot\underline{J})^2 + \frac{3}{2}\underline{I}\cdot\underline{J} - I(I+1)J(J+1)\right] \qquad \text{(III-43)}$$

A derivation of this equation from Eqs. (III-42) is given in Appendix C. The relation can also be proved by expanding Eq. (III-43) in terms of the I_z and $I_{\pm}$ components and comparing it with the similar expansion of Eqs. (III-26), (III-39), and (III-40). In either of these two derivations the components of $\underline{J}$ should be commuted among themselves only in accordance with the usual commutation rules (CON35)

$$[J_z, J_\pm] = J_zJ_\pm - J_\pm J_z = \pm J_\pm$$

$$[J_+, J_-] = 2J_z \qquad \text{(III-44)}$$

$$[J_i, J_j] = iJ_{ixj}$$

The subscript ixj indicates the Cartesian component perpendicular to the i and j components as in a vector product.

In diatomic molecules either with identical or non identical nuclei the quadrupole interaction can be obtained by applying either Eq.(III-42) or (III-43) to each nucleus separately. However, in the case of identical nuclei it is usually more convenient to express the interaction in terms

of the resultant angular momentum $\underset{\sim}{I}_R = \underset{\sim}{I}_1 + \underset{\sim}{I}_2$ since I_R is a good quantum number in a rotational state of definite J because of the symmetry requirements on the wave function with identical nuclei. This reexpression is calculated in Appendix C and two alternative forms, for the calculation of matrix elements diagonal in I_R are given . One of these is that the interaction is as given in Eqs. (III-3I) and (III-42) except that Q_{ij} is replaced in the homonuclear diatomic case by

$$q_{ij} = \frac{e\,q}{I_1(2I_1-1)}\left[1 - \frac{I_R(I_R+1)+4I_1(I_1+1)}{(2I_R-1)(2I_R+3)}\right]\left[3\frac{I_{Ri}I_{Rj}+I_{Rj}I_{Ri}}{2} - \delta_{ij}I_R(I_R+1)\right] \quad \text{(III-44a)}$$

The other alternative form derived in Appendix C is that the interaction in the diatomic homonuclear case for matrix elements diagonal in I is

$$\mathcal{H}_{E_2} = \frac{e^2 q_J Q}{2I_1(2I_1-1)J(2J-1)}\left[1 - \frac{I_R(I_R+1)+4I_1(I_1+1)}{(2I_R-1)(2I_R+3)}\right]\left[3(\underset{\sim}{I}_R\cdot\underset{\sim}{J})^2 + \frac{3}{2}\underset{\sim}{I}_R\cdot\underset{\sim}{J} - I_R^2J^2\right] \quad \text{(III-44b)}$$

Since the values of I(I+I) and J(J+I) in Eqs. (III-43) and (III-44b) are independent of the molecular and nuclear orientations and since most molecular beam experiments measure differences in energy as an orientation is changed, these terms make no net contribution to the final result and are often omitted (or included in the reference energy from which the other energies are expressed). With such an omission the above referred to equations can alternatively be written as

$$\mathcal{H}'_{E_2} = hb_1 2\underset{\sim}{I}\cdot\underset{\sim}{J}\left(2\underset{\sim}{I}\cdot\underset{\sim}{J}+1\right) \quad \text{(III-44c)}$$

The above relations are applicable to the calculation of tensor matrix elements diagonal in J for both atoms and molecules. For atoms, the above forms are also those most frequently used. However, in the case of molecules it is usually convenient to express q_j in a form that reduces the implicit dependence of q_j upon J. Most of this dependence arises from the definition of q_j in Eq. (III-4I) being based on the state with $m_j = J$.

For this reason, q_j would vary with J even if the shape of the molecule were unaltered by centrifugal stretching since an angular momentum vector J in the state $m_j = J$ can lie much closer to the z axis if J is large than if J is small. This difficulty can be avoided if q_j is expressed in terms of a coordinate system fixed within the molecule .

Consider a symmetric top molecule and let z_0 be a coordinate along the

axis of symmetry while x_o is perpendicular to the axis of symmetry but in the plane of z and z_o while y_c is mutually perpendicular to x_o and z_o . . . ,

if θ'' is the angle between z and z_o,

$$q_{JJ} = \frac{1}{e}\left\langle \frac{\partial^2 V^e}{\partial z^2} \right\rangle_{JJ} = \frac{1}{e}\left\langle \frac{\partial^2 V^e}{\partial z_o^2}\cos^2\theta'' + \frac{\partial^2 V^e}{\partial x_o^2}\sin^2\theta'' \right\rangle_{JJ} \qquad \text{(III-45)}$$

since $\partial^2 V^e / \partial z_o \partial x_o$ vanishes by symmetry and since $\partial x_o / \partial z = \sin\theta''$. However, from Laplace's equation and from the symmetry between w_o and y_o,

$$\frac{\partial^2 V^e}{\partial x_o^2} = \frac{\partial^2 V^e}{\partial y_o^2} = -\frac{1}{2}\frac{\partial^2 V^e}{\partial z_o^2} \qquad \text{(III-46)}$$

so

$$q_{JJ} = \frac{1}{e}\frac{\partial^2 V^e}{\partial z_o^2}\left\langle \frac{3\cos^2\theta'' - 1}{2} \right\rangle_{JJ} \qquad \text{(III-47)}$$

For linear molecules the indicated average can be calculated from the wave function for a rotating linear molecule in the rotational state with $m_J = J$ which is (CON35, p.54)

$$\psi_{JJ}(\theta'', \phi'') = \frac{(-1)^J}{\sqrt{2\pi}}\sqrt{\frac{(2J+1)!}{2}}\,\frac{1}{2^J J!}\sin^J\theta''\, e^{iJ\phi''} \qquad \text{(III-48)}$$

The appropriate integral can be evaluated (PEI99, Eq.483) with the result that

$$q_{JJ} = -\frac{J}{2J+3}\frac{1}{e}\frac{\partial^2 V^e}{\partial z_o^2} = -\frac{J}{2J+3}\frac{1}{e}q \qquad \text{(III-49)}$$

where $q \equiv \left(\partial^2 V^e / \partial z_o^2\right)$ (III-50)

where K is the quantum number for the projection of J along the axis of symmetry. For K = 0 Eqs. (III-49) and (III-50) agree as they should since K is necessarily zero for a linear molecule. The forms for asymmetric tops have been given by Bragg (BRA48) and others (KNI49a).

With the above, Eq. (III-43), for example, can be written as

$$\mathcal{H}_{E2} = -\frac{eqQ}{2I(2I-1)(2J+3)(2J-1)}\left[1 - \frac{3K^2}{J(J+1)}\right]\left[3(I\cdot J)^2 + \frac{3}{2}I\cdot J - I(I+1)J(J+1)\right] \qquad \text{(III-5I)}$$

Unlike q_{JJ}, the quantity q is independent of J except for such small effects as centrifugal stretching .[I]

(see footnote next page)

It is important to realize that V^e in Eq. (III-49a) is not the potential V of all electrical charges from the rest of the molecule, including the electron density of the molecule inside the nuclear radius . The difference between these two can readily be evaluated. Let ρ_1 be the electronic charge density at the nucleus, and assume that in the immediate vicinity of the nucleus it is spherically symmetric . Then, as $V - V^e$ arises solely from the spherically symmetric charge distribution ρ_1, from Poisson's equation

$$\frac{\partial^2(V-V^e)}{\partial x^2}+\frac{\partial^2(V-V^e)}{\partial y^2}+\frac{\partial^2(V-V^e)}{\partial z^2}=4\pi\rho_1=3\frac{\partial^2(V-V^e)}{\partial z^2} \qquad \text{(III-52)}$$

so

$$q=\frac{\partial^2 V^e}{\partial z_0^2}=\frac{\partial^2 V}{\partial z_0^2}-\frac{4\pi}{3}\rho_1 \qquad \text{(III-53)}$$

Due to its spherical symmetry, the ρ_1 term does not affect the nuclear quadrupole interaction.

When the interaction between the vectors $\underset{\sim}{I}$ and $\underset{\sim}{J}$ is large compared to the interaction of either with any other field, Eq.(III-5I) can be expressed in an alternative form . Let $\underset{\sim}{F}$ be the vector sum of $\underset{\sim}{I}$ and $\underset{\sim}{J}$ so $\underset{\sim}{F} = \underset{\sim}{I} + \underset{\sim}{J}$ with quantum number F. Then $\underset{\sim}{I} \cdot \underset{\sim}{J}$ can be evaluated from

$$\underset{\sim}{F}^2 = (\underset{\sim}{I} + \underset{\sim}{J})^2 = \underset{\sim}{I}^2 + \underset{\sim}{J}^2 + 2\underset{\sim}{I} \cdot \underset{\sim}{J} \qquad \text{(III-54)}$$

so

$$\underset{\sim}{I} \cdot \underset{\sim}{J} = \tfrac{1}{2}(\underset{\sim}{F}^2 - \underset{\sim}{I}^2 - \underset{\sim}{J}^2) = \tfrac{1}{2}(F(F+I)-I(I+I)-J(J+1)) = \tfrac{1}{2}C \qquad \text{(III-55)}$$

This may be used in Eq. (III-5I), whence the quadrupole energy W_{E2}(when K=0) becomes

$$W_{E2} = -\frac{eqQ}{4(2I-I)(2J+3)(2J-1)}\left[\frac{3}{8}C(C+I)-\tfrac{1}{2}I(1+1)J(J+1)\right] \qquad \text{(III-56)}$$

---

[1] There is much notational confusion in the literature on q_j and related quantities. Kellogg, Rabi, Ramsey, and Zacharias (KEL40) use the symbol q for what is here called q_j and Casimir (CAS36) uses C_e for the same quantity. Feld(FEL47) and Gordy(GOR53) use $\partial^2 V/\partial z^2$ for the quantity here designated as q and Ramsey (RAM53e) uses $\partial^2 V^e/\partial z_0^2$ for the same quantity. Townes and Bardeen (TOW54) use q in the same sense as here but the quantity they call q_j is eq_j in the present notation. Nordsieck(NOR40) and Nierenberg(NIE47) have sometimes used a still different quantity q' which is related to the present notation by $q' = (1/2e)q$. Although q will be the notation dominantly used in the present book, the more explicit $\partial^2 V^e/\partial z_0^2$ will occasionally be used .

Often magnetic interactions with an external field are large compared with the interactions coupling $\underset{\sim}{I}$ to $\underset{\sim}{J}$. In such cases a representation based on the magnetic quantum numbers m_I and m_J is often most suitable. In such a representation $\mathcal{H}_{E2}$ is not diagonal. However, the appropriate diagonal and non-diagonal matrix elements can be calculated from Eqs. (III-39) etc. or from Eq. (III-5I) . These matrix elements are calculated in Eq. (C-24) of Appendix C by the procedure used by Kellogg, Rabi, Ramsey, and Zacharias (KEL40). Their paper (KEL40) also gives a detailed example of the application of the above quadrupole interaction expressions to a practical problem, the evaluation of the quadrupole moment of the deuteron.

In order that the electrical quadrupole moment Q may be experimentally determined from the observed interaction energy, an evaluation of q_j or of $\partial^2 V^e / \partial z_o^2$ is required. Various methods for obtaining these quantities are available. The methods will be discussed in Chapter XI

III-2.4. Multipole moments of arbitrary order . Matrix elements diagonal in I of general multipole moments can be calculated by rewriting them as proportional to corresponding expressions based on the angular momentum I with the use of the general theorem given in the previous section.

On the other hand, it is often more convenient to use direct general expressions for the matrix elements that have been derived by Racah (RAC42) and Wigner (WIG3I). These expressions also give matrix elements non-diagonal in I; they are as follows for the components of say $Q_{k,q}$ in an $\alpha I m_I$ representation.

$$(\alpha I m_I | Q_{k,q} | \alpha' I' m_I') = (-1)^{I+m_I} (\alpha I \| Q_k \| \alpha' I') \, V(II'k; -m_I m_I' q) \quad \text{(III-57)}$$

where

$$V(abc; \alpha\beta\gamma) = \left[(a+b-c)!(a+c-b)!(b+c-a)!/(a+b+c+1)!\right]^{\frac{1}{2}} v(abc; \alpha\beta\gamma) \quad \text{(III-58)}$$

and

$$v(abc; \alpha\beta\gamma) = \delta(\alpha+\beta+\gamma, 0) \sum_z (-1)^{c-\gamma+z} \frac{\left[(a+\alpha)!(a-\alpha)!(b+\beta)!(b-\beta)!(c+\gamma)!(c-\gamma)!\right]^{\frac{1}{2}}}{z!(a+b-c-z)!(a-\alpha-z)!(b+\beta-z)!(c-b+\alpha+z)!(c-a-\beta+z)!} \quad \text{(III-59)}$$

The summation parameter z takes on all integral values consistent with the factorial notation, the factorial of negative numbers being meaningless . $(\alpha I \| Q_k \| \alpha' I')$ indicates the proportionality factor which is independent of m_I. Similar expressions are available for matrix elements of $F_{k,q}$ except that J

and m_J replace I and m_I in the above . The above formulae can be derived directly from the commutation rules for vectors (RAC42).

Alternatively,Racah (RAC42) has given formulae for the matrix elements of the matrix scalar products as Eq.(III-I5) in a $\gamma JIFm$ representation with the result that

$$(\gamma J I F m | F_k \cdot Q_K | \gamma' J' I' F' m') = \delta(F,F')\,\delta(m,m')(-1)^{J+I'-F} \sum_{\gamma''} (\gamma J \| F_k \| \gamma'' J')(\gamma'' I \| Q_K \| \gamma' I') W(JI\,J'I';Fk)$$

where W (abcd ; ef) is a so-called Racah coefficient defined by

$$W(abcd;ef) = \left[\frac{(a+b-e)!(a+e-b)!(b+e-a)!(c+d-e)!(c+e-d)!(d+e-c)!}{(a+b+e+1)!(c+d+e+1)!(a+c+f+1)!(b+d+f+1)!} \right]^{\frac{1}{2}} w(abcd;ef) \qquad \text{(III-6I)}$$

with

$$w(abcd;ef) = \sum_z (-1)^z \frac{(a+b+c+d+1-z)!}{(a+b-e-z)!(c+d-e-z)!(a+c-f-z)!(b+d-f-z)!(e+f-a-d+z)!(e+f-b-c+z)!}$$

(III-62)

Numerical values of the Racah coefficients for different values of the parameters have been given by several authors (BIE52,OBI54).

III-3 Magnetic Interaction

III.3.I. Magnetic multipoles. All electronic and nuclear magnetic effects presumably arise from circulating currents including the various anomalous moments associated with spins (SCH37,SCH48). Therefore, the mutual potential energy of magnetic interaction between the nucleus with nuclear current density j_n and the electrons with current density j_e producing a vector potential A_e at the nucleus may be taken (CAS36,CAS42,ROS48) as

$$\mathcal{H}_M = -\frac{1}{c}\int_{\tau_n} \underline{j}_n \cdot \underline{A}_e \, d\tau_n = -\frac{1}{c}\int_{\tau_e} \underline{j}_e \cdot \underline{A}_n \, d\tau_e \qquad \text{(III-63)}$$

However,from the equation of continuity for stationary current distributions

$$\nabla \cdot \underset{\sim}{j} = -\dot{\rho} = 0 \qquad \text{(III-64)}$$

whence $\underset{\sim}{j}_n$ can be derived from a vector potential $\underset{\sim}{m}_n$ so that

$$\underset{\sim}{j} = c \nabla \times \underset{\sim}{m}_n \qquad \text{(III-65)}$$

and $\underset{\sim}{m}_n$ may be taken as zero outside of the nucleus. As a result of this, the following equation applies since the first term equals zero if the surface S_n surrounds but lies outside the nucleus

$$\begin{aligned} \mathcal{H}_M &= -\int_{S_n} \underset{\sim}{A}_e \times \underset{\sim}{m}_n \cdot d\underset{\sim}{S}_n - \int (\nabla_n \times \underset{\sim}{m}_n) \cdot \underset{\sim}{A}_e \, d\tau_n \qquad \text{(III-66)} \\ &= -\int_{\tau_n} \nabla_n \cdot (\underset{\sim}{A}_e \times \underset{\sim}{m}_n) d\tau_n - \int \underset{\sim}{A}_e \cdot (\nabla_n \times \underset{\sim}{m}_n) d\tau_n \\ &= -\int_{\tau_n} \underset{\sim}{m}_n \cdot \nabla_n \times \underset{\sim}{A}_e \, d\tau_n = -\int_{\tau_n} \underset{\sim}{m}_n \cdot \underset{\sim}{H}_e \, d\tau_n \end{aligned}$$

The standard vector and tensor relations summarized in Appendix B are used repeatedly in the preceding and following equations .

For that portion of $\underset{\sim}{H}_e$ which arises exclusively from electron currents in a region $\tau_e^{\,e}$ external to the nucleus, as indicated by the superscript e in $\tau_e^{\,e}$ and $\underset{\sim}{H}_e^{\,e}$, one can write if r equals $|\underset{\sim}{r}_e - \underset{\sim}{r}_n|$

$$\begin{aligned} H_e^{\,e} &= \nabla_n \times \underset{\sim}{A}_e^{\,e} = \nabla_n \times \frac{1}{c} \int_{\tau_e^{\,e}} \frac{\underset{\sim}{j}_e}{r} d\tau_e \\ &= \nabla_n \times \int_{\tau_e^{\,e}} \frac{\nabla_e \times \underset{\sim}{m}_e}{r} d\tau_e \qquad \text{(III-67)} \\ &= \nabla_n \times \left[\int_{\tau_e^{\,e}} \nabla_e \times \frac{\underset{\sim}{m}_e}{r} d\tau_e - \int_{\tau_e^{\,e}} \left(\nabla_e \frac{1}{r}\right) \times \underset{\sim}{m}_e \, d\tau_e \right] \\ &= \nabla_n \times \left[-\int_{S'_e + S''_e} \frac{\underset{\sim}{m}_e}{r} d\underset{\sim}{S}_e + \int_{\tau_e^{\,e}} \left(\nabla_n \frac{1}{r} \times \underset{\sim}{m}_e \, d\tau_e \right) \right. \end{aligned}$$

S_e' in the above indicates an inner surface defining the electron current distribution, which for H_e^e is assumed wholly external to the nucleus, and S_e'' an outer surface to the electron distribution. If $\underset{\sim}{m}_e$ vanishes over both S_e' and S_e'' the first integral on the right in the last of the above expressions vanishes. Then

$$H_e^e = \int_{\tau_e^e} \nabla_n \times \nabla_n \left(\frac{\underset{\sim}{m}_e}{r}\right) d\tau_e = \int_{\tau_e^e} \left(\nabla_n \nabla_n \cdot - \nabla_n^2\right) \frac{\underset{\sim}{m}_e}{r} \, d\tau_e$$

$$= \nabla_n \int_{\tau_e^e} \nabla_n \cdot \frac{\underset{\sim}{m}_e}{r} \, d\tau_e = \nabla_n \int_{\tau_e^e} \underset{\sim}{m}_e \cdot \nabla_n \frac{1}{r} \, d\tau_e$$

$$= -\nabla_n \int_{\tau_e^e} \underset{\sim}{m}_e \cdot \nabla_e \frac{1}{r} \, d\tau_e$$

$$= -\nabla_n \int_{\tau_e^e} \nabla_e \cdot \frac{\underset{\sim}{m}_e}{r} \, d\tau_e + \nabla_n \int_{\tau_e^e} \frac{\nabla_e \cdot \underset{\sim}{m}_e}{r} \, d\tau_e \qquad \text{(III-68)}$$

$$= -\nabla_n \int_{S_e' + S_e''} \frac{\underset{\sim}{m}_e}{r} \cdot d\underset{\sim}{S}_e + \nabla_n \int_{\tau_e} \frac{\nabla_e \cdot \underset{\sim}{m}_e}{r} \, d\tau_e$$

$$= \nabla_n \int_{\tau_e} \frac{\nabla_e \cdot \underset{\sim}{m}_e}{r} \, d\tau_e$$

Therefore,

$$\mathcal{H}_n^e = -\int_{\tau_n} \underset{\sim}{m}_n \cdot \left[\nabla_n \int_{\tau_e^e} \frac{\nabla_e \cdot \underset{\sim}{m}_e}{r} \, d\tau_e\right] d\tau_n$$

$$= -\int_{\tau_n} \nabla_n \cdot \left[\underset{\sim}{m}_n \int_{\tau_e^e} \frac{\nabla \cdot \underset{\sim}{m}_e}{r} \, d\tau_e\right] d\tau_n + \int_{\tau_n} \left(\nabla_n \cdot \underset{\sim}{m}_n\right)_{\tau_e^e} \frac{\nabla_e \cdot \underset{\sim}{m}_e}{r} \, d\tau_e$$

(III-69)

If the divergence theorem is applied to the first term on the right it vanishes since m_n has already been taken equal to zero over a surface surrounding the nucleus . Therefore,

$$\mathcal{H}_M^e = \int_{\tau_e} \int_{\tau_n} \frac{(-\nabla_n \cdot \underset{\sim}{m}_n)(-\nabla_e \cdot \underset{\sim}{m}_e)}{r} d\tau_e d\tau_n \qquad (III-70)$$

It should be noted that this is exactly the same as the electric interaction of Eq.(III-5) except for the correspondence

$$\rho_e \longrightarrow -\nabla_e \cdot \underset{\sim}{m}_e$$

$$\rho_n \longrightarrow -\nabla_n \cdot \underset{\sim}{m}_n \qquad (III-7I)$$

As a result,all the theorems on multipole expansion proved in Section III-2 are equally applicable here except for the substitutions of Eq.(III-7I).The various multipole moments are then defined by a combination of Eq.(III-7I) with Eqs.(III-I6),(III-22),and (III-$\frac{3}{3}$). It should be noted,however,that Eq. (III-70) is correct only when m_e vanishes throughout the nuclear volume as indicated by the superscript e whereas Eqs. (III-5) and (III-63) apply generally. For the corrections which enter when the contributions from m_e inside the nuclear volume are included,the reader is referred to Chapter IX and to the original papers of Bohr and Weisskopf (BOH50,BOH5I,BOH53).

From the above derived correspondence,the restrictions on multipole order obtained in Section III-2.2 can be applied to magnetic multipoles. One important modification,however,arises from the fact that the electrical charge density ρ_n is normally assumed to be a scalar in which case $-\nabla_n \cdot \underset{\sim}{m}_n$ is a pseudoscalar (since from Eq. III-65 $\underset{\sim}{m}_n$ is a pseudo vector or axial vector). As a result $-\nabla_n \cdot \underset{\sim}{m}_n$ reverses sign upon inversion of coordinates while ρ_n does not . Therefore the parity argument of Section III-2.2 is exactly reversed and it is even (even* .) magnetic multipoles which cannot exist. The restrictions on magnetic multipoles therefore are : (a) For a nuclear spin I,it is impossible to observe a nuclear magnetic multipole moment of order 2^k greater than that corresponding k = 2I. (b) No even (k even) magnetic multipole moment can exist if all nuclear magnetic effects arise from the circulation of electrical charges,if there is no degeneracy of nuclear states with different parity,and if the nuclear Hamiltonian is unaltered by an inversion of coordinates. Corresponding theorems apply to the atomic field gradients

with which the nuclear multipole moments interact.

Magnetic multipole moments for a definite nuclear spin I can be expressed in terms of the nuclear angular momentum operators with the aid of the general theorem given in Section III-2.3. Alternatively, matrix elements for magnetic multipole moments or for magnetic multipole interactions can be evaluated with the general expressions given in Section III-2.4.

The lowest magnetic multipole moment by the above theorems is the magnetic dipole moment and the next lowest is the magnetic octupole moment . Pauli (PAU24) first suggested that atomic hyperfine structure might be due to nuclear magnetic dipole moments and until recently only magnetic dipole moments have been observed . For this reason the most detailed discussion of individual magnetic multipole interactions will be given for magnetic dipoles, as in the next section. However, Jaccarino, Zacharias, King, Statten, and Stroke (JAC54, TOLI3, CAS42, SCH54) and others (KUS54, DAL54) have recently found evidence for magnetic octupole interactions in iodine, indium, and gallium, as discussed in detail in Section IX-9 . The theory of magnetic octupole interactions has been given in detail by Casimir and Karreman(CAS42) and by Schwartz (SCH54). A summary of the results of the magnetic octupole theory is given in Section IX-9.

In the use of an equation such as E . (III-63), an expression for the electron current density j e is required . The electron current density consists of two parts j_{e1} and j_{e2} of which the first is the orbital current density which non-relativistically is given (PAU32) as usual by

$$j_{e1} = -\frac{e\hbar}{2mi}\left(\psi^* \nabla \psi - \psi \nabla \psi^*\right) \qquad \text{(III-7Ia)}$$

The other j_{e2}, is the spin current density and from Eq. (III-6), it is given in a non-relativistic theory (PAU32)

$$j_{e2} = c\,\nabla \times \underset{\sim}{m}_n = c\nabla \times \left[\psi^* \left(-\frac{e\hbar}{2mc}\,\underset{\sim}{\sigma}\right)\psi\right]$$

$$= -\frac{e\hbar}{2mc}\,\nabla \times \left[\psi^*\,\underset{\sim}{\sigma}\,\psi\right] \qquad \text{(III-7Ib)}$$

Similar expressions apply to the neutron and proton current densities but the value of e in Eq. (III-7Ia) must be taken as zero for neutrons and -e for protons while Eq. (III-7Ib) must be multiplied by approximate factors to

allow for the anomalous nucleon magnetic moments . Relativistically, the total probability current density is given (PAU32) by $S^{\mu} = \bar{\psi} \gamma^{\mu} \psi$ but Gordon has shown this also may be interpreted as consisting of a convection current and spin current part (PAU32, SCH49).

III-3.2. Magnetic Dipoles. From Eqs. (III-22), (III-30), and (III-7I) an expression for the nuclear magnetic dipole moment is

$$\underset{\sim}{\mu}_1 = \int_{\tau_n} (-\nabla_n \cdot \underset{\sim}{m}_n) \underset{\sim}{r}_n \, d\tau_n \qquad \text{(III-72)}$$

With the aid of the vector transformation in Appendix B, this can alternatively be written as

$$\begin{aligned} \underset{\sim}{\mu}_1 &= -\int_{\tau_n} \nabla_n \cdot (\underset{\sim}{m}_n \underset{\sim}{r}_n) \, d\tau_n + \int_{\tau_n} (\underset{\sim}{m}_n \cdot \nabla_n) \underset{\sim}{r}_n \, d\tau_n \\ &= -\int_{S_n} d\underset{\sim}{S}_n \cdot (\underset{\sim}{m}_n \underset{\sim}{r}_n) + \int_{\tau_n} \underset{\sim}{m}_n \, d\tau_n \\ &= \int_{\tau_n} \underset{\sim}{m}_n \, d\tau_n \end{aligned} \qquad \text{(III-73)}$$

A still different expression can be obtained from the relations in Appendix B since

$$\begin{aligned} \frac{1}{c} \underset{\sim}{r}_n \times \underset{\sim}{j}_n &= \underset{\sim}{r}_n \times \nabla_n \times \underset{\sim}{m}_n = \nabla_n (\underset{\sim}{r}_n \cdot \underset{\sim}{m}_n) - (\underset{\sim}{r}_n \cdot \nabla_n) \underset{\sim}{m}_n - (\underset{\sim}{m}_n \cdot \nabla_n) \underset{\sim}{r}_n - \underset{\sim}{m}_n \times (\nabla \times \underset{\sim}{r}_n) \\ &= \nabla_n (\underset{\sim}{r}_n \cdot \underset{\sim}{m}_n) - (\underset{\sim}{r}_n \cdot \nabla_n) \underset{\sim}{m}_n + \underset{\sim}{m}_n \\ &= \nabla_n (\underset{\sim}{r}_n \cdot \underset{\sim}{m}_n) - \nabla_n \cdot (\underset{\sim}{r}_n \underset{\sim}{m}_n) + 2 \underset{\sim}{m}_n \end{aligned}$$

(III-74)

Therefore,

$$\vec{\mu}_{nI} = \int_{\tau_n} \vec{m}_n \, d\tau_n = \frac{1}{2c} \int_{\tau_n} \vec{r}_n \times \vec{j}_n \, d\tau_n - \frac{1}{2} \int \nabla_n (\vec{r}_n \cdot \vec{m}_n) \, d\tau_n$$

$$+ \frac{1}{2} \int \nabla_n \cdot (\vec{r}_n \vec{m}_n) \, d\tau_n$$

$$= \frac{1}{2c} \int_{\tau_n} \vec{r}_n \times \vec{j}_n \, d\tau_n \qquad \text{(III-75)}$$

since two of the terms in the next to last step vanish when transformed to surface integrals .

From Eqs. (III-30) and (III-72), the magnetic dipole interaction can be taken as

$$\mathcal{H}_{Ml} = -\vec{\mu}_{nI} \cdot (\vec{H}_J + \vec{H}_0) \qquad \text{(III-76)}$$

where $\vec{H}_0$ is the externally applied magnetic field and $\vec{H}_j$ the magnetic field at the nucleus arising from the rest of the atom or molecule which has angular momentum J in units of $\hbar$. As discussed in the preceding section, this form of magnetic interaction implies the assumtion that the electron distribution inside the nucleus does not contribute appreciably to the interaction energy . It can be shown, as discussed in Chapter IX (CAS35, BOH50, BOH5I), that this contribution is only a few per cent even in the worst case - heavy nuclei and S electrons . However, this contribution is responsible for the important hyperfine structure anomally discussed in Chapter IX.

By the general theorem of Section III-2.3, the magnetic moment for a nucleus of definite spin I can be written as proportional to $\vec{I}$ as already anticipated in Eqs. (III-2) and (III-4) so

$$\vec{\mu}_I = (\mu_I / I) \vec{I} \qquad \text{(III-77)}$$

with the notation of Section III-1 . From this the magnetic dipole energy can be written as

$$\mathcal{H}_{Ml} = -(\mu_I / I) \vec{I} \cdot (\vec{H}_J + \vec{H}_0 \qquad \text{(III-78)}$$

If $\vec{H}_j$ is zero as in ^1S atomic states or $^1\Sigma$ states of non-rotating molecules, μ_I may readily be evaluated from a measurement of the energy level

separations, from a knowledge of the applied field $\underset{mo}{H}$, and from Eq.(III-78). However, in other cases a knowledge of $\underset{mj}{H}$ is required . Just as $\underset{m}{\mu}_I$ in Eq . (III-77) could be taken as proportional to $\underset{m}{I}$, so here $\underset{mj}{H}$ can be taken as proportional to $\underset{m}{J}$ for matrix elements diagonal in $\underset{m}{J}$. Therefore if the external field $\underset{mo}{H}$ is zero, Eq (III-78) becomes

$$\mathcal{H}_{Mt} = h \underset{mn}{a} \; \underset{m}{I} \cdot \underset{m}{J} \qquad \text{(III-79)}$$

where $\underset{m}{a}$ contains all the proportionality constants and is defined by

$$h \underset{mn}{a} = -\left(\frac{\mu_I}{I}\right) \frac{\underset{m}{H}_J}{\underset{m}{J}} = -\left(\frac{\mu_I}{I}\right) \frac{\underset{mJ}{H} \cdot \underset{m}{J}}{\underset{m}{J} \cdot \underset{m}{J}} \qquad \text{(III-80)}$$

From Eq. (III-55), the energy for the above may be written in an Fm representation as

$$W_{Mt} = \frac{h a}{2}\left[F(F+1) - I(I+1) - J(J+1)\right] = \frac{a}{2} C \qquad \text{(III-8I)}$$

In order that μ_I may be experimentally determined from the energy levels it is necessary according to Eqs. (III-80) and (III-8I) to have an estimate for $\underset{mj}{H}$, the affective magnetic field at the nucleus due to the rest of the atom . For hydrogen like atoms this effective field has been calculated by Fermi (FER30), Goudsmit (GOU29), and others (KOP40). A simplified version of these calculations is given in the author's Nuclear Moments (RAM53e, p.25) . The results are that for an S electron

$$h a = \frac{16\pi}{3} \mu_0 \frac{\mu_I}{I} \left|\psi_{no}(0)\right|^2 = \frac{8}{3} \frac{2\pi \hbar c \, R_y \, \alpha^2 Z^3 g_I}{n^3 (M/m)} \qquad \text{(III-82)}$$

and for an electron of orbital angular momentum $L \neq 0$

$$h a = \mu_0 (\mu_I / I) \langle 1/r^3 \rangle \, 2L(L+1) \Big/ J(J+1)$$
$$= \frac{2\pi \hbar c \, R_y \, \alpha^2 Z^3 g_I}{n^3 \left(L+\frac{1}{2}\right) J(J+1) (M/m)} \qquad \text{(III-83)}$$

where Ry is the Rydberg constant ($me^4/4\pi\hbar^3 c$ cm^{-1}), α is the fine structure constant $e^2/\hbar c$, M/m is the ratio of the proton to the electron mass, and $\psi_{no}(0)$ is the wave function of the S electron at the position of zero radius. Small corrections to the above have been introduced by Breit (BRE28, BRE30, BRE33, BRE48, BRE49) and Margenau (MAR40) to allow for the reduced mass effect of the nucleus and for the relativistic effects of the atomic electron .

The reduced mass effect introduces an extra factor of $(1 + m/M)^{-3}$ into Eq. (III-82). Low and Salpeter (LOW5I) have introduced a small additional reduced mass correction corresponding to recoil of the nucleus from the virtual photons that are interchanged . The correction to Eq. (III-82) from the anomalous magnetic moment of the electron is discussed in Chapters IX and XV .

An approximate formula for the calculation of the interaction constant in alkali metals has been derived by Goudsmit (GOU33) and is identical with Eq. (III-83) except that Z^3 is replaced by $Z_o^2 Z_i$ and n is replaced by n_o where Z_i is the effective charge when the valence electron is inside the core of closed electron shells, Z_o is the effective charge when it is outside the core, and n_o is the effective principal quantum number outside the core . Therefore, for alkali atoms

$$h a = \frac{2\pi \hbar c R_y \alpha^2 Z_o^2 Z_i g_I}{n_o^3 \left(L+\frac{1}{2}\right) J(J+1)(M/m)} \qquad \text{(III-84)}$$

If the fine structure separation δ of the same term is known, it can be used empirically to eliminate the Z_o and n_o dependence, since by atomic theory

$$\delta = \frac{2\pi \hbar c R_y \alpha^2 Z_o^2 Z_i^2}{n_i^3 L(L+1)} \qquad \text{(III-85)}$$

so

$$h a = \frac{L(L+1)\,\delta}{\left(L+\frac{1}{2}\right) J(J+1) Z_i (M/m)} \qquad \text{(III-86)}$$

Various corrections to Eq. (III-84) have been calculated by Fermi and Segré (FER33), Breit and Wills (BRE30, BRE33), Racah (RAC3I), and Goudsmit and Bacher (GOU29) and the results have been summarized by Kopfermann (KOP40). In addition to Eq. (III-84) applying for alkali atoms with one valence electron, it can also be used for halogens where just one electron is missing from the outer shell. It can further be applied approximately to atoms of the aluminum group, since the first two valence electrons form a more or less closed shell . More complicated relationships hold for two valence electrons such as alkaline earths. These have been summarized by Kopfermann (KOP40) .

All the states of a single hyperfine structure group have the same values of the quantum numbers I and J in Eq. (III-8I) while F takes on values from I + J to $|I - J|$. The energy difference between states with F equal to F and $(F - 1)$ is, according to Eq . (III-8I),

$$W_{F+1} - W_{F} = h a F \qquad \text{(III-87)}$$

This regularity is called the interval rule and for this reason a is often called the interval factor. However, even with magnetic dipole interactions, exceptions to this rule can occur if two atomic states of different J are separated by an amount which is not large compared to the hyperfine structure in which case J in Eq. (III-8I) is no longer a good quantum number and second order perturbations must be taken into account . Calculations of the effect of perturbations have been made by Fermi and Segrè (FER33), Casimir (CAS32), and Goudsmit and Bacher (GOU33a), and others (KOS52) .

If $J = \frac{1}{2}$, the separation of the hyperfine structure levels is, from Eq. (III-87),

$$\Delta W = W(F) - W(F-1) = \frac{1}{2} \Delta\nu = h a F = h a (I + \tfrac{1}{2}) = h a \frac{2I+1}{2} \qquad \text{(III-88)}$$

The quantity $\Delta\nu$ defined in this way is frequently referred to as the hyperfine structure separation, or the hfs $\Delta\nu$.

III-3.3 . Magnetic dipole interactions with internal and external magnetic fields . So far in the detailed discussion of Eq.(III-78) it has been assumed either that there was no intramolecular interaction or that the external field was zero . However, in many of the practical cases which arise in molecular beam measurements the magnetic interactions of the electrons and the nuclei with each other and of one or both with the external field are of comparable importance .

The effect of an external magnetic field on the hyperfine structure of an atomic energy level illustrates such a combined interaction . In this case, if the interaction energy with the external field of the electrons as well as of the nucleus is included and if μ_J is the resultant electronic magnetic moment defined analogously to the definitions in Section III-1, Eqs. (III-78) and (III-80) become

$$\mathcal{H}_{H} = h a \underset{m}{I} \cdot \underset{m}{J} - \frac{\mu_I}{J} \underset{m}{J} \cdot H_0 - \frac{\mu_I}{I} \underset{m}{I} \cdot H_0 \qquad \text{(III-88a)}$$

(a) Weak external field. If the magnetic field is very weak corresponding to a weak field Zeeman effect, the term $h a \underset{m}{I} \cdot \underset{m}{J}$ in the above

is the largest, so that $\underset{\sim}{I}$ and $\underset{\sim}{J}$ are tightly coupled together to form a resultant $\underset{\sim}{F}$ as in Fig. III-2 while $\underset{\sim}{F} = \underset{\sim}{I} + \underset{\sim}{J}$. As a consequence, only the components of the electron and nuclear magnetic moments along $\underset{\sim}{F}$ are effective since the perpendicular components average to zero . The above then becomes approximately

$$\mathcal{H}_{MI} = a \, \underset{\sim}{I} \cdot \underset{\sim}{J} - \frac{\mu_J}{J} \underset{\sim}{J} \cdot \frac{\underset{\sim}{F}}{|F|} \frac{\underset{\sim}{F}}{|F|} \cdot \underset{\sim}{H}_o - \frac{\mu_I}{I} \underset{\sim}{I} \cdot \frac{\underset{\sim}{F}}{|F|} \frac{\underset{\sim}{F}}{|F|} \cdot \underset{\sim}{H}_o \qquad \text{(III-89)}$$

But

$$\underset{\sim}{I} \cdot \underset{\sim}{I} = (\underset{\sim}{F}-\underset{\sim}{J}) \cdot (\underset{\sim}{F}-\underset{\sim}{J}) = |F|^2 + |J|^2 - 2\underset{\sim}{J} \cdot \underset{\sim}{F} \qquad \text{(III-90)}$$

so the energy for the state specified by quantum numbers F and m, where m is the magnetic quantum number for $\underset{\sim}{F}$, is

$$W_{MI}(F,m) = (Fm | \mathcal{H}_{MI} | Fm)$$

$$= a (Fm | \underset{\sim}{I} \cdot \underset{\sim}{J} | Fm)$$

$$- \frac{\mu_J}{I} \frac{F(F+1) + I(I+1) - J(J+1)}{2F(F+1)} (Fm | \underset{\sim}{F} \cdot \underset{\sim}{H}_o | Fm)$$

$$- \frac{\mu_I}{I} \frac{F(F+I) + J(J+I) - I(I+1)}{2F(F+1)} (Fm | \underset{\sim}{F} \cdot \underset{\sim}{H}_o | Fm)$$

$$= \frac{a}{2} [F(F+1) - I(I+1) - J(J+1)] \qquad \text{(III-9I)}$$

$$- \left[\frac{\mu_J}{2} \frac{F(F+1) + J(J+1) - (I+1)}{2F(F+1)} \right.$$

$$\left. + \frac{\mu_I}{I} \frac{F(F+1) + I(I+1) - J(J+1)}{2F(F+1)} \right] H_o m$$

In the low field limit the term proportional to μ_J can often be neglected due to its small size .

(b) Strong external field . In the limit of a very strong magnetic field, corresponding to a Paschen-Back effect, $\underset{\sim}{I}$ and $\underset{\sim}{J}$ will each separately precess rapidly about $\underset{\sim}{H}_o$ as in Fig. III-3, so that F is no longer a good quantum number, but instead the good quantum numbers are the magnetic quantum numbers m_I and m_J of $\underset{\sim}{I}$ and $\underset{\sim}{J}$ respectively . Then from Eq. (III-88a)

$$W_{MI}(m_I, m_J) = (m_I \, m_J | \mathcal{H}_{MI} | m_I \, m_J)$$

$$= a \left(m_I m_J \left| \underset{\sim}{(I)}_x \underset{\sim}{(J)}_x + \underset{\sim}{(I)}_y \underset{\sim}{(J)}_y + \underset{\sim}{(I)}_z \underset{\sim}{(J)}_z \right| m_I m_J \right)$$

$$- \frac{\mu_J}{J} H_o m_J - \frac{\mu_I}{I} H_o m_I$$

$$= - \frac{\mu_J}{J} H_o m_J - \frac{\mu_I}{I} H_o m_I + a m_I m_J \qquad \text{(III-92)}$$

since the diagonal elements of $\underset{\sim}{(I)}_x$ and $\underset{\sim}{(I)}_y$ are zero .

(c) Intermediate external field. The case of intermediate external field is of particular importance in molecular beam research, but it is more difficult to work out the energy levels since a secular equation must be solved as in the corresponding fine structure case (CON35) . The procedure for this solution is given below in Section III-4 . The result of the derivation there given in the case of $J = \frac{1}{2}$ but arbitrary I is

$$W_{MJ} \left(F, m \right) = - \frac{\Delta W}{2(2I + 1)} - \frac{\mu_I}{I} H_o m \pm \frac{\Delta W}{2} \sqrt{1 + \frac{4m}{2I+1} x + x^2} \qquad \text{(III-93)}$$

where

$$\Delta W = \frac{h a}{2} \left(2I + 1 \right) = 2\pi\hbar \Delta\nu$$

$$x = \frac{\left(-\mu_J / J + \mu_I / I \right) H_o}{\Delta W} \qquad \text{(III-94)}$$

and the + is used $F = I + \frac{1}{2}$ and the minus for $F = I - \frac{1}{2}$. This energy dependence for the case of $I = \frac{1}{2}$ and $J = \frac{3}{2}$ is plotted in Fig. III-4. The quantity $\Delta\nu$ of Eq. (III-94) is the hyperfine structure separation of Eq. (III-88) . The derivation of Eq. (III-93) is given in Section III-4 .

III-3.4. Magnetic interactions in molecules . Interaction energies in rotating $^1\Sigma$ diatomic molecules depend in general on three different angular momentum vectors $\underset{\sim}{I}$, of the first nucleus, $\underset{\sim}{I}_2$ of the second and $\underset{\sim}{J}$ of the molecular rotational angular momentum with corresponding magnetic moments $\underset{\sim}{\mu}_I$, $\underset{\sim}{\mu}_2$, and $\underset{\sim}{\mu}_J$. These vectors can then interact among themselves in pairs and with an external magnetic field with interactions of the type included in Eq. (III-88a) . The calculation of matrix elements for the interaction Hamiltonian and of interaction energies then proceeds for these in just the same fashion as discussed above in Section III-3.3 and subsequently

in Section III-4 .

However,in addition,Kellogg,Rabi, Ramsey,and Zacharias (KEL39a, KEL40) have pointed out that there is a mutual interaction between the three angular momentum vectors which involves all three of them together . In particular the mutual magnetic interaction $\mathcal{H}_S$ between the two nuclear magnetic moments is just the classical interaction of two dipoles so

$$\mathcal{H}_S = \frac{I}{R^3}\left\{ \underset{\sim}{\mu}_I \cdot \underset{\sim}{\mu}_2 - 3\,(\underset{\sim}{\mu}_1 \cdot \underset{\sim}{R})\,(\underset{\sim}{\mu}_2 \cdot \underset{\sim}{R})\,/R^2 \right\} \qquad \text{(III-95)}$$

where $\underset{\sim}{R}$ is a vector from nucleus 1 to nucleus 2 .

For the calculation of matrix elements of the above interaction $\underset{\sim}{\mu}_1$ and $\underset{\sim}{\mu}_2$ can be written as proportional to $\underset{\sim}{I}_1$ and $\underset{\sim}{I}_2$ in accordance with Eq. (III-2) and (III-4) . However,the dependence on the angular momentum operator $\underset{\sim}{J}$ is not immediately apparent . For this reason it is best to proceed in analogy with the method used in Section III-2.3 to express the quadrupole interaction in terms of angulat momentum operators as in Eq. (III-43) .

Eq. (III-95) can be rewritten as

$$\mathcal{H}_S = -\frac{(\mu_I/I_1)(\mu_2/I_2)}{R^3} \sum_{ij} [I_{1i} I_{2j}]\,[3\,R_i R_j/R^2 - \delta_{ij}] \qquad \text{(III-96)}$$

where the subscripts i and j indicate the different Cartesian components . By the general theorem of Section III-2.3 in the calculation of matrix elements diagonal in J one may write

$$[3\,R_i R_j/R^2 - \delta_{ij}] = C\left[3\,\frac{J_i J_j + J_j J_i}{2} - \delta_{ij}\,J^2\right] \qquad \text{(III-97)}$$

C can then be evaluated by multiplying the preceding equation by J_i on the left and J_j on the right and summing over i and j . This then gives

$$3(\underset{\sim}{J} \cdot \underset{\sim}{R})^2\,/R^2 - \underset{\sim}{J}^2 = C\left[\frac{3}{2}\,(\underset{\sim}{J}^2)^2 + \frac{3}{2} \sum_{ij} J_i J_j I_i J_j - (\underset{\sim}{J}^2)^2\right] \qquad \text{(III-98)}$$

On the left side one may use that $\underset{\sim}{J} \cdot \underset{\sim}{R}$ vanishes in diatomic molecules since the angular momentum is always perpendicular to the internuclear axis . The right side may be transformed by exactly the same procedures that in Appendix C transformed the right side of Eq. (C-I2) to the right side of Eq.(C-I9) so

$$-\underset{\sim}{J}^2 = \tfrac{1}{2} C \underset{\sim}{J}^2 \left[(2J-1)(2J+3)\right] \qquad \text{(III-99)}$$

and

$$C = -\frac{2}{(2J-1)(2J+3)} \qquad \text{(III-100)}$$

An alternative evaluation of C can be obtained by taking the $J, m_J = J$ diagonal matrix element of the zz component of both sides of Eq. (III-97) . The matrix element of the right side is immediately obvious and that of the left side can be evaluated by integration and from the fact that the only Θ dependence of Ψ_{JJ} is as $\sin^J\Theta$ (CON35,p.5). This alternative derivation of course also leads to Eq. (III-101).

From Eqs (III-97) and (III-100)

$$\left[3R_iR_j/R^2 - \delta_{ij}\right] = \frac{2}{(2J-1)(2J+3)}\, 3\left[\frac{J_iJ_j+J_jJ_i}{2} - \delta_{ij}J^2\right] \qquad \text{(III-101)}$$

When this is combined with Eq. (III-96),there results

$$\mathcal{H}_S = \frac{(\mu_I/I_1)(\mu_2/I_2)}{R^3}\, \frac{2}{(2J-1)(2J+3)} \left[3(\underset{\sim}{I}_1\cdot\underset{\sim}{J})(\underset{\sim}{I}_2\cdot\underset{\sim}{J}) - (\underset{\sim}{I}_1\cdot\underset{\sim}{I}_2)\,\underset{\sim}{J}^2\right] \qquad \text{(III-102)}$$

This expression is applicable both to homonuclear and heteronuclear diatomic molecules . However,it is the most convenient expression to use when the two nuclei are not identical as in the molecule HD. The matrix elements diagonal in I and J are then directly obtained from Eq. (III-102) and the known ways in which the angular momentum components operate on the wave functions (CON35) as in Appendix C.

If on the other hand,the two nuclei are identical,as in D_2, it is most convenient to use the resultant angulat momentum $\underset{\sim}{I}_R = \underset{\sim}{I}_1 + \underset{\sim}{I}_2$ since I_R is a good quantum number in a rotational state of definite J because of the symmetry requirements on the wave function . For this purpose,the combination of Eqs. (III-96) and (III-101) can be written in the equivalent form

$$\mathcal{H}_S = \frac{(\mu_I/I_1)(\mu_2/I_2)}{R^3}\, \frac{2}{3(2J-1)(2J+3)} \sum_{ij} 3I_{1i}I_{2J} - \delta_{ij}\underset{\sim}{I}_1\cdot\underset{\sim}{I}_2 \times \left[3\,\frac{J_iJ_j + J_jJ_i}{2} - \delta_{ij}J\right] \qquad \text{(III-103)}$$

Then if I_R is a good quantum number,by the general theorem of Section III-2,3 we can write

$$\left[3I_{1i}I_{2j} - \delta_{ij} I_1 \cdot I_2\right] = D\left[3\,\frac{I_{Ri}I_{Rj}+I_{Rj}I_{Ri}}{2} - \delta_{ij} I_R^2\right] \qquad \text{(III-I04)}$$

D can then be evaluated by multiplying both sides of the above by I_i on the left and I_j on the right and summing over i and j . Then, since the operations on the right side of Eq. (III-I04) are identical with those on the right side of Eq. (C-II) in Appendix C, it can be transformed in the same way to the form of the right side of Eq. (C - I9). Eq. (III-I04) with these modifications then becomes

$$\left[3(I_R \cdot I_1)(I_2 \cdot I_R) - (I_1 \cdot I_2)\, I_R^2\right] = \frac{D}{2} I_R^2 (2I_R-1)(2I_R+3) \qquad \text{(III-I05)}$$

But in Eq. (C-I7) of Appendix C it is shown that

$$I_R \cdot I_1 = I_R \cdot I_2 = \tfrac{1}{2} I_R^2 \qquad \text{(III-I06)}$$

and

$$I_1 \cdot I_2 = \tfrac{1}{2}\,(I_1+I_2)^2 - I_1^2 - I_2^2 = \tfrac{1}{2} I_R^2 - 2I_1^2 = \tfrac{1}{2} I_R^2 - I_1^2 \qquad \text{(III-I07)}$$

With these values, Eq. (III-I05) immediately yields

$$D = \tfrac{1}{2}\,\frac{I_R(I_R+1) + I_1(I_1+1)}{(2I_R-1)(2I_R+3)} \qquad \text{(III-I08)}$$

Therefore Eq. (III-I03) may be written as

$$S = \frac{(\mu_1/I_1)(\mu_2/I_2)}{R^3} \cdot \frac{I_R(I_R+1) + I_1(I_1+1)}{3(2I_R-1)(2I_R+3)(2J-1)(2J+3)} \times \sum_{ij}\left[3\,\frac{I_{Ri}I_{Rj}+I_{Rj}I_{Ri}}{2} - \delta_{ij} I_R^2\right]\left[3\,\frac{J_iJ_j+J_jJ_i}{2} - \delta_{ij} J^2\right] \qquad \text{(III-I09)}$$

This is sometimes the most convenient form to use . However, it is often to convenient to express it in an alternative form analogous to Eq.(III-43). For this, the derivation in Appendix C transforming Eq. (C-1) to Eq.(C-9) is directly applicable whence

$$\mathcal{H}_S = \frac{(\mu_1/I_1)(\mu_2/I_2)}{R^3}\,\frac{I_R(I_R+1)+4I_1(I_1+1)}{(2I_R-1)(2I_R+3)(2J-1)(2J+3)}\left[3(I_R \cdot J)^2 + \frac{3}{2} I_R \cdot J - I_R^2 J^2\right] \qquad \text{(III-II0)}$$

Note that the dependence of this on the operators $\underset{\sim}{I}_R$ and $\underset{\sim}{J}$ is exactly the same as that of Eq. (III-44b) for the quadrupole interaction of a homonuclear diatomic molecule with two identical nuclei . Consequently by measurements only in a single homonuclear molecule which possesses both a nuclear quadrupole moment and a spin-spin magnetic interaction it is impossible to distinguish between the spin-spin interaction and the nuclear quadrupole interaction . On the other hand, the spin-spin interaction can often be inferred from a similar measurement in a different molecule whence the quadrupole interaction can be inferred, as in a comparison of H_2 and D_2 . This difficulty does not arise with a molecule such as HD whose nuclei are not identical so the spin-spin and nuclear quadrupole interactions can separately be inferred from measurements on a molecule such as HD.

In addition to the magnetic interactions discussed above, Ramsey (RAM40, RAM52b) has pointed out that in precision experiments with molecules account must be taken of the molecular diamagnetic interaction with the external magnetic field and of the dependence of this interaction on molecular orientation . He has also pointed out (RAM50d, RAM51c) that a small correction must be made for the magnetic shielding of the nuclei by the electrons of the molecule and for the variation of this shielding with molecular orientation . The detailed discussion of these magnetic interactions and of their theoretical form is given in Chapter VIII.

III-4. Intermediate Coupling .

III-4.1. Introduction . In general the nuclear interactions of an atom or molecule in a magnetic field are combinations of several of the interactions described in Sections III-2 and III-3 . With several such interactions present simultaneously it is not in general possible to select a priori a representation for which the energy matrix is diagonal . Often it is possible to select a representation in which the largest matrix elements are diagonal, in which case a standard form of perturbation calculation can be used to make corrections for the effect of the small non-diagonal matrix elements (CON35) . The procedure for such a calculation is given in Section III-4.3. In many cases, however, the non-diagonal matrix elements are comparable to the diagonal ones in which case the perturbation procedure is unsatisfactory . In this case the best procedure is to calculate all the matrix elements from the

interaction Hamiltonian in any convenient representation . The energy levels of the system may then be obtained by a solution of the corresponding secular equation .

III-4.2. Solution of secular equation . An example of the solution of the secular equation for an intermediate coupling problem is the solution of the intermediate field problem of Section III-3.3. In this case and if the z axis is taken in the direction of the external field H_o, the Hamiltonian of the system as given in Eq. (III-88a) is

$$\mathcal{H} = - dI_z - cJ_z + h a \underset{\sim}{I} \cdot \underset{\sim}{J}$$

$$= - dI_z - cJ_z + haI_zJ_z + \tfrac{1}{2}I_+J_- + \tfrac{1}{2}h a I_-J_+ \qquad \text{(III-III)}$$

where

$$d = (\mu_I/J)H_o \qquad \text{and} \qquad c = (\mu_J/J)H_o \qquad \text{(III-II2)}$$

As discussed in Section III-3.3, an Fm representation is most convenient when H_o is large and an m_Im_J representation is best when H_o is small . However, in intermediate magnetic fields the matrix is not even approximately diagonal in either representation so that the secular equation must be solved . As a result it makes little difference as to which representation is chosen in the calculation of the complete matrix and the seculat equation will, of course, be independent of this choice even though the matrices are quite different in form . For the present calculation an m_Im_J representation will arbitrarily be chosen, so the typical matrix element will be $(m_Im_J|\mathcal{H}|m_I'm_J')$; a calculation with an Fm representation has been published elsewhere(MIL38) . The calculation will be limited to atomic states for which $J = \frac{1}{2}$ since this greatly simplifies the arithmetic while retaining an example which illustrates the important principles and is in fact the specific example needed in the majority of molecular beam experiments (BRE3I) .

From the commutation rules of Eq. (III-44), it is easy to show that $F_z = I_z + J_z$ commutes with the Hamiltonian of Eq. (III-III) . Therefore, all matrix elements of $\mathcal{H}$ not diagonal in $m = m_I + m_J$ will necessarily vanish . Therefore, the total energy matrix will consist of a number of submatrices, each corresponding to a different value of m and with no non-diagonal matrix elements connecting two submatrices of different m . Since by assumption $J = \frac{1}{2}$ these submatrices at most will be 2 x 2 corresponding to the two

possible states giving the value of m. In an $m_I m_J$ representation these states are ($m_I = m - \frac{1}{2}$, $m_J = \frac{1}{2}$) and ($m_I = m + \frac{1}{2}$, $m_J = -\frac{1}{2}$) .

Let us first consider the submatrix with $m = I + \frac{1}{2}$. In this case, the submatrix is only 1 x 1, since only one state ($m_I = I$, $m_j = \frac{1}{2}$) can give the desired m whence there are no non-diagonal matrix elements . The same thing is true of the $m = -I - \frac{1}{2}$ submatrix . Since m_I is completely determined by m and m_J let the energy of the level be designated by $W_{m\,,\,m_J}$. Then, from Eq . (III-35a)

$$W_{\pm\, I+\frac{1}{2},\pm\frac{1}{2}} = (\pm I,\ \pm\tfrac{1}{2}\left|\mathcal{H}\right|\pm I,\ \pm\tfrac{1}{2}) = \mp dL\ \mp\tfrac{1}{2}\ c + \tfrac{1}{2}\hbar\, a\, I \qquad \text{(III-II3)}$$

However, the other submatrices are 2 x 2 . The secular equation is then of the form

$$\begin{vmatrix} A - E_{m,\pm} & C \\ D & B - E_{m,\pm} \end{vmatrix} = 0 \qquad \text{(III-II4)}$$

where from Eq. (III-35a)

$$\begin{Bmatrix} A \\ B \end{Bmatrix} = (m\pm\tfrac{1}{2},\ \mp\tfrac{1}{2}\left|\mathcal{H}\right|m\pm\tfrac{1}{2},\ \mp\tfrac{1}{2}) = -d\ (m\pm\tfrac{1}{2})\pm\tfrac{1}{2}\ c\ \mp\tfrac{1}{2}\hbar\, a\ \ (m\pm\tfrac{1}{2}) \qquad \text{(III-II5)}$$

and

$$C = D = (m\pm\tfrac{1}{2},\ \mp\tfrac{1}{2}\left|\mathcal{H}\right|m\mp,\ \pm\tfrac{1}{2}) = \hbar\frac{a}{2}\,(m\ \pm\tfrac{1}{2}\left|\ I_{\pm}\right|m\mp\tfrac{1}{2})(\mp\tfrac{1}{2}/\ J_{\mp}\ \pm\ \tfrac{1}{2})$$

$$\text{(III-II6)}$$

$$= \hbar\frac{a}{2}\sqrt{(I-m+\tfrac{1}{2})(I+m+\tfrac{1}{2})}\ \sqrt{(\tfrac{1}{2}+\tfrac{1}{2})(\tfrac{1}{2}+\tfrac{1}{2})} = \hbar\frac{a}{2}\sqrt{(I+\tfrac{1}{2})^2 - m^2}$$

With the above and by straightforward expansions Eq. (III-II5) can be written as

$$a'\,W_{m,\pm}^{\ 2} + b'W_{m,\pm} + c' = 0 \qquad \text{(III-II7)}$$

where

$$a' = 1$$

$$b' = 2dm + \hbar\frac{a}{2}$$

$$c' = \left[-d(m+\tfrac{1}{2})+\tfrac{1}{2}c - \tfrac{1}{2}a\hbar(m+\tfrac{1}{2})\right]\left[-d(m-\tfrac{1}{2})-\tfrac{1}{2}c + \tfrac{1}{2}a\hbar(m-\tfrac{1}{2})\right]$$

$$- \frac{a^2\hbar^2}{4}\left[(I+\tfrac{1}{2})^2 - m^2\right] \qquad \text{(III-II8)}$$

Therefore, with some straightforward simplification

$$W_{m,\pm} = \frac{-b' \pm \sqrt{b'^2 - 4a'c'}}{2a'} \qquad \text{(III-II9)}$$

$$= -\tfrac{1}{4}a\hbar - dm \pm \sqrt{\hbar^2 a^2 (I+\tfrac{1}{2})^2 + 2\hbar(d-c)m + (d-c)^2}$$

When H = 0 so d and c are also zero the above equation gives W_{m+} as equal to the energy of Eq. (III-9I) when $F = I + \frac{1}{2}$ and W_{m-} is the energy for $F = I - \frac{1}{2}$. Hence we can use the designation W (F,m) provided it is understood that the quantum numbers in this case are those which apply at the H = 0 limit and that the upper signs are used for $F = I + \frac{1}{2}$ while the lower ones are used for $F = I - \frac{1}{2}$. The constants a,d, and c can be reexpressed from Eqs .(III-88) and (III-II2) . With those changes and some elementary rearrangements Eq. (III-II9) becomes

$$W(F,m) = -\frac{\Delta W}{2(2i+1)} - \frac{\mu_I}{I} H_o m \pm \frac{\Delta W}{2}\sqrt{1 + \frac{4m}{2I+1}x + x^2} \qquad \text{(III-I20)}$$

where

$$x = \frac{(-\mu_J/J + \mu_I/I)\ H_o}{\Delta W} \qquad \text{(III-I2I)}$$

$$= (g_j - g_I)\,\mu_o H_o / W \qquad \text{(III-I2Ia)}$$

with

$$g_j = -\mu_j/J_o$$

where μ_o is the Bohr magneton . From the nature of the above derivation, the preceding equation would seem to apply for all values of m except $\pm(I+\frac{1}{2})$ for which Eq. (III-II3) is the appropriate expression . However, if the value of Eq. (III-I20) is calculated fot the states of $F = I + \frac{1}{2}$ [whence the plus sign is used in Eq. (III_I20)] and $m = \pm(I+\frac{1}{2})$ the square root becomes the square root of a perfect square and the resulting

linear equation is just the same as Eq. (III-II3) . Hence Eq. (III-iéo) may be taken to apply to all states. This is just the relation assumed earlier without proof in Eq. (III-94). The energy variation is shown graphically in Fig. III-4 .

With more complicated Hamiltonians including such features as quadrupole moment interactions, the procedure is similar except that additional matrix elements must be computed and if the atomic or molecular J is greater than $\frac{1}{2}$ the secular equation to be solved in general is of higher degree than quadratic . For example, from Eqs. (III-44c) and (III-88a), for an atom with both magnetic and quadrupole interactions the Hamiltonian may be written as

$$\mathcal{H} = -\left(\frac{\mu_J}{J}\underset{\sim\sim}{J} + \frac{\mu_I}{I}\underset{\sim\sim}{I}\right)\cdot\underset{\sim\sim}{H}_0 + h\,a\,\underset{\sim\sim}{I}\cdot\underset{\sim\sim}{J} + b\,2\underset{\sim\sim}{I}\cdot\underset{\sim\sim}{J}\left(2\underset{\sim\sim}{I}\cdot\underset{\sim\sim}{J} + 1\right) \quad \text{(III-I22)}$$

Fig . III-5 illustrates the energy levels of an atom, Ga^{69} , whose nuclear interactions are those of the above Hamiltonian and which is in a $^2P_{3/2}$ atomic state (BEC48). Other examples of such calculations will appear in later chapters, notably Chapter VIII. In Appendix C, Eqs.(C-24),(C- 26), and (C627), useful formulae are given for the diagonal and non-diagonal matrix elements needed in such calculations .

III-4.3. Perturbation theory . Often the energy levels are desired only in regions in which some terms of the Hamiltonian are much larger than others. For example, if H_o in E_q. (III-I22) is large, the first term may be much larger than the last two . Conversely for small H_o the last two terms are large compared to the first. In such cases the desired energy levels can often be more simply obtained by the approximate methods of perturbation theory than from a complete solution of the secular equation . This is especially true if the nature of the Hamiltonian is such that the secular equation is of higher degree than the second, unlike the example of the preceding section .

In this case, one can write

$$\mathcal{H} = \mathcal{H}_o + \mathcal{H}' \quad \text{(III-I23)}$$

where $\mathcal{H}' \ll \mathcal{H}_o$. In the case of large H_o, for example, $\mathcal{H}_o$ would correspond to the first term of Eq. (III-I22) and $\mathcal{H}'$ to the last two . For simplicity

it will be assumed that there is no degeneracy to the problem with Hamiltonian $\mathcal{H}_o$ (if this assumption is not valid the more complicated procedure given for example by Condon and Shortly (CON35,p. 34) may be followed) . Then as shown in any standard quantum mechanics text (CON35), the energy of the n'th level is given by

$$W_n = W_n^{\,o} + W_n^{\,1} + W_n^{\,2} + \cdots \qquad \text{(III-I24)}$$

where

$$W_n^{\,o} = \int \psi_n^{*o} \mathcal{H}_o \psi_n^{o} d\tau = (n|\mathcal{H}|_o\, n)$$

$$W_n^{\,1} = (n|\mathcal{H}'|n) \qquad \text{(III-I25)}$$

$$W_n^{\,2} = \sum_{n'}{}' \frac{(n|\mathcal{H}'|n')(n'|\mathcal{H}'|\,n)}{W_n^{\,o} - W_{n'}^{\,o}}$$

where the $\psi_{n'}^{\,o}$'s involved in the matrix elements are the n' eigenfunctions of the unperturbed $\mathcal{H}_o$ and where the prime on the last summation symbol indicates n' is not to be taken equal to n . Expressions for the third and higher order perturbations are given by Condon and Shortly (CON35,p.34) . In molecular beam problems it is sometimes convenient to use as high as third order perturbation theory (RAM52a) but rarely higher . An easy example of the perturbation procedure can be obtained by applying it to the problem of the preceding section,in which case the desired matrix elements can be easily obtained from the appropriate parts of Eqs.(III-II5) and (III-II6), with the d and c terms being taken as contributing to $\mathcal{H}_o$ while the a term is part of $\mathcal{H}'$.

IX.-9 . Magnetic Octupole Interactions

Jaccarino,King,Satten,Stroke,and Zacharias (JAC54) have recently detected a nuclear magnetic octupole interaction with iodine in an atomic beam magnetic resonance experiment . An octupole moment interaction in the same atom was earlier reported by Tolansky (TOL39) in optical hyperfine structure experiments . However,Casimir and Karreman (CAS42) showed that an octupole interaction anywhere near as large as that reported by Tolansky was very unlikely ; the experiment of Jaccarino and Zacharias indeed confirmed Casimir and Karreman's theoretical analysis since their observed octupole interaction was much smaller than that reported by Tolansky . Nevertheless, they did find a small octupole moment interaction .

The theory of nuclear magnetic octupole interactions is included implicitly in Chapter III above as can be seen by the combination of Eqs. (III-35),(III-57), and (III-61). A detailed and explicit discussion of the theory of nuclear magnetic octupole interactions has been given by Casimir and Karreman (CAS42) and by Jaccarino,*et al* (JAC54). They show that the octupole interaction can be written in an F,m_F representation as

$$W_{E3} = hc\,\frac{5}{4}\,\frac{C^3+4C^2+4/5C\cdot\left[4I(I+1)J(J+1)+I(I+1)+J(J+1)+3\right]-4I(I+1)J(J+1)}{I(I-1)(2I-1)J(J-1)(2J-1)} \qquad \text{(IX-39)}$$

where C is defined by Eq. (III-55) and where c is the nuclear magnetic octupole constant . The quantity c can be expressed as

$$hc = \left\langle r_n^{\,3} P_{3n}\, \nabla \cdot \underset{\sim}{m}_n \right\rangle_{II} \left\langle r_e^{\,-4} P_{3e}\, \nabla \cdot \underset{\sim}{m}_e \right\rangle_{JJ} \qquad \text{(IX-40)}$$

where P_{3n} is the Legendre polynomial of degree 3 in the angular coordinates of the nuclear charge density . The quantity Ω defined by

$$\Omega = \left\langle r^3 P_3\, \nabla \cdot \underset{\sim}{m}_n \right\rangle_{II} = \left\langle \frac{5z_n^{\,3} - 3z_n r_n^{\,2}}{2}\, \nabla \cdot \underset{\sim}{m}_n \right\rangle_{II} \qquad \text{(IX-41)}$$

is called the octupole moment .

Jaccarino,Zacharias,and their associates (JAC54) have therefore analyzed their experimental results on atomic I^{127} with a Hamiltonian of the form of Eq. (IX-36) plus Eq. (IX-39) . In this way they obtained

values for a,b, and c . However,as discussed in detail by Schwartz (SCH54), second order perturbations from nuclear dipole and quadrupole interactions also contribute to c. Schwartz has calculated these second order perturbations both from the $^2P1/2$ state and from configuration interaction . When these corrections are made in c so that the remaining value of c is presumably due exclusively to a nuclear octupole interaction,the results (SCH54) are

$$a = 827.265 \pm 0.003 \text{ Mc/sec.}$$

$$b = 1146.356 \pm 0.010 \text{ Mc/sec.} \qquad \text{(IX-42)}$$

$$c = 0.00287 \pm 0.00037\text{-}0.00053 \text{ Mc/sec} = 0.00234 \pm 0.0037 \text{ Mc/sec}$$

The first term of the first expression for c is the value uncorrected for perturbations and the last term in the first expression is the theoretical correction . These values are also listed in Table XI-III . The value of the octupole moment corresponding to the above value of c is (SCH54) .

$$Q_{127} = (0.17 \pm 0.03) \times 10^{-24} \text{ N cm}^2 \qquad \text{(IX -43)}$$

Kusch and Exk (KUS54) have found evidence for a nuclear magnetic octupole interaction with In^{115} while Daly and Holloway (DAL54) have done likewise with Ga^{69} and Ga^{71} . Their corrected values for a,b,and c are tabulated on Table XI-III . The corresponding value of the nuclear magnetic octupole moments are

$$Q_{115} = (0.31 \pm 0.01) \times 10^{-24} \text{ N cm}^2$$

$$Q_{69} = (0.107 \pm 0.004) \times 10^{-24} \text{ N cm}^2$$

$$Q_{71} = (0.146 \pm 0.004) \times 10^{-24} \text{ N cm}^2$$

The magnitudes of the experimental magnetic octupole interactions in Eg. (IX-42) and Table XI-III are qualitatively very reasonable as can be seen from the following discussion . The order of magnitude of the magnetic dipole interaction is $W_{M1} \sim \mu_N \langle 1/r^3 \rangle$. On the other hand the order of magnitudes of the magnetic octupole interaction is $W_{M3} \sim \mu_N R^2 \langle 1/r^5 \rangle$

where R is the nuclear radius . Hence, one would expect

$$W_{M3} / W_{M1} \qquad R^2 \langle 1/r^5 \rangle / \langle 1/r^3 \rangle \qquad R^2/a^2$$

where a is the effective radius for the electron which provides the interaction. At first sight one might take for a the Bohr radius Ao but on further consideration it is apparent that the octupole interaction occurs chiefly when the electron is close to the nucleus in which region its orbit is only slightly affected by the shielding of the other electrons ; in other words a should be taken to have its approximately unshielded value $a = A_o/Z$ s•

$$W_{M3}/W_{M1} \qquad Z^2 R^2 / A_o^2 \qquad 10^{-5}$$

for large Z .

Schwartz (SCH54) has generalized the calculations of the Schmidt lines derived in Section VI-10.6 and has obtained corresponding expressions for any nuclear multipole moment . In particular he has obtained theoretical values for nuclear octupole moments as functions of nuclear spin I on the assumption that the octupole moment is due to a single odd nucleon . Just as in the case of the Schmidt linetwo predictions are made for each value of I depending on whether $I = L + 1/2$ or $I = L - 1/2$. Schwartz (SCH54) finds that the experimental nuclear octupole moments fall between these two theoretical curves in much the same way that the nuclear magnetic dipole moments fall between the Schmidt lines .

APPENDIX C

Quadrupole Interactions

From Equations (III-32) and (III-42),

$$\mathcal{H}_{E2} = \frac{1}{6} \frac{e^2 q_J Q}{I(2I-1)J(2J-1)} \sum_{ij} \left[3 \frac{I_i I_j + I_j I_i}{2} - \delta_{ij} \underset{\sim}{I}^2 \right] \times \left[3 \frac{J_i J_j + J_j J_i}{2} - \delta_{ij} \underset{\sim}{J}^2 \right] \quad \text{(C-1)}$$

Now consider typical terms in this equation. Thus, as I and J commute with each other

$$\sum_{ij} I_i I_j J_i J_j = \left[\sum_i I_i J_i \right] \left[\sum_j I_j J_j \right] = \left(\underset{\sim}{I} \cdot \underset{\sim}{J} \right)^2 \quad \text{(C-2)}$$

Likewise,

$$\sum_{ij} I_i I_j \delta_{ij} \underset{\sim}{J}^2 = \left[\sum_i I_i^2 \right] \underset{\sim}{J}^2 = \underset{\sim}{I}^2 \underset{\sim}{J}^2 \quad \text{(C-3)}$$

and

$$\sum_{ij} \delta_{ij} \underset{\sim}{I}^2 \underset{\sim}{J}^2 = 3 \underset{\sim}{I}^2 \underset{\sim}{J}^2 \quad \text{(C-4)}$$

The only complicated terms are

$$\sum_{ij} I_j I_i J_i J_j = \sum_{ij} I_i I_j J_j J_i \quad \text{(C-5)}$$

In this, from the commutation rules of Equation (III-44)

$$\sum_{ij} I_i I_j J_j J_i = \sum_{ij} I_j I_i J_j J_i + i \sum_{ij} I_{i\times j} J_j J_i$$

$$= \left(\underset{\sim}{I} \cdot \underset{\sim}{J}\right)^2 + \sum_{ij} I_{i\times j} J_j J_i$$

But

$$J_j J_i = \frac{J_j J_i + J_j J_i}{2} = \frac{J_i J_i + J_i J_j + i J_{j\times i}}{2}$$

so

$$i \sum_{ij} I_{i\times j} J_j J_i = \frac{i}{2} \sum_{ij} I_{i\times j} \left[J_j J_i + J_i J_j \right] - \frac{1}{2} \sum_{ij} I_{i\times j} J_{j\times i}$$

However, the first terms on the right of the above equation are antisymmetric in i and j so their sum over all i and j vanish. Therefore,

$$i \sum_{ij} I_{i\times j} J_j J_i = -\frac{1}{2} \sum_{ij} I_{i\times j} J_{j\times i} = \frac{1}{2} \sum_{ij} I_{i\times j} J_{i\times j}$$

$$= \frac{1}{2} \sum_{k} 2\, I_k J_k = \underset{\sim}{I} \cdot \underset{\sim}{J} \qquad \text{(C-6)}$$

Therefore,

$$\sum I_i I_j J_j J_i = \left(\underset{\sim}{I} \cdot \underset{\sim}{J}\right)^2 + \underset{\sim}{I} \cdot \underset{\sim}{J} \qquad \text{(C-7)}$$

and

$$\mathcal{H}_{E2} = \frac{1}{6}\frac{e^2 g_J Q}{I(2I-1)J(2J-1)}\left\{\frac{9}{2}(\underset{\sim}{I}\cdot\underset{\sim}{J})^2 + \frac{9}{2}\left[(\underset{\sim}{I}\cdot\underset{\sim}{J})^2 + \underset{\sim}{I}\cdot\underset{\sim}{J}\right]\right.$$
$$\left. - 3\underset{\sim}{I}^2\underset{\sim}{J}^2 - 3\underset{\sim}{I}^2\underset{\sim}{J}^2 + 3\underset{\sim}{I}^2\underset{\sim}{J}^2\right\} \qquad \text{(C-8)}$$

$$= \frac{e^2 g_J Q}{2I(2I-1)J(2J-1)}\left[3(\underset{\sim}{I}\cdot\underset{\sim}{J}) + \frac{3}{2}\underset{\sim}{I}\cdot\underset{\sim}{J} - \underset{\sim}{I}^2\underset{\sim}{J}^2\right] \qquad \text{(C-9)}$$

In diatomic molecules either with identical or non-identical nuclei the quadrupole interaction can be obtained by applying either (C-1) or (C-9) to each nucleus separately. However, in the case of identical nuclei it is usually most convenient to express the result in terms of the resultant angular momentum $I = I_1 + I_2$ since I is a good quantum number in a rotational state of definite J because of the symmetry requirements on the wave function with identical nuclei. This re-expression is obtained in the following paragraphs.

From Equation (III-42) applied to the two identical nuclei with $I_1 = I_2$.

$$Q_{ij} = \frac{eQ}{I_1(2I_1-1)}\left[3\,\frac{I_{1i}I_{1j} + I_{1j}I_{1i}}{2} - \delta_{ij}\underset{\sim}{I}_1^2\right.$$
$$\left. + 3\,\frac{I_{2i}I_{2j} + I_{2j}I_{2i}}{2} - \delta_{ij}\underset{\sim}{I}_2^2\right]$$

(C-10)

By the general theorem of Section III-2.3, for calculation of matrix elements diagonal in I the symmetric and traceless tensor in the brackets of the above equation can be written as proportioned to a similar tensor based on I. Therefore

$$\left[3\frac{I_{1i}I_{1j}+I_{1j}I_{1i}}{2}-\delta_{ij}I_{m1}^2+3\frac{I_{2i}I_{2j}+I_{2j}I_{2i}}{2}-\delta_{ij}I_{m2}^2\right] \qquad (C\text{-}11)$$

$$=E\left[3\frac{I_{Ri}I_{Rj}+I_{Rj}I_{Ri}}{2}-\delta_{ij}I_{mR}^2\right]$$

Then E can be evaluated by multiplying both sides on the left by I_{Ri} and on the right by I_{Rj} and summing over i and j. This gives

$$\left[\frac{3}{2}\left\{(I_{mR}\cdot I_{m1})^2+\sum_{ij}I_{Ri}I_{1j}I_{1i}I_{Rj}+(I_{mR}\cdot I_{m2})^2+\sum_{ij}I_{Ri}I_{2j}I_{1i}I_{Rj}\right\}-2I_{mR}^2I_{m1}^2\right]$$

$$=E\left[\frac{3}{2}(I_{mR}^2)^2+\frac{3}{2}\sum_{ij}I_{Ri}I_{Rj}I_{Ri}I_{Rj}-(I_{mR}^2)^2\right] \qquad (C\text{-}12)$$

The typical summation term in the above can be evaluated as follows from the commutation rules of Equation (III-44):

$$\sum_{ij}I_{Ri}I_{1j}I_{1i}I_{Rj}=\sum_{ij}I_{Ri}I_{1i}I_{1j}I_{Rj}+i\sum_{ij}I_{Ri}I_{1jxi}I_{Rj}$$

$$=(I_{mR}\cdot I_{m1})^2+i\sum_{ij}I_{1jxi}I_{Ri}I_{Rj}-\sum_{ij}I_{1ix(jxi)}I_{Rj} \qquad (C\text{-}13)$$

However

$$i \times (j \times i) = (i \cdot i) j - (i \cdot j) i = j \qquad \text{(C-14)}$$

and in the summation over i this can be achieved twice, once from each non j Cartesian component, whence an extra factor of two appears in the last term of the following equation. The next to last term in Equation (C-14) is evaluated by Equation (C-6). Therefore

$$\sum_{ij} I_{Ri} I_{1j} I_{1i} I_{Rj} = \left(\underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1}\right)^2 + \underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1} - 2\, \underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1}$$

$$= \left(\underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1}\right)^2 - \underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1} \qquad \text{(C-15)}$$

Equation (C-15) and its obvious variants can be used in Equation (C-12) with the result that

$$\left[3\left(\underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1}\right)^2 + 3\left(\underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m2}\right)^2 - \frac{3}{2} \underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1} - \frac{3}{2} \underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m2} - 2 I_{mR}^2 I_{m1}^2\right]$$

$$= E\left[3\left(I_{mR}^2\right)^2 - \frac{3}{2} I_{mR}^2 - \left(I_{mR}^2\right)^2\right] \qquad \text{(C-16)}$$

However, $\underset{\sim}{I}_{mR} \cdot \underset{\sim}{I}_{m1} = \frac{1}{2}\left[\left(\underset{\sim}{I}_{mR} - \underset{\sim}{I}_{m1}\right)^2 - I_{mR}^2 - I_{m1}^2\right]$

$$= \frac{1}{2}\left[I_{m2}^2 - I_{mR}^2 - I_{m1}^2\right] = \frac{1}{2} I_{mR}^2 \qquad \text{(C-17)}$$

since $I_{m1}^2 = I_{m2}^2$. Therefore (C-16) becomes

$$\left[\frac{3}{2}\left(I_{mR}^2\right)^2 - \frac{3}{2} I_{mR}^2 - 2 I_{mR}^2 I_{m1}^2\right] = \frac{1}{2} E I_{mR}^2 \left[4 I_R \left(I_R + 1\right) - 3\right] \qquad \text{(C-18)}$$

or

$$\frac{1}{2} I_{mR}^2 \left[3 I_R (I_R+1) - 3 - 4 I_1 (I_1+1) \right] = \frac{1}{2} E I_{mR}^2 \left[(2I_R - 1)(2I_R + 3) \right]$$

so

$$E = 1 - \frac{I_R (I_R+1) + 4 I_1 (I_1+1)}{(2I_R - 1)(2I_R + 3)} \qquad \text{(C-20)}$$

With this result then, Equation (C-10) and (C-11) yield

$$Q_{ij} = \frac{eQ}{I_1 (2I_1 - 1)} \left[1 - \frac{I_R (I_R+1) + 4 I_1 (I_1+1)}{(2I_R - 1)(2I_R + 3)} \right] \times \left[\frac{3 I_{Ri} I_{Rj} + I_{Rj} I_{Ri}}{2} - \delta_{ij} I_{mR}^2 \right] \qquad \text{(C-16)}$$

This expression is in the desired form and is dependent on the resultant angular momentum operator I. Since the factor containing the operator I is exactly the same as the factor containing the operator I in Equation (C-1), precisely the same reduction that led to the equation (C-9) is applicable so that the total quadrupole interaction energy for a homonuclear diatomic molecule may also be written as

$$\mathcal{H}_{E2} = \frac{e^2 g_I Q}{2 I_1 (2I_1 - 1) J (2J - 1)} \left[1 - \frac{I_R (I_R+1) + 4 I_1 (I_1+1)}{(2I_R - 1)(2I_R + 3)} \right] \times \left[3 (I_{mR} \cdot J_{mR})^2 + \frac{3}{2} I_{mR} \cdot J_{mw} - I_{mR}^2 J_{mw}^2 \right] \qquad \text{(C-17)}$$

From Equations (C-9) and (C-17) it is apparent that it is important to

determine the matrix elements of the quantity f defined by

$$f = 3\left(\underset{\sim}{I}\cdot\underset{\sim}{J}\right)^2 + \frac{3}{2}\left(\underset{\sim}{I}\cdot\underset{\sim}{J}\right) - \underset{\sim}{I}^2\,\underset{\sim}{J}^2 \qquad \text{(C-18)}$$

As discussed in Section III-4, these matrix elements may be desired either in an M_I, M_J representation or in an F m representation where F and m are the total and magnetic quantum numbers of

$$\underset{\sim}{F} = \underset{\sim}{I} + \underset{\sim}{J} \qquad \text{(C-20)}$$

The $m_I\, m_J$ representation is suitable to pertutbation calculations in strong external magnetic fields while the F m representation is suitable in the weak field limit. In the F m representation the calculation is trivial. The quantity $(\underset{\sim}{I}\cdot\underset{\sim}{J})$ and hence the operator f commutes with $\underset{\sim}{F}^2$ and $(\underset{\sim}{F})_z$. Therefore the matrix elements of f are diagonal in the F m representation. As shown in Equation (III-56) they are given by

$$\left(Fm\left|f\right|Fm\right) = \frac{3}{4}C\left(C+1\right) - I\left(I+1\right)J\left(J+1\right) \qquad \text{(C-21)}$$

where

$$C = \frac{1}{2}\left[F\left(F+1\right) - I\left(I+1\right) - J\left(J+1\right)\right] = \underset{\sim}{I}\cdot\underset{\sim}{J} \qquad \text{(C-22)}$$

In the $m_I\, m_J$ representation on the other hand the calculation is more difficult. However, as shown by Kellogg, Babi, Ramsey and Zacharias (KEL 40) they can easily be derived by writing f in the equivalent form

$$f = \tfrac{1}{2}\left[3J_z^2 - J(J+1)\right]\left[3I_z^2 - I(I+1)\right]$$
$$+ \tfrac{3}{4}\left[J_z J_+ + J_+ J_z\right]\left[I_z I_- + I_- I_z\right]$$
$$+ \tfrac{3}{4}\left[J_z J_- + J_- J_z\right]\left[I_z I_+ + I_+ I_z\right]$$
$$+ \tfrac{3}{4}\left[I_+^2 J_-^2 + I_-^2 J_+^2\right] \qquad \text{(C-23)}$$

The equivalence of Equations (C-18) and (C-23) can be seen by expansion and suitable use of the commutation rules of Equation III-44. With this the desired matrix elements can be obtained directly with the aid of Equations (III-35a). They are

$$(m_J, m_I | f | m_J, m_I) = \tfrac{1}{2}\left[3m_I^2 - I(I+1)\right]\left[3m_J^2 - J(J+1)\right],$$
$$(m_J, m_I | f | m_J - 1, m_J + 1) = 3/4(2m_J - 1)(2m_J + 1)\left[(J+m_J)(J-m_J+1)(I-m_I)(I+m_I+1)\right]^{\frac{1}{2}},$$
$$(m_J, m_I | f | m_J + 1, m_I - 1) = 3/4(2m_J + 1)(2m_J - 1)\left[(J-m_J)(J+m_J+1)(I+m_I)(I-m_I+1)\right]^{\frac{1}{2}},$$
$$(m_J, m_I | f | m_J - 2, m_I + 2) = 3/4\left[(J+m_J)(J+m_J-1)(J-m_J+1)(J-m_J+2)\right.$$
$$\left.\times (I-m_I)(I-m_I-1)(I+m_I+1)(I+m_I+2)\right]^{\frac{1}{2}},$$
$$(m_J, m_I | f | m_J + 2, m_I - 2) = 3/4\left[(J-m_J)(J-m_J-1)(J+m_J+1)(J+m_J+2)\right.$$
$$\left.(I+m_I)(I+m_I-1)(I-m_I+1)(I-m_I+2)\right].$$

(C-24)

In calculations of interactions of nuclear magnetic moments with external magnetic fields it is often desirable to calculate the matrix elements of $\underset{\sim}{I}$ in an F m representation. These matrix elements have been tabulated by Field and Lamb (FEL 45). The matrix elements diagonal

in F are relatively easy to calculate. For example, this may be done by using the general theorem of Section III-2.3 as in Equation (III-36) Alternatively, by the simple vector addition argument as used in Equation (III-89) it can be seen that the application of general theorem will give for matrix elements diagonal in F

$$\underset{\sim}{I} = \underset{\sim}{I} \cdot \frac{\underset{\sim}{F}}{|\underset{\sim}{F}|} \frac{\underset{\sim}{F}}{|\underset{\sim}{F}|} = \frac{I(I+1) - J(J+1) + F(F+1)}{2F(F+1)} \underset{\sim}{F}$$

$$= \frac{I(I+1) - J(J+1) + F(F+1)}{2F(F+1)}$$

$$\times \left[F_z \underset{\sim}{k} + \frac{1}{2} F_+ \left(\underset{\sim}{i} - i \underset{\sim}{j} \right) + \frac{1}{2} F_- \left(\underset{\sim}{i} + i \underset{\sim}{j} \right) \right]$$

(C-25)

The desired matrix elements can then be calculated immediately with the aid of Equation (III-35a). They are

$$\left(F, m \left| \underset{\sim}{I} \right| F, m+1 \right) = \frac{I(I+1) - J(J+1) + F(F+1)}{2F(F+1)} \cdot \frac{1}{2} \left\{ (F-m)(F+m+1) \right\}^{\frac{1}{2}} (\hat{i} + i\hat{j}),$$

$$\left(F, m \left| \underset{\sim}{I} \right| F, m \right) = \frac{I(I+1) - J(J+1) + F(F+1)}{2F(F+1)} m \underset{\sim}{k}$$

$$\left(F, m \left| \underset{\sim}{I} \right| F, m-1 \right) = \frac{I(I+1) - J(J+1) + F(F+1)}{2F(F+1)} \frac{1}{2} \left\{ (F+m)(F-m+1) \right\}^{\frac{1}{2}} (\hat{i} - i\hat{j})$$

(C-26)

In the case of matrix elements non-diagonal in F , the calculation is more difficult. One procedure is with the use of Equation (III-57). Alternatively they can be obtained from the formulae of Condon and Shortly (CON 35 pp. 63,64,67). The results are

$$(F,m\,|\,\underset{\sim}{I}\,|\,F+1,m+1)$$
$$=-\tfrac{1}{2}\left\{\frac{(F+1-I+J)(F+1+I-J)(I+J+2+F)(I+J-F)(F+m+1)(F+m+2)}{4(F+1)^2(2F+1)(2F+3)}\right\}^{\frac{1}{2}}(\mathbf{i}+i\mathbf{j}),$$

$$(F,m\,|\,\underset{\sim}{I}\,|\,F+1,m)$$
$$=\left\{\frac{(F+1-I+J)(F+1+I-J)(I+J+2+F)(I+J-F)\left[(F+1)^2-m^2\right]}{4(F+1)^2(2F+1)(2F+3)}\right\}^{\frac{1}{2}}\underset{\sim}{k},$$

$$(F,m\,|\,\underset{\sim}{I}\,|\,F+1,m-1)$$
$$=+\tfrac{1}{2}\left\{\frac{(F+1-I+J)(F+1+I-J)(I+J+2+F)(I+J-F)(-m+1)(F-m+2)}{4(F+1)^2(2F+1)(2F+3)}\right\}^{\frac{1}{2}}(\mathbf{i}-i\mathbf{j}),$$

$$(F,m\,|\,\underset{\sim}{I}\,|\,F-1,m+1)$$
$$=+\tfrac{1}{2}\left\{\frac{(F-I+J)(F+I-J)(I+J+1+F)(I+J+1-F)(F-m)(F-m-1)}{4F^2(2F-1)(2F+1)}\right\}^{\frac{1}{2}}(\mathbf{i}+i\mathbf{j}),$$

$$(F,m\,|\,\underset{\sim}{I}\,|\,F-1,m)$$
$$=\left\{\frac{(F-I+J)(F+I-J)(I+J+1+F)(I+J+1-F)(F^2-m^2)}{4F^2(2F-1)(2F+1)}\right\}^{\frac{1}{2}}\underset{\sim}{k} \qquad (C\text{-}27).$$

$$(F,m\,|\,\underset{\sim}{I}\,|\,F-1,m-1)$$
$$=-\tfrac{1}{2}\left\{\frac{(F-I+J)(F+I-J)(I+J+1+F)(I+J+1-F)(F+m)(F+m-1)}{4F^2(2F-1)(2F+1)}\right\}^{\frac{1}{2}}(\mathbf{i}-i\mathbf{j}).$$

BIE 52 Biedenham, Blatt, and Rose; Re . Mod. Phys. 24, 249 (1952)

BOH 50 Bohr and Weisskopf; Phys. Rev. 77, 94 (1950)

BOH 51 A. Bohr; Phys. Rev. 81, 134 and 331 (1951)

BOH 53 Bohr and Mottelson; Phys. Rev. 89, 316 (1953) & 90, 717 (1953) & Dan. Mat. Fys. Med. 26 & 27

BRA 48 J.K.Bragg; Phys. Rev. 74, 533 (1948)

BRE 28 S. Breit; Nature 122, 649 (1928)

BRE 30 S. Breit; Phys. Rev. 3 , 1447 (1930)

BRE 31 Breit and Rabi; Phys. Rev. 38, 2082 (1931)

BRE 32 Breit and Rosenthal; Phys. Rev. 41, 459 (1932) & 42, 348 (1932)

BRE 33 Breit and Wills; Phys. Rev. 44, 470 (1933)

BRE 48 Breit and Brown; Phys. Rev. 74, 1278 (1948)

BRE 49 Breit, Brown and Arfken; Phys. Rev. 76, 1299 (1949)

CAS 32 H.B.S.Casimir; Zeits. f. Physik, 77, 811 (1932)

CAS 35 H.B.S.Casimir; Physica 2, 719 (1935)

CAS 36 H.B.S.Casimir; Teyler's Tweede Genootschap, 11, 36 (1936)

CAS 42 Casimir and Karreman; Physica 9, 494 (1942)

CON 35 Condon and Shortly; Theory of Atomic Spectra

DIR 48 P.A.M.Dirac; Phys. Rev. (1948)

FAL 50 Falkoff and Uhlenbeck; Phys. Rev. 79, 323 (1950)

FEE 53 Feenberg and Pake; Angular Momenta, Addison Wesley Press (1953)

FEL 45 Feld and Lamb; Phys. Rev. 67, 15 (1945)

FEL 47 Feld; Phys. Rev. 72, 1116 (1947)

FER 30 E.Fermi; Zeits. f. Physik, 60, 320 (1930)

FER 33 Fermi and Segre; Z. Physik 82, 729 (1933)

GOU 29 Goudsmit and Bacher; Phys. Rev. 34, 1499 (1929) & 37, 663 (1931)

GOU 33 S.Goudsmit; Phys.Rev. 43, 636 (1933)

GOU 33a Goudsmit and Bacher; Phys. Rev. 43, 894 (1933)

GUN 51 Gunther-Mohr, Geschwind, and Townes; Phys. Rev; 81, 289 (1951)

JAC54 Jaccarina, et alia; Phys. Rev. 1954.

KEL 39a Kellogg, Babi, Ramsey & Zacharias; Phys. Rev. 56, 728 (1939)

KEL 40 Kellogg, Babi, Ramsey & Zacharias, Phys. Rev. 57, 677 (1940)

KNI 49a Knight and Feld; Report 123, Research Lab. Electronics (M.I.T.)

KOP 40 H.Kopfermann; Kernmomente, Leipzig (1940)

KOS 52 G.J.Koster; Phys. Rev. 86, 148 (1952)

LOW 51 Low and Salpeter; Phys. Rev. 77, 361 (1950) & 83, 478 (1951)

MAR 40 H.Margenan; Phys. Rev. 57, 383 (1940)

MIL 38 Millman, Babi, & Zacharias; Phys. Rev. 53, 384 (1938)

NIE 47 Nierenberg and Ramsey; Phys. Rev. 72, 1075 (1947)

NIE 48 Nierenberg, Babi and Slotnick; Phys. Rev. 73, 1430 (1948).

NOR 40 A.Nordsieck; Phys. Rev. 58, 310 (1940)

PAU 24 W.Pauli; Naturwiss, 12, 741 (1924)

PAU 32 W.Pauli; Handbuch der Physik 24 (1932)

PEI 99 B.O.Peirce; Short Table of Integrals, Ginn & Co. (1899)

POU 50 R.V.Pound; Phys.Rev. 79, 683 (1950)

PUR 50 Purcell and Ramsey; Phys. Rev. 78, 699 (1950)

RAC 31 G.Racah; Zeits. F.Physik 76, 431 (1931)

RAC 42 G.Racah; Phys. Rev. 62, 438 (1942)

RAM 40 N.J.Ramsey; Phys. Rev. 58, 226 (1940)

RAM 50d N.J.Ramsey; Phys. Rev. 78, 699 (1950)

RAM 51c N.J.Ramsey; Phys. Rev. 83, 540 (1951)

RAM 52a N.J.Ramsey; Phys. Rev. 85, 60 (1952)

RAM 52b N.J.Ramsey; and Purcell; Phys. Rev. 85, 143 (1952)

RAM 53e N.J.Ramsey; Nuclear Moments, Wiley & Sons (1953)

RAM 53f N.J.Ramsey, Malenka and Kruse; Phys. Rev. 91, 1162 (1953)

ROS 48 L.Rosenfeld; Nuclear Forces, vols. I & II, Interscience Press 1948

SCH 37 J.Schwinger; Phys. Rev. 51, 545 (1937)

SCH 46 J.Schwinger; and J.Blatt; Lecture Notes on Nuclear Physics

SCH 49 L.Schiff; Quantum Mechanics, McGraw Hill (1949)

SCH 52 J.Schwinger; On Angular Momentum, Nucl. Dev. Assoc. (1952)

SCH 54 C.Schwartz; Phys. Rev. 97, 380 (1955)

SMI 51a Smith, Purcell & Ramsey; private communication

SMI 51b J.Smith; Thesis, Harvard University

WEY 39 H.Weyl; The Classical Groups, Princeton Univ. Press (1939)

WIG 31 E.Wigner; Gruppentheorie, (1931)

QUANTUM MECHANICS

by

Julian SCHWINGER

Harvard University



Inv. 1654

ECOLE d'ÉTÉ de PHYSIQUE THÉORIQUE

Université de Grenoble

Cours professé à

L'ECOLE D'ETE DE PHYSIQUE THEORIQUE

Les Houches (Haute-Savoie) ,France

Juillet 1955

I. - The ALGEBRA OF MEASUREMENT

The classical theory of measurement is built upon the conception that the interaction between the system of interest and the measuring apparatus can be made arbitrarily small, or at least precisely compensated so that one can speak meaningfully of an idealized measurement that disturbs no property of the system . But it is characteristic of atomic phenomena that the interaction between system and instrument is: notarbitrarily small, nor can the disturbance produced by the interaction be compensated precisely since, to some extent, it is uncontrollable and unpredictable. Accordingly, a measurement of one property can produce unavoidable changes in the value previously assigned to another property and it is without meaning to speak of a microscopic system possessing precise values for all its attributes . This contradicts the classical representation of all physical quantities by numbers . The laws of atomic physics must therefore be expressed in a nonclassical mathematical framework that constitutes a symbolic expression of the properties of microscopic measurement .

In the simplest type of measurement, an ensemble of independent similar systems is sorted by the apparatus into subensembles distinguished by definite values $a' = a^1, a^2, \ldots, a^n$ of the physical quantity A. Let M (a') symbolize the selective measurement that accepts systems possessing the value a' of property A and reject all others . We define the addition of such symbols to signify measurements that select systems with any of the values contained in the sum, and the multiplication of the measure-

ment symbols to represent the successive performance of measurements (read from right to left) . It follows from the physical meaning of these operations that addition is commutative and associative while multiplication is associative and distributive . With 1 and 0 symbolizing the measurements that respectively accept and reject all systems, the properties of these elementary selective measurements are expressed by

$$M(a')\,M(a') = M(a') \qquad (I.1)$$

$$M(a')\,M(a'') = 0 \qquad a' \neq a'' \qquad (I.2)$$

$$\sum_{a'} M(a') = 1, \qquad (I.3)$$

for M (a') accepts every system produced by the measurement M (a') and rejects every system produced by M (a"), a" ≠ a', while every system appears in some subensemble . The notation 1 and 0 is consistant with the algebraic properties of these symbols . The introduction of the numbers 1 and 0 as multipliers permits (I 1) and (I 2) to be combined in

$$M(a')\,M(a'') = \delta(a',a'')\,M(a')$$

where

$$\delta(a',a'') = 1 \text{ for } a' = a''$$

$$= 0 \text{ for } a' \neq a''$$

Two physical quantities A_1 and A_2 are said to be compatible when the measurement of one does not destroy the knowledge gained by previous measurement of the other . The selective measurements $M(a'_1)$ and $M(a'_2)$, performed in either order, produces an ensemble of systems

to which one can simultaneously assign the values a'_1 to A_1 and a'_2 to A_2. The symbol of this compound measurement is :

$$M(a'_1, a'_2) = M(a'_1)\, M(a'_2) = M(a'_2)\, M(a'_1).$$

By a complete set A of compatible physical quantities $A_1, \ldots A_k$, we mean that every pair of these quantities is compatible and that no other quantities exist, apart from functions of A, that are compatible with every member of the set . The measurement symbol

$$M(a') = \prod_{i=1}^{k} M(a'_i)$$

then describes a complete measurement, which is such that the systems chosen possess definite values for the maximum number of attributes . Any attempt to determine the value of still another independent physical quantity will produce uncontrollable changes in one or more of the previously assigned values . Thus the optimum state of knowledge concerning a given system is realized by the complete measurement M(a') are said to be in the state a'. The symbolic properties of complete measurements are also given by (I.1) , (I.2) , (I.3)

A more general type of measurement incorporates a disturbance that produces a change of state. The symbol $M(a', a'')$ indicates a selective measurement in which systems are accepted only in the state a" and emerge in the state a' . The measurement process $M(a')$ is the special case in which no change of state occurs :

$$M(a') = M(a', a') .$$

The properties of successive measurements of the type $M(a',a'')$ are symbolized by

$$M(a',a'')\ M(a''',a'''') = \delta(a'',a''')\ M(a',a'''') \qquad (I.4)$$

for, if $a'' \neq a'''$, the second stage of the compound apparatus accepts none of the systems that emerge from the first stage, while if $a'' = a'''$, all such systems enter the second stage and the compound measurement serves to select systems in the state a'''' and produce then in the state a'. Note that if the two stages are reversed, we have

$$M(a''',a'''')\ M(a',a'') = \delta(a',a'''')\ M(a''',a'')$$

which differs in general from (I.4). Hence the multiplication of measurement symbols is non-commutative.

The physical quantities contained in one complete set A do not comprise the totality of physical attributes of the system. One can form other complete sets B, C, ..., which are mutually incompatible. For each choice of non-interfering physical characteristics there is a set of selective measurements referring to systems in the appropriate states: $M(b',b'')$, $M(c',c'')$, ... The most general selective measurement involves two incompatible sets of properties. We symbolize by $M(a',b')$ the measurement process that rejects all impinging systems except those in the state b', and permits only systems in the state a' to emerge from the apparatus. The general compound measurement $M(a',b')\ M(c',d')$ serves to select systems in the state d' and produces them in the state a',

as symbolized by $M(a',d')$. Now the first stage supplies systems in the state c',while the second stage accepts only systems in the state b' . The examples compound measurements that we have already considered involve the passage of all systems,or no systems between the two stages,as represented by the multiplication numbers 1 or 0. But measurement of the properties B performed on a system in the state c',referring to characteristics incompatible with B,will generally yield a statistical distribution of the possible values . Hence only a determinate fraction of the systems emerging from the first stage will be accepted by the second stage . We symbolize this in the multiplication law

$$M(a',b')\ M(c',d') = \langle b' \mid c' \rangle\ M(a',d') \qquad I.5)$$

where $\langle b' \mid c' \rangle$ is a number characterizing the statistical relation between the states b' and c' . In particular :

$$\langle a' \mid a'' \rangle = \delta(a',a'')$$

Special exemples of (I.5)) are

$$M(a')\ M(b',c') = \langle a' \mid b' \rangle\ M(a',c') \qquad (I.6))$$

$$M(a',b')\ M(c') = \langle b' \mid c' \rangle\ M(a',c')$$

We infer from (I.3) that

$$\sum_{a'} \langle a' \mid b' \rangle\ M(a',c') = \sum_{a'} M(a')\ M(b',c') = M(b',c')$$

and

$$\sum_{c'} M(a',c') \langle b' \mid c' \rangle = M(a',b')$$

which show that measurement symbols of one type can be expressed as a linear combination of the measurement symbols of another type . The general relation is

$$M(c',d') = \sum_{a'b'} M(a')\, M(c',d')\, M(b')$$

$$= \sum_{a'b'} \langle a' \mid c' \rangle\, M(a',b')\, \langle d' \mid b' \rangle \qquad (I.7)$$

From its role in effecting such connections, the totality of numbers $\langle a' \mid b' \rangle$ is called the transformation function relating the a and b descriptions, where the phrase " a description " signifies the description of a system in terms of the states produced by selective measurements of the complete set of physical quantities A.

A fundamental composition property of transformation functions is obtained on comparing

$$\sum_{b'} M(a')\, M(b')\, M(c') = \sum_{b'} \langle a' \mid b' \rangle\, \langle b' \mid c' \rangle\, M(a',c')$$

with

$$M(a') \sum_{b'} M(b')\, M(c') = M(a')\, M(c') = \langle a' \mid c' \rangle\, M(a',c')$$

namely

$$\sum_{b'} \langle a' \mid b' \rangle\, \langle b' \mid c' \rangle = \langle a' \mid c' \rangle$$

On identifying the a and c descriptions, this becomes

$$\sum_{b'} \langle a' \mid b' \rangle\, \langle b' \mid a'' \rangle = \langle a' \mid a'' \rangle = \delta(a',a'')$$

and similarly

$$\sum_{a'} \langle b' | a' \rangle \langle a' | b'' \rangle = \delta (b', b'')$$

As a consequence, we observe that

$$\sum_{a'} \sum_{b'} \langle a' | b' \rangle \langle b' | a' \rangle = \sum_{a'} 1 = \sum_{b'} \sum_{a'} \langle b' | a' \rangle \langle a' | b' \rangle = \sum_{b'} 1$$

which means that N, the total number of states, obtained in a complete measurement, is independent of the particular choice of compatible physical quantities that are measured . Hence the total number of measurement symbols, in additive combination, thus form the elements of a linear algebra of dimensionality N^2: the algebra of measurement . The elements of the measurement. The elements of the measurement algebra are called operators.

The number $\langle a' | b' \rangle$ can be regarded as a linear numerical function of the measurement symbol (operator) $M (b' , a')$. We call this linear correspondence between operators and numbers the trace,

$$\langle a' | b' \rangle = \text{tr } M (b' , a') \qquad (I.8)$$

and observe from the general linear relation (I.7) that

$$\text{tr } M (c' , d') = \sum_{a' b'} a' \; c' \; \text{tr } M (a' , b') \langle d' | b' \rangle$$

$$= \sum_{a' b'} \langle d' | b' \rangle \langle b' | a' \rangle \langle a' | c' \rangle = \langle d' | c' \rangle$$

which verifies the consistency of (I.8) . In particular,

$$\text{tr } M (a' , a'') = \delta (a' , a'')$$

$$\text{tr } M (a') = 1 .$$

The trace of a measurement symbol product is

$$\text{tr } M (a', b') M (c', d') = \langle b' | c' \rangle \text{ tr } M (a', d')$$

$$= \langle b' | c' \rangle \langle d' | a' \rangle \qquad (I.9)$$

which is to be compared with

$$\text{tr } M(c', d')\ M(a', b') = \langle d' | a' \rangle \langle b' | c' \rangle$$

Hence, despite the non-commutativity of multiplication, the trace of a product of two factors is independent of the multiplication order . This applies to any two elements X, Y, of the measurement algebra,

$$\text{tr } X\ Y = \text{tr } Y\ X.$$

A special case of (I.9) is

$$\text{tr } M(a')\ M(b') = \langle a' | b' \rangle \langle b' | a' \rangle$$

Now let us observe that the general multiplication law and the definition of the trace are preserved when we make the substitutions

$$M(a', b') \rightarrow \lambda(a')^{-1} M(a', b') \lambda(b')$$
$$\langle a' | b' \rangle \rightarrow \lambda(a') \langle a' | b' \rangle \lambda(b')^{-1}, \qquad \text{(I.10)}$$

in which the numbers $\lambda(a')$ and $\lambda(b')$ can be given arbitrary non-zero values . The appearance of the inverse numbers is required to maintain the unit value of $\langle a' | a' \rangle$, or equivalently, to leave unaltered the elementary measurement symbols $M(a')$. In view of this arbitrariness, a transformation function cannot, of itself, possess a direct physical interpretation, but must enter in some combination that remains invariant under the substitution (I.10)

The appropriate basis for the statistical interpretation of the transformation function can be inferred from the sequence of selective measurements $M(b')\ M(a')\ M(b')$, which differs from $M(b')$ in

virtue of the disturbance attended upon the intermediate A measurement . Only a fraction of the systems selected by the initial B- measurement is transmitted through the complete apparatus . Correspondingly,we have the symbolic equation

$$M(b')\ M(a')\ M(b') = p(a',b')\ M(b')$$

where the number

$$p(a',b') = \langle a' \mid b' \rangle \langle b' \mid a' \rangle$$

exhibits the required characteristics of invariance under the transformation (I.IO)

If the A-measurement selects systems in either of two states,there is a corresponding additivity of the number $p(a',b')$,

$$M(b')\ (M(a') + M(a''))\ M(b') = (p(a',b') + p(a'',b'))\ M(b'),$$

and for the A-measurement that selects all systems,there appears

$$M(b') \sum_{a'} M(a')\ M(b') = M(b')$$

whence

$$\sum_{a'} p(a',b') = 1$$

These properties qualify $p(a',b')$ for the role of the probability that one observes the state a' in a measurement performed on a system known to be in the state b' . But a probability is a real,non-negative number. Hence,we must impose a restriction on the numbers of the measurement algebra,by requiring that $\langle a' \mid b' \rangle$ and $\langle b' \mid a' \rangle$ form a pair of

complex conjugate numbers,

$$\langle b' \mid a' \rangle = \langle a' \mid b' \rangle^{*} , \qquad (I.II)$$

for then

$$p\ (a',b') = \left| \langle a' \mid b' \rangle \right|^{2} \geq 0 .$$

To maintain the complex conjugate relation (I.II), the numbers $\lambda(a')$ of (I.I0) must obey

$$\lambda^{*}(a') = \lambda(a')^{-1}$$

and thus have the form

$$\lambda(a') = e^{i\varphi(a')}$$

in which the phases $\varphi(a')$ can assume arbitrary real values .

Another satisfactory aspect of the probability formula is the symmetry property

$$p\ (a',b') = p\ (b',a') .$$

Let us recall the arbitrary convention that accompanies the interpretation of the measurement symbols and their products - the order of events is read from right to left (sinistrally) . But any measurement symbol equation is equally valid if interpreted in the opposite sense (dextrally) and no physical result should depend upon which convention is employed . On introducing the dextral interpretation $\langle a' \mid b' \rangle$ acquires the meaning possessed by $\langle b' \mid a' \rangle$ with the sinistral convention . We conclude that the probability connecting states a' and b' in given sequence must be constructed symmetrically from $\langle a' \mid b' \rangle$ and $\langle b' \mid a' \rangle$.

We shall speak of the introduction of the opposite convention for measurement symbol equations as the adjoint operation, with the symbol $\dagger$.

Thus

$$M\ (a',b')^{\dagger} = M\ (b',a')$$

and

$$M\ (a',a'')^{\dagger} = M\ (a'',a')$$

In particular

$$M\ (a')^{\dagger} = M\ (a')\ ,$$

which characterizes $M\ (a')$ as a self-adjoint or Hermitian operator .

For measurement symbol products, we have

$$(M\ (a',b')\ M\ (c',d'))^{\dagger} = M\ (d',c')\ M\ (b',a')$$

$$= M\ (c',d')^{\dagger}\ M\ (a',b')^{\dagger}$$

or, equivalently,

$$(\langle b'|\ c'\rangle\ M\ (a',d'))^{\dagger} = \langle c'|\ b'\rangle\ M\ (d',a')$$

$$= \langle b'|\ c'\rangle\ M\ (a',d')^{\dagger}$$

The significance of addition is uninfluenced by the adjoint procedure, which extends these properties to all elements of the measurement algebra,

$$(X + Y)^{\dagger} = X^{\dagger} + Y^{\dagger}\ ,\quad (X\ Y)^{\dagger} = Y^{\dagger}\ X^{\dagger}\ ,\quad (\lambda\ X)^{\dagger} = \lambda^{*} X^{\dagger}\ ,$$

in which λ is an arbitrary number .

The use of complex numbers in the measurement algebra imples the existence of a dual algebra in which all numbers are replaced by the

complex conjugate numbers . No physical result can depend upon which algebra is employed . If the operators of the dual algebra are written $X^{\ast}$, the correspondence between the two algebras is governed by the laws :

$$(X + Y)^{\ast} = X^{\ast} + Y^{\ast} , \; (X Y) ^{\ast} = X^{\ast} Y^{\ast} , \; (\lambda X)^{\ast} = \lambda^{\ast} X^{\ast} .$$

The formation of the adjoint within the complex conjugate algebra is called transposition

$$X^{T} = X^{\ast \dagger} = X^{\dagger \ast} ,$$

It has the algebraic properties

$$(X + Y)^{T} = X^{T} + Y^{T}, \; (X Y) ^{T} = Y^{T} X^{T}, \; (\lambda X) ^{T} = \lambda X^{T} .$$

The measurement symbols of a given description provide a basis for the representation of an arbitrary operator by N^2 numbers, and the abstract properties of operators are realized by the combinatorial laws of these arrays of numbers, which are those of matrices . Thus

$$X = \sum_{a'a''} \langle a' | X | a'' \rangle \; M (a', a'')$$

defines the matrix of X in the a description or a representation, and the product

$$X Y = \sum \langle a' | X | a'' \rangle \; M (a', a'') \sum \langle a''' | Y | a'''' \rangle \; M (a''', a'''')$$

$$= \sum \langle a' | X | a'' \rangle \; \delta (a'', a''') \langle a''' | Y | a'''' \rangle \; M (a', a'''')$$

shows that

$$\langle a' | X Y | a'''' \rangle = \sum_{a''} \langle a' | X | a'' \rangle \langle a'' | Y | a'''' \rangle$$

The elements of the matrix that represents X can be expressed as

$$\langle a' | X | a'' \rangle = \mathrm{tr} \quad (X \; M (a'', a')) ,$$

and in particular

$$\langle a' | X | a' \rangle = \mathrm{tr} \quad (X \; M (a'))$$

The sum of the diagonal element of the matrix is the trace of the operator. The corresponding basis in the dual algebra is $M (a', a'')^{\star}$, and the matrices that represent $X^{\star}$ and X^{T} are the complex conjugate and transpose, respectively, of the matrix that represents X. The operator $X^{\dagger} = X^{T \star}$, an element of the same algebra as X, is represented by the tranposed complex conjugate, or adjoint matrix.

Th matrix of X in the mixed a b description is defined by

$$X = \sum \langle a' | X | b' \rangle \; M \; (a', b')$$

where

$$\langle a' | X | b' \rangle = \mathrm{tr} \quad (X \; M (b', a')$$

The rule of multiplication for matrices in mixed representations is

$$\langle a' | X Y | c' \rangle = \sum_{b'} \langle a' | X | b' \rangle \langle b' | Y | c' \rangle$$

The connection between the matrices of a given operator in various representations follows from the general linear relation between measurement symbols. Thus

$$\langle a' | X | d' \rangle = \mathrm{tr} \; X \, M (d', a')$$

$$= \mathrm{tr} \; X \sum \langle c' | d' \rangle \; M \; (c', b') \; \langle a' | b' \rangle$$

$$= \sum \langle a' \mid b' \rangle \langle b' \mid X \mid c' \rangle \langle c' \mid d' \rangle$$

If X is displayed in the a b basis, the adjoint operator appears in the b a basis, with the matrix

$$\langle b' \mid X^{+} \mid a' \rangle = \langle a' \mid X \mid b' \rangle^{*}$$

As an application of mixed representations, we present an operator equivalent of the general properties of transformation functions,

$$\sum_{b'} \langle a' \mid b' \rangle \langle b' \mid c' \rangle = \langle a' \mid c' \rangle$$
$$\langle a' \mid b' \rangle^{*} = \langle b' \mid a' \rangle \qquad \text{(I.I2)}$$

which is achieved by a differential characterisation of the transformation functions . If $\delta \langle a' \mid b' \rangle$ and $\delta \langle b' \mid c' \rangle$ are any conceivable infinitesimal alterations of the corresponding transformation functions, the implied variation of $\langle a' \mid c' \rangle$ is

$$\delta \langle a' \mid c' \rangle = \sum \left[\delta \langle a' \mid b' \rangle \langle b' \mid c' \rangle + \langle a' \mid b' \rangle \delta \langle b' \mid c' \rangle \right] \qquad \text{(I.I3)}$$

while (I.I2) asserts that

$$\delta \langle a' \mid b' \rangle^{*} = \delta \langle b' \mid a' \rangle$$

One can regard the array of numbers $\delta \langle a' \mid b' \rangle$ as the matrix of an operator in the a b representation . We therefore write

$$\delta \langle a' \mid b' \rangle = i \langle a' \mid \delta W_{ab} \mid b' \rangle$$

which is the definition of an infinitesimal operator δW_{ab} . If infinitesimal operators δW_{bc} and δW_{ac} are similarly defined, the differential property (I.I3) becomes the matrix equation

$$\langle a' | \delta W_{ac} | c' \rangle = \sum_b \left[\langle a' | \delta W_{ab} | b' \rangle \langle b' | c' \rangle + \langle a' | b' \rangle \langle b' | \delta W_{bc} | c' \rangle \right]$$

from which we infer the operator equation

$$\delta W_{ac} = \delta W_{ab} + \delta W_{bc}$$

Thus the multiplicative composition law of transformation functions is expressed by an additive composition law for the infinitesimal operators δW .

On identifying the a and b descriptions, we learn that

$$\delta W_{aa} = 0$$

or

$$\delta \langle a' | a'' \rangle = 0$$

which expresses the fixed numerical values of the transformation functions,

$$\langle a' | a'' \rangle = \delta(a', a'') .$$

Indeed, the latter is not an independent condition on transformation functions but is implied by the composition property and the requirement that transformation functions as matrices be nonsingular . If we identify the a and c description we are informed that

$$\delta W_{ba} = - \delta W_{ab}$$

Now

$$\delta \langle a' | b' \rangle^* = - i \langle a' | \delta W_{ab} | b' \rangle^* = - i \langle b' | \delta W_{ab}^\dagger | a' \rangle$$

which must equal

$$\delta \langle b' | a' \rangle = i \langle b' | \delta W_{ba} | a' \rangle$$

and therefore

$$\int W_{ab}{}^{\dagger} = - \int W_{ba} = \int W_{ab} .$$

The complex conjugate property of transformation functions is thus expressed by the statement that the infinitesimal operators $\int W$ are Hermitian .

The expectation value of property A for the state b' is the average of the possible values of A weighted by the probabilities of occurence that are characteristic of state b' .
On writing the probability formula as

$$p\ (a',b') = \mathrm{tr}\ M\ (a')\ M\ (b') ,$$

the expectation value becomes

$$\langle A \rangle_{b'} = \sum_{\sim} a'\ p\ (a',b') = \mathrm{tr}\ A\ M\ (b') = \langle b' | A | b' \rangle$$

where the operator A is

$$A = \sum_{\sim} a'\ M\ (a') .$$

The correspondence thus obtained between operators and physical quantities is such that the function f (A) of property A is assigned the operator f (A), and the operators associated with a complete set of compatible physical quantities form a complete set of commuting Hermitian operators. In particular the function of A that exhibits the value unity in the state a' and zero otherwhere is characterized by the operator M (a') .

2. - THE GEOMETRY OF STATES.

The uncontrollable disturbance attendant upon a measurement implies that the process of measurement is indivisible . That is to say, any attempt to trace the history of a system during a measurement process usually changes the nature of the measurement that is being performed . Hence to conceive of a given selective measurement $M(a',b')$ as a compound measurement is without physical implication . It is only of significance that the first measurement process selects systems in the state b' and the last produces therein the state a' ; the interposed stages are without meaning for the measurement as a whole . Indeed,we can even invent a non-physical state to serve as the intermediary . We shall call this mental construct the null state 0, and write

$$M(a',b') = M(a',0)\ M(0,b') . \qquad (2.1)$$

The measurement process that selects a system in the state b' and produces it in the null state

$$M(0,b') = \overline{\Phi}(b')$$

can be described as the annihilation of a system in the state b' , and the production of a system in the state a' following its selection from the null state

$$M(a',0) = \overline{\Psi}(a')$$

can be characterized as the creation of a system in the state a . Thus the content of (2.1) is the indiscernibility of $M(a',b')$ from

the compound process of the annihilation of a system in the state b' followed by the creation of a system in the state a' ,

$$M\ (a', b') = \underline{\Psi}(a')\ \underline{\Phi}(b') \qquad (2.2)$$

The extension of the measurement algebra to include the null state supplies the properties of the $\underline{\Psi}$ and $\underline{\Phi}$ symbols . Thus

$$\underline{\Psi}(a')^{\dagger} = \underline{\Phi}(a'), \quad \underline{\Phi}(b')^{\dagger} = \underline{\Psi}(b')$$

and

$$\begin{aligned} \underline{\Psi}(a')\underline{\Psi}(b') &= \underline{\Phi}(a')\underline{\Phi}(b') = 0, \\ M(a',b')\,\underline{\Phi}(c') &= \underline{\Psi}(a')\ M(b',c') = 0, \end{aligned} \qquad (2.3)$$

whereas

$$\begin{aligned} M(a',b')\,\underline{\Psi}(c') &= \langle b' \mid c' \rangle\ \underline{\Psi}(a') \\ \underline{\Phi}(a')\ M(b',c') &= \langle a' \mid b' \rangle\ \underline{\Phi}(c') \end{aligned} \qquad (2.4)$$

and

$$\underline{\Phi}(a')\ \underline{\Psi}(b') = \langle a' \mid b' \rangle\ M(0).$$

Some properties of $M(0)$ are

$$\begin{aligned} \underline{\Psi}(a')\ M(0) &= \underline{\Psi}(a') \\ M(0)\,\underline{\Phi}(a') &= \underline{\Phi}(a') \end{aligned}$$

and

$$M(0)\,\underline{\Psi}(a') = \underline{\Phi}(a')\ M(0) = 0.$$

We must also observe that, in the extended measurement algebra, the completeness of the states is symbolized by

$$1 = \sum_{a'} M(a') + M(0).$$

The fundamental arbitrariness of measurement symbols expressed by the substitution (I.IO)

$$M\ (a',b') \rightarrow e^{-i\varphi(a')} M(a',b') \ e^{i\varphi(b')} \qquad (2.5)$$

implies the accompanying substitution

$$\underline{\Psi}(a') \rightarrow e^{-i\varphi(a')} \underline{\Psi}(a'), \quad \underline{\Phi}(b') \rightarrow e^{i\varphi(b')} \underline{\Phi}(b'), \qquad (2.6)$$

in which we have effectively removed $\varphi(0)$ by expressing all the other phases relative to it .

The characteristics of the operators $M(a',b')$ can now be derived from those of the $\underline{\Psi}$ and $\underline{\Phi}$ symbols . Thus

$$M(a',b')^{\dagger} = \underline{\Phi}(b')^{\dagger}\, \underline{\Psi}(a')^{\dagger} = \underline{\Psi}(b')\, \underline{\Phi}(a') = M(b',a')$$

and

$$\mathrm{tr}\ M(b',a') = \mathrm{tr}\, \underline{\Psi}(b')\, \underline{\Phi}(a') = \mathrm{tr}\, \underline{\Phi}(a')\, \underline{\Psi}(b')$$

$$= \langle a' | b' \rangle$$

while

$$M(a',b')\, M(c',d') = M(a',b')\, \underline{\Psi}(c')\, \underline{\Phi}(d') = \langle b' | c' \rangle\, \underline{\Psi}(a')\, \underline{\Phi}(d')$$

$$= \langle b' | c' \rangle\ M(a',d') .$$

The substitution (2.6) transforms the measurement symbols in accordance with (2.5).

The various equivalent statements contain in (2.3) show that the only significant products (those not identically zero) are of the

form $\underline{\Psi}\,\underline{\Phi}$, $\underline{\Phi}\,\underline{\Psi}$ and $X\underline{\Psi}$, $\underline{\Phi}X$, in addition to X Y, where the latin symbols are operators, elements of the physical measurement algebra .

According to the construction (2.2), all operators are linear combination of products $\underline{\Psi}\,\underline{\Phi}$,

$$X = \sum \underline{\Psi}(a') \langle a' \mid X \mid b' \rangle \underline{\Phi}(b') ,$$

and the evaluation of the class of products $X\underline{\Psi}$, $\underline{\Phi}X$ reduces to the ones contained in (2.4) ,

$$\underline{\Psi}(a')\,\underline{\Phi}(b')\,\underline{\Psi}(c') = \underline{\Psi}(a') \langle b' \mid c' \rangle$$
$$\underline{\Phi}(a')\,\underline{\Psi}(b')\,\underline{\Phi}(c') = \langle a' \mid b' \rangle \underline{\Phi}(c') .$$

Hence in any combination of interest, a product $\underline{\Phi}\,\underline{\Psi}$ is effectively equal to a number

$$\underline{\Phi}(a')\,\underline{\Psi}(b') = \langle a' \mid b' \rangle ,$$

and, in particular :

$$\underline{\Phi}(a')\,\underline{\Psi}(a'') = \delta(a', a'') .$$

Il must also be observed that in any application of the completeness property as an operator equation, we have, in effect

$$I = \sum_{a'} M(a') = \sum_{a'} \underline{\Psi}(a')\,\underline{\Phi}(a'),$$

Thus

$$X = \sum \underline{\Psi}(a')\,\underline{\Phi}(a')\,X\,\underline{\Psi}(b')\,\underline{\Phi}(b'),$$

which shows that

$$\underline{\Phi}(a')\,X\,\underline{\Psi}(b') = \langle a' \mid X \mid b' \rangle .$$

The bracket symbols

$$\underline{\Phi}\,(a') = \langle a'|\,, \quad \underline{\Psi}(b') = |b'\rangle$$

are designed to make this result an automatic consequence of the notation .

We have associated a $\underline{\Psi}$ and a $\underline{\Phi}$ symbol with each of the N physical states of a description . Now the symbols of one description are linearly related to those of another description . Thus

$$\underline{\Psi}\,(b') = \sum_{a'} \underline{\Psi}(a')\,\underline{\Phi}(a')\,\underline{\Psi}(b') = \sum \psi(a')\langle a'|b'\rangle \qquad (2.7)$$

and

$$\underline{\Phi}\,(a') = \sum \underline{\Phi}(a')\,\underline{\Psi}(b')\,\underline{\Phi}(b') = \sum \langle a'|b'\rangle \underline{\Phi}(b') \qquad (2.8)$$

which also implies the linear relation between measurement symbols of different types . Arbitrary numerical multiples of $\underline{\Psi}$ and $\underline{\Phi}$ symbols thus form the elements of two mutually adjoint algebras of dimensionality N, which are vector algebras since there is no significant multiplication of elements within each algebra . We are thereby presented with an N-dimensional geometry - the geometry of states - from which the measurement algebra can be derived, with its properties characterized in geometrical language . This geometry is metrical since the number $\underline{\Phi}\,\underline{\Psi}$ defines a scalar product . The vectors $\underline{\Psi}\,(a')$ and $\underline{\Phi}\,(a')$ of the a description provide an orthonormal basis or coordinate system and the vector transformation equations (2.7) , (2.8) describe a change of coordinate system . The product of an operator with a vector expresses a mapping upon another

vector in the same space,

$$X\Psi(b') = \sum_{a'} \Psi(a')\Phi(a')\, X\, \Psi(b') = \sum_{a'} \Psi(a') \langle a'|X|b'\rangle$$

$$\Phi(a')\, X = \sum_{b'} \Phi(a')\, X\, \Psi(b')\, \Phi(b') = \sum_{b'} \langle a'|X|b'\rangle \Phi(b')$$

In particular the effect of the operator symbolizing property A,

$$A = \sum a' \Psi(a')\, \Phi(a'),$$

upon the vector of the a coordinate system is

$$A\Psi(a') = a'\Psi(a'), \quad \Phi(a')A = \Phi(a')a',$$

which characterizes $\Psi(a')$ and $\Phi(a')$ as the right and left eigenvectors, respectively, of the complete set of commuting operators A with the eigenvalues a' . There is a dual algebra associated with each vector algebra in which all numbers are replaced by their complex conjugates .

The eigenvectors of a given description provide a basis for the representation of an arbitrary vector by N numbers . The abstract properties of vectors are realized by these sets of numbers, which are known as wave functions . We write

$$\Psi = \sum \Psi(a')\, \psi(a')$$

$$\psi(a') = \Phi(a')\Psi = \langle a'|\Psi$$

and

$$\Phi = \sum \phi(a')\, \Phi(a')$$

$$\phi(a') = \Phi\Psi(a') = \Phi|(a')\rangle$$

If Φ and Ψ are in adjoint relation, $\Phi = \Psi^{\dagger}$ the wave functions are connected by

$$\phi(a') = \psi(a')^{*} .$$

The scalar product of two vectors is

$$\underline{\Phi}_1 \underline{\Psi}_2 = \sum_{a'} \phi_1(a') \psi_2(a'),$$

and in particular

$$\underline{\Psi}^{\dagger} \underline{\Psi} = \sum \psi(a')^{*} \psi(a') \geq 0,$$

which characterizes the geometry of states as unitary . The operator $\underline{\Psi}_1, \underline{\Phi}_2$ is represented by the matrix

$$\langle a' | \underline{\Psi}_1 \underline{\Phi}_2 | b' \rangle = \psi_1(a') \phi_2(b'),$$

and wave functions that represent $\underline{\Psi}_1 = X \underline{\Psi}$ and $\underline{\Phi}_1 = \underline{\Phi} X$ are

$$\psi_1(a') = \sum \langle a' | X | b' \rangle \psi(b')$$

and

$$\phi_1(b') = \sum \phi(a') \langle a' | X | b' \rangle .$$

on placing $X = 1$, we obtain the relation between the wave function of a given vector in two different representations

$$\psi(a') = \sum \langle a' | b' \rangle \psi(b')$$

$$\phi(b') = \sum \phi(a') \langle a' | b' \rangle .$$

Note that the wave function that represents $\underline{\Psi}(b')$ in the a description is

$$\psi_{b'}(a') = \langle a' | \underline{\Psi}(b') = \langle a' | b' \rangle$$

$$= \phi_{a'}(b') .$$

From the viewpoint of the extended measurement algebra, ϕ and ψ wave

functions are matrices with but a single row and a column respectively . .

It is a convenient fiction to assert that every (Hermitian) operator symbolizes a physical quantity,and that every (unit) vector symbolizes a state . Then the expectation value of property X in the state $\underline{\Psi}$ is given by

$$\langle X \rangle_{\underline{\Psi}} = \underline{\Psi}^{\dagger} X \underline{\Psi} = \sum \psi(a')^{*} \langle a' \mid X \mid a'' \rangle \psi(a'') .$$

In particular,the probability of observing the values a' in an A-measurement performed on a system in the state $\underline{\Psi}$ is

$$p\ (a' , \underline{\Psi}) = \langle M\ (a') \rangle_{\underline{\Psi}} = \left| \psi(a') \right|^{2} .$$

The automorphisms of the unitary geometry of states are produced by the unitary transformations

$$\overline{\underline{\Phi}} = \underline{\Phi}\, U , \qquad \overline{\underline{\Psi}} = U^{-1} \underline{\Psi} , \quad \overline{X} = U^{-1} X\, U$$

applied to every vector operator,where the unitary operator U obeys

$$U^{\dagger} = U^{-1} ,$$

All algebraic relations and adjoint connections among vectors and operators are preserved by this transformation . Two successive unitary transformations form a unitary transformation and the inverse of a unitary transformation is unitary - unitary transformations form a group . The application of a unitary transformation to the orthonormal basis vectors of the a description,which are characterized by the eigenvector equation

$$\underline{\Phi}\ (a')\ (A - a') = 0 ,$$

yields the orthonormal vectors

$$\underline{\Phi}\,(\bar{a}') = \underline{\Phi}(a')\;U$$

that obey the eigenvector equation

$$\underline{\Phi}\,(\bar{a}')\;(\bar{A} - a') = 0\,.$$

Hence the $\underline{\Phi}(\bar{a}')$ are the states of a new description associated with quantities $\bar{A}$ that possess the same eigenvalue spectrum as the properties A . Since all relations among operators and vectors are preserved by the transformation , we have

$$\langle \bar{a}' | \bar{X} | \bar{a}'' \rangle = \langle a' | X | a'' \rangle$$

$$\langle \bar{a}' | \bar{\Psi} = \langle a' | \Psi \quad , \quad \bar{\Phi} | \bar{a}' \rangle = \Phi | a' \rangle$$

The equivalent forms

$$\left.\begin{aligned} \langle \bar{a}' | X | \bar{a}'' \rangle &= \langle a' | U \;\; X \;\; U^{-1} | a'' \rangle \\ \langle \bar{a}' | \Psi &= \langle a' | U \Psi \quad , \quad \Phi | \bar{a}' \rangle = \Phi\, U^{-1} | a' \rangle \quad , \end{aligned}\right\} \quad (2.9)$$

exhibit the $\bar{a}$ representatives of operators and vectors as the a representatives of associated operators and vectors .

The basis vectors of any two given descriptions, with each set placed in a definite order, are connected by a unitary operator . Thus

$$\Psi\,(a^i) = U_{ab}\;\Psi(b^i) \qquad i = 1, \ldots N$$

$$\Phi\,(b^i) = \Phi\,(a^i)\;U_{ab}$$

where

$$U_{ab} = \sum_{i=1}^{N} \Psi\,(a^i)\;\Phi(b^i)$$

obeys

$$U_{ab}^{\dagger} = U_{ba} = U_{ab}^{-1} .$$

An elementary example is the numerical phase transformation

$$\underline{\Psi}(b^i) = e^{-i\varphi(a^i)}\underline{\Psi}(a^i)$$

where

$$U_{ab} = \sum \underline{\Psi}(a^i)\; e^{i\varphi(a^i)}\,\underline{\Phi}(a^i)$$
$$= e^{i\varphi(A)} .$$

The transformation function connecting the a and b representations can be exhibited as a matrix referring to the a or b representations,

$$\langle a^i \mid b^j \rangle = \langle a^i \mid U_{ba} \mid a^j \rangle = \langle b^i \mid U_{ba} \mid b^j \rangle ,$$

and all quantities of the b representation can be expressed as a representatives of associated operators and vectors ,

$$\langle b^i \mid X \mid b^j \rangle = \langle a^i \mid U_{ab}\; X\; U_{ba} \mid a^j \rangle$$

$$\langle b^i \mid \underline{\Psi} = \langle a^i \mid U_{ab}\,\underline{\Psi} , \quad \underline{\Phi}\,|b^i\rangle = \underline{\Phi}\,U_{ba}\,|a^i\rangle .$$

If the two sets of compatible properties A and B possess the same spectra of values, the operators A and B are connected by a unitary transformation . With the ordering of basis vectors established by corresponding eigenvalues we have

$$B = \sum b^i\,\underline{\Psi}(b^i)\;\underline{\Phi}(b^i) = \sum a^i\; U_{ba}\,\underline{\Psi}(a^i)\,\underline{\Phi}(a^i)\,U_{ab}$$
$$= U_{ba}\; A\; U_{ab} .$$

The definition of a unitary operator, expressed as

$$(U-1)^\dagger (U-1) + U - 1 + (U-1)^\dagger = 0$$

shows that a unitary operator differing infinitesimally from unity has the general form

$$U = 1 + i G \ , \quad U^\dagger = U^{-1} = 1 - i G$$

where G is an infinitesimal Hermitian operator . The coordinate system transformation described by this operator is indicated by

$$\delta_a \underline{\Psi}(a') = \underline{\Psi}(\bar{a}') - \underline{\Psi}(a') = -i G \underline{\Psi}(a')$$
$$\delta_a \underline{\Phi}(a') = \underline{\Phi}(\bar{a}') - \underline{\Phi}(a') = \underline{\Phi}(a')\, i G$$

Now according to (2.9), the change of coordinate system, in its effect upon the representatives of operators and vectors, is equivalent to a corresponding change of the operators and vectors relative to the original coordinate system . Hence

$$\delta_a \langle a' | X | a'' \rangle = \langle \bar{a}' | X | \bar{a}'' \rangle - \langle a' | X | a'' \rangle$$
$$= \langle a' | X | a'' \rangle$$

and

$$\delta_a \langle a' | \underline{\Psi} = \langle a' | \delta \underline{\Psi}, \qquad \delta_a \underline{\Phi} | a'' \rangle = \delta \underline{\Phi} | a'' \rangle,$$

where

$$\delta X = U X U^{-1} - X = \frac{1}{i} [X, G] \qquad (2.10)$$

and

$$\delta \underline{\Psi} = (U - 1) \underline{\Psi} = i G \underline{\Psi}$$
$$\delta \underline{\Phi} = \underline{\Phi}(U^{-1} - 1) = - \underline{\Phi} i G .$$

The rectangular bracket represents the commutator

$$\left[A, B \right] = A\,B - B\,A \ .$$

Since all algebraic relations are preserved, the operator and vector variations are governed by rules of the type

$$\delta (X\ Y) = \delta X\ Y + X\ \delta Y$$

$$\delta (X \bar{\psi}) = \delta X \bar{\psi} + X\ \delta \bar{\psi}$$

One must distinguish between $X + \delta X$ and

$$\bar{X} = U^{-1} X U = X - \delta X \ ;$$

the latter is the operator that exhibits the same properties relative to the $\bar{a}$ description that X possesses in the a description . Thus the basis vectors $\bar{\Phi} (\bar{a}')$ are the eigenvectors of $A - \delta A$ with the eigenvalues a' .

In the successive application of unitary transformations it must be observed that the operator symbolizing a given transformation depends upon the coordinate system to which it is applied . Thus , let U_1 and U_2 be the operators describing two different transformations on the same coordinate system . When the first transformation has been applied, the operator that symbolizes the second transformation, in its effect upon the coordinate system that has resulted from the unitary transformation, is $U_1^{-1}\ U_2\ U_1$. Hence the operator that produces the complete transformation is

$$U_1\ (U_1^{-1}\ U_2\ U_1) = U_2\ U_1 \ .$$

The same form, with the operators of successive transformations multiplied from right to left, applies to any number of transformations . In particular, if one follows two transformations by their inverses, applied in the contrary sense, the unitary operator for the resulting transformation is :

$$U_{[12]} = U_2^{-1}\ U_1^{-1}\ U_2\ U_1 = (U_1\ U_2)^{-1}\ U_2\ U_1 , \qquad (2.11)$$

When both transformation are infinitesimal,

$$U_{[1,\ 2]} = 1 + i\ G_{[1,2]} ,$$

the transformation

$$U_{[1\ 2]} = 1 + i\ G_{[1\ 2]}$$

is infinitesimal to the second order ,

$$G_{[1\ 2]} = \frac{1}{i}\left[G_1 , G_2\right] \qquad (2.12)$$

The change produced in an operator is

$$\delta_{[1\ 2]}\ X = \delta_2\ \delta_1\ X - \delta_1\ \delta_2\ X$$

which, expressed in terms of commutators,implés the operator identity(Jacobi)

$$\left[\left[X,\ G_1\right] ,\ G_2\right] - \left[\left[X,\ G_2\right] ,\ G_1\right] = \left[\left[X,\ G_1\right],\ G_2\right]$$

The continued repetition of an infinitesimal unitary transformation generates a finite unitary transformation . On writing the infinitesimal Hermitian operator G , the generator of the unitary transformation, as $\delta\tau\ \ G_{(1)}$, we find that the application of the infinitesimal transformation a number of times expressed by $\tau/\delta\tau$ yields, in the limit $\delta\tau \to 0$,

$$U(\tau) = \lim_{\delta\tau \to 0} (1 + i\delta\tau\, G_{(1)})^{\frac{\tau}{\delta\tau}} = e^{i\tau G_{(1)}}$$

These operators form a one-parameter continuous group of unitary transformation,

$$U(\tau_1)\, U(\tau_2) = U(\tau_1 + \tau_2) ,$$

$$U(\tau)^{-1} = U(-\tau) , \quad U(0) = 1 .$$

A number of Hermitian operators $G_{(1)}$, ... $G_{(k)}$ generate a k-parameter continuous group of unitary transformations if these operators form a linear basis for an algebra that is closed under the unitary transformations of the group . This requires that all commutators $\left[G_{(i)} , G_{(j)} \right]$ be linear functions of the generating operators .

There is a fundamental continuous group of unitary transformations based upon the significance of measurements as physical operations in three-dimensional space . A measurement apparatus defines a spatial coordinate system with respect to which physical properties are specified . We express the uniformity of space by asserting that two coordinate systems, differing only in location or orientation, are intrinsically equivalent . In particular, physical quantities that are analogously defined with respect to different coordinate systems exhibit the same spectra of possible values . Hence the associated operators must be related by a unitary transformation . Since the totality of translations and rotations of a coordinate system form a (six-parameter) group, we infer the existence of isomorphic group of unitary transformations . An infinitesimal change of coordinate vector r in the initial system is assigned the coordinate vector $r - \delta r$ in

the new system , where

$$\delta r = \delta\epsilon + \delta\omega \times r$$

The infinitesimal generator of the corresponding unitary transformation is written

$$G_r = \delta\epsilon \cdot P + \delta\omega \cdot J$$

$$= \sum_i \delta\epsilon_i P_i + \tfrac{1}{2} \sum_{i,j} \delta\omega_{ij} J_{ij} , \delta\omega_{ij} = -\delta\omega_{ji}$$

where the Hermitian operators P and J , as the generators of infinitesimal translations and rotations, respectively, are identified with the operators of total linear and angular momentum . On comparing the two ways in which a pair of infinitesimal coordinate changes can be performed, in the manner of (2.II) ,. we find that

$$\delta_{[1\,2]}\epsilon = \delta\omega_1 \times \delta\epsilon_2 - \delta\omega_2 \times \delta\epsilon_1$$

$$\delta_{[1\,2]}\omega = \delta\omega_1 \times \delta\omega_2$$

which requires that the associated infinitesimal generators obey the commutation relations (2.I2)

$$\frac{1}{i}\left[\delta\epsilon_1 \cdot P + \delta\omega_1 \cdot J, \delta\epsilon_2 \cdot P + \delta\omega_2 \cdot J\right]$$

$$= (\delta\omega_1 \times \delta\epsilon_2 - \delta\omega_2 \times \delta\epsilon_2) \cdot P + \delta\omega_1 \times \delta\omega_2 \cdot J \quad (2.I3)$$

Hence

$$\frac{1}{i} \cdot \left[P_j, P_k\right] = 0$$

$$\frac{1}{i} \cdot \left[\bar{P}_j, J_{kl} \right] = \delta_{jl} P_k - \delta_{jk} P_l$$

$$\frac{1}{i} \cdot \left[\bar{J}_{jk}, J_{lm} \right] = \delta_{km} J_{jl} - \delta_{jm} J_{kl} + \delta_{jl} J_{km} - \delta_{kl} J_{jm} .$$

The continuous group of transformations among equivalent coordinate systems can be enlarged by the operation of reflecting the positive sense of each coordinate axis . We associate with this change of description the unitary reflection operator R , which obeys

$$R^2 = 1 .$$

A reflection, followed by the infinitesimal displacement $(\delta\epsilon, \delta\omega)$ is equivalent to first performing the displacement $(-\delta\epsilon, \delta\omega)$ and then the reflection . Accordingly,

$$R P = - P R, \qquad R J = J R$$

or

$$R P_j R^{-1} = - P_j , \quad R J_{jk} R^{-1} = J_{jk} .$$

We must remark upon the extension of our considerations to systems with infinite numbers of states and properties possessing a continuous spectrum of possible values . In any measurement of such a property, systems within a certain range of values are selected, and the concept of state will now refer to the specification of a complete set of compatible quantities within arbitrarily small intervals about prescribed values . We symbolize such states by $\bar{\Psi}_{\Delta}(a')$ and state their

completeness in

$$\sum_{a'} \underline{\Psi}_{\Delta}(a') \underline{\Phi}_{\Delta}(a') = 1$$

The change of normalization

$$\underline{\Psi}_{\Delta}(a') = (\Delta a')^{\frac{1}{2}} \underline{\Psi}(a')$$

in which $\Delta a'$ is the product of eigenvalue intervals for each continuous property (we assume all members of set A to be continuous for simplicity) now yields

$$\int \underline{\Psi}(a')\, d\, a'\, \underline{\Phi}(a') = 1$$

in the limit $\Delta a' \longrightarrow 0$. The value of the scalar product

$$\underline{\Phi}(a')\, \underline{\Psi}(a'') = \langle a' \mid a'' \rangle$$

is implied by the linear independence of the continuum vectors $\underline{\Psi}(a')$ and their completeness . Thus

$$\underline{\Psi}(a'') = \int \underline{\Psi}(a')\, d\, a' \langle a' \mid a'' \rangle$$

yields

$$\langle a' \mid a'' \rangle = \delta(a' - a'')$$

where the delta function obeys

$$f(a'') = \int f(a')\, d\, a'\, \delta(a' - a'')$$

for arbitrary functions f . This is the continuum analogue of the equation

$$f(a'') = \sum_{a'} f(a')\, \delta(a', a'')$$

and generally, in all formal relations referring to continuous spectra,

integrals replace summations . In particular, the probability that a measurement on the state $\underline{\psi}$ will yield a system in one of a group of states appears as $\int d\ a' \ |\psi(a')|^2$ so that

$$d\ a'\ p(a', \underline{\psi}) = d\ a'\ |\psi(a')|^2$$

is the probability of encountering a system with properties in the infinitesimal range $d\ a'$ about a' .

3 . - THE DYNAMICAL PRINCIPLE

A measurement is a physical operation in space and in time . The properties of a system are described in relation to measurements at a given time and no value of the time can be distinguished from another by measurements performed on an isolated physical system . Hence the operators symbolizing analogous properties at different times must be related by a unitary transformation . The propagation in time of the disturbance produced by a measurement implies that physical quantities at different times are incompatible, in general . Accordingly, complete sets of compatible properties will refer to a common time, and the characterization of a state requires the specification of the values of these quantities (together with a spatial coordinate system) and of the time . The transformation function relating to arbitrary descriptions thus

appears as $\langle a'_1\ t_1 \mid a''_2\ t_2 \rangle$ special cases of which are $\langle a'\ t \mid b'\ t \rangle$, connecting two different sets of compatible quantities at the same time, and $\langle a'\ t_1 \mid a''\ t_2 \rangle$, which relates analogous properties at different times. The connection between states at two different times involve the entire dynamical history of the system in the interval . Hence the properties of specific systems must be contained completely in a dynamical principle that characterizes the general transformation function .

Any infinitesimal alteration of the transformation function $\langle a'_1\ t_1 \mid a'_2\ t_2 \rangle$ can be expressed as

$$\delta \langle a'_1\ t_1 \mid a'_2\ t_2 \rangle = i \langle a'_1\ t_1 \mid \delta W_{1\,2} \mid a'_2\ t_2 \rangle ,$$

where $\delta W_{1\,2}$ is an infinitesimal Hermitian operator with the additivity property

$$\delta W_{1\,2} + \delta W_{2\,3} = \delta W_{1\,3}$$

for consecutive transformations . We now state our fundamental dynamical postulate : There exists a class of infinitesimal alterations for which the associated operators $\delta W_{1\,2}$ are obtained by appropriate variations of a single operator, the action operator $W_{1\,2}$,

$$\delta W_{1\,2} = \delta \left[W_{1\,2} \right] .$$

It is consistent with the properties of the infinitesimal operators to assert that the action operator is Hermitian

$$W_{1\,2}{}^{+} = W_{1\,2} ,$$

and has the additive combinatorial property

$$W_{1\,2} + W_{2\,3} = W_{1\,3} .$$

If one views the transformation from the $a_2\ t_2$ description to the $a_1\ t_1$

description as occurring continuously in time through an infinite succession of infinitesimally differing descriptions, the additivity property asserts that

$$W_{1\,2} = \sum_{t_2}^{t_1} W_{t + dt,\, t'}$$

where

$$W_{t,\, t} = 0$$

since $\langle a'\ t \,|\, a''\ t \rangle$ has fixed numerical values . Hence the action operator has the general form

$$W_{1\,2} = \int_{t_2}^{t_1} dt\ L(t),$$

in which $L(t)$, the Lagrangian operator, is a Hermitian function of fundamental dynamical variables $x_a(t)$ in the infinitesimal neighborhood of time t ,and of the time . There is no loss of generality in taking the dynamical variables $x_a(t)$ to be Hermitian . For our present purposes, we suppose the number of such variables to be finite, which is appropriate for the non-relativistic description of a fixed number of particles . The objects of possible variation in the action operator are the terminal times t_1 and t_2, the dynamical variables $x_a(t)$, and the structure of the Lagrangian operator .

For a given dynamical system, the relation between analogous states at different times is completely determined . Hence any alteration in the transformation function must be associated with a change of the states, including the times to which they refer .Such infinitesimal eigenvector transformations are generated by Hermitian operators, that can depend only

upon dynamical variables at the stated time,

$$\delta \Phi (a'_1 t_1) = i \Phi (a'_1 t_1) G_1$$

$$\delta \Psi (a'_2 t_2) = - i G_2 \Psi (a'_2 t_2)$$

and therefore

$$\delta \langle a'_1 t_1 | a'_2 t_2 \rangle = i \langle a'_1 t_1 | G_1 - G_2 | a'_2 t_2 \rangle \quad (3.1)$$

Thus, with a given dynamical system ,

$$\delta W_{1\,2} = \delta \left[\int_{t_2}^{t_1} d t \; L(t) \right] = G_1 - G_2 ,$$

which is the operator principle of stationary action, for it asserts that the variation of the action integral involves only dynamical variables at the terminal times . This principle implies equations of motion for the dynamical variables and yields specific forms for generators of infinitesimal unitary transformations . Note that the Lagrangian operator cannot be determined completely by the dynamical nature of the system, since we must be able to produce a variety of infinitesimal transformations for a given system . Indeed, if two Lagrangians differ by a time derivative,

$$\bar{L} = L - \frac{d}{dt} W , \quad (3.2)$$

the action operators are related by

$$\bar{W}_{1\,2} = W_{1\,2} - (W_1 - W_2) ,$$

and

$$\delta \bar{W}_{1\,2} = \delta W_{1\,2} - \delta W_1 + \delta W_2$$

$$= \bar{G}_1 - \bar{G}_2$$

which also satisfies the stationary action requirement but implies new generators of infinitesimal transformations at times t_1 and t_2, as given by

$$\delta W_1 = G_1 - \bar{G}_1 , \quad \delta W_2 = G_2 - \bar{G}_2 \qquad (3.3)$$

The specification of a state at a given time determines the state at any other time . To be consistent with this deterministic requirement, as applied to an infinitesimal time interval , the equations of motion for the dynamical variables must be differential equations of finite order . By suitable adjunction of variables, such equations of motion can be presented as first order differential equations . A general form of Lagrangian operator that yields explicit first order equations of motion is

$$L = \tfrac{1}{2} \sum_{a,b} \left(x_a A_{ab} \frac{dx_b}{dt} - \frac{dx_a}{dt} A_{ab} x_b \right) - H (x_a, t)$$

$$= \tfrac{1}{2} \left(x A \frac{dx}{dt} - \frac{dx}{dt} A x \right) - H (x, t) \qquad (3.4)$$

in which A_{ab} is a numerical matrix . We shall speak of the two parts of this Lagrangian operator as the kinematical and the dynamical part, respectively . Note that the kinematical part has been symmetrized with respect to transference of the time derivative . The Hermitian requirement on L applies to the two parts separately . Hence H , the Hamiltonian operator , must be Hermitian , and the matrix A must be skew- Hermitian ,

$$A^{T *} = - A .$$

For an isolated dynamical system, there is no explicit dependence upon time in H .

The action operator implied by the Lagrangian operator (3.4) is

$$W_{1\,2} = \int_{t_2}^{t_1} \left[\tfrac{1}{2} (x A\, dx - dx A\, x) - H\, dt \right].$$

In this form, the limits of integration are objects of variation . However, we can think of introducing an auxiliary variable τ , $t = t(\tau)$, and produce the variations δt_1, δt_2 by altering the functional dependence of t upon τ, with fixed limits τ_1, τ_2 . This procedure places the time on somewhat the same footing as the dynamical variables x_a . Now

$$\delta \left[\tfrac{1}{2} (x A\, dx - dx A\, x) - H\, dt \right] - d \left[\tfrac{1}{2} (x A \delta x - \delta x A\, x) - H \delta t \right]$$
$$= \delta x A\, dx - dx A \delta x - \delta H\, dt + dH\, \delta t ,$$

and thus the stationary action principle asserts that

$$\delta x A\, dx - dx A \delta x = \delta H\, dt - dH \delta t ,$$

or

$$\delta H = \frac{dH}{dt} \delta t + \delta x A \frac{dx}{dt} - \frac{dx}{dt} A \delta x , \qquad (3.5)$$

and yields the infinitesimal generators

$$G = \tfrac{1}{2} (x A \delta x - \delta x A\, x) - H \delta t = G_x + G_t$$

at the terminal times .

The structure of the Hamiltonian operator is as yet unspecified . If its variation is to possess the form (3.5), with the δx_a appearing only on the left and on the right, these variations must possess elementary operators properties , characterizing the class of variations to which the

dynamical principle refers . Thus, we should be able to displace δx_a entirely to the left , or to the right , in the structure of δH, which defines the left and the right derivatives of H ,

$$\delta H - \frac{\partial H}{\partial t} \delta t = \delta x \frac{\partial_l H}{\partial x} = \frac{\partial_r H}{\partial x} \delta x$$

In view of the complete symmetry between left and right in tne multiplication process , we infer that the terms in (3.5) with δx on the left, and on the right, are identical . We thus obtain

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \qquad (3.6)$$

and

$$\left.\begin{aligned} 2A \frac{dx}{dt} &= \frac{\partial_l H}{\partial x} \quad , \\ -\frac{dx}{dt} 2A &= -2A^T \frac{dx}{dt} = \frac{\partial_r H}{\partial x} \end{aligned}\right\} \qquad (3.7)$$

where the latter must be equivalent forms of the equations of motion . Similarly, the infinitesimal generator G_x should possess the equivalent forms

$$G_x = -\delta x A x$$

and

$$G_x = x A \delta x = (A^T x) \delta x .$$

If the matrix A is completely reducible into two submatrices , we speak of the two corresponding sets of variables as being kinematically independent . Now the skew-Hermitian matrix A can always be written as the sum of an anti-symmetrical real matrix a , and of a symmetrical imaginary matrix s . In view of the manner in which A and A^T enter

in equivalent versions of the equations of motion and of G_x we conclude that the construction of a consistent theory requires that the expression

$$A = a + s$$

$$a = a^{\dagger} = -a^{T} \quad , \quad s = s^{T} = -s^{\dagger}$$

be a complete reduction into two submatrices . Accordingly there exists two kinematically independent types of variables : variables of the first kind , z_i , associated with a , and variables of the second kind , ζ_α , associated with s . The Lagrangian operator can now be written as

$$L = \tfrac{1}{2} \left\{ z, a \frac{dz}{dt} \right\} + \tfrac{1}{2} \left\{ \zeta , s \frac{d\zeta}{dt} \right\} - H (z, \zeta , t) ,$$

where the curly bracket symbolizing the symmetrical product or anti-commutator,

$$\left\{ A, B \right\} = A B + B A .$$

The equations of motion appear in two sets

$$\left.\begin{aligned} 2 a \frac{dz}{dt} &= \frac{\partial_l H}{\partial z} = \frac{\partial_r H}{\partial z} \quad , \\ 2 s \frac{d\zeta}{dt} &= \frac{\partial_l H}{\partial \zeta} = \frac{\partial_r H}{\partial \zeta} \quad , \end{aligned}\right\} \qquad (3.8)$$

and

$$G_x = G_z + G_\zeta \quad ,$$

with

$$G_z = -\delta z \; a \; z = -(a z) \delta z$$

$$G_\zeta = -\delta \zeta \; s \zeta = (s \zeta) \delta \zeta$$

The identity of left and right derivatives of H for variables of the first kind, and the equivalent forms of G_z , indicate that the δz_i must commute with all operators ,

$$\left[\delta z_i , z_j\right] = \left[\delta z_i , \xi_\alpha\right] = 0 .$$

The two versions of G_ξ show that the $\delta \xi_\alpha$ must anticommute with the variables of the second kind ,

$$\left\{\delta \xi_\alpha , \xi_\beta\right\} = 0 .$$

The opposite sign of left and right derivatives of H for variables of the second kind then requires that H be an even function of the ξ_α , but it remains an arbitrary function of the z_i , provided that

$$\left[\delta \xi_\alpha , z_i\right] = 0 .$$

In these operator properties of δz_i and $\delta \xi_\alpha$ we have the precise specification of the class of variations to which the fundamental dynamical principle refers .

For the purposes of the present discussion , we shall assume that (3. 7) are true equations of motion, which can be solved for each of the dx_i / dt . This requires that the matrix A be non singular, and the latter restriction applies in turn to the submatrices a and s . If the determinant of the antisymmetrical matrix a is not to vanish identically , the number of dynamical variables of the first kind must be even ($2 n_1$) . By means of a real linear transformation, the antisymmetrical real matrix 2 a can be brought into a canonical diagonal form in which

the n_1 diagonal elements are the two-dimensional matrices

$$\begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix}.$$

With this transformation the dynamical variables are partitioned into n_1 kinematically independent pairs . If the Hermitian dynamical variables associated with the k^{th} two dimensional matrix are called p_k and q_k , the kinematical part of the Lagrangian referring to variables of the first kind reads

$$\tfrac{1}{2} \cdot \sum_{k=1}^{n_1} \left(\tfrac{1}{2} \left\{ p_k, \frac{dq_k}{dt} \right\} - \tfrac{1}{2} \left\{ \frac{dp_k}{dt} , q_k \right\} \right) ,$$

the equations of motion exhibit the canonical form

$$\frac{dq_k}{dt} = \frac{\partial H}{\partial p_k} , \qquad - \frac{dp_k}{dt} = \frac{\partial H}{\partial q_k} ,$$

and the infinitesimal generator G_z becomes

$$G_{q,p} = \tfrac{1}{2} \sum_{k=1}^{n_1} (p_k \, \delta q_k - \delta p_k \, q_k) \qquad (3.9)$$

in which δq_k and δp_k commute with all operators . Alternative forms are obtained by adding suitable time derivatives to L with a corresponding alteration of the infinitesimal generator in the manner of (3.2) and (3.3) . Thus, with

$$W = - \tfrac{1}{2} \sum \tfrac{1}{2} \left\{ p_k , q_k \right\}$$

the kinematical term in L becomes

$$\sum \tfrac{1}{2} \cdot \left\{ p_k , \frac{dq_k}{dt} \right\}$$

and the associated generator is

$$G_q = \sum_{k=1}^{n_1} p_k \, \delta q_k$$

On reversing the sign of W these quantities become

$$-\sum \tfrac{1}{2} \left\{ \frac{dp_k}{dt} , q_k \right\}$$

and

$$G_p = -\sum_{k=1}^{n_1} \delta p_k \, q_k$$

The infinitesimal Hermitian operator G_q evidently generates the infinitesimal unitary transformation in which, at a given time , the operators q_k are independently subjected to the variation δq_k , with no change of the variables p_k , while G_p has the converse interpretation. Hence, if F is any function of q_k and p_k , the canonical variables of the first kind , at the stated time , we infer from (2.10) that

$$\frac{1}{i} \left[F , G_q \right] = \sum_k \frac{\partial F}{\partial q_k} \delta q_k$$

$$\frac{1}{i} \left[F , G_p \right] = \sum_k \delta p_k \frac{\partial F}{\partial p_k}$$

Now δq_k and δp_k commute with all quantities and are otherwise arbitrary .

Therefore

$$1/i \left[F, \; p_k \right] = \frac{\partial F}{\partial q_k} \quad , \qquad 1/i \left[q_k \, , \; F \right] = \frac{\partial F}{\partial p_k} \quad , \qquad (3.10)$$

which , in particular , yields the commutation properties of the canonical variables of the first kind ,

$$\left[q_k \, , \; q_l \right] = \left[p_k \, , \; p_l \right] = 0, \qquad \left[q_k \, , \; p_l \right] = i \cdot \delta_{kl} \, .$$

The infinitesimal generator (3 . 9) ,

$$G_{q,p} = \tfrac{1}{2} \, G_q + \tfrac{1}{2} \, G_p$$

clearly produces the transformation in which q_k and p_k are changed by $\frac{1}{2} \delta q_k$ and $\frac{1}{2} \delta p_k$, respectively . Since this refers symmetrically to both sets of canonical variables the same interpretation applies to the equivalent generator G_z , that it describes variations of z by $\frac{1}{2} \delta z$. Hence

$$\frac{1}{i} \left[F, \; G_z \right] = \tfrac{1}{2} \delta z \, \frac{\partial F}{\partial z} \quad ,$$

or

$$\frac{1}{i} \left[2a \, z, \; F \right] = \frac{\partial F}{\partial z} \quad , \qquad (3.11)$$

from which we derive the commutation relations for the $2 n_1$ dynamical variables of the first kind without specialization to the canonical form ,

$$\left[z_k \, , \; z_l \right] = i \tfrac{1}{2} \, (a^{-1})_{kl} \, .$$

What is fundamental in the canonical version is the possibility of dividing the $2 n_1$ dynamical variables into two sets of n_1 variables

that are complementary in the sense that each set comprises the generators of infinitesimal variations of the other set . This possibility is more generally expressed by the existence of transformations that bring the real, antisymmetrical, non-singular $2n_1$-dimensional matrix a into the partitioned form

$$a = \frac{1}{2}\begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix},$$

in which the real n_1-dimensional matrix α is symmetrical and positive definite . If the dynamical variables are separated correspondingly into the two sets of n_1 variables $z_k^{(1)}$, $z_k^{(2)}$, the kinematical term in the Lagrangian acquires the bilinear form

$$\frac{1}{2}\left\{ z^{(1)} , \alpha \frac{d z^{(2)}}{dt} \right\} , \qquad (3.12)$$

to within a time derivative . The ensuing generators and commutation relations are

$$G_z(2) = z^{(1)} \alpha \delta z^{(2)} \quad , \quad G_z(1) = -\delta z^{(1)} \alpha z^{(2)}$$

and

$$\left[z_k^{(1)} , z_l^{(1)} \right] = \left[z_k^{(2)} , z_l^{(2)} \right] = 0,$$

$$\left[z_k^{(1)} , z_l^{(2)} \right] = -i \, (\alpha^{-1})_{kl} . \qquad (3.13)$$

The explicit canonical form then emerges on subjecting both sets of variables to a common real linear transformation that reduces α to the unit matrix. We now observe that the validity of the bilinear structure (3.12) does not depend upon $z_k^{(1)}$ and $z_k^{(2)}$ being Hermitian operators . Thus, to within an additive time derivative (3.12) can be replaced by

$$\tfrac{1}{2}\left\{ y^{+},\ i\alpha\frac{dy}{dt}\right\}, \qquad (3.14)$$

where

$$y = 2^{-\frac{1}{2}}\left(z^{(1)} - i\, z^{(2)}\right), \qquad (3.15)$$

and the form (3.14) will persist under arbitrary complex linear transformation of the y variables if α is described as a positive-definite Hermitian matrix . The commutation relations for the non-Hermitian variables , deduced from the infinitesimal generators ,

$$G_y = i\, y^{+}\alpha\, \delta y, \qquad G_{y^{+}} = -i\, \delta y^{+}\, \alpha\, y ,$$

are

$$\left[y_k , y_l\right] = \left[y_k^{+} , y_l^{+}\right] = 0 , \quad \left[y_k ,\ y_l^{+}\right] = (\alpha^{-1})_{kl} ,$$

which agree with those derived directly from (3.13) and (3.15) . The canonical form obtained by reducing α to the unit matrix is identical with that for Hermitian dynamical variables, except that the non-Hermitian canonical variables q_k and p_k are related by

$$p_k = i\, q_k^{+} .$$

Hence the canonical equations of motion become

$$i\frac{dq_k}{dt} = \frac{\partial H}{\partial q_k^{+}}$$

together with the adjoint equation .

The treatment of the variables of the second kind is simplified by anticipating that a consistent theory can be constructed only if the real , symmetrical matrix $-i\,s$ is positive-definite . Then one can

introduce new Hermitian variables ξ_α by a real linear transformation that reduces $-2is$ to the unit matrix. The generator of infinitesimal changes in these variables reads

$$G_\xi = \frac{i}{2} \cdot \sum \xi_\alpha \delta\xi_\alpha ,$$

which shows that the various ξ_α are kinematically independent. The operator expression of this statement is

$$[\xi_\alpha , \xi_\beta \delta\xi_\beta] = 0, \qquad \alpha \neq \beta$$

or

$$\{\xi_\alpha , \xi_\beta\} = 0 \qquad \alpha \neq \beta \qquad (3.16)$$

since $\delta\xi_\beta$ anticommutes with all variables of the second kind. We shall conclude from this anticommutativity of the various ξ_α, together with their kinematical independence, that the dynamical variables of the second kind are necessarily even in number ($2n_2$). If this number were odd, the product of all the variables would commute with every ξ_α. Hence the product must be a multiple of the unit operator, which contradicts the independence of the various ξ_α.

The division of the $2n_2$ dynamical variables of the second kind into two complementary sets of n_2 variables can be accomplished only by introducing non-Hermitian variables. First let us bring the $2n_2$-dimensional matrix S into the partioned form,

$$s = \frac{i}{2} \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

where the real n_2-dimensional matrix a is symmetrical and positive-

definite . If the correspondingly partioned variables are called $\zeta_\alpha^{(1)}$, $\zeta_\alpha^{(2)}$, the kinematical term in L becomes

$$\tfrac{1}{2}\left[\zeta^{(1)} , \tfrac{i}{2} a \frac{d\zeta^{(1)}}{dt}\right] + \tfrac{1}{2}\left[\zeta^{(2)} , \tfrac{i}{2} a \frac{d\zeta^{(2)}}{dt}\right],$$

which , to within an additive time derivative, can be replaced by the bilinear form

$$\tfrac{1}{2}\left[\eta^{\dagger} , ia\frac{d\eta}{dt}\right]$$

where

$$\eta = 2^{-\frac{1}{2}} (\zeta^{(1)} - i\zeta^{(2)}) \quad .$$

On reducing a to the unit matrix we obtain the canonical version ,

$$\sum_{\alpha=1}^{n_2} \tfrac{1}{2}\left[P_\alpha , \frac{dq_\alpha}{dt}\right],$$

with

$$P_\alpha = i q_\alpha^{\dagger}$$

The action principle now yields the canonical form of the equations of motion,

$$\frac{dq_\alpha}{dt} = \frac{\partial_l H}{\partial P_\alpha} , - \frac{dp_\alpha}{dt} = \frac{\partial_r H}{\partial q_\alpha}$$

or equivalently ,

$$i \frac{dq_\alpha}{dt} = \frac{\partial_l H}{\partial q_\alpha^{\dagger}}$$

and the adjoint equations ; while the generators of infinitesimal changes in the q_α or p_α are

$$G_q = \sum_{\alpha=1}^{n_2} P_\alpha \delta q_\alpha \quad , G_p = - \sum_{\alpha=1}^{m_2} \delta P_\alpha q_\alpha \, .$$

The accompanying commutation properties with an arbitrary function of the dynamical variables at a given time are

$$\frac{1}{i}\left[F\,,\,G_q\right]=\sum\frac{\partial_r F}{\partial q_\alpha}\,\delta q_\alpha$$

$$\frac{1}{i}\left[F\,,\,G_p\right]=\sum\delta p_\alpha\,\frac{\partial_\ell F}{\partial p_\alpha}\quad .$$

Since the δq_α and δp_α anticommute with all variables of the second kind, we distinguish between the classes of operators F_e and F_o, that are even and odd functions of the variables of the second kind. Then

$$\frac{1}{i}\left[F_e,\,p_\alpha\right]=\frac{\partial_r F_e}{\partial q_\alpha}=-\frac{\partial_\ell F_e}{\partial q_\alpha}$$

$$\frac{1}{i}\left[q_\alpha,\,F_e\right]=\frac{\partial_\ell F_e}{\partial p_\alpha}=-\frac{\partial_r F_e}{\partial p_\alpha}\qquad (3.17)$$

and

$$\frac{1}{i}\left\{f_o\,,\,p_\alpha\right\}=\frac{\partial_r F_o}{\partial q_\alpha}=\frac{\partial_\ell F_o}{\partial q_\alpha}$$

$$\frac{1}{i}\left\{q_\alpha\,,\,F_o\right\}=\frac{\partial_\ell F_o}{\partial p_\alpha}=\frac{\partial_r F_o}{\partial p_\alpha}\quad .$$

The application of the latter set to the variables q_α and p_α yields

$$\left\{q_\alpha\,,\,q_\beta\right\}=\left\{p_\alpha\,,\,p_\beta\right\}=0,\quad \left\{q_\alpha\,,\,p_\beta\right\}=i\,\delta_{\alpha\beta}\,,$$

the canonical commutation relations for variables of the second kind.

An equivalent form is

$$\left\{q_\alpha\,,\,q_\beta\right\}=\left\{q_\alpha^+\,,\,q_\beta^+\right\}=0,\quad \left\{q_\alpha\,,\,q_\beta^+\right\}=\delta_{\alpha\beta}\,.$$

We also learn from (3.11) that variables of the first kind commute with

variables of the second kind .

The infinitesimal operator

$$G_{q,p} = \sum_a \tfrac{1}{2} (p_a \delta q_a - \delta p_a \, q_a) = \tfrac{1}{2} G_q + \tfrac{1}{2} G_p$$

refers symmetrically to all variables of the second kind, and generates the infinitesimal transformation in which q_a and p_a are varied by $\tfrac{1}{2} q_a$ and $\tfrac{1}{2} p_a$, respectively . Hence the operator G_ξ must be interpreted as the generator of the infinitesimal change of the variables by $\tfrac{1}{2}\delta\xi$ and

$$\frac{1}{i} \left[F , G_\xi \right] = \tfrac{1}{2} \delta\xi \frac{\partial_\ell F}{\partial \xi}$$

Accordingly ,

$$\frac{1}{i} \left[2 S \xi \, , F_e \right] = \frac{\partial_\ell F_e}{\partial \xi} = - \frac{\partial_r F_e}{\partial \xi} \qquad (3.18)$$

and

$$\frac{1}{i} \left[2 S \xi , F_o \right] = \frac{\partial_\ell F_o}{\partial \xi} = \frac{\partial_r F_o}{\partial \xi} ,$$

from which we infer the commutation properties of Hermitian variables of the second kind ,

$$\left\{ \xi_\alpha , \xi_\beta \right\} = i \tfrac{1}{2} (S^{-1})_{\alpha\beta} .$$

It is in the structure of the left side that we find the basis for the statement that a consistent development of the theory is possible only if $-iS$ is a positive-definite matrix . On transforming -2 is into the unit matrix , we obtain the commutation properties of the ξ_α variables

$$\left\{ \xi_\alpha , \xi_\beta \right\} = \delta_{\alpha\beta} ,$$

which includes (3.16) and

$$\xi_\alpha^2 = \tfrac{1}{2} \quad , \qquad \alpha = 1 \ \ldots \ 2n_2 \ .$$

In addition to the infinitesimal operator G_x that we have been discussing, the principle of stationary action yields

$$G_t = - H \delta t \ ,$$

which evidently dexcribes the unitary transformation from a description at time t to the analogous one , at time $t + \delta t$ if F is a function of the dynamical variables $x(\ell)$ and of the time t, the operator F that plays the role of F in the description referring to time $t + \delta t$ is

$$F = F - \delta F = F(x(t+\delta t), t)$$

Since the numerical parameter t is not affected by operator transformations . Then ,

$$- \delta F = \left(\frac{dF}{dt} - \frac{\partial F}{\partial t}\right) \delta t$$

and therefore

$$\frac{I}{i}[F, Gt] = -\left(\frac{dF}{dt} - \frac{\partial F}{\partial t}\right) \delta t$$

or

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{I}{i}[F, H],$$

which is the general equation of motion . With $F = H$, we obtain(3.6) and the insertion of $F = x$ and ξ , combined with the commutation relations (3.11) and (3.18) yields the equations of motion for

the fundamental dynamical variables in the form (3.8). Similarly, the insertion of the canonical variables yields the canonical version of the equations of motion. This demonstrates the consistency of the theory that has been evolved from the fundamental dynamical postulate.

4. The Special Canonical Group

The operator properties of the infinitesimal variations that appear in the action principle are consistent with the commutation relations of the dynamical variables . In particular the commutativity, or anticommutativity of the δq and δp maintains the structure of the canonical commutation relations . These elementary operator properties must also apply between two independent variations, and the totality of such infinitesimal operators G_q , G_p or G_x , together with arbitrary infinitesimal multiples of the unit operator , form the infinitesimal generators of a group , which we call the special canonical transformation group . Thus ,

$$\left[G_x^{(1)}, G_x^{(2)} \right] = i \frac{\partial G_x^{(1)}}{\partial x} \tfrac{1}{2} \delta x^{(2)} = -i \tfrac{1}{2} \delta x^{(1)} \wedge \delta x^{(2)} ,$$

or , in canonical form (for either type of variable) ,

$$\left[G_q^{(1)}, G_q^{(2)} \right] = \left[G_p^{(1)}, G_p^{(2)} \right] = 0$$

$$\left[G_p^{(1)}, G_q^{(2)} \right] = -i \sum \delta p_a^{(1)} \delta q_a^{(2)} ,$$

and the resulting infinitesimal multiples of the unit operator commute with all dynamical variables since the infinitesimal variations appear quadratically . Accordingly, this set of operators is closed with respect to the formation of commutators , which is the condition for the generation of a finite group of transformations .

Complete sets of commuting Hermitian operators for a system

described by dynamical variables of the first kind are provided by the n canonical variables q_k , or by the complementary set of n operators p_k . Thus, two elementary examples of states are symbolized by $\Phi(q't)$ and $\Phi(p't)$. We consider the significance of the changes induced in these vectors by the transformations with the generators G_q and G_p , respectively,

$$\delta_q \Phi(q't) = i\Phi(q't)G_q \quad , \quad \delta_p \Phi(p't) = i\Phi(p't)G_p$$

The vector $\Phi(q't) + \delta\Phi(q't)$ is an eigenvector of the operator $q_k - \delta q_k$ with the eigenvalues q'_k . Now δq_k commutes with all operators and is therefore an arbitrary real infinitesimal multiple of the unit operator . Hence the varied vector can be described equivalently as an eigenvector of the operators q_k with the eigenvalues $q'_k + \delta q_k$, which shows that the spectra of the q_k form a continuum extending from $-\infty$ to ∞ , and that the variation $\delta_q \Phi(q't)$ can be ascribed to a change of the eigenvalues by $\delta q'_k = \delta q_k$. A similar argument applies to the variables p_k . Accordingly

$$\delta_q \Phi(q't) = \sum_k \delta q_k \frac{\partial}{\partial q'_k} \Phi(q't)$$

$$\delta_p \Phi(p't) = \sum_k \delta p_k \frac{\partial}{\partial p'_k} \Phi(p't)$$

and therefore

$$-i\frac{\partial}{\partial q'_k}\Phi(q't) = \Phi(q't)\,p_k(t)$$

$$i\frac{\partial}{\partial p'_k}\Phi(p't) = \Phi(p't)\,q_k(t)$$

The adjoint relations are

$$i\frac{\partial}{\partial q'_k}\Psi(q't) = P_k\Psi(q't) \quad , \quad -i\frac{\partial}{\partial p'_k}\Psi(p't) = q_k\Psi(p't)$$

For functions $F(q, p)$ that can be constructed algebrically from the p_k and arbitrary functions of the q_k, we establish by induction that (using the bracket notation)

$$\langle q't|F(q,p) = F\left(q', -i\frac{\partial}{\partial q'}\right)\langle q't| \qquad (4.1)$$

and

$$F(q,p)\,|q't\rangle = F^T\left(q', i\frac{\partial}{\partial q'}\right)|q't\rangle \qquad (4.2)$$

which supplies a differential operator realization for the abstract operators. Here F^T, the transposed differential operator, is obtained from F by reversing the sense of all products. Similarly, for functions derived by algebraic construction from the q_k and arbitrary functions of the p_k,

$$\left.\begin{aligned}\langle p't|F(q,p) &= F\left(i\frac{\partial}{\partial p'}, p'\right)\langle p't| \\ F(q,p)\,|p't\rangle &= F^T\left(i\frac{\partial}{\partial p'}, p'\right)|p't\rangle\end{aligned}\right\} \qquad (4.3)$$

We find an application of (4.1) and (4.2) in the significance of the generator $G_t = -H\delta t$,

$$\delta_t\Psi(q't) = -i\Psi(q't)\,H(q,p,t)\,\delta t,$$

which implies that

$$i\frac{\partial}{\partial t}\langle q't| = \langle q't|\,H(q,p,t)$$

$$= H\left(q', -i\frac{\partial}{\partial q'}, t\right)\langle q't| ,$$

and similarly

$$-i\frac{\partial}{\partial t}|q't\rangle = H^T\left(q', i\frac{\partial}{\partial q'}, t\right)|q't\rangle ,$$

The mutually conjugate wave functions that represent a given state ,

$$\Psi(q't) = \langle q't|\Psi \quad , \quad \phi(q't) = \Phi|q't\rangle$$

therefore obey

$$i\frac{\partial}{\partial t}\Psi(q't) = H\left(q', -i\frac{\partial}{\partial q'}, t\right)\Psi(q't)$$

$$-i\frac{\partial}{\partial t}\phi(q't) = H^T\left(q', i\frac{\partial}{\partial q'}, t\right)\phi(q't) ,$$

the Schrödinger differential equations for a system characterized by dynamical variables of the first kind . That the two equations must be related by complex conjugation is consistent with the Hermitian nature of the operator H .

We must also notice the effect of the transformation generated by G_q on $\Psi(p't)$, and of G_p on $\Phi(q't)$, say . Thus

$$\delta_q\Psi(p't) = -i\,G_q\Psi(p't) = -i\left(\sum p'_k\,\delta q_k\right)\Psi(p't)$$

$$\delta_p\Phi(q't) = i\,\Phi(q't)\,G_p = -i\left(\sum \delta p_k\,q'_k\right)\Phi(q't) \qquad (4.4)$$

and the result of these transformations is to multiply the initial vectors

by numerical phase factors . Hence, in response to the operations of the special canonical group , a given eigenvector in the q or p descriptions can be transformed into one with any other set of eigenvalues , and multiplied by an arbitrary phase factor .

The construction of the transformation functions connecting the q and p descriptions at a given time also involves (4.4) . Thus , in the manner of (3.1) ,

$$\langle q't \mid p't \rangle = i \langle q't \mid G_q - G_p \mid p't \rangle$$
$$= i \left[\sum p'_k \, \delta q'_k + \sum \delta p'_k \; q'_k \right] \langle q't \mid p't \rangle ,$$

whence

$$\langle q't \mid p't \rangle = C \cdot e^{i \sum p'_k \, q'_k}$$

and

$$\langle p't \mid q't \rangle = C^* \cdot e^{-i \; p'_k \, q'_k}$$

The magnitude of the constant C is fixed by the composition property

$$\delta (q' - q'') = \int \langle q' \mid p' \rangle \, dp' \langle p' \mid q'' \rangle$$
$$= |C|^2 \int dp' \cdot e^{i \sum p'_k (q'_k - q''_k)} ,$$

which yields

$$|C| = (2\pi)^{-\frac{n}{2}} \qquad (4.6)$$

according to the integral representation of the delta function in dimensions . With a simple choice of phase for C , we find that the wavefunctions representing a given state in the q and p descriptions are related by

$$\psi(q't) = (2\pi)^{-n/2}\int e^{i\sum q'_k p'_k}\, dp'\, \psi(p't)$$

$$\psi(p't) = (2\pi)^{-n/2}\int e^{-i\sum p'_k q'_k}\, dq'\, \psi(q't)\ ,$$

together with the complex conjugate equations . Accordingly, the scalar product of two vectors has several evaluations depending upon the choice of representations , as in

$$\Phi_1 \Psi_2 = \int \phi_1(q't)\, dq'\, \psi_2(q't)$$

$$= (2\pi)^{-n/2} \int \phi_1(p't)\, dp'\, e^{-i\sum p'_k q'_k}\, dq'\, \psi_2(q't)$$

One can convert the mixed representation evaluation into a form involving differentiations rather that integrations . Thus

$$\phi_1(p')\, e^{-i\sum p'_k q'_k} = \phi_1\left(i\frac{\partial}{\partial q''}\right) e^{-i\sum p'_\kappa q'_\kappa}$$

and on transferring the derivatives to the function $\psi_2(q')$, we have

$$\Phi_1 \Psi_2 = (2\pi)^{-n/2}\int dq'\, dp'\, e^{-i\sum p'_k q'_k}\, \phi_1\left(-i\frac{\partial}{\partial q'}\right)\psi_2(q')$$

$$= (2\pi)^{n/2}\, \phi_1\left(-i\frac{\partial}{\partial q'}\right)\psi_2(q')\ \Big|_{q'=0}\ , \qquad (4.7)$$

in which the evaluation at $q' = 0$ is enforced by the delta function that is encountered on performing the p' integration . A basic example of (4.7) is

$$\langle q'|p'\rangle = (2\pi)^{n/2}\, \left\langle q'\left|-i\frac{\partial}{\partial q''}\right.\right\rangle \langle q''|p'\rangle\ \Big|_{q''=0}$$

which is a composition property for the transformation function $\langle q'|p'\rangle$.

and similarly ,

$$\langle p'|q'\rangle = (2\pi)^{-n/2} \langle p'|i\frac{\partial}{\partial p''}\rangle\langle p''|q\rangle \Big|_{p''=0} \quad (4.8)$$

In close relation with the scalar product representations are the various equivalent expressions for the trace of an operator ,

$$\text{tr } X = \int dq' \langle q'|X|q'\rangle = (2\pi)^{-n/2} \int dq'\, dp' \langle q'|X|p'\rangle \cdot e^{-i\sum p'_k q'_k}$$

$$= (2\pi)^{n/2} \langle -i\frac{\partial}{\partial p'} | X | p' \rangle \Big|_{p'=0}$$

Another derivation of (4.7) is based on the following operator construction of vector in terms of wave functions ,

$$\Psi = \int |q'\rangle dq' \psi(q') = \psi(q) \int |q'\rangle dq'$$

$$= \psi(q) |p'=0\rangle (2\pi)^{n/2}$$

and

$$\Phi = \int \phi(p') dp' \langle p'| = (2\pi)^{n/2} \langle q'=0| \phi(p)$$

Thus

$$\Phi_1 \Psi_2 = (2\pi)^n \langle q'=0| \phi_1(p) \psi_2(q) |p'=0\rangle$$

and the differential operator realization (4.1) yields

$$\Phi_1 \Psi_2 = \phi_1(-i\frac{\partial}{\partial q'}) \psi_2(q') (2\pi)^n \langle q'|p'=0\rangle \Big|_{q'=0}$$

which is identical with (4. 7) . An alternative evaluation is obtained from (4.3) ,

$$\Phi_1 \Psi_2 = \psi_2(-i\frac{\partial}{\partial p'}) \phi_1(p') (2\pi)^n \langle q'=0|p'\rangle \Big|_{p'=0}$$

$$\bar{\Phi}_1 \bar{\Psi}_2 = (2\pi)^{n/2} \phi_1(p') \psi_2\left(-i \frac{\partial^T}{\partial p'}\right)\Big|_{p'=0}$$

in which ∂^T indicates that the conventional sense of differentiation is reversed .

We now discuss an extension of the eigenvector concept to non-Hermitian canonical variables of the first kind . If q_k , p_k are Hermitian operators, a set of non-Hermitian canonical variables is provided by

$$\bar{q}_k = 2^{-1/2}(q_k + i p_k) \quad , \quad \bar{p}_k = 2^{-1/2}(p_k + i q_k) = i \bar{q}_k^{\dagger}$$

and indeed

$$[\bar{q}_k , \bar{q}_\ell] = [\bar{q}_k^{\dagger} , q_\ell^{\dagger}] = 0 \ , \ [\bar{q}_k , \bar{q}_\ell^{\dagger}] = \delta_{k\ell}$$

We first remark that the commuting Hermitian operators $\bar{q}_k^{\dagger} q_k$ possess simultaneously the eigenvalue zero and that the associated mutually adjoint eigenvectors $\bar{\Psi}(0)$ and $\bar{\Phi}(0)$ obey

$$\bar{q}_k \bar{\Psi}(0) = 0 \quad , \bar{\Phi}(0) \bar{q}_k^{\dagger} = 0 \qquad (4.9)$$

which permits them to be characterized, respectively, as the right eigenvector of the $\bar{q}_k$, and the left eigenvector of the $\bar{q}_k^{\dagger}$, with null eigenvalues . The equivalence of (4.9) with the initial definition follows from

$$\bar{\Phi}(0) \bar{q}_k^{T} \bar{q}_k \bar{\Psi}(0) = (\bar{q}_k \bar{\Psi}(0))^{\dagger} \bar{q}_k \bar{\Psi}(0) = 0$$

To verify the existence of this zero eigenvalue state , we construct its wave function in the q representation , at a given time, from the

eigenvector equation

$$\langle q' | \bar{q}_k \underline{\Psi}(0) = 2^{-1/2} \left(q'_k + \frac{\partial}{\partial q'_k} \right) \psi_0(q') = 0$$

and thus

$$\psi_0(q') = A \cdot e^{-\frac{1}{2} \sum_k q'^2_k} \qquad (4.10)$$

The constant A is fixed in magnitude as $\pi^{-n/2}$ on applying a unit normalization condition to the vector $\underline{\Psi}(0)$. In contrast with the situation for Hermitian canonical variables, $\bar{q}_k$ and $\bar{p}_k$ do not appear here in a completely symmetrical way; the null eigenvalue, right eigenvector of $\bar{q}_k^{\dagger}$ and left eigenvector of $\bar{q}_k$ do not exist. Indeed, were they to exist, we should draw from the commutation relations the conclusion that the non-negative Hermitian operators $\bar{q}_k^{\dagger} \bar{q}_k$ had the eigenvalue -1. Alternatively, on constructing the wave function representing this hypothetical state, we find that the negative sign in the exponent of (4.10), which makes the latter function bounded and normalizable, is replaced by a positive sign. Hence the new wave function is fundamentally unnormalizable, ant the corresponding state does not exist in any real sense. Yet this type of state will be found to have a limited utility.

It can now be concluded that the non-Hermitian variables q_k, $q_k^{\dagger}$, possess eigenvectors, $\underline{\Psi}(q')$ and $\underline{\Phi}(q')$, for arbitrary complex number eigenvalues, which vectors are in adjoint relation when $q_k^{\dagger\prime} = q'^{*}_k$. This follows from the observation that the commutation properties of the operators

$$'q_k = q_k - q'_k \quad , \quad 'q_k^\dagger = q_k^\dagger - q_k^{\dagger\prime} \qquad (4.11)$$

are identical with those of q_k and $q_k^\dagger$. Hence the operator (4.11) possess eigenvectors associated with zero eigenvalues, or equivalently, there exist vectors obeying the equations

$$(q_k - q'_k)\,\underline{\Psi}(q') = 0, \quad \underline{\Phi}(q^{\dagger\prime})(q_k^\dagger - q_k^{\dagger\prime}) = 0$$

for arbitrary complex numbers q_k', where

$$\underline{\Phi}(q^{\dagger\prime}) = \underline{\Psi}(q')^\dagger \quad , \quad q_k^{\dagger\prime} = q_k'^{*}$$

The application of the infinitesimal operators

$$G_q = \sum_k P_k \delta q_k = i \sum q_k^\dagger \delta q_k$$

$$G_{q^\dagger} = -i \sum \delta q_k^\dagger q_k = G_q^\dagger$$

to these eigenvectors, as in

$$\delta_q \underline{\Psi}(q') = -i\, G_q \underline{\Psi}(q')$$

$$\delta_{q^\dagger} \underline{\Phi}(q^{\dagger\prime}) = i \underline{\Phi}(q^{\dagger\prime})\, G_{q^\dagger}$$

yields eigenvectors of the operators $q_k - \delta q_k$ and $q_k^\dagger - \delta q_k^\dagger$ with the given eigenvalues, and thus is equivalent to changing the eigenvalues by the infinitesimal complex number δq and $\delta q^\dagger$. Hence,

$$\frac{\partial}{\partial q'_k}\,\underline{\Psi}(q') = q_k^\dagger\,\underline{\Psi}(q')$$

$$\frac{\partial}{\partial q^{\dagger\prime}}\,\underline{\Phi}(q^{\dagger\prime}) = \underline{\Phi}(q^{\dagger\prime})\, q_k \,,$$

which are differential operator realizations of the non-Hermitian canonical variables . For an arbitrary algebraic function of these variables at time t we have

$$\langle q^{\dagger\prime}, t \mid F(q, q^{\dagger}) = F\left(\frac{\partial}{\partial q^{\dagger\prime}}, q^{\dagger\prime}\right) \langle q^{\dagger\prime} t \mid$$

and

$$F(q, q^{\dagger}) \mid q' t\rangle = F^{T}\left(q', \frac{\partial}{\partial q'}\right) \mid q' t\rangle$$

$$= \mid q' t\rangle \; F^{T}\left(q', \frac{\partial}{\partial q'}\right)$$

Thus , the Schrödinger equations for a system characterized by a Hamiltonian operator $H(q, q^{\dagger}, t)$ read

$$i \frac{\partial}{\partial t} \psi(q^{\dagger\prime} t) = H\left(\frac{\partial}{\partial q^{\dagger\prime}}, q^{\dagger\prime}, t\right) \psi'(q^{\dagger\prime} t)$$

$$-i \frac{\partial}{\partial t} \phi(q' t) = \phi(q' t) \, H\left(q', \frac{\partial^{T}}{\partial q'}, t\right).$$

The analogue of (4. 4) is

$$\delta_{q^{\dagger}} \Psi(q') = -i \, G_{q^{\dagger}} \Psi(q') = -\left[\sum \delta q_{k}^{\dagger} \, q_{k}'\right] \Psi(q')$$

$$\delta_{q} \Phi(q^{\dagger\prime}) = i \, \Phi(q^{\dagger\prime}) \, G_{q} = -\Phi(q^{\dagger\prime}) \left[\sum q_{k}^{\dagger\prime} \delta q_{k}\right],$$

but unlike the situation with Hermitian variables , these infinitesimal changes are not necessarily imaginary numerical multiples of the original vectors . Hence , the special canonical group for non-Hermitian variables constitues the totality of transformations that change an eigenvector with given complex eigenvalues into one with any other set of eigenvalues , and which is multiplied by arbitrary complex numerical factors . The transformation function $\langle q^{\dagger\prime} \mid q' \rangle$ is obtained from the differential expression

$$\delta \langle q^{t\prime} | q' \rangle = i \langle q^{t\prime} | G_{q^t} - G_q | q' \rangle$$

as

$$\langle q^{t\prime} | q' \rangle = e^{\sum q_k^{t\prime} q_k'} \qquad (4.12)$$

in which the choice of multiplicative constant is based upon the normalization adopted for the zero eigenvalue state . Note that all other eigenvectors exceed unity in length ,

$$\underline{\Psi}(q')^\dagger \underline{\Psi}(q') = e^{\sum |q_k'|^2}$$

The differential composition property of (4.12) differs from (4.8) only in the absence of the numerical factor ,

$$\langle q^{t\prime} | q' \rangle = \langle q^{t\prime} | \frac{\partial}{\partial q^{t\prime\prime}} \rangle \langle q^{t\prime\prime} | q' \rangle \Big|_{q^{t\prime\prime}=0}$$

and a similar relation is

$$\langle q^{t\prime} | q' \rangle = \langle q^{t\prime} | q'' \rangle \langle \frac{\partial^T}{\partial q''} | q' \rangle \Big|_{q''=0}$$

We can also view these results as particular representations of the completness properties

$$1 = | \frac{\partial}{\partial q^{t\prime}} \rangle \langle q^{t\prime} | \Big|_{q^{t\prime}=0} = | q' \rangle \langle \frac{\partial^T}{\partial q'} | \Big|_{q'=0} \qquad (4.13)$$

and infer the general scalar product evaluations

$$\underline{\Phi}_1 \underline{\Psi}_2 = \phi_1 \left(\frac{\partial}{\partial q^{t\prime}}\right) \psi_2 (q^{t\prime}) \Big|_{q^{t\prime}=0} = \phi_1 (q') \psi_2 \left(\frac{\partial^T}{\partial q'}\right) \Big|_{q'=0}$$

as well as the trace formulae

$$\mathrm{tr}\ X = \langle \frac{\partial}{\partial q'} | X | q' \rangle \Big|_{q'=0} = \langle q^{t\prime} | X | \frac{\partial^+}{\partial q^{t\prime}} \rangle \Big|_{q^{t\prime}=0}$$

Since we are dealing with complex eigenvalues , one can give

(4.I3) a contour integral form . Thus

$$1 = \oint \prod_{k=1}^{n} \left(\frac{1}{2\pi i} \cdot \frac{dq_k^{+\prime}}{q_k^{+\prime}} \right) \left| \frac{\partial}{\partial q^{+\prime}} \right\rangle \left\langle q^{+\prime} \right| \qquad (4.I4)$$

in which each integration path encircles the origin in the positive sense .

If we also write :

$$\frac{1}{q_k^{+\prime}} = \int d\, q'_k \quad e^{-\, q'_k\, q_k^{+\prime}} \quad ,$$

where the integration is extended from zero to infinity along a suitable path, (4.I4) becomes

$$1 = \frac{1}{(2\pi i)^n} \int d\, q' d\, q^{+\prime} \left| q' \right\rangle \quad e^{-\sum q'_k q_k^{+\prime}} \left\langle q^{+\prime} \right| ,$$

or equivalently ,

$$\Phi_1 \Psi_2 = \frac{1}{(2\pi i)^n} \int dq' dq^{+\prime} \varphi_1(q') e^{-\sum q'_k q_k^{+\prime}} \psi_2(q^{+\prime})$$

and

$$\mathrm{tr}\, X = \frac{1}{(2\pi i)^n} \int dq' dq^{+\prime} e^{-\sum q'_k q_k^{+\prime}} \left\langle q^{+\prime} \right| X \left| q' \right\rangle .$$

It is now possible to define vectors that, to some extent, play the role of right eigenvector of the q^{+} or left eigenvector of the q, namely

$$\left| q^{+\prime} \right\rangle = \int dq' \left| q' \right\rangle e^{-\sum q'_k q_k^{+\prime}} = \left| + \frac{\partial}{\partial q^{+\prime}} \right\rangle \prod_k \frac{1}{q_k^{+\prime}}$$

and

$$\left\langle q' \right| = \int dq^{+\prime} e^{-\sum q'_k q_k^{+\prime}} \left\langle q^{+\prime} \right| = \prod_k \frac{1}{q'_k} \left\langle - \frac{\partial^{\top}}{\partial q'} \right| ,$$

for then the completeness statement reads

$$1 = \frac{1}{(2\pi i)^n} \oint d q^{+\prime} \, | q^{+\prime} > < q^{+\prime} |$$

$$= \frac{1}{(2\pi i)^n} \oint d q^{\prime} \, | q^{\prime} > < q^{\prime} |$$

The wave function transformation equations

$$\psi(q') = \int dq^{+\prime} \, e^{-\sum q'_k q^{+\prime}_k} \psi(q^{+\prime}) = \psi\left(-\frac{\partial}{\partial q'}\right) \prod_k \frac{1}{q'_k}$$

and

$$\phi(q^{+\prime}) = \int dq' \, \phi(q') \, e^{-\sum q'_k q^{+\prime}_k} = \phi\left(-\frac{\partial}{\partial q^{+\prime}}\right) \prod_k \frac{1}{q^{+\prime}_k},$$

combined with

$$\psi_{q'}(q^{+\prime}) = \phi_{q^{+\prime}}(q') = e^{\sum q^{+\prime}_k q'_k}$$

show that wave functions of the type $\psi(q^{+\prime})$ and $\phi(q')$ are regular functions of the corresponding complex variables in the neighborhood of the origin, while functions of the type $\psi(q')$ and $\phi(q^{+\prime})$ are regular in the neighborhood of the point at infinity. In the evaluation of a scalar product,

$$\underline{\phi}_1 \underline{\psi}_2 = \frac{1}{(2\pi i)^n} \oint dq^{+\prime} \, \phi_1(q^{+\prime}) \, \psi_2(q^{+\prime})$$

$$= \frac{1}{(2\pi i)^n} \oint dq' \, \phi_1(q') \, \psi_2(q')$$

the integration paths are to be drawn in the common domain of regularity of the two wave functions. For example,

$$< q' | q'' > = \int dq^{+\prime} \, e^{-\sum q'_k q^{+\prime}_k} \cdot e^{\sum q^{+\prime}_k q''_k} =$$

$$= \prod_k \frac{1}{q'_k - q''_k}, \quad |q'_k| > |q''_k|,$$

considered as $\psi_{q''}(q')$ or $\phi_{q'}(q'')$, shows these regularity characteristics, and we obtain

$$\frac{1}{(2\pi i)^n} \oint dq' \, \phi(q') \, \langle q' | \, q'' \rangle = \phi(q'')$$

if the q' contours enclose the points q'', but no singularity of $\phi(q')$, while

$$\frac{1}{(2\pi i)^n} \oint dq'' \, \langle q' | \, q'' \rangle \, \psi(q'') = \psi(q') \, ,$$

if the q'' integration paths contain all the singularity of $\psi(q'')$ but not the points q'. The integral is evalued with the end of a circle about the point at infinity where $\psi(q)$ wave functions vanish. Similar properties hold for the transformation functions

$$\langle q^{t\prime} | \, q^{t''} \rangle = \prod_k \frac{1}{q_k^{t''} - q_k^{t}} \, , \quad |q_k^{t''}| > |q_k^{t\prime}|$$

The manner in which $\langle q' |$ differs from a true eigenvector is indicated by

$$\langle q' | \, (q_k - q_k') = \int dq^{t\prime} \frac{\partial}{\partial q^{t\prime}} \left(e^{-\sum q_k' \, q_k^{t\prime}} \langle q^{t\prime} | \right)$$

$$= - \prod_{\ell \neq k} \frac{1}{q_\ell'} \langle \frac{\partial^t}{dq'} | \, ,$$

or

$$\langle q' | \, q_k = \langle q' | \, q_k' - \frac{1}{2\pi i} \oint dq_k' \, \langle q' | \, ,$$

and similarly

$$q_k^t | \, q^{t\prime} \rangle = q_k^{t\prime} | \, q^{t\prime} \rangle - \frac{1}{2\pi i} \oint dq_k^{t\prime} | \, q^{t\prime} \rangle \, .$$

The importance of these additional terms is seen on comparing the two equivalent evaluations of the matrix $\langle q' \mid q_k \mid q'' \rangle$,

$$\langle q' \mid q_k \mid q'' \rangle = \langle q' \mid q'' \rangle \, q''_k = \frac{q''_k}{q'_k - q''_k} \prod_{l \neq k} \frac{1}{q'_l - q''_l}$$

and

$$\langle q' \mid q_k \mid q'' \rangle = q'_k \langle q' \mid q'' \rangle - \frac{1}{2\pi i} \oint dq'_k \langle q' \mid q'' \rangle$$

$$= \left(\frac{q'_k}{q'_k - q''_k} - 1 \right) \prod_{l \neq k} \frac{1}{q'_l - q''_l}$$

For the transformation functions $\langle q' \mid q^{t\prime} \rangle$ we find

$$\langle q' \mid q^{t\prime} \rangle = \int dq'' \prod_k \frac{1}{q'_k - q''_k} \, e^{- q''_k q^{t\prime}_k}$$

$$= \prod_{k=1}^{n} \int_0^\infty d\tau_k \, \frac{e^{-\tau_k}}{q'_k q^{t\prime}_k - \tau_k}$$

which is not uniquely defined , and indeed does not exist when $q'_k \, q^{t\prime}_k = 0$. However , for the purposes of integration , when multiplied by a wave function $\psi(q^{t\prime})$ or $\phi(q')$, all that is significant is the behavior of this transformation function in the neighborhood of the point at infinity , as represented by the asymptotic expansion .

$$\langle q' \mid q^{t\prime} \rangle \sim \prod_k \sum_{n=0}^{\infty} \frac{n!}{(q'_k \, q^{t\prime}_k)^{n+1}}$$

A concept of eigenvalue and eigenvector will now be devised for

the non-Hermitian , anticommuting canonical variables of the second kind , q_α , $q_\alpha^\dagger$. The eigenvalues of commuting variables of the first kind are quantities that commute with all operators and with each other . Correspondingly the eigenvalues of the anticommuting variables of the second kind must anti-commute with all variables of the latter type and with each other , which requires a suitable extension of the number system . Let us observe, first that

$$q_\alpha^\dagger q_\alpha q_\alpha^\dagger q_\alpha = q_\alpha^\dagger (1 - q_\alpha^\dagger q_\alpha) q_\alpha = q_\alpha^\dagger q_\alpha ,$$

since $q_\alpha^2 = q_\alpha^{\dagger 2} = 0$. Hence the spectra of the commuting Hermitian operators $q_\alpha^\dagger q_\alpha$ contain only the numbers 0 and 1 , and the operators $q_\alpha q_\alpha^\dagger$ exhibit the same spectra . The state in which all quantities $q_\alpha^\dagger q_\alpha$ have the value zero is equivalently described as a zero eigenvalue, right eigenvector of the q_α or a zero eigenvalue, left eigenvector of the $q_\alpha^\dagger$,

$$q_\alpha \Psi(0) = \Phi(0) q_\alpha^\dagger = 0 \qquad (4.15)$$

Now the operators

$$'q_\alpha = q_\alpha - q'_\alpha \quad , \quad 'q_\alpha^\dagger = q_\alpha^\dagger - q_\alpha^{\dagger\prime}$$

have the same commutation relations as the q_α , $q_\alpha^\dagger$ provided the numbers q'_α and $q_\alpha^{\dagger\prime}$ anticommute among themselves and with the second kind variables . Thus ,

$$\{ 'q_\alpha , 'q_\beta^\dagger \} = \{ q_\alpha , q_\beta^\dagger \} - \{ q'_\alpha , q_\beta^\dagger \} - \{ q_\alpha , q_\beta^{\dagger\prime} \} + \{ q'_\alpha , q_\beta^{\dagger\prime} \} = \delta_{\alpha\beta}$$

and therefore the new operators possess null eigenvalue states in the manner of (4.I5) or, there exist vectors that obey

$$(q_\alpha - q'_\alpha) \underline{\Psi}(q') = 0 , \quad \underline{\Phi} (q^{\dagger\prime}) (q^\dagger_\alpha - q^{\dagger\prime}_\alpha) = 0 ,$$

which identifies q'_α and $q^{\dagger\prime}_\alpha$ as the eigenvalues of the operators q_α and $q^\dagger_\alpha$. Furthermore ,

$$\underline{\Phi} (q^{\dagger\prime}) = \underline{\Psi}(q')^\dagger , \qquad q^{\dagger\prime}_\alpha = q'^\dagger_\alpha$$

in which the adjoint rather than the complex conjugate sign is affixed to the eigenvalue since the sense of multiplication of eigenvalues is significant.

The infinitesimal operators

$$G_q = i \sum_\alpha q^\dagger_\alpha \delta q_\alpha$$

$$G_q = - i \sum \delta q^\dagger_\alpha \; q_\alpha = G_q^\dagger$$

produce changes in the eigenvectors ,

$$\left.\begin{aligned} \delta_q \underline{\Psi} (q') &= - i G_q \underline{\Psi} (q') = \sum q^\dagger_\alpha \delta q_\alpha \underline{\Psi} (q') \\ \delta_{q^\dagger} \underline{\Phi}(q^{\dagger\prime}) &= i \underline{\Phi} (q^{\prime\prime}) \; G_{q^\dagger} = \underline{\Phi} (q^{\dagger\prime}) \sum \delta q^\dagger_\alpha \; q_\alpha \end{aligned}\right\} \quad (4.I6)$$

that yield the eigenvectors of the operators $q_\alpha - \delta q_\alpha$ and $q^\dagger_\alpha - \delta q^\dagger_\alpha$ with the eigenvalues q'_α and $q^{\dagger\prime}_\alpha$. But the commutation properties of the operator variations are identical with those of the eigenvalues , which permit us to write (4.I6) as

$$\frac{\partial_r}{\partial q'_\alpha} \underline{\Psi} (q') = q^\dagger_\alpha \underline{\Psi}(q') , \quad \frac{\partial_l}{\partial q^{\dagger\prime}_\alpha} \underline{\Phi} (q^{\dagger\prime}) = \underline{\Phi}(q^{\dagger\prime}) \; q_\alpha , \qquad (4.I7)$$

and produces a differential operator realization of the canonical variables of the second kind . In arriving at (4.17) , we have supposed that the eigenvalue variations commute with the eigenvectors . This will be correct if it is valid for the null eigenvalue states . With arbitrary algebraic functions of these variables at the given time , we find

$$\langle q' | F(q, q^{\dagger}) = F\left(\frac{\partial_e}{\partial_{q^{\dagger'}}}, q^{\dagger'}\right) \langle q^{\dagger'} |$$

and

$$F(q, q^{\dagger}) |q'\rangle = |q'\rangle F\left(q, \frac{\partial_r^{\tau}}{\partial q'}\right) .$$

Hence , for a system described by variables of the second kind and the Hamiltonian $H(q, q^{\dagger}, t)$, we have the Schrödinger equations

$$i\frac{\partial}{\partial t} \psi(q^{\dagger'} t) = H\left(\frac{\partial_e}{\partial q^{\dagger'}}, q^{\dagger'}, t\right) \psi(q^{\dagger'} t)$$

$$-i\frac{\partial}{\partial t} \phi(q' t) = \phi(q' t) \; H\left(q', \frac{\partial_r^{\tau}}{\partial q'}, t\right) .$$

The infinitesimal changes

$$\delta_{q^{\dagger}} |q'\rangle = -\sum \delta dq_{\alpha}^{\dagger} \; q'_{\alpha} \; |q'\rangle$$

$$\delta_q \langle q^{\dagger'}| = -\langle q^{\dagger'}| \sum q_{\alpha}^{\dagger'} \delta q_{\alpha}$$

involve multiplication of the original vectors by infinitesimal numbers, and thus the special canonical group for variables of the second kind contains the totality of transformations that alter eigenvalues and multiply eigenvectors by commuting numbers formed from bilinear combinations of the eigenvalues . In the construction of the transformation function $\langle q^{\dagger'} | q' \rangle$

from the differential expression

$$\delta \langle q^{+\prime} | q' \rangle = \langle q^{+\prime} | q' \rangle \delta \left(\sum q^{+\prime}_{\alpha} q'_{\alpha} \right) ,$$

the commutativity of the bilinear eigenvalue products leads immediately to

$$\langle q^{+\prime} | q' \rangle = e^{\sum_{\alpha} q^{+\prime}_{\alpha} q'_{\alpha}} ,$$

or

$$\langle q^{+\prime} | q' \rangle = \prod_{\alpha} e^{q^{+\prime}_{\alpha} q'_{\alpha}} = \prod_{\alpha} (1 + q^{+\prime}_{\alpha} q'_{\alpha}) ,$$

since the square of any eigenvalue vanishes . This transformation function has the composition properties

$$\langle q^{+\prime} | q' \rangle = \left\langle q^{+\prime} \left| \frac{\partial \ell}{\partial q^{+\prime\prime}} \right\rangle \right\langle q^{+\prime\prime} | q' \rangle \Big|_{q^{+\prime\prime} = 0}$$

and

$$\langle q^{+\prime} | q' \rangle = \langle q^{+\prime} | q'' \rangle \left\langle \frac{\partial r^{T}}{\partial q''} \middle| q' \right\rangle \Big|_{q'' = 0} ,$$

from which we infer the scalar product evaluations

$$\underline{\phi}_1 \underline{\psi}_2 = \phi_1 \left(\frac{\partial \ell}{\partial q^{+\prime}} \right) \psi_2 (q^{+\prime}) \Big|_{q^{+\prime} = 0}$$

$$= \phi_1 (q') \psi_2 \left(\frac{\partial r^{T}}{\partial q'} \right) \Big|_{q' = 0} ,$$

and

$$\mathrm{tr}\, X = \left\langle \frac{\partial r}{q'} \middle| X \middle| q' \right\rangle \Big|_{q' = 0} = \left\langle q^{+\prime} \middle| X \middle| \frac{\partial \ell^{T}}{\partial q^{+\prime}} \right\rangle \Big|_{q^{+\prime} = 0}$$

For the variables of the second kind , there is complete symmetry between q_{α} and q^{+}_{α} . Hence one can also define eigenvectors $\underline{\psi} (q^{+\prime})$

and $\overset{T}{\phi}(q')$, and repeat the above discussion with q and q^+ interchanged. The connection between the $|q'\rangle$, $\langle q^{+\prime}|$ states and the $|q^{+\prime}\rangle$, $\langle q'|$ states is described by the transformation functions $\langle q' \mid q'' \rangle$ and $\langle q^{+\prime} \mid q^{+}{}'' \rangle$. Now

$$\delta \langle q' \mid q'' \rangle = \langle q' | \sum_\alpha \delta q'_\alpha \, q^+_\alpha + \sum_\alpha q^+_\alpha \, \delta q''_\alpha \mid q'' \rangle$$

$$= \langle q' \mid q'' \rangle \sum_\alpha \frac{\partial \overset{T}{r}}{\partial q''_\alpha} \left(\delta_{q''_\alpha} - \delta_{q'_\alpha} \right)$$

and

$$0 = \langle q' \mid q_\alpha - q_\alpha \mid q'' \rangle = \langle q' \mid q'' \rangle \left(q'_\alpha - q''_\alpha \right)$$

which shows that $\langle q' \mid q'' \rangle$ is a function of the eigenvalue differences $q''_\alpha - q'_\alpha$, that vanishes on multiplication with each such difference . Hence , to within a constant factor ,

$$\langle q' \mid q'' \rangle = \prod_\alpha \left(q''_\alpha - q'_\alpha \right) . \qquad (4.18)$$

in which the product is arranged in the standard order , 1, 2 , ... n , from left to right . On taking the adjoint , the order of multiplication becomes transposed ,

$$\langle q^{+\prime} \mid q \rangle = \prod_\alpha \overset{T}{} \left(q^{+\prime}_\alpha - q^{+\prime\prime}_\alpha \right) \qquad (4.19)$$

The choice of unit numerical constants in (4.18) and (4.19) is verified , to within inverse phase factors , by the validity of the composition property

$$\langle q^{+\prime} \mid \frac{\partial \ell}{\partial q'} \rangle \langle q' \mid q'' \rangle \Big|_{q'=0} = \langle q^{+\prime} \mid q'' \rangle$$

Indeed ,

$$\prod_{\alpha}^{T} \left(q_{\alpha}^{+1} - \frac{\partial \ell}{\partial q'_{\alpha}} \right) \prod_{\alpha} \left(q''_{\alpha} - q'_{\alpha} \right) = \prod_{\alpha} \left(1 + q_{\alpha}^{+1} q''_{\alpha} \right) ,$$

in which the opposite sense of multiplication in the two products permits their combination without the intervention of the minus signs that accompagny the anticommutativity of the eigenvalues . It will be observed that $\langle q' \mid q'' \rangle$ is homogeneous of degree n in the eigenvalues . Hence if $\mid q'' \rangle$ commutes with the eigenvalues , $\langle q' \mid$ and $\mid q^{+1} \rangle$ anticommute with them , for odd n , and it is only for even n that the symmetry between the operator q and q^{+} extends completely to their eigenvectors .

The connection between the wave functions $\psi (q')$ and $\psi (q^{+\nu})$ that represent a given state is obtained from

$$\langle q^{+1} \mid \Psi = \langle q^{+1} \mid q^{+\mathrm{II}} \rangle \left\langle \frac{\partial^{T}_{\nu}}{\partial q^{+\mathrm{II}}} \mid \Psi \mid \right._{q^{+\mathrm{II}} = 0}$$

$$= \prod_{\alpha}^{T} \left(q_{\alpha}^{+1} - q_{\alpha}^{+\mathrm{II}} \right) \left\langle \frac{\partial^{T}_{\nu}}{\partial q^{+\mathrm{II}}} \mid \Psi \right|_{q^{+\mathrm{II}} = 0}$$

as

$$\psi (q^{+1}) = \prod_{\alpha}^{T} \left(q_{\alpha}^{+1} \right) \psi \left(- \frac{\partial^{T}_{\nu}}{\partial q^{+1}} \right)$$

and similarly

$$\langle q' \mid \Psi = \langle q' \mid q'' \rangle \left\langle \frac{\partial^{T}_{\nu}}{\partial q''} \mid \Psi \right|_{q'' = 0}$$

yields the inverse relation

$$\psi (q') = \prod_{\alpha} \left(- q'_{\alpha} \right) \psi \left(- \frac{\partial^{T}_{\nu}}{\partial q'} \right)$$

We also have

$$\phi(q') = \phi(-\frac{\partial \ell}{\partial q'}) \prod_{\alpha} (q'_{\alpha})$$

and

$$\phi(q^{+\prime}) = \phi(-\frac{\partial \ell}{\partial q^{+\prime}}) \prod_{\alpha}^{T} (-q^{+\prime}_{\alpha}).$$

The evaluation of the scalar product $\underset{\sim}{\phi}_1 \underset{\sim}{\psi}_2$ with the aid of the wave functions $\phi_1(q')$ and $\psi_2(q')$ is derived from the following expression of completeness,

$$1 = |q'\rangle \langle \frac{\partial r}{\partial q'}^{T} | \frac{\partial \ell}{\partial q''} \rangle \langle q''| \Big|_{q'=q''=0},$$

as

$$\underset{\sim}{\phi}_1 \underset{\sim}{\psi}_2 = \phi_1(q') \prod_{\alpha}^{T} (\frac{\partial r}{\partial q'_{\alpha}}^{T} - \frac{\partial \ell}{\partial q''_{\alpha}}) \psi_2(q'') \Big|_{q'=q''=0}$$

The results we have obtained for systems characterized by variables of the first kind, or of the second kind, can evidently be combined to deal with a physical system that requires both types of dynamical variables for its description.

5. Canonical Transformations

The application of the stationary action principle to a given Lagrangian operator yields a specific form for the generator of infinitesimal transformations at the terminal times. With a suitable labelling of canonical variables, such a generator is

$$G = \sum_{a} p_a \delta q_a - H(q,p,t) \delta t \qquad (5.1)$$

in which the summation is extended over both types of dynamical variables . From the significance of the transformations generated by this operator follow the canonical commutation relations and the canonical equations of motion . Now , we have already remarked that the addition of a time derivative to the Lagrangian produces new generators of transformations , for the same dynamical system . More generally , if

$$\delta (\bar{L}\, dt - L\, dt) = - d (\delta W)$$

we shall have

$$\delta \bar{W}_{12} - \delta W_{12} = - \delta W_1 + \delta W_2 \quad ,$$

and thereby the new generators

$$\bar{G} = G - \delta W$$

at times t_1 and t_2 , with the same equations of motion . One can then ask for the conditions under which $\bar{G}$ also displays the canonical form (5.1) , but expressed in terms of new dynamical variables $\bar{q}_a(t), \bar{p}_a(t)$ and a new Hamiltonian operator $\bar{H} (\bar{q} , \bar{p} , t)$. According to the interpretation we should extend to $\bar{G}$ the operators $\bar{q}$, $\bar{p}$ would obey the canonical commutation relations and canonical equations of motion , which characterizes such a change of fundamental dynamical variables as a canonical transformation .

We shall regard

$$\delta W (q, \bar{q} , t) = \sum p_a \delta q_a - H \delta t - \sum \bar{p}_a \delta \bar{q}_a + \bar{H} \delta t$$

as the result of applying the analogous elementary operator variations

δq_a and $\delta \bar{q}_a$ to the variables contained in the action operator W, and of an infinitesimal time alteration. If W contains other superfluous canonical variables v_a, in addition to the q_a and $\bar{q}_a$, we infer the differential equations

$$\left.\begin{aligned} p_a &= \frac{\partial_r W}{\partial q_a} , \quad -\bar{p}_a = \frac{\partial_r W}{\partial \bar{q}_a} , \quad \frac{\partial W}{\partial v_a} = 0 \\ \bar{H} &= H + \frac{\partial W}{\partial t} . \end{aligned}\right\} \quad (5.2)$$

A canonical transformation is obtained should these implicit operator equations possess a solution for the $\bar{q}_a$ and $\bar{p}_a$.

Canonical transformations form a group. The action operator that describes the transformation inverse to $q , p \longrightarrow \bar{q} , \bar{p}$ is

$$W (\bar{q} , q) = - W (q , \bar{q}) ,$$

while, for two successive transformations $q , p \longrightarrow \bar{q} , \bar{p} , \bar{q}, \bar{p} \longrightarrow \bar{\bar{q}} , \bar{\bar{p}}$, the generator of the composite transformation is

$$W (q , \bar{\bar{q}}) = W (q , \bar{q}) + W (\bar{q} , \bar{\bar{q}}) .$$

The latter form illustrates the idea of superfluous variables. If the individual operators $w (q , \bar{q})$ and $W (\bar{q} , \bar{\bar{q}})$ contain only the indicated variables, the sum involves $q , \bar{\bar{q}}$, and $\bar{q}$. But according to the conditions for the separate canonical transformations,

$$\frac{\partial_r}{\partial q_a} W (q , \bar{\bar{q}}) = - \bar{p}_a + \bar{p}_a = 0 ,$$

and it must be possible to exhibit $W (q , \bar{\bar{q}})$ as a function of the

variables q , $\bar{\bar{q}}$ alone . It is not always desirable , however, to eliminate the superfluous variables . This is particularly true when the canonical transformation involves relations among the variables q_a and $\bar{\bar{q}}_a$, which prevents their independent variation . By retaining the variables of a suitable intermediate transformation , one can derive the desired transformation by independent differentiation . An important example is provided by the identity transformation . Let us remark first on the familiar transformation that interchanges the roles of the complementary q and p variables, as in

$$G_q - G_p = \delta W (q , p)$$

where

$$W (q , p) = - W (p , q) = \sum_a p_a \cdot q_a ,$$

and the dot indicates the symmetrized or antisymmetrized product for the variables of the first and second kind , respectively . Now

$$G_p = -\sum_a \delta p_a \; q_a = -\sum_i q_i \, \delta p_i + \sum q_\alpha \, \delta p_\alpha ,$$

so that, considered as a canonical transformation , we have

$$\bar{q}_i = p_i \quad , \quad \bar{p}_i = - q_i$$

$$\bar{q}_\alpha = p_\alpha \quad , \quad p_\alpha = q_\alpha$$

for the two kinds of variables . On adding the action operators for the transformations $q \longrightarrow p$ and its inverse $p \rightarrow q = \bar{q}$, we obtain :

$$W (q , \bar{q}) = \sum_a p_a \cdot (q_a - \bar{q}_a)$$

which contains the p_a as superfluous variables . To eliminate the latter , we must impose the transformation equations $\bar{q}_a = q_a$, which yields $W = 0$. But , with the superfluous variables retained , the differential equations (5.2) are applicable , and the identity transformation is generated .

Transformations in the infinitesimal neighborhood of the identity - infinitesimal canonical transformations - must be described by an action operator that differs infinitesimally from the one producing the identity transformation . Such a form is

$$W (q , \bar{q} , t) = \sum_a p_a (q_a - \bar{q}_a) - G (\bar{q} , p, t) \qquad (5.3)$$

where G is an essentially arbitrary infinitesimal function of the indicated variables (it must be an even function of the variables of the second kind). On applying (5.2) , with the p_a as superfluous variables , we obtain the explicit equations of an infinitesimal canonical transformation ,

$$\bar{q}_a = q_a - \delta q_a \quad , \quad \bar{p}_a = p_a - \delta p_a$$

$$\delta q_a = \frac{\partial}{\partial p_a} G (q , p , t) \quad , \quad \delta p_a = - \frac{\partial}{\partial q_a} G (q , p , t)$$

$$\bar{H} (\bar{q} , \bar{p} , t) = H (q, p , t) - \frac{\partial}{\partial t} G (q , p, t)$$

Now , according to (3.10) and (3.11) ,

$$\delta q_a = \frac{1}{i} \left[q_a , G \right] \quad , \quad \delta p_a = \frac{1}{i} \left[p_a , G \right]$$

or , more generally ,

$$\delta x_a = \frac{1}{i} \left[x_a , G \right] .$$

which means that

$$\bar{X}_a = X_a - \delta X_a = (1 - i G) X_a (1 + i G) . \qquad (5.4)$$

If G is a Hermitian operator , this is a unitary transformation , and indeed the subgroup of canonical transformations that preserve the Hermiticity of dynamical variables (real canonical transformations) is equivalent to the group of unitary transformations . Since the transformation (5.4) maintains all algebraic relations , we have

$$\bar{H} (\bar{q} , \bar{p} , t) = (1 - i G) \bar{H} (q , p , t) (1 + i G) = H(q,p,t) + \frac{1}{i}\left[G , H\right]$$

and therefore

$$\bar{H} (q,p,t) - H(q,p,t) = -\frac{\partial}{\partial t} G - \frac{1}{i}\left[G , H\right] = - \frac{d}{dt} G ,$$

which makes explicit the functional form of the new Hamiltonian . Infinitesimal canonical transformations that leave the Hamiltonian form invariant have infinitesimal generators G that are constants of the motion .

Among the possible forms of δq_a and δp_a , are the ones with elementary operator properties that characterize the transformations of the special canonical group . Should it be necessary to distinguish the latter from the general infinitesimal transformation, we shall use the symbol ∂ rather than δ . Thus , the infinitesimal transformations of the special canonical group ,

$$\partial q_a = \frac{\partial_l G_q}{\partial p_a} \quad , \quad 0 = - \frac{\partial_r G_q}{\partial q_a}$$

and

$$0 = \frac{\partial_l G_p}{\partial p_a} \quad , \qquad = - \frac{\partial_r G_p}{\partial q_a}$$

are generated by

$$G_q = \sum p_a \; \partial q_a \quad , \qquad G_p = - \sum \partial p_a \; q_a \quad .$$

The operator $G_t = - H \delta t$ produces the canonical transformation in which the dynamical variables at time t are replaced by those at time $t + \delta t$, and the energy operator H is a constant of the motion when t does not occur explicitly, as is required for the preservation of the functional form of H under a time translation.

The change in description that accompanies a rotation of the spacial coordinate system is a unitary transformation, and a canonical transformation, with the infinitesimal generator

$$G_\omega = \delta \omega \cdot J \quad .$$

With regard to the form of this transformation, we remark that differently oriented coordinate system are intrinsically equivalent and we should expect that the kinematical term in the Lagrangian presents the same appearance in terms of the analogous variables appropriate to any coordinate system. The same comment applies to the Hamiltonian operator of a physically isolated system, for which the total angular momentum operator J, as the generator of a transformation that leaves the form of H invariant, is a constant of the motion. In view of a bilinear structure of the kinematical term in L,

$$\sum_a p_a \cdot \frac{dq_a}{dt} \quad ,$$

the change involved in an arbitrary rotation of the spatial coordinate system will be a linear transformation among suitably chosen q variables combined with the contragradient transformation of the p variables . For an infinitesimal rotation , then ,

$$\delta_\omega q_a = \frac{\partial_l G_\omega}{\partial p_a} = -i \sum_b \delta\omega \cdot J_{ab} \, q_b$$

$$\delta_\omega p_b = -\frac{\partial_r G_\omega}{\partial q_b} = i \sum_a p_a \, \delta\omega \cdot J_{ab} \quad ,$$

in which J_{ab} is a numerical vector matrix . The general form of the angular momentum operator is thereby obtained as

$$J = -i \sum_{ab} p_a \cdot J_{ab} \, q_b = -i \; p \cdot J q \quad ,$$

in which the two types of dynamical variables appear additively . If Hermitian variables of the first kind are employed , the matrix J_{ik} must be pure imaginary , while with non-Hermitian variables in the relation $p_a = i q_a^+$, the matrix J_{ab} is Hermitian . Either choice yields a Hermitian operator for J .

On applying the infinitesimal transformation generated by G_ω to the operator J , we find :

$$\delta_\omega J = \frac{1}{i} \left[J , \; \delta\omega \cdot J \right] = \delta\omega \times J$$

according to the commutation relations (2.13) . But we also have

$$\delta_\omega J = -i \; (p \cdot J \delta\omega q + \delta\omega \, p \cdot J q)$$

$$= -i\, p\, \frac{1}{i}\left[J, \delta\omega\cdot J\right] q \ ,$$

and therefore

$$\frac{1}{i}\left[J, \delta\omega J\right] = \delta\omega \times J \ .$$

The numerical matrices J obey the same commutation relations as the angular momentum operator J. It is an incidental consequence of these commutation relations that the trace of the matrix J vanishes, and accordingly, the explicit symmetrization or anti-symmetrization of factors in J is unnecessary. The decomposition of the matrix J into irreducible submatrices produces a partioning of the dynamical variables into kinematically independent groups that appear additively in the structure of J. Each such group defines a dynamical variable of several components whose rotational transformation properties are fixed by the number of components, for this integer, the dimensionality of the corresponding submatrix of J, essentially determines the structure of these matrix representations of J. An irreductible set of three variables necessarily has the rotational transformation properties of a three - dimensional space vector, for example. The discussion of the previous section indicates that the number of components possessed by a dynamical variable of the second kind is even.

Our comments concerning invariance with respect to rotations of the coordinate system apply equally to coordinate system translations, which have the infinitesimal generator

$$G_\varepsilon = \delta\varepsilon \cdot P$$

The appropriate transformation of the dynamical variables that leaves the kinematical part of L invariant is the addition of constants to(suitably chosen) variables q_a . Then G_ε will have the structure of G_q and different generators of this type do commute , as the commutation properties of the total linear momentum operator require . If we apply an infinitesimal rotation to the operator

$$P = \sum p_a \frac{\delta q_a}{\delta \varepsilon} ,$$

in accordance with

$$\delta_\omega P = \frac{1}{i}[P , \delta\omega \cdot J] = \delta\omega \times P$$
$$= \sum \delta_\omega p_a \frac{\delta q_a}{\delta \varepsilon} ,$$

we recognize that the dynamical variables p_a that make a contribution to P have the rotational transformation properties of space vectors . Hence , only three-component variables of the first kind can be affected by a translation of the coordinate system . If the dynamical variables that depend upon the location of the coordinate system are presented as the Hermitian vectors $r_1, \ldots, r_n$, with complementary variables $p_1, \ldots, p_n$, a suitable adjustement of the relative eigenvalue scales will guarantee that

$$\delta_\varepsilon r_k = \delta\varepsilon$$

and therefore

$$P = \sum_{k=1}^{n} p_k .$$

On exhibiting the contribution of these variables to the total angular momentum, we have

$$J = \sum_{k=1}^{n} r_k \times p_k + J_{int} \quad ,$$

in which the latter term contains all variables that are uninfluenced by translations . These are evidently the internal variables for a system of n particles that are localized spatially by the position coordinates r_k .

We now examine infinitesimal canonical transformations that are characterized by small changes , $- d\tau_r$, of certain parameters τ_r , so that the infinitesimal generator has the form

$$G = \sum G_{(r)} (- d \tau_r) \quad ,$$

in which the $G_{(r)}$ may depend explicitly upon the values of the parameters . According to the meaning of the canonical transformation , the transformed dynamical variables are given by

$$\bar{x}_a = x_a + \sum_r \frac{dx_a}{d\tau_r} (- d\tau_r)$$

or

$$\delta x_a = x_a - \bar{x}_a = \sum \frac{dx_a}{d\tau_r} d\tau_r = dx_a \quad .$$

For the canonical variables , then , we have the equations of motion

$$\frac{dq_a}{d\tau_r} = - \frac{\partial_1 G_{(r)}}{\partial p_a} \quad , \quad - \frac{dp_a}{d\tau_r} = \frac{\partial_r G_{(r)}}{\partial q_a} \quad . \qquad (5.5)$$

By repeated application of such infinitesimal transformations , a finite transformation can be generated in which the parameters τ are altered from τ_1 to τ_2 along a definite path . The appropriate action operator is the sum of those for the individual transformations ,

$$W_{12} = \sum_{\tau_2}^{\tau_1} W_{\tau, \tau - d\tau}$$

where, according to (5.3),

$$W_{\tau, \tau - d\tau} = \sum_a p_a(\tau) \left(q_a(\tau) - q_a(\tau - d\tau) \right) + \sum_a G_{(r)} \left(q(\tau - d\tau), p(\tau), \tau \right) d\tau_r ,$$

Hence,

$$W_{12} = \int_{\tau_2}^{\tau_1} \left(\sum_a p_a \, dq_a + \sum_r G_{(r)} \, d\tau_r \right)$$

is the action operator generating the finite canonical transformation, in which the operators referring to all the values of τ intermediate between τ_1 and τ_2 appear as superfluous variables. Since W_{12} must be independent of these intermediate variables, it is stationary with respect to infinitesimal variations of all dynamical quantities that do not refer to the terminal values of τ. Now

$$\delta \left[\sum p_a \, dq_a + \sum_r G_{(r)} d\tau_r \right] - d \left[\sum p_a \, \delta q_a + \sum_r G_{(r)} \delta \tau_r \right]$$
$$= \sum_a \left(\delta p_a \, dq_a - dp_a \delta q_a \right) + \sum_r \left(\delta G_{(r)} \, d\tau_r - dG_{(r)} \, \delta \tau_r \right) ,$$

which in turn equals

$$\sum_a \delta p_a \left(dq_a + \sum \frac{\partial_1 G}{\partial p_a} \, d\tau_r \right) - \sum_a \left(dp_a - \sum_r \frac{\partial_r G_{(r)}}{\partial q_a} d\tau_r \right) \delta q_a$$
$$+ \sum \left(\frac{\partial G_{(s)}}{d\tau_r} - \frac{d G_{(r)}}{d\tau_s} \right) \delta \tau_r \, d\tau_s \qquad (5.6)$$

and thus the stationary requirement, applied to a given parameter path

($\delta\tau_r = 0$) again yields the differential equations (5.5) , and

$$\delta W_{12} = G_1 - G_2$$

$$G = \sum p_a \delta q_a + \sum G_{(r)} \delta\tau_r$$

With a single parameter , $\tau_{(1)} = t$ and $G_{(1)} = - H$, we regain the original action principle , now appearing as the characterization of a canonical transformation from a description at time t ($= t_1$) to the analogous one referring to another time t_o ($= t_2$) , which we call the Hamilton-Jacobi transformation . With t_o held fixed , the action operator W ($= W_{12}$) obeys

$$\delta W = \sum p_a \delta q_a - H\delta t - \sum p_{ao} \delta q_{ao}$$

or

$$p_a = \frac{\partial_r W}{\partial q_a} \quad , \quad - p_{ao} = \frac{\partial_r W}{\partial q_{ao}} \quad , \quad H + \frac{\partial W}{\partial t} = 0 \ .$$

On comparison with (5. 2) we recognize that the Hamilton-Jacobi canonical transformation is such that $\overline{H} = 0$, which expresses the lack of dependence of the new dynamical variables $q (t_o)$, $p (t_o)$ on the time t . Of course , the new Hamiltonian at time t is not the same as H , evaluated at time t_o , which governs the dependence of W upon the parapeter t_o ,

$$\frac{\partial W}{\partial t_o} = H (q (t_o) \quad , \quad p (t_o) \ , t_o) \ .$$

For a canonical transformation involving several parameters , the last term of (5.6) displays the effect of altering the path along

which the parameters evolve . According to the significance of $G_{(s)}$ as an infinitesimal generator ,

$$\frac{d\, G_{(r)}}{d\, \tau_s} = \frac{\partial\, G_{(r)}}{\partial\, \tau_s} - \frac{1}{i} \left[G_{(r)} \; , G_{(s)} \right]$$

and therefore ,

$$\frac{\partial\, G_{(s)}}{\partial\, \tau_r} - \frac{d\, G_{(r)}}{d\, \tau_s} = R_{rs}$$

is antisymmetrical with respect to, the indices r and s ,

$$R_{rs} = \frac{\partial}{\partial \tau_r} G_{(s)} - \frac{\partial}{\partial \tau_s} G_{(r)} + \frac{1}{i} \left[G_{(r)} \; , G_{(s)} \right]$$

$$= - R_{sr}$$

The complete variations of W_{12} is thus

$$\delta\, W_{12} = G_1 - G_2 + \int_{\tau_2}^{\tau_1} \sum R_{rs} \; \frac{1}{2} \left(\delta \tau_r \, d\tau_s - \delta \tau_s \, d \tau_r \right) \qquad (5.7)$$

If the canonical transformation is to be independent of the integration path it is necessary that

$$R_{rs} = 0 \, .$$

The simplest situation , for which the integration path is significant , has the structure of the special canonical group ; the operators $G_{(r)}$ are independent of the parameters and every commutator $G_{(r)} \; ; \; G_{(s)}$ commutes with each generator $G_{(q)}$. In consequence ,

$$\frac{d}{d\, \tau_q} \; R_{rs} = 0$$

and an infinitesimal change of path , with fixed terminal conditions , yields

$$\delta W_{12} = \sum_{r<s} \frac{1}{i} \left[G_{(r)} , G_{(s)} \right] \delta \int \tau_r d\tau_s$$

The action operators associated with two different paths are thus related by

$$(W_{12})_{I} = (W_{12})_{II} + \sum_{r<s} \frac{1}{i} \left[G_{(r)} , G_{(s)} \right] \int_{I-II} \tau_r d\tau_s \quad (5.8)$$

in which the individual integrals represent the area included between the two curves in the various planes of the parameter space . The general condition for such an integrability of the differential form (5.7),with fixed terminal conditions , is

$$\frac{d R_{rs}}{d\tau_q} + \frac{d R_{qr}}{d\tau_s} + \frac{d R_{sq}}{d\tau_r} = 0$$

which is always satisfied , in virtue of the Jacobi identity , for

$$\frac{d}{d\tau_q} F = \left[\frac{\partial}{\partial \tau_q} - i G_{(q)} , F \right]$$

$$R_{rs} = i \left[\frac{\partial}{\partial \tau_r} - i G_{(r)} , \frac{\partial}{\partial \tau_s} - i G_{(s)} \right]$$

A particular path in the parameter space is the straight line characterized by the variation of the single parameter λ from 0 to 1 ,

$$d\tau_r = \tau_r d\lambda \quad , \qquad \tau = \tau_1 - \tau_2 ,$$

as described by the generator ,

$$G_{(\lambda)} = \sum_{r} G_{(r)} \tau_r .$$

If the $G_{(r)}$ are independent of the parameter τ , $G_{(\lambda)}$ remains constant during the evolution of λ , and the differential equation

$$\frac{\partial}{\partial\lambda}\langle\lambda| = i\langle\lambda|G(\lambda)$$

can be integrated to yield the finite transformation

$$\langle\tau_1| = \langle\tau_2| e^{i\sum G(r)\tau_r} \qquad (5.9)$$

along the straight line path . If only a single parameter is altered , the transformation is

$$\langle\tau_{r1}| = \langle\tau_{r2}| e^{i G(r)\tau_r},$$

which is illustrated by the time development of a conservative system ,

$$\langle t_1| = \langle t_2| e^{-i Ht} .$$

A variety of transformations characterized by paths formed from straight line segments parallel to the parameter axes can now be constructed . Thus for two parameters we have

$$\langle\tau_1\tau_2|_I = \langle 0\,\tau_2| e^{i G(1)\tau_1} e^{i G(2)\tau_2} = \langle 0| e^{i G(1)\tau_1} e^{i G(2)\tau_2}$$

as compared with

$$\langle\tau_1\tau_2|_{II} = \langle\tau_1 0| e^{i G(2)\tau_2} e^{i G(1)\tau_1} = \langle 0| e^{i G(2)\tau_2} e^{i G(1)\tau_1}$$

and the transformation (5.9) ,

$$\langle\tau_1\tau_2|_{III} = \langle 0| e^{i(G(1)\tau_1 + G(2)\tau_2)}$$

For the example in which $\left[G_{(1)}, G_{(2)}\right]$ commutes with $G_{(1)}$ and $G_{(2)}$, we infer from (5.8) that the action operators for the various paths are connected by

$$(W)_{III} = (W)_I + \frac{1}{i}\left[G_{(1)}, G_{(2)}\right] \frac{1}{2}\tau_1\tau_2$$

$$= (W)_{II} - \frac{1}{i}\left[G_{(1)}, G_{(2)}\right] \frac{1}{2}\tau_1\tau_2$$

which, as statements about the composition of successive transformations, imply that

$$e^{i(A+B)} = e^{iA} e^{IB} e^{\frac{1}{2}[A,B]} = e^{iB} e^{iA} e^{-\frac{1}{2}[A,B]}$$

if $[A,B]$ commutes with A and B.

We now consider the construction of a group of finite canonical transformation, where a transformation must be specified completely by the values of the parameters and thus is independent of the integration path. Apart from the elementary situation in which the commutators of all infinitesimal generators vanish (Abelian group), the operators $G_{(r)}$ must be explicit functions of the parameters if all operators R_{rs} are to be zero. We shall express these generators as linear combinations of another set, $G_{(a)}$, which do not depend explicitly upon the parameters, and obey the commutation relation

$$\left[G_{(b)}, G_{(c)}\right] = \sum_a G_{(a)} g_{abc} \qquad (5.10)$$

where the numerical coefficients g_{abc} satisfy relations, implied by the Jacobi identity, which are described by the remark that the array of numbers with fixed b,

$$g_b = (g_{abc}),$$

provide a matrix representation for $G_{(b)}$ and obey (5.10) as a matrix

equation . If the operators $G_{(b)}$ are Hermitian , the elements of the matrices g_b are imaginary , but these matrices are Hermitian only if g_{abc} is antisymmetrical in the three indices (it is necessarily antisymmetrical in b and c) . A group of the latter type with three parameters is isomorphic with the space rotation group . Now

$$G_{(r)} = \sum G_{(a)} \; \xi_{ar}(\tau),$$

which we also express by

$$\sum G_{(r)} \; d\tau_r = \sum G_a \; d\xi_a \qquad (5.11)$$

where the differentials

$$d\xi_a = \sum \xi_{ar}(\tau) d\tau_r \qquad (5.12)$$

are not generally integrable . The requirement of path independence for the finite transformations is now expressed by the differential equations ,

$$\frac{\partial}{\partial \tau_r}\xi_{as} - \frac{\partial}{\partial \tau_s}\xi_a = i\sum_{bc} g_{abc}\,\xi_{br}\,\xi_{cs} \qquad (5.13)$$

and the integrability conditions for these equations are automatically satisfied in virtue of the commutation properties of the matrices g_b .

A particularly useful set of parameters is defined in the following way . As λ varies from 0 to 1 , let a point in the τ - parameter space move out from the origin along the curve described by

$$\sum \xi_{ar}(\tau)\, d\tau_r = t_a \; d\lambda \qquad (5.14)$$

where the t_a are arbitrary constants . The point in the τ - space that

is reacged at $\lambda = 1$ is determined by the numbers t_a , which constitute a new set of parameters . The functions

$$\xi_{ab}(t) = \sum \xi_{ar}(\tau) \frac{\partial \tau_r}{\partial t_b}$$

obey equations of the same form as (5.I3) , and the inexact differential (5.I2) which we now call δt_a , appears as

$$\delta t_a = \sum \xi_{ab}(t)\, dt_b$$

If the parameters t_a are replaced by $(1 + \varepsilon)\, t_a$, the curve defined by (5.I4) is simply extended beyond the terminal point with parameters t_a to $(1 + \varepsilon)\, t_a$. Hence from the integrated form of (5.I4) , we obtain

$$\int_t^{(1+\varepsilon)t} \sum \xi_{ab}\left(\frac{1}{t}\right) dt_b = \varepsilon t_a$$

or , for infinitesimal ε ,

$$\sum \xi_{ab}(t)\, t_b = t_a \qquad (5.I5)$$

If we consider a straight line path in the t space ,

$$d\, t_a = d\, \lambda\, t_a \ ,$$

at the intermediate point λt_a , we have

$$\delta t_a = \sum \xi_{ab}(\lambda t)\, t_b\, d\lambda = dt_a \ ,$$

and therefore the infinitesimal generator (5.II) reads

$$\sum G_{(a)}\, t_a\, d\lambda = G_{(\lambda)}\, d\lambda$$

We conclude from our previous discussion that the finite transformations of the group ,

$$\langle t \mid = \langle 0 \mid U \ (t) \ , \qquad (5.16)$$

have the exponential form ,

$$U \ (t) \ = \ e^{\, i \sum G_{(a)} \ t_a}$$

in terms of the t parameters .

The relation (5.15) permits the differential equations for the functions $\xi_{ab}(t)$ to be given a linear form . Indeed

$$\sum t_c \left(\frac{\partial}{\partial t_c} \xi_{ab} - \frac{\partial}{\partial t_b} \xi_{ac} \right) = i \sum g_{ade} \ \xi_{dc} \ t_c \ \xi_{eb}$$

becomes

$$\left(\sum t_c \frac{\partial}{\partial t_c} + 1 \right) \xi_{ab} - \delta_{ab} = i \sum g_{ade} \ t_d \ \xi_{eb}$$

and , on applying this equation at the point λt , we obtain

$$\frac{\partial}{\partial \lambda} \lambda \xi_{ab}(\lambda t) - i \sum g_{ade} \ t_d \ \lambda \xi_{eb} (\lambda t) \ = \ \delta_{ab}$$

In a matrix notation this reads

$$\left(\frac{\partial}{\partial \lambda} \ - i \sum g_a \ t_a \right) \ X (\lambda) \ = \ 1$$

where the functions $\xi_{ab}(t)$ are obtained as the element of the matrix $X (1)$, and $X (0) = 0$. The formal solution is

$$X (1) = \int_0^1 d \lambda \, e^{\, i \lambda \sum g_a \ t_a} \ = \frac{e^{\, i \sum g_a t_a} - 1}{i \sum g_a \ t_a} \ = \ \xi (t)$$

and

$$\delta \ t \ = \ \xi (t) \ \ d \, t \ .$$

An infinitesimal alteration of the t parameters produces an infinitesimal vector change given by

$$d\langle t \mid = i \langle t \mid \sum G_a(t)\, \delta t_a(t) \qquad (5.17)$$

in which the operators $G_a(t)$ refer implicitly to the parameters t. Now the connection between operators and vectors associated with the parameter values t and 0 is described by $U(t)$, in accordance with (5.16) and

$$G_a(t) = U(t)^{-i} G_a U(t),$$

so that a statement equivalent to (5.17) is

$$d\,U(t) = i \sum G_a\, \delta t_a\, U(t) \qquad (5.18)$$

or

$$U(t) + d\,U(t) = \left[1 + i \sum G_a\, \delta t_a\right] U(t), \qquad (5.19)$$

which illustrates the multiplication of operators for successive transformations. More generally, we have

$$U(t_1) = U(t_2)\, U(t_3)$$

where

$$t_1 = t_1(t_2, t_3) \qquad (5.20)$$

expresses the multiplication properties of the group. An infinitesimal change of t_2 induces a change of t_1 in accordance with (5.20). But (5.19) asserts that

$$\left[1 + i \sum G_a\, \delta t_a(t_1)\right] U(t_1) = \left[1 + i \sum G_a\, \delta t_a(t_2)\right] U(t_2)\, U(t_3),$$

or

$$\delta t_a (t_1) = \delta t_a (t_2) \quad .$$

The ensuing differential forms

$$\sum \xi_{ab} (t_1) \; d\, t_{1b} = \sum \xi_{ab} (t_2) \; d\, t_{2b}$$

together with the initial condition

$$t_2 = 0 \qquad : \qquad t_1 = t_3$$

serve to determine the composition property of the group parameters .

An altermative construction is obtained by differentiating with respect to t_3 . The essential infinitesimal form is

$$U (t) + d\, U (t) = U (t) \left[1 + i \sum G_a \delta_r t_a \right]$$

where $\delta_r t_a$ are a set of inexact differentials to be distinguished from δt_a , which might now be called $\delta_\ell t_a$. The structure of these new quantities is obtained on replacing t with $- t$ in (5.I8) and remarking that

$$U(-t) = U (t)^{-1} \quad ,$$

which yields

$$\delta_r t_a = \sum \xi_{ab} \; (-t) \; d\, t_b \quad ,$$

with

$$\xi (-t) = e^{-i \sum g_a t_a} \; \xi (t) \quad .$$

We now deduce from

$$\delta_r t_a (t_1) = \delta_r t_a (t_3)$$

the differential forms

$$\sum \xi_{ab}(-t_1)\, d\, t_{1b} = \sum \xi_{ab}(-t_3)\, d\, t_{3b}$$

which are supplemented by the initial condition

$$t_3 = 0 \quad : \quad t_1 = t_2 \quad .$$

We have remarked on the conservation property which attends the generators of infinitesimal canonical transformations that leave invariant the structure of the Hamiltonian operator. It is worth noting that there exist certain canonical transformations that preserve the form of the Hamiltonian operator for an arbitrary dynamical system , and thus provide constants of the motion of a more general type . Such quantities can be called kinematical rather than dynamical constants of the motion . As an example , consider the rotation of the coordinate system about an axis through the angle $\pm 2\pi$, which cannot alter the structure of H , whether or not the system is isolated . The operator that describes a rotation through the angle φ about the k^{th} axis is

$$U_k(\varphi) = e^{i\varphi J_k}$$

and thus

$$U_k(\pm 2\pi) = e^{\pm 2\pi i J_k}$$

obeys

$$U_k(\pm 2\pi)^{-1} H U_k(\pm 2\pi) = H$$

for arbitrary H . Hence $U_k(\pm 2\pi)$ remains constant despite any modification of the dynamical properties of the system . Now , from the effect , on the components of a vector of rotation about an axis through

the angle $\pm \pi$, we conclude that , with $i \neq k$,

$$U_i (\mp \pi) \; U_k (\varphi) \; U_i (\pm \pi) = U_k \; (- \varphi)$$

and alternative evaluations of

$$U_i (\pi) \; U_k (\mp \pi) \; U_i (- \pi) \; U_k (\pm \pi)$$

yield

$$U_i (2 \pi) = U_k \; (\pm 2 \pi) \; .$$

Hence

$$U_i (2 \pi) = U_k (2 \pi) \quad , \quad U_k (2 \pi) = U_k (2 \pi)^{-1} ,$$

and the possible eigenvalues of $U_k (2\pi)$ are ± 1 , which must be the same for the operators referring to any axis . The eigenvalues of every component of J are thus either integers , or half-integers, and no external disturbance can produce a transition between the two categories , in consonance with the specification of the structure of J by the type and number of the dynamical variables of the system .

As a second example , we recall that the Hamiltonian operator must be an even function of the variables of the second kind . Hence,the Hamiltonian for any conceivable system must be invariant under the transformation

$$\bar{X}_\alpha = - X_\alpha$$

or, with canonical variables ,

$$\bar{q}_\alpha = - q_\alpha \quad , \quad \bar{q}^+_\alpha = - q^+_\alpha \quad .$$

Now this finite canonical transformation can be constructed from the infinitesimal transformation ,

$$\delta q_\alpha = - i \delta \varphi \, q_\alpha \quad , \quad \delta q^+_\alpha = i \delta \varphi \, q^+_\alpha \quad ,$$

which has the generator

$$G = \delta\varphi \sum q_\alpha^+ q_\alpha = \delta\varphi\, G(\varphi)$$

according to

$$\delta q_\alpha = \frac{\partial_l G}{\partial\, i q_\alpha^+} \quad , \quad \delta\, i q_\alpha^+ = -\frac{\partial_r G}{\partial q_\alpha} \quad .$$

Indeed , the differential equations ,

$$\frac{\partial}{\partial\varphi} q_\alpha = i\, q_\alpha \quad , \quad \frac{\partial}{\partial\varphi} q_\alpha^+ = -i\, q_\alpha^+$$

are integrated to

$$q_\alpha(\varphi) = e^{i\varphi} q_\alpha \quad , \quad q_\alpha^+(\varphi) = e^{-i\varphi} q_\alpha^+$$

and the desired transformation is obtained with $\varphi = \pm\pi$. Consequently the operator

$$e^{\pm\pi i \sum q_\alpha^+ q_\alpha} \qquad (5.21)$$

is a kinematical constant of the motion . Since the eigenvalues of $q_\alpha^+ q_\alpha$ are integers , 0 , 1 , the two forms are identical , and the eigenvalues of (5.21) are ± 1 , according as $\sum q_\alpha^+ q_\alpha$ possesses an even or an odd integral value . It is to be noted that

$$e^{\pi i \sum q_\alpha^+ q_\alpha} = e^{\pi i n_2}\, e^{\pi i \sum q_\alpha q_\alpha^+}$$

which again shows that the symmetry between q_α and q_α^+ is complete only if n_2 is even .

6. - CANONICAL TRANSFORMATION FUNCTIONS

According to the significance of the action operator $W(q, \bar{q}, t)$ that defines a canonical transformation at time t, infinitesimal alterations of the eigenvalues and of t produce a change in the canonical transformation function $\langle q't \,|\, \bar{q}'t \rangle$ given by

$$\delta \langle q' t \,|\, \bar{q}' t \rangle = i \langle q' t \,|\, \delta W(q, \bar{q}, t) \,|\, \bar{q}' t \rangle \quad . \quad (6.1)$$

For some transformations, the commutation properties of the q and $\bar{q}$ variables can be used to rearrange the operator δW so that the q's everywhere stand to the left of the $\bar{q}$'s. This ordered differential expression will be denoted by $\delta \mathcal{W}(q, \bar{q}, t)$, and

$$\delta W(q, \bar{q}, t) = \delta \mathcal{W}(q, \bar{q}, t) \quad .$$

From the manner of construction of the latter operator, the variables q and $\bar{q}$ act directly on their eigenvectors in (6.1), and this equation becomes

$$\delta \langle q' t \,|\, \bar{q}' t \rangle = i \, \delta \mathcal{W}(q', \bar{q}, t) \langle q' t \,|\, \bar{q}' t \rangle \quad .$$

Hence $\delta \mathcal{W}$ must be an exact differential, and integration yields

$$\langle q' t \,|\, \bar{q}' t \rangle = e^{i \mathcal{W}(q'; \bar{q}', t)} \quad ,$$

in which a multiplicative integration constant is incorporated additively in $\mathcal{W}$. This constant is fixed, in part, by the composition properties of the transformation functions. It is to be emphasized that the ordered operator $\mathcal{W}$ does not equal W, and, indeed, is not Hermitian, should

W possess this property . We have in effect already illustrated the ordering method by the construction of the transformation function $\langle q' | p' \rangle$. For Hermitian variables of the first kind , for example ,

$$W\ (q\ ,\ p)\ = \sum \tfrac{1}{2} \left\{ q_k\ ,\ p_k \right\}$$

whereas

$$\mathcal{W}(q\ ,\ p)\ = \sum q_k\ p_k + i\ \frac{n}{2}\ \log\ 2\pi$$

$$= W\ (q\ ,\ p\ ) + i\ \frac{n}{2}\ (\log 2\pi + 1)$$

The transformation function for an arbitrary infinitesimal canonical transformation is easily constructed in this way , if one uses the action operator

$$W(p\ ,\ \bar{q}\ ) = W(p,\ q\ ) + W(q\ ,\ \bar{q}\ )$$

$$= -\sum_a p_a\ \bar{q}_a - G(\bar{q}\ ,\ p\ )$$

for $$\delta W\ (p\ ,\ \bar{q}\ ) = -\sum (\delta p_a\ \bar{q}_a \delta \bar{q}_a) - \delta G\ (\bar{q}\ ,\ p\ )$$

is brought into the ordered form by performing the required operation on the infinitesimal quantity δ G with the aid of the known commutation relations between $\bar{q} = q$ and p . It is convenient to supplement G with a numerical factor λ so that we obtain a one-parameter family of transformations that includes the desired one ($\lambda = 1$) and the identity transformation ($\lambda = 0$) . Then

$$\delta(\lambda G) = \delta\lambda G + \lambda \delta G\ ,$$

and the ordering operation is to be applied to G as well as to δ G

From the former we obtain an equivalent operator which we call $G(p;\bar{q})$, and thus

$$\delta W = -\delta\left[\sum p_a \bar{q}_a\right] - \delta\lambda G(p;\bar{q}) - \lambda\delta G .$$

The integrability of the ordered operator δW now demands that the ordered version of δG be simply the variations of $G(p;\bar{q})$ and thus ($\lambda = 1$)

$$W(p;\bar{q}) = -\sum p_a \bar{q}_a - G(p;\bar{q}) + \text{Const} .$$

in which the additive constant is that appropriate to th. identity transformation $W(p;q)$ and depends ipon the kinds of variables employed. Hence

$$\langle p' | \bar{q}' \rangle = e^{iW(p';\bar{q}')} = C\, e^{-i\sum p'_a q'_a - iG(p';q')}, \quad (6.2)$$

or

$$\langle p' | q' \rangle = \langle p' | q' \rangle\, e^{-iG(p';q')}$$

and in view of the infinitesimal nature of G,

$$\langle p' | \bar{q}' \rangle = \langle p' | q' \rangle \left[1 - iG(p';q')\right] = \langle p' | 1 - iG | q' \rangle,$$

which simply repeats the significance of G as the generator of the infinitesimal transformation. When the transformation corresponds to an infinitesimal change of parameters, (6.2) reads

$$\langle p'\tau | q'\tau - d\tau \rangle = C\, e^{-i\sum p'_a q'_a + i\sum d\tau_r G_{(r)}(p';q',\tau)} \quad (6.3)$$

and in particular,

$$\langle p'\, t | q'\, t - dt \rangle = C\, e^{-i\sum p'_a q'_a - i\,dt\,H(p';q',t)} .$$

The composition properties of transformation functions can now be used to introduce other choices of canonical variables . For example , with Hermitian variables of the first kind ,

$$\langle q' \mid \bar{q}'' \rangle = \int \langle q' \mid p' \rangle \, dp' \, \langle p' \mid \bar{q}'' \rangle$$

$$= \int \frac{dp'}{(2\pi)^n} \, e^{\, i \sum (q'_k - q''_k) \, p'_k \, - \, i \, G \, (p' \, ; \, q'')}$$

according to (4.5) and (4. 6) . The integrations are easily performed if G is a linear function of the p variables , or a non singular quadratic function . In the first situation

$$G \ (p' \, ; \, q'') = G (0 \, ; \, q'') + \sum p'_k \ \frac{\partial G \ (p' \, ; q'')}{\partial \, p'_k}$$

and $$\langle q' \mid q'' \rangle = e^{\, - \, i \, G \, (0 \, ; \, q'')} \, \delta \, (q' - q'' - \frac{\partial G}{\partial \, p'} \,). \quad (6.4)$$

For the quadratic function of p , it is convenient to move the origin of the p' variables to the point defined by

$$q'_k \ - q''_k \ = \frac{\partial \, G \ (p \, ; \, q'')}{\partial \, p_k} \qquad (6.5)$$

Then ,

$$\langle q' \mid \bar{q}'' \rangle = e^{\, i \, w} \int \frac{dp'}{(2\pi)^n} \ e^{\, - \frac{i}{2} \sum p'_k \, p'_l \frac{\partial^2}{\partial p_k \, \partial p_l}}$$

$$= \left[(2\pi i)^n \ \det \frac{\partial^2 G}{\partial \, p \, \partial p} \right]^{-\frac{1}{2}} e^{\, i \, w}$$

where

$$w \ = \sum (q'_k - q''_k) \, p_k - G \ (p \, ; \, q'') = \sum p_k \ \frac{\partial \, G(p \, ; q'')}{\partial \, p_k} \ - G(p \, ; q'').$$

On eliminating p with the aid of (6.5) , we find

$$w = \tfrac{1}{2} \sum \, (\, q'_k - q''_k - \frac{\partial}{\partial \, p_k} \, G \ (\, 0 \, , \, q'') \ M_{kl} \ (\, q'_l - q''_l - \frac{\partial}{\partial \, p_l} \, G \ (\, 0 , q'') \, -$$

$$- \ G \ (\, 0 \, , \, q'' \,)$$

in which $M(q'')$ is the matrix inverse to $\partial^2 G / \partial p \, \partial p$. In the limit as the latter approaches zero , the delta function form (6.4) is obtained . Both results are also easily derived directly , without reference to the intermediate p representation . For the example of an infinitesimal Hamilton-Jacobi transformation in which H meets the requirement of a non-singular quadratic dependence on p , we have

$$\left\langle q't \,\middle|\, q'' \, t - dt \right\rangle = \left\langle q' \, t + dt \,\middle|\, q'' \, t \right\rangle$$
$$= \left(\det \frac{m}{2\pi i dt} \right)^{\frac{1}{2}} e^{i\omega}$$

where m is the matrix inverse to $\partial^2 H / \partial p \, \partial p$ and

$$\omega = dt \left[\frac{1}{2} \sum_{k\ell} \left(\frac{q'_k - q''_k}{dt} - \frac{\partial}{\partial p_k} H(0 ; q'' , t) \right) m_{kl} \left(\frac{q'_l - q''_l}{dt} - \frac{\partial}{\partial p_l} H(0 ; q'' , t) \right) - H(0 ; q'' , t) \right] \qquad (6.6)$$

An equivalent form for the latter is

$$\omega = dt \left[\frac{1}{2} \sum_{k\ell} \frac{q'_k - q''_k}{dt} m_{kl} \frac{q'_l - q''_l}{dt} + \sum_k \frac{q'_k - q''_k}{dt} p_{ok}(q'', t) - H(p_o ; q'' , t) \right] ,$$

in which p_o is defined by

$$\frac{\partial}{\partial p_k} H(p_o ; q'' , t) = 0 .$$

The transformation function for a finite canonical transformation can be constructed by the repeated composition of that for an infinitesimal transformation . To evaluate $\langle \tau_1 | \tau_2 \rangle$ along a particular path in the parameter space , we choose N intermediate points on the path , and compute

$$\langle \tau_1 | \tau_2 \rangle = \langle \tau_1 | \tau^{(1)} \rangle \times \langle \tau^{(1)} | \tau^{(2)} \rangle \times \ldots \times \langle \tau^{(N)} | \tau_2 \rangle \qquad (6.8)$$

in which X signifies the application of a method of composition appropriate to the canonical variables employed . In the limit $N \longrightarrow \infty$ the required path is traced as an infinite sequence of infinitesimal transformations . With Hermitian variables of the first kind and composition by integration , for example , we begin with

$$\langle q' \tau | q'' \tau - d\tau \rangle = \int \frac{d p'}{(2\pi)^n} \; e^{\, i \sum (q'_k - q''_k) p'_k + i \sum d\tau_r G_{(r)} (p' ; q'' , \tau)}$$

and obtain

$$\langle q' \tau_1 | \; q'' \tau_2 \rangle =$$

$$= \lim_{N \to \infty} \int \prod_{j=1}^{N+1} \left[\frac{dq^{(j)} dp^{(j)}}{(2)^n} \; e^{\, i \sum (q_k^{(j-1)} - q_k^{(j)}) p_k^{(j)} + i \sum (\tau_r^{(j-1)} - \tau_r^{(j)})} \right.$$

$$\left. G_{(r)} (p^{(j)} , q^{(j)} , \tau^{(j)}) \right] \delta (q^{(N+1)} - q'') ,$$

in which

$$\tau^{(0)} = \tau \quad , \quad \tau^{(N+1)} = \tau_2 \quad , \quad q^{(0)} = q'$$

and the delta function factor enforces the restriction $q^{(N+1)} = q''$. For the Hamiltonian-Jacobi transformation with a Hamiltonian that depends quadratically on p , this general form reduces to

$$\langle q' \; t' \, | \, q'' \; t_2 \rangle =$$

$$= \lim_{N \to \infty} \int \prod_{j=1}^{N+1} \left[dq^{(j)} \left(\det \frac{1}{2\pi i} \frac{\partial^2 w}{\partial q^{(j-1)} \partial q^{(j-1)}} \right)^{\frac{1}{2}} \cdot e^{\, i w (q^{(j-1)} t^{(j-1)} , \; q^{(j)} t^{(j)})} \right]$$

$$\delta (q^{(N+1)} - q'')$$

in which w is given by (6.6) or (6.7) . The generally applicable

technique of composition by differentiation can be applied directly to the infinitesimal transformation function (6. 3) and we obtain $\langle p'\tau_1 | q'\tau_2 \rangle =$

$$= \lim_{N\to\infty} \prod_{j=1}^{N+1} \left[\langle p^{(j-1)}\tau^{(j-1)} \left| i\frac{\partial \mathcal{L}}{\partial \rho^{(j)}} \cdot \tau^{(j)} \right\rangle \right] e^{-i\sum_a p_a^{(N+1)} q'_a} \Big|_{p^{(j)} = \cdots = p^{(N+1)} = 0}$$

$$= \lim_{N\to\infty} \prod_{j=1}^{N+1} \left[e^{\sum_a p_a^{(j-1)} \frac{\partial \mathcal{L}}{\partial \dot{\tau}_a^{(j)}} + i\sum_r (\tau_r^{(j-1)} - \tau_r^{(j)}) G_{(r)}\left(p^{(j-1)}; i\frac{\partial \mathcal{L}}{\partial \rho^{(j)}}, \tau_r^{(j)}\right)} \right] \times$$

$$\times e^{-i\sum_a p_a^{(N+1)} q'_a} \Big|_{p^{(j)} = \cdots = p^{(N+1)} = 0} \qquad (6.9)$$

apart from numerical factors of $(2\pi)^{n/2}$ for Hermitian variables of the first kind . The boundary values are

$$\tau^{(0)} = \tau_1 \quad , \quad \tau^{(N+1)} = \tau_2 \quad , \quad p^{(0)} = p'$$

and the sense of multiplication is that of (6.8)

On writing (6.3) as

$$\langle p'\tau | q'\tau - d\tau = \langle p' | q' \rangle \, e^{i\sum d\tau_r G_{(r)}} \; (p' \; ; q', \tau)$$

$$= \langle p' \left| e^{i\sum d\tau_r G_{(r)}} (q , p, \tau) \right| q' \rangle \qquad (6.10)$$

the transformation function is obtained as a matrix on which the states and canonical variables do not depend upon τ . The result of composing successive transformations then appears as the matrix of the product of the associated operators, and thus $\langle \tau_1 | \tau_2 \rangle =$

$$= \lim_{N\to\infty} \left\langle \left| \prod_{j=1}^{N+1} e^{i\sum_r (\tau_r^{(j-1)} - \tau_r^{(j)}) G_{(r)}(x, \tau^{(j)})} \right| \right\rangle =$$

$$= \left\langle \left| \left(e^{i\int_{\tau_2}^{\tau_1} \sum d\tau_r G_{(r)}(x,\tau)} \right)_+ \right| \right\rangle ,$$

which defines the ordered exponential operator . The $+$ subscript refers to the manner of multiplication , in which the positions of the operators correspond to the path in parameter space as deformed into straight line ,

with τ_2 on the right and τ_1 on the left . We speak of negative rather than positive ordering if the sense of multiplication is reversed , so that

$$\left(e^{i\int_{\tau_3}^{\tau_1}\sum d\tau_r\, G_{(r)}}\right)_+ = \left(e^{-i\int_{\tau_1}^{\tau_2}\sum d\tau_r\, G_{(r)}}\right)_-$$

When the generators are independent of τ , and the path is a straight line, the ordering is no longer significant and

$$\langle \tau_1 | \tau_2 \rangle = \langle | \, e^{i\sum \tau_r G_{(r)}} | \rangle \qquad (6.II)$$

in agreement with (5.9) . Should the operators $G_{(r)}$ commute , thus generating an Abelian group , the exponential form (6.II) is valid without reference to the integration path . Let us suppose that the generators constitute , or can be extended to be a complete set of commuting Hermitian operators. Then a G representation exists , and a canonical transformation function describing the transformations of the group , say $\langle p'\tau_1 | q'\tau_2 \rangle$, can be exhibited as

$$\langle p'\tau_1 | q'\tau_2 \rangle = \langle p' | \, e^{i\sum \tau_r G_{(r)}} | q' \rangle$$
$$= \sum_{G'} \langle p' | G' \rangle \, e^{i\sum \tau_r G'_{(r)}} \langle G' | q' \rangle ,$$

which shows how the canonical transformation function serves to determine the eigenvalues of the operators G as well as wave functions representing the G states .

An elementary example for $n = 1$ is provided by the Hamiltonian operator

$$G = -i\lambda pq = \lambda q^+ q$$

in which the non-Hermitian canonical variables may be of either type . A direct construction of the canonical transformation function can be obtained from (6.9) by repeated application of

$$e^{(1+i\lambda dt)\, p^{(j-1)} \frac{\partial \ell}{\partial p^{(j)}}} e^{(1+i\lambda dt) p^{(j)} \frac{\partial \ell}{\partial p^{(j+1)}}} \Big|_{p^{(j)} = 0}$$

$$= e^{(1+i\lambda dt)^2\, p^{(j-1)} \frac{\partial \ell}{\partial p^{(j+1)}}} ,$$

namely ,

$$\langle q^{+\prime} \tau_1 \,|\, q' \tau_2 \rangle = e^{i\lambda \tau_{q^{+\prime} q'}}$$

$$= \sum_{n=0}^{\infty} \frac{(q^{+\prime})^n}{(n!)^{\frac{1}{2}}} e^{in\lambda\tau} \frac{(q')^n}{(n!)^{\frac{1}{2}}} \qquad (6.12)$$

The eigenvalues thus obtained are

$$(q^{+} q)' = n = 0, 1, \ldots$$

and wave functions of these states are given by

$$\langle q^{+\prime} \,|\, n \rangle = \frac{(q^{+\prime})^n}{(n!)^{\frac{1}{2}}} , \qquad \langle n \,|\, q' \rangle = \frac{(q')^n}{(n!)^{\frac{1}{2}}}$$

The derivation applies equally to both types of variable . However , for variables of the second kind $(q')^2 = (q^{+\prime})^2 = 0$, and the summation in (6.12) terminates after $n = 0, 1$, which are indeed the only eigenvalues of $q^{+}_{\alpha} q_{\alpha}$.

A more general ordered operator form appears on decomposing the generators into two additive parts

$$G_{(r)} = G^{o}_{(r)} + G^{\cdot}_{(r)} .$$

Then (6.3) can be written

$$\langle p'\tau \mid q'\tau - d\tau \rangle_G = \langle p'\tau \mid q'\tau - d\tau \rangle_{G^0} e^{i\sum d\tau_r G'_{(r)}(p',q',\tau)}$$

$$= \langle p'\tau \mid e^{i\sum d\tau_r G'_r(x(\tau),\tau)} \mid q'\tau - d\tau \rangle_{G^0}$$

in which the subscripts indicate that the infinitesimal parameter change is governed by the generators G or G^0, respectively .

6. APPENDIX .

TRANSITION TO COMPLEX CONJUGATE ALGEBRA .

We consider one more general property of the transformation function : their behaviour when we pass to the complex conjugate algebra . This will be seen to be connected with the reversal of the sense of time .

Let $d\tau_r\, G_r(q,q^\dagger)$ be a set of generators of parameterized infinitesimal transformations where the variables are the non-Hermitian variables $q, q^\dagger$

$$\delta \langle q' \tau_1 / q'' \tau_2 \rangle = \langle q' T_1 | \int - \sum q_a^\dagger dq_a + i \sum_r d\tau_r G_r(q,q^\dagger) | q'' \tau_2 \rangle \tag{6.13}$$

$$\frac{\partial G_r}{\partial \tau} = 0$$

From our previous considerations of the dual algebra (p.11 seq.), we expect that (6.13) will have the same form in the complex conjugate algebra . However, we must examine in detail how to pass from one algebra to the other .

Consider a class of G_r such that the sign of G_r is either changed or left unchanged by taking the complex conjugate, or more generally , taking into account the linear translation of the q_a , such that :

$$G_+(r)\ (Uq, q^\dagger U^{-1})^* = G(r)_+ (q^*, q^{\dagger *})$$

$$G_-(r)\ (Uq, q^\dagger U^{-1})^* = -\, G_-(r)\ (q^*, q^{\dagger *}) .$$

Then to preserve the form of (6.13) on transition to the dual algebra, we must also make the transformation

$$\tau_{+1} \longrightarrow -\tau_{+1} ,$$

So that $\tau_{+2} \rightarrow -\tau_{+2}$

$$\delta \langle Uq', -\tau_{+1}\tau_{-1} \Big|^{*} Uq'', -\tau_{+2}\tau_{-2} \rangle^{*} =$$

$$= \langle Uq', -\tau_{+1}\tau_{-1} \Big|^{*} \int \Big(-\sum_{a} q_a^{\dagger *} \delta q_a^{*} + i \sum_{r} \delta\tau_r \tau(r) \Big| Uq'' -\tau_{+2}\tau_{-2}$$

The commutator for q_a^{*}, $q^{\dagger *}$ remains the same as for

$$q_a, q_a^{\dagger} \qquad \left\{ \left[q_a^{*}, q_b^{\dagger *} \right] \right\} = \delta_{ab} .$$

The consequence $p = i\, q^{\dagger}$ implies that the variables undergo the transformation

$$q \rightarrow q^{*}$$
$$p \rightarrow -p^{*} .$$

This is reflected in the commutation law

$$\left\{ [p, q] \right\} = -i$$

However, no i appears in the commutator for non-Hermitian variables .

The proper transition of one algebra to the other is then indicated in :

$$\begin{cases} q \\ q^{+} \end{cases} \rightarrow \begin{cases} q^{*} \\ q^{*\dagger} \end{cases} , \quad |q\tau_{+}\tau_{-}\rangle \rightarrow |U\, q'^{*} - \tau_{+}\rangle^{*} \qquad (6.14)$$

We may observe here a fundamental distinction between the $G_{+}(r)$ and $G_{-}(r)$. The operators G_{-} must always have eigenvalues in pairs , $\pm$ the same absolute value :

If : $G_{-}\psi = G'\psi$

then : $-G_{-}\psi^{*} = G'\psi^{*}$.

I.e. if there exists the eigenvector ψ belonging to the eigenvalue G' then the vector ψ^{*} is an eigenvector belonging to the eigenvalue $-G'$:

A Hamiltonian must belong to the $G+(1)$, since these must be a minimum energy . An angular momentum belongs to the $G-(3)$

$$J \times J = i J$$

implies : $(-J^{*}) \times (-J^{*}) = i(-J^{*})$.

From what we have just seen the invariance in passing to the dual algebra implies /invariance under reversal of sense of time (H belongs to class $G_{+}(1)$) .We may also derive further physical information :

The correspondence between algebras (cf .p.12) gives :

$$\langle q^{+\prime}\, t_1 \mid q''\, t_2 \rangle^{*} = \langle q^{+\prime}\, t_1 \mid^{*} q''\, t_2 \rangle^{*} \quad .$$

But

$$|q''t_2\rangle^{*} \rightarrow |U^{-1}q''^{*} - t_2\rangle$$

$$\langle q^{+\prime}\, t_1 |^{*} \rightarrow \langle q^{+\prime\,*} U - t_1 |$$

Hence

$$\langle q^{+\prime}\, t_1 \mid q''t_2 \rangle^{*} = \langle q^{+\prime\,*} U - t_1 \mid U^{-1} q''^{*} - t_2 \rangle \quad . \qquad (6.15)$$

On the other hand

$$\langle q^{+\prime}\, t_1 \mid q''\, t_2 \rangle^{*} = \langle q^{+''} t_2 \mid q' t_1 \rangle \qquad (6.16)$$

In (6.15) we have the transformation function for evolution in time through the interval $-t_1 + t_2$; in (6.16) for the same system evolving through the interval $t_2 - t_1$. The amplitude for the transition from state I to state 2 is the same

as from state 2 to state I with the same sense of time .

7. SYSTEMS OF IDENTICAL PARTICLES .

We suppose we have a system of n particles . Each of the particles has its own set of operators which we denote by x_i, the x_i meaning the whole set of commuting operators for the i th particle . If these operators are the same for each particle , the particles are kinematically similar . If in addition , the Hamiltonian H is a symmetric function of the $x_1, x_2, .., x_n$, the particles are dynamically similar and there is no physical means of distinguishing one particle from another .

Call the set of quantities which can be obtained in a physical measurement on the i th particle a_i . We have seen already the properties of the measurement algebra for the individual particles . In particular we will use here (1.4)

$$M_k(a'_k, a''_k)\, M_k(a'''_k, a''''_k) = (a''_k, a'''_k)\, M_k(a'_k a''''_k) \qquad (1.4)$$

We note here also that the measurement of the properties of one particle does not interfere with the measurements made on another at the same time :

$$\left[M_k(a'_k, a''_k), M_\ell(a'''_\ell, a^{IV}_\ell) \right] = 0$$

We now define the following single-particle operator

$$F^{(1)} = \sum_{k=1}^{n} f^{(1)}(\chi_k)$$

where it is to be noted that "single particle " does not mean that the operator

acts on a single particle , but that it is a sum of operators referring to single particles . We define also two-particle operators

$$F^{(2)} = \sum_{k<\ell} f^{(2)}(\chi_k, \chi_\ell)$$

which depend on the interactions of the particles in pairs .The form of an n-particle operator is clear . The $F^{(k)}$ have the property of being symmetrical in all the particles and do not distinguish the particles in any way .

We have seen in the development of the measurement algebra that operators are linear combinations of measurement symbols . In the present context we have :

$$f^{(1)}(x_k) = \sum_{a',a''} \langle a'_k \mid f^{(1)}(x_k) \mid a''_k \rangle \; M_k(a'_k, a''_k)$$

Since,however, the particles are identical, we may omit reference to k in the matrix element . Further, with the definition $M(a',a'') = \sum_{k=1}^{n} M_k(a'_k, a''_k)$

we have $$F^{(1)} = \sum_{a'a''} \langle a' \mid f^{(1)} \mid a'' \rangle \; M(a',a'') \qquad (7.1)$$

(The addition of measurement symbols was defined to signify measurements which select systems with any of the values contained in the sum . (p.1.)

We write $F^{(2)} = \frac{1}{2} \sum_{k \neq \ell} f^{(2)}(x_k, x_\ell)$ by the same process in two stages .

$$f^{(2)}(x_k, x_\ell) = \sum_k \langle a'_k \mid f^{(2)}(x_k, x_\ell) \mid a''_k \rangle \; M_k(a',a'')$$

where $\langle a'_k \mid f^2(x_k, x_\ell) \mid a''_k \rangle$ is a number as far as the k th particle is concerned, but still an operator for the th particle . Repeating the process, we obtain

$$f^{(2)}(\chi_k \chi_\ell) = \tfrac{1}{2} \sum_{k \neq \ell} \langle a'_k a'''_\ell \mid f^{(2)}(\chi_k, \chi_\ell) \mid a''_k a^{IV}_\ell \rangle M_k(a', a'') M_\ell(a''', a^{IV})$$

an equation that is symmetric in k and ℓ, and consequently in a', a''' and a'' and a^{IV}. In $F^{(2)}$ instead of summing with $k \neq \ell$, we may sum over k and ℓ independently and subtract the $k = \ell$ contribution. (Cf.(1.4).

$$F^{(2)} = \sum_a \langle a' a''' \mid f^{(2)} \mid a'' a^{IV} \rangle \left[M(a' a'') M(a''', a^{IV}) - \delta(a'', a''') M(a', a^{IV}) \right] \quad (7.2)$$

Using the symmetry in a', a''' and a'', a^{IV}, we may interchange a' and a''', a'' and a^{IV} in the measurement symbols and in the Kronecker symbols δ.

$$F^{(2)} = \sum_a \langle a', a''' \mid F^{(2)} \mid a'' a''' \rangle \left[M(a''', a^{IV}) M(a', a'') - \delta(a^{IV}, a') M(a''', a'') \right]$$

Subtraction of one form of $F^{(2)}$ from the other gives us the commutator for the $M(a', a'')$:

$$\left[M(a', a''), M(a''', a^{IV}) \right] = \delta(a'', a''') M(a', a^{IV}) - \delta(a', a^{IV}) M(a''', a'') \quad (7.3)$$

This is the <u>multiplication property</u> of the $M(a', a'')$ which could also be read directly from the physical meaning of the $M(a', a'')$.

The <u>adjoint property</u> $M(a', a'')^\dagger = M(a'', a')$ follows directly from the definition of $M(a', a'')$. The <u>completeness property</u> is

$$\sum_a M(a', a') = \sum_a \sum_k M_k(a', a') = \sum_k 1 = n$$

where n is the number of particles in the system.

In the measurement algebra of systems when the particles were not treated as identical, we introduced at this stage annihilation and creation

operators

$$M(0,a'') = \psi^{\dagger}(a'') \text{ and } M(a',0) = \psi(a')$$

so that :

$$M(a',a'') = \psi(a')\,\psi^{\dagger}(a'') \quad .$$

We now wish to interpret :

$$M(a',a'') = \sum_k M_k(a',a'')$$

as a two stage process in which a particle (i.e., any particle of the system) in the mode a'' is annihilated and created in the mode a' . (In identical particle systems , we shall use " state " for the whole system , " mode " for an individual particle) . As far as notation goes , history steps in . To conform to usual practice we must define

$\psi(a'')$ an operator that annihilates a particle in mode a'' ,

$\psi(a')^{\dagger}$ an operator that creates a particle in the mode a' .

Then

$$M(a',a'') = \psi(a')^{\dagger}\,\psi(a'') \quad .$$

In the non-identical particle discussion the introduction of annihilation and creation

operators required that we augment the number of states by the 0 state . But here the completeness relation gives

$$\sum_{a'} \psi(a')^{\dagger} \psi(a') = N$$

where N is the operator for the total number of particles . N is an Hermitian operator and has as eigenvalues any integer, since we may increase or diminish the number of particles by creation and annihilation .

We now derive commutation relations for N and ψ in terms of their properties already known to us. Let $\Psi(n)$ be the state with n particles . Then

$$N \Psi(n) = n \Psi(n)$$

$$\psi(a') \Psi(n) = \Psi(n-1)$$

$$N \psi(a') \Psi(n) = (n-1) \psi(a') \Psi(n)$$

or we can write

$$N \psi(a') \Psi(n) = \psi(a') (N-1) \Psi(n)$$

This gives us the operator equation

$$N \psi(a') = \psi(a') (N-1)$$

or

$$[\psi(a'), N] = \psi(a') \qquad (7.4)$$

The commutation relation $[\psi(a')^{\dagger}, N] = -\psi(a')^{\dagger}$ may be derived in the same way or by taking the adjoint of the first .

We now look for the commutation relations for the ψ and $\psi^{\dagger}$ among themselves .

$$\psi(a')^{\dagger}\ \psi(a'')^{\dagger}\ \Psi(n)$$

means a state with n particles to which two additional particles have been added, one in the mode a', the other in the mode a''. This cannot differ from

$$\psi(a'')^{\dagger}\ \psi(a')^{\dagger}\Psi(n)$$

since the only difference between the two states is the order in which the two particules are introduced ; this has no physical meaning .

The only difference can be in the phase factor which we write as C,

$$\psi(a')^{\dagger}\ \psi(a'')^{\dagger} = C\ \psi(a'')^{\dagger}\ \psi(a')^{\dagger} \qquad (7.5)$$

If we suppose that C is independent of the modes a',a'' (see the appendix to this section for a more careful discussion of this point),we may write also by reversing the order of the above considerations

$$\psi(a'')^{\dagger}\ \psi(a')^{\dagger} = C\ \psi(a')^{\dagger}\psi(a'')^{\dagger}$$

Substitution of $\psi(a'')^{\dagger}\ \psi(a')^{\dagger}$ from this expression into the previous one gives

$$\psi(a')^{\dagger}\ \psi(a'')^{\dagger} = C^2\psi(a')^{\dagger}\ \psi(a'')^{\dagger}$$

Hence C must be ± 1 ; the $\psi^{\dagger}$ either commute or anticommute . We have then automatically for the $\psi^{\dagger}$ the commutation relations of non-Hermitian variables of the first and second kind . We may repeat the derivation for the commutator of two ψ or simply take the adjoint of the relations for the $\psi^{\dagger}$.

For the commutator of a ψ and $\psi^{\dagger}$ we proceed as follows :

$$\left[\psi(a'), N\right] = \left[\psi(a'), \sum_{a''} \psi(a'')^{\dagger}\psi(a'')\right] = \psi(a')$$

A typical term of the commutator bracket will be

$$\psi(a')\,\psi(a'')^{\dagger}\,\psi(a'') - \psi(a'')^{\dagger}\,\psi(a'')\,\psi(a')$$

If the $\psi(a')$ commute, we get

$$\sum_{a''}\left[\psi(a'),\psi(a'')^{\dagger}\right]\psi(a'') = \psi(a')$$

or

$$\left[\psi(a'),\psi(a'')^{\dagger}\right] = \delta(a',a'')$$

If they anticommute, we obtain

$$\left\{\psi(a'),\psi(a'')^{\dagger}\right\} = \delta(a',a'')$$

In the discussion of systems of identical particles we see that the number of dynamical variables which create or annihilate particles in any mode must be equal to the number of possible modes of a particle . This number is infinite, since, say the position of a particle has an infinity of possible values . The ψ are $\psi(r,t)$ in order to discuss particles in space-time . We are lead naturally and inevitably to the field point of view .

We now rewrite the operator relations (7.1) (7.2) in terms of the operators $\psi, \psi^{\dagger}$.

$$F^{(1)} = \sum_{a} \psi(a')^{\dagger} \langle a' \mid f^{(1)} \mid a'' \rangle \psi(a'') ,$$

where a discrete eigenvalue description is explicitly indicated . In a continuous description we write

$$(7.6) \qquad F^{(1)} = \int dr\, dr' \sum_{\xi\xi'} \psi(r,\xi)^{\dagger} \langle r\xi \mid f^{(1)} \mid r'\xi' \rangle \psi(r',\xi')$$

where the ξ are eigenvalues that are still discrete , the inner (spin) states of a particle . (7.6) looks like the expression for an expectation value but is here the definition of an operator and not a numerical equation . It is this analogous form, however, that historically accounts for the definition of $\psi^{\dagger}$ a creation operator and ψ as an annihilation operator rather than the opposite order . The appearance of operators in place of wave functions in this equation also accounts for the name of second quantization which is applied to the description of identical particle systems . (First quantization gives wave functions, second quantization changes wave functions to operators)

For $F^{(2)}$,we have

$$F^{(2)} = \tfrac{1}{2} \sum_{a} \langle a'a''' \mid f^{(2)} \mid a''a'' \rangle \psi(a')^{\dagger} \psi(a'') \psi(a''')^{\dagger} \psi(a'') -$$

$$- \delta(a'',a''') \psi(a')^{\dagger} \psi(a'')$$

With the commutation law we may change the bracketed terms to

$\pm \psi(a')^{\dagger} \psi(a''')^{\dagger} \psi(a'') \psi(a'')$, the $\pm$ depending on whether we have variables of the first or second kind .We compensate for the sign ambiguity by a change in order and write

$$F^{(2)} = \tfrac{1}{2} \sum_{a} \psi(a')^{\dagger} \psi(a''')^{\dagger} \langle a'a''' \mid f^{(2)} \mid a''a'' \rangle \psi(a'') \psi(a'')$$

observing that the opposite order of indices must be maintained in the end of the matrix and in the last operators . Similarly

$$F^{(k)} = \frac{1}{K!} \sum_{a} \psi(a')^{\dagger} \psi(a''')^{\dagger} \ldots \psi(a^{2k-1})^{\dagger} \langle a'a''' \ldots a^{2k-1} \mid f^{(k)} \mid a^{2k} a^{2k-2} \ldots$$

$$\ldots a'' \rangle \psi(a'') \psi(a'') \ldots \psi(a^{2k})$$

Since any operator may be constructed from the ψ and $\psi^\dagger$, these are now the fundamental variables of the theory .

In the continuous description the unit matrix $\delta(a',a'')$ which occurs in the commutation relations passes over to a δ function :

$$\{[\psi(\tau,\xi,t),\psi(\tau',\xi',t)^+]\} = \delta(\tau-\tau')\delta(\xi,\xi')$$

APPENDIX .

In (7.5) , we have assumed that C is not dependent on the a',a", or at least that $C(a',a'') = C(a'',a')$. This assumption justified itself by leading very naturally to the commutation relations which are consistent with our previous developments . We now allow $C = C(a',a'')$ and see what follows .

$$\psi(a')\psi(a'') = C(a',a'')\psi(a'')\psi(a')$$

$$\psi(a'')\psi(a') = C(a'',a')\psi(a')\psi(a'')$$

Substitution gives

$$C(a',a'')\,C(a'',a') = 1 \qquad (7.7)$$

We shall also need to invert the order of $\psi(a')^\dagger\psi(a'')^\dagger$. Taking the adjoint of (7.7),we have

$$\psi(a')^\dagger\psi(a'')^\dagger = C(a'',a')^*\psi(a'')^\dagger\psi(a')^\dagger$$

The commutator $\left[\psi(a'),N\right] = \psi(a')$ written out is

$$\psi(a') \sum_{a''} \psi(a'')^{\dagger} \psi(a'') - \sum_{a''} \psi(a'')^{\dagger} \psi(a'') \psi(a') = \psi(a')$$

Reversal of the order in the second term gives

$$\psi(a') \psi(a'')^{\dagger} - C(a'',a') \psi(a'')^{\dagger} \psi(a') = \delta(a',a'')$$

which is the rule for inverting the order of a ψ and $\psi^{\dagger}$. From the adjoint of this and the same equation for $\psi(a'')^{\dagger} \psi(a')$, we obtain that

$$C(a',a'')^{*} = C(a'',a') .$$

Hence :

$$\psi(a')^{\dagger} \psi(a'')^{\dagger} = C(a',a'') \psi(a'')^{\dagger} \psi(a') ,$$

(7.3) written in terms of the $\psi, \psi^{\dagger}$ is

$$\psi(a')^{\dagger} \psi(a'') \psi(a''')^{\dagger} \psi(a^{IV}) \ldots \psi(a''')^{\dagger} \psi(a^{IV}) \psi(a')^{\dagger} \psi(a'') =$$

$$= \delta(a'',a''') \psi(a')^{\dagger} \psi(a^{IV}) - \delta(a',a^{IV}) \psi(a''')^{\dagger} \psi(a'')$$

This becomes after inverting $\psi(a'') \psi(a''')^{\dagger}$ and $\psi(a^{IV}) \psi(a')^{\dagger}$,

$$C(a''',a'') \psi(a')^{\dagger} \psi(a''')^{\dagger} \psi(a'') \psi(a^{IV}) =$$

$$= C(a',a^{IV}) \psi(a''')^{\dagger} \psi(a')^{\dagger} \psi(a^{IV}) \psi(a'')$$

Further inversions lead to

$$C(a''',a'') C(a'',a^{IV}) C(a',a''') = C(a',a^{IV})$$

Multiply by $C(a^{IV},a'')$. Then

$$C(a',a''') C(a''',a'') = C(a',a^{IV}) C(a^{IV},a'')$$

for any a', a'', a''', a''''. This can only be satisfied if $C(a',a'') = f(a'), g(a'')$, and further such that $f(a')\, g(a') =$ a constant, independent of the a'.
In (7.7) the constant occurs squared . Hence it must be ± 1 , and $f = \pm g^{-1}$.
Using this , we may write :

$$C(a',a'') = \pm\ e^{i\left[\varphi(a') - \varphi(a'')\right]}$$

This implies that the reversal of order of a pair of ψ's involves only a phase factor, which is of no physical significance . We can create particles in arbitrary phases .

Formally, we make the transformation $'\psi(a') = e^{i\varphi(a')N}\psi(a')$

Then :

$$'\psi(a')\,'\psi(a'') = e^{i\varphi(a')N}\psi(a')\ e^{i\varphi(a'')N}\psi(a'')$$

We use

$$\psi(a')\, e^{i\varphi N} = e^{i\varphi(N+1)}\psi(a')$$

(which may be verified by applying it to a state vector with particle number n .)
to write :

$$\begin{aligned}
'\psi(a')\ '\psi(a'') &= e^{i\left[\varphi(a') + \varphi(a'')\right]N}\ e^{i\varphi(a'')}\psi(a)\ \psi(a'') \\
&= \pm\, e^{i\left[\varphi(a') + \varphi(a'')\right]N}\ e^{i\varphi(a')}\psi(a'')\psi(a') \\
&= \pm\, e^{i\varphi(a'')N}\psi(a'')\ e^{i\varphi(a')N}\psi(a') \\
&= \pm\ '\psi(a'')\ '\psi(a')
\end{aligned}$$

Hence the newly defined $'\psi(a')$ fulfill the commutation laws as derived above .

However, the freedom of choice of the phase factors is of some interest .

Suppose, for example, we have two classes of particles of the second kind , say electrons and protons . For class I , we may define the phase to be $\varphi = \pi/2$ for class II, $\varphi = -\ /2$. Then any operator of either class will anticommute with all members of the same class, but commute with members of the other class. This choice is usually not made , but the present discussion indicates that it is a possible choice .

8. - DYNAMICAL THEORY FOR IDENTICAL PARTICLE SYSTEMS

In order to see how the dynamical principle may be applied to the field operators $\psi^+ \psi$ we recall that the Lagrangian for canonical Hermitian variables was of the form $L = \sum_a p_a \frac{dq_a}{dt} - H$, if we neglect symmetrization . For non-Hermitian operators we make the substitution $p_a \rightarrow i q_a^+$. We apply this immediately to the non-Hermitian continuous operators ψ^+, ψ , replacing summation by integration and the total time derivative by a partial derivative :

$$L = \int dr \sum_{\xi} i \psi (r \xi t)^+ \cdot \frac{\partial}{\partial t} \psi (r \xi t) - H$$

Where the , means symmetrize for variables of the first kind , antisymmetrize for variables of the second kind . We adopt a vector notation for the discrete indices and write

$$L = \int dr \ i \psi (r t)^+ \cdot \frac{\partial}{\partial t} \psi (r t) - H \qquad (8.1)$$

The Hamiltonian operator will be a sum of terms corresponding to no interaction among particles , interaction in pairs, threes, and so on . In general, we should

write :

$$H = \int dr\, dr'\, \psi(r,t)^{\dagger} \langle r | H^{(1)} | r' \rangle \psi(r',t) + \tfrac{1}{2} \int dr\, dr'\, dr''\, dr'''\, \psi(r)^{\dagger} \psi(r')^{\dagger} \times$$

$$\times \langle r\, r' | H^{(2)} | r''\, r''' \rangle \psi(r''') \psi(r'') + \ldots$$

In non-relativistic theory, however, where the interactions are contact interactions not action at a distance, the $\langle r\, r' | H^{(2)} | r'' r''' \rangle$ will be of the form :

$$H^{(2)}(r,r') \delta(r - r'') \delta(r' - r''') ,$$

so that the integrals simplify to

$$\int dr\, dr'\, \psi(r)^{\dagger} \psi(r')^{\dagger} H^{(2)}(r,r') \psi(r') \psi(r)$$

The variations within the interval must vanish .Varying L with respect to $\psi(r)$ we have the equation of motion :

$$i \frac{\partial}{\partial t} \psi(r,t) = H^{(1)}(r, \tfrac{1}{i} \nabla) \psi(r,t) +$$

$$+ \int dr'\, \psi(r',t)^{\dagger} H^{(2)}(r,r') \psi(r',t) \psi(r,t) + \ldots$$

If we had no interaction term, this would look exactly like the Schroedinger Equation, except that it is not a numerical equation whose solution is a wave function, but an operator equation . The interaction term is suggestive of the " average potential energy" operator in many-electron problems in first quantization, but again it is a term made up of operators , not wave functions .

From the action principle, in addition to equations of motion , we obtain generators of infinitesimal unitary transformations from the terms that do not vanish on the boundary . By analogy with $G_q = \sum_a p_a\, dq_a$ of our previous

development , we may write: $G_q = \int dr\, i\psi^{\dagger}(r,t)\delta\psi(r,t)$ which, of course , could be obtained directly by treating (8.1) exactly as we did (3.4) to obtain G_q.

The commutation relations follow immediatly (using 2.10)

$$\left[\psi(r), \int dr'\, \psi(r)^{\dagger}\, \delta\psi(r)\right] = \delta\psi(r)$$

$$\int dr'\, \frac{1}{i}\left\{\left[\psi(r), \psi(r')^{\dagger}\right]\right\}\delta\psi(r') = \delta\psi(r)$$

where the choice of $[\,]$ or $\{\,\}$ is determined by whether $\delta\psi(r)$ commutes or anticommutes with the $\psi(r)$, i.e., whether the variables are of the first or second kind . The commutators are then

$$\frac{1}{i}\left\{\left[\psi(r)\,,\, \psi(r')^{\dagger}\right]\right\} = \delta(r - r')$$

We now want to consider the physical invariance of the system under translations and rotations to derive expressions for the total linear and angular momentum operators as we did for the systems treated in section 2 (p.30-32) .

After a translation and rotation a position vector in the new coordinate system is $\vec{r} = \vec{r} - \delta\vec{r}$ in terms of vectors of the old coordinate system, where

$$\delta r = \delta\varepsilon + \delta\omega \times r$$

(if we suppose that $\psi(1)$ has no intrinsec coordinates , so $\psi(r) = \bar{\psi}(r - \delta r)$ The operator $\bar{\psi}(r) = \psi(r + \delta r)$ has the same eigenvalue spectrum in the new system as did $\psi(r)$ in the old system . As indicated by (2.10)

$$\bar{\psi}(r) = \psi(r) - \delta\psi$$

$$= \psi + \delta r.\nabla\psi$$

Then for continuous variables , we write in general :

$$\delta\psi = -(\delta\varepsilon + \delta\omega \times r)\cdot\nabla\psi$$

But we must also take into account the effect of rotation on the ψ inasmuch as they are also vectors as far as the discrete indices are concerned . Hence , we add an infinitesimal transformation among the components of ψ :

$$\delta\psi = -(\delta\varepsilon + \delta\omega\times r)\cdot\nabla\psi - i\,\delta\omega\cdot s\psi$$

Here s is a numerical matrix acting on the ψ_ξ (Cf.G_r on p. 31.)

In order to write the generator of $\delta\psi$, we must take into account the fact that the summations have become integrals . We introduce the functional derivative $\delta/\delta\chi$ so that corresponding to $\delta_\chi G = \int dr\, \delta\psi^+ \frac{\delta_\xi G}{\delta\psi^+}$ we may write $\delta\psi = \frac{\delta_\xi G_\eta}{\delta_\zeta \psi^+}$. Correspondingly :

$$G = \int dr\, \psi^+(\delta\varepsilon + \delta\omega\times r)\cdot\tfrac{1}{i}\nabla\psi + \int dr\, \psi^+\delta\omega\cdot s\psi$$

$$= \delta\varepsilon\cdot P + \delta\omega\cdot J$$

with the definitions of the total linear and angular momentum operators :

$$P = \int dr\, \psi^+\tfrac{1}{i}\nabla\psi \qquad ; \quad J = \int dr\, \psi^+(r\times\tfrac{1}{i}\nabla + s)\psi$$

In the discussion of J in systems where the particles were not identical , we saw that the commutator law $J\times J = iJ$ imposed the same condition on the numerical matrix J (cf. p. 83) . Here we make the same argument for the matrix s .

$$\delta_\omega J = \delta\omega \times J$$

$$= \int dr\, \delta_\omega\psi^+(r\times\tfrac{1}{i}\nabla + s)\psi + \int dr\, \psi^+(r\times\tfrac{1}{i}\nabla + s)\,\delta_\omega\psi$$

$$\delta_\omega \psi = \frac{\delta_\varepsilon G_\omega}{\delta i \psi^\dagger} = -i\delta\omega\cdot\left(r\times\frac{1}{i}\nabla + s\right)\psi$$

$$i\,\delta_\omega \psi^\dagger = -\frac{\delta_\pi G_\omega}{\delta\psi} = -\psi^\dagger\,\delta\omega\cdot\left(r\times\frac{1}{i}\nabla + s\right)$$

$$\delta_\omega J = -\int dr\, i\psi^\dagger\left[\left(r\times\frac{1}{i}\nabla + s\right),\ \delta\omega\cdot\left(r\times\frac{1}{i}\nabla + s\right)\right]\psi$$

Hence for the part of $\delta_\omega J$ that depends on s , we have :

$$\delta_\omega J_{(s)} = \int dr\ \psi^\dagger(\delta\omega\times s)\psi \quad \text{(i.e. } \delta\omega\times J(s)\text{) if } s\times s = i\,s\ .$$

As in the case of J , the dimensionality of the completely reduced matrix s will classify systems of particles according to their spin .

It is useful here to introduce a more general type of variation to illustrate a method of argument that will be useful in the discussion of relativistic fields . Let the symbol $(\delta\psi)$ indicate a more complicated type of variation : one in which the coordinate transformations are a function of time :

$$(\delta\psi) = \delta\psi + \left(\delta\varepsilon^{(r)} + \delta\omega^{(r)}\times r\right)\cdot\nabla\psi - i\,\delta\omega^{(r)}\,s\,\psi$$

For an isolated system H remains invariant under coordinate transformations even when they are time-dependent . The only effect of the time dependence is to add new terms to the variation of the kinetic part of the Lagrangian :

$$(\delta L) = \delta L + \dot{\delta\varepsilon}\cdot P + \dot{\delta\omega}\cdot J$$

In the action principle these terms become

$$\int dt\,\frac{d}{dt}\left[\delta\varepsilon\cdot P + \delta\omega\cdot J\right] - \int dt\left[\delta\varepsilon\cdot\dot{P} + \delta\omega\cdot\dot{J}\right]$$

Here the first integral gives the generators of the coordinate transformations . The second integral indicates that P and J are constants of motion .

Additional information concerning general characteristics of a system of identical particles may be obtained by considering another kind of variation. Let $\delta\psi = -i\delta\alpha.\psi$ where α is a real number . This is simply a change of phase .

Then $\bar{\psi} = \psi - \delta\psi = (1 + i\delta\alpha)\psi$ and the adjoint variation is :

$$\delta\psi^{+} = i\delta\alpha\,\psi^{+}$$

Imposing the conditions for a canonical transformation, we have

$$-i\delta\alpha\,\psi = \frac{\delta_{\ell} G_{\alpha}}{\delta i\psi^{+}} \quad ; \quad i\delta\alpha\,\psi^{+} = -\frac{\delta_{r} G_{\alpha}}{\delta i\psi}$$

$$G_{\alpha} = \delta\alpha \int dv\; \psi^{+}\psi = \delta\alpha\, N$$

Hence, if for a given system H is invariant under the transformation $\psi \to e^{i\alpha}\psi$ $\psi^{+} \to e^{-i\alpha}\psi^{+}$ then for the system $\frac{dN}{dt} = 0$. The number of particles is a constant of motion . That this depends on the invariance of H under the transformation indicates that the number of particles in a system is in general a dynamical question , not one that is purely kinematic .

But for F.D. fields we may go further . Since the Hamiltonian of a F.D. system must be even in the variables ψ^{+}, ψ , the transformation $\psi \to -\psi$ $\psi^{+} \to -\psi^{+}$ leaves H unchanged . We has studied this transformation in section 5 p. 99 -100 . Here, we have again the finite transformation $\psi \to e^{i\alpha}\psi$,

$\psi' \to e^{-i\alpha} \psi'$ with $\alpha = \pi$.

The unitary operator generating this transformation is $e^{i\pi N_{FD}} = \pm 1$ and as we say in section 5 p. 100, this is a kinematical invariant . For a given system , it is either + 1 or - 1 , and cannot change in sign . This implies that the number of particles is either even or odd . Hence for a F.D. system , particles must be created or annihilated in even numbers .

The behavior of F.D. systems under rotation also gives us a first suggestion , even in non-relativistic theory, of the relation between spin and statistics .

Carry out a rotation about the z-axis through the angle 2π . H is invariant . For infinitesimal rotations

$$\bar{\psi} = \psi + i\delta\omega_3 \cdot [r \times \tfrac{1}{i}\nabla + s]\psi$$

If the finite rotation through the angle 2π is carried out by repeated infinitesimal rotations , the orbital part remains unchanged since $\vec{r}$ returns to its initial value . But for the s operator , we have as the total effect

$$\psi \longrightarrow e^{2\pi i s_3} \psi$$

For a particle with spin s , s_3 has the possible values s, s-1,...,-s +1, -s . If the spin s is integral , after the rotation , ψ has its original value . But if s is half-integral

$$\left.\begin{aligned} \psi &\to -\psi \\ \psi^+ &\to -\psi^+ \end{aligned}\right\}$$ The operators are double-valued .

Then for the required invariance $H(\psi, \psi^+) = H(-\psi, -\psi^+)$. Hence H must be an even function of the ψ- operators. In general it would be a severe restriction on the form of H for particles of half-integral spin to require that it be an even function of the field operators. However, if we can identify systems with half-integral spin with systems of the second kind, there is no further restriction on H. This is, of course, no proof that half-integral spin $\longrightarrow$ variables of the second kind, but an indication of simplicity and consistency. The proof will be given in the relativistic theory.

So far we have taken it for granted that our variables of the first kind describe a B.E. field and those of the second kind a F.D. field. We will now calculate the eigenfunctions and eigenvalues for fields of non-interacting particles, i.e. for fields in which

$$H = \int dr\, \psi^+ H^{(1)} \psi = \int dr\, \psi(r,x) + H^{(1)}\left(r, \frac{1}{i}\nabla\right) \psi(r,x)$$

in which as usual we suppress spin indices by a vector notation.

We wish to construct the transformation function in time $\langle \psi^{+\prime} t_1 | \psi' t_2 \rangle$ by reordering δW to give a $\delta \Omega$. By the dynamical principle

$$\delta \langle \psi^{+\prime} t_1 | \psi' t_2 \rangle = i \langle \psi^{+\prime} t_1 | \delta W | \psi' t_2 \rangle$$

where

$$\begin{aligned} \delta W &= G\psi^+ t_1 - G\psi t_2 - H^{(1)} \delta t_1 \qquad (t_2 \text{ is left unchanged}) \\ &= -i \int dr\, \delta\psi^+(r t_1) \psi(r t_2) - i \int dr\, \psi(r t_2)^+ \delta\psi(r, t_2) \\ &\quad - \int dr\, \psi(r,t)^+ H^{(1)} \psi(r,t)\, dt_1 \end{aligned}$$

To construct $\delta \mathcal{W}$ we must first rewrite δW in a form in which the operators act on their corresponding vectors at the proper times , i.e., we must have δW as a function of $\psi(r, t_2)$ and $\psi(r, t_1)^+$. To do this we make use of the equation of motion :

$$\left(i \frac{\partial}{\partial t} - H^{(1)} \right) \psi = 0$$

The solution with the requisite boundary conditions is

$$\psi(r, t_1) = \int dr' \, \langle r t_1 | r' t_2 \rangle \, \psi(r' t_2)$$

or

$$\psi(r, t_1) = i \int dr' \, G_{ret}(r, t_1, r' t_2) \, \psi(r', t_2)$$

G_{ret} is chosen since $t_1 > t_2$

$$i \, G_{ret}(t = t' + 0) = \delta(r - r') = \langle r \; t_1 | \, r' t_2 \rangle$$

for $t_1 = t_2$

We now have

$$\delta W = \int dr \, dr' \, \delta\psi^+ G(r, t_1, r' t_2) - i \int dr \, \psi(r, t_2)^+ \, \delta\psi$$
$$- i \int dr \, dr' \, \psi(r, t_1)^+ H^{(1)}(r, t_1) \, G(r t_1, r' t_2) \, \psi(r', t_2)$$

We could also solve directly for $\psi(r, t_1)$ in terms of $\psi(r, t_2)^+$ but it is only necessary to observe (a) that G_{ret} satisfies the Schroedinger equation hence $H^{(1)} G_{1\,2} = i \frac{\partial}{\partial t} G_{1\,2}$

and (b) that by the dynamical postulate δW is a perfect differential . We have now what we are looking for :

$$\delta W = \delta \int dr \, dr' \, \psi(r, t_1)^+ G(r t_1, r' t_2) = \delta \mathcal{W}$$

since the operators are already properly ordered .

Then : $\Omega = \int \psi(r t_1)^+ G_{ret}(r t_1, r' t_2) \psi(r' t_2) + \text{const.}$

and $\langle \psi^+ t_1 | \psi' t_2 \rangle = e^{i \int dr\, dr'\, \psi(r,t_1)^+ G_{ret}(r t, r' t_2) \psi(r' t_2)'}$

where the constant of integration has been evaluated from the limiting case $t_1 = t_2$ where $i G = \delta(r - r')$ and the transformation function takes the form for non-Hermitian variables :

$$\langle \psi^{+\prime} | \psi' \rangle = e^{\int dr\, \psi^+ \psi}$$

What we have done so far is valid also if $H = H(t)$, but we do not need that generality here .

We now go to a discrete representation , diagonal in the energy :

$$\langle \psi(a', t_1)^{+\prime} | \psi(a'', t_2)' \rangle = e^{i \sum \psi(a', t_1)^+ G(a' t_1, a'' t_2) \psi(a'' t_2)}$$

From the definition of G , $i G(a' t_1, a'' t_2) = \langle a' t_1 | a'' t_2 \rangle = \delta(a', a'') e^{-i E_{a'}(t_1 -$

since the representation is diagonal in the energy .

$$\langle \psi^{+\prime} t_1 | \psi' t_2 \rangle = e^{\sum \psi(a')^{+\prime} e^{-i E_a t} \psi(a')'} \qquad \text{where } t = t_1 - t_2$$

We may write the exponential as a product since the eigenvalues taken in pairs will commute for both kinds of variables .

$$\langle \psi^{+\prime} t_1 | \psi' t_2 \rangle = \prod_{a'} e^{\psi(a')^{+\prime} e^{-i E_a t} \psi(a')}$$

$$= \prod_{a} \sum_{n_{a'}=0}^{\infty} \frac{(\psi(a')^{+\prime} \psi(a'))^{n_{a'}}}{n_{a'}!} e^{i n_{a'} E_{a'}}$$

The effect of the product $\prod_a$ on the exponential is to give $e^{-i\sum_{a'} n_{a'} E_{a'} t} = e^{-iE(n)t}$

where $E(n)$ is the sum of the energy of the n particles in the mode a' .

For systems of the first kind :

$n_{a'} = 0, 1, \ldots, \infty$. There may be any number of particles in a given mode . This is the occupation number for Bose-Einstein statistics . The wave functions are

$$\langle \psi^{\dagger\prime} | n \rangle = \prod_{a'} \frac{\left[\psi(a')^{\dagger\prime}\right]^{n_{a'}}}{\sqrt{n_{a'}!}} \; ; \quad \langle n | \psi' \rangle = \prod_{a'} \frac{\left[\psi(a')'\right]^{n_{a'}}}{\sqrt{n_{a'}!}}$$

For systems of the second kind , the series terminates after the linear term $n_a = 0, 1$. The exclusion principle is operating . The occupation number is that of Fermi-Dirac statistics . The wave functions are

$$\langle n | \psi' \rangle = \prod_{a'} \left[\psi(a')'\right]^{n_{a'}}$$

where the terms in the product are arrange in an arbitrary standard order .

$$\langle \psi^{\dagger\prime} | n \rangle = \prod_{a'}{}^{T} \left[\psi(a')'\right]^{n_{a'}}$$

where the $\prod^T$ means the product must be arranged in the order opposite to that in $\langle n | \psi' \rangle$

9 . - PERTURBATION CALCULATIONS .

We will repeat the calculation we have just made from a different point of view which leads toward general methods of perturbation calculation in field theory .

We consider the system with the Lagrangian :

$$L = \int dr \psi^{\dagger} \left(i \frac{\partial}{\partial t} - H^{(1)} \right) \psi$$

We introduce also a canonical transformation corresponding to a displacement of the canonical variables .

$$L = \int dr \psi^{\dagger} \left(i \frac{\partial}{\partial t} - H^{(1)} \right) \psi - \int dr \left[\psi^{\dagger} \eta + \eta^{\dagger} \psi \right]$$

where the η and $\eta^{\dagger}$ are operators defined to commute or anticommute according as the field operator ψ $\psi^{\dagger}$ are variables of the first or second kind . The equation of motion becomes

$$\left(i \frac{\partial}{\partial t} - H^{(1)} \right) \psi(r,t) = \eta(r,t)$$

From the field point of view η then acts as a source of field as , say, the current and charge in Maxwell theory .

We now define that $\eta(r,t)$ is to act at the instant $t = t_o$, as an impulsive force does in mechanics .

$$\eta(r, t) = i \psi(r)' \delta(t - t_o)$$

Hence the field jumps discontinuously at $t = t_o$:

$$\psi(r,t) \Big|_{t_o - o}^{t_o + o} = \psi(r)' \qquad (9.1)$$

The adjoint equation of motion is :

$$\psi^{\dagger} \left(-i \frac{\partial^{\dagger}}{\partial t} - H^{(1)} \right) = \eta^{\dagger}$$

We write also

$$\eta^{\dagger} = i\psi^{\dagger\prime} \delta (t - t_0) \qquad (9.2)$$

This last equation is not the adjoint of the definition of η but the sign here is chosen so that the field discontinuity is in the sense opposite to that caused by η , i.e.

$$\psi^{\dagger} \Big|_{t_0 - o}^{t_0 + o} = -i\psi^{\dagger\prime}$$

By using the " field sources" , we can now construct $\langle \psi^{\dagger\prime} t_1 | \psi' t_2 \rangle$ from $\langle 0\, t_1 | 0\, t_2 \rangle$. The usefulness of this technique appears from the following: $|0\rangle$ for a system of independent particles is the state with occupation number 0, that is the vacuum of the system . Since

$$\langle n | \psi(a')' \rangle = \overline{|_{a'}|} \left[\psi(a')'\right]^{n_{a'}} / n_{a'}!$$

will vanish when the eigenvalue 0 appears in the product unless raised to the 0 th power , $\langle n | 0 \rangle = \delta(n, 0)$. We may construct any transformation function from the vacuum transformation function by introducing sources which create the physical situation we are interested in , and destroy it later , leaving the vacuum again . $\langle 0\, t_1 | 0\, t_2 \rangle$ will mean the vacuum-vacuum transformation function with the " source" η at t_2 and the " sink" η' at t_1 .

$$\delta_{\eta\eta^{\dagger}} \langle 0\, t_1 | 0\, t_2 \rangle = -i \langle 0\, t_1 | \int_{t_1}^{t_2} dt\, dr \left[\delta\eta\, \psi^{\dagger} + \delta\eta^{\dagger} \psi\right] | 0\, t_2 \rangle$$

$$= - i \int dt\, dr \left[\delta\eta^{\dagger} \langle 0|\psi|0\rangle + \langle 0|\psi^{\dagger}|0\rangle \delta\eta \right]$$

We consider the term $\frac{\langle 0|\psi|0\rangle}{\langle 0|0\rangle}$ by itself . This is a number which is a solution of the differential equation

$$\left(i \frac{\partial}{\partial t} - H^{(1)} \right) \left[\frac{\langle 0|\psi(r,t)|0\rangle}{\langle 0|0\rangle} \right] = \eta(r,t)$$

We write the solution immediatly , using the retarded Green's function .

$$\frac{\langle 0|\psi(r,t)|0\rangle}{\langle 0|0\rangle} = \int_{t_2}^{t_1} G_{ret}(rt, r't') \eta(r't') \, dr' dt'$$

The retarded Green's function is a solution of

$$\left(i \frac{\partial}{\partial t} - H^{(1)} \right) G = \delta(r - r') \, \delta(t - t')$$

with the boundary conditions

$$G_{ret}(r\, t_1 r' t') = \frac{1}{i} \langle r\, t | r'\, t' \rangle \quad t > t'$$

$$= 0 \qquad t < t' .$$

G_{ret} is appropriate here since $t_1 > t_2$.

$$\frac{\delta_{\eta\eta'} \langle 0\, t_1 | 0\, t_2 \rangle}{\langle 0\, t_1 | 0\, t_2 \rangle} = - i \int dt\, dt'\, dr\, dr' \left[\delta\eta^{\dagger} G(rt, r't') \eta + \eta^{\dagger} G(r\,t, r't') \delta\eta \right]$$

$$\langle 0\, t_1 | 0\, t_2 \rangle_{\eta\eta'} = e^{-i \int dt\, dt'\, dr\, dr' \left[\eta(r,t)^{\dagger} G_{ret}(r\,t, r't') \eta(r't') \right.}$$

The constant of integration is unity since in the limit $t_1 = t_2$

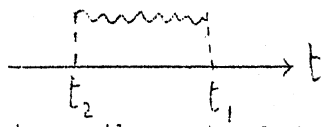
$$\langle \psi^{\dagger\prime} | \psi' \rangle = e^{\int \psi^{\dagger\prime} \psi' dr}$$

We have not yet imposed the conditions (9.1) (9.2) on the source functions . We could produce transformation functions for more complicated Hamiltonians by differentiation with respect to η and $\eta^\dagger$.

We now make use of $\eta, \eta^\dagger$ in the form (9.1), (9.2) to get

$$\langle \psi^\dagger t_1 | \psi' t_2 \rangle = e^{i \int dr\, dr'\, \psi^\dagger(r)'\, G(r\, t_1, r' t_2)\, \psi(r')'} \qquad (9.3)$$

Instead of introducing sources into the vacuum state , we can use the method to calculate the effect of a small perturbation on a system when we know the transformation function for the unperturbed system .



Let the Hamiltonian for the complete system be $H^{(1)} = H_o + H_1$, where we suppose H_1 is a small time-dependent perturbation that acts on the system between t_2 and t_1 . The transformation function for the system without the perturbation is presumed to be known and diagonal in the energy :

$$\langle a' t_1 \mid a'' t_2 \rangle_{H_o} = \delta(a', a'')\, e^{-iE_{a'}(t_1 - t_2)}$$

With the Green's function we can express $\langle \psi^{\dagger\prime} t_1 | \psi' t_2 \rangle$ in the form (9.3). By definition $i G_{ret} = \langle a' t_1 \mid a'' t_2 \rangle_H$. The problem is to express $\langle a' t_1 | a'' t_2 \rangle_{H^{(1)}}$ in terms of $\langle a' t_1 | a'' t_2 \rangle_{H_o}$. The two differ by the differential transformation function generated by $H_1(t)$

$$\langle a' t_1 | a'' t_2 \rangle = \langle a' t_1 \mid a'' t_2 \rangle_{H_o} - i \left\langle a' t_1 \left| \int_{t_2}^{t_1} dt\, H_1(t) \right| a'' t_2 \right\rangle_{H_o}$$

$$= \delta(a',a'')\, e^{-i E_{a'}(t_1-t_2)} - i\int_{t_2}^{t_1} dt\, e^{-i E_{a'}(t_1-t)} \langle a' | H_1(t) | a'' \rangle$$
$$e^{-i E_{a''}(t - t_2)} \qquad (9.4)$$

Since the perturbation is confined to the interval $t_2 \to t_1$ and vanishes out of the interval , we may extend the integration limits to infinity . Now with the Green's function $G = \frac{1}{i} \langle a' t_1 | a'' t_2 \rangle_H$ from (9.4) in (9.3) we can expand the exponential to obtain

$$\langle \psi^{t'} t_1 | \psi' t_2 \rangle = \sum_{n\,n'} \langle \psi^{t'} | n \rangle \langle n t_1 | n' t_2 \rangle \langle n' \psi' \rangle \qquad (9.5)$$

We know what the $\langle \psi^{t'} | n \rangle$ and $\langle n' | \psi \rangle$ are . The $|\langle n t_1 | n' t_2 \rangle|^2$ gives the probability of a transition between the states with the occupation numbers n and n' .

$$\langle \psi^{t'} t_1 | \psi' t_2 \rangle = e^{\sum_{a'} \psi(a')^{t'} [e^{-iE_{a'}(t_1-t_2)} \delta(a',a'') - i \int_{-\infty}^{\infty} dt\, e^{-iE_{a'}(t_1-t)} \langle a' | H_1(t) | a'' \rangle \times}$$
$$e^{-iE_{a''}(t-t_2)}] \psi(a'')'$$

However before making the expansion we can absorb the time dependence into the $\psi^{t'}, \psi'$ by a change of phase :

$$\psi(a')^{t'} \; e^{-i E_{a'} t_1} \to \psi(a')^{t'}$$

$$\psi(a'')' \; e^{i E_{a''} t_2} \to \psi(a'')'$$

At the same time in the occupation number representation, we change the phase of the $\langle \psi^{t'} | n \rangle$:

$$\langle \psi^{t'} | n \rangle \to \langle \psi^{t'} | n \rangle \, e^{iE(n) t_1}$$

and remove the time dependence from the matrix element by defining $\langle n | S | n' \rangle$ so that :

$$\langle n t_1 | n' t_2 \rangle = e^{-E(n)t_1} \langle n | S | n' \rangle e^{iE(n')t_2}$$

We now have

$$\langle \psi^{+\prime} t_1 | \psi' t_2 \rangle = e^{\sum \psi(a')^{+\prime} \, s(a',a'') \psi'(a'')} \qquad (9.6)$$

where

$$s = \delta(a',a'') - i\int_{-\infty}^{\infty} dt\, e^{-i(E_{a'} - E_{a''})t} \langle a_1 | H_1(t) | a'' \rangle .$$

$$\equiv \delta(a',a'') + \langle a' | t | a'' \rangle$$

and

$$\langle \psi^{+\prime} t_1 | \psi' t_2 \rangle = \sum \langle \psi^{+\prime} | n \rangle \langle n | S | n' \rangle \langle n' | \psi' \rangle \qquad (9.7)$$

We expand (9.5), omitting the factor $\psi^{+\prime}(a')$ which would give us only $\langle \psi^{+\prime} | n \rangle$ which we know already. The $n_{a'}$ th power term is

$$\prod_{a'} \sum_{a''} \frac{\left(s(a,a'')\psi(a'')'\right)^{n_{a'}}}{\sqrt{n_{a'}!}}$$

This corresponds to the $\langle n | S | n' \rangle \langle n' | \psi' \rangle$ part of (9.6). Since the perturbation is small we write :

$$\prod_{a'} \frac{\left(\psi(a')' + \sum_{a''} \langle a' | t | a'' \rangle \psi(a'')\right)^{n_{a'}}}{\sqrt{n_{a'}!}} = \langle n | \psi' \rangle + \prod_{a'} \frac{[\psi(a')']^{n_{a'}-1}}{\sqrt{n_{a'}!}} n_{a'} \sum_{a''} \langle a' | t | a'' \rangle \psi(a'')' \qquad (9.8)$$

Here we have as factors terms that are almost the wave functions of B.E. fields for non-interacting particles in the occupation number representation. We would have $\langle n_{a'} - 1 | \psi' \rangle$ if the factorial in the denominator were $(n_{a'} - 1)$. This we obtain by canceling $\sqrt{n_{a'}}$ with the factor in the

numerator . We have also $\psi'(a'')$ which combines with the $\prod_{a'} \psi'(a')$ when $a' = a''$ to raise the power of the $\psi(a'')$ by 1 . To have the appropriate denominator we need a factor $\sqrt{a''+1}$. When we combine these factors , we have for the perturbation term

$$\sum_{a''} \sqrt{n_{a'}\,(n_{a''}+1)}\ \langle a'|t|a''\rangle \langle n - 1_a + 1_{a''}|\psi'\rangle$$

where $\langle n - 1_{a'} + 1_{a''}|\psi'\rangle$ means a state with the occupation number n from which one particle in the a' th mode has been taken and to which one in the a'' th mode has been added . By comparison with (9.7) :

$$\langle n \mid S \mid n - 1_{a'} + 1_{a''}\rangle = \sqrt{n_{a'}\,(n_{a''}+1)}\ \langle a'|t|a''\rangle ,$$

The matrix is written in the wrong historical order . We put it in the right time sequence :

$$\langle n + 1_{a'} - 1_{a''} \mid S \mid n\rangle = \sqrt{(n_{a'}+1)\ n_{a''}}\ \langle a'|t|a''\rangle$$

The second order term would clearly involve trans'itions in which two particles were created and two annihilated . Since the probability of a transition depends on $|\langle n \mid S \mid n'\rangle|^2$

$$P(n + 1_{a'} - 1_{a''}, n) = (n_{a'}+1)\, n_{a''} \left| \int dt\, e^{-it\left[Ea' - Ea''\right]} \langle a'|H_1(t)|\ a''\rangle \right|^2$$

The integral is what we get from simple perturbation theory . The factors $n_{a'}+1$ and $n_{a''}$ show that the transition is proportional to the number of particles in the state from which the transition is made and also to the

number in the final state (stimulated emission of light) .

For F.D. fields the series expansion in (9.8) breaks off after the linear terms . $n_{a'} = 0,1$. The coefficient is easily seen to be $\sqrt{(1 - n_{a'}) n_{a''}}$ which expresses the fact that the initial state must be occupied , and final state unoccupied if there is to be a transition .

IO. TRANSFORMATION FROM FIELD DESCRIPTION TO PARTICLE DESCRIPTION .

So far we have described the identical particle system by means of the field operators of annihilation and creation : $\psi(r,t)$, $\psi(r,t)^{+}$. We want now to see how we can pass from this description to one with operators in ordinary space , i.e., the Schroedinger Equation description in the 3n coordinates of configuration space .

The point of contact is the interpretation of the operator

$$\psi(r,t)^{+}\psi(r,t) = \rho(r,t)$$

$$N = \int dr\, \rho(r,t)$$

The particle description corresponds to a description in terms of the eigenvalues of $\rho(r,t)$.

$\rho(r,t)$ is Hermitian and $[\rho(x), \rho(x')] = 0$. Hence we can look for the eigenvalues and eigenfunctions of ρ and expect them to be physically significant .

What is the transformation generated by these operators ?

$$G = \int dr \cdot \delta\lambda(r)\,\rho(r)$$

For a canonical transformation

$$\delta\psi(r) = \frac{\delta_R G}{\delta i\psi^+} = -i\,\delta\lambda(r)\,\psi(r)$$

so that

$$\psi(r) \to \bar{\psi}(r) = (1 + i\,\delta\lambda(r))\,\psi(r)$$

This is a phase factor transformation . In general for a parameterized transformation

$$\langle \psi^{+'} \lambda \,|\, \psi\, 0\rangle = \langle \psi^{+'}|\, e^{i\int dr\, \lambda(r)\rho'(r)}\,|\psi'\rangle \qquad (10.1)$$

(Cf. (5.9)

In the present case we may first transform the $\psi'(r)$ and then write the transformation function for non-Hermitian canonical variables)

$$\langle \psi^{+}\lambda \,|\, \psi\lambda\rangle = e^{\int dr\, \psi(r)^{+'}\psi'(r)}$$

Here

$$\langle \psi^{+}\lambda \,|\, \psi\, 0\rangle = e^{\int dr\; \psi(r)^{+'} e^{i\lambda(r)}\psi'(r)} \qquad (10.2)$$

If we expand (10.2) and compare with (10.1) , we have the eigenvalues and eigenfunctions of ρ

$$e^{\int dr\, \psi^{+'} e^{i\lambda(r)}\psi} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\int dr\, \psi(r)^{+'}\, e^{i\lambda(r)}\, \psi'(r)\right)^n \qquad (10.3)$$

$$= \sum_{n=0}^{\infty} \int \frac{dr \ldots dr}{n!}\; e^{i\left[\lambda(r_1) + \lambda(r_2) + \ldots + \lambda(r_n)\right.}$$

The comparison with (10.1) shows that

$$\int dr\, \lambda(r)\, \rho'(r) = \sum_{k=1}^{n} \lambda(r_k)$$

or

$$\rho'(r) = \sum_{k=1}^{n} \delta(r - r_k)$$

Hence $\rho'(r)$ indicates how many particles are in the system and where they are .

In (10.3) the r_i enter symmetrically . We may interpret the 1/n ! as compensating for the repititious counting of physically indistinguishable states . With this understanding, we rewrite (10.3) in a way that makes it manifestly a transformation function :

$$\langle \psi^{+'} \lambda | \psi' 0 \rangle = \sum_{n=0}^{\infty} \int \frac{dr_1 \ldots dr_n}{n!} \langle \psi^{+'} | r_1 \ldots r_n \rangle\, e^{i \int dr\, \lambda \rho'} \langle r_1 \ldots r_n | \psi' \rangle$$

The eigenfunctions are the transformation functions from the field to the particle representation :

$$\langle \psi^{+'} | r_1 \ldots r_n \rangle = \psi(r_n)^{+'} \ldots \psi(r_1)^{+'}$$

$$\langle r_1 \ldots r_n | \psi \rangle = \psi(r_1)' \ldots \psi(r_n)'$$

In particular for n = 0, i.e., no particles , we may define :

$$\langle \psi^{+} | 0 \rangle = 1 \quad ; \quad \langle 0 | \psi' \rangle = 1$$

We can use these vacuum state wave functions to give an explicit operational form for a state with n particles .

$$\langle \psi^{+'} | r_1 \ldots r_n \rangle = \psi(r_n)^{+'} \ldots \psi(r_1)^{+'} \langle \psi^{+'} | 0 \rangle$$

in conformity with the original definition of the ψ^+ as creation operators.

In this form the commutation properties of the $\psi(r_n)^{+'}$ also give the symmetry rules for wave functions of B.D. and F.D. systems . Two states that differ only by the interchange $r_1 \longrightarrow r_2$, $r_2 \longrightarrow r_1$ are physically indistinguishable . For B.E. systems , the interchange leaves the sign of the wave function unchanged . In F.D. systems the commutation rules mean a change of sign of the wave function .

The discussion of the configuration space description also requires that we consider the orthogonality and normalization properties of the $| r_1 \cdots r_n \rangle$

$$\langle r_1 \cdots r_n | r_1' \cdots r_n' \rangle = \langle r_1 \cdots r_n | \Psi' \rangle \times \langle \psi^{+'} | r_1 \cdots r_n \rangle$$

where the $\times$ indicates some mode of composition (i.e. integration , summation for discrete eigenstates , or differentiation .) We use here composition by differentiation :

$$\langle r_1 \cdots r_n | r_1' \cdots r_n' \rangle = \langle r_1 \cdots r_n | \frac{\delta e}{\delta \psi^{+'}} \rangle \langle \psi^{+'} | r_1' \cdots r_n' \rangle \Big|_{\psi^{+'} = 0}$$

We write the right hand side in the field representation :

$$\frac{\delta e}{\delta \psi(r_1)^{+'}} \cdots \frac{\delta e}{\delta \psi(r_n)^{+'}} \psi(r_n')^{+'} \cdots \psi(r_1')^{+'} \Big|_{\psi^{+'} = 0}$$

If $n > n'$, there will be differentiation with respect to variables that do not appear on the right . Hence the states will be orthogonal . If $n' > n$, the operation $|_{\psi^{+'} = 0}$ will make the product vanish : the states will be orthogonal . The two statements can be put into a factor $\delta_{nn'}$.

Consider now the case $n = n'$, treating the cases $n=1$, $n=2$ to see how it goes .

For $n = n' = 1$,

$$\langle r | r' \rangle = \frac{\delta e}{\delta \psi(r)^{+\prime}} \psi(r')^{+\prime} = \delta(r - r')$$

For $n = n' = 2$,

$$\langle r_1 r_2 | r_1' r_2' \rangle = \frac{\delta e}{\delta \psi(r_1)^{+\prime}} \frac{\delta e}{\delta \psi(r_2)^{+\prime}} \psi(r_2')^{+\prime} \psi(r_1')^{+\prime}$$

Consider first the differentiation with respect to $\psi(r_2)^{+\prime}$. We are differentiating a product , so there will be two terms . In order to bring $\psi(r_1)^{+\prime}$ adjacent to $\frac{\delta e}{\delta \psi(r_2)^{+\prime}}$, we must use the commutation properties of the $\psi(r)^{+\prime}$. Hence :

$$\frac{\delta}{\delta \psi(r_2)^{+\prime}} \mp \left\{ \psi(r_2')^{+\prime} \psi(r_1')^{+\prime} \right\} = \delta(r_2 - r_2') \psi(r_1')^{+\prime} \pm \delta(r_2 - r_1') \psi(r_2)^{+\prime}$$

Since only single $\psi(r)^{+\prime}$ terms remain, the second differentiation can be carried out with no complications :

$$\langle r_1 r_2 | r_1' r_2' \rangle = \delta(r_1 - r_1') \delta(r_2 - r_2') \pm \delta(r_1 - r_2') \delta(r_2 - r_1')$$

In the general case we have for F.D. systems :

$$\langle r_1 \ldots r_n | r_1' \ldots r_n' \rangle = \prod_{\text{permut.}} \pm \delta(r_1 - r_2') \delta(r_2 - r_3') \ldots$$

where $\prod_{\text{permut.}} \pm$ means all permutations of the factors are to be taken , the + with even permutations , the - with odd permutations . This is a determinate of the n th order . We may write :

$$\langle r_1 \ldots r_n | r_1' \ldots r_n' \rangle = \det_{(n)} \delta(r_i - r_j')$$

For B.E. systems, the structure is similar to the determinant but all terms have a + sign .

The final problem in the configuration space representation is to discuss the dynamics . We first ask what the Schroedinger equation looks like in the particle description . In the Schroedinger description the wave functions are functions of the time :

$$\langle r_1 \dots r_n, t | \Psi \rangle = \psi (r_1 \dots r_n, t)$$

We start with a given field with

$$H = \int dr \, \psi^{\dagger}(r) H^{(1)} \psi(r) + \tfrac{1}{2} \int dr dr' \psi^{\dagger}(r') \, H^{(2)} \psi(r') \psi(r) + \dots$$

We can write the Schroedinger equation immediately :

$$i \frac{\partial}{\partial t} \langle r_1 .. r_n \, t | \Psi \rangle = \langle r_1 \dots r_n t | H \Psi \rangle$$

(We would write first $i \frac{\partial}{\partial t} \langle r_1 .. r_n \, t | \Psi \rangle = H \langle r_1 \dots r_n t | \Psi \rangle$. But from the nature of our linear vector space, the projection of $H \Psi$ on $\langle r_1 .. r_n |$ is just the same as H applied to the projection of Ψ on $\langle r_1 \dots r_n |$. The form given is more convenient to get the results below .)

$$i \frac{\partial}{\partial t} \langle r_1 \dots r_n, t | \Psi \rangle = \int \frac{dr' \dots dr'}{n!} \langle r_1 \dots r_n | H | r_1' .. r_n' \rangle \langle r_1' \dots r_n' | \Psi \rangle$$

To express $\langle r_1 .. r_n | H | r_1' \dots r_n' \rangle$ in terms of the known field Hamiltonian, we make a transformation :

$$\langle r_1 \dots r_n | H | r_1' \dots r_n' \rangle = \langle r_1 \dots r_n | \psi' \rangle \times \langle \psi' | H | \psi' \rangle \times \langle \psi' | r_1' \dots r_n' \rangle$$

Composition by differentiation introduces δ-functions . After integration over the r_i' coordinates we are left with

$$i\frac{\partial}{\partial t}\langle r_1 \ldots r_n, t|\Psi\rangle = \left[\sum_{k=1}^{n} H_{(k)}^{(1)} + \tfrac{1}{2}\sum_{k\neq l} H_{(kl)}^{(2)}\right]\langle r_1 .. r_n, t|\Psi\rangle$$

This is the Schroedinger equation in the form we expect . The operators act on the coordinates of the individual particles .

The transformation function $\langle \psi^{+'} t_1 | \psi' t_2\rangle$ is :

$$\langle \psi^{+'} t_1|\psi' . t_2\rangle = \sum_{nn'}\int \frac{dr_1 \ldots dr_n}{n!}\frac{dr_1' \ldots dr_n'}{n!} .$$

$$\langle \psi^{+'}|r_1 \ldots r_n\rangle\langle r_1 \ldots ; r_n t_1|\, r_1' \ldots r_n' t_2\rangle\langle r_1' \ldots r_n'|\psi'\rangle \qquad (10.4)$$

which corresponds to (9.5) . If we expand $\langle \psi^{+} t_1|\psi' t_2\rangle$ in eigenvalues we obtain by comparison with (10.4) the Green's function or transformation function for an arbitrary number of particles .

11. - LORENTZ INVARIANCE .

It is only when we have operators that depend both on space and time that we are in a position to discuss Lorentz invariance . We have seen infinitesimal coordinate transformations in three dimensions . We now wish to generalize to Lorentz transformations in space-fime .

For a transformation of the space coordinates :

$$\delta r = \delta \varepsilon + \delta\omega \times r ,$$

$$\delta t = 0 .$$

Now, however, the coordinates system is to move in time with a velocity δV . If we write :

$$\delta r = \delta \varepsilon + \delta\omega \times r + \delta V\, t$$

with

$$\delta t = 0 ,$$

we have the Galilean transformation . We know however that $(r_1 - r_2)^2 - (t_1 - t_2)^2$ must be left invariant under the transformation to the moving coordinate system (We choose units so that $c = 1$) . This means that the space and time variations must compensate . For the required invariance

$$\delta t = \delta V . r$$

Equivalent operators for observers in different inertial frames must have the same spectra . The considerations of invariance under the transformations generated by P and J must be extended to transformations in space and time. Corresponding to the extension of δr by the term $\delta V\, t$, we must include a term in the generator of space-time transformations :

$$G = \delta \varepsilon . P + \delta\omega . J + \delta V . N - \delta t . H .$$

The first three terms are the generators of the infinitesimal Lorentz transformation, the fourth of the evolution of the system in time . Together there are two generators of translations (P,H) and two of rotations (J,N) . We amalgamate these to have a four-dimensional translation and rotation .

The four dimensional notation we shall use is the following :the four vector :

$$x^{\mu} = x^{o}, x^{1}, x^{2}, x^{3} \quad , \quad \text{where} \quad x^{o} = t \ .$$

The metrical tensor $g_{\mu\nu}$ with :

$$g_{oo} = -1, \quad g_{kk} = 1 \ (k = 1,2,3) \quad g_{\mu\nu} = 0 \ (\mu \neq \nu) \ .$$

(Latin indices will mean summation over the three space indices) .

Correspondingly :

$$x_{o} = - x^{o} = - t \ .$$

$g_{\mu\lambda} x^{\mu} x^{\lambda}$ is the fundamental invariant .

We shall also on occasion use : $x_{4} = i x^{o} = i t$.

In this notation , we write for the generators of translations :

$$- \delta t \, . \, H + \delta \varepsilon \, . \, P = \delta \varepsilon^{\mu} P_{\mu}$$

which defines

$$P_{o} = - H = - P^{o}$$

Or, we may write $\delta \varepsilon_{\mu} P_{\mu}$, where the fact that both indices are below indicates that we are using $\delta \varepsilon = i \, \delta t$ and hence $P_{4} = i H = i P^{o}$.

For the generators of rotations , we write

$$\delta \omega \, . \, J + \delta V \, . \, N = \tfrac{1}{2} \, \delta \varepsilon^{\mu\nu} J_{\mu\nu}$$

where $\delta \varepsilon^{12} = \delta \omega_{3}$, $\delta \varepsilon^{ko} = \delta V^{k}$

$$J_{ko} = N_{k}$$

(Since $\delta\varepsilon^{\mu\nu}$ is antisymmetric, no diagonal terms appear .)

The use of the imaginary term x_4 and the cooresponding use of operators multiplied by i loses for us the Hermitian property . The rule is that when the index o appears, all the operators remain Hermitian ; when the index 4 appears , some of the operators are skew-Hermitian .

The commutation relations of P_μ and $J_{\mu\nu}$ are obtained from the generator

$$G = \delta\varepsilon^{\mu} P_\mu + \tfrac{1}{2}\delta\varepsilon^{\mu\nu} J_{\mu\nu}$$

by the methods of section 2 (p.31) and the results for four dimensions are identical with those for three (pp.31,32) :

$$\left[P_\mu, P_\nu\right] = 0$$

$$\frac{1}{i}\left[P_\mu, J_{\nu\lambda}\right] = \delta_{\mu\lambda} P_\nu - \delta_{\mu\nu} P_\lambda$$

$$\frac{1}{i}\left[J_{\mu\nu}, J_{\kappa\lambda}\right] = \delta_{\nu\kappa} J_{\lambda\mu} - \delta_{\mu\lambda} J_{\kappa\nu} - \delta_{\nu\kappa} J_{\lambda\mu} + \delta_{\mu\kappa} J_{\lambda\nu} .$$

If we consider separately commutators containing a time component , we have

$$\frac{1}{i}\left[P_1, J_{14}\right] = -P_4$$

In three dimensional motation this is (cancelling an i) :

$$\frac{1}{i}\left[P_1, N_1\right] = H$$

or in general :

$$\frac{1}{i}\left[P_k, N_\ell\right] = \delta_{k\ell} H$$

For J, N commutators , we have :

$$\frac{1}{i}\left[J_1, J_2\right] = J_3$$

$$\frac{1}{i}\left[J_1, N_2\right] = N_3$$

$$\frac{1}{i}\left[N_1, N_2\right] = -J_3 \qquad (11.1)$$

We may make linear combinations of the J_k, N_k , so that the combinations have the commutation properties of the three dimensional angular momental : if we define

$$J_k^{(1)} = \tfrac{1}{2}\left(J_k + i N_k\right) ; \quad J_k^{(2)} = \tfrac{1}{2}\left(J_k - i N_k\right) ,$$

then , invirtue of the commutation relations (11.1) ,

$$\frac{1}{i}\left[J_1^{(1)}, J_2^{(2)}\right] = 0$$

$$\frac{1}{i}\left[J_1^{(1)}, J_2^{(1)}\right] = J_3^{(1)}$$

Hence the (2j + 1) dimensional real matrices that form a representation of the three dimensional angular momentum operators (where the angular momentum is j) also are a representation of the $J^{(1)}$, $J^{(2)}$. We now express J and N in terms of $J^{(1)}$ and $J^{(2)}$:

$$J = J^{(1)} + J^{(2)}$$

$$N = \frac{1}{i}\left(J^{(1)} - J^{(2)}\right) \qquad (11.2)$$

Since $J^{(1)}$ and $J^{(2)}$ are Hermitian , N is skew-Hermitian ; there exists no finite dimensional Hermitian representation for N . Hence , we have a further complication for field theory . We must not only describe the system by continuous operators in space-time with an infinite number of states , but we must also take into account the non-Hermiticity of the internal spin .

A further property of the $J_{\mu\nu}$ is their behavior under reflection of the space coordinate axes . The J_k are invariant , but since N enters the generator multiplied by δV (which becomes $-\delta V$) , $N \to -N$. This implies that under reflection in space $J^{(1)} \rightleftarrows J^{(2)}$. Hence , if we have a representation in which $J^{(1)}(k)$ and $J^{(2)}(\ell)$ occur (where $J^{(1)}(k)$ means that $J^{(1)}$ is represented by a matrix of dimensionality ($2k + 1$) , then for invariance $J^{(1)}(\ell)$ and $J^{(2)}(k)$ must also occur .

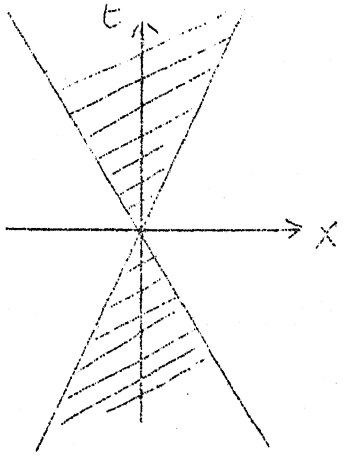
Let $\tilde{J} = J^{(1)}(k) + J^{(2)}(\ell)$, and R be the reflection operator . Then the invariant is : $J = \tilde{J} + R\tilde{J} = J^{(1)}(k) + J^{(2)}(\ell) + J^{(1)}(\ell) + J^{(2)}(k)$

The simplest non-trivial case will be for $k = 0$, $\ell = \frac{1}{2}$. This requires that $\tilde{J}$ be represented by a two dimensional matrix , J by a four dimensional matrix . For spin $\frac{1}{2}$ there is no representation in relativistic theory simpler than the Dirac (4×4) - representation .

12 - RELATIVISTIC DYNAMICAL THEORY .

The transformation functions we have considered until now were of the form $\langle a' t_1 | a'' t_2 \rangle$. Evidently in singling out the coordinate t , they are not

invariant . We must still discuss the evolution of a system from $t_2 \rightarrow t_1$, but the time as observed at t_1 will in general be measured along a time axis that is rotated with respect to the time axis at t_2 . However , both time measurements are made by real observers, consequently in ordinary three-dimensional space . In the four-dimensional coordinate system this is on a plane surface σ , whose normal is time-like .



Note : Points in space-time that can be connected by a light signal (inside the light cone) are separated by a time-like interval : $r^2 - t^2 < 0$. Points between which no signals can be transmitted are separated by a space-like interval (outside the light cone : $r^2 - t^2 > 0$.) Since $r^2 - t^2$ is invariant , space or time-like surfaces remain so under a Lorentz transformation .

The transformation functions will then be between an initial state on a plane space-like surface σ_2 to a final state on another plane space-like surface σ_1 : $\langle a' \sigma_1 | a'' \sigma_2 \rangle$.

The fundamental dynamical principle takes the form :

$$\delta \langle a' \sigma_1 | a'' \sigma_2 \rangle = i \langle a' \sigma_1 | \delta W_{12} | a'' \sigma_2 \rangle$$

where δW_{12} is the differential of a function

$$W_{12} = \int_{\sigma_2}^{\sigma_1} (dx) \mathcal{L}$$

(dx) means the volume element where the integration is over the four coordinates x^μ. We call $\mathcal{L}$ the Lagrange function operator and observe that W_{12} reduces properly to the non-relativistic case :

$$W_{12} = \int_{t_2}^{t_1} dt \int_{\tau} dx_1, dx_2\, dx_3\, \mathcal{L} = \int_{t_2}^{t_1} dt\, L \quad .$$

$\delta W_{12} = G_1 - G_2$, depending only on the initial and final states of the system . As in the non-relativistic case , $\mathcal{L}$ must give first order equations of motion . The variables of the non-relativistic theory , x(t) , become $\chi(x)$, where x indicates the four space-time coordinates . The proper form of $\mathcal{L}$ to give the linear equations is :

$$\mathcal{L} = \tfrac{1}{2} \left(\chi A^\mu \frac{\partial \chi}{\partial x^\mu} - \frac{\partial \chi}{\partial x^\mu} A^\mu \chi \right) - \mathcal{H}$$

the gradients replacing the time derivative of non-relativistic theory , since the four coordinates must enter symmetrically . (Observe that we are using at present $x^\mu = (x^0, x^1, x^2, x^3)$.

We now discuss the Lorentz invariance of the Lagrange function . $\mathcal{L}$ appears in :

$$\delta \langle a' \sigma_1 \mid a'' \sigma_2 \rangle = i \langle a' \sigma_1 \mid \delta \int_{\sigma_2}^{\sigma_1} (dx) \mathcal{L} \mid a'' \sigma_2 \rangle . \qquad (12.1)$$

σ_1 and σ_2 must be defined in some common coordinate system . In (12.1) , σ_1 , σ_2 and (dx) are invariant . Hence $\mathcal{L}$ must be invariant under Lorentz transformations , at least those that preserve the sense of time . As we shall see immediately below , there is no <u>continuous</u> Lorentz transformation

that goes through the center of the light cone , from future to past . We choose the χ to be Hermitian . $\mathcal{L}$ must be Hermitian . Hence , as in the non-relativistic theory , A^{μ} will be skew-Hermitian ; $\mathcal{H}$ is Hermitian .

We now introduce the Lorentz transformation explicitly . Under the Lorentz transformation

$$x^{\mu} \longrightarrow \bar{x}^{\mu}$$

$$\bar{x}^{\mu} \longrightarrow \ell^{\mu}_{\nu} x^{\nu}$$

The metric must be preserved :

$$g_{\mu\nu} \bar{x}^{\mu} \bar{x}^{\nu} = g_{\mu\nu} \ell^{\mu}_{\lambda} \ell^{\nu}_{\sigma} x^{\lambda} x^{\sigma}$$

$$g_{\mu\nu} = g_{\lambda k} \ell^{k}_{\mu} \ell^{\lambda}_{\nu}$$

$$- g_{00} = 1 = (\ell^{0}_{0})^2 - (\ell^{k}_{0})^2 \qquad \text{(Sum over k .)}$$

Hence $(\ell^{0}_{0})^2 = 1 + (\ell^{k}_{0})^2 \geqslant 1$

(since the transformation is real) . Then $\frac{\partial \bar{x}^{0}}{\partial x^{0}} = \ell^{0}_{0} \geqslant 0$. Hence the x^{0} cannot change sign by a continuous Lorentz transformation .

By the introduction of the new variables $\bar{x}^{\mu}$, the operators are also transformed . We indicate the operator which transforms the field variables as L :

$$\chi(\bar{x}) = L \chi(x)$$

The Lorentz transformation must leave the χ Hermitian . Hence L is real :

$$L^{*} = L .$$

We introduce the notation $\frac{\partial}{\partial x^{\mu}} = \partial_{\mu}$. Then : $\bar{\partial}_{\mu} = \ell^{\nu}_{\mu} \partial_{\nu}$

In $\mathcal{L}$ (where the A^μ are not changed since they are numerical matrices) we must have :

$$\bar{\chi}\, A^\mu\, \partial_\mu\, \bar{\chi} = \chi A^\mu \partial_\mu \chi$$
$$= \chi L^T A^\mu L\, \ell^\nu_\mu\, \partial_\nu \chi$$

This imposes on A^μ the condition

$$L^T A^\mu L\, \ell^\nu_\mu = A^\nu$$

or

$$L^T A^\mu L = \ell^\mu_\nu A \qquad (12.2)$$

Also , of course,

$$\mathcal{H}(L\chi) = \mathcal{H}(\chi)$$

If we pass to an infinitesimal Lorentz transformation, the equations above become :

$$\bar{x}^\mu = x^\mu + \varepsilon^\mu_\nu x^\nu$$

The $\varepsilon^{\mu\nu}$ have the property $\varepsilon^{\mu\nu} = -\varepsilon^{\nu\mu}$, i.e, in a Euclidian space they are antisymmetric .

$$L = 1 + \frac{i}{2}\varepsilon^{\mu\nu} S_{\mu\nu}$$

Since L is real , iS is real . Therefore S is pure imaginary . (12.2) puts the following relation between the S and A^μ :

$$i\,(A^\lambda S_{\mu\nu} + S_{\mu\nu} T\, A^\lambda) = \delta^\lambda_\mu A_\nu - \delta^\lambda_\nu A_\mu \qquad (12.3)$$

(The δ^λ_μ are the unit tensor .)

(The intermediate steps to derive (12.3) are :

$$(1 + \frac{i}{2}\varepsilon^{\mu\nu} S^T_{\mu\nu})\, A^\lambda\, (1 + \frac{i}{2}\varepsilon^{\mu\nu} S_{\mu\nu}) = A^\lambda + \varepsilon^\lambda_K A^K$$

We wish to equate the coefficients of $\varepsilon^{\mu\nu}$. Hence the right-hand side must be transformed :

$$\frac{i}{2}\varepsilon^{\mu\nu}\left[S^{T}_{\mu\nu} A^{\lambda} + A^{\lambda} S_{\mu\nu}\right] = \delta^{\lambda}_{\mu}\ \varepsilon^{\mu}_{\kappa} A^{\kappa} = \delta^{\lambda}_{\mu}\varepsilon^{\mu\nu} A_{\nu}$$

Taking into account that the factor $\frac{1}{2}$ on the left hand is to compensate for repetition in the double summation and that right hand side is still to be summed , we have (12.3) .

We now consider explicitly the action principle . In

$$\int_{\sigma_2}^{\sigma_1} (dx)\ \mathcal{L}$$

we shall vary the operators χ and the coordinates simultaneously . The variation of the coordinates is arbitrary, except that σ_1 and σ_2 must remain plane space-like surfaces . That is , on σ_1 and σ_2

$$\left.\begin{aligned} \delta x^{\nu} &= \varepsilon^{\nu} + x^{\mu}\varepsilon_{\mu}^{\ \nu} \\ \delta x_{\nu} &= \varepsilon_{\nu} + x^{\mu}\varepsilon_{\mu\nu} \end{aligned}\right\} \quad \varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu}$$

This implies

$$\partial_{\mu}\,\delta x_{\nu} + \delta_{\nu}\,\delta x_{\mu} = 0$$

The total variation in the operators will come from the arbitrary variations of the operators themselves and from the variation induced by the coordinate transformation . We write the total variation as the sum of an intrinsic variation of the field operators χ and as a variation coming from the rotation of the coordinates :

$$\delta(\chi) = \delta\chi + \frac{i}{2}\,\delta^{\nu}\delta^{\mu}_{\kappa}\, S_{\mu\nu}\,\chi \qquad (12.4)$$

that $S_{\mu\nu}$ is antisymmetric insures that only the rotational part of the coordinate transformation occurs in (12.4) .

Since the coordinate transformation is unrestricted except at the surfaces σ_1, σ_2 , there can be a volume dilation :

$$(d\,x) = (d\,x)\,\partial_\mu \delta x^\mu \qquad \text{(the divergence !)}$$

The coordinate transformation also induces a change in the gradient operator :

$$\delta(\partial_\mu) = \frac{\partial}{\partial(x^\mu + \delta x^\mu)} - \frac{\partial}{\partial x^\mu} = -\partial_\mu \delta x^\nu \partial_\nu$$

so that

$$\delta(\partial_\mu \chi) = \partial_\mu(\delta\chi) - (\partial_\mu \delta x^\nu)\partial_\nu \chi \qquad (12.5)$$

where it is to be noted that the first term on the right hand side is the gradient of the total variation (12.4)

The coordinate transformation is more general than a Lorentz transformation . However, since the Lagrange function is invariant under Lorentz transformations , we have considerable simplification : If the terms : $\partial_\mu \delta x^\nu (= \varepsilon^\nu_\mu$ for an infinitesimal Lorentz transformation) were constant and antisymmetric , the transformation would be a Lorentz transformation and the Lagrange function invariant . Hence, we need only retain terms in which the symmetric part of $\partial_\mu \delta x^\nu$ occurs, i e. $\frac{1}{2}(\partial_\mu \delta x^\nu + \partial_\nu \delta x^\mu)$ or terms in which $\partial_\mu \delta x^\nu$ is differentiated . These terms are contained entirely in (L2.5)

The variation gives :

$$\delta\int\mathcal{L}(dx) = \int(dx)\Big[\delta\chi A^\mu\partial_\mu\chi - \partial_\mu\chi A^\mu\delta\chi - \delta\mathcal{H} + \partial_\mu\tfrac{i}{2}(\chi A^\mu\delta\chi - \delta\chi A^\mu\chi) + (\partial_\mu\delta x_\nu)T^{\mu\nu}$$

where $T^{\mu\nu}$ is defined :

$$T^{\mu\nu} = T^{\nu\mu} = g^{\mu\nu}\mathcal{L} - \tfrac{i}{2}\Big[(\chi A^{(\mu}\partial^{\nu)}\chi) - \partial^{(\nu}\chi A^{\mu)}\chi - \partial_\lambda\tfrac{i}{2}(\chi A^{(\mu}S^{\nu)\lambda}\chi - S^{(\nu\lambda}A^{\mu)}\chi\Big]$$

$T^{\mu\nu}$ comes entirely from the variation of the coordinates . The symmetrization of tensors has been introduced : $A^{(\mu}B^{\nu)} \equiv \frac{1}{2}(A^\mu B^\nu + A^\nu B^\mu)$.

Notes : Most of the terms come quite directly . $g^{\mu\nu}\mathcal{L}$ comes from the variation of (dx) . We have first $\partial_\mu\delta x^\mu\mathcal{L}$ and lower the index by writing $\partial_\mu\delta x_\nu g^{\mu\nu}\mathcal{L}$.

Symmetrization of $\partial_\mu\delta x_\nu$ has been taken into the tensor indices.

The terms containing $S^{\mu\lambda}$ require an integration by parts. Consider the term coming from $\frac{1}{2}\chi A^\lambda\partial_\lambda\chi$:

$$\frac{1}{2}\frac{1}{2}\partial_\lambda\partial_\mu\delta x_\nu\chi A^\lambda S^{\mu\nu}\chi = \frac{i}{4}\partial_\lambda\partial_\mu\delta x_\nu\chi A^\mu S^{\lambda\nu}\chi$$
$$= \frac{i}{4}\frac{1}{2}\Big(\partial_\lambda\big[\partial_\mu\delta x_\nu + \partial_\nu\delta x_\mu\big]\Big)\chi A^\mu S^{\lambda\nu}\chi$$

Here, upon integration by parts , the integrated term will contain

$$\partial_\mu\delta x_\nu + \partial_\nu\delta x_\mu = 0 \qquad \text{on} \quad \sigma_1, \sigma_2$$

Hence under the integral sign we are left with only

$$-\frac{i}{4}\partial_\mu\delta x_\nu\partial_\lambda(\chi A^{(\mu}S^{\nu)\lambda}\chi)$$

As will appear below $T^{\mu\nu}$ is the stress-energy-momentum tensor .

When we integrate the last term by parts, we obtain

$$\int (dx)\left[\partial_\mu (\delta x_\nu T^{\mu\nu}) - \delta x\, \partial_\mu T^{\mu\nu}\right]$$

From the vanishing of the coefficient of δx_ν we have the conservation laws

$$\partial_\mu T^{\mu\nu} = 0 \tag{12.6}$$

The first terms give :

$$\delta \mathcal{H} = \delta\chi\, A^\mu \partial_\mu \chi - \partial_\mu \chi\, A^\mu \delta\chi \tag{12.7}$$

from which below we will derive the equations of motion .

The divergence terms contribute to the surface integrals and give us the generators for infinitesimal changes in χ and x^μ

Gauss' theorem in four dimensions gives us :

$$\int (dx)\, \partial_\mu B^\mu = \int_{\sigma_1} d\sigma_\mu B^\mu - \int_{\sigma_2} d\sigma_\mu B^\mu$$

Here for our physical interpretation the surface element is to be chosen to be a volume element in ordinary space . i.e., $d\sigma_\mu$ is the time component of $d\sigma_\mu$.

From the divergence terms we have then

$$G = G_\chi + G_x$$
$$G = \int d\sigma_\mu \tfrac{1}{2}(\chi A^\mu \delta\chi - \delta\chi A^\mu \chi)$$
$$G = \int d\sigma_\mu T^{\mu\nu} \delta x_\nu$$

where on the surfaces σ_1, σ_2 ,

$$\delta x_\nu = \xi_\nu + x^\mu \epsilon_{\mu\nu}$$

Hence :

$$G_\chi = \int d\sigma_\mu \, T^{\mu\nu} (\varepsilon_\nu + x^\lambda \varepsilon_{\lambda\nu})$$

$$= \varepsilon_\nu P^\nu + \varepsilon_{\mu\nu} J^{\mu\nu}$$

where

$$P^\nu = \int d\sigma_\mu \, T^{\mu\nu} \; ; \quad J^{\mu\nu} = \int d\sigma_\lambda \, (x^\mu T^{\lambda\nu} - x^\nu T^{\lambda\mu})$$

P^ν is the energy-momentum four vector . The vanishing divergence (12.6) states the conservation laws of energy and momentum .

$P^0 = \int d\sigma_0 \, T^{0\sigma}$ is the energy component of P^μ . Hence T^{00} is the energy density .

We now return to (12.7) to write explicitly the equations of motion . Analogously to (3.6) and (3.7) we write :

$$2 A^\mu \partial_\mu \chi = \frac{\partial_\varepsilon \mathcal{H}}{\partial \chi}$$

$$-2 A^{\mu T} \partial_\mu \chi = \frac{\partial_x \mathcal{H}}{\partial \chi}$$

as equivalent equations of motion .

Similarly for G_χ :

$$G_\chi = \int d\sigma_\mu \, \chi \, A^\mu \delta\chi = -\int d\sigma_\mu \, (A^{\mu T} \delta\chi) \chi$$

Again, by dividing A^μ into an antisymmetric and symmetric part, we divide the variables into two classes . We shall write $A^\mu = a^\mu + s^\mu$ and the variables of the first kind associated with a^μ we shall write ϕ ;

those of the second kind associated with S^μ, we write ψ.

As in (3.8) we now decompose the equations of motion into the two sets :

$$2\, a^\mu \partial_\mu \phi = \frac{\partial_t \mathcal{H}}{\partial \phi} = \frac{\partial_x \mathcal{H}}{\partial \phi}$$

$$2\, s^\mu \partial_\mu \psi = \frac{\partial_t \mathcal{H}}{\partial \psi} = -\frac{\partial_x \mathcal{H}}{\partial \psi}$$

$$G_\chi = G_\phi + G_\psi$$

$$G_\phi = \int d\sigma_\mu\, \phi\, a^\mu\, \delta\phi = \int d\sigma_\mu\, (a^\mu\, \delta\phi)\, \phi$$

$$G_\psi = \int d\sigma_\mu\, \psi\, s^\mu\, \delta\psi = -\int d\sigma_\mu\, (s^\mu\, \delta\psi)\, \psi$$

For an observer in terms of a coordinate system intrinsic to the surface σ, the generators are :

$$G_\phi = \int d\sigma_0\, \phi\, a^0 \delta\phi \; ; \quad G_\psi = \int d\sigma_0\, \psi\, s^0\, \delta\psi$$

where the time-like surface element $d\sigma_0$ normal to the plane three dimensional surface, is the volume element in ordinary space. If we decompose the equations of motion so as to make the time dependence explicit, we have :

$$2\, a^0\, \partial_0 \phi = \frac{\partial \mathcal{H}}{\partial \phi} - 2\, a^\kappa\, \partial_\kappa \phi \qquad (12.8)$$

Hence in both the generators and the equations of motion, a^0 requires special consideration. Only if a^0 is non-singular, so that we can solve (12.8) for $\partial_0 \phi$ are the equations (12.8) true equations of motion. If a^0 is singular, some are equations of motion, some equations of contraint. Similarly, in the generator, if a^0 is singular, not all the field variables

are independent .

We now repeat the arguments we made when we first introduced our variables of the first and second kind (p.4o sq.) . From the symmetry of left and right derivatives and the equivalent forms of G_ϕ :

$$\left[\delta\phi, \phi\right] = 0 \qquad ; \qquad \left[\delta\phi, \psi\right] = 0$$

From the equivalent forms of G_ψ . $\left\{\delta\psi, \psi\right\} = 0$. The opposite signs of the right and left derivatives then requires that $\mathcal{H}$ be an even function of the variables ψ , but it may be arbitrary in the ϕ if

$$\left[\delta\psi, \phi\right] = 0 .$$

We now make the explicit assumption that the equations (12.8) are true equations of motion ($\det a^0 \neq 0$) . (This assumption is not in general verified for relativistic fields !) Then by the arguments of p.42., we may introduce the canonical form :

$$\frac{1}{i}\left[\phi , G_\phi\right] = \tfrac{1}{2}\delta\phi$$

(See p.45 for the factor $\frac{1}{2}$)

or :

$$\frac{1}{i}\left[\phi(x), \int_\sigma d\sigma'_0\, \phi(x')\ a^0\ \delta\phi(x')\right] = \tfrac{1}{2} \quad (x)$$

where x is a point on the surface σ . By the commutation law for $\phi, \delta\phi$

$$\int_\sigma d\sigma'_0\ \frac{1}{i}\left[\phi(x), \phi(x')\ a^0\right]\delta\phi(x') = \tfrac{1}{2}\delta\phi(x)$$

or

$$\frac{1}{i}\left[\phi(x), \phi(x')\ 2\ a^0\right] = \delta^0(x - x')$$

where $\delta^o(x - x')$ is defined by

$$\int d\sigma_o \; \delta^o(x - x') f(x') = f(x)$$

$\delta^o(x - x')$ is the time-like component of a four vector on a space-like surface, actually the three dimensional δ - function :

$$\delta (x^1 - x^{1'}) \delta (x^2 - x^2) \delta (x^3 - x^{3'})$$

Hence for Hermitian variables of the first kind :

$$\left[\phi (x) , \phi(x') \right] = \frac{i}{2} (a^o)^{-1} \delta^o(x - x')$$

For Hermitian variables ψ on a space-like surface

$$\left[\psi (x) , \psi(x') \right] = \frac{i}{2} (s^o)^{-1} \delta^o(x - x')$$

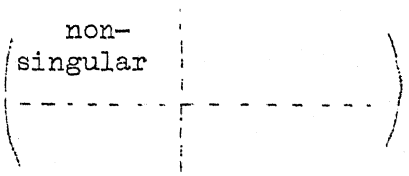
where we assume as on p.47 sq. that the matrix $i s^o$ is non-singular and positive definite .

The discussion just above was based on the assumption that a^o and S^o were non-singular . This is in general not true, and we must modify our interpretation of the theory to take this into account .

If a^o is singular, some linear combinations of $\delta\phi$ are zero and do not appear in the generator .

$$G_o = \int d\sigma_o \phi \, a^o \, \delta\phi$$

By a linear transformation we can arrange a^o so that the non-singular part is separated from the rest . Then the variations connected with the singular part will disappear .



In what follows by a^0 we mean the matrix arranged with the non-singular part separated from the rest . The equations of motion (12.8) then break into two sets . Some are true equations of motion for the variables associated with the non-singular part of a^0 . For the variables associated with the singular part of a^0 we do not have equations of motion but relations among the variables (equations of constraint)

In the simple cases the equations of motion and the equations of constraint are sufficient to determine all the variables . But there exist more complicated cases where some of the variables are not determined by the field equations . An example of this will be given in the following section .

The situation for variables of the second kind is similar . s^0 may be arranged so that the non-singular part is separated from the rest . The non-singular part of is^0 is positive definite , the whole is non-negative (eigenvalues either positive or zero) . We shall not meet any situations where the field variables are not determined by the field equations .

13 - EXAMPLES OF RELATIVISTIC FIELDS .

We find a major divergence from the non-relativistic theory in the fact that not all field components are kinematically independent . We see this now in some concrete examples .

I. Consider a field of the first kind with one scalar and four vector components $\phi(x)$, $\phi_\mu(x)$. The simplest invariant Lagrange function is :

$$\mathcal{L} = -\tfrac{1}{2}\left(\phi^{\mu}\,\partial_{\mu}\phi - \phi\,\partial_{\mu}\phi^{\mu}\right) - \frac{m}{2}\left(\phi^{\mu}\phi_{\mu} - \phi^{2}\right)$$

In the kinematic part the second term is obtained from the first by a partial integration . The invariance is clear . Variation with respect to ϕ^{μ} and ϕ give the equations of motion :

$$\delta\phi^{\mu} \longrightarrow \partial_{\mu}\phi = m\,\phi_{\mu}$$

$$\delta\phi \longrightarrow \partial_{\mu}\phi^{\mu} = m\,\phi$$

From the boundary terms :

$$G = \int_{\sigma} d\sigma_{\mu}\,\frac{1}{2}\left(\phi\,\delta\phi^{\mu} - \phi^{\mu}\,\delta\phi\right)$$

We have five field variables . Which are the independent variables for which we have true equations of motion ?

$$G = \int d\sigma_{0}\,\frac{1}{2}\left(\phi\,\delta\phi^{0} - \phi^{0}\,\delta\phi\right)$$

Only $\delta\phi^{0}$ and $\delta\phi$ appear here . Hence they are the only ones that are kinematically independent . The field equations decompose into :

$$\left.\begin{aligned} \partial_{0}\phi &= m\,\phi_{0} \\ \partial_{0}\phi^{0} &= m\,\phi - \partial_{k}\phi^{k} \end{aligned}\right\} \quad \text{(equations of motions)}$$

and

$$\partial_{k}\phi = m\,\phi_{k} \qquad \text{(equation of constraint)}$$

The commutators of ϕ and ϕ^{0} may be determined from the generator . (ϕ and ϕ^{0} are canonically conjugate variables .)

$$G_{\phi^{0}} = \int d\sigma_{0}\;\phi\,\delta\phi^{0} \qquad \text{(analogous to } \textstyle\sum p\,\delta q \text{)}$$

so

$$[\phi(x), \phi(x')] = 0$$

$$[\phi^0(x), \phi^0(x')] = 0$$

$$\frac{1}{i}[\phi(x), \phi^0(x')] = \delta^0(x - x')$$

Then from the field equations :

$$[\phi(x), \phi_k(x')] = 0$$

$$\frac{1}{i}[\phi_k(x), \phi^0(x')] = \frac{1}{im}[\partial_k \phi(x), \phi^0(x')] = \frac{1}{im}\partial_k \delta^0(x - x')$$

The more complicated structure arises from the fact that we have a commutator of variables that are not canonically conjugate . By elimination among the field equations , we may obtain :

$$(\partial_\mu^2 - m^2)\phi = 0 \quad ; \quad (\partial_\mu^2 - m^2)\phi_\mu = 0$$

II. We now go to a more complicated example : we have IO field components ,the vector ϕ_μ and the antisymmetric tensor $F_{\mu\nu} = -F_{\nu\mu}$. We write for the Lagrange function the simplest invariant structure linear in the derivatives :

$$\mathcal{L} = -\frac{1}{2}\left[F^{\mu\nu}\cdot\frac{1}{2}(\partial_\mu\phi_\nu - \partial_\nu\phi_\mu) - \phi_\nu\,\partial_\mu F^{\mu\nu}\right] + m\left[\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \phi^\mu\phi_\mu\right] \quad (13.1)$$

(Again the first two terms are connected by an integration by parts .)

The equations of motion are :

$$\left.\begin{aligned} \delta F^{\mu\nu} &\longrightarrow \partial_\mu\phi_\nu = m F_{\mu\nu} \\ \delta\phi_\nu &\longrightarrow \partial_\mu F^{\mu\nu} = m\phi^\nu \end{aligned}\right\} \quad (13.2)$$

and the generator is

$$G = \int d\sigma_\mu \, \tfrac{1}{2} \left[\phi_\nu \delta F^{\mu\nu} - F^{\mu\nu} \delta \phi_\nu \right]$$

Since in

$$G = \int d\sigma_{(0)} \, \tfrac{1}{2} \left[\phi_\nu \delta F^{0\nu} - F^{0\nu} \delta \phi_\nu \right]$$

($F^{00} = 0$) , the only independent components are three components ϕ_k and the and the three components F^{0k} , six out of ten .

The field equations decompose :

$$\left.\begin{aligned} \partial_0 \phi_k &= \partial_k \phi_0 + m F_{0k} \\ \partial_0 F^{0k} &= m \phi^k - \partial_\ell F^{\ell k} \end{aligned}\right\} \quad \text{(six equations of motion)}$$

and

$$\left.\begin{aligned} \partial_k \phi_\ell - \partial_\ell \phi_k &= m F_{k\ell} \\ \partial_k F^{k0} &= m \phi^0 \end{aligned}\right\} \quad \text{(four equations of constraint)}$$

We are still in the simple category where the equations serve to determine all the variables .

For the canonical variables the commutation relations are :

$$\left[\phi_k(x), \phi_\ell(x') \right] = 0$$

$$\left[F^{0k}(x), F^{0\ell}(x') \right] = 0$$

$$\frac{1}{i} \left[\phi_k(x), F^{0\ell}(x') \right] = -\delta_{\ell k} \, \delta^0(x - x')$$

and for the non-canonical variables :

$$\frac{1}{i}\left[\phi_k(x)\,,\phi^{o}(x')\right] = \frac{1}{im}\partial_{\ell}\left[\phi_k(x),\, F^{\ell o}(x')\right] = \frac{1}{m}\,\partial'_k\,\delta^{o}(x-x')$$

We note also at this point for comparison with the next case, that the three independent spatial components of ϕ_k and F^{ok} imply three independent components of the angular momentum vector .

If the divergence of ϕ_k vanishes, we may eliminate $F_{\mu\nu}$ from (13.2) to obtain the equation :

$$\partial^2_{\mu}\phi_{\nu} = m^2\,\phi_{\nu}$$

III. We have an example of the more complicated situation if we pass to degenerate form of the vector-antisymmetric tensor field by letting $m = 0$. (photon field) . We carry out the limiting process in the Lagrange function after redefining $F_{\mu\nu}$:

$$m\,F_{\mu\nu} \longrightarrow F_{\mu\nu}$$

and multiplying $\mathcal{L}$ by m . We also adopt a conventional notation by renaming the vectors : $\phi_{\mu} \longrightarrow A_{\mu}$. (13.1) becomes :

$$\mathcal{L} = -\tfrac{1}{2}\left[F^{\mu\nu}\cdot\tfrac{1}{2}\left(\partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}\right) - A_{\nu}\partial_{\mu}F^{\mu\nu}\right] + \frac{1}{4}\,F^{\mu\nu}F_{\mu\nu}$$

and (13.2) :

$$\left.\begin{aligned} F_{\mu\nu} &= \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \\ \partial_{\mu}F^{\mu\nu} &= 0 \end{aligned}\right\} \qquad (13.3)$$

(Maxwell's equations in empty space) .

Then

$$G = \int d\sigma_\mu \, \tfrac{1}{2} \left[A_\nu \, \delta F^{\mu\nu} - F^{\mu\nu} \, \delta A_\nu \right]$$

What are now the independent variables ?

$$G = \int d\sigma_0 \, \tfrac{1}{2} \left[A_k \, \delta F^{0k} - F^{0k} \, \delta A_k \right]$$

Of the ten field components we have six in G . We expect the six equations of motion and four of constraint to determine the ten field components . But what actually happens ?

$$\left.\begin{aligned} \partial_0 A_k &= F_{0k} + \partial_k A_0 \\ \partial_0 F^{0k} &= -\partial_\ell F^{\ell k} \end{aligned}\right\} \qquad \text{(equations of motion)}$$

$$\left.\begin{aligned} \partial_k A_\ell - \partial_\ell A_k &= F_{k\ell} \\ \partial_k F^{k0} &= 0 \end{aligned}\right\} \qquad \text{(equations of constraint)}$$

In the generator F^{0k} and A_k appear . $F^{\ell k}$, $\ell \neq 0$ does not appear , but is determined by the equations of constraint . A^0 does not appear in G and there is no equation of motion or constraint to determine A^0 . Furthermore, the equation of constraint which formerly served this purpose ($\partial_k F^{k0} = m\phi^0$) now constrains the F^{k0} . The three components F^{k0} are no longer independent. The three independent components of angular momentum now degenerate to two .

The indetermination of A^0 corresponds to the condition of gauge invariance In virtue of (13.3) , A_μ does not appear in $\mathcal{L}$, but only $\partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}$ (Recall that the first two terms in $\mathcal{L}$ are connected

by

by an integration /parts . The field strength are physicall significant,not the potential .) Hence $\mathcal{L}$ is invariant under the transformation

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \qquad (13.4)$$

where Λ is a scalar . From the very beginning we do not have a completely determined A_μ .

We have seen that only two components of F^{ok} are independent . Now we may choose $A^o = 0$, since it does not enter G, and with the freedom indicated by (13.4) set

$$\partial_k A^k = 0 .$$

We have then only two independent components of A^k . Hence in G we have four independent components (two p's and two q's for two degrees of freedom = = polarization) . We must then admit only variations that are consistent with these constraints .

In asymmetric form we write :

$$G = \int d\sigma_o \, A_k \, \delta F^{ok}$$

and

$$\frac{1}{i}\left[F^{ok}(x), \int d\sigma_o' A_\varrho(x') \delta F^{o\varrho}(x') \right] = \delta F^{ok}(x) \qquad (13.5)$$

The constraint on the variation leads to

$$\frac{1}{i}\left[F^{ok}(x), A_\varrho(x') \right] = \delta^k_\varrho \, \delta^o(x - x') - \partial^k \partial'_\varrho \frac{1}{4\, r/x - x'/} \qquad (13.6)$$

(We first rewrite the right-hand side of (13.5) as :

$$\delta F^{ok}(x) = \int d\sigma_o' \delta^k_\varrho \, \delta^o(x - x') \, \delta F^{o\varrho}(x')$$

and bring it to the left . We introduce the constraint $\partial_\varrho \delta F^{o\varrho} = 0$ with a

Lagrangian multiplier λ : ($\lambda = \lambda(x,x')$)

$$\int d\sigma' \left(\left[F^{ok}(x), A_\ell(x') \right] - i\delta^k_\ell \delta^o(x-x') \right\} \delta F^{o\ell}(x') + i\lambda^k \partial_\ell \delta F^{o\ell}(x') \Big)$$

The last term may be integrated by parts . The integrated term (volume integral in three dimensions !) vanishes if $\delta F^{o\ell}$ vanishes on a properly chosen surface . Then :

$$\frac{1}{i} \left[F^{ok}(x), A_\ell(x') \right] = \delta^k_\ell \delta^o(x - x') + \partial_\ell \lambda^k$$

λ^k is further determined by

$$\partial_\ell A^\ell(x) = 0$$

Then :

$$\lambda_k = -\partial_k \mathcal{D}_o(x - x')$$

where

$$\mathcal{D}_o = \frac{1}{4\pi |x - x'|}$$

Hence

$$\frac{1}{i} \left[F^{ok}(x), A_\ell(x') \right] = \delta^k_\ell \delta^o(x-x') - \partial^k \partial_{\ell'} \mathcal{D}_o (x - x')$$

The commutator for the divergenceless F^{ok} and A_ℓ contains only the divergenceless part of the δ -function .)

The additional terms in the commutator appear because of A^μ . If one takes the curl of A^μ to get $F^{\mu\nu}$, the additional term disappears . The commutator for E and H is not dependent on the choice of gauge .

14. REVERSAL OF SENSE OF TIME

The connection between spin and statistics is purely relativistic . The invariance under all Lorentz transformations (including the reflexion of the time axis) gives just enough information to connect particles of integral spin with variables of the first kind (B.E. fields) and particles of half-integral spin with variables of the second kind (F.D. fields)

As we have seen, the reverseal of the sense of time is an exceptional Lorentz transformation in that no continuous Lorentz transformation changes the sign of the x^0 axis . We introduce the time reversal as follows .

We suppose we have a Lorentz transformation

$$\bar{x}^\mu = \ell^\mu_\nu x^\nu$$

in which to now we have restricted $\ell^0_0 \geqslant 1$

$$\bar{\chi} = L\chi$$

For infinitesimal Lorentz transformations

$$L = 1 + \frac{i}{2}\varepsilon^{\mu\nu} S_{\mu\nu}$$

We now also introduce a transformation which reflects one axis, e.g.

$$\chi^1 \longrightarrow -\chi^1$$

Our system must be invariant under reflection of the spatial axes . We indicate the operator that inverts x^i by L_1.

In what follows we shall treat of Lorentz transformations that rotate the 0-axis to the 1-axis and then reflections L_1 . If we rotate the 0-axis to the position of the 1-axis and then reflect it, we have the same thing

as if we inverted the 1-axis and then rotated the 0-axis in the opposite sense to the 1-axis :

$$L_i \left(1 + i \varepsilon^{01} S_{01} \right) = \left(1 - \varepsilon^{01} S_{01} \right) L_1$$

$$L_1^{-1} S_{01} L_1 = - S_{01} \qquad (14.1)$$

When we use the imaginary time, we have a Euclidean space and the four axes are equivalent . One way to reverse the time axis is to rotate the 4-axis to the 1-axis, reflect by L_1 and then rotate back . The operator for the finite rotation is $e^{\frac{\pi}{2} i S_{14}}$

$$e^{-\frac{\pi i}{2} S_{14}} L_1 \; e^{\frac{\pi i}{2} S_{14}} = L_4$$

where L_4 has now the formal properties of time reflection . By (14.1)

$$L_4 = L_1 e^{\pi i S_{14}}$$

We now investigate the properties of L_4 . L_1 as a space reflection is a real (Hermitian) operator : $L_1^{\dagger} = L_1$

$$S_{14} = - i S_{10}$$

S_{10} is imaginary, since it ɔoccurs with a factor i in the real Lorentz transformation . Hence S_{14} is real .

Therefore

$$L_4^{\dagger} = L_1 \; e^{-\pi i S_{14}} = L_4 \; e^{-2\pi i S_{14}} = e^{2\pi i S_{14}} L_4$$

This is a rotation through an angle of 2π . Now S_{14} must have the same

eigenvalues as, say, S_{12} . If the spin is integral :

$$e^{2\pi i S_{12}} = L \cdot 1$$

If the spin is half integral

$$e^{2\pi i S_{12}} = -1 .$$

Hence for integral spin $L_4^* = L_4$;
For half-integral spin $L_4^* = -L_4$. } (14.2)

For half-integral spin, the operators change from Hermitian to skew-Hermitian .

So far we have the formal properties of time reversal in its relation to spin . We must now see its relation to statistics , i.e,the relations of the variables of the two kinds to the time-reversal formalism .

For any Lorentz transformation

$$L^T A^\mu L = \ell^\mu_\nu A^\nu$$

For the time reversal, this becomes :

$$L_4^T A^0 L_4 = - A^0 \qquad \left(\begin{array}{l} \chi^0 = -\chi^0 \\ \chi^k = \chi^k \end{array} \right)$$

For the matrix a^0 associated with variables of the first kind this requires that a real antisymmetric matrix change sign . This can be done with a real transformation .

For the matrix $-is^0$ associated with variables of the second kind, the situation is quite different . $-is^0$ is non-negative (the eigenvalues are $\geqslant 0$.) The sign cannot be reserved by a transformation with real numbers .

$$L_4^T \, (- i s^0) \, L_4 = - (- i s^0)$$

demands that L_4 be imaginary for F.D. fields (just the condition for half-integral spin of (14.2) !

The reversal of the sense of time must also leave the action principle invariant . In

$$\delta \langle 1 \rangle = i \langle 1 | \delta \int_{\sigma_1}^{\sigma_2} (dx)\, \mathcal{L} | 1 \rangle$$

the time reversal reverses the sense of the integration . By going to the complex conjugate algebra, this is compensated for in the change of the sign of i . We must examine what happens in the complex conjugate algebra when we make the transformation

$$x^0 \longrightarrow -x^0 \quad \text{and} \quad \chi \rightarrow L_4 \chi$$

The case of variables of the first kind offers no difficulty since the variables remain Hermitian . But the variables ofthe second kind become skew-Hermitian , and we must see what this implies . Since the ψ become skew-Hermitian in the complex-conjugate algebra, we write as the condition of invariance

$$\mathcal{L}(\phi, i\psi)^* = \mathcal{L}(\phi^*, \psi^*)$$

For variables of the first kind the kinetic term is

$$(\phi\, a^\mu \partial_\mu \phi)^* = \phi^* a^\mu \partial_\mu \phi^*$$

since the a^μ are real . The theory is kinematically invariant .

For variables of the second kind , the condition on the kinematic term is

$$(i\psi s^\mu \partial_\mu \psi)^* = \psi^\mu s^\mu \partial$$

This is fulfilled since s^μ is imaginary ; the change of sign of s^μ just compensates for the factor -1 .

In $\mathcal{H}$:

$$\mathcal{H}(\phi, i\psi)^* = \mathcal{H}(\phi^*, \psi^*)$$

Again, the matrices associated with variables of the first kind are real, and the invariance is obvious . Variables of the second kind appear quadratically, associated with an imaginary matrix . Hence again the sign changes compensate.

We may write also for the invariance

$$\mathcal{H}(\phi, i\psi)^* = \mathcal{H}^*(\phi^*, -i\psi^*)$$

where the star on $\mathcal{H}$ is to remind us that the signs of the imaginary matrices are to be changed . We can then rename the variables, since ϕ and ψ are here Hermitian :

$$\mathcal{H}^*(\phi, \psi) = \mathcal{H}(\phi, i\psi)$$

We may make a further statement . Since $\mathcal{H}$ is Hermitian

$$\mathcal{H}^*(\phi, \psi) = \mathcal{H}^{\mathsf{T}}(\phi, \psi)$$

$$\mathcal{H}^{\mathsf{T}}(\phi, \psi) = \mathcal{H}(\phi, i\psi)$$

where $\mathcal{H}^{\mathsf{T}}$ means transpose the individual operators and invert their order.

The ψ occur in pairs . Hence going from $\psi \rightarrow i\psi$ means changing the sign of a term . But $\mathcal{H}$ will always be constructed so as to be antisymmetric in the ψ variables .

15. - THE ELECTRON FIELD .

We now consider specific properties of special kinds of systems .Suppose that $\mathcal{L}$ is invariant under the transformation :

$$\chi \rightarrow \overline{\chi} = e^{i\lambda\varepsilon}\chi$$

where λ is an arbitrary real number , not a function of x^μ. Although the operator is analogous to a rotational operator, it is not a rotation in space-time . We suppose ξ commutes with all other matrices of the theory since it is not connected with space-time properties .

$e^{i\lambda\xi}$ must be real to preserve Hermiticity, ξ imaginary .

For invariance of $\mathcal{L}$:

$$\chi A^\mu \partial_\mu \chi = \bar{\chi} A^\mu \partial_\mu \bar{\chi} = \chi e^{i\lambda\xi^\dagger} A^\mu e^{i\lambda\xi} \partial_\mu \chi$$

Since ξ commutes with A^μ,

$$e^{i\lambda\xi^\dagger} \cdot e^{i\lambda\xi} = 1$$

$$\xi^T = -\xi = \xi^* \qquad \xi^\dagger = \xi$$

We now go to the infinitesimal transformation

$$\bar{\chi} = (1 + \delta\lambda\xi)\chi = \chi - \delta\chi$$

$$\delta\chi = -i\delta\lambda\xi\chi$$

We introduce this variation into the action principle . We set $\lambda = \lambda(\chi)$ and carry out the infinitesimal transformation at each point in space-time . As before with time dependent rotation (p.128) and the Lorentz invariance, the invariance of $\mathcal{L}$ under the transformation makes only a small addition appear in the variation : only those terms appear where $\delta\lambda$ is differentiated.

$$\delta\mathcal{L}_{\delta\lambda} = -i\tfrac{1}{2}\chi A^\mu \xi \chi \partial_\mu \delta\lambda + i\tfrac{1}{2}\xi\chi A^\mu \chi \partial_\mu \delta\lambda = -i\chi A^\mu \xi\chi \partial_\mu \delta\lambda$$

We define the vector $j^\mu = -i\chi A^\mu \xi \chi$,

$$\delta\mathcal{L}_{\delta\lambda} = \partial_\mu(\delta\lambda j^\mu) - \partial_\mu j^\mu \delta\lambda$$

In the interior of the region, the factor of $\delta\lambda$ must vanish . Hence the vector j^{μ} has the conservation law

$$\partial_{\mu} j^{\mu} = 0 .$$

On the boundary, the divergence term gives :

$$G_{\lambda} = \int d\sigma_{0}\ j^{0}$$

We define the operator $Q = \int d\sigma_{\mu}\ j^{\mu}$

In a particular coordinate system with

$$Q = \int d\sigma_{0}\ j^{0} ,$$

j^{0} is the charge density .

From the generator G_{λ} we can derive commutation relations :

$$\delta_{\lambda}\chi = -i\,\delta\lambda\,\varepsilon\chi$$

$$\frac{1}{i}\left[\chi, G_{\lambda}\right] = \delta_{\lambda}\chi$$

$$\frac{1}{i}\left[\chi, Q\right] = -i\,\varepsilon\chi$$

$$\left[\chi, Q\right] = \varepsilon\chi$$

Since ε is Hermitian , its eigenvalues are real . Since $\varepsilon^{T} = -\varepsilon$ the eigenvalues appear in pairs, $\pm$ the same absolute magnitude .

$$\left(\varepsilon\chi = \varepsilon'\chi \longrightarrow \varepsilon^{*}\chi^{*} = \varepsilon'\chi^{*} \longrightarrow -\varepsilon\chi^{*} = \varepsilon'\chi^{*}\right.$$

We have automatically the symmetry between $\pm$ charge .

More in detail : There must exist a linear transformation that reshuffles the eigenvalues so that $\varepsilon \longrightarrow -\varepsilon$. Let the matrix be c. Then :

$$c^T \mathcal{E} c = -\mathcal{E}$$

$$\chi \longrightarrow c\chi$$

$$\mathcal{L} \longrightarrow \mathcal{L}$$

$c^T c = 1$ since $\chi\chi$ terms in $\mathcal{H}$ must be invariant .

But $Q \rightarrow -Q$ since :

$$j^{\mu} \rightarrow -i\chi c^T A^{\mu} \mathcal{E} c \chi$$

and

$$c^T \mathcal{E} c = -\mathcal{E}$$

$$j^{\mu} \longrightarrow -j^{\mu}; \quad Q \rightarrow -Q .$$

Let C now be a unitary operator generating this transformation (not the numerical matrix c) .

$$C^{-1} \chi C = c\chi$$

$$C^{-1} Q C = -Q .$$

The transformation is very much like the reflection of an axis .

$$C^2 = 1 \text{ - its eigenvalues are } \pm 1 .$$

The states of a system can be specified by the ± 1 eigenvalues of C .

The simplest example of the invariance under this transformation will be the case where $\mathcal{E}$ is of dimensionality 2 . $\mathcal{E}$ is antisymmetric and pure imaginary .

Therefore

$$\mathcal{E} = e \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

The χ must decompose into two sets, corresponding to the two rows of ε

$$\chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}$$

$$\varepsilon\chi = ie \begin{pmatrix} -\chi_2 \\ \chi_1 \end{pmatrix}$$

(Here we mean by χ_1 and χ_2 two sets of components , equal in number . The components χ_1 will have among themselves different space-time properties .)

The eigenvalues of ε are then $\pm$ e .

$$j^\mu = -i\chi A^\mu \varepsilon \chi = -e\left[\chi_1 A^\mu \chi_2 - \chi_2 A^\mu \chi_1\right]$$

Terms $\chi A^\mu \partial_\mu \chi$ in $\mathcal{L}$ will split into pairs :

$$\chi_1 A^\mu \partial_\mu \chi_1 + \chi_2 A^\mu \partial_\mu \chi_2$$

The effect of ε is to double everything in the space-time world . We work most conveniently with ε diagonal . This requires the use of non-Hermitian variables . Since the variables occur in pairs, we can use canonical formalism. We simultaneously diagonalize ε and go to canonical variables :

$$\chi_{+-} = \chi_1 - i\chi_2$$

$$\chi_- = \chi_1 + i\chi_2 = \chi_+^\dagger$$

$$\left[\chi\,;\,Q\right] = \varepsilon\chi \longrightarrow \left[\chi_+\,,\,Q\right] = e\chi_+$$

$$\left[\chi_-\,,\,Q\right] = -e\chi_-$$

This is analogous to the commutation properties of annihilation and creation operators :

$$\left[\psi\,,\,N\right] = \psi$$

$$\left[\psi^\dagger\,,\,N\right] = -\psi^\dagger$$

However, here only charge is annihilated or created, not the particle .

χ_+ reduces charge by e ; χ_- creates charge of e :

$$Q\chi_+ = \chi_+(Q - e) \qquad \text{(from the commutator.)}$$

Annihilation of charge may mean annihilation of a charge of + e or creation of a charge of - e .

In terms of the variables χ_+ , χ_- :

$$j^\mu = -\frac{ie}{2}\left[\chi_- A^\mu \chi_+ - \chi_+ A^\mu \chi_-\right]$$

$$\mathcal{L}_{kin} = \frac{1}{4}\left[\chi_- A^\kappa \partial_\mu \chi_+ - \partial_\mu \chi_- A^\kappa \chi_+\right] + \frac{1}{4}\left[\chi_+ A^\kappa \partial_\mu \chi_- - \partial_\mu \chi_+ A^\kappa \chi_-\right] \quad (15.1$$

We may get the generators directly from (15.1)

$$G_{\chi\pm} = \int d\sigma_\mu \tfrac{1}{2}\left[\chi_- A^\mu \delta\chi_+ - \delta\chi_- A^\mu \chi_+\right]$$

and the asymmetric form

$$G_{\chi\pm} = \int d\sigma_0 \chi_- A^0 \delta\chi_+ \qquad \text{, analogous to (3.9) and the following equations .}$$

For the commutation relations we use the submatrix of A^0 that is non-singular and obtain for the kinematically independent field components :

$$\left\{\left[\chi_+(x), \chi_-(x')\right]\right\} = i\,(A^0)^{-1}\,\delta^3(x - x')$$

$$\left\{\left[\chi_\pm(x), \chi_\pm(x')\right]\right\} = 0 .$$

The eigenvalues of Q are integral multiples of e .(The operator form $\chi_- A^\mu \chi_+$ is the same as $q^\dagger q$. We found the eigenvalues of $\Lambda q^\dagger q$ on p.108 sq

The invariance under the transformation $\chi \longrightarrow e^{i\Lambda\xi}\chi$ introduced charge

We now go to considerations which introduce spin in examining the nature of the matrices A^μ.

A^μ is a matrix which in products of operators gives vector terms in $\mathcal{L}$. Clearly scalar terms should also appear, say in $\mathcal{H}$. Let B be a matrix associated with scalar terms so that in $\mathcal{H}$ a term $\chi B \chi$ appears. (We are now speaking in terms of Hermitian operators as in our treatment of Lorentz invariance .) B must be Hermitian in order that $\mathcal{L}$ be Hermitian :

$$(\chi B \chi)^\dagger = \chi B \chi \qquad B^\dagger = B .$$

Just as A^μ decomposes into parts associated with variables of the first and second kind, B must also decompose :

$$B = B^{(1)} + B^{(2)}$$

where $B^{(1)}$ is real and symmetric , $B^{(2)}$ antisymmetric and imaginary to correspond to the commutation properties and invariance under reflection of the sense of time .

Under Lorentz transformations $\chi B \chi$ must be invariant :

$$\chi \longrightarrow L \chi$$

$$L^T B L = B .$$

We can use B to put $L^T A^\mu L = \ell^\mu_\nu A^\nu$ in a more convenient form, i.e., in the form of a similarity transformation (which preserves algebraic relation

If B exists as a non-singular matrix

$$L^T = B L^{-1} B^{-1} . \qquad (15.2)$$

Then :

$$L^{-1} \left(B^{-1} A^{\mu} \right) L = \ell^{\mu}_{\nu} B^{-1} A^{\nu} .$$

Our problem is now to investigate the properties of the matrix $B^{-1} A^{\mu}$.

$$\left(B^{-1} A^{\mu} \right)^{+} = - A^{\mu} B^{-1} = - B \left(B^{-1} A^{\mu} \right) B^{-1} .$$

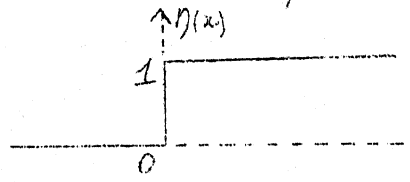
We now consider properties that the A^{μ} must have and combine A^{μ} with B to make the invariance relations convenient, eventually deriving explicit conditions for $B^{-1} A^{\mu}$.

We first consider the field radiating from a point at the origin of the space-time coordinate system . The radiation field will be a solution of the field equation

$$2 A^{\mu} \partial_{\mu} \chi = \frac{\partial_{\mu} \mathcal{H}}{\partial \chi} \qquad (15.3)$$

The A^{μ} are by no means arbitrary but must be such that the field has a discontinuity at the surface of the light cone $r = t$, for the field is zero outside the light cone, finite inside the light cone . Since $\partial_{\mu} \chi$ will be infinite at $-g_{\mu\nu} x^{\mu} x^{\nu} = 0$ $\left(t^2 - r^2 = 0 \right)$, the A^{μ} must vanish there . We show this formally as follows .

If $\chi(x)$ is a solution of (15.3), then $\eta \left(-g_{\mu\nu} x^{\mu} x^{\nu} \right) \chi(x)$ is also a solution, where $\eta(x)$ is the Heaviside function defined as :



$$\eta(-0) = 0 , \qquad \eta(+0) = 1 .$$

$$\partial_{\mu} \eta \left(-g_{\mu\nu} x^{\mu} x^{\nu} \right) = -2 x_{\mu} \delta' \left(-g_{\mu\nu} x^{\mu} x^{\nu} \right) .$$

Substitution of $\eta \left(-g_{\mu\nu} x^{\mu} x^{\nu} \right) \chi$ into (15.3) leads to

$$2A^{\mu}\left[-2x_{\mu}\chi^{\delta}\left(-g_{\mu\nu}\ x^{\mu}x^{\nu}\right)\right]+2A^{\mu}\eta\left(-g_{\mu\nu}\ x^{\mu}x^{\nu}\right)\partial_{\mu}\chi\ =$$
$$=\frac{\partial_{\rho}\mathcal{H}}{\partial\chi}$$

or using (15.3)

$$A^{\mu}\ x_{\mu}\chi^{\delta}\left(-g_{\mu\nu}\ x^{\mu}x^{\nu}\right)=0\ .$$

Hence $A^{\mu}x_{\mu}\chi$ must vanish on the light cone . We might expect $A^{\mu}x_{\mu}\chi$ to be closely relatdd to $-g_{\mu\nu}\ x^{\mu}x^{\nu}$ since this vanishes on the light cone .

We multiply on the left by B^{-1} and consider the matrix $B^{-1}A^{\mu}x_{\mu}$. Let : $X=B^{-1}A^{\mu}x_{\mu}$ and

$$X^{m}+\lambda_{1}X^{m-1}+\ldots\lambda_{m}=0 \qquad (15.4)$$

be the equation of lowest degree satisfied by the matrix X . We carry out a Lorentz transformation of all the variables in (15.4) . The X are invariant, hence the λ_{i} are also invariant . The λ_{i} must then be functions of the invariant $g_{\mu\nu}x^{\mu}x^{\nu}$.

The x_{μ} in X is arbitrary . As a consequence, (15.4) is an identity in the x_{μ} , each term containing x_{μ} to the same power . This means that the powers of X must differ by two (λ is a function of $g_{\mu\nu}x^{\mu}x^{\nu}$) .(15.4) is either odd or even .

Let us suppose $m=2n$. Then the last term λ_{n} is a function of $(\ g_{\mu\nu}x^{\mu}x^{\nu}\ )^{n}$. The last term of an equation such as (15.4) is always the product of the eigenvalues of X . Hence the eigenvalues are proportional to $g_{\mu\nu}x^{\mu}x^{\nu}$. (The eigenvalues then vanish on the light cone, thus corresponding

to the discontinuity in the solution of the equations of motion .)

The simplest system has an equation of degree 2 , the next simplest, 3 . The simplest case is then :

$$(B^{-1} A^{\mu} x_{\mu})^2 = g_{\mu\nu} x^{\mu} x^{\nu} ,$$

where a proportionality constant is absorbed into A^{μ} . (This is consistent with $B^{-1} A^{\mu} x_{\mu} \chi \delta (-g_{\mu\nu} x^{\mu} x^{\nu}) = 0$.

We now consider explicitly F.D. systems and define :

$$B^{-1} A^{\mu} = i \gamma^{\mu}$$ to conform with the usual notation .

$$(\gamma^{\mu} x_{\mu})^2 = - g^{\mu\nu} x_{\mu} x_{\nu}$$

$$\tfrac{1}{2} (\gamma^{\mu} \gamma^{\nu} + \gamma^{\nu} \gamma^{\mu}) x_{\mu} x_{\nu} = - g^{\mu\nu} x_{\mu} x_{\nu}$$

so that

$$\gamma^{\mu} \gamma^{\nu} + \gamma^{\nu} \gamma^{\mu} = - 2 g^{\mu\nu}$$

$$(\gamma^{0})^2 = 1 \qquad (\gamma^{\kappa})^2 = - 1 .$$

For Fermi-Dirac systems $-i A^{0}$ ($= -i s^{0}$) must be positive definite (or non-negative) . The simplest possibility is $- i A^{0} = 1$. Then $B^{-1} A^{0} = i \gamma^{0}$ reduces to $B = \gamma^{0}$.

We had supposed B to occur in the Hamiltonian function as $\chi B \chi$. With the non-Hermitian variables of a charged field, this would correspond to $\chi_{-} B \chi_{+}$ (with appropriate symmetrization) . We now rename the $\chi_{\pm}$:

$$\left. \begin{aligned} \chi_{+} &= \psi \\ \chi_{-} &= \psi^{+} \end{aligned} \right\} \quad \chi_{-} B \chi_{+} \longrightarrow \psi^{+} \gamma^{0} \psi$$

We introduce $\psi^- = \psi^\dagger \gamma^0$

The invariant term becomes : $\bar{\psi}\,\psi$

From a geometrical point of view, γ^0 is the metric . We do not have Hermitian variables in a representation of finite order, but the metric serves to restore the necessary properties of reality .

Consistent with $B^{-1} A^0 = i\gamma^0$

$$A^\mu = i\gamma^0\gamma^\mu$$

The Lagrange function becomes

$$\mathcal{L} = \frac{1}{4}\left[\bar{\psi}\,,\, i\gamma^\mu\partial_\mu\psi\right] - \frac{1}{4}\left[\psi\,,\, i\gamma^\mu\partial_\mu\bar{\psi}\right] - m\tfrac{1}{2}\left[\bar{\psi},\psi\right]$$

where the kinematic terms are derived by substitution of $\psi^\dagger,\psi$, $i\gamma^0\gamma^\mu$ respectively for χ_+,χ_- , A^μ and the Hamiltonian is chosen as the simplest invariant .

The field equations are then

$$\left(\gamma^\mu \frac{1}{i}\partial_\mu + m\right)\psi = 0$$

$$\bar{\psi}\left(-\gamma^\mu \frac{1}{i}\partial_\mu^\dagger + m\right) = 0$$

The current operator is

$$j^\mu = e\,\frac{i}{2}\left[\bar{\psi}\,,\,\gamma^\mu\psi\right]$$

and the commutation relations

$$\left\{\psi(x)\,,\,\bar{\psi}(x')\gamma^0\right\} = \delta^0(x-x')\,, \qquad (15.5)$$

or

$$\left\{\psi(x)\,,\,\psi^\dagger(x')\right\} = \delta^0(x-x')\,.$$

16. - INTERACTION BETWEEN FIELDS.

We treat the interaction between fields in the case of the Dirac and Maxwell fields as a fundamental example .

$$\mathcal{L} = -\tfrac{1}{2}\left[\bar{\psi}\cdot i\gamma^{\mu}\partial_{\mu}\psi - i\partial_{\mu}\bar{\psi}\gamma^{\mu}\cdot\psi\right] - m\,\bar{\psi}\cdot\psi$$
$$+\tfrac{1}{2}F^{\mu\nu}(\partial_{\mu}A_{\nu}-\partial_{\nu}A_{\mu}) - A_{\nu}\cdot\partial_{\mu}F^{\mu\nu} + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + j^{\mu}A_{\mu}$$

Here the first line is the Lagrange function for a Dirac field, the next three terms for a Maxwell field, as we have already seen . As a usual term $\bar{\psi}\cdot\psi$ means half the commutator for the Dirac field, half the anticommutator for the Maxwell field . The last term is the coupling term, the simplest invariant in the operators of both fields .

$$j^{\mu} = e\,\bar{\psi}\cdot\gamma^{\mu}\psi = \frac{e}{2}\left[\bar{\psi},\gamma^{\mu}\psi\right]$$

By varying the operators, we get the set of field equations :

$$\begin{aligned}
\delta\bar{\psi} &\longrightarrow \gamma^{\mu}\left(\frac{1}{i}\partial_{\mu} - eA_{\mu}\right)\psi + m\psi = 0\\
\delta\psi &\longrightarrow \bar{\psi}\left[\gamma^{\mu}\left(-\frac{1}{i}\partial_{\mu}^{T} - eA_{\mu}\right) + m\right] = 0\\
\delta A_{\mu} &\longrightarrow \partial_{\nu}F^{\mu\nu} = j^{\mu}\\
\delta F^{\mu\nu} &\longrightarrow \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} = F_{\mu\nu}
\end{aligned} \qquad (16.1)$$

The Lagrange function has the property of general gauge invariance . (This will be defined below .) After the introduction of non-Hermitian variables , the Dirac field is invariant under the transformation

$$\psi \rightarrow e^{i e\lambda}\psi$$
$$\bar{\psi} \rightarrow e^{-i e\lambda}\bar{\psi}$$

λ may be a function of x, (This is the property of general gauge invariance) provided that there is an accompanying transformation of the Maxwell field . From the differentiation of the $\lambda(x)$ we get from the kinematic part of the electron field an extra term in the Lagfange function :

$$\mathcal{L} \rightarrow \mathcal{L} - j^{\mu} \partial_{\mu} \lambda$$

We may compensate for this term and have an invariant Lagrange function if we make simultaneously the transformation :

$$A_{\mu} \longrightarrow A_{\mu} + \partial_{\mu} \lambda$$

As we have seen before, this does not affect the $F^{\mu\nu}$ since the curl of a gradient vanishes .

Even with the coupled fields, we still have the fundamental arbitrariness associated with the Maxwell field by itself and we must choose the kinematically independent field components . Since the generators come only from divergence terms in $\mathcal{L}$, they are not changed by the coupling term .

$$G = \int d\sigma_o \left[\bar{\psi} i \gamma^{o} \delta\psi - F^{ok} \delta A_k \right]$$

For the Maxwell field we are in the same situation as with the uncoupled field as far as the equations of motion and equations of constraint are concerned, except that one equation of restraint now implies that the Maxwell /is field/partly determined by the Dirac field :

$$\partial_k F^{ok} = j^{o} \qquad (16.2)$$

As in the case of the uncoupled field, we choose the gauge so that for a particular observer

$$\partial_k A^k = 0, \qquad (16.3)$$

but here we shall see what is the significance of this choice .

In general, we may write a field vector V as the sum of a transverse part and a longitudinal part (where by longitudinal part we mean parallel to the direction of propagation .)

$$V^k = V_L^k + V_T^k$$

Here V_L^k can be derived as the gradient of some scalar function :

$$V_L^k = \partial^k \Lambda$$

so that

$$\partial_K V_\ell^L - \partial_\ell V_k^L = 0,$$

and the divergence of V_T^k vanishes :

$$\partial_K V_T^k = 0 .$$

Hence the choice of gauge (16.3) means we are choosing A_k to be transverse and are eliminating the longitudinal part . We are not, however, free in our choice of A_o . (16.3) implies :

$$\partial_k (\partial^\circ A^k - \partial^K A^\circ) = j^\circ$$

or, since $\partial^\circ (\partial_k A^k) = 0$,

$$\partial_k \partial^K A^\circ = - j^\circ.$$

The solution of this Poisson equation for A° is :

$$A^\circ(x) = \int d\sigma_o \frac{j^\circ(x')}{4\pi(x-x')}$$

Hence the choice of gauge (16.3) determines A^0 uniquely .

In the same way, we can separate F^{0k} into a transverse and a longitudinal part :

$$F^{0k} = F_L^{0k} + F_T^{0k}$$

where again

$$F_L^{0k} = \partial^k f$$

where f is a scalar function . Then in the generator G, an integration by parts brings in the factor $\partial_k \delta A^k = 0$ to eliminate the contribution of F_L^{0k} to G . This is consistent with :

$$\partial_k F^{0k} = \partial_k F_L^{0k} + \partial_k F_T^{0k} = j^0$$

where $\partial_k F_T^{0k}$ vanishes . F_L^{0k} is not free to be varied arbitrarily .

The kinematically independent field components are then :

$$\psi, \bar{\psi} \quad ; \quad A_T^k \ , \quad F_T^{0k} \quad .$$

The Maxwell components must be varied subject to the condition that they remain transverse :

$$\partial_k A_T^k = \partial_k F_T^{0k} = 0 \ .$$

The commutation relations will be those we determined before :

$$\left[\psi(x) \ , \ \bar{\psi}(x') \gamma^0 \right] = \delta^0 (x - x') \qquad (15.5)$$

$$\frac{1}{i} \left[A_k^T (x), \ - F_T^{0\ell}(x') \right] = \left[\delta_k^{\ell} \ \delta^0 (x - x') \right]^T \qquad (13.6)$$

where we write the transverse part of the δ-function for its divergenceless part .

We now turn to practical methods for calculating the effect of the coupling

of the fields. We treat an uncoupled Dirac field with :

$$\mathcal{L} = \tfrac{1}{2}\left[\bar{\psi} \cdot i\gamma^{\mu}\partial_{\mu}\psi - i\partial_{\mu}\bar{\psi}\cdot\gamma^{\mu}\psi\right] - m\bar{\psi}\cdot\psi$$

and examine the effect of including sources in the variation, i.e.the problem with

$$\mathcal{L}' = \mathcal{L} + \bar{\psi}\eta + \bar{\eta}\psi$$

The η are given quantities with the properties

$$\{\eta,\bar{\eta}\} = \{\eta,\bar{\psi}\} = \text{etc.} = 0.$$

The field equation becomes :

$$\left(\gamma^{\mu}\frac{1}{i}\partial^{\mu} + m\right)\psi = \eta$$

If the source η is localized in time, it gives an abrupt change of the field . We choose :

$$\eta = \frac{1}{i}\gamma^{0}\psi(x)\delta(x^{0})$$

which ensures that :

$$\psi\Big]_{-o}^{+o} = \psi'$$

Now :

$$\delta_{\eta}\langle\sigma_1|\sigma_2\rangle = i\int_{\sigma_2}^{\sigma_1}(dx)\langle\sigma_1|\bar{\psi}\,d\eta + d\bar{\eta}\,\psi|\sigma_2\rangle$$

$$\frac{\delta_{\eta}\langle\sigma_1|\sigma_2\rangle}{\langle\sigma_1|\sigma_2\rangle} = i\int_{\sigma_2}^{\sigma_1}(dx)\left[\frac{\langle\sigma_1|\bar{\psi}(x)|\sigma_2\rangle}{\langle\sigma_1|\sigma_2\rangle}\delta\eta + \delta\bar{\eta}\,\frac{\langle\sigma_1|\psi(x)|\sigma_2\rangle}{\langle\sigma_1|\sigma_2\rangle}\right] \quad (16.4)$$

Look at the last term by itself : $\dfrac{\langle\sigma_1|\psi|\sigma_2\rangle}{\langle\sigma_1|\sigma_2\rangle}$ is a number which is a solution of the equation

$$\left(\gamma^{\mu}\frac{1}{i}\delta_{\mu} + m\right)\frac{\langle\sigma_1|\psi|\sigma_2\rangle}{\langle\sigma_1|\sigma_2\rangle} = \eta$$

We now choose the representation so that $\langle \sigma_1 | \psi | \sigma_2 \rangle = 0$ when $\eta = 0$. The simplest way would be to have $\psi' = 0$ on σ_2 . The solution of the equation in terms of the source is then :

$$\frac{\langle \sigma_1 | \psi(x) | \sigma_2 \rangle}{\langle \sigma_1 | \sigma_2 \rangle} = \int_{\sigma_2}^{\sigma_1} (dx') \, G(x,x') \eta(x')$$

where G satisfies :

$$\left(\gamma^\mu \frac{1}{i} \partial_\mu + m \right) G(x,x') = \delta(x - x')$$

The form of G depends on what the boundary conditions are, i.e. $\psi' = 0$ on σ_1 or σ_2 , or something else . We shall examine this more in detail immediately .

The solution of (16.4) is then

$$\langle \sigma_1 | \sigma_2 \rangle_\eta = e^{i \int_{\sigma_2}^{\sigma_1} (dx)(dx') \bar{\eta}(x) G(x,x') \eta(x')} \langle \sigma_1 | \sigma_2 \rangle_{\eta=0} \qquad (16.5)$$

We now look at possible boundary conditions to determine which G we would use . Let the transformation function be

$$\langle \bar{\psi}' \sigma_1 | \psi' \sigma_2 \rangle$$

With $\psi' = 0$ on σ_2 in

$$\frac{\langle \bar{\psi}' \sigma_1 | \psi | \psi' \sigma_2 \rangle}{\langle \bar{\psi}' \sigma_1 | \psi' \sigma_2 \rangle} = \int_{\sigma_2}^{\sigma_1} (dx) \, G(x,x') \eta(x')$$

we choose $G_{ret.}$, since $G_{ret.}$ is defined so that :

$$G_{ret} \begin{cases} = \frac{1}{i} \langle \bar{\psi}\sigma_1 \,|\, \psi'\sigma_2 \rangle & \text{if } x^0 > x^{0'} \\ = 0 & \text{if } x^0 < x^{0'} . \end{cases}$$

On the other hand, if $\bar{\psi}' = 0$ on σ_1, we choose $G_{adv.} = 0$ if $x^{0'} < x^0$. In the relativistic world, however, neither of these representations is the most convenient .

$\psi(x)$ satisfied the Dirac equation and varies in time with various frequencies . A solution which contains a term $\sim \exp(-ip^0 x_0)$ must also contain a term $\sim \exp(ip^0 x_0)$, since the system is invariant under the reversal of the sense of time . (There is no counterpart in non-relativistic theory since there p^0 = energy, always > 0.) We divide $\psi(x)$ into a positive frequency part and a negative frequency part :

$$\psi(x) = \psi^{(+)}(x) + \psi^{(-)}(x)$$

The important step will now be to interpret these two operators as particle annihilation and creation operators . The particles of the system carry always positive energy and the lowest state of the system will be the vacuum of the system . (From the commutator with H :)

$$\left[\chi^{(+)}, p^0\right] = i\partial_0 \chi^{(+)}$$
$$= p^0 \chi^{(+)} \qquad \text{(from a Fourier term.)}$$

(Apparently from the whole $\chi^+ = \int_0^\infty e^{-ip^0 x_0} \chi^+(p^0)\, d p^0$

we get

$$i\,\partial_0 \chi^{(+)} = \bar{p}^0 \int_0^\infty e^{-ip^0 x_0}\, \chi(p^0)\, dp^0$$

where $\bar{p}^0$ is an average value of p^0 .)

$$p^0 \chi^{(+)} = \chi^{(+)}(P^0 - p^0) \quad , \qquad p^0 > 0 .$$

(This is of the same form as the annihilation operator identity for $\psi(a')$:

$$N\,\psi(a') = \psi(a')\,(N - 1)$$

concerned directly with the number of particles and not their energy .)

Hence after the operator $\chi^{(+)}$ has acted on a state, the energy has been decreased by the amount p^0 :

$$\chi^{(+)}_{p^0}\, \Psi(P^{0'}) = \Psi\,(P^{0'} - p^0)$$

The corresponding equations for $\chi^{(-)}$ show that it creates energy p^0 in a system .

The vacuum will be defined as the 0-eigenstate of the annihilation operator :

$$\chi^{(+)}\, \Psi(0) = 0 .$$

It is natural to introduce another representation which distinguishes the $\psi^{(+)}$ and $\psi^{(-)}$. We express both ψ and $\bar{\psi}$ as the sum of positive and negative frequency parts . In the generatorthis becomes :

$$G = \int d\sigma_0\, \bar{\psi}\, i\gamma^0 \delta\psi = \int d\sigma_0 \left(\bar{\psi}^{(+)} + \bar{\psi}^{(-)}\right) i\gamma^0 \left(\delta\psi^{(+)} + \delta\psi^{(-)}\right)$$

Using the fact that $\bar{\psi}^{(+)} = \overline{\psi^{(-)}}$ and the orthogonality of the positive and negative frequency parts of two operators , we get :

$$G = \int d\sigma_0 \left(\bar{\psi}^{(+)} i\gamma^0 \delta\psi^{(-)} + \bar{\psi}^{(-)} i\gamma^0 \delta\psi^{(+)} \right) \qquad (16.6)$$

(Let $f(x)^{(+)} = \int_0^\infty e^{-ixt} f(t)\, dt$, $\quad g(x)^{(-)} = \int_{-\infty}^0 e^{-ixt} g(t) dt$. Then :

$$\int \overline{f(x)^{(+)}}\, g(x)^{(-)} dx = \int_0^\infty \overline{f(t)} dt \int_{-\infty}^0 g(t') dt' \int_{-\infty}^\infty e^{-ix(t'-t)} dx = 0 \,.$$

since the last integral is a representation of $\delta(t'-t)$ and there is no common region of integration of t, t') .

We add a total variation to (16.6) to get :

$$G = \int d\sigma_0 \left(\bar{\psi}^{(-)} i\gamma^0 \delta\psi^{(+)} - \delta\bar{\psi}^{(+)} i\gamma^0 \psi^{(-)} \right)$$

In this canonical form, the field is made up of ψ and $\bar{\psi}$ together . The q-operators are the positive frequency part, the p-operators the negative frequency part. We use the representation :

$$\langle \chi^{(-)'}\sigma_1 \,|\, \chi^{(+)'}\sigma_2 \rangle$$

together with the interpretation that $|\chi^{(+)'}\sigma_2'\rangle$ for zero eigenvalues is the vacuum state .

We now return to the choice of the Green's function of (16.5) in this representation . The boundary conditions

$$\chi^{(-)'} = 0 \quad \text{on} \quad \sigma_1$$

$$\chi^{(+)'} = 0 \quad \text{on} \quad \sigma_2$$

imply that the field may contain only negative frequency components prior to t_2 , only positive frequency components after t_1 . Translated into boundary conditions for $G(x,x')$ this becomes :

on σ_2 : $G_+(x,x')$, $x^0 < x^{0'}$ contains only negative frequency components in its dependence on x .

on σ_1 : $G_+(x,x')$, $x^0 > x^{0'}$, contains only positive frequency components in its dependence on x .

G_+ represents a field radiating from a point source with positive frequencies forward in time, with negative frequencies backward in time . Hence, we write :

$$\langle 0\sigma_1 | 0\sigma_2 \rangle_\eta = e^{i \int_{\sigma_2}^{\sigma_1} (dx)(dx') \bar{\eta}(x) G_+(x,x') \eta(x)}$$

Without the source , $\langle 0\sigma_1 | 0\sigma_2 \rangle_0 = 1 \cdot 1 = 1$ a vacuum to vacuum transformation function without change .

This representation is found to be the most convenient for the Dirac field .

We now turn to the Maxwell field and consider the field coupled to an external source :

$$\mathcal{L} = \tfrac{1}{2} \left[F^{\mu\nu} \cdot \tfrac{1}{2} (\partial_\mu A_\nu - \partial_\nu A_\mu) - A_\nu \cdot \partial_\mu F^{\mu\nu} + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + A^\mu J_\mu(x) \right.$$

From the field equations we have :

$$\partial_\nu F^{\mu\nu} = J^\nu \qquad (16.7)$$

The kinematically independent components are F_T^{ok} , A_T^{ℓ} . As in the case of the Dirac field we divide the operators into positive and negative frequency parts which are Hermitian conjugates , one of the other :

$$F_T^{ok} = F_T^{ok\,(+)} + F_T^{ok(-)} \quad .$$

$F_T^{ok(+)}$ is a photon annihilator, $F_T^{ok(-)}$ a photon creator . We choose the representation :

$$\langle F_T^{(-)'}\sigma_1 \mid F_T^{(+)'}\sigma_2 \rangle$$

with $F_T^{(-)'} = F_T^{(+)'} = 0$ for the vacuum-vacuum transformation function . Variation of the source gives :

$$\delta_J \langle 0\sigma_1 | 0\sigma_2 \rangle = i \int (dx) \langle 0\sigma_1 | \delta J_\mu A^\mu | 0\sigma_2 \rangle$$

δJ^μ is constrained by $\partial_\mu \delta J^\mu = 0$. This,however,(by methods similar to those used in deriving (13.6) leads only to adding a gradient of a scalar to A_μ . Since this is just the condition of gauge invariance, we omit explicit consideration of this constraint .

We may write a numerical equation for $\frac{\langle 0\sigma_1|A^\mu|0\sigma_2\rangle}{\langle 0\sigma_1|0\sigma_2\rangle}$ from (16.7) (up to a gradient term which we may omit) :

$$-\Box \frac{\langle 0\sigma_1|A^\mu|0\sigma_2\rangle}{\langle 0\sigma_1|0\sigma_2\rangle} = J^\mu$$

The solution in terms of a Green's function G , satisfying

$$-\Box G(x,x') = \delta(x - x')$$

with the boundary conditions discussed above is

$$\frac{\langle 0\sigma_1|A^\mu|0\sigma_2\rangle}{\langle 0\sigma_1|0\sigma_2\rangle} = \int (dx')\ G_+(x,x')\ J^\mu(x')$$

Hence

$$\frac{\delta\langle 0\sigma_1|0\sigma_2\rangle}{\langle 0\sigma_1|0\sigma_2\rangle} = i \int (dx)(dx')\ \delta J_\mu(x)\ G_+(x,x')\ J^\mu(x')$$

and

$$\langle 0\sigma_1|0\sigma_2\rangle = e^{\frac{1}{2} i \int (dx)(dx')\ J^\mu(x)\ G_+(x,x')\ J_\mu(x')}$$

The transformation function for the photon and electron fields not interacting is the product of the transformation functions of each field . To calculate the transformation function for the interacting fields, we introduce the interaction term $\lambda\, j^{\mu} A_{\mu}$ into the Lagrange function and differentiate with respect to λ :

$$\frac{\partial}{\partial \lambda} \langle o\sigma_1 | o\sigma_2 \rangle = i \int (dx) \langle o\sigma_1 | j^{\mu} A_{\mu} | o\sigma_2 \rangle$$

where $j^{\mu} = e\, \bar{\psi} \cdot \gamma^{\mu} \psi$. To write j^{μ} more conveniently for what follows we remark that

$$\bar{\psi} \cdot \gamma^{\mu} \psi = \bar{\psi}_{\alpha} \cdot \gamma^{\mu}_{\alpha\beta} \psi_{\beta} = - \gamma^{\mu}_{\alpha\beta} \psi_{\alpha} \cdot \bar{\psi}_{\beta}$$

We now consider the $\psi_{\alpha} \cdot \bar{\psi}_{\beta}$ as a matrix and write :

$$j^{\mu} = -e \operatorname{tr} \gamma^{\mu} \psi \cdot \bar{\psi}$$

The Lagrange function for the two fields already contains

$$\bar{\psi} \eta + \bar{\eta} \psi + J_{\mu} A^{\mu}$$

We produce $\frac{\partial}{\partial \lambda} \langle o\sigma_1 | o\sigma_2 \rangle$ by differentiation, i.e. following the process

$$\delta_{\eta} \langle o\sigma_1 | o\sigma_2 \rangle = i \int (dx) \langle o\sigma_1 | \bar{\psi}\, d\eta | o\sigma_2 \rangle$$

or

$$\frac{1}{i} \frac{\delta e}{\delta \eta(x)} \langle o\sigma_1 | o\sigma_2 \rangle = \langle o\sigma_1 | \bar{\psi}(x) | o\sigma_2 \rangle$$

$$\frac{1}{i} \frac{\delta \varepsilon}{\delta \bar{\eta}(x)} \langle o\sigma_1 | o\sigma_2 \rangle = \langle o\sigma_1 | \psi(x) | o\sigma_2 \rangle$$

$$\frac{1}{i} \frac{\delta}{\delta J^{\mu}(x)} \langle o\sigma_1 | o\sigma_2 \rangle = \langle o\sigma_1 | A_{\mu}(x) | o\sigma_2 \rangle$$

To produce $\langle 0\sigma_1 | j^\mu A_\mu | 0\sigma_2 \rangle$ we apply the three operators :

$$\langle 0\sigma_1 | j^\mu A_\mu | 0\sigma_2 \rangle = i\, e\, \mathrm{tr}\, \gamma \frac{\delta_e}{\delta\bar{\eta}(x)} \frac{\delta_\eta}{\delta\eta(x)} \frac{\delta}{\delta J^\mu(x)} \langle 0\sigma_1 | 0\sigma_2 \rangle$$

To get the right order of factors and the antisymmetrization, the $\frac{\delta}{\delta\eta(x)}, \frac{\delta}{\delta\bar{\eta}(x)}$ must be applied in both senses and averaged from future and past. We now have

$$\frac{\partial}{\partial\lambda}\langle 0\sigma_1 | 0\sigma_2 \rangle = -e \int (dx)\, \mathrm{tr}\, \gamma \frac{\delta_e}{\delta\bar{\eta}} \frac{\delta_\eta}{\delta\eta} \frac{\delta}{\delta J^\mu} \langle 0\sigma_1 | 0\sigma_2 \rangle$$

Then

$$\langle 0\sigma_1 | 0\sigma_2 \rangle_{int.} = e^{-e\int (dx)\, \mathrm{tr}\, \gamma \frac{\delta_e}{\delta\bar{\eta}} \frac{\delta_\eta}{\delta\eta} \frac{\delta}{\delta J^\mu}} e^{i\int (dx)\bar{\eta} G_+^0 \eta} e^{\frac{i}{2}\int J G_+^0 J} \qquad (16.8)$$

One method of calculating with this solution is to expand as a series in e as a perturbation approach. However this is not possible for bound states; one needs more exact solutions.

Another way of treating the coupling of two fields is the following. One begins with the uncoupled fields, but with a given external vector potential A_μ^e :

$$\mathcal{L} = \frac{1}{2}\Big[F^{\mu\nu} \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu) - A_\nu \cdot \partial_\mu F^{\mu\nu} \Big] + \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$
$$- \frac{1}{2}\Big[\bar{\psi}\cdot\gamma^\mu \big(\frac{1}{i}\partial_\mu - e A_\mu^e\big)\psi - \big(\frac{1}{i}\partial_\mu + A_\mu^e\big)\psi\cdot\gamma^\mu\psi \Big] - m\bar{\psi}\cdot\psi$$
$$+ \bar{\psi}\eta + \eta\psi$$

Following the process used before for the Dirac field with sources we obtain :

$$\langle 0\sigma_1 | 0\sigma_2 \rangle_\eta = e^{i\int (dx)(dx')\bar{\eta}(x)\, G_+(x,x')\, \eta(x')} \langle 0\sigma_1 | 0\sigma_2 \rangle_{\eta=0}$$

where now $G_+(x, x')$ satisfies

$$\left[\gamma^\mu \left(\frac{1}{i} \partial_\mu - e A^e_\mu \right) + m \right] G = \delta(x - x').$$

$\langle 0\sigma_1 | 0\sigma_2 \rangle_{\eta=0}$ is the vacuum-vacuum transformation function with A^e_μ acting in the time interval between σ_2 and σ_1 (which we will calculate below).

We now use this to calculate the effect of the coupling by the following device. We introduce into the Lagrange function a term

$$\mathcal{L}_{\text{interact.}} = \lambda\, j_\mu (A^\mu + A^{\mu e})$$

where A^μ is the A^μ of the Maxwell field, $A^{\mu e}$ an external, superimposed potential. By varying only the external field we still get the effect of the total current. Following the process we used above to introduce terms by differentiation we get the effect of the coupling as

$$\langle 0\sigma_1 | 0\sigma_2 \rangle_{\text{full coupling}} = e^{-i\int (dx) \frac{\delta}{\delta J_\mu} \frac{\delta}{\delta A^e}} \langle 0\sigma_1 | 0\sigma_2 \rangle_{\text{no coupling}}$$

This form is more useful than (16.8).

We still need to calculate $\langle 0\sigma_1 | 0\sigma_2 \rangle$ for the Dirac field in order to have the complete transformation function before the coupling.

We calculate for the Dirac field

$$\delta_{A^e_\mu} \langle o\sigma_1 | \; o\sigma_2 \rangle = i \int (dx) \langle o\sigma_1 | \, j^\mu \, | o\sigma_2 \rangle \delta A^e_\mu$$

where

$$j^\mu = - e \; tr \; \gamma^\mu \psi . \bar{\psi}$$

We express $\frac{\langle o\sigma_1 | \psi . \bar{\psi} | o\sigma_2 \rangle}{\langle o\sigma_1 | o\sigma_2 \rangle}$ in terms of $G_+ (x, x')$ as follows :

$$\frac{\delta_r}{\delta \eta} \langle o\sigma_1 | \; o\sigma_2 \rangle = i \langle o\sigma_1 | \bar{\psi}(x) | \; o\sigma_2 \rangle$$

$$\frac{\delta_e}{\delta \bar{\eta}} \langle o\sigma_1 | \; o\sigma_2 \rangle = i \langle o\sigma_1 | \psi(x) | \; o\sigma_2 \rangle$$

But

$$\langle o\sigma_1 | \; o\sigma_2 \rangle = e^{\, i \int_{\sigma_2}^{\sigma_1} (dx)(dx') \, \bar{\eta}(x) \, G_+ (x, x') \, \eta(x')}$$

By differentiating the right hand side (with respect to η , $\bar{\eta}$ and taking the limit as $x \longrightarrow x'$) we get

$$\frac{\delta \langle o\sigma_1 | \; o\sigma_2 \rangle}{\langle o\sigma_1 | \; o\sigma_2 \rangle} = - e \int (dx) \; tr \; \gamma^\mu \; G_+ (x, x') \; \delta A^e_\mu (x) \qquad (16.9)$$

In order to integrate we must first express $G_+ (x, x')$ as a function of A^e_μ .

The Green's function satisfies

$$\left[\gamma^\mu (p_\mu - e A^e_\mu) + m \right] G = 1 \qquad (16.10)$$

where we have written $\frac{1}{i} \partial_\mu = p_\mu$ and $\delta (x - x') = 1$ (the unit matrix). To solve this formally we introduce the inverse operator to $\gamma^\mu p_\mu + m$, i.e. G° such that

$$(\gamma^\mu p_\mu + m) \, G^\circ = 1 .$$

(16.10) becomes :

$$\left[(G^{\circ})^{-1} - e \gamma^{\mu} A^{e}_{\mu} \right] G = 1$$

$$(1 - G^{\circ} e \gamma^{\mu} A^{e}_{\mu}) G = G^{\circ}$$

$$G = (1 - G^{\circ} e \gamma^{\mu} A^{e}_{\mu})^{-1} G^{\circ} .$$

(The expansion of $(1 - G^{\circ} e \gamma^{\mu} A^{e}_{\mu})^{-1}$ at this point gives a Born solution expansion).

We can now write (16.9) as

$$\frac{\delta \langle 0 \sigma_1 | 0 \sigma_2 \rangle}{\langle 0 \sigma_1 | 0 \sigma_2 \rangle} = - e \, \mathrm{tr} \frac{1}{1 - G^{\circ} e \gamma A} G^{\circ} \gamma \delta A$$

where a vector and operator product notation is used to suppress indice and integration signs.

$\overline{\mathrm{Tr}} \frac{1}{X} \delta X$ can be shown to be $\delta \log \det X$ from the law of differentiation of determinants. Hence

$$\delta \log \langle 0 \sigma_1 | 0 \sigma_2 \rangle = \delta \log \det (1 - G^{\circ} e \gamma A^{e})$$

so that the transformation function for the electron field with the source A^{e} is

$$\langle 0 \sigma_1 | 0 \sigma_2 | = \det (1 - G^{\circ} e \gamma A^{e})$$

(The constant of integration is unity since for $A^{e} = 0$, $\langle 0 \sigma_1 | 0 \sigma_2 \rangle = 1$).

We may now write

$$\langle 0\sigma_1 \mid 0\sigma_2\rangle_{\text{full coupling}} = e^{i\frac{\delta}{\delta J}\frac{\delta}{\delta A^e}} \det(1-Ge\gamma A^e)\; e^{i\bar{\eta} G_+^{A^e}\eta}\; e^{\frac{i}{2}JG_+^{o}J}$$

where indices and integration are suppressed. ($G_+^{A^e}$ is a reminder that G_+ is calculated with A^e present).

We may carry the calculation formally one step furthet. Since the operator $e^{\Delta x\frac{\partial}{\partial x}}$ applied to $f(x)$ gives

$$e^{\Delta x\frac{\partial}{\partial x}} f(x) = f(x+\Delta x),$$

we may interpret $exp\left[i\frac{\delta}{\delta J}\frac{\delta}{\delta A^e}\right]$ as a displacement operator and set $A^e = 0$. Then in the equation $A^e \longrightarrow i\frac{\delta}{\delta J}$

$$\langle 0\sigma_1 \mid 0\sigma_2\rangle_{\text{full inter.}} = \det\left(1+G^{o}e\gamma\frac{\delta}{\delta J}\right) e^{i\bar{\eta}\left(1+iG^{o}e\gamma\frac{\delta}{\delta J}\right)^{-1}G^{o}\eta}$$

$$\times\; e^{\frac{i}{2}JG_+^{o}J}$$

Since we differentiate with respect of J, the source of the Maxwell field is not to be given a determined value until after the differentiations are carried out. J may be split into two parts $J = J^{int} + J^{e}$ where J^{int} is the internal current, J^{e} an arbitrary external part to simulate given conditions.

Although we have a formal solution to the problem of the interact fields, it is not known how one can handle the derivative operators without resorting to expansion in series.

N O T E S O N T H E T H E O R Y

O F P H A S E T R A N S I T I O N S

by

George E. UHLENBECK

University of Michigan

These notes were prepared by Dr. G.J.Hooyman, Instituut voor Theoretische Natuurkunde , Utrecht .

Université de Grenoble

Cours professé à

L'ECOLE D'ETE DE PHYSIQUE THEORIQUE

Les Houches (Haute-Savoie) , France .

Juillet 1955 .

- 1 - THE STATISTICAL THEORY OF PHASE TRANSITIONS.

The statistical derivation of phase transitions always involves the evaluation of the partition function (or sum-over-states) Z

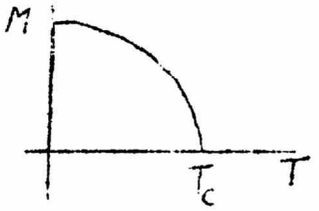
$$Z = \sum_{(i)} G_i e^{-E_i/kT}$$

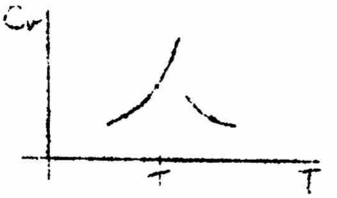
of the system over its possible levels i with energy E_i and multiplicity G_i , from which the thermodynamical functions can be derived, since Z is connected with the free energy Ψ by

$$Z = e^{-\Psi/RT}$$

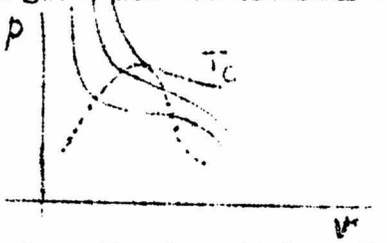
For systems of interacting particles the summation $\sum_{(i)}$ can only be performed in very few cases .

As examples of phase transitions we mention the behaviour of ferromagnetic substances (occurrence of a Curie transition temperature) :





and the condensation of a gas below the critical temperature T_c :



For a finite number N of particles G_i is finite and Z is an analytic function of T . Transitions can only occur in the limiting case $N \longrightarrow \infty$. In the case of the gas for instance one has to consider

N , $v \to \infty$ with finite number density N/v and for the free energy this yields

$$\lim_{N,\, v \to \infty} \psi = N \varphi (T, v),$$

when φ (the free energy per particle) should only depend on the intensive variables T and v .

The exact solution of this limiting problem has only been carried out for the 2-dimensional Ising-model of a ferromagnetic substance and the condensation of a Bose-Einstein gas .

The Ising-model consists of a given lattice on each of the sites of which a spin is situated . The spin parameter S_i can take the values ± 1 . Each spin is assumed to interact only with its nearest neighbors (4 in the 2-dim. case) . The energy of a certain configuration is then

$$E = -\tfrac{1}{2} J \sum_n S_i S_j - \mu H \sum_i S_i ,$$

where $\sum_n$ is taken over all interacting pairs . Note that this is different from the actual ferromagnetic case , where the interaction involves the scalar product ($S_i \cdot S_j$) and where the S_i are q-numbers . The interaction energy J is the increase in energy if two neighbouring spins change from parallel to anti-parallel (in an anti- ferromagnetic lattice $J < 0$). H is the magnetic field and μ the magnetic moment of each atom . With

$$L = J / 2 k T \quad \text{and} \quad C = \mu H / k T$$

the partition function is

$$\sum_{(S_i)} e^{L \sum_n S_i S_j + C \sum_i S_i}$$

In the case of a gas , we suppose the intermolecular forces to be additive and central so that the potential energy can be written as

$$U = \sum_{i<j} \phi(r_{ij})$$

and the (classical) partition function then is

$$Z = e^{-\Psi/kT} = \frac{1}{N!\,h^{3N}} \int_{-\infty}^{\infty} \cdots \int d\vec{p}_1 \ldots d\vec{p}_N \int_V \cdots \int d\vec{r}_1 \ldots d\vec{r}_N \, e^{-\left(\sum_i p_i^2/2m + U\right)/kT}$$

$$= \frac{(2\pi m k T/h^2)^{3N/2}}{N!} \int_V \cdots \int e^{-U/kT} \, d\vec{r}_1 \ldots d\vec{r}_N \qquad (1)$$

(the factor h^{-3N} makes Z dimensionless) .The remaining configurational integral in general cannot be evaluated exactly .

Of course, approximate theories exist for the magnetic problem (Weiss) and the condensation problem (e.g, Van der Waals) . But , apart from their approximate character , such theories presuppose thermodynamics (for instance in the use of the so-called Maxwell rule) which from the point of view of statistics is rather unsatisfactory , since statistical mechanics should provide the basis for thermodynamics .

In the following we will outline the main features of the 2-dimensional Ising-problem and the Bose-Einstein problem . It will turn out that the mathematical mechanism in the existing theories is completely different in these two cases. The question of a unified mathematical method then arises .

- 2. THE ISING-PROBLEM.

We first consider the one domensional case of N lattice points (linear chain) taking the lattice points on a circle, we can identify the points $N+1$ and 1. If we do not take into account the magnetic field, the partitism function is

$$Z = e^{-\Psi/kT} = \sum_{S_1=-1}^{+1} \cdots \sum_{S_N=-1}^{+1} e^{L S_1 S_2} e^{L S_2 S_3} \cdots e^{L S_{N-1} S_N} e^{S_N S_1}$$

Considering the S_i as matrix indices and introducing the matrices $V_{S_i S_j} = e^{L S_i S_j}$ one has $Z = \sum V_{S_1 S_2} V_{S_2 S_3} \cdots V_{S_N S_1} = Trace\,(V^N)$,

where $$V = \begin{pmatrix} e^{L} & e^{-L} \\ e^{-L} & e^{L} \end{pmatrix}$$

Let λ_1 and λ_2 be the eigenvalues of V. On diagonalizing V, one can write

$$Z = \lambda_1^N + \lambda_2^N$$

The secular equation is

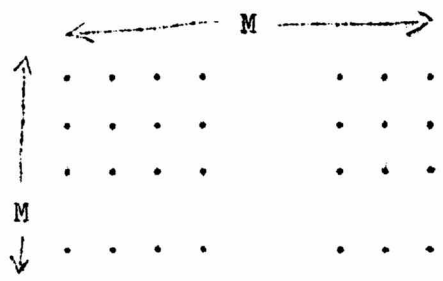
$$\begin{vmatrix} e^{L}-\lambda & e^{-L} \\ e^{-L} & e^{L}-\lambda \end{vmatrix} = 0 \longrightarrow \begin{cases} \lambda_1 = 2\ \cosh\ L \\ \lambda_2 = 2\ \sinh\ L \end{cases}$$

Since N is very large , only the largest eigenvalue λ_1 is important and therefore

$$\Psi = -kTN \log \lambda_1(T) \qquad (\sim N)$$

from which the entropy , the energy etc .. follow . λ and therefore Ψ are analytic functions of T , so here the Ising-model does not lead to a phase transition (in the 1-dimensional case a Curie point never occurs) .

The two-dimensional Ising-model has been treated along the same lines by Onsager . With $N = M^2$, we consider each column of Matoms as a unit, interacting with neighbouring colums . To get rid of boundary conditions we now identify the (M +1) - th column with the first one and the (M + 1) -th row with the first row by winding the lattice on a torus . Each column interacts with neighbouring ones ,each unit has 2^M states , which can be denoted by a matrix index $(s_{1i}, s_{2i}, \cdots s_{Mi})$.



Again one can use the matrix method and

$$Z = \text{Trace } (V^M) = \sum_k^{2^M} \lambda_k^M$$

where now V is a $2^M \times 2^M$ - matrix , but the determination of the eigenvalues now is a major problem . Onsager developed a method (simplified by Onsager and Kaufman) to determine the largest eigenvalue . For $N \to \infty$ the result is again $\Psi(T) = N\psi(T)$ with

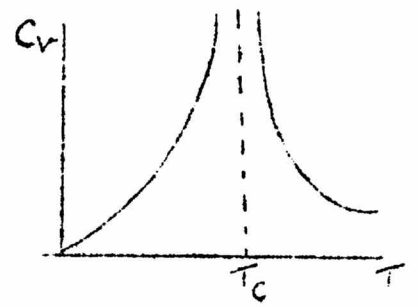
$$-\psi/kT = \frac{1}{N}\log Z = \log(2\cosh 2L) + \frac{1}{\pi}\int_0^{\pi/2} \log\tfrac{1}{2}\left\{1+\sqrt{1-K^2\sin^2\varphi}\right\} d\varphi,$$

where

$$K = 2\sinh 2L / \cosh^2 2L .$$

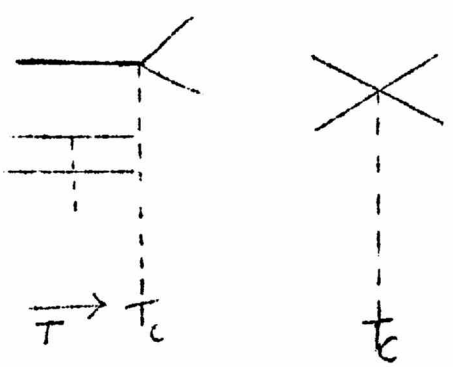
Now a transition point occurs since $\psi(T)$ has a singular point . The critical temperature T_c is determined by

$$\sinh 2L_c = 1 \qquad (L_c = 0.4407\ldots)$$



At the transition point T_c the specific heat C_V becomes infinite .

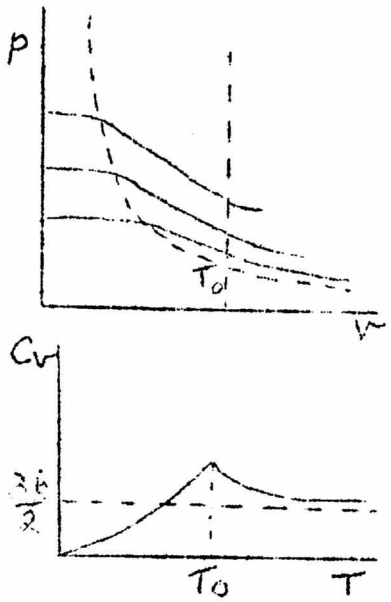
In Onsager's treatment the largest eigenvalue turned out to be 2-fold degenerate up to the temperature T_c. One might expect that in other cases, the discontinuity arises from a crossing of eigenvalues at T_c in such a way that , in taking the largest eigenvalues one has to jump over from one to the other at T_c .



The 3-dimensional Ising-problem is still unsolved , it is even unknown if C_V remains finite in the transition region .

- 3 - THE BOSE -EINSTEIN CONDENSATION .

For a system of N identical non-interacting particles, obeying Bose-Einstein statistics and enclosed in a volume V , the statistical treatment leads for $N \longrightarrow \infty$ with constant density to a condensation



phenomenon ; with decreasing specific volume v at constant temperature the pressure turns out to be constant below a critical volume v_c . The (p , v)- curve has a discontinuity in the second derivative at v_c . The condensation occurs for every isotherm , the locus of the transition points being $p \sim v^{-5/3}$. With decreasing temperature at constant v a transition occurs at $T_0(v)$, where the specific heat C_v shows a discontinuity in the first derivative .

The treatment is slightly different from the discussion of the ordinary gas .(For the treatment, starting from an integral like 1 see Kahn and Uhlenbeck , Physica 5 (1938) 399) If ε_i are the translatory energy levels of a particle in a volume V and n_i the occupation numbers of the levels , the energy E is $\sum_i n_i \varepsilon_i$ and the partition function is

$$Z = \sum{}' e^{-(n_1\varepsilon_1 + n_2\varepsilon_2 + \cdots)/kT}$$

where the prime in $\sum'$ means summation over all occupations ,satisfying $\sum_i n_i = N$. Then Z is the coefficient of z^N in the generating function ($\sum$ without restriction and $\beta = 1/kT$) :

$$F(z) = \sum e^{-\beta(n_1\varepsilon_1 + n_2\varepsilon_2 + \cdots)} z^{n_1 + n_2 + \cdots}$$

$$= \prod_i \sum_{n_i=0}^{\infty} (z e^{-\beta\varepsilon_i})^{n_i} = \prod_i (1 - z e^{-\beta\varepsilon_i})^{-1} = e^{-\sum_i \log(1 - z e^{-\beta\varepsilon_i})}$$

With $N, V \to \infty$ at constant $v = V/N$, the energy spectum becomes continuous as one can therefore replace the sum by an integral :

$$\sum_i \to \frac{V}{h^3}\int d\vec{p}\, \log\left(1 - z e^{-\beta p^2/2m}\right) = -\frac{V}{h^3}\sum_{k=0}^{\infty}\frac{z^k}{k}\int d\vec{p}\, e^{-k\beta p^2/2m}$$

$$= -\frac{4\pi V}{h^3}\sum_k \frac{z^k}{k}\int_0^\infty dp\, p^2 e^{-k\beta p^2/2m} = V\left(\frac{2\pi m k T}{h^2}\right)^{3/2}\chi(z)$$

where $$\chi(z) = \sum_{k=1}^{\infty}\frac{z^k}{k^{5/2}}$$

Since Z is the coefficient of z^N in $F(z)$, we have, introducing the de Broglie-wavelength

$$\lambda = h/\sqrt{2\pi m k T}$$

and taking z to be a complex variable,

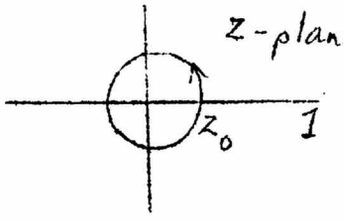
$$Z = \frac{1}{2\pi i}\oint \frac{dz}{z^{N+1}} e^{V\chi(z)/\lambda^3} \simeq \frac{1}{2\pi i}\oint dz\, e^{N\{v\chi(z)/\lambda^3 - \log z\}}$$

according to Cauchy's theorem . The closed contour of integration should enclose $z = 0$.

The integral is known as <u>Kramer's integral</u> (Kramers, Leiden comm ., suppl. № 83 , 1936) and can be evaluated by the method of steepest descent. We rewrite ($a = v/\lambda^3$)

$$Z = \frac{1}{2\pi i}\oint e^{N g(z,a)} \quad \text{with} \quad g(z,a) = a\chi(z) - \log z$$

and consider first the case :



<u>I. a >> 1</u> . $\chi(z)$ is increasing and g has a maximum at z_0 given by

so the saddlepoint $z_0 << 1$. Taking the contour through z_0 , we can replace the integral by an integral over $z = z_0 + i\eta$ with η from $-\infty$ to $+\infty$. Here

$$g(z_0, a) = g(z_0, a) - \tfrac{1}{2}\eta^2 g''(z_0, a) + \cdots$$

with $g''(z_0, a) > 0$ and

$$Z \simeq \frac{e^{Ng(z_0,a)}}{2\pi i} i\int_{-\infty}^{\infty} e^{-\frac{1}{2}Ng''\eta^2} d\eta = \frac{1}{2\pi}\sqrt{\frac{2\pi}{Ng''}}\, e^{Ng(z_0,a)}$$

$$-\Psi/kT = Ng(z_0, a) = N\psi(T, v)$$

so for $N \to \infty$

$$\Psi = -kTNg(z_0, a) = N\psi(T, v)$$

where $\psi(T, v)$ is continuous . With $p = -\partial\psi/\partial v$ this leads to the decreasing portion of the (p, v) - curve for large v .

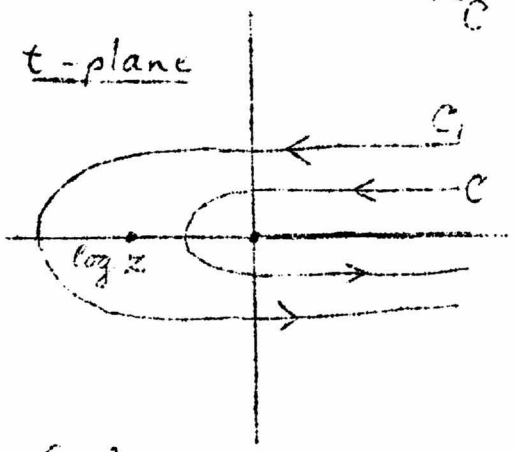
II. Small v . With decreasing v z_0 will go to 1 , which leads to a critical value of v (or a) since $\chi(z)$ singular at $z = 1$. The critical value $a_c = v_c/\lambda^3$ is given by

$$a_c^{-1} = \chi(1) = \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} = \zeta\left(\tfrac{3}{2}\right) = 2.612\ldots = \gamma .$$

We now have to consider the contour integral through z_0 for $z_0 \to 1$.

Since $\chi(z)$ converges within the unit circle it has an integral representation which is (cf. W. Opechowski, Physica (1937) 715) .

$$\chi(z) = \frac{1}{2}\frac{z}{\Gamma(\frac{5}{2})}\int_C \frac{t^{3/2}}{e^t - z}\,dt$$



The integrand is double-valued and has a pole at $t = \log z$, so one has to make a cut along the positive real t-axis and C should not include the point $\log z$ For small z the integrand can be developed

$$\int_C \frac{t^{3/2}}{e^t - z}\,dt = \int_C \frac{t^{3/2}e^{-t}\,dt}{1 - ze^{-t}} = \sum_{k=1}^{\infty} z^{k-1}\int_C t^{3/2}e^{-kt}\,dt$$

and since the last integral is the Hankel integral for $2\,\Gamma(\frac{5}{2})/k^{5/2}$; we again find the original series

$$\chi(z) = \sum_{k=1}^{\infty} z^k/k^{5/2}$$

If $z \to 1$, $\log z \to 0$ and C would enclose the pole . But as long as $z < 1$ we have , replacing the path C by C_1 ,

$$\int_{C_1} = \int_C + (\text{residu in } t = \log z) = \int_C + 2\pi i(\log z)^{3/2}$$

or

$$\chi(z) = \frac{z}{2\Gamma(\frac{5}{2})}\int_{C_1}\frac{t^{3/2}\,dt}{e^t - z} + \frac{4\sqrt{\pi}}{3}z(-\log z)^{3/2}$$

The integral is regular around $Z = 1$, and can be developed in powers of $(1 - Z)$. Thus one gets

$$\chi(z) = \delta - \gamma(1-z) + \cdots + \frac{4\sqrt{\pi}}{3}(1-z)^{\frac{3}{2}} + \cdots$$

where $\delta = \zeta\left(\frac{1}{2}\right)$ and $\gamma = \chi'(1) = \zeta\left(\frac{3}{2}\right)$. In the Z-plane $Z = 1$ is a branchpoint.

With $y = \sqrt{1-z}$ we then find

$$g(z,a) = a\delta + (1 - a\gamma)y^2 + \frac{4a\sqrt{\pi}}{3}y^3 + \cdots$$

Since $a_c \gamma = 1$, we have $1 - a\gamma > 0$ for $a < a_c$, so the integrand e^{Ng} has a maximum along the positive real axis and $y = 0$ is a saddlepoint in the y-plane. We thus can put $y = i\eta$ and have

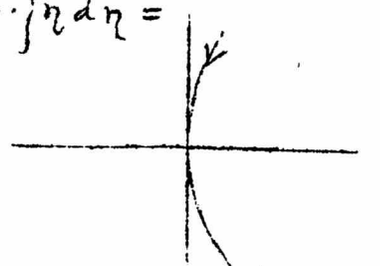
$$Z \cong \frac{2}{2\pi i}\int_{-i\infty}^{+i\infty} y\,dy\; e^{Ng(a,y)}$$

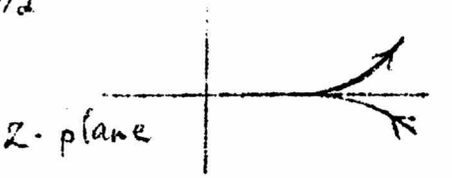
$$= -\frac{e^{N\delta a}}{\pi i}\int_{-\infty}^{+\infty} \eta\,d\eta\; e^{N\left\{(1-a\gamma)y^2 + 4a\sqrt{\pi}y^{3/3} + \cdots\right\}} =$$

$$= -\frac{e^{N\delta a}}{\pi i}\int_{-\infty}^{+\infty} e^{-N(1-a\gamma)\eta^2}\left(1 + \frac{4a\sqrt{\pi}N}{3}i^3\eta^3 + \cdots\right)\eta\,d\eta =$$

$$= -\frac{e^{N\delta a}}{\pi i}\int_{-\infty}^{+\infty} -i\eta^4 d\eta\,\frac{4a\sqrt{\pi}}{3}\,e^{-N(1-a\gamma)\eta^2} =$$

$$= \frac{e^{N\delta a}}{\pi}\,\frac{4aN\sqrt{\pi}}{3}\cdot\frac{3}{4}\,\frac{\sqrt{\pi}}{\{N(1-a\gamma)\}^{5/2}}$$





The saddle point $z_0=1$ is a turning point for the path of steepest descent . It is reached for $v = v_c$. For $v < v_c$ the saddle point sticks to $z_0 = 1$. From the obtained expression for Z in the case $v < v_c$ we find

$$\log Z = N\delta a + \text{terms of order } \log N$$

so for $N \to \infty$

$$\Psi = N\psi \text{ where } \psi = -a\delta kT = -v\delta kT/\lambda^3$$

and

$$p = -\partial\psi/\partial v = \delta kT/\lambda^3,$$

so for $v < v_c$ the <u>pressure is independent of v</u>. At $v = v_c$ the pressure and $\partial p/\partial v$ are continuous , but $\partial^2 p/\partial v^2$ is discontinuous . Since $v_c = \lambda^3/\gamma \sim T^{-3/2}$ the locus of transition points is $\sim T^{-3/2}$. There exists no critical temperature since at any T condensation occurs for sufficiently small v .

From the foregoing it is clear that the mathematical mechanisms of the Ising-problem and the Bose-Einstein condensation are completely different . A unitary mathematical formalism for both cases of phase transitions (the only cases which have been solved in an exact way) might be found in the theory of linear graphs , which we will discuss now .

- 4 - THE THEORY OF LINEAR GRAPHS .

I. <u>Introduction</u> . A linear graph is a collection of points and of lines , joining these points . A graph can be connected or disconnected .

Examples are :

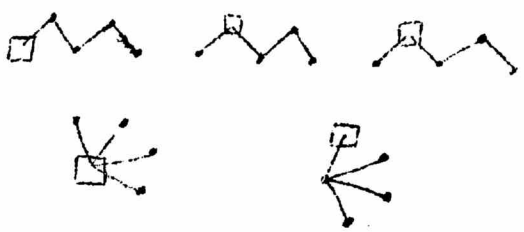
a . Cayley trees : linear graphs without cycles .

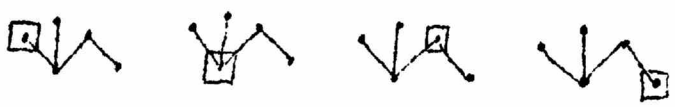


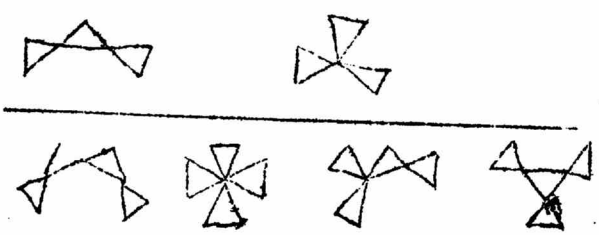
With 5 points there are only 3 topologically different Cayley trees .

For all kinds of graphs, the general problem is to determine the number of topologically different graphs . The answer will be different for distinguishable points and for indistingouishable points .

Cayley investigated the number of isomers $C_n H_{2n+2}$. The carbon chains of these isomers form Cayley trees with the restriction that the maximum number of lines arriving at each point (the " branching number ") is 4 . In the case of isomers of $C_n H_{2n+1} OH$ and similar compounds there is one preferred C-atom , the carbon chain forms a to-called <u>rooted Cayley-tree</u> . For $n = 5$, there are evidently 9 rooted by trees .





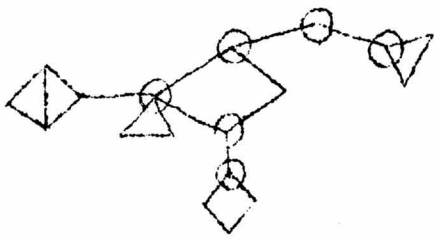


b . Cacti : Cayley trees with triangles as units instead of lines . With 3 or 4 triangles there are resp. 2 and 4 cacti .

c. Husimi trees : Cayley trees arbitrary polygons as units or also : linear graphs in which each line belongs to at most one cycle . In pure Husimi trees the units are equal, in mixed trees they are different .



In a general linear graph one can distinguish articulation points . An articulation point is such that by omitting it , the graph is divided in two or more parts . A connected graph without articulation points we will call a star . Clearly a general connected graph is divided by its articulation points in stars . If one omits in the stars all the internal lines and draws only the " outline ", which is a polygon, one gets a mixed Husimi tree . In this sense, the general connected graph is a generalization of a Husimi tree, just as a star is a generalization of a polygon.



One of the general problems of graph theory is a combinatorial problem . It arises in various fields of physics and is characteristic for successive approximation methods , e.g. in the virial development . An analogous case is the quantum mechanical perturbation problem , where the so-called Feynman graphs appear .

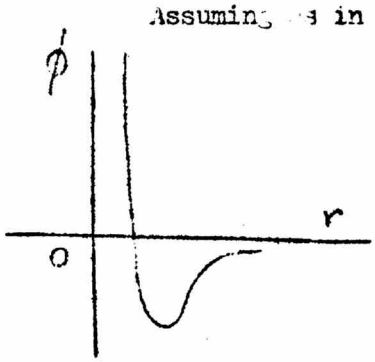
- 5 - APPLICATION TO PHASE TRANSITIONS.

I. The condensation problem . As we saw in parag.1 , the central problem is the evaluation of the partition function and in particular the

the configuration integral

$$Q_N = \int_V \cdots \int e^{-U/kT} \, d\vec{r}_1 \cdots d\vec{r}_N \qquad (1)$$

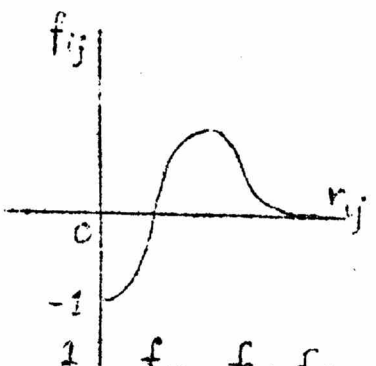
Assuming as in § 1 $U = \sum_{i<j} \phi(r_{ij})$ we can write (Mayer)



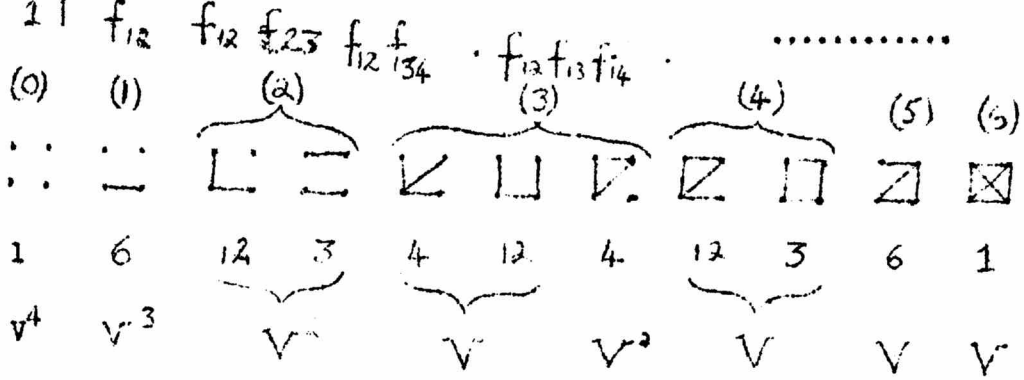
$$e^{-U/kT} = \prod_{i<j} e^{-\phi(r_{ij})/kT} = \prod_{i<j} (1 + f_{ij}) \qquad (2)$$

where

$$f_{ij} = e^{-\phi(r_{ij})/kT} - 1$$



The problem of developing $\prod$ is clearly connected with the theory of graphs , since one can represent all terms of a certain type by a linear graph and then determine the number of these terms. For N = 4 (2^b terms) the different types of terms are represented by the following graphs :



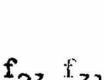
We have indicated the number of lines in each graph (= number of factors f in the terms) and the numbers of terms of each type .

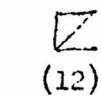
Q_N can now be expanded in powers of V. We consider a certain " partitio " of the N (numbered) molecules in m, single molecules, m_2 pairs, m_3 triples ..., m_l sets of molecules where

$$\sum_{l=1}^{N} l\, m_l = N \qquad (3)$$

Two special types for $N = 4$ are for instance (1,2) (3,4) ($m_2 = 2$, $m_l = 0$ for $m \neq l$) and (1,2,3) (4) ($m_1 = 1$, $m_3 = 1$). In (2) we take together all terms belonging to the considered partitio, that is in which the given pairs, triples, ... each form a connected graph in the graph representation. The cluster function $U_l(\vec{r}_1, \ldots, \vec{r}_l)$ is defined as the sum of all terms represented by connected graphs of l points. For instance:

$$U_2(1,2) = f_{12}$$

$$U_3(1,2,3) = f_{12} f_{23} + f_{13} f_{32} + f_{21} f_{13} + f_{12} f_{23} f_{31} =$$  + 

$$U_4(1,2,3,4) =$$  +  +  +  +  + 

It is clearly a symmetric function of $\vec{r}_1, \ldots, \vec{r}_l$. The cluster integral is defined by

$$b_l = \frac{1}{l!\, V} \int \cdots \int d\vec{r}_1 \ldots d\vec{r}_l \; U_l \; .$$

In the l-fold integral over the connected graph we first can perform the integration over $l-1$ molecules and the result is practically independent of the position of the l-th molecule since each f is only different from zero for small distances of the molecules. The integration over the l-th

molecule then leads to a factor V . Therefore for large V (and fixed ℓ) the b_ℓ will become asymptotically independent of V and are then only functions of the temperature . Define b_ℓ 1 . For a definite partitio we have the general contribution

$$\prod_\ell (V \; b_\ell \; \ell !)^{m_\ell} \qquad (4)$$

Since the partitio with given numbers m can be realized in

$$N \; ! \Big/ \{ (1!)^{m_1} (2!)^{m_2} \; \ldots\ldots \; m_1 \; ! \; m_2 ! \; \ldots\ldots \} \qquad (5)$$

different ways, we find

$$Q_N = \sum{}' \frac{N \; !}{(1!)^{m_1} (2!)^{m_2} \ldots m_1 ! \; m_2 ! \; \ldots} \prod_\ell (V \ell ! b_\ell)^{m_\ell} = \sum{}' \frac{N \; !}{m_1 ! \; m_2 \ldots} \prod_\ell (V b_\ell)^{m_\ell} \qquad (6)$$

where $\sum'$ means summation over all values m_ℓ , obeying (3).

For instance , for N = 4 the partitio 4= 2+2, represented by the single graph  , gives a contribution (V_2 ! b_2)^2 and it occurs 3 times . This follows also from (4) and (5) for $m_2 = 2$, $m_\ell = 0$ for $\ell \neq 2$. The partitio $4 = 3 + 1$ is represented by the graphs and . It gives a contribution (V_3 ! b_3) (V_1 ! b_1) and it occurs four times (since all the permutations of the three connected points are included in U_3) . This follows again from (4) and (5) with N=4, m_1 = 1, m_3 =3 , all other $m_\ell = 0$.

Q_N has now been expanded in powers of V and we are interested in its behaviour for large N . With the assumed intermolecular potential ϕ the function f_{ij} will have a large positive part if the temperature is

not too high, and this will lead to $b_\ell > 0$. We suppose this to be the case, so all terms in (6) are positive . The terms in (6) are then of the same form as the volume in the Γ - space of a gas of N molecules, corresponding to a certain occupation of cells in the μ-space and, like there, for one special distribution (for one set of the m) the term is extremely large . We therefore can replace Q_N by this maximum term , which is the " Maxwell - Boltzmann distribution " for the m_ℓ . To find this term we calculate the maximum of

$$\log. F = \log. \frac{N!}{m_1! m_2! \ldots} (Vb_1)^{m_1} (Vb_2)^{m_2} \ldots =$$

$$= N \log. N - N - \sum_\ell (m_\ell \log m_\ell - m_\ell) + \sum_\ell m_\ell \log Vb_\ell ,$$

(here we have applied the Stirling formula to N ! and the m_ℓ ! which is allowed since the small values of m_ℓ do not play a part) with the auxiliary condition (3) ,

$$\delta \log. F = - \sum_\ell \log m_\ell \delta m_\ell + \sum_\ell \log Vb_\ell \delta m_\ell = 0$$

with

$$\sum_\ell \ell \, \delta m_\ell = 0 .$$

Applying the Lagrange method of undetermined multipliers, we find

$$- \log \bar{m}_\ell + \log Vb_\ell + \beta \ell = 0$$

or

$$\bar{m}_\ell = Vb_\ell e^{\beta \ell} \equiv Vb_\ell z^\ell ,$$

where the parameter z (which depends on the volume) is determined by the condition (3) for the $\bar{m}_\ell$:

$$\boxed{\frac{N}{V} = \frac{1}{v} = \sum_{\ell=1}^{N} \ell\, b_\ell z^\ell} \qquad (7)$$

This is the first Mayer equation . By taking the second variation it can be verified that $F(\bar{m}_\ell)$ is indeed a very sharp maximum . Replacing Q_N by $F(\bar{m}_\ell)$, we find with the use of (7) from the partition function

$$\Psi = -kT \log Z = -\frac{3}{2} R T \log T + \text{const.}\, NT + N k T \log z - k T V \sum_\ell b_\ell$$

hence

$$p = -\left(\frac{\partial \Psi}{\partial V}\right)_T = -\frac{N k T}{z}\frac{\partial z}{\partial V} + k T \sum_\ell b_\ell z^\ell + k T V \sum_\ell \ell b_\ell z^{\ell-1}\frac{\partial z}{\partial V}$$

which in view of (7) yields the second Mayer equation

$$\frac{p}{kT} = \sum_{\ell=1}^{N} b_\ell z^\ell \qquad (8)$$

The equation of state is now obtained by eliminating z from (7) and (8) . Since (7) cannot explicitly be solved for z , this has to be done by successive approximation and one finds p/kT as a series in $1/v$, that is the virial expansion for the equation of state .

For large v , z is small . If we only take the first terms in (7) and (8) , we find the first approximation

$$\frac{p}{kT} = z = \frac{1}{v}$$

i.e., the ideal gas law . The first correction is obtained by inserting this first approximation for z in the quadratic terms of (7) and (8) :

$$(7) \longrightarrow \frac{1}{v} = z + 2 b_2 \left(\frac{1}{v}\right)^2 \quad \text{or} \quad z = \frac{1}{v} - 2 b_2 \left(\frac{1}{v}\right)^2$$

then $(8) \rightarrow \frac{p}{RT} = \frac{1}{v} - 2\,b_2 \left(\frac{1}{v}\right)^2 + b_2 \left(\frac{1}{v}\right)^2 = \frac{1}{v}\left(1 - \frac{b_2}{v}\right)$

so that in the virial expansion

$$p = \frac{RT}{V}\left(1 + \frac{B}{V} + \frac{C}{V^2} + \ldots\right)$$

the second virial coefficient is $B = -N b_2$. From our previous definition of b_2 , we can express B in terms of the intermolecular potential

$$B(T) = -N b_2 = -\frac{N}{2V} \iint \left(e^{-\phi(r_{12})/kT} - 1\right) d\vec{r}_1\, d\vec{r}_2 =$$

$$= -\frac{N}{2V} \int d\vec{r}_2 \int \left(e^{-\phi(\rho)/kT} - 1\right) d\left(\vec{r}_1 - \vec{r}_2\right) .$$

where $\rho = \left| \vec{r}_1 - \vec{r}_2 \right|$, or

$$B(T) = 2\pi N \int_0^\infty \left(1 - e^{\phi(\rho)/kT}\right) \rho^2 \, d\rho .$$

From measurements of B at various temperatures one can infer the parameters , determining the intermolecular potential .

The question now arises if the above treatment leads to a condensation . The answer is determined by the behaviour of the series $\chi(z) = \sum_l b_l z^l$ occuring in the second Mayer equation , the discussion goes along the same lines as that of the Kramers integral and is given in the paper of Kahn and Uhlenbeck and in Kahn's thesis (On the theory of the equation of state , Utrecht, 1938) . The series plays the same part as the series $\sum_l l / l^{5/2}$ in the case of the Bose-Einstein condensation . Kahn was able to prove that condensation only occurs if $\chi(z)$ fulfilled the following conditions (analogous to the properties of the series in the case of Bose- Einstein condensation) :

1) $\chi(z)$ has a singular point z_0 on the positive real axis (according to Hadamard's theorem this will be the case if $b_\ell > 0$ and if a singular point exists) ,

2) $\chi(z_0)$ and $\chi'(z_0)$ finite ,

3) some additional conditions which are satisfied if in a region around z_0

$$\chi(z) = f(z) + (z - z_0)^\alpha g(z)$$

where α is not integer and > 1 ($\alpha = \frac{3}{2}$ in the Bose-Einstein case) and $f(z)$ and $g(z)$ are analytic in this region .

For this case Kahn proved that the isotherm shows a horizontal portion . But the investigation of $\chi(z)$ involves the study of the cluster integrals b_ℓ and this problem is in general still far from solved .

From the foregoing, it is clear that the theory of condensation leads to two different problems :

a . The combinatorial (or topological) proble : how many terms contribute to the cluster integral ?

b . The integral problem : the evaluation of the different " irreducible " integrals .

For the condensation problem it will be of special interest to investigate the situation for large ℓ , since this will determine the convergence properties of the series $\chi(z)$.

The first problem is now completely solved . It is a problem of

the number of topologically different graphs of a certain type . Cayley was the first to deal with such problems in a systematic way , but he did not succeed in finding the complete answer , even for the case of the Cayley trees . Much progress was achieved by a paper of Polya (Acta Math . 68 (1938) 145) and the case of Cayley trees was solved by Otter (Ann. Math. 49 (1948) 583)

Further literature :

R.J. Riddell, Dissertation , Univ. of Michigan (1951) .

R.J. Riddell and G.E. Uhlenbeck , J. Chem. Phys. 21 (1953) 2056 .

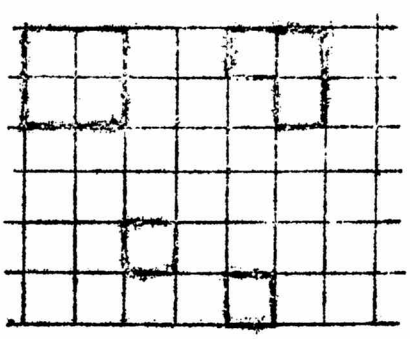
G.W. Ford , Dissertation , Univ. of Michigan (1954.)

- § 6 . OTHER APPLICATIONS OF THE THEORY OF GRAPHS .

As was already mentioned before , the determination of the n-th term in successive approximation methods generally leads to a combinatorial problem (how many contributions ?) and an integral problem . Examples are :

a. The Ising problem . The connection with graph theory was given by van der Waerden (Zs. f. Physik 118 (1941) 473) ; see also the review article by Newell and Montroll (Rev. Mod.Phys. 25 (1953) 353) .

The combinatorial problem is : how many different graphs with given length are possible in a certain lattice ? Consider a square lattice .



The integral problem is very easy : the integral is 1 for every closed graph and zero for every non-closed graph . The combinatorial problem is to find the number of closed graphs (1 of length 4 , 3 of

length 8 , etc) .

b . The connection with the Bose-Einstein condensation , especially for systems of interacting particles . The application of graph theory is presumably possible , though this has not yet been verified (Butler and Friedman , Phys. Rev. 98 (1955) 287,294 ; also Luttinger and Yang).

c . Perturbation problems in field theory . Here the problem is to determine the number of irreducible Feyman diagrams . It has been solved by Hurst (Ploc.Roy.Soc. 214 (1952) 44) and Riddell (Phys. Rev. 91 (1953) 1241) .

§ 7 . THE PRINCIPAL THEOREM .

Passing on to the general method of graph theory, we now give the principal theorem, due to Polya (reference above) .

Think of a collection of " figures", each with a certain " content" which is described by a set of integers. The content may for instance be a number k of red balls and a number ℓ of red of white balls . Let us assume that there are $a_{k\ell}$ different figures of content (k, ℓ) and let the numbers $a_{k\ell}$ be given in the form of a generating function :

$$f(x,y) = \sum_{k,\ell} a_{k\ell}\, x^k y^\ell \qquad (1)$$

(in general a function of n variables if the content is described by n integers). We now consider s points in space, to each of which we will associate 1 figure so that we obtain a certain configuration . Let H_s be some permutation group of s points , of degree s and order h .

Two configurations then will be considered as equivalent if they can be transformed into each other by a permutation belonging to H_s . The content of a configuration is the sum of the contents of the figures .

Problem : Given the function $f(x,y)$, find the number $A_{k\ell}$ of non-equivalent configurations with total content (k, ℓ) , expressed by the generating function

$$F(x,y) = \sum_{k,\ell} A_k \; x^k y^\ell \qquad (2)$$

Solution : Each permutation of H_s can be written uniquely in its cyclic representation (such that each object occurs in one and only one cycle) and then consists of j_1 cycles of 1, j_2 cycles of 2,.. j_s cycles of s , where

$$\sum_{k=1}^{s} k j_k = s, \qquad (3)$$

With the variables $f_1, \ldots, f_s$, we introduce the polynomial

$$\mathcal{Z}(H_s) \equiv \frac{1}{h} \sum_{(j)}{}' g(j_1, j_2 \ldots j_s) f_1^{j_1} f_2^{j_2} \ldots f_s^{j_s} \qquad (4)$$

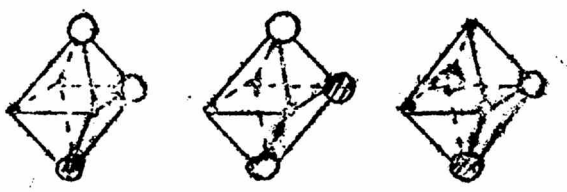
the cycle index of H_s , where $g(j_1,..j_s)$ is the number of permutations in H_s with j_1 cycles of one , ... j_s cycles of s ar where S' should be consistent with (3) .Also $S' g(j_1, j_2, \ldots j_s) = h$.

The solution of the problem is then

$$F(x,y) = \frac{1}{h} S' g(j_1 \ldots j_s) f^{j_1}(x,y) f^{j_2}(x^2, y^2) \ldots f^{j_s}(x^s, y^s) \equiv \mathcal{Z}\{H_s, f(x,y)\} \qquad (5)$$

We will not give the proof (cf. Polya) , but give some examples .

1) Let the s points be the vertices of an <u>octahedron</u> and the content of a configuration be 3 red , 2 blue and 1 white ball. In how many different ways can the 6 balls be distributed over the 6 vertices if two arrangements which are transformed into each other by a rotation are considered the same ? The answer is 3 . The figures are the 3 types of balls, the content is 1 red ball (1,0,0) 1 blue ball (0,1,0) or 1 white ball (0,0,1).



$$f(x,y,z) = x + y + z$$

s = 6 and H_s is the octahedron group of rotations , with h = 24 , for which

$$\mathcal{J}(H_s) = \frac{1}{24}\left(f_1^6 + 6 f_1^2 f_4 + 3 f_1^2 f_2^2 + 6 f_2^3 + 8 f_3^2 \right)$$

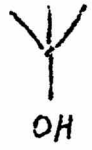
so

$$F(x,y,z) = \frac{1}{24}\left\{ f^6(x,y,z) + 6 f^2(x,y,z)\, f(x^4, y^4, z^4) + \ldots \right\}$$

The answer is then the coefficient of $x^3 y^2 z$, that is 3 .

2) The same problem if no rotation is allowed . This is the elementary problem to distribute the balls over 6 points and the answer is $\frac{6!}{1!\,2!\,3!} = 60$. Again $f(x, y, z) = x + y + z$ and $s = 6$. H_s is now the unit element , $h = 1$. $\mathcal{J}(H_s) = f_1^6$ and $F(x,y,z) = (x + y + z)^6$. The coefficient of $x^3 y^2 z$ is $\frac{6!}{1!\,2!\,3!}$

3) . Number of saturated alcohols $C_n H_{2n+1} OH$ (not including the stereo-isomers) . The counting series id $F(x) = \sum_{n=0}^{\infty} A_n x^n$ with $A_0 \equiv 1$. Starting with the C-atom that carries the - OH , we see that A_n is the number of rooted Cayley trees with n points each of branching number 4.



The figures to be placed on the 3 remaining valencies ($s = 3$) are again rooted carbon chains, including zero chains, so clearly $f(x) = F(x)$. Since we do not pay attention to stereo-isomers , H_s is the complete symmetric permutation group G_3 of 3 objects , which is of the order $h = 3 ! = 6$ and for which

$$\mathcal{Z}(G_3) = \frac{1}{6} (f_1^3 + 3 f_1 f_2 + 2 f_3) .$$

Applying Polya's theorem we have to bear in mind that in $\mathcal{Z}\{G_3 , f(x)\} = \mathcal{Z}\{G_3, F(x)\}$ only the n-1 remaining C-atoms (to be placed at the 3 valencies of $\rightarrow C - OH$) are counted . To count also the root we have to shift all the coefficients in this series to the next pow of x . This is done by multiplying the series with x (then adding 1 to account for the first term $A_0 \equiv 1$) . So Polya's theorem yields

$$F(x) = 1 + x \, \mathcal{Z}\{G_3, F(x)\} = 1 + \frac{x}{6}\{F^3(x) + 3 F(x)F(x^2) + 2 F(x^3) ,$$

from which $F(x)$ can be found by inserting the counting series $\sum_n A_n x^n$ and successively equating the coefficients of equal powers on both sides :

$$F(x) = 1 + x + x^2 + 2x^3 + 4x^4 + 8x^5 + 17x^6 + \ldots .$$

4) Number of alcohols $C_n H_{2n+1} O H$, including stereo isomers.

The derivation is the same as in 3) apart from the fact the valencies are

now placed according to a tetrahedron and that now only rotations over 120° around one valency axis are allowed . H_s is now the cyclic group S_3 of 3 objects , of order 3 . This leads to

$$\mathcal{Z}(S_3) = \frac{1}{3}\,(f_1^3 + 2 f_3)\ ,$$

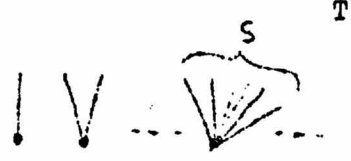
$$F(x) = 1 + \frac{x}{3}\left\{ F^3(x) + 2 F(x^3) \right\}\ ,$$

which yields

$$F(x) = 1 + x + x^2 + 2x^3 + 5x^4 + 11x^5 + \ldots\ .$$

5) The Cayley formula . We now ask for the counting series $T(x)$ for the number T_n of rooted Cayley trees (where the " branching number " of each point is no longer restricted to 4 as was the case for the saturated alcohols) . T_n is the number of these trees with n lines ($n+1$ points) and

$$T(x) = \sum_{n=0}^{\infty} T_n x^n$$



with $T_0 = 1$, $T_1 = 1$, $T_2 = 2$,... The lines starting from the root we call main branches ; let s be the number of main branches . The figures to be placed at the s main branches are again rooted Cayley trees and the figure series is again the counting series for these rooted trees . H_s is the full symmetric group G_s of s objects, of order $s!$ For this group the number $g(j_1, \ldots j_s)$ of eq. (4) is

$$g(j_1, \ldots j_s) = \frac{s!}{1^{j_1} 2^{j_2} 3^{j_3} \ldots j_1!\, j_2!\, j_3! \ldots}$$

The factors $j_1!, j_2!$ ... account for the permutations of the j_1, cycles among each other, of the j_2 cycles among each other etc . and the factors k^{jk} account for the fact that (12 ...k) = (23 ...k 1) = ... (k...321). The cycle index therefore is

$$Z(G_s) = \frac{1}{s!} \sum' \frac{s!}{1^{j_1} j_1! \, 2^{j_2} j_2! \ldots} f_1^{j_1} f_2^{j_2} \ldots \qquad (6)$$

($\sum'$ according to eq. (3)) . Applying Polya's theorem we find :

$$T(x) = 1 + \sum_{s=1}^{\infty} x^s Z\{G_s, T(x)\} \qquad (7)$$

The factor x^s ensures that we also count the s lines of the root (cf. example 3) , the summation is over all possible numbers of main branches . Inserting (6) into (7) we find

$$T(x) = 1 + \sum_{s=1}^{\infty} \sum' \frac{x^{j_1 + 2j_2 + \ldots + sj_s}}{1^{j_1} j_1! \, 2^{j_2} j_2! \ldots} T^{j_1}(x) \, T^{j_2}(x^2) \ldots$$

and since $s \longrightarrow \infty$ the prime can be omitted :

$$T(x) = \prod_{r=1}^{\infty} \sum_{j_r=0}^{\infty} \frac{1}{j_r!} \left\{ \frac{x^r \, T(x^r)}{r} \right\}^{j_r} = \prod_{r=1}^{\infty} e^{x^r \, T(x^r)/r}$$

or

$$T(x) = e^{\sum_{r=1}^{\infty} x^r \, T(x^r)/r} \qquad (8)$$

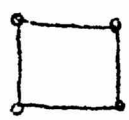
the Cayley equation , which yields after development

$$T(x) = 1 + x + 2x^2 + 4x^3 + 9x^4 + 20x^5 + 48x^6 + \ldots .$$

One can raise the question whether such a series has only a

formal meaning or it has really a region of convergence . We shall later see that the series has a region of convergence and that the singular point which is nearest to the origin in the complex plane lies on the positive real axis (as it should for a series with positive coefficients according to Hadamard's theorem) . We then also return to the connection of the series with the mathematical formation of the condensation phenomenon .

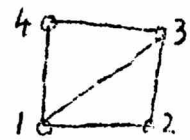
§ 8 . GENERALIZATIONS . We will now use Polyà's theorem to count more complicated graphs . We first introduce the notion of the group of a graph , denoted by Γ: this is the group of covering operations (all operations which transform the graph into itself) . The order of Γ is the symmetry number of the graph . We consider some examples .



If the graph is a rectangle , Γ is the dieder group D_4 of rotations and reflections of a rectangle (order 8) . For a rectangle with 2 diagonals (a complete graph of 4 points) any permutation of the 4 points is allowed and is the symmetric group G_4 of permutations of four objects (order 4 ! = 24) .



For a rectangle with one diagonal , Γ consists of the unit element, the 2" flips " around the lines 1-3 and 2-4 , and their product, so is the four group E,(13),(24), (13)(24) .



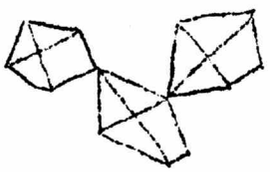
If we want to keep a given point fixed , the number of covering operations will be restricted . They form a subgroup of Γ , which we will call the derived group Γ_q where q denotes the fixed

point . In the first two examples Γ_q is the same whatever the choice of the point q . In the third example Γ_q is different for $q = 1$ (or 3) and $q = 2$ (or 4) . Without proof we mention the theorem

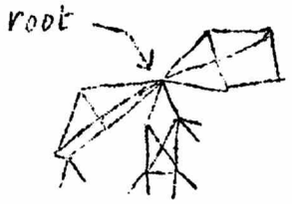
$$Z'(\Gamma) \equiv \frac{\partial}{\partial f_1} Z(\Gamma) = \sum_q Z(\Gamma_q) \qquad (1)$$

where the sum is taken over different derived groups (N.B. In evaluating $Z(\Gamma_q)$ the fixed point q should of course not be included in the cyclic representation of Γ_q . One can easily verify the theorem for the above examples .

In considering now more complicated graphs we first suppose all the " building blocks " to be the same (as the lines in Cayley trees or the triangles in cacti and so consider in general pure star trees (generalisation of Husimi trees) .



a) Pure rooted star trees . Let T_n be the number of different pure rooted star trees of n (equal) stars and let us again introduce the counting series $t(x) = \sum_{n=0} T_n x^n$ with $T_0 \equiv 1$ We will derive an implicit equation for $T(x)$ in two steps, using in each step Polya's theorem .



I . As in ex. 5) of §7 , we call stars starting from the root (e.g. the pentagons with 2 diagonals in the figure) the main branches . If $T_{n\,m}$ is the number of trees with n stars and

m main branches, we can introduce the generating function

$$T_m(x) = \sum_{n=0}^{\infty} T_{n\,m}\, x^n \qquad (2)$$

and clearly have

$$T(x) = \sum_{m=0}^{\infty} T_m(x) \qquad (T_0(x) \equiv 1) \qquad (3)$$

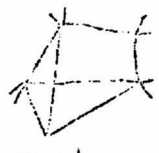
We now can reduce $T_m(x)$ to $T_1(x)$ by considering the root as an m-fold point on each of which 1 rooted tree of 1 main branch is hung. For the configuration the figure series is $T_1(x)$, the number of points $s = m$, the group H_s is the full symmetric group G_m (any permutation of the m rooted subtrees is allowed) and Polya's theorem gives us the configuration series

$$T_m(x) = \mathfrak{Z}\{ G_m, T_1(x) \} \qquad (4)$$

If we now insert (4) in (3), we find (in the same way as in § 7 eq. (8) was derived from eq. (7)

$$T(x) = e^{\sum_{r=1}^{\infty} T_1(x^r)/r} \qquad (5)$$

II. To find $T_1(x)$, we again apply Polya's theorem. The s points for the figures are the vertices of the 1 main branch, except the root (in the example $s = 4$) the figures series is $T(x)$, the group H_s is formed by all the permutations of the points of the main branch which conserve its structure and keep the root fixed, that is, Γ_q of the star. Polya's theorem gives then the configurational series



$$T_1(x) = \sum^{\Gamma} x\, \mathcal{Z}\{\Gamma_q,\ T(x)\}, \qquad (6)$$

where the factor x compensates the fact that the main branch is not counted by $\mathcal{Z}$ and where the sum should be taken over the topologically different possible choices of a point of the main branch as the root .

Applying theorem (1) we can rewrite (6) as

$$T_1(x) = x\, \mathcal{Z}'\{\Gamma, T(x)\} \qquad (7)$$

which combines with (5) to yield

$$T(x) = e^{\sum_{r=1}^{\infty} x^r \mathcal{Z}'\{\Gamma, T(x^r)\}/r} \qquad (8)$$

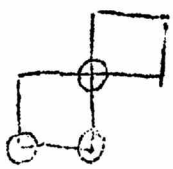
As an example, consider the case where the stars are rectangles . $\Gamma = D_4$ and $\mathcal{Z}(\Gamma) = \frac{1}{8}\{f_1^4 + 2f_1^2 f_2 + 3f_2^2 + 2f_4\}$ so $\mathcal{Z}'(\Gamma) = \frac{1}{2}\{f_1^3 + f_1 f_2\}$

The trees are pure Husimi trees and the counting series $Q(x)$ satisfies

$$Q(x) = e^{\sum_{r=1}^{\infty} x^r \{Q^3(x^r) + Q(x^r)\, Q(x^{2r})\}/2r}$$

from which one finds

$$Q(x) = 1 + x + 3x^2 + 11x^3 + 46x^4 + 208x^5 + \ldots.$$



For $n=2$ there are 3 trees , arising from the 3 different possible positions of the root .

b) Mixed rooted star trees . The stars can now be chosen out of a given finite collection (think of polygons with given numbers of diagonals) As an example we consider rooted Husimi trees of n_2 lines and n_3 triangles

and introduce a counting series of 2 variables

$$H\ (\ x,\ y\ ) = \sum_{n_2, n_3} H\ (n_2\ , n_3)\ \ x^{n_2}\ y^{n_3}, \quad (H(0,0\ ) = 1\ \ )$$

The reasoning is the same as for pure trees : one considers the ccunting series $H_{k\ell}(x,y\ )$ for this type of trees with k lines and ℓ triangles together forming $k + \ell$ main branches . At these $k+\ell$ main branches, we hang the figures . Consider first $\ell = 0$ (only k main lines) The figures series is $H\ (x,y\ )$ and thus are k points , $_3 = G_k$ and we would get $H_{ko}\ (\ x,y\ ) = x^k\ \mathfrak{Z}\left\{G_k\ ,\ H\ (\ x,y\ )\right\}$. Then consider $k = 0$ (only main triangles) . For <u>each</u> main triangle , the figure series would then be the configurational series of the mixed rooted trees hung on the 2 remaining corners of that main triangle, that is

$$\mathfrak{Z}\left\{G_2\ ,\ H\ (\ x,y\ )\right\}\ = \frac{1}{2}\left\{H^2\ \ (\ x,y\ ) + \ H\ (\ x^2\ ,y^2)\right\}.$$

Using this as the figure series for the figures to be hung on the ℓ main triangles, we find

$$H_{o\ell}\ (x,y\ ) = y^{\ell}\ \mathfrak{Z}\left[G_\ell, \mathfrak{Z}\left\{G_2\ ,\ H\ (\ x,y\ )\right\}\right]\ .$$

For arbitrary values of K and ℓ , we have :

$$H_{k\ell}\ (x,y) \ = \ H_{ko}\ (\ x,y\ )\ \ H_{o\ell}\ (x,y)$$

and with

$$H\ (\ x,y) = \sum_{k,\ell} H_{k\ell}\ (\ x,y)$$

we find (along the same lines as was done in § 7, ex.5)

$$H(x,y) = \exp\left[\sum_{r=1}^{\infty} \frac{x^r}{r} H(x^r, y^r) + \sum_{s=1}^{\infty} \frac{y^s}{25}\left\{H^2(x^s,y^s) + H(x^{2s}, y^{2s})\right\}\right] \qquad (9)$$

which leads to

$$H(x,y) = 1 + x + y + 2x^2 + 3xy + 2y^2 + 4x^3 + 10x^2y + 10xy^2 + 5y^3 + \ldots$$

c) Pure free star trees . We denote the counting series as $t(x) = \sum_n t_n x^n$.

1) Gayley trees (star is line) . For this case Otter (Ann. of Math.49 (1948) 583) derived

$$t(x) = T(x) - \frac{1}{2} x\left\{T^2(x) - T(x^2)\right\}$$

T (x) is the corresponding counting series for rooted trees . Note that in this section both t (x) and T (x) are counted according to the number of stars .

2) Cacti (star is triangle) . Harary and Uhlenbeck (Proc. Nat. Ac. U.S. 39 (1953) 315) showed for this case .

$$t(x) = T(x) - \frac{1}{3} T^3(x) - T(x^3)$$

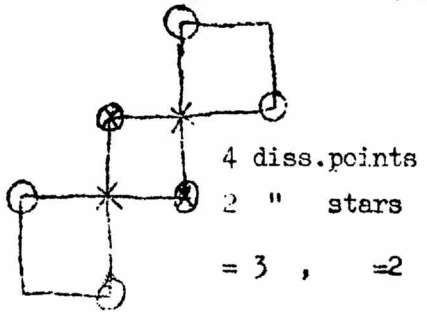
3) Husimi trees with rectangles . Harary and Uhlenbeck derived

$$t(x) = T(x) - x\left\{\frac{3}{8} T^4(x) + \frac{1}{4} T^2(x)\, T(x^2) - \frac{3}{8} T^2(x^2) - \frac{1}{4} T(x^4)\right\}$$

Of course, in each case $t(x) < T(x)$. The general answer for arbitrary stars has been given by Norman (Michigan Univ. diss.1954) :

$$t(x) = T(x) - x\, T(x)\, Z_1\{\Gamma, T(x)\} + x\, Z_2\{\Gamma, T(x)\} \qquad (10)$$

Proof : Consider a definite star tree H of n equal stars . We call two points similar if there exists a covering operation of H which transforms these points into each other . Let $p(H)$ be the number of dissimilar points of H , likewise s(H) the total number of dissimilar stars in H and p_i the number of dissimilar points in the i-th class of dissimilar stars . We have the



Lemma : $$p(H) - 1 = \sum_{i=1}^{s(H)} (p_i - 1) \quad (11)$$

Proof : Consider the endpoint stars of H . They are of various classes . Remove all similar endpoint stars of a definite class, say class 1. Each of them contains p_1 dissimilar points . In this way one removes $p_1 - 1$ dissimilar points because an endpoint star by definition has only one articulation point, and this articulation point must be one of t dissimilar points and it is not removed by removing the endpoint star . In the remaining tree, one removes the next class of similar endpoint stars, by which one takes away $p_2 - 1$ dissimilar points , etc . One finally is left with a tree of only one class of stars in which there are now $p_{s(H)}$ dissimilar points left . This proves the Lemma .

Consider next the collection of all free pure trees of n stars . For each of them (11) applies . To make them rooted trees, each of the dissimilar points can be choosen as the root . Summing eq. (11) over all the free trees we therefore obtain

$$T_n - t_n = \sum_{H_n} \sum_{i=1}^{s(H_n)} p_i \quad - \sum_{H_n} s(H_n) \equiv A_n - B_n \qquad (12)$$

or for the corresponding counting series ($A(x) = \sum_n A_n x^n$, etc)

$$t(x) = T(x) - A(x) + B(x) \qquad (13)$$

A_n is the total number of dissimilar points in all the free pure trees of n stars , each counted k -fold if it occurs in k dissimilar stars . B_n is the total number of dissimilar stars in this collection of n-starred trees . It is equal to the number of ways of hanging rooted star trees (altogether n-1 stars) at the corners of one bas . star, where two configurations are counted as one if they are transformed into each other by a covering operation of the group Γ of the star . We therefore can apply Polya's theorem to find B(x) . The figure series is T (x) (rooted star trees !), s= number of points of the star and the group H_s is precisely Γ . The configuration series is therefore $Z\{\Gamma, T(x)\}$ and

$$B(x) = x\, Z\{\Gamma, T(x)\}$$

(the factor x ensures that the basic star is also counted ; one could also add unity to include $B_0 = 1$, but it would be cancelled by a term -1 in (13) resulting from - A (x) . We thus see that B (x) of (13) gives rise to the last term on the right-hand side of (10) .

In a similar way ,Norman should that - A (x) leads to the second term on the right-hand side of (10) .

§ 9. CONVERGENCE AND ASYMPTOTIC BEHAVIOUR OF COUNTING SERIES .

We will examplify the underlying problems for the case of Cayley trees for the counting series of which we derived (§ 7, eq.(8) in the case of rooted trees .

$$T(x) = \sum_{n=0}^{\infty} T_n x^n = e^{\sum_{r=1}^{\infty} x^r T(x^r)/r} \qquad (T_0 \equiv 1) \qquad (1)$$

The convergence and the analytic behaviour of T (x) were first investigated by Otter .

a) T(x) converges . Since the coefficients T_n are positive , $T(x^r) < T^r(x)$ for $x > 0$. From this it follows that for $x > 0$, $T(x)$ has as a majorant the function $y(x)$ which satisfies

$$y(x) = e^{\sum_{r=1}^{\infty} x^r y^r(x)/r}$$

Now

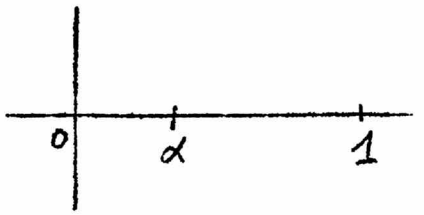
$$y(x) = e^{-\ln(1-xy)} = \frac{1}{1-xy}$$

or

$$y(x) = (1 - \sqrt{1-4x}) / 2x .$$

Hence the convergence radius of y(x) is $\frac{1}{4}$. The convergence radius of T(x) is therefore at least $\frac{1}{4}$. Furthermore , clearly $\alpha < 1$, so $\frac{1}{4} \leq \alpha < 1$.

b) T(α) = a is finite . According to Hardy's theorem for a series with positive coefficients , the first singular point lies on the positive real axis, it is the point α . In order to prove that Lim T(x) $x \to \alpha$



exists and has a finite value a , we remark that $\overline{T}(x) > \exp\left(x\,T(x)\right)$ (countains only the first term of $\sum$ in (1) for $x > 0$, so

$$\frac{T(x)}{\ln T(x)} < \frac{1}{x}$$

Since t (x) is monotonically increasing , it follows that T(x) bounded for $x \leq \alpha$

c) $T(\alpha) = \alpha^{-1}$ Consider the function

$$F(x,y) \equiv e^{xy + \frac{1}{2}x^2 T(x^2) + \ldots} - y \qquad (2)$$

For $x < \alpha$ the equation $F(x,y) = 0$ has the unique solution $y = T(x)$ and $F(\alpha, a) = 0$. Around $x = \alpha$, $y = a$ the function $F(x,y)$ is analytic in x and in y (in x since around $x = \alpha < 1$, we have $x^2 < \alpha$, so $T(x^2)$ is analytic , etc) . From this it follows that

$$\left(\frac{\partial F}{\partial y}\right)_{\substack{x=\alpha \\ y=a}} = 0 \qquad (3)$$

Suppose that $(\partial F/\partial y)$ were $\neq 0$. We could then develop $F(x, y)$ around $x = \alpha$, $y = a$ (where F is analytic) and from the theory of implicit functions it then would follow that y exists as an analytic function of x in this region . But this is contradictory to the fact that $x = \alpha$ is a singular point for $y = T(x)$.

From (2) and (3) we find

$$\left(\frac{\partial F}{\partial y}\right)_{\substack{x=\alpha \\ y=a}} = \left\{ x\, e^{x y + \frac{1}{2}x^2 T(x) + \ldots} - 1 \right\}_{\substack{x=\alpha \\ y=a}} = 0$$

d) Value of α. Inserting $x = \alpha$ in (1) , we obtain

$$t(\alpha) = \frac{1}{\alpha} = e^{1 + \frac{1}{2}\alpha^2 \; T(\alpha^2) + \ldots}$$

which can be solved for α by successive approximation . The first approximation would be $\alpha = e^{-1} = 0,368 \ldots$ Otter computed α to 7 decimals and found $\alpha = 0,3383219 \ldots$

e) α is a branch point of finite order . From (2) it follows that

$$\left(\frac{\partial F}{\partial x}\right)_{\substack{x=\alpha \\ y=a}} = T(\alpha)\left\{ a + \alpha T(\alpha^2) + \alpha^3 T'(\alpha^2) + \ldots \right\} \neq 0 \quad (4)$$

Since $F(x,y)$ is analytic in x and in y around $x = \alpha$, $y = a$, this implies that in this region x is an analytic function of y , and α is a branchpoint of finite order .

$$\left(\frac{dx}{dy}\right)_{\substack{x=\alpha \\ y=a}} = -\left(\frac{\partial F / \partial y}{\partial F / \partial x}\right)_{\substack{x=\alpha \\ y=a}} = 0$$

in view of (3) and (4) .

f) α is a branchpoint of order 2 . One easily computes $\partial^2 F(\alpha, a) / \partial y^2 = \alpha$. Therefore

$$\left(\frac{d^2x}{dy^2}\right)_{\substack{x=\alpha \\ y=a}} = -\left(\frac{\partial^2 F / \partial y^2}{\partial F / \partial x}\right)_{\substack{x=\alpha \\ y=a}} \neq 0$$

and so $T(x)$ must branch as $\sqrt{x-\alpha}$ at $x = \alpha$ and around this point

$$T(x) = \frac{1}{\alpha} + b\sqrt{\alpha - x} + R(x)(\alpha - x) + \ldots \quad , \qquad (5)$$

where $R(x)$ is regular . To calculate b we take the logarithmic derivative of (1)

$$T'(x) / T(x) = x\,T'(x) + x^3\,T'(x^2) + \ldots + T(x) + x\,T(x^2) + \ldots$$

so that

$$x\,T'(x)\left\{1 - x\,T(x)\right\} = T(x)\sum_{\nu=2}^{\infty}\left\{x^{\nu}\,T(x^{\nu}) + x^{2\nu+2}\,T'(x^{\nu+1})\right\}$$

With (5) this gives

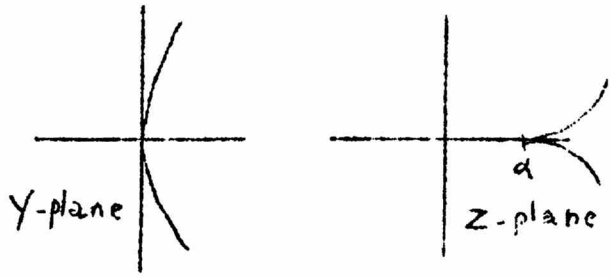
$$\frac{1}{2}\alpha^2 b^2 = \frac{1}{\alpha}\left[1 + \sum_{\nu=2}^{\infty}\left\{\alpha^{\nu}\,T(\alpha^{\nu}) + \alpha^{2\nu}\,T'(\alpha^{\nu})\right\}\right]$$

From this Otter computed $b = 7{,}924780$.

g) Asymptotic behaviour . In order to find the asymptotic behaviour of the T_n we consider the Cauchy integral (around the origin)

$$T_{n-1} = \frac{1}{2\pi i}\oint \frac{T(z)}{z^n}\,dz$$

and apply the method of steepest descent . The discussion is the same as that of the Kramers integral : putting $y = \sqrt{\alpha - z}$ around $z = \alpha$, one sees that in the y-plane $y=0$ is a steep maximum along the real axis , so the path of steepest descent is the imaginary axis . α is again a turning point for the path of steepest descent in the z - plane . Putting $y = i\eta$ one finds



$$T_{n-1} \simeq \frac{b}{2\sqrt{\pi}} \cdot \frac{\alpha^{-n+\frac{3}{2}}}{n^{3/2}}$$

so that the numbers increase exponentially . Already for $n = 10$ this formula is quite good : it gives $T_9 \simeq 708$, whereas the exact result is $T_9 = 719$

h) Free Cayley trees . As a special case of the Norman formula , we have for the counting series (see § 8, c 1)

$$t(x) = T(x) - \frac{1}{2} x \left\{ T^2(x) - T(x^2) \right\} \qquad (6)$$

So $t(x)$ has the same radius of convergence α as $T(x)$. The behaviour around $x = \alpha$ is sligtly different . Inserting (5) into (6) one sees that the coefficient of $\sqrt{\alpha - x}$ cancels and

$$t(x) = t(\alpha) - D(\alpha - x)^{3/2} + \ldots \qquad (7)$$

Again , α is a branchpoint of the order 2 , but the behaviour of $t(x)$ is different . For D one finds $\frac{1}{3} \alpha^2 b^3$ and for the asymptotic behaviour

$$t_{n-1} \simeq \frac{3D}{4\sqrt{\pi}} \cdot \frac{\alpha^{-n+\frac{5}{2}}}{n^{5/2}}$$

i) Generalization . The above treatment can be generalized to include such cases as cacti and pure Husimi trees . The counting numbers then rise faster and α is smaller . Ford (diss. 1954; Michigan) showed, that the general behaviour remains the same for arbitrary mixed Husimi trees (counted acoording to the number of points) . Even for this case the behaviour of the counting series is as $\sqrt{\alpha - x}$ for rooted trees and as $(\alpha - x)^{3/2}$ for free trees . And even for all mixed star trees (counted again according to the number of points) this result is found with the

restriction that the constituting stars consists of a finite and fixed number of linearly independent cycles .

§ 10 . THE POSSIBLE CONNECTION WITH THE CONDENSATION PROBLEM .

In answering the question of the connection between graph theory and the condensation problem, we recall first the Mayer equations (§5) , in which the cluster integrals

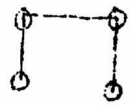
$$b_\ell(T)\ \frac{1}{\ell!\,V}\int\cdots\int U_\ell(\vec{r}_1\ \ldots\vec{r}_\ell)\ d\vec{r}_1 \ldots d\vec{r}_\ell \qquad (1)$$

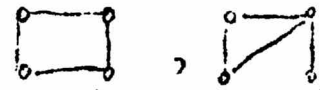
occured . The cluster function U_ℓ was defined as the sum of all products of functions f_{ij} represented by connected graphs of ℓ points and therefore U_ℓ is symmetric in the $\vec{r}_1 \ldots,\vec{r}_\ell$. This is the first point where graph theory comes in .

Furthermore , in Kahn's treatment of the condensation problem the necessary conditions for the behaviour of the series $\chi(z) = \sum_1^N b_\ell z^\ell$ (occuring in the second Mayer equation) were given in order that condensation should occur . As a matter of fact, we saw that the generating functions (the counting series) investigated in the preceding section , precisely satisfy these conditions . This suggests a connection with the theory of graphs .

Let us first investigate the cluster integrals . We then have to consider all connected graphs of ℓ points . If we order them first according to their number of lines . Let this number be k . With given ℓ then : $\ell - 1 \leq k \leq \frac{1}{2}\ell(\ell-1)$. With $k = \ell - 1$, the graph contains no cycles

and is a Cayley tree . For $k = \ell$, there is one cycle in the graph , which now is a mixed Husimi tree with 1 cycle (polygon) . Going on in this way , we see that if the graphs are ordered according to increasing k , they are at the same time arranged ccording to the number of cycles , which runs from





0 (for $k = -1$) up to $\frac{1}{2}(\ell - 1)(\ell - 2)$ (for $k = \frac{1}{2}\ell(\ell - 1) - 1$)

With given ℓ and k , we can further distinguish the different types of such graphs . Let $\gamma(\ell, k)$ be the number of topologically different connected graphs of ℓ points and k lines . For $\gamma(\ell, k)$ a functional relation is known .

Consider finally a definite graph i out of the $\gamma(\ell, k)$ different ones . The corresponding contribution to b_ℓ contains $\ell! / s_i(\ell, k)$ equal terms , where $s_i(\ell, k)$ is the symmetry number of the connected graph i with ℓ points and k lines ($\equiv$ the order of the group of that graph) , since there are $\ell!$ permutations of the points and $s_i(\ell, k)$ covering operations of that graph . So, we can rewrite (1) as

$$b_\ell(T) = \frac{1}{\ell! V} \sum_{k=-1}^{\frac{1}{2}\ell(\ell-1)} \sum_{i=1}^{\gamma(\ell,k)} \frac{\ell!}{s_i(\ell, k)} J_i(\ell, k) \qquad (2)$$

An extensive discussion of this cluster integrals and the virial coefficients , to which they give rise , is given by Riddell and Uhlenbeck , J.Chem.Phys. 21 (1953) 2056 .

Consider first the coefficient $c_i(\ell,k) = \ell!/s_i(\ell,k)$, that is the number of connected graphs with ℓ individualized points and k lines of topological class i. Then

$$c(\ell,k) = \sum_{i=1}^{\gamma(\ell,k)} c_i(\ell,k) \qquad (3)$$

is the total number of connected graphs with ℓ individualized points and k lines. This number can be found. We introduce the series

$$N(x,y) = \sum_{\ell=1}^{\infty} \sum_{k=0}^{\frac{1}{2}\ell(\ell-1)} \frac{x^\ell y^k}{\ell!} \binom{\frac{1}{2}\ell(\ell-1)}{k} \qquad (4)$$

which is the counting series for all graphs (connected and disconnected) with ℓ individualized points and k lines, since the binomial coefficient clearly gives the number of possible ways to take k lines out of $\frac{1}{2}\ell(\ell-1)$, which is the maximum number of lines for ℓ points. Analogously, we introduce

$$C(x,y) = \sum_{\ell=1}^{\infty} \sum_{k=\ell-1}^{\frac{1}{2}\ell(\ell-1)} c(\ell,k) \frac{x^\ell y^k}{\ell!} \qquad (5)$$

For these two series the relation

$$C(x,y) = \log\{1 + N(x,y)\} \qquad (6)$$

can be proved (cf. the above mentioned paper ; a simpler proof has been given by Ford in his dissertation). The numbers $c(\ell,k)$ are therefore determined by (4), (5) and (6) and one finds

$$c(\ell,k) = \binom{\frac{1}{2}\ell(\ell-1)}{k} - \ell\binom{\frac{1}{2}(\ell-1)(\ell-2)}{k} + \text{correction terms}.$$

For large p , and k not near to its limits $p-1$ and $\frac{1}{2}p(p-1)$ the second term is very small with respect to the firstone , which means that in that case the majority of the graphs are connected , a plausible result .

In order to find an asymptotic expression for $g(p,k)$ we consider the number $\pi(p,k)$ of topological different graphs of p points and k lines, connected or disconnected . It has been determined by Polya in the form of the counting polynomial

$$F(y) = \sum_{k=0}^{\frac{1}{2}p(p-1)} \pi(p,k)\, y^k \qquad (7)$$

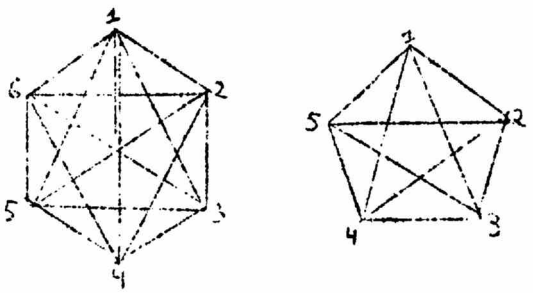
by means of the Polya theorem . The derivation is as follows .

Since we consider free graphs , the p points are equivalent so their group is the full symmetric group G_p . We can say that our figure collection consists of two objects : a line and no line . The figure counting series is therefore $1 + y$. However , we do not hang these figures on the p points. Now every permutation of the p points induces a permutation of the $\frac{1}{2}p(p-1)$ pairs of points . These permutations therefore form a permutation group of degree (number of objects) $\frac{1}{2}p(p-1)$. We will call it the pair group $\mathcal{H}_p$. According to Polya's theorem then

$$F_p(y) = Z\{\mathcal{H}_p, 1+y\} \qquad (8)$$

To find the cycle index, one has to determine what permutation of type $\{i_1, i_2, \ldots, i_{\frac{1}{2}p(p-1)}\}$ (i_1 one-cycles , i_2 two-cycles , etc) corresponds to a permutation of type $\{j_1, j_2, \ldots, j_p\}$ in G_p . One has to distinguish :

a . Points , occuring in one cycle , say of length ν . Arrange

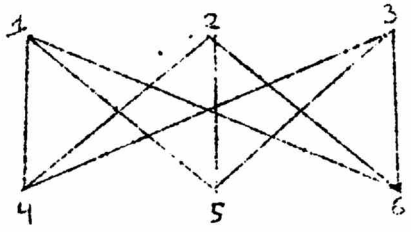


the ν points on a polygon and number them . Now under a cyclic permutation (1,2,..., ν) the point pair 1-2 goes into the pair 2 - 3 , etc, which percisely gives rise to a point pair cycle of length ν . Further, pair 1-3 goes into 2-4 , etc, which gives rise to another point-pair cycle of length ν . So one can go on .Clearly one gets :

- for ν odd : $\frac{1}{2}(\nu-1)$ cycles of length ν in the pair group ,
- " ν even: $\frac{1}{2}(\nu-1)$ " " " " " "

and one cycle (consisting of the main diagonals) of length $\frac{1}{2}\nu$.

b . Points occuring in different cycles .



If the cycles are all of equal length (say ν) then they clearly induce ν cycles of length ν in the pair group . If the cycles are of unequal length (say ν_1, ν_2) then their permutations will induce cycles of length m (ν_1, ν_2) = least common multiplier of ν_1, ν_2 and the number of such cycles will be d (ν_1, ν_2) = largest common divisor of ν_1, ν_2. Remember $m(\nu_1,\nu_2)\, d\,(\nu_1,\nu_2) = \nu_1 \nu_2$

Since there are $\ell\,! / \pi\, k^{j_k}\, j_k\,!$ permutations of type ($j_1, j_2 \ldots$ j_ℓ) in G_ℓ one gets for the cycle index of $\mathfrak{H}_\ell$:

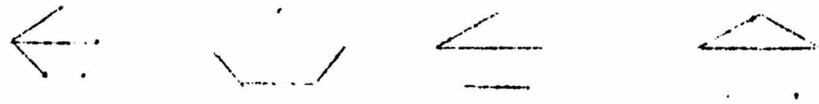
$$Z(\mathcal{H}_\ell) = \frac{1}{\ell!} \sum_{\{j(k)\}} \frac{\ell!}{\prod_{k=1} j_k! \, k!^{j_k}} \prod_{k \text{ odd}} b_k^{\frac{1}{2} j_k (k-1)} \prod_{k \text{ even}} b_k^{\frac{1}{2} j_k (k-2)} b_{\frac{1}{2}k}^{j_k} \times$$

$$\times \prod_{k} b_k^{\frac{1}{2} \sigma_k (j_k - 1) k} \prod_{k<\ell} b_{m(k,\ell)}^{\sigma_k j_\ell \, d(k,\ell)} \qquad (9)$$

$F_\ell(y)$ follows from (8) and (9) . For $\ell = 5$, one finds , e.g.,

$$F_5(y) = 1 + y + 2y^2 + 4y^3 + 6y^4 + 6y^5 + 6y^6 + 4y^7 + 2y^8 + y^9 + y^{10}$$

The 4 different graphs of 5 points and 3 lines are :



From the explicit expression for $F_\ell(y)$ one can also derive the asymptotic behavior of $\pi(\ell, k)$. For large ℓ , and k not near to the end points of its domain $\left(0 \leftrightarrow \frac{1}{2}\ell(\ell-1)\right)$ one finds :

$$\pi(\ell, k) \cong \frac{1}{\ell!} \binom{\frac{1}{2}\ell(\ell-1)}{k} \qquad (10)$$

which means that the majority of the graphs have no symmetry (the binomial coefficient gives the number of connected + disconnected graphs with ℓ individualized points and k lines , the factor $1/\ell!$ removes the distinguishability of the points) .

We see therefore , that for large ℓ and " average " values of k the majority of the graphs are connected and have no symmetry , so that

$$\gamma(\ell, k) \cong \frac{1}{\ell!} \, c(\ell, k) \cong \pi(\ell, k) \qquad (11)$$

Returning to eq.(2), let us suppose that $J_i(\ell, k)/V$ is approximately constant (independent of k and ℓ) . The behaviour of b_ℓ would then be

$$b_\ell \sim \frac{J}{\ell!} \sum_{k=\ell-1}^{\frac{1}{2}\ell(\ell-1)} C(\ell, k) \cong J \sum_{k=0}^{\frac{1}{2}\ell(\ell-1)} \frac{1}{\ell!} \binom{\frac{1}{2}\ell(\ell-1)}{k} = \frac{J}{\ell!} 2^{\frac{1}{2}\ell(\ell-1)}$$

and

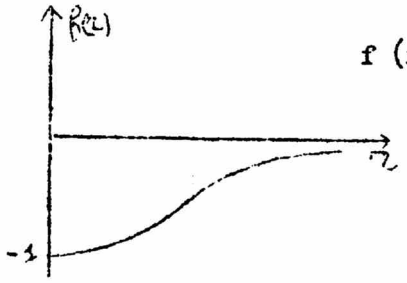
$$\chi(z) \sim \sum_\ell \frac{1}{\ell!} 2^{\frac{1}{2}\ell(\ell-1)} z^\ell$$

therefore $\chi(z)$ would always be divergent for $z > 0$. So we have to know something about the dependence of $J_i(\ell, k)$ on k . The integral should decrease for larger values of k in order that $\chi(z)$ be convergent . This will actually be the case since adding a line between two points means that we introduce a factor in the integrand which requires the two points to be less than a certain distance apart . Thus with increasing k the integrand will differ from zero over a smaller region of the 3ℓ- dimensional phase space .

§ II . THE INTEGRAL PROBLEM . The integrals $J_i(\ell, k)$ should cause the convergence of $\chi(z)$ by suppressing a large number of configurations . To study the behaviour of the integral ,we have to introduce a special type of intermolecular potential, or rather a special choice for the function $f(r)$. For this purpose, it is useful to consider :

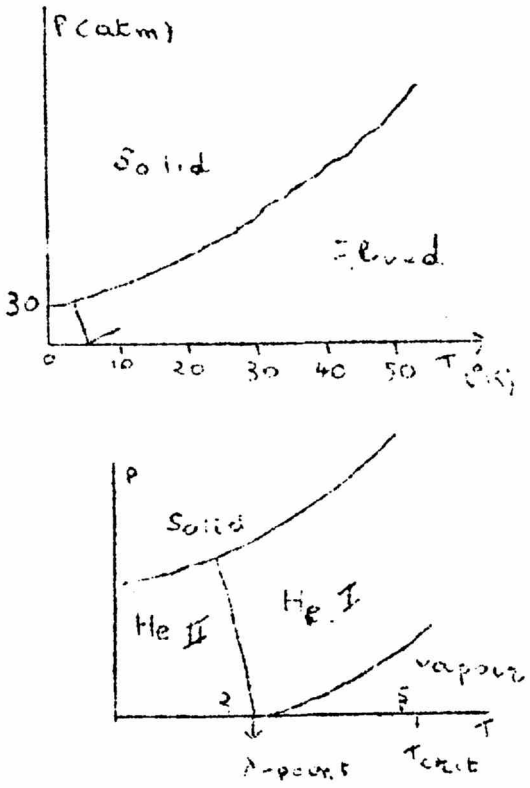
A. The Gaussian model . In § 5 , we gave the general behaviour of $f(r)$ for short-range repulsion + long-range attraction (e.g. the Lennard-Jones potential) . If we would take $f(r) = A e^{-\alpha r^2}$ with $A > 0$

(corresponding to attractive forces only) we would suppress the negative part of f(r) and clearly we could not expect $\chi(z)$ to be convergent . To include the repulsive part of the intermolecular force , we could add a term $-(1+A)\,e^{-\beta r^2}$ (remember that f(o) = - 1) . However, it is of interest to consider only one Gaussian



$$f(r) = -e^{-\alpha r^2}$$

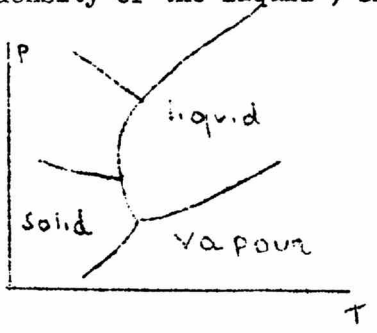
corresponding to repulsive forces only . The motivation of this choice is the striking and paradoxical result of Kirkwood and co-workers, who showed from the behaviour of the phase integral and on the basis of the so-called " superposition approximation " that a gas of molecules with only repulsive forces (e.g. elastic spheres) shows a phase transition : at high pressure the system is split up into two phases with different density and entropy , a solid phase (arrangements of ordered molecules) surrounded by a liquid phase (series of articles from Kirwood and Monroe, J. Chem. Phys. 9 (1941) 514 up to Kirwood , Maun and Alder , J.Chem. Phys. 18 (1950) 1040) Although perhaps hard to believe , one must admit that there is no rigorous argument which disproves the existence of such a " condensation " . In addition there is the suggestion that perhaps the Kirwood transition has something to do with the solidification of helium , which is known to occur at temperatures many times the critical temperature (5 º k) if the pressure is high enough .



PHASE DIAGRAM OF HELIUM

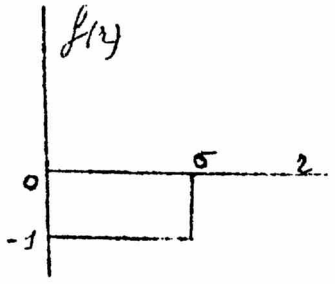
This can hardly be ascribed to the weak attractive forces between the helium atoms which are of the order of k $T_{crit.}$ The solidification of a gas is perhaps a general consequence of the sharp repulsive forces. (For helium the intermolecular force is very well known ; as for all chemically non-active molecules it consists of a Van der Waals attraction on which a sharp repulsive core is superimposed) .

It might well be a general fact of nature that any assembly of molecules can at any temperature be brought into the solid state if one raises the pressure high enough . Even the phase diagram of water (where the density of the solid state is less than the density of the liquid) shows this behaviour .



PHASE DIAGRAM OF WATER.

If this is indeed a general fact of nature the explanation in terms of the intermolecular forces should be found in some general feature of these forces . Such a feature is the presence of a sharp repulsive core .



With only repulsive forces one would think of elastic spheres of diameter σ . However, the cluster integrals are hard to calculate for this case and the virial coefficients have only been calculated up to the 4 th one . Since the kirkwood transition will according to the above discussion be independent of the special form of the repulsive potential, we will consider the Gaussian model (1) for the function $f(\;)$. We then have to consider for a connected graph of type i

$$\frac{1}{V} J_i(\ell,k) = \frac{A^k}{V^\ell} \int \cdots \int e^{-\sum_{n,m}^{\ell} |\vec{r}_n - \vec{r}_m|^2} d\vec{r}_1 \cdots d\vec{r}_\ell \qquad (2)$$

with $A = -1$, where the sum is over those lines n, m which occur in the graph . This sum can be written as a quadratic form

$$\sum_{n,m}^{\ell} |\vec{r}_n - \vec{r}_m|^2 = \sum_{n=1}^{\ell} \sum_{m=1}^{\ell} \Delta_{nm} r_n r_m \qquad (3)$$

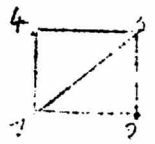
where Δ_{nm} is given by

$$\Delta_{n\,m} \begin{cases} = -1 & \text{if the line } (n,m) \text{ occurs} \\ = 0 & \text{" " does not occur} \\ = \text{number of lines attached to the point } (= \text{branching number if} & n = m \end{cases}$$

We then introduce the matrix with elements Δ_{nm} , the graph matrix for this connected graph . $\Delta_{n\,m}$ clearly is symmetric . The diagonal elements are the branching numbers of the ℓ points , the off-diagonal

elements are either 0 or -1 . As an example , we give the graph matrix for one special graph of 4 points :

$$\begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 3 & -1 \\ -1 & 0 & -1 & 2 \end{pmatrix}$$



B . Properties of the graph matrix .

a. The determinant $\|\Delta\|$ is zero . This is immediately clear from the definition of $\Delta_{n\,m}$ since by adding all columns to the firstone the elements of the first column become (branching number - branching number) = 0 .

b. The minors of rank $\ell - 1$ are all equal (in absolute value). (For disconnected graphs the complexity is zero) . Their value is called the graph complexity $d_i(\ell, k)$ We will show this for the ℓ principal minors (obtained by striking out the n-th row and column) . If we take the position of the first point as the origin , $r_1 = 0$ and the first row and column of Δ do not occur in (3) , the quadratic form thus being determined by the (1,1) - minor . The integration over r_1 then leads to a factor V and one gets

$$\frac{1}{V} J_i(\ell,k) = A^k \int_V \cdots \int e^{-\alpha \sum_{n=2}^{\ell} \sum_{m=2}^{\ell} \Delta_{n\,m}\, r_n\, r_m} \, d\vec{r}_2 \cdots d\vec{r}_\ell =$$

$$= A^k \left(\frac{\pi}{\alpha}\right)^{3/2\,(\ell-1)} \left\{ d_i(\ell,k) \right\}^{-3/2}$$

where $d_i(\ell,k)$ is the minor obtained by striking out the first row and column in Δ . Clearly , the same value is obtained by striking out the n-th row and column , since the result cannot depend on the choice of the origin . The ℓ principal minors have therefore the same absolute value .

With $(r/\alpha)^{3/2} = 2b$ (the first virial coefficient, following from the Gaussian model , then becomes equal to b) the expression for the cluster integral becomes

$$b_\ell = (2b)^{\ell-1} \sum_{k=\ell-1}^{\frac{1}{2}\ell(\ell-1)} A^k \sum_{i=1}^{\gamma(\ell,k)} \frac{1}{s_i(\ell,k)} \left\{\frac{1}{d_i(\ell,k)}\right\}^{3/2} \qquad (4)$$

c . For connected graphs with articulation points (star trees) the complexity is the product of the complexities of the constituing stars :

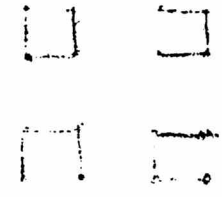
$$d_i(\ell,k) = \prod_{\text{all stars}} d^{\text{star}}$$

This is again obvious from the integral representation .

d . The complexity of a star is equal to the number of different Cayley trees of ℓ individualized points that can be formed from the ℓ points and k lines , occuring in the star . This theorem was already known to Kirchhoff (collected Works) , who derived it in connection with the theory of electric circuits . The general proof is given by Ford (diss.1954, Univ. of Michigan) . The idea of the proof is to write out the complexity, then each term corresponds to one Cayley tree . We give an example :



$$\Delta = \begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{pmatrix} \quad d = 4$$



e. The maximum value of $d_i(\ell,k)$ for a given value of ℓ is $\ell^{\ell-2}$ and is obtained for a complete graph (all pairs connected, $k = \frac{1}{2}\ell(\ell-1)$). For a complete graph $\Delta_{nm} = \ell - 1$ if $n = m$ and $= -1$ for $n \neq m$. The matrix Δ has ℓ rows and columns and $|\Delta| = 0$ whereas forming the principal minor, we get a cyclic determinant, the value of which is $\ell^{\ell-2} = d$. Using the Kirchhoff theorem one can therefore conclude that there are $\ell^{\ell-2}$ Cayley trees with ℓ individualized points, which is a classical result.

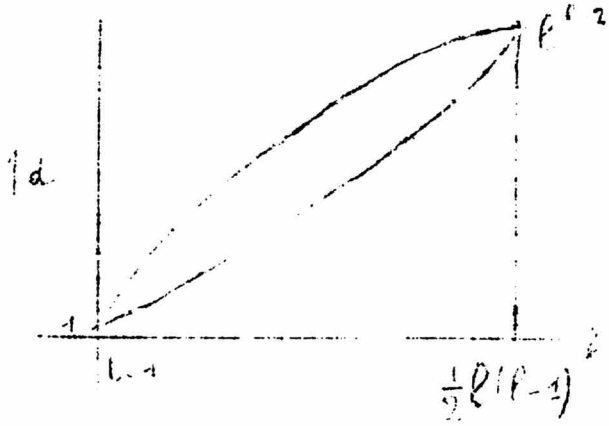
$$\begin{pmatrix} \ell-1 & -1 & -1 & \dots & -1 \\ -1 & \ell-1 & -1 & \dots & -1 \\ -1 & -1 & \ell-1 & \dots & -1 \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -1 & -1 & \dots & \ell-1 \end{pmatrix}$$

f.
$$\sum_{i=1}^{\gamma(\ell,k)} c_i(\ell,k)\, d_i(\ell,k) = \ell^{\ell-2} \; \frac{\frac{1}{2}(\ell-1)(\ell-2)}{k-\ell+1} \qquad (5)$$

The proof of this theorem (due to Ford) is simple. The left-hand side is the total number of all possible Cayley trees of ℓ individualized points and k lines (including disconnected trees since then $d_i = 0$) This number can also be found by starting with a complete graph of ℓ points, $\frac{1}{2}\ell(\ell-1)$ lines. From this graph, we can form $\ell^{\ell-2}$ different Cayley trees of ℓ points, $\ell - 1$ lines. Consider one of them. From the remaining $\frac{1}{2}(\ell-1)(\ell-2)$ lines, discard $k - \ell + 1$ lines. This will lead to a connected graph which

contains the chosen Cayley tree and any way of discarding leads to a different graph . Therefore the chosen Cayley tree will be contained in $\binom{\frac{1}{2}(\ell-1)(\ell-2)}{k-\ell+1}$ different graphs . This leads to the number on the right- and side of (5).

C. The integral problem . At the end of § 10 , we saw that the integral $J_i(\ell,k)$ is decreasing if k approaches the maximum of its domain. From (4) we see that for the Gaussian model one needs to know something about the distribution of the values of (Cayley tree) and that $d = \ell^{\ell-2}$ for $k = \frac{1}{2}\ell(\ell-1)$ (complete graph) . For values of k between these limits there will be a certain range of d_i- values . Suppose we can introduce a distribution function $n(\ell, k, d)$ for the values of d .



We know :

$\sum_d n(\ell, k, d) = C(\ell, k)$ = total number of connected graphs of ℓ individualized points and k lines , asymptotically $\simeq \binom{\frac{1}{2}\ell(\ell-1)}{k}$,

$\sum_d d\, n(\ell, k, d) = \ell^{\ell-2}\binom{\frac{1}{2}(\ell-1)(\ell-2)}{k-\ell+1}$ according to (5)

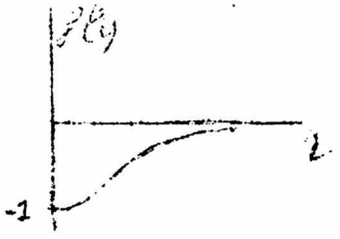
So, we know the zeroth and first moments of this distribution function and if the distribution were a Gaussian one , the knowledge of the second moment (or the spread) would be sufficient to determine the complete distribution function .

The Gaussian character has been investigated by tests (diss.Ford) . For $\ell = 7$ there are about 40 values of d for k near the middle of its domain (remember that for increasing ℓ the overwhelming majority of all graphs have k in the middle) and the histogram of these values showed indeed a Gaussian behaviour .

But a rigorous proof of the Gaussian character is prc bly a very fundamental problem .

§ 12 . RESULTS . The final answers are meagre .

I - For the Gaussian model of repulsive forces the first 7 cluster integrals b can be calculated . For the virial expansion of the equation of state .

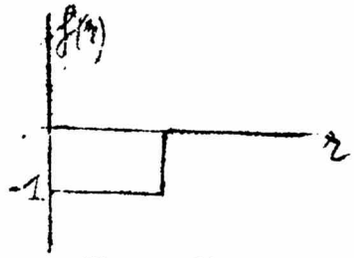


$$\frac{p\,v}{k\,T} = 1 + \frac{B_2}{v} + \frac{B_3}{v^2} + \ldots.$$

one finds

$$\frac{p\,v}{k\,T} = 1 + \frac{b}{v} + 0.287\,\frac{b^2}{v^2} - 0.125\,\frac{b^3}{v^3} + 0.013\,\frac{b^4}{v^4} + 0.038\,\frac{b^5}{v^5} - 0.030\,\frac{b^6}{v^6} + \ldots$$

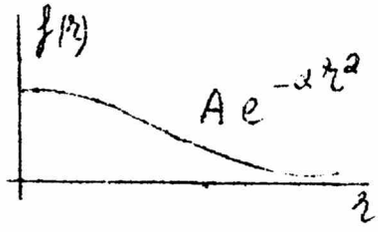
As far as calculated, the series of cluster integrals b_ℓ is alternating (both for the Gaussian model and for elastic spheres) .



$$\frac{pv}{kT} = 1 + \frac{b}{v} + 0.625\,\frac{b^2}{v^2} + 0.287\,\frac{b^3}{v^3} + 0.116\,\frac{b^4}{v^4} + \ldots$$

The coefficient B_3 was already given by Boltzmann, B_4 by van Laar and checked by Nyboer and Van Hove (Phys.Rev. 85 (1952) 777) . B_5 was given by Rosenbluth (J.Chem. Phys. 22 (1954) 884) .

II - One can also consider the Gaussian model for purely attractive forces . With $A > 0$, the integral $J_i(\ell, k)$ and also b_ℓ are always positive.

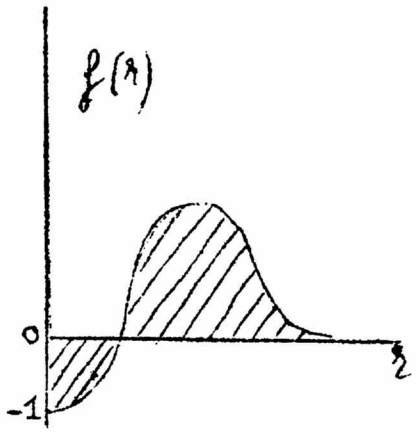


With increasing k (increasing complexity of the graph) $J_i(\ell,k)$ decreases . Taking the maximum value $\ell^{\ell-2}$ for d in (4) one then can find lower limits for the b_ℓ and it turns out that the series $\chi(z) = \sum_\ell b_\ell z^\ell$ is always divergent . So the integrals $J_i(\ell,k)$ do not decrease fast enough to overcome the increase in the total number of connected graphs with ℓ points .

With purely attractive forces the free energy Ψ would always be proportional to the number of pairs of molecules (so $\sim N^2$) instead of proportional to N .

Van Hove (Physica 15 (1949) 951) has shown that the sharp repulsive core is necessary for the proportionality of all thermodynamic quantities with N and therefore for the existence of an equation of state . The divergence of $\chi(z)$ for the attractive Gaussian model is in agreement with Van Hove 's result .

III - For an attractive intermolecular force with a repulsive core



f (r) is of the form given in the figure . At low temperature the area under the positive part is large with respect to the area under the negative part . The integrals for k in the beginning of its domain will then certainly be positive (graphs which are Cayley trees, Husimi trees with one cycle , with 2 cycles etc) . For larger values of k the contributions of smaller values of r become more and more important (the graph is " clustered up " , new factors f_{nm} require that the molecules are less than a certain distance apart) . This will allow changes in sign of the $J_i(\ell, k)$, and for higher values of k they will be alternating in sign . With many cycles , it seems therefore likely that the contributions of all the graphs for the largest values of k will roughly cancel each other . This would lead to a certain " cut off" . In his thesis , Ford showed that for a reasonable cut off the series $\chi(z)$ has exactly the Kahn properties for condensation as a consequence of the theorem mentioned in § 9 , i on the counting series of mixed star trees . But, of course, this should not be considered as a rigorous proof for the occurrence of condensation.

IV - With <u>purely repulsive forces</u> there remains the Kirkwood conjecture . In the Gaussian model $A = -1$, so the integrals $J_i(\ell, k)$ are alternating in sign with increasing k . Because of the strong cancellation it is hard to make asymptotic estimates for the b_ℓ . Ford derived a number of inequalities and estimates . The obtained estimated values of b_ℓ gave rise

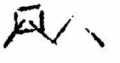
to series $\chi(z)$ which had the first singular point on the negative real axis , which would contradict the possibility of a Kirkwood condensation . But a definite conclusion would only be possible with estimates of b_e, more exact than those given by Ford .

NOTES ON THE THEORY OF PHASE TRANSITIONS

by

George E. UHLENBECK
University of Michigan .

E R R A T A

Page	Line	Read :
2	1	$N,V \to \infty$ ... N / V
	3	$\lim_{N,V \to \infty} \Psi$
5	1	λ_1 and therefore ...
	2 from below	$\int_0^{\pi/2}$
7	13	... like (1) see ...
8	5	add : $(\mid z \mid < 1)$
	1 from below	... given by
		$g'(z_0,a) = a\chi'(z_0) - z_0 = 0$ or $z_0\chi'(z_0) = \frac{1}{a}$
9	7	$-\Psi / kT = N g(z_0,a) - \frac{1}{2} \log N + \ldots$
	3 from below	$a_c^{-1} = \chi'(1) = \ldots$
10	2	Physica 4
13	5 from below	9 rooted Cayley trees .
		figure :  instead of 

Page	Line	Read
14	1	Cayley trees with arbitrary polygons .
15	1o	$N = 4$ (2^6 terms)
	3 from below	$\sim V^4 \quad V^3 \ldots$
16	2	... in m_1 single ...
	3	m_ℓ sets of ℓ molecules .
17	3	$b_1 \equiv 1$
	10 from below	$(V 2 ! b_2)^2$
	7 from below	$(V 3 ! b_3) (V 1 ! b)$
18	5	for one set of the m_ℓ
19	5	$- k T V \sum_\ell b_\ell z^\ell$
24	7	$A_{s,s}$
	1 from below	$\equiv \mathcal{Z}\left\{ H_s, f(x,y) \right\}$
26	2	$\sum_{n=0}^{\infty} A_n x^n$
28	3	k^{j_k}
30	6 from below	$T(x) = \sum_{n=0}^{\infty} T_n x^n$
33	7	and there are k points , $H_s = G_k$
	9	(only ℓ [illegible] triangles)
34	1	$\frac{y^s}{2\,s}$

Page	Line	Read
34	6 from below	$t(x) = T(x) - \frac{1}{3} x \left\{ T^3(x) - T(x^3) \right\}$
35	7,8	$p_1 = 3$, $p_2 = .2$
36	1 from below	Norman showed
37	6 " "	The convergence radius α of
38	1 " "	add :
		so $\alpha T(\alpha) = 1$ or $a = T(\alpha) = \frac{1}{\alpha}$
39	2	$T(\alpha) = \frac{1}{\alpha} =$
43	6 from below	of the ℓ points
	4 " "	$\sum_{k=\ell-1}^{\frac{1}{2}\ell(\ell-1)}$
44	7 from below	$x^{\ell} y^{k}$
45	8	$F_{\ell}(y)$
	9/8 from below	.. on the ℓ points, but on the $\frac{1}{2}\ell(\ell-1)$ pairs of points . Now ...
	5 " "	pair group $\mathfrak{H}_{\ell}$
46	4	$(1, 2, \ldots, u)$
	10	for u even : $\frac{1}{2}(u-2)$ cycles
47	1	$S' \frac{\ell!}{(j_k)! \prod_{k=1}^{\ell} j_k! k^{j_k}}$

Page	Line	Read
47	9 from below	$\left(\cdots + \frac{1}{2} \ell \, (\ell - 1) \right)$
48	4	$\sum_{k=\ell-1}^{\frac{1}{2}\ell\,(\ell-1)}$
52	1 from below	$A^k \left(\frac{\pi}{\alpha}\right)^{3(\ell-1)/2}$
53	5	With $(\pi/\alpha)^{3/2} = 2\,b$
55	7	about the distribution of the values of $d_i(\ell, k)$ for given ℓ and k . We know that $d = 1$ for $k = \ell - 1$ (Cayley tree) ...
57	1	For comparison we give the virial expansion for elastic spheres :

THE STATISTICAL MECHANICS OF NON-EQUILIBRIUM PHENOMENA

by

George E. UHLENBECK

University of Ann Harbor, Michigan

These lecture notes were previously published for a series of lectures given at Princeton University in the fall of 1954.

Université de Grenoble

Cours professé à

l'ECOLE D'ETE DE PHYSIQUE THEORIQUE

Les Houches (Haute-Savoie), France

18 juillet - 15 août 1955

THE STATISTICAL MECHANICS OF NON-EQUILIBRIUM PHENOMENA

1. The problem of Boltzmann.

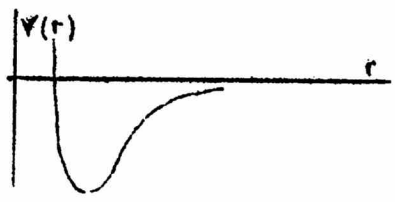
How can one possibly "explain" the irreversible behaviour of thermodynamical system from a strictly reversible mechanical model?

This will be the theme for my first set of lectures. It is properly called the problem of Boltzmann, since he struggled with it during his whole life. Of course after his work and the further developments by Gibbs, Einstein, Smoluchowski, Ehrenfest and others, the essential feature of the solution have become quite clear and well known. However, it may be good to review the situation, because especially for non-ideal gases the concrete questions of how to deduce the non-equilibrium properties of the system from its molecular constitution are still far from a satisfactory solution.

I will restrict myself by using :

(a) Classical mechanics. It is time that the quantum mechanics, especially thru the act of measurement, adds a new feature (perhaps a "time" irreversibility) to the problem, and it is also true that often the quantum theory is formally simpler, but I think the essential questions arise already in the classical theory, and since I am most familiar with it I will restrict myself to the classical theory.

(b) Point molecules interacting thru short-range central forces which have the additivity property (no many-body forces). One must think of a van der Waals or molecular potential but I will often consider the case of repulsive forces only (elastic spheres or repulsion $\sim 1/r^5$, the Maxwell molecules).



2. The kinetic method and the H-theorem.

The central question is the explanation of the second law of thermodynamics; however, it is simple to see that if one can explain one typical irreversible event, and specifically the approach to thermal equilibrium, then everything follows. This is the reason why the proof of the Maxwell-Boltzmann distribution law has such importance. Can one show for a gas that the distribution functions

$f(t, \vec{r}, \vec{v})\, d\vec{r}\, d\vec{v}$ = number of molecules in μ-space cell $d\vec{r}\, d\vec{v}$

approaches in course of time the M. B. distribution

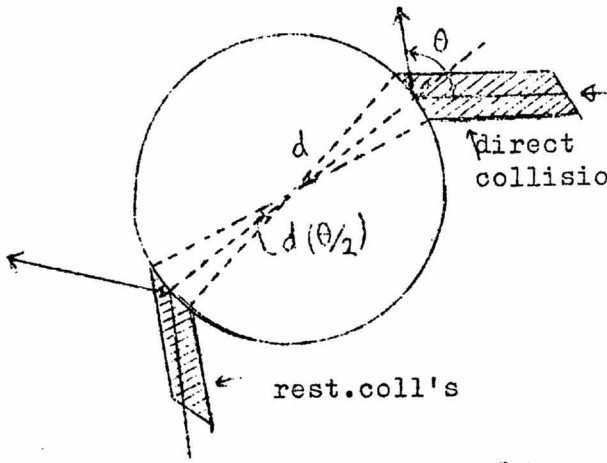
$$f_o = A\, e^{-\frac{1}{kT}\left\{\frac{mv^2}{2} + U(\vec{r})\right\}}$$

where U = potential of outside forces, and A is determined by total number of molecules?

I now have to remind you of the so-called kinetic proof of Boltzman The starting point is the basic continuity. Equation :

$$\frac{Df}{Dt} \equiv \frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_r f + \vec{a} \cdot \nabla_v f \equiv \int d\vec{v}_1 \int d\Omega\, I(g,\theta)(f'f'_1 - f f_1) \qquad (1)$$

where $\vec{a}$ = acceleration of outside forces; index 1 and prime refers to the velocity variable only, so that $f_1 \equiv f(r, v_1, t)$ etc.; $(\vec{v}, \vec{v}_1) \rightleftarrows (\vec{v}', \vec{v}'_1)$ are the direct and restituting collisions; $g \equiv |\vec{v}_1 - \vec{v}| = |\vec{v}'_1 - \vec{v}'|$; $I(g,\theta)$ is differential cross section for turning the rolative velocity over an angle θ "in solid angle dr". For instance for elastic spheres :



Volume cylinder :

$$g\, d^2 \cdot \cos\frac{\theta}{2}\; 2\pi \sin\frac{\theta}{2}\, d\,\frac{\theta}{2}\; d\phi$$

$$= g\,\frac{d^2}{4}\, d\Omega .$$

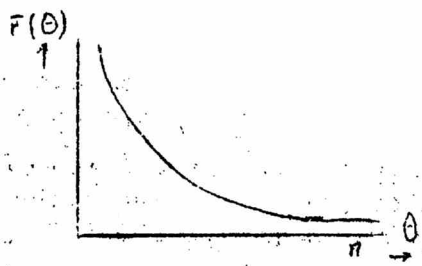
So $I(g,\theta) = \frac{d^2}{4}$.

For Maxwell molecules, force law k/r^5, one finds :

$$g\, I(g,\theta) = \sqrt{\frac{2k}{m}}\; F(\theta),$$

independent of g. $F(\theta)$ is complicated, is monotonic decreasing; for small θ ,

$$F(\theta) = \frac{\sqrt{3\pi}}{15}\; \frac{1}{\theta^{5/2}}\; \left(1 + \frac{35}{24\pi}\theta + \ldots\right)$$



It is well approximated by :

$$F(\theta) \sim \frac{\sqrt{3\pi}}{\sqrt{2}}\; \frac{1}{\sin^5\sqrt{\theta/2}}$$

The usual justification of (1) is familiar; f changes for two reasons : the streaming in μ -space, and the collisions. The collision term is crucial; it produces the irreversibility. It is based on the Stoszzahl Ansalz, which is an assumption of a non-mechanical nature. It is where the statistics comes in! Note the following limitations :

(a) Some uniformity is assumed. Strictly, the positions of the colliding molecules should not be the same, and variations over distances of the order of a mean free path are perhaps already doubtful. Thus very fast varying phenomena as high frequency sound (freq. of order of collision frequency, or wavelength of order of free path) may not be correctly described.

(b) Only binary collisions are assumed, and no correlations of the velocity directions is taken into account, although forces are present.

Clearly Boltzmann had some sort of a successive approximation idea in mind. Already the description of the state of the gas by $f(\vec{r}, \vec{v}, t)$ implies that in zero. approx. the forces are neglected and the gas is ideal. The collisions depending on pairs of molecules are the next approxi ation, and are responsible for the transport phenomena (viscosity, heat cond.) and also for the first deviation (2nd virial coefficient) of the ideal gas law.

The approach to equilibrium is proved by means of the H-theorem. Boltzmann defines :

$$H(t) \equiv \iint F \log f \, d\vec{r} \, d\vec{v} \tag{2}$$

and he proves

$$1) \quad \frac{d}{dt} H = \iint \frac{\partial f}{\partial t} \log f \, d\vec{r} \, d\vec{v} = \left(\frac{d}{dt} H\right)_{\text{streaming}} + \left(\frac{d}{dt} H\right)_{\text{collisions}}$$

Due to the streaming of the H function does not change. This is a special use of a general theorem. If there are no collisions (so no forces) then the gas is an ensemble (copies of identical systems each in different states of motion and each sending one point into the phase space; systems are now the molecules, phase space is μ-space). The Liouville theorem states that the streaming in such a case is as an incompressible fluid (or is an "equivolumic transformation"). Therefore any integral of the form

$$\int \cdots \int F(p) \, dq_1 \cdots dp_n$$

must be independent of the time.

2) Due to the collisions H always decreases, and dH/dt is only zero if for all possible collisions :

$$f f_1 = f' f'_1 \tag{3}$$

This determines the dependence of f on the velocities; log f must be a linear function of the quantities conserved in a collision, which leads to

$$f = A e^{-\beta \frac{m}{2} \left\{ (u - u_0)^2 + \cdots \right\}} \tag{4}$$

In every volume elem. in velocity space the approach to (4) is monotonic. The approach to the M. B. distribution in coordinate space is much more complicated. In (4) the quantities A, u_o, v_o, w_o can still be functions of x, y, z and t, which must follow from

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_r f + \vec{a} \cdot \nabla_v f = 0 \tag{5}$$

Only for sufficiently general outside forces (which include shape of the vessel) is the solution unique and corresponds to β = const, $u_o = v_o = w_o = 0$ and $A = A_o e^{-\beta U}$. For special forces there are a host of special solutions, which can even be time dependent. See

Boltzmann, Wiss. Abh.II, p.83

Since such solutions are not well known, allow me to mention one of them. Consider the case of a harmonic attractive force so that $\vec{a} = w^2 \vec{r}$. Writing :

$$\log f = -hv^2 + \vec{k} \cdot \vec{v} + n$$

$$\left[h = \frac{\beta m}{2} ; \quad \vec{k} = \beta m \vec{v}_o ; \quad n = \log A - \frac{\beta m}{2} v_o^2 \right].$$

One gets from (5) :

$$\left.\begin{aligned} &\frac{\partial n}{\partial t} - \omega^2 \chi_\alpha K_\alpha = 0 && \text{(a)} \\ &\frac{\partial K_i}{\partial t} + \frac{\partial n}{\partial \chi_i} + 2\omega^2 \chi_i h = 0 && \text{(b)} \\ &\frac{1}{2}\left(\frac{\partial K_i}{\partial \chi_j} + \frac{\partial K_j}{\partial \chi_i}\right) - \delta_{ij} \frac{\delta h}{\delta t} = 0 && \text{(c)} \\ &\frac{\partial h}{\partial \chi_i} = 0 && \text{(d)} \end{aligned}\right\} \qquad (6)$$

So h is a function of t alone, and :

$$\vec{K} = \left[\vec{\Omega} \times \vec{r}\right] + c\vec{r} + \vec{d} ; \quad \frac{dh}{dt} = +c ; \qquad (7)$$

the motion is combination of rotation, radial expansion, and translation; Ω c, $\vec{d}$ can still be functions of t. Eq (7) is still general, independently of outside forces. From (6a), (6b), one gets : Ω = const. and

$$\ddot{d}_i + \omega^2 d_i = 0, \qquad \ddot{c} + 4\omega^2 c = 0$$

by eliminating n. Take for example the case :

$$\vec{d} = 0 \qquad c = c_o \cos 2\omega t$$

This gives :

$$h = h_o + \frac{c_o}{2\omega} \sin 2\omega t,$$

and then :

$$n = -\omega^2 h_o r^2 + \frac{\omega c_o r^2}{2} \sin 2\omega t.$$

If h_0 is large enough, one gets a Gaussian density distribution with a curious "breathing" motion, accompanied by periodic temperature changes. In addition there can be rotations and translatory motions in which the center of the Gaussian gas cloud oscillates around the origin.

Already for an anisotropic hermonic potential $U = \frac{1}{2}\omega_1^2 x_1^2 + \frac{1}{2}\omega_2^2 + \ldots$ One proves from eqs.(6) that h is constant in time, and $\vec{v}$ and c are zero. Oscillatory solutions are still possible. They will disappear if anharmonic terms are added.

The reason for the curious solutions is clearly the isochronic property of harmonic forces, which prevent them from producing the proper mixing.

3) Excursion on the linearized Boltzmann equation.

Since the H-theorem does not tell us the approach to equilibrium in detail, and also to show the different nature of the approach to equilibrium in the velocity and the space coordinates, it is good to look at the case in which the disturbance from equilibrium is small. Putting :

$$f = f_0 \{ 1 + h(\vec{r}, \vec{v}, t) \} \tag{8}$$

one gets the linear problem :

$$\frac{\partial h}{\partial t} = -\vec{v}\cdot\nabla_r h - \vec{a}\cdot\nabla h + \int d\vec{v}_1 \int d\omega\, g\, I(g,\theta) f_0 f_{01} (h' + h_1' - h - h_1)$$
$$= S(h) + C(h) \tag{9}$$

Consider first the case of no outside forces, and a disturbance dependent on v and t alone. $f_0 = A e^{-\frac{\beta m}{2} v^2}$; C(h) is an isotropic operator, in fact one can show that it has the form ;

$$C(h) = -\sigma(v) h + \int d\vec{v}_1 K(v, v_1, \cos\phi)\, h(\vec{v}_1)$$

As a result the eigenfunctions of C, defined by

$$C(\psi) = \lambda \psi, \tag{10}$$

must have the form :

$$\psi_{rlm} = f_{rl}(v)\, Y_{lm}(\theta, \phi),$$

using polar coordination in vel. space. The eigenvalues λ_{rl} are at least 21 + 1 fold degenerate. There are 5 zero eigenvalues λ_{00}, $\lambda_{1\phi}$, and λ_{10} corresponding to the eigenfunctions $1, v_x, v_z$ and $v^2 - \frac{3}{\beta}$, which are orthogonal. This is a consequence of the five conservation theorem. All other eigen values are negative. Proof :

$$\lambda_{rl} = \frac{\int d\vec{v} f_0 \psi_{rl} C(\psi_{rl})}{\int d\vec{v} f_0 \psi_{rl}^2} \tag{11}$$

Numerator is $\int d\vec{v}\, d\vec{v}_1 \int d\Omega\, g\, I(g,\theta)\, f_0 f_{01} \psi(\psi' + \psi_1' - \psi - \psi_1)^6 -$

$$-\frac{1}{4}\int d\vec{v}\int d\vec{v}_1 \int d\Omega\, g\, I(g,\theta)\, f_0 f_{01}(\psi' + \psi_1' - \psi - \psi_1)^2 < 0$$

using the same trick as used in the H-theorem. This shows that the disturbance decays monotonically to zero; the λ_{re} are the decay constants of the different modes. Developing :

$$h(\vec{v},t) = \sum a_{re}(t)\, \psi_{re}(t),$$

one gets

$$a_{re} = \lambda_{re}\, a_{re}$$

For Maxwell molecules, $q\,I(g,\theta)$ is independent of g, and since the transformation of the velocities in a collision is linear, it is clear that if h is a homogeneous polynomial in v_x, v_y, v_z, then C(h) must also be a homogeneous polynomial of the same degree. So it should be possible to find the eigen-polynomials, which turn out to be the Sonine polynomia in C^2, so :

$$\psi_{re} \sim S^{r}_{l+\frac{1}{2}}(c^2)\cdot c^l\, Y_{lm}(\theta\phi) \qquad (12)$$

$S^{(m)}(x)$ is the coefficient of Z^m in $(1-Z)^{-n-1}\, e^{-\frac{x}{1-z}\frac{z}{z}}$, and the corresponding eigenvalues are :

$$\lambda_{re} = 2\pi \int_0^{\pi} d\theta \sin\theta\, F(\theta) \Big[\cos^{2r+l}\frac{\theta}{2}\, P_l\Big(\cos\frac{\theta}{2}\Big) + \sin^{2r+l}\frac{\theta}{2}\, P_l\Big(\sin\frac{\theta}{2}\Big) - (1+\delta_{r0}\delta_{l0})\Big] \qquad (13)$$

They form a discrete set going to $-\infty$. Note that in (12) and (13) all velocities are measured in units of $\sqrt{2kT/m}$ and the collision operator is defined with the dimensionless $F(\ )$. As a result, the times are measured in units $\frac{1}{n}\sqrt{\frac{m}{2k}}$ and since the cross sections are now (x =kappa) $\sim \sqrt{\frac{x}{m}}\cdot\frac{1}{\sqrt{kT}} = \sqrt{\frac{x}{kT}}$, the time unit is $\frac{1}{n\sigma}\sqrt{\frac{m}{2kT}}$, which is of the order of the time between collisions. The first negative eigenvalue

$$\lambda_{20} = \lambda_{11} = \nu_1\sqrt{\frac{2k}{m}}\cdot\frac{1}{2}\int_0^{\pi} d\theta \sin^3\theta\, F(\theta) \qquad (14)$$

is the relaxation time of the gas already computed by Maxwell.

Whether for other force laws the eigen-value spectrum has the same character is not known but it is likely. However the approach of $\lambda_{\nu} \to \infty$ is probably due to the infinite range of the x/r^5 force and the corresponding infinite value of the total cross section. I expect, say for elastic spheres, a finite lower bound, but I can't prove it.

Turning now to the streaming operator S (h), clearly the eigen-values will in general not be real since S (h) is not self adjoint. In fact for the harmonic force one can write :

$$\left[\left(v_x \frac{\partial}{\partial(\omega x)} - \omega_x \frac{\partial}{\partial v_x} \right) + \ldots \right] h = \frac{\lambda}{\omega} h \qquad (15)$$

At the left are three rotation operators in the (v_x, x), (v_y, y) and (v_z, z) so clearly :

$$\lambda^{st.} = i\,(m_1, + m_2 + m_3)\, w \qquad (16)$$

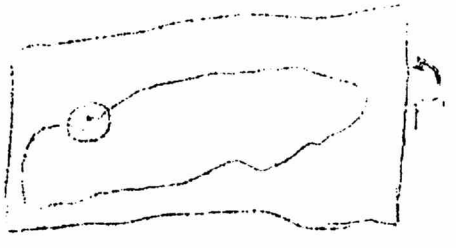
so that they are pure imaginary.

For Maxwell molecules it may be possible to determine the combined eigen functions and eigen-values of S (h) + C (h), since polynomials in r and ν remain polynomials in r and ν of the same degree.

The main qualitative point to remember is, that the approach in velocity and in space are quite different. The approach to the local Maxwell distribution is monotonic and quite fast (time of order of time between collisions) while in space the approach will usually have a damped periodic character (depending on outside forces) and goes slowly.

4. The criticism of the kinetic method and the answer of Boltzmann.

The criticism is directed toward the H-theorem, and is an expression of the feeling of discomfort that somehow, using reversible mechanical models, Boltzmann succeeded in showing that H decreases monotonically till the equilibrium state is reached, that is, a typical irreversible behavior. These criticisms have been crystallized in two famous objections, the Umkehr-eindwand of Loschmidt and the Wiederkehr-eindwand of Zermelo. Especially thelatter, based on theorem of Poincaré which says that a bounded mechanical system is always quasi-periodic, seems decisive. On the energy surface the point in Γ space represents the state of the gas will return after the Poincaré cycle to any surrounding of the initial positio The HH must therefore sometime increase.



Boltzmann's answer was not well understood in his time, probably because of his rather uncritical use of probabilistic terms. It amounts to this : note that a distribution in μ-space of a finite number of molecules required finite cells w_1, w_2.... The number in cell w_i

$$m_i = f_i \omega_i \longrightarrow f(r, p, t)\, d\vec{r}\, d\vec{p}$$

must be appreciable. With a distribution Z in μ-space corresponde in Γ-space a whole volume (3 N-dim.)

$$W(Z) = \frac{N!}{m_1!\, m_2! \ldots} \omega_1^{n_1} \omega_2^{n_2} \ldots \qquad (17)$$

Now one proves that with the accessory conditions

$$\sum n_i = N \qquad \sum \epsilon_i n_i = E$$

the M.B. distribution $\overline{n_i} = A \omega_i e^{-\beta \epsilon_i}$

overwhelmingly maximixes W(Z). If one therefore assumes that there is no priori preference for the motion of the Γ-point for any portion of the energy surface, then one can conclude that :

1. If the gas is not in the M.B. distribution it almost always will go into it.

2. Once it has the M.B. distribution it almost always will stay there.

Clearly in this way there is no conflict with the Poincaré theorem. The further development of this statistical method of Boltzmann leads to a completely satisfactory description of the equilibrium properti in the first place of gases, and then by considering more complicated systems as a large molecule to the prescription for the calculation of all thermodynamic properties from the partition function of the system.

However, the details of the approach to the equilibrium state are lost. The relation to the kinetic method is dark, and there is therefore no general method for complicated systems (say dense gaeas for instance) to derive the non-equilibrium properties from the molecular picture.

5. Excursion on stochastic processes.

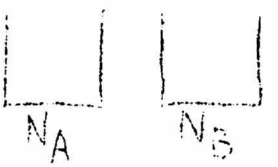
The whole situation has been further clarified by an example of Ehrenfest and by the work of Smoluchowski on the concentration fluctuation of colloidal suspensions.

The example of Ehrenfest : Consider two urns A nd B, containing

N_A and N_B balls respectively. Suppose

$$N_A + N_B = 2R$$

$$N_A - N_B = 2K$$



At fixed intervals a ball is chosen at random from among the $N_A + N_B$ and moved from its urn to the other. This is a kind of model for temperature or concentration equalization between two reservoirs. It is also interpretable as a random walk problem with attracting center. Most complete discussion :

M. Kac : Amer. Math. M. 54, 369, 1947.

Analogue to Boltzmann statistical method :

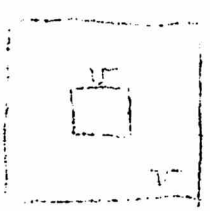
$$W(K) = \frac{2R!}{N_A! N_B!} \left(\frac{1}{2}\right)^{2R} \qquad (18)$$

Most probable state (= Equilibrium state) $N_A = N_B$ or $k = 0$ and this is very sharp if R is large. In fact, then :

$$W(K) = \frac{1}{\sqrt{\pi R}} e^{-\frac{K^2}{R}}$$

One can now go farther, because there is a "probability mechanism" which allows one to describe how k changes; the different k-values form a discrete random series.

Analogous is the situation in the concentration fluctuations :



Smoluchowski, Phys. ZS.. 17, 557-585 (1916)

$$W(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad \text{with } p = \frac{v}{V}$$

Equilibrium state $n = N\frac{v}{V}$. Transition by the "mechanism" of a jumping chance.

Clearly the k or n correspond to the distribution $(n_1\ n_2 \ldots)$ for the gas, and the behavior of k is just like the behavior according to Boltzmann of the state of the gas. Any value of k between - R to R will occur, but of course extremely rarely (and the Poincaré cycle is correspondingly long) if k is far from equilibrium. And there is no distinction between the future and the past. Irreversibility is a human illusion, or as Smoluchowski expresses it :

A process will be considered either irreversible or reversible according to whether the initial state has a recurrence time (Poincaré cycle) which is long or short compared to the time of observation. And

one should say in addition, that there is such a sharp variation of the length of the recurrence time (due to the large number of molecules) with the degree of deviation from equilibrium that the two alternatives are almost a discontinuous function of the deviation.

6. Discrete Markoff series.

Of course if one knows the "probability mechanism" one can go further, as I said before, and one can get some idea of how fast the equilibrium state is reached. The situation is especially simple for the so-called Markoff processes, of which the Ehrenfest model (but not the concentration fluctuation) is an example.

In general, one describes a random series k_1, k_2 ... by the set of distribution functions :

$W_1(k)$ = probability of finding k

$W_2(k_1, k_2; s)$ = probability of finding a pair k_1, k_2 s steps apart; etc.

A Markoff series is the next simplest to the purely random series (where all k's are independent of each other, so that $W_1(k)$ gives all the information). In a Markoff series all information is contained in the second distribution function which of course implies the first. One can also say that the probability of finding a value k depends only on the previous value of k. Introducing the conditional probability distribution P(n/m, s), the probability of finding m after s steps starting with n, one has for a Markoff series :

$$P(n/m, s) = \sum_K P(n/K, s-1)\, P(K/m, 1) \qquad (19)$$

Call P(k/m, 1) = Q(k, m); then since $\sum_m Q(k, m) = 1$, one can write (19) in the form (suppressing n) :

$$P(m, s) = \sum_K P(K, s-1)\, Q(K, m)$$

$$= \sum_K P(K, s-1)\, Q(K, m) + P(m, s-1)\left[1 - \sum_k Q(m, k)\right]$$

or

$$P(m, s) - P(m, s-1) = \sum_K P(K, s-1)\, Q(K, m) - P(m, s-1)\sum_k Q(m, k) \qquad (20)$$

This smells like the Boltzmann equation; the rate of change of P with "time" (= s) is equal to the gains due to transitions from k to m minus the losses due to transitions from m to any value different from m. However, note that the variable m corresponds in thegas case to the

set of numbers (n_1, n_2 ...). It is therefore really an equation on a higher level than the Boltzmann equation (which tells how each n_1 changes because of the collisions). Equation (20) really corresponds to a kind of "master equation" which would have to tell how the probability of a distribution (n_1, n_2...) changes with time. Note also that (20) is a linear equation. In equation 520) one has to suppose that Q(k, m) is knwon. It describes the basic "probability mechanism". One then has to solve Equation (20) for P(m, s) assuming an "initial" condition, say P(m, o) = (n, m).

For the Ehrenfest model in each step the k can only change with one, and

$$Q(K_0, K) = \frac{R + K_0}{2R} \delta(K_0 - 1, K) + \frac{R - K_0}{2R} \delta(K_0 + 1, K) \tag{21}$$

so that (19) becomes

$$P(m, s) = \frac{R + m + 1}{2R} P(m + 1, s - 1) + \frac{R - m + 1}{2R} P(m - 1, s - 1) \tag{22}$$

One proves easily that

$$\langle m(s) \rangle = \sum_m m P(m, s) = \left(1 - \frac{1}{R}\right) \langle m(s - 1) \rangle$$

so that if initially one is sure that m = n :

$$\langle m(s) \rangle = n \left(1 - \frac{1}{R}\right)^s$$

This exponential approach of the average value from n to the equilibrium value zero is the analogue of the monotonic approach of any velocity distribution to the Maxwell distribution. One can now of course also calculate the square average and show :

$$\langle m^2(s) \rangle = n^2 \left(1 - \frac{2}{R}\right)^s + \frac{1}{2} R \left[1 - \left(1 - \frac{2}{R}\right)^s\right]$$

so that for $s \to \infty$, $\langle m^2(s) \rangle \to \frac{R}{2}$ which is the value which follows from $W_1(m)$. In fact, one can show that any distribution P(m, s) approaches for $s \to \infty$ the first probability distribution $W_1(m)$ monotonically. In general, the proof goes as follows : from the basic equation :

$$P(n \mid m, s) = \sum_K P(n \mid K, s - 1) \quad Q(K, m)$$

one sees that one gets P(n / m, s) by successive matrix multiplication of

Q(k, m), which fulfills the conditions :

$$Q(k,m) \geq C \qquad \sum_m Q(k,m) = 1$$

It is obvious to introduce the eigenvalues of Q(k, m). Note, however, that Q(k, m) is <u>not</u> symmetric, so one must consider the linear equations:

$$\left.\begin{aligned} \lambda x_k &= \sum_s \lambda_s Q_{sk} \\ \text{together with the adjoint equations :} \\ \lambda y_h &= \sum_t Q_{th} + y_t \end{aligned}\right\} \qquad (24)$$

The secular determinant

$$\| \lambda_{ik} - Q_{ik} \| = 0 \qquad (25)$$

leads to the eigenvalue λ_i, and to λ_i corresponds

$$\lambda_i \rightarrow \begin{cases} x_{ik} & k = 1, 2 \ldots \\ y_{hi} & h = 1, 2 \ldots \end{cases}$$

and one has the two orthogonality and normalization relations :

$$\sum_k x_{ik}\, y_{kj} = \delta_{ij} \qquad \sum_s y_{is}\, x_{sj} = \delta_{ij}$$

By adding all columns to the first, one sees that $\lambda = 1$ is an eigenvalue; call it λ_1, then from (24) one verifies that

$$x_{1k} = W_1(K) \,; \; y_{h1} = 1$$

One can prove further that all other eigenvalues are in absolute value <u>less than one</u>; in fact, for the Ehrenfest model Kac has proved that

$\lambda_i = \frac{i}{R}$, $-R \leq i \leq R$, and he has determined also all the eigenvectors:

$$x_{ik} = C^{i}_{k+R} \qquad y_{hi} = \frac{(-1)^{R+h}}{2^{2R}} C^{-h}_{i+R}$$

where the symbol C_m^n is defined by :

$$(1 - Z)^{R-n} (1 + Z)^{R+n} = \sum_{m=0}^{2R} C_m^n Z^m$$

From the orthogonality relations follows then in the standard way :

$$Q_{nm} = \sum_j \lambda_j \gamma_{nj} \chi_{jm}$$

$$P(n \mid m, s) = \sum_j \lambda_j^s \gamma_{nj} \chi_{jm} \qquad (26)$$

$$= W_1(m) + \sum_{j>1} \lambda_j^s \gamma_{nj} \chi_{jm}$$

so that for $s \to \infty$ $\quad P(n \mid m, s) \to W_1(m)$.

7. The mean recurrence time.

One can give now also further details about the recurrence and persistence time. Basic paper is

Smoluchowski, Wiener Ber. 124, 339, 1915.

For Markoff processes everything can be done very elegantly. Define :

$P'(n \mid n, s)$ = probability, given n, that after s steps n occurs for the first time

Then clearly :

$$P(n \mid n, s) = P'(n \mid n, s) + \sum_{k=1}^{s-1} P'(n / n, k)\, P(n / n, s-k) \qquad (27)$$

Introduce now the generating functions :

$$h(n \mid n, z) = \sum_{s=1}^{\infty} P(n \mid n, s)\, z^s$$

$$g(n \mid n, z) = \sum P'(n \mid n, s)\, z^s$$

Then from (27) :

$$g(n \mid n, z) = h(n \mid n, z) = \sum_{s=1}^{\infty} \sum_{k=1}^{s-1} z^s P'(n \mid n, k) P(n \mid n, s-k)$$

In double sum, interchange and call s-k = r :

$$\sum_{s=1}^{\infty} \sum_{K=1}^{s-1} = \sum_{K=1}^{\infty} \sum_{s=K+1}^{\infty} = \sum_{K=1}^{\infty} \sum_{r=1}^{\infty}$$

so it becomes :

$$\sum_{K=1}^{\infty} P'(n/n, K) z^K \sum_{r=1}^{\infty} P(n/n, r) z^r = g(n/n, z) h(n/n, z)$$

and therefore :

$$g(n/n, z) = \frac{h(n/n, z)}{1 + h(n/n, z)} \qquad (28)$$

$$= 1 - \frac{1}{1 + h(n/n, z)}$$

Using now the form (26) for P(n n, z), then clearly :

$$1 + h(n/n, z) = \frac{W_1(n)}{1 - z} + p(z)$$

where p(z) is regular around z = 1, due to the fact that all eigenvalues λ are < 1, except λ_1. From (28) one thus sees that

$$\lim_{z \to 1} g(n/n, z) = \sum_{s=1}^{\infty} P'(n/n, s) = 1$$

P'(n/n, s) is therefore a bona fide probability. Further :

$$\frac{dg}{dz} = \frac{\frac{dh}{dz}}{(1+h)^2} = \frac{\frac{W_1(n)}{(1-z)^2} + p'(z)}{\left[\frac{W_1(n)}{1-z} + p(z)\right]^2}$$

so that :

$$\lim_{z \to 1} \frac{dg}{dz} = \frac{1}{W_1(n)}$$

This gives the mean recurrence time $\overline{T}(n)$, since

$$\overline{T}(n) = \tau \sum_{s=1}^{\infty} s\, P(n|n,s) = \tau \lim_{z \to 1} \frac{dg}{dz} = \frac{\tau}{W_1(n)} \qquad (29)$$

The mean square recurrence time depends already on all the eigenvalues and can therefore only be calculated for special models like the Ehrenfest model.

The concentration fluctuations of a colloidal suspensions as analysed by Smoluchowski do not form a Markoff process. However, the result (29) (with appropriate definition of $\overline{T}(n)$ is still valid. It is interesting to see how it compares with experiment (Svedberg). In these experiments τ = 1/33 min.; the average number $\langle n \rangle = \nu$ = 1.55; the first probability distribution $W_1(n) \cong \frac{\rho^{-\nu} \nu^n}{n!}$ (Poisson)

n	$\overline{T}_{obs}$	$\overline{T}_{calc}$
0	6.08	5.54
1	3.13	3.16
2	4.11	4.05
3	7.85	8.07
4	18.6	20.9

The check is therefore pretty good. The value n = 7 was observed only once; $\overline{T}(7)$ would be 1105 τ = 27 min., while the whole series of observations was about 600. For n = 17, $\overline{T}$ = 500,000 years!, so the transitions from n = 17 would certainly be judged to be irreversible.

The Ideas of Gibbs

8. The notion of an ensemble.

In the Ehrenfest example I pointed out already that the random variable k corresponds in the case of the ideal gas to the whole distribution $(n_1, n_2, \ldots)$ in μ-space. To describe the change of k in the Ehrenfest model we introduced the probability distribution P(k, t) of which we proved that it approached monotonically the "equilibrium" distribution $W_1(k) = \binom{2R}{R-k}\left(\frac{1}{2}\right)^{2R}$. In order therefore to describe in more detail how the Maxwell-Boltzmann distribution is reached for the gas, one must introduce the notion of the probability $P(n_1, n_2..t)$ of a certain state $(n_1, n_2\ldots)$; or, using the description of the state of the gas by a point Q in Γ-space, one must introduce a probability distribution $P(Q, t) = P(x_1 \ldots x_N, t)$ in Γ-space. (Notation : $x_i = (\vec{q}_i, \vec{p}_i)$ = coordinate and momentum of the i^{th} particle. This can only mean that one considers not only the given system but a large number of copies of it; an ensemble, in the language of Gibbs.

Similarly, in the Ehrenfest model, to test the result (say, of Z) for $P(k_o/k, t)$, one must consider an ensemble of experiments. In each the initial distribution is determined by k_o, but the successive values would be different and one would obtain Z sequences of k-values, from which at any time the distribution $P(k_o/k, t)$ can be found. Analogously, if one starts with a distribution $(n_1, n_2\ldots)$, corresponding to a region on the energy surface in the Γ-space of the gas, the different members of the ensemble (represented by points in this region) would move quite differently and provide different series of successive states; or, in the language of Gibbs, one has a streaming of the ensemble "fluid" over the energy surface. Of course there is also an essential difference; in the Ehrenfest model the "probability mechanism" is known, while in the case of the gas the motion of the point in Γ-space is in principle completely determined by the laws of mechanics and the differences in behavior are due to the different initial positions on which the motion depends very sensitively because of the sharp short-range interactions between the molecules.

In the Ehrenfest model $P(k_o/k, t)$ approaches $W_1(k)$. If one expresses the basic idea of the statistical method of Boltzmann for the gas, now, by saying that the a priori probability of some state of the gas is determined by the volume in Γ-space to which the state corresponds, then one can expect that the $P(n_1, n_2\ldots;t)$ will, for $t \to \infty$, monotonically approach

$$\frac{N!}{n_1!\,n_2!}\left(\frac{\omega_1}{\Omega}\right)^{n_1}\left(\frac{\omega_2}{\Omega}\right)^{n_2}$$

(where Ω is the volume of the whole phase (μ)space), so that the overwhelming majority of the members of the ensemble will have a Maxwell Boltzmann distribution. Or, in the language of Gibbs, one can expect

that any initial density distribution of the ensemble over the energy surface (which for an ideal gas is of course the hypersphere $\sum \frac{p_i^2}{2m} = E$ will approach the uniform distribution (i.e., uniform between the surfaces E and E + dE), or, as Gibbs call it, the micro-canonical distribution. That the volume in phase space, or the "statistical factor" in the problem of partition of the energy E between N independent particles leads for large N to the Maxwell Boltzmann distribution is perhaps not quite familiar (although given by Maxwell). The proof is as follows : since $p^2 dp \sim \sqrt{\varepsilon}\, d\varepsilon$, the probability that one particle has energy between ε_1 and $\varepsilon_1 + d\varepsilon_1$ is :

$$f(\varepsilon_1)\, d\varepsilon_1 \sim d\varepsilon_1 \sqrt{\varepsilon_1} \int \dots \int d\varepsilon_2 \dots d\varepsilon_N \sqrt{\varepsilon_2 \varepsilon_3 \dots \varepsilon_N}$$

where the integral is to be taken over all values so that $\varepsilon_2 + \varepsilon_3 + \dots \varepsilon_N = E - \varepsilon_1$. By induction or otherwise one gets (dropping index 1) :

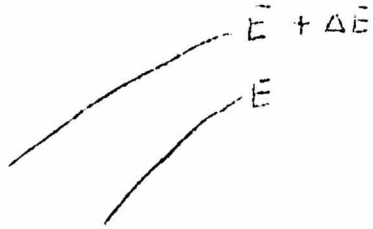
$$f(\varepsilon)\, d\varepsilon = A_N \sqrt{\varepsilon}\, (E - \varepsilon)^{\frac{3N-3}{2}} d\varepsilon$$

where A_N is a normalization factor. This gives $\bar{\varepsilon} = E/N$ (Maxwell's proof of equipartition), and it is easy to prove that for large N :

$$f(\varepsilon)\, d\varepsilon = A \sqrt{\varepsilon}\, e^{-\frac{3\varepsilon}{2\bar{\varepsilon}}} d\varepsilon$$

the Maxwell-Boltzmann distribution.

Of course one now can generalize to more complicated systems than ideal gases, and in fact Gibbs introduced the microcanonical ensemble (now better defined as the uniform distribution between two neighboring energy surfaces E and $E + \Delta E$ in the Γ -space :



- note that these surfaces need not be parallel) immediately for general systems. It was one of his aims to give a more deductive and lucid presentation of the methods of Boltzmann, and in this he certainly succeeded! I will not discuss the further developments, such as the other special ensembles (canonical, grand ensemble, the string of theorems connecting them with each other, the thermodynamic analogies etc.). They lead again to the basic relation between the free energy of

the system and the partition function , which gives in principle the solution of all the problems of statistical physics (connection between macroscopic properties and the molecular structure and interactions) for systems in thermodynamic equilibrium.

It should be emphasized that, while in the Ehrenfest model the equal a priori probability of the two boxes (from which $W_1(k)$ follows) is put into the problem, for the gas the assumption of Boltzmann must be considered to be a property of the mechanical system. One calls it the ergodic property. What it implies for the mechanical motion can be explained as follows : in the Ehrenfest model, clearly the different series of k-values of the ensemble can also be considered as successive large chunks of a very long single series in time; therefore for any quantity the time average will be the same as the ensemble average, and the same must be assumed for the gas. That is,

$$\langle f(x_1 \dots x_N) \rangle_{\text{time}} = \lim_{T \to \infty} \frac{1}{T} \int_0^T f[x_1(t) \dots x_N(t)]\, dt$$

$$= \int \dots \int dx_1 \dots dx_N \, f \, \sigma(E) = \langle f(x_1 \dots x_N) \rangle_{\text{ensemble}}$$

where $\sigma(E)$ is the microcanonical probability distribution. This property, although very plausible because of the erratic motion of the Γ-point, is of course not proved. It can be reduced to the assumption of metrical transitivity (from which the ergodic property then follows) but although an advance, it does not help the physicist!

9. The approach to the microcanonical distribution and the Gibbs H-theorem.

See Gibbs, Chapter 12.

As I said, it is very plausible that any distribution of the ensemble "fluid" between the nergy surfaces E and $E + \Delta E$, will monotonically become the uniform or microcanonical distribution. It is appropriate here to make some comments on the question of the approach.

a) First, the monotonic approach is not in conflict with the Poincaré theorem. Each member of the ensemble has a quasi periodic motion, and one can prove (see M. Kac, Bull. Amer. Math. Soc. 53, 1002, 1947) that for intermittent observation and for a metrically transitive mechanical system the average recurrence time (or Poincaré cycle) is $\sim \frac{\Omega}{\omega}$ (where ω is the area on the energy surface), quite analogous to the result $\overline{T} = W_1(n)$ for the Ehrenfest model.

b) It is not true that the uniform distribution is the only stationary distribution. To find the most general stationary distribution I have to remind you of a general characteristic of the streaming, expressed by the Liouville theorem : the streaming of the "ensemble fluid" is like that of an incompressible fluid. The proof is so simple that I shall give it :

Proof : The "velocity" $\vec{V}$ (with components $\dot{q}_1 \dots \dot{p}_N$) fulfills the condition :

$$\mathrm{DIV}\,\vec{V} = \sum_i \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = \sum_i \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} \right) \equiv 0$$

by virtue of Hamilton's equations. Since one always has the continuity equation :

$$\frac{D\rho}{Dt} + \rho \, \mathrm{div}\, \vec{V} = 0$$

one has therefore :

$$\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + \sum_i \left(\dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i} \right) = 0$$

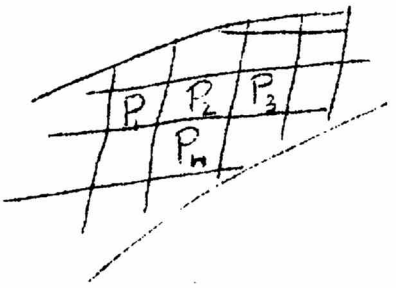
i.e. :

$$\frac{\partial \rho}{\partial t} = \{ H, \rho \} \qquad (1)$$

(the Poisson bracket of H and).

The most general stationary density distribution is therefore a distribution where ρ is constant along stream lines (lines of constant H), but not necessarily the same constant along different stream lines.

c) Since ρ remains constant if one moves with the fluid, the distribution can only become uniform in a "coarse-grained" sense, as pointed out by Ehrenfest : If the density initially is constant in different regions (say P_1, $P_2 \dots$), then because of the distortion due to the streaming the average density



$$\bar{\rho} = \frac{\int \dots \int \rho \, dx_1 \dots dx_N}{\Delta x_1 \dots \Delta x_N}$$

over the fixed regions will become equal. This can be shown analytically

by the Gibbs H-theorem.

We define this H-function by :

$$H_G \equiv \sum_m P_m \log P_m (\Delta x_1 \ldots \Delta x_N)_m = \int \ldots \int dx_1 \ldots dx_N P \log P$$

$$= \int \ldots \int dx_1 \ldots dx_N \, \rho \log \rho \qquad (2)$$

$$\delta \tau_m = (\Delta x_1 \ldots \Delta x_N)_m$$

is the volume of the m^{th} cell.

Now start with a non-uniform ensemble, in which the density is actually constant in the set of regions $\delta \tau_1$, $\delta \tau_2$ ... So at t = 0, the fine-grained and coarse-grained densities are identical and

$$H = H_0 = \int \rho_0 \log \rho_0 \, dx_1 \ldots dx_N$$

Now consider at a later time t :

$$H - H_t = \int \ldots \int (\rho_0 \log \rho_0 - \rho_t \log \rho_t) \, dx_1 \ldots dx_N$$

$$= \int \ldots \int (\rho_t \log \rho_t - \rho_t \log \rho_t - \rho_t + \rho_t) \, dx_1 \ldots dx_N ,$$

where we have used $\int \ldots \int \rho_t \, dx_1 \ldots dx_N = \int \ldots \int P_t \, dx_1 \ldots dx_N$

$$= 1$$

by normalization and :

$$\int \ldots \int \rho_t \log \rho_t \, dx_1 \ldots dx_N = \int \ldots \int \rho_0 \log \rho_0 \, dx_1 \ldots dx_N$$

because of Liouville.

Using the lemma that F (x,y= = x log x - x log y - x + y $\geq$ 0 (zero only for x = y), one sees that

$$H_0 - H_t \geq 0$$

and that H_G is smallest for the coarsely uniform ensemble.

10. - Criticism of the Gibbs Theorem.

Ehrenfest[1] expressed the opinion that the time for the approach to the uniform distribution may well be several Poincaré cycles. If so, it clearly would have nothing to do with the actual approach to equilibrium. I must confess that this is one of the few points I do not understand in the Encyclopedia article. I think the time of approach has nothing to do with the Poincaré cycle. It will, rather, depend on the size of the coarse-grained cells; the larger they are chosen the sooner they will become roughly uniform.

In my opinion the important objection is that the theorem has little to do with real irreversible processes since it does not tell which coarse-grained density "corresponds to our knowledge" of a certain initial non-equilibrium situation. Of course this is connected with the difficulty of using an ensemble for the description of a single system. The justification must lie in the fact that a macroscopic description of a system is so rough that a very large number of microscopically different states of the system correspond to it. One must then assume that the actually observed macroscopic changes of state are the same for the overwhelming majority of the ensemble, so that the average of some quantity over the (non-stationary) ensemble as a function of time will represent the change of the quantity with time for the single system.

The trouble is, not only that such statements are hard to prove, but there is no hint or guess on how to construct such non-equilibrium ensembles. As a result there is no general theory for non-equilibrium phenomena in Gibb's scheme. Of course, in view of the enormous variety of characteristic times involved (especially if solids are present), such a general theory is probably too much to ask for. However it should be possible, say for dense gases, to show how in principle typical transport quantities (like for instance, the heat conductivity) are related to the intermolecular forces.

11. - "Rückblick" on the kinetic method of Boltzmann.[2]

In particular for nearly ideal gases one would like to know how the ..."Stoszzahl Ansatz" and the Boltzmann equation discussed in the first lecture fit into the general scheme of Gibbs.

A rather obvious idea is to interpret the Stoszzahl Ansatz as an assumption for the transition probability between the different states (n_1, n_2) of the ideal gas. I proposed to do this in some analogous statistical problem in the theory of cosmic ray showers (Physica, 7, 344, 1940; Phys. Rev. 62, 497, 1942), and it was then taken up by Siegert (Phys. Rev. 76, 1708, 1949) and recently much further developed by M. Kac.

1. Enziclopedia, P.61
2. Footnote on next page.

Let n_1, n_2... be the number of molecules in the different velocity-cells ω_1, ω_2... Assume the distribution in space to be uniform, and no outside forces. Let the probability per unit time of a collision $(i, j) \rightarrow (k, l)$ be : $a_{ij}^{kl} n_i n_j$; then it follows that the probability distribution $P(n_1, n_2 \ldots ; t)$, which, recall, is analogous to $P(k, t)$ in the Ehrenfest model, will satisfy the "master equation" : (Analogous to the Smoluchowski equation for the Ehrenfest case (20) p.10)

$$\frac{\partial P(n_1, n_2 \ldots, t)}{\partial t} = \frac{1}{2} \sum_{(ij)(kl)} \left[a_{kl}^{ij} (n_k + 1)(n_l + 1) P(n_1, \ldots, n_k + 1, n_l + 1, \ldots \right.$$

$$\left. \ldots n_i - 1, n_j - 1, \ldots, t) - a_{ij}^{kl} n_i n_j P(n_1, \ldots ; t) \right] \qquad (1)$$

Note that this is a linear equation that is given by the Stosszahl Ansatz assumption. By summing over all values of n_1, n_2... with the condition $\sum n_i = N$, and assuming microscopic reversibility : $a_{ij}^{kl} = a_{kl}^{ij}$, and noting the symmetry in (ij) et (k l); furthermore $a_{ij}^{kl} = 0$ and in (1) i, j $\neq$ k, l, one sees that

$$\frac{d}{dt} S' P(n_1, n_2, \ldots, t) = 0 \qquad (2)$$

(where S' means that one is to sum over all n_1, n_2 ... subject to $\sum n_i = N$) so that the normalization S'P $(n_1, n_2 \ldots ; t = 1$ is conserved. For the rate of change of the average number $\bar{n}_s = S' n_s P(n_1, n_2 \ldots ; t)$ in cell ω_1 one then finds :

$$\frac{d\bar{n}_s}{dt} = \frac{1}{2} \sum_{ijkl} a_{ij}^{kl} S' \left[n_s (n_k + 1)(n_l + 1) P(\ldots, n_k + 1, n_l + 1, \ldots, n_i - 1, n_j - 1, \ldots \right.$$

$$\left. - n_s n_i n_j P(n_1, n_2, \ldots) \right]$$

One must distinguish in the "straight" sum the cases where s is different from i, j, k, l (which are all different from each other) and where it is equal to any one of the i, j, k, l. Changing the summations in the first "round" sum back to n_k, n_l, n_i, n_j, one then easily finds that the third

---

[2] (Foot-note from page 21). This constrats strongly with the kinetic method of Boltzmann, where the approach to equilibrium can be found in detail, since the process is probalistic, rather than dynamical. The natural question then is : Can we "fix up" Boltzmann's method now that we have more insight into the nature of the difficulties? Gibbs, in his famous chapter 12, tells us nothing on this point.

order terms in the n's always cancel, and one gets :

$$\frac{dn_i}{dt} = \sum_{j,k,l} \alpha_{ij}^{kl} \left(n_k n_l - n_i n_j \right) \qquad (3)$$

This would be just the Boltzmann equation if one could put $\overline{n_k n_l} = \bar{n}_k \bar{n}_l$ and $n_i n_j = n_i n_j$ that is, if one could neglect the correlations, which presumably would be permissible if all the occupation numbers are large.

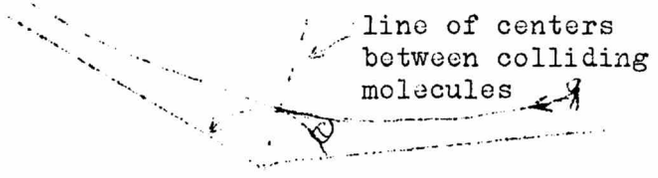
Siegert has shown that (1) leads, for t , to the equilibrium distribution $\frac{N!}{n_1! \, n_2! \ldots} \left(\frac{\quad}{\quad} \right)^{n_i}$, just as we saw in the Ehrenfest example that P(k, t) for t $\longrightarrow$ approaches the first probability distribution $W_1(k)$ which is proportional to $N!/N_A! \, N_B!$

12. - Formulation of the master equation on the energy surface.

It is nicer and more concrete to express the same idea in the "Gibbs" way. For the ideal gas, the energy surface is a 3N-1 dimensional hyperspherical surface imbedded in 3N-dimensional space. Assume that on the energy surface the -point has a random walk motion where the probability of a step in time dt is again determined by the "Stoszzahl Ansatz". Since momentum is also conserved, one can put = 0, so the motion is really on a 3N-3 dimensional sphere (again only the velocity distribution will be considered).

Let (R, t) be the probability that the -point is at R at time t, where R ≡ ( ).

Consider a collision between molecules i and j. To represent it, we use a "collision operator" A. A collision is rotation of the vector define $A_{ij}(\vec{S}) \vec{R} = (\vec{r}_1, \ldots, \vec{r}_i - \vec{g}_{ij} \cdot \vec{S}) \vec{S}, \ldots, \vec{v}_N)$



as the rotated vector; $\vec{g}_{ij} = \vec{v}_i - \vec{v}_j$

$\vec{S}$ unit vector in the perihelium direction. If total volume is V, then according to the Stoszzahl Ansatz, the problem per second of a specific collision is $\frac{1}{V} I(g\theta) \, d\Omega$ and one has therefore

$$\frac{\partial \phi(\vec{R}, t)}{\partial t} = \frac{1}{2V} \sum_{ij} \int d\Omega \, g_{ij} I(g_{ij}, \theta) \left\{ \phi A_{ij} \vec{R}, t) - \phi(\vec{R}, t) \right\} \qquad (4)$$

Note again that this is a linear equation. Furthermore if the N particles are indistinguishable one should only allow symmetric functions $\phi(v_1 \ldots v_N, t)$. By contraction (= integration over all v's except, say, $v_1 \ldots v_k$, where k is arbitrary) one gets the partial velocity distribution of groups of particles. In particular for

$$f^{(N)}(v_1, t) = \int \cdots \int dv_2 \cdots dv_N \, \phi(\vec{R}, t)$$

(in the integral, energy and momentum conditions must of course be taken into account) one gets the Boltzmann-like equation :

$$\frac{\partial f_1^{(N)}(\vec{v}_1, t)}{\partial t} = \frac{N-1}{V} \int d\vec{v}_2 \int d\Omega \, g \, I(g, \theta)$$

$$\times \left\{ f_2^{(N)}(\vec{v}_1 - (\vec{g}\cdot\vec{s})\vec{s}, \vec{v}_2 + (\vec{g}\cdot\vec{s})\vec{s}) - f_2^{(N)}(\vec{v}_1, \vec{v}_2, t) \right\} \qquad (5)$$

(We have chosen the subscript 2 for the collision "partner" of 1; of course it is arbitrary). In the proof one must distinguish again the cases i, j $\neq$ 1 which give no contribution, and i = 1, j $\neq$ 1; i $\neq$ 1, j = 1, which give the two terms in (5). Eq. (5) would become the Boltzmann equation if one could assume

$$f_2^{(N)}(v_1, v_2) = f^{(N)}(v_1) \, f^{(N)}(v_2)$$

and if one could then introduce the average number of particles $f(\vec{v}, t) = \frac{N}{V} f_1^{(N)}(\vec{v}, t)$ instead of $f_1^{(N)}$.

Again it can be expected that for large t the $\phi(\vec{R}, t)$ becomes the uniform distribution (microcanonical ensemble), which as we saw, leads for large N to the M.B. distribution for $f_1(v)$.

Marc Kac has been able to go further. To do this, he considered a simplified version of the basic equation (4); it is a kind of one-dimensional Maxwell model. Let $\vec{R} = x_1 x_2 \ldots, x_N$ and

$$A_{ij}(\theta) R = x_1, x_2, \ldots x_{i-1}, x_i \cos\theta + x_j \sin\theta, \ldots, -x_i \sin\theta + x_j \cos\theta, \ldots, x_N$$

$$= e^{i\theta L_{ij}} R$$

where L_{ij} is the infinitesimal rotation operator in the ij plane. Assume that the probability for such a rotation depends only on θ . The random walk is executed on the sphere $x_1^2 + x_2^2 \ldots + x_N^2 = N$, and the master equation then is :

$$\frac{\partial \phi(\vec{R}, t)}{\partial t} = \frac{1}{2N} \sum_{i,j} \int_{-\pi}^{+\pi} \gamma(\theta) \, d\theta \left\{ \phi(A_{ij}(\theta)\vec{R}, t) - \phi(\vec{R}, t) \right\} \qquad (4a)$$

We can write (4a) as $\partial_t \phi = \Omega \phi$,

where

$$\Omega = \frac{v}{2N} \sum_{ij} \left(e^{\partial L_{ij}} - 1 \right).$$

The Boltzmann equation analogous to (4a) is :

$$\frac{\partial f(x,t)}{\partial t} = v \int_{-\infty}^{+\infty} dy \int_{-\pi}^{+\pi} F(\theta) d\theta \{ f(x \cos\theta + y \sin\theta, t) f(-x \sin\theta + y \cos\theta, t) - f(x,t) f(y,t) \}$$

The "collision cross section" will be $\int_{-\pi}^{+\pi} F(\theta) d\theta$, and we will assume it to be <u>finite</u>, and it is also appropriate to assume $F(\theta)$ an even function of θ .

FURTHER RESULTS OF KAC; THE BASIC QUESTION

13. - The Propagation of the "Boltzmann property".

Let $\Phi_N(\vec{R},0)$ be a sequence of probability density functions, defined on the sphere $\sum_{i=1}^{N} x_i^2 = N$, which have the "Boltzmann property"

$$\lim_{N\to\infty} f_K^{(N)}(x_1 \dots x_K, 0) = \prod_{j=1}^{K} \lim_{N\to\infty} f_1^{(N)}(x_j, 0). \qquad (1)$$

Then Kac was able to prove that the sequence of functions $\Phi_N(\vec{R},t)$ which are the solutions of the master equations with $\Phi_N(\vec{R},0)$ as initial values, also have the Boltzmann property, i.e. :

$$\lim_{N\to\infty} f_K^{(N)}(x_1 \dots x_K, t) = \prod_{j=1}^{K} \lim_{N\to\infty} f_1^{(N)}(x_j, t). \qquad (1\,a)$$

This theorem, of which the proof is still rather complicated, is of interest because it elucidates what previously was hidden under the assumption of molecular chaos. Clearly $\Phi_N(\vec{R},t)$ contains in general much more information than $f_1(x,t)$. However one must expect that all correlations are "broken" very quickly, and the state of molecular chaos (for which the Boltzmann property holds) is established and then persists (in time). The non-linearity of the Boltzmann equation is from this point of view, due only to the special initial distribution; the basic problem (= the initial value problem for themaster equation) is a linear problem.

14. - The approach to the uniform distribution.

It is again very plausible that for $t \to \infty$, $\Phi(R,t)$ will go over into the microcanonical distribution

$$\Phi_0(R) = \frac{1}{S_N(\sqrt{N})} \qquad (2)$$

($S_N(r)$ = surface area of N-dimensional sphere of radius r), which leads for the one-particle distribution to :

$$f_1^{(N)}(x) \sim \left(1 - \frac{x^2}{N}\right)^{\frac{N-3}{2}} \quad \text{as } t \to \infty$$

which in turm gives the M.B. distribution for N $\longrightarrow$ 00

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} . \tag{3}$$

The approach to (2) will be monotonic as can be seen by an analogue of the H-theorem. In fact one does not need the function

$$H(t) = \int \cdots \int \phi \log \phi \, dS : \tag{4}$$

one can show the monotonic decrease already for the simpler quantity

$$K(t) \equiv \int \cdots \int \phi^2 dS \tag{5}$$

where dS = dx_1 .... dx_2 = element of surface area (on the sphere).

In fact

$$\frac{dK}{dt} = \frac{v}{N} \int_{-\pi}^{+\pi} d\theta F(\theta) \sum_{ij} \int \cdots \int dS \, \phi(R,t) \{ \phi(A_{ij}(\theta)R,t) - \phi(R,t) \}$$

Put:

$$A_{ij}(\theta) R = R' \quad , \qquad R = A_{ij}(-\theta) R' .$$

Use R' and $\theta' = -\theta$ as new variables, and add the new result to the original form and divide by two; one gets then :

$$\frac{dK}{dt} = -\frac{v}{2N} \int_{-\pi}^{+\pi} d\theta F(\theta) \sum_{ij} \int \cdots \int dS \{ \phi(A_{ij} R,t) - \phi(R,t) \}^2$$

$$\Rightarrow \qquad \frac{dK}{dt} \leq 0, \tag{6}$$

and the equality sign is only obtained for the uniform distribution.

The proof for H(t) goes exactly the same way; in fact any function of ϕ which is concave upward would do. The exceptional property of the H-function is that with the Boltzmann property, clearly

$$\int \cdots \int dS \; \phi \log \phi \longrightarrow \text{const.} + N \int_{-\infty}^{+\infty} dx \; f \log f \qquad (7)$$

which is what one expects for the entropy of an ideal gas, and which sh ows the connection between the Gibbs and the Boltzmann H-functions.

It still needs further argument to show rigorously that "any" $\phi(R,t)$ goes over into the uniform distribution as $t \to \infty$. Kac was able to do this in the following sense : he proved that for every integrable function $\chi(\vec{R})$ on the sphere

$$\lim_{t \to \infty} \int \phi(\vec{R},t)\,\chi(\vec{R})\,dS = \frac{\int \chi(R)\,dS}{S_N(\sqrt{N})}$$

15. - The relaxation time.

It would be of interest to know the spectral decomposition of the operator Ω in the master equation,

$$\frac{\partial \phi}{\partial t} = \Omega \phi,$$

which can be written as :

$$\Omega = \frac{1}{2N} \int_{-\pi}^{+\pi} d\theta \, F(\theta) \sum_{ij} \left(e^{i\theta L_{ij}} - 1 \right), \qquad (8)$$

where $L_{ij} = \frac{1}{i}\left(x_i \frac{\partial}{\partial x_j} - x_j \frac{\partial}{\partial x_i}\right)$ is the infinitesimal rotation operator. The question of the spectrum of Ω has not been quite settled yet. Clerly zero is an eigenvalue, with eigen-function ϕ = const., and the other eigen functions must be parical harmonics on the N-dimensional sphere. Let $H_K(x_1 \ldots x_N)$ be the solid harmonics of degree K,

which are in addition symmetric in $x_1, x_2, \ldots x_N$. There may be sevreal of them which are linearly independent, which is indicated by the index . Clearly will be a linear combination of such spherical harmonics of the same degree, so that for each k one has to solve a finite secular problem to find the spectrum. This seems feasible, but has not been accomplished yet. One can argue that if functions (R) which fulfill the Boltzmann property are a sufficiently complete set of functions, then the spectrum of should bethe same as the spectrum of the Boltzmann non-linear collision operator. This spectrum in turn should be, for the first eigenvalues (which determine the last stages of the approach to equilibrium when the deviation from equilibrium has already become small), identical with the spectrum of the <u>linearized Boltzmann</u> equation, obtained by putting

$$f(x,t) = \frac{1}{\sqrt{\pi}} e^{-x^2/2} \left\{ 1 + h(x,t) \right\}$$

and neglecting quadratic terms in h. This gives :

$$\frac{\partial h}{\partial t} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dy\, e^{-y^2/2} \int_{-\pi}^{\pi} d\theta\, F(\theta) \left\{ h(x\cos\theta + y\sin\theta, t) + h(-x\sin\theta + y\cos\theta, t) - h(x) - h(y) \right\} \equiv B_1(h)$$

(9)

This is the linearized form of the Boltzmann equation

$$\frac{\partial f}{\partial t} = B(h) = \nu \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\, F(\theta) \left\{ f(x',t) f(y',t) - f(x) f(y) \right\}.$$

It is easy to prove that the eigen-valu of $B_1(h)$ are the Hermite polynomials

$$h_k(x) = H_k\left(\frac{x}{\sqrt{2}}\right) = k! \sum_{m} \frac{(-)^m (2x)^{k-2m}}{m!\,(k-2m)!}$$

(defined with e^{-x^2} as weight function for $H_k(x)$).

Using the addition theorem :

$$H_k(a\cos\theta + b\sin\theta) = \sum_{l=0}^{k} \binom{k}{l} \cos^l\theta \sin^{k-l}\theta\, H_k(a) H_{k-l}(b),$$

one easily finds for the spectrum :

$$\lambda_k = -\int_{-\pi}^{+\pi} d\theta\, F(\theta) \left\{ 1 - \delta_{k0} + \cos^k\theta + (-1)^k \sin^k\theta \right\}. \tag{10}$$

The eigenvalue zero is doubly degenerate (k = 0, k = 2) corresponding to the sonservation of number and energy. All other eigenvalues are negative, and have the value $-\int_{-\pi}^{+\pi} d\theta\, F(\theta)$ as limit point. For instance with $F(\theta) = 1$ for $-\pi/2 < \theta < \pi/2$ zero otherwise, one gets :

$$\lambda_0 = \lambda_1 = 0, \qquad \lambda_{2n} = -\pi \left[1 - \frac{2}{\sqrt{\pi}} \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \right]$$

$$\lambda_{2n+1} = -\pi \left[1 - \frac{1}{\sqrt{\pi}} \frac{\Gamma(n+1)}{\Gamma(n+\frac{1}{2})} \right].$$

The even and odd λ's form monotonic decreasing sequences with $-\pi$ as limit point. The eigenvalue nearest to zero is $\lambda_2 = -\pi/4$ which is therefore the analogue of the inverse of the relaxation time.

Larger disturbances form equilibrium will die down with exponential factors which will consist of linear combinations of the eigenvalues λ_k . This may well produce an almost continuous spectrum beyond a value which must be bigger than $2\lambda_2$. A few discrete values, and especially λ_2 itself, will clearly remain.

16. - Final Remarks.

The master equation approach shows conclusively, in my opinion, that the kinetic method of Boltzmann, properly interpetated, is in harmony with the statistical method and with the ideas of Gibbs. It leads to interesting mathematical problems, which still need a lot of work. However, there remains the basic question :

What is the relation between the master equation $\dot{\phi} = \Omega\phi$ and the

basic dynamical equation $\partial_t \rho = \{H, \rho\}$ (Liouville theorem), which determines the streaming over the energy surface ?

In other words : how can one "derive" the probability assumptions form the special features which caracterizes the streaming under the influence of the short range intermolecular forces? Clearly what the master equation implies, is that with sufficient dilution, one can omit the intermolecular forces in H, and replace the streaming by the "jumping" with a probability determined by the Stozzahl Ansa tz, i.e. the "master equation" approach is, strictly, applicable only to an ideal gas.

Because the connection with the dynamics is <u>not</u> established, the master equation approach does not allow one to answer the questions of physical interest :

1) What are the limitations of the kinetic method, especially with regard to rapidly varying phenomena (high frequency sound, for instance)?

2) How should one extend the kinetic method so as to take triple and higher multiple collisions into account, (i.e.) to dense systems)?

The Hydrodynamical Equations

17. - The general conservation laws.

Going back to the Boltzmann equation with outside forces :

$$\frac{\partial f}{\partial t} + \xi_\alpha \frac{\partial f}{\partial x_\alpha} + X_\alpha \frac{\partial f}{\partial \xi_\alpha} = \int d\vec{\xi}_1 \int d\Omega\, q\, I(q,\theta)(f' f_1' - f f_1) \qquad (1)$$

I would like to outline the further development (due to Maxwell-Chapman, Lorentz-Hilbert-Enskog) partially because it still is the only rigorous theory for a class of non-equilibrium phenomena, and also in order to point out the peculiar features of the approximation method.

The first step is to derive from (1) the so-called transport equation for a quantity $\phi(\xi, \eta, \zeta)$ which a molecule can carry "on its back". Defining :

$$\bar{\phi}(x, y, z, t) \equiv \frac{\int \phi f \, d\vec{\xi}}{\int f \, d\vec{\xi}} = \frac{1}{n} \int \phi f \, d\vec{\xi},$$

where n(x, y, z; t) = number density ($n\,m = \rho$), one gets :

$$\frac{\partial}{\partial t}(n\bar{\phi}) + \frac{\partial}{\partial x_\alpha}(n \bar{\xi}_\alpha \bar{\phi}) - n X_\alpha \frac{\overline{\partial \phi}}{\partial \xi_\alpha}$$

$$= \frac{1}{4} \int d\vec{\xi} \int d\vec{\xi}_1 \int d\Omega\, q\, I(q,\theta)(\phi + \phi_1 - \phi' - \phi_1')(f' f_1' - f f_1). \qquad (2)$$

For ϕ equal to any of the five quantities $m, m\xi, m\eta, m\zeta, \frac{m}{2}(\xi^2 + \eta^2 + \zeta^2)$ which are conserved in a collision, the right-hand side of (2) vanishes, and one gets the five general hydrodynamical equations :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{u}) = 0$$

$$\rho \frac{Du_\alpha}{Dt} = \rho \left(\frac{\partial u_\alpha}{\partial t} + \vec{u} \cdot \operatorname{grad} u_\alpha \right) = \rho X_\alpha - \frac{\partial P_{\alpha\beta}}{\partial x_\beta}$$

$$\rho \frac{D}{Dt}\left(\frac{Q}{\rho}\right) + \operatorname{div} \vec{q} = -P_{\alpha\beta} D_{\alpha\beta} \qquad (3)$$

where

$x_i = x, y, z$: coordinates

$\xi_i = \xi, \eta, \zeta$: velocity components

$u_i = u, v, w$: average velocity

$u_i = \overline{\xi_i}$

$U_i = U, V, W$ = thermal velocity

$= \xi_i - u_i$

$X_i = X, Y, Z$ = outside force (per unit mass)

$Q = \frac{1}{2} \rho \overline{U^2}$ = thermal energy density

$P_{ij} = \rho \overline{U_i U_j}$ = stress tensor

$q_i = \frac{1}{2} \rho \overline{U_i U^2}$ = heat current density

$D_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$ = rate of strain (deformation) tensor

and, as indicated,

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u_\alpha \frac{\partial}{\partial x_\alpha}$$

The equations (3) express the conservation laws of number, impulse and energy, and are still an empty frame since we do not know yet how to express P_{ij}, q_i in terms of the average (flow) velocity, the density, andthe temperature. In other words, one still has to derive the Newton and Fourier phenomenological laws for friction and heat conduction.

18. - The Hilbert-Enskog development (1).

To do this one has to solve the Boltzmann equation. Hilbert introduces a formal parmater θ, and writes :

$$f = \frac{1}{\theta} f^{(0)} + f^{(1)} + \theta f^{(2)} + \ldots$$

In addition (this is Enskog's contribution) it is necessary to decompose

(1) Chapman and Cowling, "The Mathematical Theory of Non-uniform Gases".

the time

the time derivative in parts of successive order of approximation by writing :

$$\frac{\partial}{\partial t} = \frac{\partial_0}{\partial t} + \theta \frac{\partial_1}{\partial t} + \cdots$$

Calling $D \equiv \xi_\alpha \frac{\partial}{\partial x_\alpha} + X_\alpha \frac{\partial}{\partial \xi_\alpha}$, and the collision operator $J(f f_1)$, one gets from (1), by equating equal powers of θ,

$$J(f^{(0)} f_1^{(0)}) = 0 \tag{4}$$

$$\left(\frac{\partial_0}{\partial t} + D\right) f^{(0)} = J(f^{(0)} f_1^{(0)} + f^{(1)} f_1^{(0)}) \tag{5}$$

$$\frac{\partial_1 f^{(0)}}{\partial t} + \left(\frac{\partial_0}{\partial t} + D\right) f^{(1)} = J(f^{(0)} f_1^{(2)} + f^{(1)} f_1^{(1)} + f^{(2)} f_1^{(0)}) \tag{6}$$

etc.

a) First approximation.

The development is clearly so arranged that (4) is the first approximation, of which the general solution is :

$$f^{(0)} = n \left(\frac{m}{2\pi kT}\right)^{3/2} e^{-\frac{m}{2kT}\left[(\xi - u)^2 + (\eta - v)^2 + (\zeta - w)^2\right]} \tag{7}$$

the local Maxwell distribution, which contains still the five macroscopic quantities n, u_i, T which are functions of x, y, z; t and remain undetermined.

b) Second approximation.

Writing $f^{(1)} = f^{(0)} \phi^{(1)}$ one easily sees that (5) becomes :

$$\left(\frac{\partial_0}{\partial t} + D\right) f^{(0)} = -I(\phi^{(1)}) \tag{5a}$$

where $I(\varphi)$ is the linear isotropic operator

$$I(\phi) = \int d\vec{\xi}_1 \int d\Omega \, I(g, \theta)(\phi + \phi_1 - \phi' - \phi_1') f^{(0)} f_1^{(0)}.$$

Equation (5a) is an inhomogeneous integral equation for $\phi^{(1)}$. The homogeneous part has clearly as solutions the five conserved quantities $\psi_1 = 1$, $\psi_2 = \xi$, $\psi_3 = \eta$, $\psi_4 = \zeta$, $\psi_5 = \xi^2 + \eta^2 + \zeta^2$

In order that the inhomogeneous equation have a solution, the inhomogeneous part must be orthogonal to the five solutions ψ_λ. This leads to the five integrability conditions :

$$\left.\begin{aligned} \frac{\partial_0 \rho}{\partial t} &= -\operatorname{div}(\rho \vec{u}) \\ \frac{\partial_0 \vec{u}}{\partial t} &= -(\vec{u}\cdot\operatorname{grad})\vec{u} + \vec{x} - \frac{1}{\rho}\operatorname{grad} p \\ \frac{\partial_0 T}{\partial t} &= -\vec{u}\cdot\operatorname{grad} T - \frac{2}{3} T \operatorname{div}\vec{u} \end{aligned}\right\} \qquad (5\,b)$$

which are the ideal fluid, or first order hydrodynamical, equations for n, u_i, T, or as Enskog expresses it : in first order the time derivatives $\frac{\partial_0 \rho}{\partial t}, \frac{\partial u_i}{\partial t}, \frac{\partial_0 T}{\partial t}$ must be taken to be given by the ideal hydrodynamical equations. One can also say that in first order one can calculate the average values in the general hydrodynamical equations (3) with the local Maxwell-Boltzmann distribution (7), giving the Euler equations

$$P^{(0)}_{ij} = p\,\delta_{ij}, \qquad q^{(0)}_i = 0.$$

Note that $p = nkT$ is given by the ideal gas law. If (5b) is fulfilled (5a) has a solution, which is determined except for an arbitrary linear combination of the five solutions ψ_α of the homogeneous equation. However, we will omit this combination and make $f^{(1)}$ definite by requiring :

$$\int \psi_\alpha f^{(1)} d\tau = 0 \qquad (5\,c)$$

c) Solution of (5 a).

The left-hand side becomes, using (7) and eliminating the $\partial_0/\partial t$ terms with (5 b),

$$\left(\frac{\partial_0}{\partial t} + D\right) f^{(0)} =$$

$$= f^{(0)}\left[\frac{\partial \ln T}{\partial x_\alpha}\frac{1}{4}\left(\frac{mc^2}{kT}\right)\left(\frac{mc^2}{kT} - \frac{5}{2}\right) + \frac{m}{kT} D_{\alpha\beta}\left(c_\alpha c_\beta - \frac{1}{3}\delta_{\alpha\beta} c^2\right)\right]$$

From the linearity of the operator I it follows that one can decompose $\Phi^{(1)}$. writing $f^{(0)} = Ae^{-\frac{m}{2kT}u^2}$ one can put

$$\Phi^{(1)} = -\frac{1}{AT}\frac{\partial T}{\partial x_\alpha} \Pi_\alpha - \frac{1}{A}\frac{p}{kT} D_{\alpha\beta}\,\Phi_{\alpha\beta}$$

where Π_i , Φ_{ij} are functions of the U_i determined by :

$$I'(\Pi_i) = U_i\left(\frac{m}{2kT}U^2 - \frac{5}{2}\right)e^{-\frac{m}{2kT}U^2}$$

$$I'(\Phi_{ij}) = \left(U_iU_j - \frac{1}{3}\delta_{ij}U^2\right)e^{-\frac{m}{2kT}U^2} \qquad (5\,d)$$

and :

$$I'(\Psi) \equiv \int d\vec{U}_1 \int d\Omega\, g(g,\theta)\left(\Psi + \Psi' - \Psi_1' - \Psi_1\right)e^{-\frac{m}{2kT}(U^2+U_1^2)}$$

is recognized as the linearized Boltzmann operator. Since I' is an isotropic operator, Π_i and Φ_{ij} must be, respectively, an isotropic vecotor and an isotropic tensor in the velocity space U_i, so that they must have the form :

$$\Pi_i = U_i\,\chi_a(U^2), \qquad \Phi_{ij} = \left(U_iU_j - \frac{1}{3}\delta_{ij}\right)\chi_b(U^2) \qquad (5e)$$

The two functions χ_a and χ_b are determined by (5 d). The solution is immediate if one knows the eigenfunctions and eigenvalues of the operator I. This is the case for Maxwell molecules; for other interactions one develops χ_a and χ_b in Sonine polynomials, which are the orthogonal eigenfunctions of $I(\psi)$ for Maxwell molecules. (See Equation (9), page5). The resulting infinite set of linear equations can be solved by convergent series expansions, of which the parameter is a measure of the deviation of the force law from the $1/r^5$ repulsion.

d) Third approximation.

Putting $f^{(2)} = f^{(0)}\,\phi^{(2)}$, and keeping in mind that $(f^0 f_1^0 \psi') = f^0 f_1^0 \psi'$, equation (6) becomes :

$$\frac{\partial_1 f^{(0)}}{\partial t} + \left(\frac{\partial_0}{\partial t} + D\right)f^{(1)} - J\left(f^{(0)} f_1^{(1)}\right) = -I\left[\phi^{(2)}\right] \qquad (6\,a)$$

Again the left- hand side is known and must be orthogonal to the five solutions

Ψ_α of the homogeneous equation. Since clearly :

(by the "Boltzmann" transformation) and

$$\int \Psi_\alpha \frac{\partial f^{(0)}}{\partial t} d\vec{\xi} = 0$$

because of (5a), one only has to fulfill :

$$\int \Psi_\alpha \left(\frac{\partial f^{(0)}}{\partial t} + D f^{(0)} \right) d\vec{\xi} = 0 \qquad (6b)$$

One can say that (6b) determines (note that $f^{(1)}$ is completely known) the first order time derivatives of n, u_i and T. One finds that the zero and first time derivatives together fulfill the Stokes-Navier or second order hydrodynamic equations, which therefore appear as the integrability conditions for the third approximation. The same equations are obtained by calculating the average values in the general hydrodynamical equations (3) using for the distribution function $f^{(0)} + f^{(1)}$. One finds:

$$P_{ij}^{(1)} = p\delta_{ij} - 2\mu \left(D_{ij} - 1/3\, D_{kk} \delta_{ij} \right), \quad q_i^{(1)} = \nu \frac{\partial T}{\partial x_i}$$

which are the Newton and Fourier laws. When the stress tensor given by Newton's law is inserted in the equations of motion (the middle ones of (3)), the Navier-Stokes equations result. The viscosity coefficient μ and heat conduction coefficient ν are certain integrals over the functions $\chi_a(U^2)$, $\chi_b(U^2)$ and in first approximation (exact for Maxwell molecules!) are given by :

$$\mu = \frac{5}{8} \sqrt{\pi m kT} \, \frac{1}{\int_0^\infty dg\, g^7 e^{-g^2} Q_{tr}(g)} \qquad (8)$$

$$\nu = 5/2\, c_v \mu$$

where g is the relative velocity in units $\sqrt{4\,kT/m}$ and Q_{t_r} is the so-called transport cross section :

$$Q_{t_r}(g) = 2\pi \int_0^\pi d\theta \, \sin^3\theta \, I(g,\theta).$$

If the first order time derivatives of n, u_i and T are fixed by (6b), then

(6a) has a solution determined up to a linear combination of the ψ_α, which again we omit, fixing the $f^{(2)}$ by the requirements :

$$\int \psi_\alpha f^{(2)} d\vec{\xi} = 0$$

The $f^{(2)}$ can then be determined by the same method as used for $f^{(1)}$, and in this way one can go on. The integrability conditions for the fourth approxmmation are the Burnett, or third order hydrodynamical equations, also obtainable from (3) by using $f^{(0)} + f^{(1)} + f^{(2)}$ in the calculation of the average values. They are still of the first order in the time derivatives but now contain higher order space derivatives of the macroscopic quantities ρ, u_i and T, and new gas constants analogous to μ and κ appear.

19. General discussion of the development

a. - The development can be said to be in powers of $\lambda \nabla$ where λ is of the order of the mean free path and the gradient operator acts on the macroscopic quantities. In fact, from (5d) one sees that $\phi^{(1)}$ contains terms of order :

$$\frac{1}{n\sigma} \frac{1}{T} \frac{\partial T}{\partial x} \cong \frac{\lambda}{T} \frac{\partial T}{\partial x}$$

and

$$\frac{1}{n\sigma} \sqrt{\frac{m}{2kT}} \frac{\partial u}{\partial x} \cong \frac{\lambda}{\text{sound vel.}} \frac{\partial u}{\partial x}$$

where σ is some kind of collision cross section. The development is therefore in the uniformity of the macroscopic quantities ρ, u_i, T and one can say that the Hilbert parameter λ is a measure of the uniformity.

b. - The development is so arranged that at any stage the equations are of first order in the time derivatives of ρ, u_i and T (by leaving off the homogeneous solutions). Therefore the initial values of ρ, u_i and T determine the "state" of the gas. This macroscopic causality theorem (Hilbert) is of course not a theorem but an "Ansatz". It seems paradoxical because from the Boltzmann equation it clearly follows that one needs initially $f(\xi_1, \xi_2, \xi_3, xyz, 0)$ (which contains of course much more information than the five moments $n(xyz)$, $u_i(xyz)$, $T(xyz)$ do.) to determine the further state of the gas.

One must expect, therefore, that an arbitrary initial velocity distribution $f(\xi_1, \xi_2, \xi_3; xyz, 0)$ in a very short time (of order of the time between collisions, $\lambda \sqrt{\frac{kT}{m}}$) degenerates into a normal

state" determined through the local Maxwell-Boltzmann distribution by the macroscopic quantities ρ, u_i, T, and that the further slow equalizatio or adjustmant of the spatial non-uniformities proceeds according to the hydrodynamical equations and in harmony with the macroscopic causality requirement.

c. - The development does not distinguish between the magnitude of the disturbance and the scale of the disturbance from equilibrium. In any physical problem these can be well distinguished. For instance, in the propagation of sound they are measured by the intensity and the wavelength of the sound wave, respectively. Especially for the discussion of what happens at low pressures (that is, for large mean free paths) the Hilbert-Enskog development is not appropriate. For the theory of the transition between the "Clausius gas regime" (at moderate pressures) and the "Knudsen gas regime" (at very low pressures) it is better to make a straightforward perturbation expansion of the Boltzmann equation by putting :

$$f = f_o (1 + h)$$

where f_o is the complete equilibrium distribution and the perturbation h satisfies the linear homogeneous equation :

$$\frac{\partial h}{\partial t} + \xi_\alpha \frac{\partial h}{\partial x_\alpha} = J(h)$$

in the absence of outside forces. For various special cases this can be solved in detail, allowing for a discussion of the dependence on the scale or Knudsen number ($\sim \lambda / L$, where L = representative length) for small disturbances or small Mach number.

20. The idea of the virial expansion.

The classical result that the viscosity and heat conduction coefficients are independent of the pressure of the gas is clearly a consequence of the limitation to binary collisions. Presumably for higher densities one may expect that for slowly varying phenomena the Stokes-Navier equations remain valid, but that the μ (and ν) will become functions of the density, which are developable in the form :

$$\mu = \mu_0 + \mu_1 \rho + \mu_2 \rho^2 + \ldots \qquad (1)$$

analogous to the well-known virial development :

$$\frac{pv}{kT} = 1 + \frac{B}{v} + \frac{C}{v^2} + \ldots \qquad (2)$$

of the equation of state, with $v = \frac{V}{N} = \frac{m}{\rho}$. In (2) one knows how to express the virial coefficients in terms of the intermolecular force $\Phi(r)$. One has :

$$B = -\frac{1}{2}\int d\vec{r}\, f(r) = 2\pi \int_0^\infty dr\, r^2 \left(1 - e^{-\frac{\phi(r)}{kT}}\right)$$

$$\left(= -\tfrac{1}{3}\right) \iint d\vec{r}\, d\vec{r}'\, f(r)\, f(r')\, f(|\vec{r} - \vec{r}'|) \qquad (3)$$

And so on! B involves the interaction of a pair of molecules, C of a triple, and so forth.

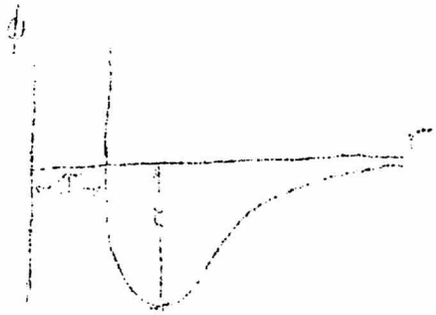
The μ_0 is the viscosity coefficient derived in the previous section :

$$\mu_0 = \frac{5}{8}\sqrt{\pi m kT}\,\frac{1}{\int_0^\infty dg\, g^7 e^{-g^2} Q_{tr}(g)}$$

(g = relative velocity in units $\sqrt{4kT/m}$; $Q_{tr} = 2\pi \int_0^\pi d\theta \sin^3\theta\, I(g,\theta)$)

It involves through the $I(g,\theta)$ the intermolecular force $\phi(r)$ between a pair of molecules. It is therefore related to the second virial coefficient B. The relation is very implicit, but by assuming a form for $\phi(r)$, as for instance a two-parameter Lennards-Jones potential

$$\phi(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$



one can calculate B(T) and $\mu_0(T)$ and see whether one can reproduce the data adapting ε and σ . Such calculations have been made by Hirschfelder and coworkers with success. The agreement is very good, especially for the inert gases.

Analogously one must expect that μ_1 involves the interaction of the molecules in triples, so that it would be related to the third virial coefficient C. However, no general formulae for μ_1 , μ_2 , ... are known.

21. <u>The virial expansion for the equilibrium state</u>.

Although this is well known, and a little aside from our main

topic, I will give a derivation of (2) and (3) by following Bogolubov (Journal of Phys. 10, 256, 1946). It is an introduction to the method he proposes in a following paper (Journal of Phys. 10, 265, 1946) for the virial expansion of non-equilibrium properties. Start from the canonical distribution in co-ordinate space

$$D_N(\vec{q}_1 \cdots \vec{q}_N) = \frac{1}{Q_N} e^{-U_N/kT}, \qquad (4)$$

$$U_N \equiv \sum_{i<j} \phi(|\vec{q}_i - \vec{q}_j|); \quad Q_N = \int_V \cdots \int d\vec{q}_1 \cdots d\vec{q}_N \, e^{-U_N/kT}$$

D_N is symmetric in $q_1 \cdots q_N$, and normalized to unity. One can form partial distribution functions by integrating: let $\frac{1}{V^s} F_s(q_1 \cdots q_s)$ be the probability of finding an s-tuple in $q_1 \cdots q_s$:

$$F_s = V^s \int \cdots \int D_N \, dq_{s+1} \cdots dq_N .$$

Now observe that D_N fulfills

$$\frac{\partial D_N}{\partial q_{i\alpha}} + \frac{1}{kT} \frac{\partial U_N}{\partial q_{i\alpha}} D_N = 0 \qquad (\alpha = x, y, z) \qquad (5)$$

by logarithmic differentiation of (4). Integrate (5) over $q_{s+1} \cdots q_N$; then one gets for F_s in the limit $N \longrightarrow \infty$, $V \longrightarrow \infty$, $v = V/N$ finite:

$$\frac{\partial F_s}{\partial q_{i\alpha}} + \frac{1}{kT} \frac{\partial U_s}{\partial q_{i\alpha}} F_s + \frac{1}{kT} \int d\vec{q}_{s+1} \frac{\partial \phi(|\vec{q}_i - \vec{q}_{s+1}|)}{\partial q_{i\alpha}} F_{s+1} = 0 \qquad (6)$$

where :

$$U_s = \sum_{i<j=1}^{s} \phi(|q_i - q_j|) \equiv \sum \phi_{ij}$$

Proof :

$$U_N = U_s + \sum_{s<i<j\le N} \phi_{ij} + \sum_{i\le s<j} \phi_{ij}$$

From the U_s part of U_N in (5) one gets the first two terms in (6). If i, j are both $> s$, differentiation by $q_{i\alpha}$ gives zero; the only terms

remaining which give a non-vanishing contribution are ϕ_{ij}, with $j > s$.

Because of the symmetry of D_N, one gets for all the N - s values of j the same expression, namely :

$$\frac{1}{kTV} \int d\vec{q}_{s+1} \frac{\partial \phi_{i,s+1}}{\partial \vec{q}_{i\alpha}} F_{s+1}$$

so that in the limit we have (6) $\left(\frac{N-s}{V} \longrightarrow \frac{1}{v} \right)$.

One can now say that the hierarchy of eqs. (6) for the successive distribution functions must be solved with the conditions :

$$F_s(\vec{q}_1 \cdots \vec{q}_s) - \prod_{i=1}^{s} F_1(\vec{q}_i) \longrightarrow 0 \qquad (7)$$

if all $|\vec{q}_i - \vec{q}_j| \longrightarrow \infty$

Now make the development in powers of $\frac{1}{v}$:

$$F_s = F_s^{(0)} + \frac{1}{v} F_s^{(1)} + \frac{1}{v^2} F_s^{(2)} + \cdots \qquad (8)$$

Then one gets from (6) :

$$\left.\begin{aligned} &\frac{\partial F_s^{(0)}}{\partial \vec{q}_{i\alpha}} + \frac{1}{kT} \frac{\partial \phi_{is}}{\partial \vec{q}_{i\alpha}} F_s^{(0)} = 0 \\ &\frac{\partial F_s^{(r)}}{\partial \vec{q}_{i\alpha}} + \frac{1}{kT} \frac{\partial \phi_{is}}{\partial \vec{q}_{i\alpha}} F_s^{(r)} + \frac{1}{kT} \int d\vec{q}_{s+1} \frac{\partial \phi_{i,s+1}}{\partial \vec{q}_{i\alpha}} F_{s+1}^{(r-1)} = 0 \end{aligned}\right\} \qquad (9)$$

which can now be solved successively. The development (8) must also be used in (7), giving :

$$\left.\begin{aligned} &F_s^{(0)}(\vec{q}_1 \cdots \vec{q}_s) - \prod_{i=1}^{s} F_1^{(0)}(\vec{q}_i) \longrightarrow 0, \\ &F_s^{(1)} \longrightarrow \sum_{i=1}^{s} F_1^{(1)}(\vec{q}_i) \prod_{k \neq i} F_1^{(0)}(\vec{q}_k), \end{aligned}\right\} \qquad (10)$$

etc. Finally, one must require, since $\int F_1(q_1)\,dq = V$, that :

$$\lim_{V\to\infty} \frac{1}{V} F_1^{(1)} dq = 1 \qquad \lim_{V\to\infty} \frac{1}{V} \int_V F_1^{(r)}(q_1)\,dq = 0 \qquad (11)$$

Eqs. (9), (10), (11), suffice to determine all $F_s^{(r)}$.

For instance in <u>zeroth approximation</u> : from (9) it follows, putting:

$$F_s^{(0)} = C_s^{(0)}(q_1 \cdots q_s)\exp(-U_s/kT)$$

that

$$\frac{\partial C_s^{(0)}}{\partial q_{s+1}} = 0$$

so $C_s^{(0)}$ = const. (because of symmetry requirement). From boundary condition (10),

$C_s^{(0)} = (C_1^{(0)})^s$, and from normalization (11) $C^{(0)}$, so :

$$F^{(0)} = 1 , \quad F_s^{(0)} = \exp(-U_s/kT). \qquad (12)$$

In <u>first approximation</u>, putting again $F_s^{(1)} = C_s^{(1)}\exp(-U_s/kT)$, one gets from (9), using the symmetry requirement :

$$C_s^{(1)} = \int dq_{s+1}\left\{ e^{-\frac{1}{kT}\sum_{i=1}^{s}\phi_{i,s+1}} - 1 \right\} + B_s .$$

B_s is a constant. If $|q_i - q_s| \to \infty$, for $i,j = 1,2,\ldots,s$, one sees that :

$$F_s^{(1)} \to \sum_{i=1}^{s} \int dq_{s+1}\left(e^{-\frac{1}{kT}\phi_{i,s+1}} - 1 \right) + B_s$$

Since :

$$F_1^{(1)} = C_1^{(1)} = \int dq_2 \left\{ e^{-\frac{1}{kT}\phi_{1,2}} - 1 \right\} + B_1 ,$$

one gets from the boundary condition (10), that B_s = sBr . Finally from the normalization condition (11) :

$$B_1 = -\int dq_2 \left\{ e^{-\frac{1}{kT}\phi_{1,2}} - 1 \right\}$$

giving :

$$F_1^{(1)} = 0, \quad F_s^{(1)} = \epsilon^{-\frac{U_s}{kT}} \int dq_{s+1} \left\{ \prod_{i=1}^{s} \left(1 + f(q_i - q_{s+1})\right) - 1 - \sum_{i=1}^{s} f(q_i - q_{s+1}) \right\} \tag{13}$$

where for abbrevation :

$$f(r) = e^{-\phi(r)/kT} - 1$$

And so on. To get the equation of state one only has to use the general virial theorem which leads to :

$$\frac{PV}{kT} = 1 - \frac{1}{6kT} \cdot \frac{1}{2} \int d\vec{r}\, r\, \phi(r) F_2(r)$$

where $r = q_1 - q_2$, in which one must then substitute the successive approximation for $F_2(q_1, q_2)$ which one has found. Using the zeroth and 1st approximations (12) and (13) gives the equation of state up to the third virial coefficient.

22. <u>The Kirkwood - Born - Green development</u>.

For <u>non-equilibrium situations</u> the probability distribution $D_N(x_1, \ldots x_N, t)$ in Γ-space $(x_i \equiv \vec{q}_i, \vec{p}_i)$ is of course not known.

All we do know is that D_N must be symmetric in $x_1, x_2 \ldots x_n$ and that it fulfills Liouville's equation :

$$\frac{\partial D_N}{\partial t} = \{H, D\} \tag{14}$$

$$\left(\{H, D\} = \sum_{i=1}^{N} \frac{\partial H}{\partial q_i}\frac{\partial D}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial D}{\partial q_i}\right), \quad H = \sum_{i=1}^{N}\left[T(p_i) + U(q_i)\right] + \sum_{i<j}\phi(r_{ij}),$$

$T(p) = \frac{p^2}{2m}$; $U(q)$ = outside force, in which one must include the "wall potential" is the gas is enclosed in a vessel of volume V.

One still can introduce the partial distributions :

$$\frac{1}{V^s} F_s(t, x_1, \ldots x_s) = \int \ldots \int D_N\, dx_{s+1} \ldots dx_N$$

and then (14) is equivalent to a hierarchy of equations. These were written out at about the same time by Kirkwood, Born and Green, and Bogolubov, and I will follow the last. In the limit $N \to \infty$, $V \to \infty$ $v = V/N$ finite, one gets :

$$\frac{\partial F_s}{\partial t} = \{H_s, F_s\} + \frac{1}{v}\int dx_{s+1} \left\{\sum_{i=1}^{s} \phi_{i,s+1}, F_{s+1}\right\}. \qquad (15)$$

H_s is the Hamiltonian of the s-tuple of particles :

$$H_s = \sum_{i=1}^{s} T(p_i) + \sum_{1 \le i < j \le s} \phi_{ij}$$

where there are no outside forces, and the wall potential can also be omitted if V is very large.

The proof of (15) is quite similar to that of (6). In fact one gets immediately :

$$\frac{\partial F_s}{\partial t} = \{H_s, F_s\} + V^s \int \cdots \int dx_{s+1} \cdots dx_N \left\{\sum_{i=1}^{s} [T(p_i) + U(q_i)] + \sum_{\substack{i,j=1 \\ i<j \le s}} \phi_{ij}, D\right\}.$$

Since D must be assumed to vanish for large p_i and q_i, one easily sees that in the last term only the terms with the intermolecular potential with $i \le s$ and $j > s$ will give a contribution. For such a term :

$$V^s \int \cdots \int dx_{s+1} \cdots dx_N \{\phi(r_{ij}), D\} =$$
$$= V^s \int dx_j \frac{\partial \phi}{\partial q_i} \frac{\partial}{\partial p_i} \int \cdots \int D\, dx_{s+1} \cdots ' dx_N$$

where the prime means that dx_j is omitted. Because of the symmetry of D, one can always put $j = s + 1$ and one has $N - s$ equal terms :

$$V^s \int dx_{s+1} \frac{\partial \phi(|q_i - q_{s+1}|)}{\partial q_i} \frac{\partial}{\partial p_i} \int \cdots \int D\, dx_{s+2} \cdots dx_N$$

$$= \frac{1}{V} \int dx_{s+1} \{\phi(|q_i - q_{s+1}|), F_{s+1}\}.$$

Altogether one gets therefore :

$$\frac{N-s}{V}\int dx_{s+1}\left\{\sum_{i=1}^{s}\phi_{i,s+1},\ F_s\right\}$$

which in the limit gives the last term in (15).

Of course no real advance has been made. The set of equations (15) is equivalent to the Liouville theorem. Note that for s = 1, (15) becomes :

$$\frac{\partial F_1(q,p,t)}{\partial t} = -\frac{p_\alpha}{m}\frac{\partial F_1}{\partial q_\alpha} + \int d\vec{p}_1 \int d\vec{q}_1 \frac{\partial \phi(|\vec{q}-\vec{q}_1|)}{\partial q_\alpha}\frac{\partial F_2(q,q_1,p,p_1)}{\partial p_\alpha} . \qquad (15a)$$

which "smells" like the Boltzmann equation, but the connection is still far from clear.

The ideas of Bogolubov . II.

23. General ideas.

One is tempted to imitate the virial expansion for the equilibrium state, starting now from eq. (15). This is not a good idea, because in the zeroth approximation one would get :

$$\frac{\partial F_s^{(0)}}{\partial t} = \{H_s, F_s^{(0)}\}$$

which means that $F_s^{(0)}$ changes in time according to the interactions between the s particles, without collisions with other particles. Such an expansion can therefore be valid only for very short times, small compared to the time $t_o = \lambda / U_{av}$ between collisions. In fact the order of magnitude of successive terms in such an expansion after time t would be :

$$\frac{F_2}{F_1} \frac{t}{r_o/U_{av}} \cong n r_o^3 U_{av} t \cong \frac{U_{av} t}{\lambda} = \frac{t}{t_o}$$

(n = 1/v = number of particles / cc ; r_o = range of intermolecular forces; $\lambda \cong \frac{1}{n r_o^2}$ = mean free path).

Yet one must somehow obtain the Boltzmann equation in first approximation. One might think that at low densities somehow the state of the gas can be described by the first distribution function F_1 (x, t) and that all higher ones are products of F_1; at higher densities one would need both F_1 and F_2, the higher distribution should be expressible in these, and so on. I thought along these lines for some time, and also Kirkwood follows this idea. Success has been very meagre; at best one can "derive" the Boltzmann equation from (15a), but no one has really gone any further.

Bogolubov argues in a different way. He distinguishes in the temporal development of the gas three stages :

a. Initial mixing : from an arbitrary initial distribution $D(x_1 \ldots x_n, 0)$ he supposes that very quickly (in a time of the order $\tau_o = r_o/U_{av}$ = the interaction time, or time of a collision; τ_o is very small compared to $t_o = \lambda / U_{av}$ which is the mean free time or time between collisions, at least for moderate densities), a (first) "smoothing" process occurs, and that from then on the development depends only on $F_1(x, t)$. The initial detailed knowledge is lost, so that after this initial period the further development depends only on F_1.

b. Kinetic stage. This describes the development of F_1, for which one assumes a kinetic equation of the form :

$$\frac{\partial F_1}{\partial t} = A(x, F_1) = A_0(x, F_1) + \frac{1}{U} A_1(x, F_1) + \frac{1}{U^2} A_2(x, F_1) + \ldots$$

(1)

while all the higher probability distributions depend on the time only through F_1. One puts :

$$F_s(t, x_1 \ldots x_s) = F_s(x_1 \ldots x_s ; F_1) = F_s^{(0)}(x_1 \ldots x_s ; F_1) + \frac{1}{\sigma} F_s^{(1)}(x_1 \ldots x_s ; F_1) + \ldots \qquad (2)$$

for $s \geq 2$. One keeps the eqs. (15) relating the distribution functions, and the problem is to find $A(x, F_1)$. It must of course be so arranged that A_0 describes the streaming, A_1 the Boltzmann binary collision term, A_2 the triple collision term, etc.

After a while, in a time of order t_0 (the mean free time), but small compared to macroscopic relaxation times, a second smoothing process occurs, so that from then on one does not need F_1, but only the five moments ρ, $\vec{u}$, and T. Again one can say that some knowledge is lost, and that the further development now depends only on the values of ρ, $\vec{u}$, and T after this second period.

c. Hydrodynamical stage. This describes the development of p, $\vec{u}$, and T, for which one assumes hydrodynamical eqs. of the general form :

$$\frac{\partial \rho}{\partial t} = R(\vec{q}, \rho, \vec{u}, T)$$

$$\frac{\partial \vec{u}}{\partial t} = \vec{U}(\vec{q}, \rho, \vec{u}, T)$$

$$\frac{\partial T}{\partial t} = \Phi(\vec{q}, \rho, \vec{u}, T). \qquad (3)$$

One keeps the kinetic equation (1) previously derives, but one assumes that F_1 depends on the time only through p, $\vec{u}$, T, so :

$$F_1(t, \vec{q}, \vec{p}) = F_1(\vec{q}, \vec{p}, \rho, \vec{u}, T).$$

The problem is to find the functions R, $\vec{U}$, Φ . Like Hilbert and Enskog, Bogolubov uses an uniformity parameter in all functions are developed, so that this part is rather analogous to the Hilbert-Enskog development, except that now in principle the virial type of development can be obtained for the viscosity and heat conduction coefficients also. There are no results as yet. But there is a whole program !

24. The kinetic equation.

Using the general equation (15) :

$$\frac{\partial F_s}{\partial t} = \{H_s, F_s\} + \frac{1}{\sigma} \int dx_{s+1} \left\{ \sum_{i=1}^{s} \Phi(q_i - q_{s+1}), F_{s+1} \right\}$$

and the expansions (1) and (2), one gets for S = 1 :

$$A_0(x_1; F_1) = \{T(p_1), F_1\} = -\frac{p_1}{m_1}\frac{\partial F_1}{\partial q_1}$$

$$A_1(x_1; F_1) = \int dx_2 \{\phi_{12}, F_2^{(0)}(x_1, x_2; F_1)\} \qquad (4)$$

etc. A_0 is therefore the streaming term, as expected. For $s \geq 2$, one gets :

$$\{H_s, F_s^{(0)}\} - D_0 F_s^{(0)} = 0 \qquad (5a)$$

$$\{H_s, F_s^{(1)}\} - D_0 F_s^{(1)} - D_1 F_s^{(0)} + \int dx_{s+1} \{\sum_{(i)}^{s} \phi_{i,s+1}, F_{s+1}^{(0)}\} \qquad (5b)$$

etc. Here the differentiation operation D_r acting on any function $\Psi(x_1 \ldots x_s, F_1)$ means

$$D_r \Psi \equiv \frac{\partial \Psi}{\partial F_1} A_r(x_1, F_1) \qquad (6)$$

since all time dependences occur through F_1. The program is now to solve equation (5a), which then from (4b) would give A_1, then solve (5b) which would give A_2, etc.

However, to solve (5a) one needs some boundary conditions. Bogolubov demands that for any function F_1 and for all F_s $(s \geq 2)$:

$$S_{-\tau}^{(s)} F_s(x_1 \ldots x_s, S_\tau^{(1)} F_1) - \prod_{(i)} (S_\tau^{(1)} F_1(x_i)) \} \longrightarrow 0$$

for $\tau \longrightarrow \infty$ (7)

Excursion on the operator $S_t^{(s)}$:

Suppose that the motion of the s particles under the influence of their mutual interactions (with Hamiltonian H_s) is solved, so that one knows the phases :

$$x_i = x_i(t, x_1 \ldots x_s) \qquad i = 1, 2 \ldots s$$

of the s points as functions of t and the initial phases $x_1 \ldots x_s$.

Define the "streaming operator" $S_t^{(s)}$ acting on a function $f(x_1 \dots x_s)$ by :

$$S_t^{(s)} f(x_1 \dots x_s) \equiv f(x_1, x_2, \dots x_s). \tag{8}$$

In particular, of course, $S_t^{(1)} x_i = x_i$, and also clearly $S^{(s)}_{t_1+t_2} = S^{(s)}_{t_1} \cdot S^{(s)}_{t_2}$. With the help of $S^{(s)}_t$ one can "integrate" partial differential equations of the form :

$$\frac{\partial \phi(t, x_1 \dots x_s)}{\partial t} = \{H_s, \phi\} + f(t, x_1 \dots x_s) \tag{9}$$

where f is known. Namely, one can write equation (9) in the form $\frac{D\phi}{Dt} = f(t, x_1 \dots x_s)$ ($\frac{D}{Dt}$ = substantial differentiation quotient). Therefore, $\phi(t, x_1 \dots x_s) = \phi(0, x_{1,0} \dots x_{s,0}) + \int_0^t (\tau, x_{1,\tau} \dots x_{s,\tau})$, where the initial positions $x_{1,0} \dots x_{s,0}$ and the positions at time τ, $x_{1,\tau} \dots x_{s,\tau}$, must be considered as functions of t and the $x_1 \dots x_s$ (which are the positions at time t). Therefore :

$$\phi(0, x_{1,0} \dots x_{s,0}) = \phi(0, S^{(1)}_{-t} x_1, \dots S^{(1)}_{-t} x_s)$$
$$= S^{(s)}_{-t} \phi(0, x_1 \dots x_s)$$

and analogously :

$$f(\tau, x_{1,\tau} \dots x_{s,\tau}) = S_{-(t-\tau)} f(\tau, x_1 \dots x_s).$$

Hence the solution of (9) can be written in the form :

$$\phi(t, x_1 \dots x_s) = S^{(s)}_{-t} \phi(0, x_1 \dots x_s) + \int_0^t d\tau\, S^{(s)}_{-(t-\tau)} f(\tau, x_1 \dots x_s). \tag{10}$$

Condition (7) somehow says that before the s-tuple collision the F_s is a product of the $F_1(x_i)$. I must say that it is not clear to me, but it works, and it is perhaps best to show how one can now determine the $A_i(x; F_1)$ successively.

First note that with the expansion (2), the condition (7) is equivalent to :

$$S^{(s)}_{-\tau} \left\{ F_s^{(s)}(x_1 \dots x_s; S^{(1)}_{\tau} F_1) - \prod_{(i)} S^{(1)}_{\tau} F_1(x_i) \right\} \to 0 \quad \text{as } \tau \to \infty,$$

$$S^{(s)}_{-\tau} F_s^{(s)}(x_1 \dots x_s; S^{(1)}_{\tau} F_1) \to 0 \quad \text{as } \tau \to \infty \tag{7a}$$

etc. Now one can start.

First approximation.

We have to solve (5a). Since (5a) must hold for any function F_1, put in $F_s^{(0)}$ the function $S_{-\tau}^{(1)} F_1$ for F_1. Then one sees that

$$D_0 F_s^{(0)}(x_1 \cdots x_s; S_{-\tau}^{(1)} F_1) = \frac{\partial}{\partial \tau} F_s^{(0)}(x_1 \cdots x_s; S_{-\tau}^{(1)} F_1)$$

because

$$\frac{\partial}{\partial \tau} F_s^{(0)} = \frac{\partial F_s^{(0)}}{\partial (S_\tau^{(1)} F_1)} \cdot \frac{\partial}{\partial \tau} S_{-\tau}^{(1)} F_1$$

and

$$\frac{\partial}{\partial \tau} S_{-\tau}^{(1)} F_1 = \{H_1, F_1\} = \{T(p_1) F_1\} = A_0(x; F_1)$$

so one just gets $D_0 F_s^{(0)}$ according to the definition (6) of the operator D_0. Therefore (5a) becomes :

$$\frac{\partial}{\partial \tau} F_s^{(0)}(x_1 \cdots x_s; S_{-\tau}^{(1)} F_1) = \{H_s, F_s^{(0)}(x_1 \cdots x_s; S_{-\tau}^{(1)} F_1)\}$$

whence (using the first term of the general solution (10)),

$$F_s^{(0)}(x_1 \cdots x_s; S_{-\tau}^{(1)} F_1) = S_{-\tau}^{(s)} F_s^{(0)}(x_1 \cdots x_s; F_1).$$

so :

$$F_s^{(0)}(x_1 \cdots x_s; F_1) = S_{-\tau}^{(s)} F_s^{(0)}(x_1 \cdots x_s; S_\tau^{(1)} F_1).$$

This holds for arbitrary τ, and since the left-hand side is independent of τ, one can go to $\tau \longrightarrow \infty$, and using (7a) one gets :

$$S_\tau^{(1)} F_1(x_i) = F_1\left(q_i + \frac{p_i}{m} \tau, p_i\right),$$

On the other hand :

$$S_{\tau}^{(1)} F_i(x_i) = F_i\left(q_i + \frac{p_i}{m}, p_i\right),$$

so :

$$S_{-\tau}^{(s)} \prod_i S_{\tau}^{(1)} F(x_i) = \prod_i F_i\left(S_{-\tau}^{(s)} q_i + \frac{\tau}{m} S_{-\tau}^{(s)} p_i, S_{-\tau}^{(s)} p_i\right).$$

Call now :

$$P_i^{(s)}(x_1 \cdots x_s) \equiv \lim_{\tau \to \infty} S_{-\tau}^{(s)} p_i, \qquad (11)$$

$P_i^{(s)}$ is the constant initial impulse of the i-th particle in the s-tuple collision (governed by H_s), which leads to the configuration $x_1 \ldots x_s$ at time zero. The transition from p_i to P_i goes fast (in a time of order of the interaction time $T_o = r_o/U_{av}$), at least for small s), and since :

$$S_{-\tau}^{(s)} q_i = q_i - \frac{1}{m}\int_0^{\tau} d\tau' \, S_{-\tau'}^{(s)} p_i$$

one can see that for $\tau \to \infty$, $S_{-\tau}^{(s)} q_i + \frac{\tau}{m} S_{-\tau}^{(s)} p_i$ will go to the limit :

$$Q_i^{(s)}(x_1 \cdots x_s) = q_i + \frac{1}{m}\int_0^{\infty} \left\{ P_i^{(s)} - S_{-\tau}^{(s)} p_i \right\} d\tau \qquad (12)$$

$Q_i^{(s)}$ is the position the i-th particle in the s-tuple collision at time zero would have had if the particle had gone on with the initial momentum $P_i^{(s)}$. With the definitions (11) and (12), one has therefore:

$$F_s^{(0)}(x_1 \cdots x_s ; F_1) = \prod_{i=1}^{s} F_1\left(Q_i^{(s)} P_i^{(s)}\right)$$

and according to (4b) this leads to :

$$A_1(x_1, F_1) = \int dx_2 \left\{ \Phi_{1,2} \, F_1\left(Q_1^{(2)} P_1^{(2)}\right) F_1\left(Q_2^{(2)}, P_{(2)}^{(2)}\right) \right\} \qquad (14)$$

We will show the relation of this expression to the Boltzmann collision operator presently. However, let us first look at the :

Second approximation.

To determine $F_s^{(1)}$ ($s \geq 2$), one has from (5b) :

$$D_0 F_s^{(1)}(x_1 \dots x_s ; F_1) = \{H_s , F_s^{(1)}\} + \Psi_s(x_1 \dots x_s ; F_1), \quad (15)$$

where we have put

$$\Psi_s \equiv - D_1 F_s^{(0)} + \int dx_{s+1} \left\{ \sum_{i=1}^{s} \phi_{i,s+1} F_{s+1}^{(0)} \right\}, \quad (16)$$

Clearly Ψ_s is known. In fact from the definition (6) of D, one gets :

$$\Psi_s(x_1 \dots x_s ; F_1) = - \sum_{i=1}^{s} A_i(Q_i^{(s)}, P_i^{(s)} ; F_i) \prod_{J=1}^{s} F_1(Q_J^{(s)}, P_J^{(s)})$$

$$- \int dx_{s+1} \left\{ \sum_{i=1}^{s} \phi_{i,s+1} \prod_{J=1}^{s+1} F_1(Q_j^{(s+1)}, P_j^{(s+1)}) \right\} \quad (16a)$$

Now put in (15) $S_{-\tau}^{(1)} F_1$ for F_1 again; then one gets

$$\frac{\partial}{\partial \tau} F_s^{(1)}(x_1 \dots x_s , S_{-\tau}^{(1)} F_1) = \{H_s , F_s^{(1)}\} + \Psi_s(x_1 \dots x_s ; S_{-\tau}^{(1)} F_1).$$

Using now the complete solution (10), one gets :

$$F_s^{(1)}(x_1 \dots x_s ; S_{-\tau}^{(1)} F_1) = S_{-\tau}^{(s)} F_s^{(1)}(x_1 \dots x_s ; F_1) +$$

$$+ \int_0^{\tau} d\tau' S_{(\tau - \tau')}^{(s)} \Psi_s(x_1 \dots x_s , S_{-\tau}^{(1)} F_1),$$

or (also changing $\tau - \tau'$ to τ') :

$$F_s^{(1)}(x_1 \dots x_s ; F_1) = S_{-\tau}^{(s)} F_s^{(1)}(x_1 \dots x_s , S_{\tau}^{(1)} F_1) +$$

$$+ \int_0^{\tau} d\tau' S_{-\tau'}^{(s)} \Psi_s(x_1 \dots x_s ; S_{\tau'}^{(1)} F_1)$$

The left-hand side is again independent of τ, so by going to the limit $\tau \to \infty$, and using the boundary condition (7a), one gets :

$$F_s^{(1)}(x_1 \dots x_s , F_1) = \int_0^{\infty} d\tau S_{-\tau}^{(s)} \Psi_s(x_1 \dots x_s ; S_{\tau}^{(1)} F_1) \quad (17)$$

which inserted in (4) gives for the triple collision term in the basic

kinetic equation (1) :

$$A_1(x_1; F_1) = \int dx_2 \left\{ \Phi_{1,2}, \int_0^\infty d\tau\, S_{-\tau}^{(2)} \Psi_2(x_1, x_2; S_\tau^{(1)} F_1) \right\} \tag{18}$$

It is now clear how the successive approximation method clicks along.

25. The Boltzmann equation.

We conclude with Bogolubov's proof of the equivalence of the expression (14) for $A_1(x_1; F_1)$ with the Boltzmann collision operator. This equivalence is only rigorously true for the case where F_1 depends only on the impulse. In this case we will write $w(t, p)$ for $F_1(t, \vec{x})$. Of course in this case $A_0(x; F_1) = \{ T(p), w(t,p) \} = 0$ Now note that in general

$$\{ H_s(x_1 \ldots x_s), f(P_1^{(s)} \ldots P_s^{(s)}) \} = 0 \tag{19}$$

where f is an arbitrary function and the $P_i^{(s)}$ are defined by (11). The proof follows by observing that the Poisson bracket (19) must have the same form whether expressed in the variables $x_1 \ldots x_s$ or in the variables $S_{-\tau}^{(s)} x_i$, because the transition from x_1 to $S_{-\tau}^{(s)} x_i$ is a contact transformation and the Poisson bracket is invariant under such transformations. But for $\tau \to +\infty$, H_s will not depend on the coordinates since the particles are then far apart, and the Poisson bracket of two functions of the momenta is zero; hence (19) follows.

Apply (19) for two particles with $H(x_1, x_2) = T(p_1) + T(p_2) + \Phi_{1,2}$ and $f(P_1^{(2)}, P_2^{(2)}) = w(t, P_1^{(1)}) = w(t, P_2^{(1)}) = F_2(x_1, x_2; F_1)$;

then one gets from (14)

$$A_1(x_1; F_1) = \int dx_2 \left\{ \Phi_{1,2}, w(t, P_1^{(2)}) w(t, P_2^{(2)}) \right\}$$

$$= -\int dx_2 \left\{ T(p_1) + T(p_2), w(t, P_1^{(2)}) w(t, P_2^{(2)}) \right\}$$

$$= \int\int d\vec{q}_2\, d\vec{p}_2 \left\{ \frac{P_{1\alpha}}{m} \frac{\partial}{\partial q_{1\alpha}} + \frac{P_{2\alpha}}{m} \frac{\partial}{\partial q_{2\alpha}} \right\} w(t, P_1^{(2)}) w(t, P_2^{(2)}).$$

Since $P_1^{(2)}$ and $P_2^{(2)}$ depend only on $|\vec{q}_2 - \vec{q}_1|$ (central forces!)

$\frac{\partial}{\partial q_{1\alpha}} = -\frac{\partial}{\partial q_{2\alpha}}$. Take cylindrical coordinates now for the dq_2 integration, with the origin in $\vec{q}_1$ and axis in the direction of the relative velocity $\frac{1}{m}(\vec{P}_2 - \vec{P}_1)$; call the coordinate along this axis ξ ; then :

$$\frac{p_{1\alpha}-F_{1\alpha}}{m}\frac{\partial}{\partial q_{1\alpha}}w(t,P_1^{(1)})w(t,P_2^{2})=\frac{|\vec{p}_2-\vec{p}_1|}{m}\frac{\partial}{\partial\xi}w(t,P_1^{(2)})w(t,P_2^{(2)})$$

One now can integrate over ξ . Remember that $P_1^{(2)}(x_1, x_2)$ and $P_2^{(2)}(x_1, x_2)$ are the impulses with which the two bodies start off in the collision which goes through the phase x_1, x_2. Therefore

$$\left\{P_1^{(2)}(x_1,x_2)\right\}_{\xi=-\infty}=p_1\ ,\quad \left\{P_2^{(2)}(x_1,x_2)\right\}_{\xi=-\infty}=p_2$$

and

$$\left\{P_1^{(2)}(x_1,x_2)\right\}_{\xi=+\infty}=p_1^*\ ,\quad \left\{P_2^{(2)}(x_1,x_2)\right\}_{\xi=+\infty}=p_2^*$$

where p_1^*, p_2^* are the impulses of the "restituting collision" (p_1^*, p_2^*) (p_1, p_2). Since because of the axial symmetry $d\,q_2 = 2\pi b\ db\ d\xi$ (where b is the impact parameter), one gets :

$$A_1(x_1,F_1)=2\pi\int d\vec{p}_2\int g\,b\,db\left\{w(t,p_1^*)w(t,p_2^*)-w(t,p_1)w(t,p_2)\right\} \qquad (20)$$

with :

$$g=\frac{1}{m}\left|\vec{p}_2-\vec{p}_1\right|,\quad 2\pi b\,db=2\pi I(g,\theta)\sin\theta\,d\theta$$

; so (20) is just the Boltzmann collision operator.

For the general case in which F_1 depends also on the coordinates, Eq. (14) is only equivalent to the Boltzmann collision operator if one neglects the difference in the positions of the two collision "partners" during a collision. Otherwise one gets correction terms (described by Bogolubov as a kind of "inter-ference" between the streaming and the collision terms), which contribute to the equation of state (remember that the Boltzmann equation leads to the ideal gas law!) and also presumably to the ρ^2 correction in the viscosity. For elastic spheres thsi was already noticed by Enskog (see Chapter 16 in Chapman and Cowling).

For the discussion of the hydrodynamical stage, I have to refer to the book

N. Bogolubov , Problemy Dinamicheskoi Teorii v Statistichestoi Fizike

Addition to Ch. 9

An except of the paper of Kac. Consider first.

a . Stationary, discrete stochastic series , which need not be Markoffian . We will extend the result of Ch. 7 for this more general case .

Denote by $\tilde{n}_i$ " not n_i " , so that for instance :

$$W_1 (\tilde{n}_i) = 1 - \sum_j W (n_j)$$

$$W_2 (n_i, \tilde{n}_k) = \sum_{j \neq k} W_2 (n_i, n_j)$$

etc . Drop the index i , and choose a fixed value of n . Then,from the stationarity , and the general consistency relations between the successive distribution functions , one sees :

$$W_{k+2} (n, \underbrace{\tilde{n} \ldots \tilde{n}}_{k}, n) = W_{k+1} (n, \underbrace{\tilde{n} \ldots \tilde{n}}_{k}) - W_{k+2} (n, \underbrace{\tilde{n} \ldots \tilde{n}}_{k+1}) \qquad (a)$$

$$W_{k+} (n, \underbrace{\tilde{n} \ldots \tilde{n}}_{k}) = W_{k-} (\underbrace{\tilde{n} \ldots \tilde{n}}_{k}) - W_{k+1} (\underbrace{\tilde{n} \ldots \tilde{n}}_{k+1}) \qquad (b)$$

$$W_{k+2} (n, \underbrace{\tilde{n} \ldots \tilde{n}}_{k+1}) = W_{k+1} (\underbrace{\tilde{n} \ldots \tilde{n}}_{k+1}) - W_{k+2} (\underbrace{\tilde{n} \ldots \tilde{n}}_{k+2}) \qquad (c)$$

Calling for abbreviation :

$$w_s = W_s (\underbrace{\tilde{n} \ldots \tilde{n}}_{s})$$

one gets by combining (a) , (b) and (c) , the basic lemma :

$$W_{k+2}(n,\tilde{n}..\tilde{n},n) = w_k - 2\,w_{k+1} + w_{k+2} \qquad (1)$$

For $k = 0$, (a) and (c) are still valid ; (b) becomes :

$$W_1(n) = W_o(\tilde{n}) - W_1(\tilde{n})$$

also true if W_o 1 . So (1) remains valid for $k = 0$.

Theorem 1_a : For each n :

$$\sum_{k=0}^{\infty} P_{k+2}\left(n \,\middle|\, \tilde{n} \ldots \tilde{n}, n\right) = 1 \qquad (2)$$

Proof : The conditional probability P_{k+2} is defined by :

$$P_{k+2}\left(n \,\middle|\, \tilde{n} \ldots \tilde{n}, n\right) = \frac{W_{k+2}(n;\tilde{n}\ldots\tilde{n};n)}{W_1(n)}$$

From (1) one gets easily :

$$\sum_{k=o}^{N} W_{k+2}(n,\tilde{n}\ldots\tilde{n},n) = 1 - W_1(\tilde{n}) + w_{N+2} - w_{N+1}$$

Now $w_{N+1} \geqslant w_{N+2}$, since N+2 successive $\tilde{n}$ -values must be less probable then N+1 values, and the w's are not negative and bounded, so $\lim_{N\to\infty} w_N$ exists, and therefore $\lim_{N\to\infty}(w_{N+2} - w_{N+1}) = 0$, from which (2) follows .

Eq. (2) is the Wiederkehr theorem for discrete stochastic series. It says , that it is certain that an n-value will come back after some time $k\tilde{\tau}$.

Theorem 2a . If : $\lim_{N\to\infty} W_N (\tilde{n} \ldots \tilde{n}) = 0$ (3)

then the average recurrence time of the state n is given by :

$$\tau(n) = \frac{\tau}{W_1(n)} \qquad (4)$$

Proof : Since by definition :

$$\tau(n) = \sum_{k=0}^{\infty} (k+1) \tau P_{k+2} (n / \tilde{n} \ldots \tilde{n}, n)$$

one has to show :

$$\sum_{k=0}^{\infty} (k+1) W_{k+2} (n, \tilde{n} \ldots \tilde{n}, n) = 1$$

From (1) one gets easily :

$$\sum_{k=0}^{N-1} W_{k+2} (n, \tilde{n} \ldots \tilde{n}, n) = 1 - \left\{ N (w_N - w_{N+1}) + w_N \right\}$$

Now , since the left side is a non decreasing sequence , $N (w_N - w_{N+1}) + w_N$ forms a non increasing sequence of positive terms, so that the limit for $N \to \infty$ exists . Since lim w_N exists according to (3) , $\lim N(w_N - w_{N+1})$ exists . Now, the series of non negative terms $\sum (w_N - w_{N+1})$ converges $\left[\sum_{N=0}^{M} (w_N - w_{N+1}) = 1 - w_{M+1} \right.$, and $\left. \lim_{M\to\infty} w_{M+1} = 0 \right]$, and therefore if the limit of $N (w_N - w_{N+1})$ exists (and we know that it does) the limit must be zero . This completes the proof . Consider now :

b. The streaming in Γ - space , say between two energy surfaces E and $E + \Delta E$. Call the total volume $\Omega = 1$, and let A be a small region . (We will speak about a region, by which we mean the small

cylindrical volume between the two energy surfaces) . Let T be a volume preserving transformation of the Γ -space into itself ; say T is the streaming of the ensemble fluid in the time τ . Successive transformations will be denoted by T^n . One then has :

Theorem 1_b . (Poincaré's theorem) : For almost all points ω in A there exists an $n \geq 1$, such that $T^n \omega$ is again in A .

Proof : Let $f(\omega)$ be the characteristic function of the region A (which means : $f(\omega) = 1$ if ω is in A , and zero otherwise) . The sequence : $f(\omega)$, $f(T\omega)$, $f(T^2\omega)$ ... can be considered as a discrete, stationary stochastic series . Each variable can have the values one or zero with probability $[A]$ and $1 - [A]$, if $[A]$ is the volume of A . $W_{k+2}(1,\underbrace{0 \ldots 0}_{k},1)$ is the volume of those points in A which after $k+1$ steps return to A . Theorem 1_a applies , and the statement $\lim_{k \to \infty} W_{k+2}(1,0 \ldots 0,1) = W_{11}$ means that the total volume of points ω which eventually will return to A is equal to the volume of A , so that almost all points will return (" almost " because a set of points of measure zero may be an exception)

Theorem 2_b . If T is metrically transitive (or if the motion is ergotic) , and if $n(\omega)$ is the first $n \geq 1$ for which , if ω is in A , $T^n \omega$ is again in A , then :

$$\int_A n(\omega)\, d\tau = 1$$

or the ensemble average of the length of the Poincaré cycles :

$$\Theta = \frac{1}{[A]} \int_A n(\omega)\, d\tau = [A]$$

which is analogous to 4

Proof : One has to show that the metrical transitivity implies condition 3 . Now condition (3) is in the dynamical case

$$\lim_{n\to\infty} \int_{\Omega} d\omega \prod_{k=0}^{n} \left(1 - f\left(T^k \omega\right)\right) = 0 \qquad (5)$$

which means that the volume of the points in Ω which are never reached by any point in A is zero . Clearly this will be the case if the motion is ergodic . According to the Kirchhoff ergodic theorem it implies that the time average of any function of the motion is equal to the ensemble average. Therefore especially , for any point ω in Ω :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(T^k \omega\right) = [A] \qquad (6)$$

Let B be the set of points , which is never reached by any point in A . So , if a point ω is in B $f\left(T^k \omega\right) = 0$ and therefore $\lim_{n\to\infty} \frac{1}{n} f\left(T^k \omega\right) = 0$; so B must have measure zero, since we know from (6) that for all ω (except for a possible set of measure zero) the limit is $[A]$.

Additions to Ch. I8 and 19 .

I. - The Chapman - Enskog theory is limited because of the following three basic assumptions which have been made :

a - Central forces , which limits the theory to mono- atomic gases .

b - Small value of the development parameter L $\frac{\nabla(\text{macr.qu.}}{\text{macr. qu.}}$, which limits the theory to slowly varying phenomena (no shock waves , or very high frequency sound , where the wavelenght becomes comparable with L). Also Knudsen gases, where L is comparable with macroscopic dimensions , can not be treated .

c - Binary collisions , which limits the theory to gases of low density . This last restriction is the most difficult to remove , since one has to extend the Boltzann equation in a fundamental and ill Bogolubov unknown way . It is dealt with in the last part of the notes .

2 . - The extension to poly atomic gases . The interest lies especially in the application to the propagation of sound . Already in I881 Lorentz remarked that for polyatomic one must expect besides the viscosity and heatconduction coefficients μ and λ , a third gas constant connected with the transfer of translational energy to internal energy , and that this should have some effect on the propagation of sound . For low frequency where the equilibrium between translational and internal degrees of freedom has time to adjust itself all the time , the velocity of sound will be given by the classical formula :

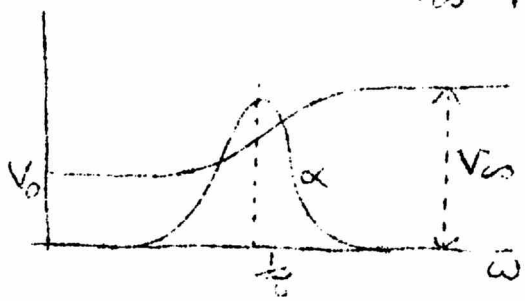
$$V_o = \sqrt{\frac{C_p}{C_v} \frac{k T}{m}}$$

If we assume however that the transfer of energy needs on the average a time τ , so that one has an equation of the form :

$$\frac{d U_{int}}{dt} = \frac{U_{transl.} - U\ int.}{\tau} \qquad (1)$$

(analogous to Newton's law of cooling) , then for frequencies $\omega \gg \frac{1}{\tau}$, the internal energy U_{int} will remain unaffected , and the gas will behave as a mono-atomic gas , so that :

$$V_\infty = \sqrt{\frac{5}{3}\frac{kT}{m}}$$



One gets a dispension region around $\frac{1}{\tau}$ accompanied with an absorption α .

This effect was actually discovered by Kneser in the late twenties for CO_2 , and since the theory has been considered by various authors (Rutgers, Herzfeld- Rice, Bourgin , Landau- Teller a . o) See the review of W.T.Richards , Rev. of Mod.Ph.II , 36 (1939) . There remained the task to incorporate these developments into the general Chapman-Enskog theory . Already in 1922 this was tried by Pidduck (see Chapman and Cowling, Ch.11) . Pidduck assumed that the molecules were rough elastic spheres, so that in a collision translational energy could be transformed into rotational energy . For this model the Chapman-Enskog theory can be generalized, but the model is so artificial that it is hard to see what it has to do with reality . Furthermore , the dispersion of sound is usually caused by the difficulty of the transfer of translational into vibrational energy , and only recently the so-called rotational sound dispersion has been observed for hydrogen .

The origin of the difficulty to extend the theory to more realistic

molecular models lies in the fact that in the classical theory with non-spherical molecules the probabilities for the direct and restituting collisions are in general <u>not</u> equal to each other . It is well known that this fact (also first pointed out by Lorentz) already causes difficulties in the proof of the Maxwell-Boltzmann distribution . It forced Boltzmann to generalize his proof of the H-theorem by considering cycles of collisions (see last chapter of Vol. II in the Gas theorie of Boltzmann) . Dr. C.S. Wang Chang and myself have evaded this difficulty by describing the internal degrees of freedom of the molecule quantum mechanically .[1] Instead of <u>one</u> distribution function f , we now describe the state of the gas by a series of distribution functions :

$$f_i \equiv f\left(\vec{r}, \vec{\xi}, E_i, t\right)$$

giving the number of molecules in the phase-cell $d\vec{r}\, d\vec{\xi}$ which are in the i^{th} quantum state with internal energy E_i . The Boltzmann equation becomes:

$$\frac{D f_i}{D t} = \sum_{j,k\ell} \int d\vec{\xi}_1 \int d\Omega\, g\, I^{k}_{ij}(g, \theta, \varphi) \left[f'_k f'_{j1} - f_i f_{ij} \right] \qquad (2)$$

The $I^{k\ell}_{ij}(g, \theta, \varphi)$ is the differential collision cross-section for a collision of two molecules in the states E_i and E_j , where after collision they are in the states E_K , E_ρ and where the relative velocity has turned over the angles θ, φ , in the solid angle $d\Omega$. The magnitude of the relative velocity changes in general too (because of conservation of energy) , say to g ! However the principle of microscopic reversibility

remains valid , so that :

$$g \, I^{k\ell}_{ij} = g' \, I^{ij}_{k\ell} \qquad (3)$$

This relation is the reason why the semi quantum theoretical description is so much simpler than the classical description . Finally in (2) $f_{ij} \equiv f(\vec{r}, \vec{\xi}, E_j, t)$ and f'_k, f'_ℓ are defined analogously .

Especially for the application to the propagation of sound it is convenient to distinguish the two limiting cases :

a - The inelastic collision cross section is of the same order of magitude as the elastic cross section .

b - $I_{inelast} \ll I_{elast.}$

and to arrange the successive approximations in these two cases differently . We begin with :

a . Transfer of translational to internal energy goes easy .

The treatment is then very similar to the mono-atomic case . In zero th approximation , one starts from the complete, local equilibrium distribution:

$$f_i^{(o)} = n\left(\frac{m}{2\pi kT}\right)^{3/2} \frac{1}{\sum_s \exp(-E_s/kT)} \exp\left[-\frac{1}{kT}\left(\frac{1}{2} m (\vec{\xi} - \vec{u})^2 + E_i\right)\right]$$

dependent on the five macroscopic quantities n, $\vec{u}$, and T , which are still functions of x,y,z and t . In this approximation one again gets the Euler equations, only now the therman energy Q depends more complicatedly on T , since the internal specific heat $C_{int.}$ is in general temperature dependent.

In first approximation one also gets the Navier-Stokes equations

but now :

$$\left.\begin{aligned} P_{ij} &= p\delta_{ij} - 2\mu\left(D_{ij} - \tfrac{1}{3} D_{\alpha\alpha}\delta_{ij}\right) - \kappa\delta_{ij} D_{\alpha\alpha} \\ q_i &= -(\lambda_{tr} + \lambda_{int.}) \frac{\partial T}{\partial x_i} \end{aligned}\right\} \quad (4)$$

The only differences with the case of the mono-atomic gas are, that the heat conductivity consists of two parts , and that , as Lorentz said , a new constant , the so-called dilatational viscosity coefficient κ appears . The constants λ_{tr} , λ_{int}, and κ can be expressed in the I^k_{ij} ,which are supposed to be known . They can be in principle calculated if the molecular model and the interaction law are known ,although in practice this is usually a very difficult task . As example I only give the expression for κ:

$$\kappa = \left(\frac{c_{int}}{c_{tot}}\right)^2 \frac{1}{2} \sqrt{\pi mkT} \; \frac{\left(\sum_i e^{-\varepsilon_i} \varepsilon_i\right)^2}{\sum_{ijk\ell} e^{-\varepsilon_i - \varepsilon_j} \iint d\Omega \, d G \, (\Delta\varepsilon)^2 G^3 e^{-G^2} I^k_{ij}} \quad (5)$$

where $\varepsilon_i = E_i/kT$ and $\Delta\varepsilon = \varepsilon_k + \varepsilon_\ell - \varepsilon_i - \varepsilon_j$. Because of the factor $(\Delta\varepsilon)^2$ clearly only the inelastic collisions give a contribution to κ.

Applied to the propagation of sound , in the validity of this first or Navier-Stokes approximation , one does <u>not</u> get any dispersion of the sound. The dilatational viscosity only gives a contribution to the classical (Kirchoff) or "ω^2" absorption . One finds for the absorption coefficient :

$$\frac{\omega^2}{\rho V_o^3} \left\{ \left(\tfrac{4}{3}\mu + \kappa\right) + \frac{mk}{c_p c_v} \lambda \right\} \quad (6)$$

The effect of κ can therefore not be separated from the classical causes of sound absorption through the viscosity and the heat conduction .

b - Transfer of translational to internal energy is difficult .

In this case, it can occur easily that in perturbations of the equilibrium t state the temperature of the translational motion differs from the temperature of the internal degrees of freedom . It is therefore better to describe the state of the gas by six instead of five macroscopic quantities , namely ρ , u_i , $T_{transl.}$ and $T_{int.}$ .By taking the inelastic collisions into account in a one higher approximation than before, one gets in zero^{-th} approximation again the Euler equations together with $DT_{int}/DT = 0$. In first approximation however one does not get the Navies-Stokes equations , but the equations first put forward by Herzfeld and Rice (Phys. Rev. 31 , 691 , 1928) on phenemenological grounds . One obtains , besides the continuity equation and the equation of motion (in which the pressure tensor has now the same form as for a mono-atomic gas , therefore without a dilatational viscosity term), two energy equations , which determine the changes of $T_{tr.}$ and $T_{int.}$ The really new equation is :

$$C_{int.} \frac{\partial T_{int}}{\partial t} = \frac{m \lambda_{int}}{\rho} \Delta T_{int.} + \frac{C_{int}}{\tau} (T_{tr} - T_{int}) \qquad (7.)$$

The physical interpretation is clear . There are again four gas constants λ_{tr} , λ_{int} , μ and τ , which can be expressed in I^k_{ij}. In the λ_{tr} and μ occurs now only the elastic collision cross-section, so that just as in the mono-atomic case one has with good approximation :

$$\lambda_{tr} = \frac{5}{2} \mu C_{tr.}$$

In the relaxation time τ occurs only the inelastic collision cross-section , and the expression for τ is very analogous to the expression (5) for κ. Formally in fact :

$$\kappa = \frac{c^3_{int.}}{c^2_{tot.}} \; n\,k\,T \qquad (8)$$

Applied to the propagation of sound (as done already by Herzfeld and Rice) one gets a dispersion region around the frequency i/ , accompanied by an absorption band , which is well separated from the " " absorption region caused by the viscosity and heat conduction , for which one gets eq.(6) with $k = 0$.

II - Extension to fast varying phenomena . The question is , what to do when the Chapman-Enskog development parameter : $\left(\frac{\nabla(\text{macr.qu})}{\text{macr. qu.}} \right)$ is not more small , as will be the case for Knudsen gases , and for fast varying phenomena (structure of shock waves , propagation of high frequency sound) . Of special interest are the experiments of Greenspan (J.Acoust. Soc. 22 , 56 (1950) on the dispersion and absorption of high frequency (up to 1 mH) sound in Helium at low pressures (about 0.1 mm of Hg.) . Greenspan plots his result as function of the dimension less parameter $G = \frac{\mu\omega}{\rho \cdot V_0^2}$, which is $\sim \frac{\omega}{\omega}$, where the collision frequency $\omega_{coll} \sim \frac{1}{\tau}$. For Maxwell molecules, with L the transport mean free path, one gets $G = \frac{3\pi}{2} \frac{\omega}{\omega_{coll}}$ Biggest value of G is about 10 , so one really has extreme conditions . For G = 10 , the velocity has increased to about 3 V_0 , while for G=2 the absorption coefficient is already 0.5 , so that the intensity drops by

a factor 1/e in 2 cm .

a. One can try to extend the theory by carrying the development further. It is to be noted that the successive order hydrodynamical equations (Euler, Navier-Stokes, Burnett, etc) are all of the first order in the time derivatives of the macroscopic quantities ρ, u_i and T , and are of successive order in the space derivatives . Therefore to solve the initial value problem it is always sufficient to give the initial values of ρ, u_i and T This fact is called the macroscopic causality theorem (Hilbert) . It is paradoxical because from the Boltzmann equation it would appear that one would need the initial value of the whole distribution f, which is much more than the first five moments (in the velocity) ρ , u_i and T . One must say the causality theorem is not a theorem but an " Ansatz " . Starting from an arbitrary initial velocity distribution one must assume that in a short time (of order $L / \bar{v}$) the gas reaches a " normal " state, from which the further temporal development proceeds according to the macroscopic causality theorem (Compare the hydrodynamical stage in the Bogolubov theory , Ch. 23) .

b. There are few practical applications of the Burnett equations . The propagation of sound has been considered by Primakoff and by C.S. Wang Chang . The result can be found simpler in an other way , which we will discuss in a moment . It is clear anyway ,that as soon as the parameter $L \frac{\nabla (\text{macr.qu.})}{\text{macr.qu.}}$ is comparable with one, the Chapman-Enskog development will be of little use .

c. It is good to remember a remark made by Osborne Reynolds . One knows

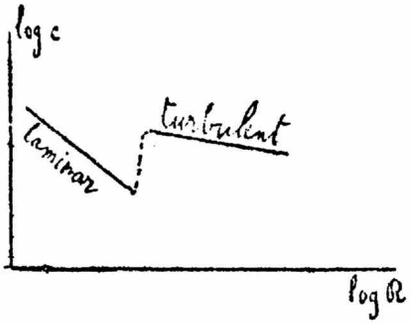
that for viscous and incompressible fluids all dimension less numbers (of drag, lift , etc) can only be function of the Reynolds number $\mathcal{R}$. For instance for the flow through a circular tube the resistance coefficient C , defined by :

$$p_1 - p_2 = C \frac{L}{R} \frac{\rho V^2}{2}$$

(p_1, p_2 are the pressures at the ends of the tube of length L , radius R ; ρ, V are density and average velocity of the fluid) is a function of

$$R = \frac{\rho V R}{\mu}$$

if μ = viscosity coefficient . The function $C(\mathcal{R})$ summarizes all the experiments about the flow ; for low speed one has (Poisemble)



$$C(\mathcal{R}) = \frac{16}{\mathcal{R}}$$

while for large $\mathcal{R}$ one gets the turbulent regime with a different law for C . Now, from the kinetic theory for a gas

$$\mu \sim \rho L v$$

if v is an average molecular velocity . Therefore :

$$\mathcal{R} \sim \frac{V}{v} \cdot \frac{R}{L}$$

is really the product of two dimensionlless numbers . Only if $V/v \ll 1$ and $R/L \gg 1$ will they enter the theory as the product .

In general (at high speeds , or for low pressures (Knudsengas) one must expect that both numbers will enter independent of each other . We will call L/R the Knudsen number ; since the sound velocity is of same

order as the average molecular velocity $\sqrt{v}$ can be taken as the Mach number .

d. In general, one can say that in any non equilibrium situation one can distinguish between :

1. The magnitude of the disturbance from equilibrium , and measured by a " Mach " number, and 2 the scale of the disturbance from equilibrium , which is measured by the Knudsen number . Examples : a propagation of sound, magnitude of the disturbance is determined by the intensity , the scale by the wave length of the sound ; b motion of a sphere through a gas, magnitude of disturbance is determined by the speed (measured by the usual Mach number) , the scale by the size of the sphere comp. to the mean free path ; c heatflux between two parallel plates with temperatures T_1 and T_2 ; magnitude of disturbance is determined by the " Mach" number $(T_2 - T_1) / \frac{1}{2} (T_1 + T_2)$, the scale by the Knudsen number L / d , if d = distance between plates .

It seems to me unlikely , that there exist macroscopic equations , which would be valid for all values of both the Mach and the Knudsen number . It also seems better to make developments in either the Mach or in the Knudsen number , and not in the combination M / K as is done in the Chapman-Enskog development .

e . For all disturbances , one can always write :

$$f = f_o (1 + h (\vec{r}, \vec{\xi}, t)$$

where f_o is the complete (therefore not the local) equilibrium

distribution and h is the disturbance . If $h \ll 1$, one gets the homogeneous equation :

$$\sqrt{\frac{m}{2kT}}\,\frac{\partial h}{\partial t} + c_\alpha \frac{\partial h}{\partial x_\alpha} = n\, \mathcal{I}(h) \qquad (1)$$

with :

$$\mathcal{I}(h) = \frac{1}{\pi^{3/2}} \int d\vec{c}_1 e^{-c_1^2} \int d\Omega g \mathcal{I}(g,\theta)\,(1 + h'_1 - h - h_1)$$

is the linearized collision operator , and where we use again the dimensionless velocity $c_i = \sqrt{\frac{m}{2kT}} \cdot \xi_i$. No assumption is made about the scale of the disturbance .

For example , for the propagation of sound of low intensity but arbitrary wavelength one has to solve (1) , if h has the form :

$$h(\vec{c})\; e^{i(\omega t - \sigma z)}$$

with ω real and $\sigma = \sigma_1 - i\sigma_2$. Developing $h(\vec{c})$ in the eigenfunctions of $\mathcal{I}(h)$:

$$h = \sum a_{r\ell}\, \psi_{r\ell}(\vec{c})$$

(1) becomes the infinite system of linear homogeneous equations :

$$i\sqrt{\frac{m}{2kT}}\;\omega\, a_{r\ell} - i\sigma \sum_{r'\ell'} M_{r\ell;\, r'\ell'}\, a_{r'\ell'} = n\lambda_{r\ell}\, a_{r\ell} \qquad (2)$$

where :

$$M_{r\ell,\, r'\ell'} = \int d\vec{c}\; c_z \psi_{r\ell} \psi_{r'\ell'}$$

is the matrix element of c_z and $\lambda_{r\ell}$ are the eigenvalues of $\mathcal{I}(h)$. For Maxwell molecules , the $\psi_{r\ell}$, $\lambda_{r\ell}$ are known (see Ch. 3) and the $M_{r\ell,\, r'\ell'}$ can easily be evaluated . The condition of solubility of the set (2) is that

the determinant :

$$\| (i\omega\sqrt{\frac{m}{2kT}} - n\lambda_{r\ell}) \delta_{r\ell,r'\ell'} - i\sigma M_{r\ell,r'\ell'} \| = 0 \qquad (3)$$

which gives a relation between ω and σ , which is the dispersion law of the gas . One finds (always for Maxwell molecules , force law k/r^5 , using dimension less frequency $\omega_0 = \omega / n\sqrt{\frac{2k}{m}}$ and propagation constant $\sigma_0 = \sqrt{\frac{kT}{k}} \cdot \frac{\sigma}{n}$ for the beginning of the infinite determinant (3) :

$2r+\ell$	$2r+\ell$	0	1	2		3	
	ℓ	0	1	0	2	1	3
	ℓ	0	0	1	0	1	0
0	0	0	ω_0	$-\frac{1}{\sqrt{2}}\sigma_0$	0	0	0
1	1	0	$-\frac{1}{\sqrt{2}}\sigma_0$	ω_0	$\frac{1}{\sqrt{3}}\sigma_0$	$-\sqrt{\frac{2}{3}}\sigma_0$	0
	0	1	0	$\frac{1}{\sqrt{3}}\sigma_0$	ω_0	0	$\sqrt{\frac{5}{6}}\sigma_0$
2	2	0	0	$-\sqrt{\frac{2}{3}}\sigma_0$	0	$\omega_0 + i\lambda_{02}$	$\frac{2}{\sqrt{15}}\sigma_0$
	1	1	0	0	$\sqrt{\frac{5}{6}}\sigma_0$	$\frac{2}{\sqrt{15}}\sigma_0$	$\omega_0 + i\lambda_{11}$
3	3	0	0	0	0	$-\frac{3}{\sqrt{10}}\sigma_0$	0
							$\omega_0 + i\lambda_{03}$

We were only able to go further by using a breaking - off procedure , in which we were guided by the Chapman-Enskog development . We used successively

3×3, 5×5, 8×8, 11×11 terms . At each stage one has an algebraic relation between ω_0 and σ_0 . One can then treat these relations in two different ways . The first and really consistent way is to develop $\sigma_0 = \sigma_{01} - i\sigma_{02}$ as a function of ω in powers of ω , which amounts to a development in powers of the Greenspan parameter G . In each approximation one retains those powers of G of which the coefficients are not changed in the next approximation . One thus gets :

$$\sigma_1 = \frac{\omega}{V} = \frac{\omega}{V_0}\left[1 - \underset{\text{"}}{\frac{215}{72}}\ G^2 + \underset{\text{"}}{\frac{4,115,101}{27.34.7}}\ G^4 - \ldots\ldots\right]$$

$$2.99 \qquad 56.7$$

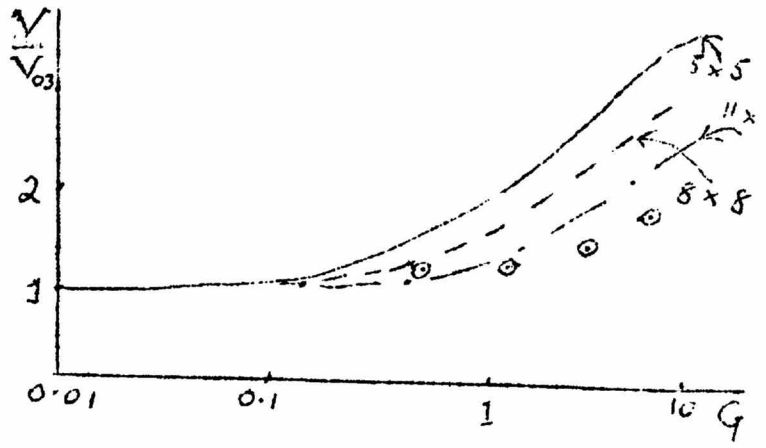
$$\sigma_2 = \frac{\omega}{V_0}\, G\left[\frac{7}{6} - \underset{\text{"}}{\frac{5155}{432}}\ G^2 + 310.5\ G^4 - \ldots\right]$$

$$11.9$$

These series are the same as obtained from the successive order hydrodynamical equations , but the derivation is <u>much</u> simpler . Because ? the expected slow convergence they are of little use for the comparison with Greenspan's experiments .

The second and less consistent way is to solve in each approximation σ as function of ω exactly and compare the successive curves with experiment .

With the 11×11 determinant one then gets approximate agreement with experiment (see figure ; ⊙ are the experimental points.) However



since one has no idea about the convergence of the procedure, this agreement may not mean anything .

- ADDITION TO CH.22 .

From the Born-Green-Kirkwood- Yvon - Bogolubov hierarchy of equations (15) it is possible to derive general hydrodynamic equations , analogous to but more general than Eqs. (3), Ch.17 . Especially from the first equation (15a) of the hierarchy follows (if one includes outside forces) almost immediatly the continuity equation :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{u}) = 0 \qquad (1)$$

$$\rho = mn = mN \int F_1(p,q)\, d\vec{p}$$

$$u_i = \frac{1}{n} \int \frac{p_i}{m} F_1(p,q)\, d\vec{p}$$

and the equation of motion :

$$\rho \frac{D u_i}{D t} = \rho X_i - \frac{\partial P_{i\alpha}}{\partial q_\alpha} - \int d\vec{Q}\, \frac{\partial \phi(|\vec{q}_1 - \vec{q}_2|)}{\partial q_i}\, n_2(\vec{q},\vec{Q}) \qquad (2)$$

with

$P_{ij} = \rho \overline{u_i u_j}$ = kinetic stress tensor and :

$$n_2 (\vec{q}, \vec{Q}) = \frac{N^2}{2} \int\int d\vec{p}\, d\vec{P}\, f_2 (q,p ; Q,P)$$

is the pair density distribition . It is possible to write the last term in (2) as the divergence of the symmetric tensor :

$$P'_{ij} = \frac{1}{2} \int_0^\infty dR R^2 \frac{d\phi(R)}{dR} \iint d\vec{k} \int_0^R d\lambda k_i k_j n_2 (\vec{q} + \vec{k}\lambda - \vec{k}R, \vec{q} + \vec{k}\lambda) \quad (3)$$

where $\vec{k}$ is a unit vector pointing from $\vec{q}$ to $\vec{Q}$. This was first pointed out by Enskog . As a result one can combine the last two terms in (2) to $\partial P_{i\alpha} / \partial q_\alpha$, where $P_{ij} = \delta_{ij} + P'_{ij}$ is the total stress tensor consisting of a kinetic and an intermolecular force part . The physical interpretation of (3) is clear . The formal proof is as follows :

Last term of (2) is of the form :

$$A_i = \int d\vec{Q}\, \frac{\partial}{\partial Q_i} F (\vec{q}, \vec{Q})$$

where F is a scalar and symmetric function of $\vec{q}$ and $\vec{Q}$. Now : $A_i = \frac{\partial S_{i\alpha}}{\partial q_\alpha}$ with :

$$S_{ij} = S_{ji} = \frac{1}{2} \int_0^\infty dR R^2 \iint d\vec{k} \int_0^R d\lambda k_i k_j F (\vec{q} + \vec{k}\lambda - \vec{k}R, \vec{q} + \vec{k}\lambda)$$

Proof : Verify :

$$\frac{\partial S_{i\alpha}}{\partial q_\alpha} = \frac{1}{2} \int_0^\infty dR R^2 \iint d\vec{k} \int_0^R d\lambda k_i k_\alpha \frac{\partial F}{\partial q_\alpha}$$

F is a function of $\vec{q} + \vec{k}\lambda$ and $\vec{k}R$, so that

$$\frac{\partial F}{\partial \lambda} = k_\alpha \frac{\partial F}{\partial q_\alpha}$$

Therefore :

$$\frac{\partial S_{i\chi}}{\partial q_\alpha} = \frac{1}{2} \int_0^\infty dR\, R^2 \iint d\vec{k}\, k_i \left\{ F\left(\vec{q}, \vec{q} + \vec{k}R\right) - F\left(\vec{q} - \vec{k}R, q\right) \right\}$$

In the last integral , one interchanges the two points , which changes the direction of $\vec{k}$. Since F is symmetric one thus gets the same result as the first integral , so that :

$$\frac{\partial S_{i\alpha}}{\partial q_\alpha} = \int_0^\infty dR\, R^2 \iint d\vec{k}\, k_i\, F\left(\vec{q}, \vec{q} + \vec{k}R\right)$$

which is just A_i [put origin of $\vec{Q}$ integration in $\vec{q}$, and use polar coordinates .]

In an analogous way , one can generalize the energy equation (Comp . last of eqs. (3) , Ch 17) to :

$$\rho \frac{D}{Dt}\left(\frac{Q}{\rho} + \bar{\phi}\right) + \operatorname{div} \eta = - P_{\alpha\beta} D_{\alpha\beta} \qquad (4)$$

with $D_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial q_j} + \frac{\partial u_j}{\partial q_i}\right)$, P_{ij} = total stress tensor , and :

$Q = \frac{1}{2}\rho \overline{u_\alpha u_\alpha}$ = density of the kinetic thermal energy .

$$\bar{\phi} = \frac{1}{n_2} \iiint \phi\left(|\vec{q} - \vec{Q}|\right) F_2(q, p ; Q, P)\, d\vec{p}\, d\vec{P}\, d\vec{Q}$$

= potential energy density .

Finally , in (4) enters the total energy flux density η , which just as the P_{ij} consists of two parts : $\eta_i = q_i + q_i$

where $q_i = \frac{1}{2}\rho \overline{u_i u_\alpha u_\alpha}$ = kinetic thermal energy flux density , and

q_i' is the contribution due to the intermolecular forces . One finds :

$$q_i' = \frac{1}{2m}\int d\vec{Q}\,\frac{d\phi}{dR}\,k_i\,k_\alpha \iint d\vec{p}\,d\vec{P}\int_0^R d\lambda\,(p_\alpha - \bar{p}_\alpha)\,F_2\,(\vec{q}+\vec{k}\lambda-\vec{k}R,$$

$$\vec{q}+k\lambda, p, P) + \frac{1}{2m}\int d\vec{Q}\,\phi(R)\iint d\vec{p}\,d\vec{}\,(p_i - \bar{p}_i)$$

$$F_2\,(q, p; Q, P) \qquad (5)$$

with $d\vec{Q} = R^2\,dR\,d\vec{k}$. For the physical interpretation of (5) see Kirkwood .