

Sophus Lie's Approach to Differential Equations

IAP lecture 2006 (S. Helgason)

1 Groups

Let X be a set. A one-to-one mapping φ of X onto X is called a *bijection*. Let $B(X)$ denote the set of all bijections of X onto X .

Let $\varphi, \psi \in B(X)$. We define a product $\varphi\psi$ by $(\varphi\psi)(x) = \varphi(\psi(x)) \quad x \in X$. Then $\varphi\psi \in B(X)$. The identity map $I : X \rightarrow X$ belongs to $B(X)$. If $\varphi \in B(X)$ we define φ^{-1} by $\varphi^{-1}(\varphi(x)) = x$. Then $\varphi^{-1} \in B(X)$ and $\varphi\varphi^{-1} = \varphi^{-1}\varphi = I$. Also, the product $\varphi\psi$ is associative, i.e.,

$$(\varphi\psi)\tau = \varphi(\psi\tau).$$

In fact,

$$(\varphi\psi)\tau(x) = (\varphi\psi)(\tau(x)) = \varphi(\psi(\tau(x))) = \varphi(\psi\tau(x)).$$

In general, $\varphi\psi \neq \psi\varphi$. For this consider the example $X = \{1, 2, 3\}$. Then $B(X)$ consists of the permutation of 1, 2, 3. Let $\varphi : X \rightarrow X$ fix 1 but exchange 2 and 3. Let $\psi : X \rightarrow X$ fix 3 but exchange 1 and 2. Then $\varphi\psi(1) = 3$ but $\psi\varphi(1) = 2$ so $\varphi\psi \neq \psi\varphi$.

The set $B(X)$ is a prototype of a group. A *group* is a set G such that for any $a, b, c \in G$ is an associated new element $ab \in G$ such that

$$a(bc) = (ab)c.$$

One also assumes the existence of an element $e \in G$ such that $ea = ae$ for all $a \in G$. One also assumes that for each $a \in G$ there exists an element $a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = e$.

Example. $B(X)$ is a group.

If G is a group a subset $H \subset G$ is called a *subgroup* if $h \in H, k \in H$ implies $hk, h^{-1} \in H$. Then H is a group. H is called a *normal subgroup* if $gHg^{-1} \subset H$ for all $g \in G$. If H is a normal subgroup the family of cosets $gH = \{gh : h \in H\}$ can be made into a group, denoted G/H , (the *factor group*) by the product definition

$$g_1H \cdot g_2H = g_1g_2H \quad (\text{well-defined}).$$

G is said to be *abelian* if $xy = yx$ for all $x, y \in G$. Let G' denote the subgroup of G generated by all commutators $xyx^{-1}y^{-1}$ ($x, y \in G$). The

G' is a normal subgroup and G/G' is abelian. We can form the sequence

$$G \subset G' \supset (G')' \supset \dots .$$

G is said to be *solvable* if this ends with e .

As mentioned, $B(X)$ is a group and its subgroups are called *transformation groups* of X .

2 Polynomials

It is well known from algebra that each polynomial equation $p(x) = 0$ of degree 1, 2, 3, 4 is solvable by radicals. Through Abel's work one knows that this is no longer true for $p(x)$ of degree 5.

In Galois' theory the theory becomes deeper. Here to each polynomial $p(x)$ is associated a certain finite group, a subgroup of the group of permutations of the roots. *The solvability of this group is equivalent to $p(x) = 0$ being solvable by radicals.* The equation $x^5 - x - \frac{1}{3} = 0$ can be shown to have Galois group, the full permutation group S_5 of five letters which can be shown not to be solvable (in contrast to S_i ($i < 5$)). Thus the equation is not solvable by radicals.

3 Lie's Program.

Inspired by Galois' theory, Lie got the idea of doing something in this spirit for differential equations. First let us look at a first order equation

$$(3.1) \quad \frac{dy}{dx} = F(x, y),$$

a *solution* being by definition a function $y = u(x)$ such that $u'(x) = F(x, u(x))$.

Thus a solution is a curve in \mathbf{R}^2 (an *integral curve*). A transformation $T \in B(X)$ ($X = \mathbf{R}^2$) is said to leave the equation (3.1) *stable* if it permutes the integral curves.

Special case. $F(x, y) = g(x)$. Then (3.1) becomes

$$\frac{dy}{dx} = g(x),$$

with the solutions

$$u(x) = \int_0^x g(t) dt + C \quad C = \text{constant.}$$

The integral curves are all parallel so each translation $T_t : (x, y) \rightarrow (x, y + t)$ leaves the equation stable. Note also that

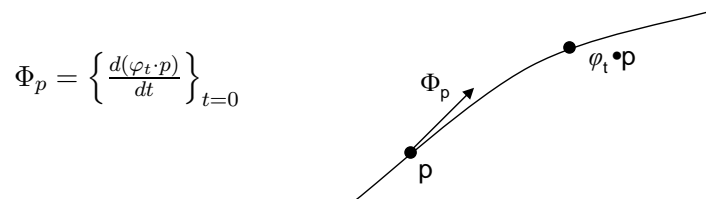
$$T_{s+t} = T_s T_t,$$

that is, $t \rightarrow T_t$ is a 1-parameter group. More generally, consider the differential equation

$$(3.2) \quad \frac{dy}{dx} = \frac{Y(x, y)}{X(x, y)} \quad (\text{Lie's notation})$$

and assume we have a 1-parameter group $\varphi_t (t \in \mathbf{R})$ of differentiable bijections of \mathbf{R}^2 leaving (3.2) stable.

Consider the *vector field* on \mathbf{R}^2



$$\Phi_p = \left\{ \frac{d(\varphi_t \cdot p)}{dt} \right\}_{t=0}$$

Here Φ_p is the tangent vector to the orbit $\varphi_t \cdot p$ at $p \in \mathbf{R}^2$. Thinking of a vector at $p = (x, y)$ as a directional derivative we write the vector field in the form

$$(3.3) \quad \Phi_p = \Phi_{x,y} = \xi(x, y) \frac{\partial}{\partial x} + \eta(x, y) \frac{\partial}{\partial y}.$$

Theorem 3.1 (Lie (1874)). *Assume $\varphi_t (t \in \mathbf{R}^2)$ is a 1-parameter group leaving (3.2) stable. The \exists function $U(x, y)$ such that*

$$(3.4) \quad \frac{\partial U}{\partial x} = \frac{-Y}{X\eta - Y\xi} \quad , \quad \frac{\partial U}{\partial y} = \frac{X}{X\eta - Y\xi}$$

and $U(x, y) = \text{const}$ is the solution to (3.2).

A proof will be indicated later.

The statement is equivalent to the statement that

$$\frac{X dy - Y dx}{X\eta - Y\xi}$$

is an exact differential

$$\frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy$$

or equivalently that $(X\eta - Y\xi)^{-1}$ is an *integrating factor* for the equation

$$X dy - Y dx = 0.$$

The theorem means that the solution to (3.2) can be found by integrating the equations (3.4) which is done in Calculus.¹

We sketch the method. We write (3.4) in the form

$$\frac{\partial U}{\partial x} = M(x, y), \quad \frac{\partial U}{\partial y} = N(x, y).$$

Assuming U a solution we have

$$U(x, y) = \int_a^y N(x, z) dz + g(x)$$

for some function $g(x)$. Then

$$M(x, y) - \frac{\partial}{\partial x} \int_a^y N(x, z) dz = g'(x).$$

Since $\partial/\partial y$ of the left hand side equals

$$M_y - N_x = 0$$

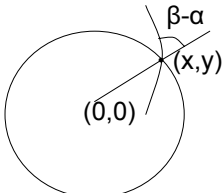
the function $g(x)$ does exist and the formula for $U(x, y)$ gives a solution. In the old days this was expressed: Under the assumption of the theorem, equation (3.2) can be solved by *quadratures*. This was to emphasize the analogy with solving an algebraic equation in radicals when the Galois group was abelian. Later we discuss the solvable case.

Lie's remarkable theorem seems unfortunately ignored in recent books on ordinary differential equations. Before a proof let us consider a few examples. A favorite of mine is the equation

$$(3.5) \quad \frac{dy}{dx} = \frac{y + x(x^2 + y^2)}{x - y(x^2 + y^2)}$$

¹At least in the good old Thomas.

which at first looks rather formidable. However, we can write it

$$(3.6) \quad \frac{\frac{dy}{dx} - \frac{y}{x}}{1 + \frac{y}{x} \frac{dy}{dx}} = x^2 + y^2$$


The slope of the ray from $(0,0)$ to (x,y) is $\frac{y}{x} = \tan \alpha$ and the slope of the tangent to the integral curve through (x,y) is $\frac{dy}{dx} = \tan \beta$. Since

$$\tan(\beta - \alpha) = \frac{\tan \beta - \tan \alpha}{1 + \tan \alpha \tan \beta}$$

(3.6) states that

$$\tan(\beta - \alpha) = x^2 + y^2.$$

This means that the angle $\beta - \alpha$ is constant as (x,y) varies on a circle with center $(0,0)$. Thus each rotation

$$(3.7) \quad \varphi_t : (x,y) \rightarrow (x \cos t - y \sin t, x \sin t + y \cos t)$$

maps each integral curve into another integral curve, in other words leaves the equation (3.5) stable. Also the rotations φ_t form a group ($\varphi_{t+s} = \varphi_t \varphi_s$). Here we have from (3.7)

$$\Phi_p = \left(\frac{d(\varphi_t \cdot p)}{dt} \right)_{t=0} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}$$

so by Lie's theorem

$$[(x - y(x^2 + y^2))x - (y + x(x^2 + y^2))(-y)]^{-1} = (x^2 + y^2)^{-1}$$

is an integrating factor. Also

$$\frac{X dy - Y dx}{x^2 + y^2} = \left(\frac{x}{x^2 + y^2} - y \right) dy - \left(\frac{y}{x^2 + y^2} + x \right) dx$$

and by inspection we see that this is

$$\frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy$$

for

$$U(x, y) = \text{Arctan}\left(\frac{y}{x}\right) - \frac{x^2 + y^2}{2}.$$

The solution $U(x, y) = c$ can be written

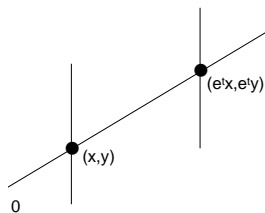
$$y = x \tan\left(\frac{1}{2}(x^2 + y^2) + c\right).$$

Exercise. (an example from Lie's paper)

Consider the differential equation

$$(3.8) \quad \frac{dy}{dx} = f\left(\frac{y}{x}\right).$$

On each ray



consider the tangent to the integral curve through (x, y) . At the point $(e^t x, e^t y)$ the slope of the tangent is the same. Thus the map

$$\varphi_t : (x, y) \rightarrow (e^t x, e^t y)$$

leaves the equation stable. Also, $\varphi_{t+s} = \varphi_t \varphi_s$. Here

$$\Phi_p = \left\{ \frac{d(\varphi_t \cdot p)}{dt} \right\}_{t=0} = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}.$$

By the theorem $-f\left(\frac{y}{x}\right) dx + dy$ has integrating factor $(y - f\left(\frac{y}{x}\right)x)^{-1}$, in other words

$$\frac{\partial}{\partial x} \left(\frac{1}{y - f\left(\frac{y}{x}\right)x} \right) = \frac{\partial}{\partial y} \left(\frac{-f\left(\frac{y}{x}\right)}{y - f\left(\frac{y}{x}\right)x} \right)$$

and

$$\frac{-f\left(\frac{y}{x}\right)}{y - f\left(\frac{y}{x}\right)x} dx + \frac{1}{y - f\left(\frac{y}{x}\right)x} dy = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy.$$

The solution is $U(x, y) = c$.

To find U we know from the last equation that

$$U(x, y) = \int_1^y \frac{dz}{z - f\left(\frac{z}{x}\right)} + g(x)$$

and as explained before the relation

$$\frac{\partial U}{\partial x} = \frac{-f\left(\frac{y}{x}\right)}{y - f\left(\frac{y}{x}\right)x}$$

gives a formula for $g'(x)$ in terms of x alone. Thus the solution to (3.8) is given by two quadratures.

Exercise. Carry this out for $f(z) = z^2 + 2z$. Solution is $y = \frac{x^2}{c-x}$.

We still do not have a differential equation analog to the algebraic theorem for solvable Galois group. For this we consider groups depending on more parameters. Let $X = \mathbf{R}^2$ and let G be the subgroup of $B(X)$ preserving distances and orientation. If $\sigma \in G$ let t be the translation such that $\sigma \cdot 0 = t \cdot 0$. Then $t^{-1}\sigma \cdot 0 = 0$ so $t^{-1}\sigma$ is a rotation k around the origin. Let $\theta(\sigma)$ be the angle between the x-axis ℓ and $k \cdot \ell$ and let $t = (x(\sigma), y(\sigma))$. Then

$$\sigma : \begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x(\sigma) \\ y(\sigma) \end{pmatrix} + \begin{pmatrix} \cos \theta(\sigma) & \sin \theta(\sigma) \\ -\sin \theta(\sigma) & \cos \theta(\sigma) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

A simple computation shows

$$\begin{aligned} x(\sigma\tau^{-1}) &= x(\sigma) - x(\tau) \cos(\theta(\sigma) - \theta(\tau)) + y(\tau)(\sin \theta(\sigma) - \theta(\tau)) \\ y(\sigma\tau^{-1}) &= y(\sigma) - x(\tau) \sin(\theta(\sigma) - \theta(\tau)) - y(\tau) \cos(\theta(\sigma) - \theta(\tau)) \\ \theta(\sigma\tau^{-1}) &= \theta(\sigma) - \theta(\tau) \pmod{2\pi}. \end{aligned}$$

Thus the elements of G are parameterized by three parameters such that the parameters of product and inverse are smooth functions of the parameters of the factors. This suggests the definition of a *Lie group*.

What is a Lie algebra?

Consider a transformation group of \mathbf{R}^n depending effectively on r parameters:

$$\begin{aligned} T : (x_1, \dots, x_n) &\rightarrow (x'_1, \dots, x'_n) \text{ where} \\ x'_i &= f_i(x_1, \dots, x_n; t_1, \dots, t_r), \\ T = I &\text{ for } (t_1, \dots, t_r) = (0, \dots, 0). \end{aligned}$$

Assume that if S is the transformation corresponding to the parameters (s_1, \dots, s_r) then

$$TS^{-1} : x'_i = f_i(x_1, \dots, x_n; u_1, \dots, u_r)$$

and the u_i are smooth functions of the t_i, s_j above. Generalizing Φ_p above Lie defined the vector fields on \mathbf{R}^n :

$$(3.9) \quad T_k = \sum_{i=1}^n \left(\frac{\partial f_i}{\partial t_k} \right)_{t=0} \frac{\partial}{\partial x_i}.$$

Now if X and Y are any vector fields on \mathbf{R}^n they map functions into functions so $X \circ Y$ and $Y \circ X$ are well defined maps of functions. So is the bracket

$$[X, Y] = X \circ Y - Y \circ X$$

and it is easily seen to involve only first order derivatives (the second derivatives cancel) so $[X, Y]$ is another vector field.

Lie proved the fundamental fact that the vector fields (3.9) satisfy

$$(3.10) \quad [T_k, T_\ell] = \sum_{p=1}^r c_{k\ell}^p T_p,$$

where the coefficients $c_{k\ell}^p$ are *constants*, satisfying

$$(3.11) \quad c_{k\ell}^p = -c_{\ell k}^p, \quad \sum_{q=1}^r (c_{kq}^p c_{\ell m}^q + c_{mq}^p c_{k\ell}^q + c_{\ell q}^p c_{mk}^q) = 0.$$

If we put $\mathfrak{g} = \{X = \sum_1^r a_p T_p \mid a_p \in \mathbf{R}\}$, relations (3.11) can be stated: For $X, Y, Z \in \mathfrak{g}$ we have

$$(3.12) \quad [X, Y] = -[Y, X]$$

$$(3.13) \quad [X, [Y, Z]] + [Y, [Z, X]] + [[Z, X], Y] = 0.$$

A vector space with a rule of composition $(X, Y) \rightarrow [X, Y]$ satisfying (3.12) – (3.13) is called a *Lie algebra*. Thus \mathfrak{g} is a Lie algebra.

Lie also proved a converse that to a finite-dimensional Lie algebra there exists (locally) a transformation group corresponding to it.

A Lie algebra \mathfrak{g} is *abelian* if $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$. Putting $\mathcal{D}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and $\mathcal{D}^S \mathfrak{g} = \mathcal{D}(\mathcal{D}^{S-1} \mathfrak{g})$, \mathfrak{g} is said to be *solvable* if $\mathcal{D}^S \mathfrak{g} = 0$ for some S .

We can now state an analog to the solvability of $p(x) = 0$ in radicals in terms of the Galois group.

Theorem 3.2. *Suppose the system of differential equations*

$$(3.14) \quad \frac{dy_j}{dx} = f^j(x, y_1, \dots, y_r) \quad 1 \leq j \leq r$$

has a solvable r -dimensional stability group in (x, y_1, \dots, y_r) with r -dimensional orbits. Then the solution can be found by repeated quadratures, explicitly given by the group.

For the proof a good reference is P. Olver, Applications of Lie Groups to Differential Equations (Springer 1986).

The case $r = 1$ is the one in Theorem 3.1. Let us verify the theorem for an example (from Olver's book).

Example. Consider the differential equation

$$(3.15) \quad x^2 \frac{d^2 y}{dx^2} = f\left(x \frac{dy}{dx} - y\right).$$

Putting $z = \frac{dy}{dx}$ we can put the equation in the form (3.14)

$$(3.16) \quad x^2 \frac{dz}{dx} = f(xz - y), \quad \frac{dy}{dx} = z.$$

Here the transformations

$$T_{s,t} : (x, y, z) \rightarrow (sx, y + tx, \frac{z}{s} + \frac{t}{s}) \quad s > 0, t \in \mathbf{R}$$

leave the system (3.16) stable. Also

$$T_{\sigma,\tau} \circ T_{s,t} = T_{\sigma s, t + \tau s}.$$

The one-parameter subgroups $T_{s,0}$ and $T_{1,t}$ generate the vector fields

$$X_1 = x \frac{\partial}{\partial x} - z \frac{\partial}{\partial z}, \quad X_2 = x \frac{\partial}{\partial y} + \frac{\partial}{\partial z}.$$

Then $[X_1, X_2] = X_2$ so the Lie algebra $\mathbf{R}X_1 + \mathbf{R}X_2$ is solvable. According to the theorem, (3.15) is solvable by quadratures. Let us verify this. We put $w = xz - y$. Then $\frac{dw}{dx} = x \frac{dz}{dx}$ so

$$x \frac{dw}{dx} = f(w)$$

giving

$$\int \frac{dw}{f(w)} = \log |x| + C.$$

Writing this in the form $w = g(x)$ we arrive at the equation

$$\frac{dy}{dx} - \frac{y}{x} = \frac{g(x)}{x}.$$

This has the symmetry group $(x, y) \rightarrow (x, y + tx)$ with vector field $x \frac{\partial}{\partial y}$.

By Theorem 3.1 x^{-1} is an integrating factor for

$$dy - \frac{y + g(x)}{x} dx = 0$$

that is

$$\frac{dy}{x} - \frac{y + g(x)}{x^2} dx = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy.$$

Thus

$$\frac{\partial U}{\partial y} = \frac{1}{x} \quad \text{so} \quad U = \frac{y}{x} + h(x)$$

$$\frac{\partial U}{\partial x} = -\frac{y}{x^2} + h'(x) = -\frac{y}{x^2} - \frac{g(x)}{x^2}.$$

Thus

$$h(x) = -\int_1^x \frac{g(t)}{t^2} dt$$

and the solution to (3.15) takes the form

$$y = -x \int_1^x \frac{g(t)}{t^2} dt + cx.$$

4 The Heat Equation

This is the equation

$$(4.1) \quad \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}.$$

Here a transformation $(x, t, u) \rightarrow (x_1, t_1, u_1)$ leaves (4.1) stable if it permutes the solutions. Lie determined the stability group (cf. Olver's book above). It is a six-dimensional group times an (uninteresting) infinite-dimensional group. One 1-parameter subgroup is quite interesting. It gives the result that if $f(x, t)$ is a solution to (4.1) then so is the function

$$u(x, t) = \frac{1}{(1 + 4st)^{1/2}} e^{\frac{-sx^2}{1+4st}} f\left(\frac{x}{1 + 4st}, \frac{t}{1 + 4st}\right).$$

5 Proof of Lie's Theorem

So far as I know this interesting theorem does not occur in most recent books on ordinary differential equations. Older proofs seem a bit obscure (but take a look at Lie's original proof (Collected Works, Vol. 3)). The proof in Olver's book is clean and rigorous but contained in a longer theory of prolongations.

Below is a short proof. Suppose φ_t is a 1-parameter group leaving the equation

$$(5.1) \quad \frac{dy}{dx} = \frac{Y(x, y)}{X(x, y)}$$

stable. If $U(x, y) = c$ is a solution we have with $\varphi_t(x, y) = (x_t, y_t)$,

$$U(x_t, y_t) = c(t) \quad (\text{all } t)$$

so

$$\frac{\partial U}{\partial x} \frac{dx_t}{dt} + \frac{\partial U}{\partial y} \frac{dy_t}{dt} = c'(t)$$

and by (3.3)

$$(5.2) \quad \frac{\partial U}{\partial x} \xi + \frac{\partial U}{\partial y} \eta = c'(0).$$

Secondly

$$\frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} \frac{dy}{dx} = 0$$

so

$$(5.3) \quad \frac{\partial U}{\partial x} X + \frac{\partial U}{\partial y} Y = 0.$$

If $c'(0) \neq 0$ we can normalize U such that $c'(0) = 1$. Then (5.2) and (5.3) imply

$$\frac{\partial U}{\partial x} = \frac{-Y}{X\eta - Y\xi}, \quad \frac{\partial U}{\partial y} = \frac{X}{X\eta - Y\xi}$$

so $(X\eta - Y\xi)^{-1}$ is an integrating factor for $X dy - Y dx = 0$ as claimed.

On the other hand, if $c'(0) = 0$, (5.2)–(5.3) imply $dy/dx = \eta/\xi$ so the integral curves are just the orbits of φ_t .

6 Other Contacts with Group Theory. An Example.

Many of you are familiar with the mean-value theorem for harmonic functions, that is solutions to Laplace's equation

$$Lu = \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) u = 0.$$

The theorem states that the solutions are characterized by

$$(6.1) \quad u(x_1, x_2) = \frac{1}{2\pi} \int_0^{2\pi} u(x_1 + r \cos \theta, x_2 + r \sin \theta) d\theta$$

for all x_1, x_2 and all $r \geq 0$. Geometrically this means that for each $p \in \mathbf{R}^2$, $r > 0$, we have

$$(6.2) \quad u(p) = (M^r u)(p),$$

the mean value of u on a circle with center p and radius r . (Exercise: What is the 1-dimensional version of this result?)

Invoking the above group G of isometries of \mathbf{R}^2 let x denote the translation by (x_1, x_2) , let $k = k(\theta)$ the rotation by θ and y the translation by $(r, 0)$. Then (6.1) can be stated

$$(6.3) \quad u(x \cdot 0) = \frac{1}{2\pi} \int_0^{2\pi} u(xky \cdot 0) dk.$$

A Riemannian manifold X with distance function d has an analog to L above, the *Laplace-Beltrami operator*. We assume X is two-point homogeneous, that is given any two pairs (p, q) , (p', q') in X with $d(p, q) = d(p', q')$ there exists an element g from the isometry group $I(X)$ such that $g \cdot p = p'$, $g \cdot q = q'$. Fix $o \in X$ and let K be the subgroup of $I(X)$ fixing o . Then the solutions to $Lu = 0$ are characterized by

$$(6.4) \quad u(x \cdot o) = \int_K u(xky \cdot o) dk \quad x, y \in I(X),$$

dk being the Haar measures on K (Godement).

Since the set $\{xky \cdot o : k \in K\}$ is the sphere in X with center x and radius $d(o, y)$ formula (6.4) can also be written

$$u(p) = (M^r u)(p) \quad p \in X, r \geq 0$$

in exact analogy with (6.2).