# An Algebraic Theory of Integration Methods 

By J. C. Butcher


#### Abstract

A class of integration methods which includes Runge-Kutta methods, as well as the Picard successive approximation method, is shown to be related to a certain group which can be represented as the family of real-valued functions on the set of rooted trees. For each integration method, a group element is defined corresponding to it and it is shown that the numerical result obtained using the method is characterised by this group element. If two methods are given, then a new method may be defined in such a way that when it is applied to a given initial-value problem the result is the same as for the successive application of the given methods. It is shown that the group element for this new method is the product of the group elements corresponding to the given methods. Various properties of the group and certain of its subgroups are examined. The concept of order is defined as a relationship between group elements.


1. Introduction. In the study of Runge-Kutta methods for the integration of ordinary differential equations, a complicated set of algebraic conditions arises which must be satisfied for a method to have some specified order. These algebraic conditions can be written as a set of equations in which the left-hand sides are certain polynomials in the coefficients of the method and the right-hand sides are certain rational numbers. Discussions of the properties of general Runge-Kutta methods and the derivation of particular methods can typically be formulated in such a way that the properties of these polynomial equations come under close scrutiny. In the author's papers on this subject, a combinatorial interpretation of the equations has been used. That is, use has been made of the relationship between the graphs known as arborescences (or rooted trees) and the forms of the various equations.

In this paper, the same point of view is taken and it is shown that, in a certain sense, a given method is characterised by a real-valued function on the set of rooted trees (the word "trees" will be used as an abbreviation for "rooted trees" for the rest of this paper). In fact, the images of the various trees describe the method in much the same way as moments describe a measure in integration theory.

For the purposes of this paper, it is convenient to generalise considerably the concept of a Runge-Kutta method. Not only will it not be assumed that the method is explicit but it will not be assumed that the number of stages in the method is finite. One consequence of this is that the theoretical solution to a differential equation itself may be thought of as being produced by a particular method (the Picard method). Also, it becomes possible to study the properties of the various methods in terms of a certain group that arises naturally. To say a given method is of order $n$, for example, will be equivalent to saying that the method and the Picard method have group elements in the same cosets with respect to a certain normal subgroup.

Received November 26, 1968, revised August 31, 1970.
AMS 1970 subject classifications. Primary 65L05; Secondary 05C05, 34A45, 34A50.
Key words and phrases. Runge-Kutta methods, initial-value problems, Picard method, order of methods, group, graph, rooted tree.

In the next section of this paper, certain notations and properties of trees will be developed. Section 3 will be concerned with a formulation of the general type of initial-value problems we are considering in this paper and Section 4 will prove some analytical results necessary for the later sections. In Section 5, we shall prove the basic result that the tree functions referred to above characterise the integration method and in Section 6, a group which is relevant to the theory is introduced and some of its properties are studied there and in Section 7.

Applications of the various theoretical results are not considered in this paper, but an outline of such applications is given in Section 8. The appendix, Section 9, gives various tables for reference purposes.
2. Some Properties of Trees. Consider a nonempty finite set $\sigma$ and a binary relation $\rho$ on $\sigma$. The pair ( $\sigma, \rho$ ) will be called a labelled graph and the elements of $\sigma$ will be known as labels. $(\sigma, \rho)$ and $\left(\sigma^{\prime}, \rho^{\prime}\right)$ will be said to be isomorphic if there is a bijection $\varphi: \sigma \rightarrow \sigma^{\prime}$ such that $\rho^{\prime}$ contains $(\varphi(x), \varphi(y))$ if and only if $\rho$ contains $(x, y)$. We will think of the class of all labelled graphs isomorphic to a given labelled graph as the graph corresponding to the given labelled graph. We will speak of a labelled graph as a labelling of the graph corresponding to it. Usually, we will not distinguish explicitly between a graph and a labelling of it.

Table 2.1

| $\sigma$ | $\rho$ |  |
| :---: | :---: | :---: |
| \{0\} | \{ \} | 0. $\quad \tau$ |
| \{0, 1\} | $\{(0,1)\}$ | ${ }_{00}^{1 p} \pi \tau$ |
| \{0, 1, 2\} | $\{(0,1),(1,2)\}$ |  |
| $\{0,1,2\}$ | $\{(0,1),(0,2)\}$ | ${ }_{0}^{1}{ }^{2} \quad \pi \cdot \tau$ |
| $\{0,1,2,3\}$ | $\{(0,1),(1,2),(2,3)\}$ |  |
| \{0, 1, 2, 3\} | $\{(0,1),(0,2),(2,3)\}$ | $\int_{0}^{10} \int_{\tau \tau}^{3 \rho} \tau \tau=(\tau \cdot \tau \tau) \tau$ |
| \{0, 1, 2, 3\} | $\{(0,1),(1,2),(1,3)\}$ | $\int_{0}^{20} \int_{\tau(\tau \tau \cdot \tau)}^{3}$ |
| $\{0,1,2,3\}$ | $\{(0,1),(0,2),(0,3)\}$ | $(\tau \tau \cdot \tau) \tau$ |

If $\gamma_{1}=\left(\sigma_{1}, \rho_{1}\right)$ and $\gamma_{2}=\left(\sigma_{2}, \rho_{2}\right)$ are graphs such that $\sigma_{1} \cap \sigma_{2}=\varnothing$ and such that $\sigma_{1} \cup \sigma_{2}=\sigma, \rho_{1} \cup \rho_{2}=\rho$, where $\gamma=(\sigma, \rho)$ is a graph, then we call $\left\{\gamma_{1}, \gamma_{2}\right\}$ a disconnection of $\gamma$. If $\gamma$ has no disconnection, then we say that $\gamma$ is a connected graph.

A connected graph $\gamma=(\sigma, \rho)$ is called a tree (as has been already pointed out, we are not using this word in the usual way) if there is an element $r$ of $\sigma$ ( $r$ will be called the root of $\gamma$ ) such that for each $y \in \sigma \backslash\{r\}$ there is a unique $x \in \sigma$ such that $(x, y) \in \rho$ and there is no $x \in \sigma$ such that $(x, r) \in \rho$. The order of a tree is the number of elements in $\sigma$. We denote by $T$ the set of all trees. In Table 2.1, the eight trees of order not exceeding 4 will be listed together with a diagrammatic representation in which the elements of $\sigma$ are represented by points and the elements of $\rho$ are represented by arcs. The root in each case is labelled with the symbol 0 . The final column will be explained later.

We now define the product of two trees. Let $\left(\sigma_{1}, \rho_{1}\right)$ and $\left(\sigma_{2}, \rho_{2}\right)$ be labellings of two trees $\gamma_{1}$ and $\gamma_{2}$ such that $\sigma_{1}$ and $\sigma_{2}$ are disjoint and let $r_{1}$ and $r_{2}$ be the roots of ( $\sigma_{1}, \rho_{1}$ ) and ( $\sigma_{2}, \rho_{2}$ ). We define the product $\gamma_{1} \cdot \gamma_{2}$ as the tree which may be labelled as

$$
\left(\sigma_{1} \cup \sigma_{2}, \rho_{1} \cup \rho_{2} \cup\left\{\left(r_{1}, r_{2}\right)\right\}\right)
$$

It is easy to see that this is a labelled tree with root $r_{1}$. If the order of a tree is defined as the number of elements in $\sigma$ for a labelling ( $\sigma, \rho$ ), then we see that the order of a product is the sum of the orders of the factors. Also, we see that if $\gamma_{1}, \gamma_{2}, \gamma_{3}$ are trees then

$$
\begin{equation*}
\left(\gamma_{1} \cdot \gamma_{2}\right) \cdot \gamma_{3}=\left(\gamma_{1} \cdot \gamma_{3}\right) \cdot \gamma_{2} \tag{2.1}
\end{equation*}
$$

We now prove a factorisation principle in terms of this product.
Lemma 2.1. If $t$ is a tree of order greater than 1 then there are trees $u$, $v$ such that $t=u \cdot v$.

Proof. We take $(\sigma, \rho)$ to be a labelling of $t$ and we shall prove that there are $\left(\sigma_{1}, \rho_{1}\right)$ and ( $\sigma_{2}, \rho_{2}$ ) such that $\sigma_{1} \cap \sigma_{2}=\varnothing, \sigma_{1} \cup \sigma_{2}=\sigma, \rho=\rho_{1} \cup \rho_{2} \cup\left\{\left(r_{1}, r_{2}\right)\right\}$, where $r_{1}, r_{2}$ are the roots of $\left(\sigma_{1}, \rho_{1}\right)$ and ( $\sigma_{2}, \rho_{2}$ ). This will be proved by induction on the order of $t$. For $t$ of order 2 , we have if $(\sigma, \rho)=(\{0,1\},\{(0,1)\})$ then $\left(\sigma_{1}, \rho_{1}\right)=$ ( $\{0\},\{ \}$ ) and $\left(\sigma_{2}, \rho_{2}\right)=(\{1\},\{ \})$. We now suppose the result to be true for orders less than $n$ where $n>2$ and prove the result for $t$ of order $n$.

Since $\rho^{-1}$ is a function on $\sigma \backslash\{r\}$ to $\sigma$, where $r$ is the root of ( $\sigma, \cdot \rho$ ), and since $\sigma$ has 1 more element than $\sigma \backslash\{r\}, \rho^{-1}$ is not surjective. Hence, there is an $x \in \sigma$ such that there is no $y \in \sigma$ such that $(x, y) \in \rho$. Let $z$ be the unique element in $\sigma$ such that $(z, x) \in \rho$. We define $\left(\sigma^{\prime}, \rho^{\prime}\right)=(\sigma \backslash\{x\}, \rho \backslash\{(z, x)\})$. It is easy to see that $\left(\sigma^{\prime}, \rho^{\prime}\right)$ is a tree with root $r$ and order $n-1$. Let $\left(\sigma_{1}^{\prime}, \rho_{1}^{\prime}\right),\left(\sigma_{2}^{\prime}, \rho_{2}^{\prime}\right)$ be formed corresponding to a factorisation of ( $\sigma^{\prime}, \rho^{\prime}$ ). We distinguish two cases:

Case I. $z \in \sigma_{1}^{\prime}$. In this case we define $\left(\sigma_{1}, \rho_{1}\right)=\left(\sigma_{1}^{\prime} \cup\{x\}, \rho_{1}^{\prime} \cup\{z, x\}\right),\left(\sigma_{2}, \rho_{2}\right)=$ $\left(\sigma_{2}^{\prime}, \rho_{2}^{\prime}\right)$.

Case II. $z \in \sigma_{2}^{\prime}$. We define $\left(\sigma_{1}, \rho_{1}\right)=\left(\sigma_{1}^{\prime}, \rho_{1}^{\prime}\right),\left(\sigma_{2}, \rho_{2}\right)=\left(\sigma_{2}^{\prime} \cup\{x\}, \rho_{2}^{\prime} \cup\{z, x\}\right)$.
In each case, we see that $\left(\sigma_{1}, \rho_{1}\right)$ and $\left(\sigma_{2}, \rho_{2}\right)$ are trees satisfying the required conditions.

Let $\tau$ denote the unique tree with order 1 . Then, using Lemma 2.1, we see that all trees are generated using $\tau$ and the product that has been defined. We also see that an induction principle can be used for recursive definitions of functions of
trees and in the proofs of certain results. The principle is that if a result is true for $\tau$ and for $t \cdot u$ whenever it is true for $t, u \in T$, then the result is true for all trees. Stated formally, this is the next result.

Lemma 2.2. If $U$ is a subset of $T$ with the properties that $\tau \in U$ and $t \cdot u \in U$ for all $t, u \in U$ then $U=T$.

Proof. We shall prove by induction on $n$ that every tree of order $n$ is a member of $U$. For $n=1$, we have only the tree $\tau$ and $\tau \in U$. For $n>1$, we assume the result for orders less than $n$. If $t$ is of order $n$, then, by Lemma 2.1, there are trees $u, v$ such that $t=u \cdot \sigma$. But by the additive property of orders, $u, v$ are of order less than $n$ so that $u, v \in U$. Hence, $t=u \cdot v \in U$.

At this point, we make a comment on notation. The product of trees which we have denoted by . will also be denoted by juxtaposition. We will use the two notations interchangeably except that we conventionally suppose the juxtaposition operations to be performed first. This convention is intended to reduce the number of times that parentheses must be used. For example, $\left(t_{1} t_{2} \cdot t_{3} t_{4}\right) t_{5} \cdot t_{6} t_{7}$ will denote the same tree as $\left(\left(\left(t_{1} \cdot t_{2}\right) \cdot\left(t_{3} \cdot t_{4}\right)\right) \cdot t_{5}\right) \cdot\left(t_{6} \cdot t_{7}\right)$. Using this notational device, the various trees of Table 2.1 are shown in the last column written in the form of products generated by $\tau$. Note that the trees $\tau \tau \cdot \tau \tau$ and $(\tau \cdot \tau \tau) \tau$ are identical by Eq. (2.1).

For many purposes, it is convenient to label trees with some standard type of symbol. Thus, we introduce the set $S$ defined as the union, for $n=1,2,3, \cdots$, of the sets of $n$-tuples of positive integers. That is,

$$
S=\{(1),(2), \cdots,(1,1),(1,2), \cdots,(2,1), \cdots,(1,1,1),(1,1,2), \cdots\}
$$

If $x \in S$ is an $n$-tuple, then we define $x^{-}$as the ( $n-1$ )-tuple formed by omitting the final integer from $x$. Conventionally, we write $x^{-}=() \notin S$ if $x$ is a 1-tuple. Let $F$ denote the set of finite subsets of $S$.

If $\sigma \in F$, then we can associate with $\sigma$ the labelled graph $(\sigma, \rho)$ where

$$
\rho=\left\{\left(x^{-}, x\right): x, x^{-} \in \sigma\right\} .
$$

We define a labelled forest as a labelled graph ( $\sigma, \rho$ ) such that ( $\sigma, \rho$ ) can be disconnected into $\left(\sigma_{1}, \rho_{1}\right), \cdots,\left(\sigma_{m}, \rho_{m}\right)$ (with $m$ possibly equal to 1 ) where each of $\left(\sigma_{1}, \rho_{1}\right)$, $\cdots,\left(\sigma_{m}, \rho_{m}\right)$ is a labelled tree or such that $(\sigma, \rho)=(\varnothing, \varnothing)$. The following lemma is stated without proof.

Lemma 2.3. Any finite subset of $S$ represents a forest and any forest can be represented by a finite subset of $S$.

Now, consider the set $M$ of real-valued functions on $F$. We will define a binary operation on $M$. However, we first define a partial ordering $\leqq$ on $F$ by writing $\sigma_{1} \leqq \sigma_{2}$ if $\sigma_{1} \subseteq \sigma_{2}$ and $x \in \sigma_{1}, x^{-} \in \sigma_{2} \Rightarrow x^{-} \in \sigma_{1}$. If $\alpha, \beta \in M$ we define the product $\alpha \beta$ by

$$
\begin{equation*}
(\alpha \beta)(\sigma)=\sum_{\sigma^{\prime} \leq \sigma} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \beta\left(\sigma^{\prime}\right) \tag{2.2}
\end{equation*}
$$

for all $\sigma \in F$ and the summation is over all members $\sigma^{\prime}$ of $F$ such that $\sigma^{\prime} \leqq \sigma$.
If we write $\sigma_{1}+\sigma_{2}+\cdots=\sigma$ to indicate that $\sigma_{1}, \sigma_{2}, \cdots$ are disjoint sets whose union is $\sigma$ then we may rewrite (2.2) as

$$
(\alpha \beta)(\sigma)=\sum_{\sigma_{1}+\sigma,-\sigma ; \sigma \leq \sigma} \alpha\left(\sigma_{1}\right) \beta\left(\sigma_{2}\right)
$$

We now show that the operation we have just defined is associative.
Lemma 2.4. For $\alpha, \beta, \gamma \in M,(\alpha \beta) \gamma=\alpha(\beta \gamma)$.
Proof. We have, for all $\sigma \in F$,

$$
\begin{align*}
& ((\alpha \beta) \gamma)(\sigma)=\sum_{\sigma_{1}+\sigma_{2}+\sigma_{2}-\sigma ; \sigma_{3} \leq \sigma ; \sigma_{2} \leq \sigma_{1} \cup \sigma_{2}} \alpha\left(\sigma_{1}\right) \beta\left(\sigma_{2}\right) \gamma\left(\sigma_{3}\right),  \tag{2.3}\\
& (\alpha(\beta \gamma))(\sigma)=\sum_{\sigma_{1}+\sigma_{2}+\sigma_{1}-\sigma ; \sigma_{2} \cup \sigma_{2} \leq \sigma_{i}, \sigma_{3} \leq \sigma_{2} \cup \sigma_{2}} \alpha\left(\sigma_{1}\right) \beta\left(\sigma_{2}\right) \gamma\left(\sigma_{3}\right) . \tag{2.4}
\end{align*}
$$

To show that the right-hand sides of (2.3) and (2.4) are identical, it is sufficient to show that if $\sigma_{1}, \sigma_{2}, \sigma_{3}$ are disjoint then

$$
\begin{equation*}
\sigma_{3} \leqq \sigma_{1} \cup \sigma_{2} \cup \sigma_{3} \tag{2.5}
\end{equation*}
$$

and

$$
\begin{equation*}
\sigma_{2} \leqq \sigma_{1} \cup \sigma_{2} \tag{2.6}
\end{equation*}
$$

if and only if

$$
\begin{equation*}
\sigma_{2} \cup \sigma_{3} \leqq \sigma_{1} \cup \sigma_{2} \cup \sigma_{3} \tag{2.7}
\end{equation*}
$$

and

$$
\begin{equation*}
\sigma_{3} \leqq \sigma_{2} \cup \sigma_{3} \tag{2.8}
\end{equation*}
$$

The relations (2.5)-(2.8) with $\leqq$ replaced by $\subseteq$ are automatically true. Hence, we may write the following equivalent form of these four statements.

$$
\begin{align*}
x \in \sigma_{3} & \Rightarrow x^{-} \notin \sigma_{1} \cup \sigma_{2},  \tag{2.9}\\
x \in \sigma_{2} & \Rightarrow x^{-} \notin \sigma_{1},  \tag{2.10}\\
x \in \sigma_{2} \cup \sigma_{3} & \Rightarrow x^{-} \notin \sigma_{1},  \tag{2.11}\\
x \in \sigma_{3} & \Rightarrow x^{-} \notin \sigma_{2} . \tag{2.12}
\end{align*}
$$

That (2.9), (2.10) are equivalent to (2.11), (2.12) is immediately clear.
With the operation defined by (2.2), $M$ is a semigroup with identity $e$ defined by $e(\varnothing)=1, e(\sigma)=0$ if $\sigma \in F \backslash\{\varnothing\}$. The subset $M_{0}=\{\alpha: \alpha \in M, \alpha(\varnothing) \neq 0\}$ with the operation is a group. To show this, we see that if $\alpha, \beta \in M_{0}$ then $(\alpha \beta)(\varnothing)=\alpha(\varnothing) \beta(\varnothing)$ $\neq 0$ so that $\alpha \beta \in M_{0}$ and we show how to construct an element $\alpha^{-1} \in M_{0}$ so that $\alpha \alpha^{-1}=e$. We do this by defining $\alpha^{-1}(\sigma)$ recursively on the number of elements in $\sigma$. We have

$$
\begin{align*}
\alpha^{-1}(\varnothing) & =\alpha(\varnothing)^{-1}  \tag{2.13}\\
\alpha^{-1}(\sigma) & =-\left(\sum_{\sigma, \leq ; \sigma^{\prime}, \sigma} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \alpha^{-1}\left(\sigma^{\prime}\right)\right) / \alpha(\varnothing) \tag{2.14}
\end{align*}
$$

for $\sigma \in F \backslash\{\varnothing\}$.
We consider further the subset $M_{1}$ of $M_{0}$ defined so that $\alpha \in M_{1}$ if and only if, whenever $\sigma_{1}$ and $\sigma_{2}$ correspond to the same forest, then $\alpha\left(\sigma_{1}\right)=\alpha\left(\sigma_{2}\right)$. It is easy to see that $M_{1}$ is a subgroup of $M_{0}$. However, our main concern will be with a subgroup $M_{2}$ of $M_{1}$ which we now describe.

An element $\alpha \in M_{1}$ is said to be multiplicative if $\alpha(\varnothing)=1$ and if, whenever
( $\sigma_{1}, \sigma_{2}$ ) is a disconnection of $\sigma$ (that is $\sigma_{1}+\sigma_{2}=\sigma, \sigma_{1} \leqq \sigma, \sigma_{2} \leqq \sigma$ ), then $\alpha(\sigma)=$ $\alpha\left(\sigma_{1}\right) \alpha\left(\sigma_{2}\right)$. We define $M_{2}$ as the set of multiplicative elements of $M_{1}$. This means in effect that members of $M_{2}$ can be represented by functions on $T$ to $R$. That $M_{2}$ is a subgroup is shown by the following lemma.

Lemma 2.5. If $\alpha, \beta \in M_{1}$ are multiplicative then $\alpha \beta$ and $\alpha^{-1}$ are multiplicative.
Proof. Suppose ( $\sigma_{1}, \sigma_{2}$ ) is a disconnection of $\sigma$ then there is a bijection between the set of pairs $\left(\sigma_{1}^{\prime}, \sigma_{2}^{\prime}\right)$ such that $\sigma_{1}^{\prime} \leqq \sigma_{1}, \sigma_{2}^{\prime} \leqq \sigma_{2}$ and the set $\left\{\sigma^{\prime} \in F: \sigma^{\prime} \leqq \sigma\right\}$. In fact, the bijection is such that $\sigma^{\prime}=\sigma_{1}^{\prime} \cup \sigma_{2}^{\prime}$ corresponds to ( $\sigma_{1}^{\prime}, \sigma_{2}^{\prime}$ ). In the inverse relation, $\left(\sigma_{1}^{\prime}, \sigma_{2}^{\prime}\right)=\left(\sigma_{1} \cap \sigma^{\prime}, \sigma_{2} \cap \sigma^{\prime}\right)$ is a disconnection of the corresponding $\sigma^{\prime}$ and $\left(\sigma_{1} \backslash \sigma_{1}^{\prime}, \sigma_{2} \backslash \sigma_{2}^{\prime}\right)$ is a disconnection of $\sigma \backslash \sigma^{\prime}$.

We now have

$$
\begin{aligned}
(\alpha \beta)(\sigma) & =\sum_{\sigma^{\prime} \leq \sigma} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \beta\left(\sigma^{\prime}\right) \\
& =\sum_{\sigma_{1}^{\prime} \leq \sigma_{1}} \sum_{\sigma_{2} \leq \sigma_{2}} \alpha\left(\sigma_{1} \backslash \sigma_{1}^{\prime}\right) \alpha\left(\sigma_{2} \backslash \sigma_{2}^{\prime}\right) \beta\left(\sigma_{1}^{\prime}\right) \beta\left(\sigma_{2}^{\prime}\right) \\
& =(\alpha \beta)\left(\sigma_{1}\right) \cdot(\alpha \beta)\left(\sigma_{2}\right)
\end{aligned}
$$

so that $\alpha \beta$ is multiplicative.
To prove $\alpha^{-1}$ is multiplicative, we proceed by induction. Thus, we suppose $\alpha^{-1}\left(\sigma^{\prime}\right)=\alpha^{-1}\left(\sigma_{1}^{\prime}\right) \alpha^{-1}\left(\sigma_{2}^{\prime}\right)$ whenever $\left(\sigma_{1}^{\prime}, \sigma_{2}^{\prime}\right)$ is a disconnection of the forest $\sigma^{\prime}$ with fewer elements than $\sigma$. We now have

$$
\begin{aligned}
\alpha^{-1}(\sigma) & =-\sum_{\sigma^{\prime} \leq ; ; \sigma^{\prime} \neq \sigma} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \alpha^{-1}\left(\sigma^{\prime}\right) \\
& =\alpha^{-1}\left(\sigma_{1}\right) \alpha^{-1}\left(\sigma_{2}\right)-\sum_{\sigma_{1}^{\prime} \leq \sigma_{1}} \sum_{\sigma^{\prime} \leq \sigma_{2}} \alpha\left(\sigma_{1} \backslash \sigma_{1}^{\prime}\right) \alpha\left(\sigma_{2} \backslash \sigma_{2}^{\prime}\right) \alpha^{-1}\left(\sigma_{1}^{\prime}\right) \alpha^{-1}\left(\sigma_{2}^{\prime}\right) \\
& =\alpha^{-1}\left(\sigma_{1}\right) \alpha^{-1}\left(\sigma_{2}\right)
\end{aligned}
$$

We now observe that if $\sigma \in F$ is a labelled tree and $\sigma^{\prime} \leqq \sigma$, then either $\sigma^{\prime}$ is a labelled tree or $\sigma^{\prime} \neq \varnothing$. Hence, in this case, (2.2) takes the form

$$
\begin{align*}
(\alpha \beta)(\sigma) & =\sum_{\sigma^{\prime} \leq \sigma} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \beta\left(\sigma^{\prime}\right)  \tag{2.15}\\
& =\alpha(\sigma)+\sum_{\sigma^{\prime} \leq \sigma ; \sigma^{\prime} \neq \varnothing} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \beta\left(\sigma^{\prime}\right)
\end{align*}
$$

where the summation involves values of $\beta$ evaluated only at trees.
Let $a, b$ be functions on $T$ to $R$ corresponding to $\alpha, \beta \in M_{2}$. That is, if $\varphi$ is a function on the nonempty connected members of $F$ to $T$ defined so that for each $\sigma$, $\varphi(\sigma)$ is the tree corresponding to $\sigma$, then $a=\alpha \circ \varphi^{-1}, b=\beta \circ \varphi^{-1}$. If $t \in T$ is given by $t=\varphi(\sigma)$, then (2.15) becomes

$$
\begin{equation*}
(a b)(t)=a(t)+\lambda(a, t)(b) \tag{2.16}
\end{equation*}
$$

where for a given $a$ and $t, \lambda(a, t)$ is a certain linear functional on the set of functions on $T$ to $R$.

In later sections of this paper, the function $\lambda$ will assume some prominence so we now consider its main properties. First, however, we make some definitions.

We define $G$ as the set of functions on $T$ to $R$ with the composition rule given by (2.16) so that $G$ is a group isomorphic to $M_{2}$. The identity of $G$ maps all $t \in T$ to 0 .

We also define for $t \in T$, the linear functional $\hat{t}$ (or $\hat{t}$ ) on $G$ defined by $\hat{t}(a)=a(t)$ for all $a \in G$. $\hat{T}$ will denote the set $\{\hat{t}: t \in T\}$ and $G^{*}$ will denote the set of linear functionals on $G$ spanned by $\hat{T}$. That is, if $p \in G^{*}$ then there is a function $c_{p}$ on $T$ to $R$ such that $\left\{t \in T: c_{p}(t) \neq 0\right\}$ is finite and so that for all $a \in G$,

$$
p(a)=\sum_{i \in T} c_{p}(t) a(t)
$$

For $\hat{t}, \hat{u} \in \hat{T}$ we define the product $\hat{t} \cdot \hat{u}$ by $\hat{t} \cdot \hat{u}=(t u)^{\wedge}$ and if $p, q \in G^{*}$ we define $p \cdot q \in G^{*}$ in such a way that it is a bilinear extension of the product defined for $\hat{T}$. . That is, if $p=\sum_{t \in T} c_{p}(t) \hat{t}, q=\sum_{t \in T} c_{q}(t) \hat{t}$ then we define $p q$ as

$$
p q=\sum_{t, u \in T} c_{p}(t) c_{a}(u) \hat{t} \hat{u}
$$

We note that $\lambda$ is a mapping on $G \times T$ to $G^{*}$. The following result enables us to determine $\lambda$ using the induction principle for $T$ given by Lemma 2.2.

Theorem 2.6. For $a \in G$, and $u, v \in T$ we have

$$
\begin{align*}
\lambda(a, \tau) & =\hat{\tau}  \tag{2.17}\\
\lambda(a, u v) & =\lambda(a, u) \lambda(a, v)+a(v) \lambda(a, u) \tag{2.18}
\end{align*}
$$

Proof. (2.17) is equivalent to $(\alpha \beta)(\sigma)=\alpha(\sigma)+\beta(\sigma)$ in the case when $\sigma \in F$ has exactly one member. This is clear because if $\sigma^{\prime} \leqq \sigma$ then $\sigma^{\prime}=\varnothing$ or $\sigma^{\prime}=\sigma$.

The proof of (2.18) is more involved. Let $\sigma \in F$ correspond to $t=u v$ and let $\sigma_{1}, \sigma_{2}$ correspond to $u$ and $v$, respectively. Without loss of generality, we may suppose that the root of $\sigma$ is (1) and the root of $\sigma_{2}$ is $(1,1)$. The only connection between $\sigma_{1}$ and $\sigma_{2}$ is $\left((1),(1,1)\right.$ ) so any subset of $\sigma_{1}$ not containing (1) is not connected to any subset of $\sigma_{2}$ and any subset of $\sigma_{2}$ not containing $(1,1)$ is not connected to any subset of $\sigma_{1}$. Furthermore, if $\sigma_{1}^{\prime} \leqq \sigma_{1}, \sigma_{2}^{\prime} \leqq \sigma_{2}$ and $\sigma_{1}^{\prime}=\varnothing$ only if $\sigma_{2}^{\prime}=\varnothing$ then $\sigma_{1}^{\prime} \cup \sigma_{2}^{\prime} \leqq \sigma$. Also we observe that the relationship between such pairs ( $\sigma_{1}^{\prime}, \sigma_{2}^{\prime}$ ) as have these properties and $\sigma^{\prime}=\sigma_{1}^{\prime} \cup \sigma_{2}^{\prime}$ is a bijection. Hence, we have

$$
\begin{aligned}
(\alpha \beta)(\sigma)= & \sum_{\sigma^{\prime} \leq \sigma} \alpha\left(\sigma \backslash \sigma^{\prime}\right) \beta\left(\sigma^{\prime}\right) \\
= & \alpha(\sigma)+\sum_{\varnothing \neq \sigma_{1}^{\prime} \leq \sigma_{1}} \alpha\left(\sigma_{1} \backslash \sigma_{1}^{\prime}\right) \alpha\left(\sigma_{2}\right) \beta\left(\sigma_{1}^{\prime}\right) \\
& +\sum_{\varnothing \neq \sigma_{1}: \leq \sigma_{1}: \varnothing \neq \sigma_{2}^{\prime} \leq \sigma_{2}} \alpha\left(\sigma_{1} \backslash \sigma_{1}^{\prime}\right) \alpha\left(\sigma_{2} \backslash \sigma_{2}^{\prime}\right) \beta\left(\sigma_{1}^{\prime} \cup \sigma_{2}^{\prime}\right)
\end{aligned}
$$

and this is equivalent to (2.18).
A final property of $\lambda$ we will consider is
Lemma 2.7. If $a, b, c \in G$ and $t \in T$ then

$$
\begin{equation*}
\lambda(a b, t)(c)+\lambda(a, t)(b)=\lambda(a, t)(b c) \tag{2.19}
\end{equation*}
$$

Proof. This result is simply a reformulation of the associative law for $G$. We have

$$
\begin{aligned}
((a b) c)(t) & =\lambda(a b, t)(c)+(a b)(t) \\
& =\lambda(a b, t)(c)+\lambda(a, t)(b)+a(t), \\
(a(b c))(t) & =\lambda(a, t)(b c)+a(t),
\end{aligned}
$$

and (2.19) follows by equating these expressions.
3. The General Integration Method. Let $H$ denote an arbitrary set and $B(H)$ the set of bounded real-valued functions on $H . A(H)$ will denote the set of bounded linear operators on $B(H)$ to $B(H)$. If we are given $H$ and $a \in A(H)$, then we can consider an "initial-value problem" defined by a Lipschitz-continuous function $f$ with sufficiently small Lipschitz constant on a Banach space $X$ to $X$ and by a fixed element $\eta_{0} \in X$. The problem is to evaluate $\eta_{1}=y\left(h_{1}\right)$ for $h_{1}$ a fixed point in $H$ and $y$ a bounded function on $H$ to $X$ (we write $Y$ for the set of such functions) which satisfies

$$
\begin{equation*}
y=\eta_{0} e+a \times(f \circ y) . \tag{3.1}
\end{equation*}
$$

Here, $e$ denotes the unit in $B(H)$ so that $\eta_{0} e$ will be defined by $\left(\eta_{0} e\right)(h)=\eta_{0}$ for all $h \in H$. The notation $a \times$ denotes a linear operator on $Y$ to $Y$ defined by $p \circ(a \times y)=$ $a(p \circ y)$ for all $p \in X^{*}$ (the dual space of $X$ ). In order to ensure that a $\times$ is defined, we will assume that $X$ is reflexive. For a given $h \in H$, the mapping that takes $p$ to $a(p \circ y)(h)$ is linear and continuous and is thus equal to $\hat{x} \in X^{* *}$ for some $x \in X$. $(a \times y)(h)$ is thus defined as $x$. In later sections of the paper, we shall, in fact, suppose that $X$ is finite dimensional.

To illustrate the type of initial-value problem covered by our general method, we point out that if $H$ is a finite set, then the procedure for evaluating $\eta_{1}$ is just that of a Runge-Kutta method. On the other hand, if $H=[0,1]$ and $a$ is defined by

$$
\begin{equation*}
a(x)(h)=\int_{0}^{h} x(t) d t \tag{3.2}
\end{equation*}
$$

for $h \in H$ then $y(h)$ defined by (3.1) is the solution at $t=h_{-}$to the Cauchy problem

$$
\begin{equation*}
\frac{d y}{d t}(t)=f(y(t)), \quad y(0)=\eta_{0} . \tag{3.3}
\end{equation*}
$$

It is of interest to consider then a subset of $B(H)$ containing $e$, the (pointwise) product of any two members and $a(x)$ for any member $x$. Thus, such a subset would contain $e, a(e), a(a(e)), a(e)^{2}, a(e) \cdot a(a(e)), \cdots$. We will be interested later in a Banach algebra containing such elements but for the present, we concern ourselves only with the smallest subset of $B(H)$ with the properties described. It is convenient, for a fixed $a$ to relate the various members of this subset to $T$ and, in fact, the subset is the range of the function $\mu_{a}$ now to be defined.

For a given $a \in A(H)$, we define the mapping $\mu_{a}$ on $T$ to $B(H)$ such that

$$
\begin{align*}
\mu_{a}(\tau) & =e  \tag{3.4}\\
\mu_{a}\left(t_{1} t_{2}\right) & =\mu_{a}\left(t_{1}\right) a\left(\mu_{a}\left(t_{2}\right)\right) \tag{3.5}
\end{align*}
$$

for $t_{1}, t_{2} \in T$ where the multiplication on the right of (3.5) is pointwise. It is typical of this type of recursive definition of functions on $T$ that it is necessary to verify that the definition of the image of $t_{1} t_{2}$ is such that the images of $t_{1} t_{2} \cdot t_{3}$ and $t_{1} t_{3} \cdot t_{2}$ should be equal. In the present case, we have

$$
\mu_{a}\left(t_{1} t_{2} \cdot t_{3}\right)=\mu_{a}\left(t_{1}\right) \cdot a\left(\mu_{a}\left(t_{2}\right)\right) \cdot a\left(\mu_{a}\left(t_{3}\right)\right)=\mu_{a}\left(t_{1} t_{3} \cdot t_{2}\right)
$$

It is also convenient to introduce a mapping $\nu_{a}=\mu_{a} \circ \varphi$ where $\varphi$ is the mapping on $T$ to $T$ defined by $\varphi(t)=\tau t$. Thus,

$$
\begin{equation*}
\nu_{a}(t)=a\left(\mu_{a}(t)\right) \tag{3.6}
\end{equation*}
$$

and the elements in the range of $\nu_{a}$ can be regarded as generators of the range of $\mu_{a}$.
To illustrate the notation introduced so far, the eight trees of Table 2.1 are listed in Table 3.1 with the expressions for $\mu_{a}$ and $\nu_{a}$ in each case. The range of $\varphi$ will be

Table 3.1

| $t$ | $\mu_{a}(t)$ | $\nu_{a}(t)$ |
| :--- | :--- | :--- |
| $\tau$ | $e$ | $a(e)$ |
| $\tau \tau$ | $a(e)$ | $a(a(e))$ |
| $\tau \cdot \tau \tau$ | $a(a(e))$ | $a(a(a(e)))$ |
| $\tau \tau \cdot \tau$ | $a(e)^{2}$ | $a\left(a(e)^{2}\right)$ |
| $\tau(\tau \cdot \tau \tau)$ | $a(a(a(e)))$ | $a(a(a(a(e))))$ |
| $\tau \tau \cdot \tau \tau$ | $a(e) a(a(e))$ | $a(a(e) a(a(e)))$ |
| $\tau(\tau \tau \cdot \tau)$ | $a\left(a(e)^{2}\right)$ | $a\left(a\left(a(e)^{2}\right)\right)$ |
| $(\tau \tau \cdot \tau) \tau$ | $a(e)^{3}$ | $a\left(a(e)^{3}\right)$ |

denoted by $U$ so that $U \subset T$. We now define for each $a \in A(H)$ an equivalence relation on $H$. $h_{1}, h_{2}$ will be equivalent if and only if, for all $t \in T, \mu_{a}(t)\left(h_{1}\right)=\mu_{a}(t)\left(h_{2}\right)$. It is clear that in this definition we could replace $\mu$ by $\nu$ or, what is equivalent, we could replace $T$ by $U$. Let $P_{a}$ be the partition of $H$ corresponding to this equivalence relation. Thus, there is a function $\pi_{a}$ mapping $H$ onto $P_{a}$ such that $\pi_{a}\left(h_{1}\right)=\pi_{a}\left(h_{2}\right)$ if and only if $h_{1}, h_{2}$ are equivalent and for all $h \in H, h \in \pi_{a}(h)$.

In Section 5, we will show that, under certain conditions, the solutions to (3.1) have the same values for all equivalent members of $H$. In this sense, we may identify these members and this result will have important consequences in the rest of the paper.

However, it is first necessary to make a closer study of the operator $a$ and this we do in the next section.
4. Analytical Preliminaries. In this section, we will be concerned with a fixed member $a$ of $A(H)$ so, without risk of confusion, we will omit the subscript in $\mu_{a}$, $\nu_{a}, \pi_{a}$ and $P_{a}$.

Since for all $t \in T, \nu(t)$ maps equivalent elements to the same image, the relation $\nu(t) \circ \pi^{-1}$ is a function on $P$. We now impose on $P$ the weak topology generated by the family $\left\{\nu(t) \circ \pi^{-1}: t \in T\right\}$. That is, $P$ will have the weakest topology for which all functions in this family are continuous. With this topology, $P$ has the following property:

Lemma 4.1. $P$ is a Tychonoff space (that is, $P$ is a completely regular $T_{1}$ space).
Proof. Consider the family of closed bounded subsets of $R,\{Q(t): t \in T\}$ where $Q(t)$ is defined as the closure of the range of $\nu(t)$. Let $Q$ denote the product space of this family, so that $Q$ is compact and Hausdorff and hence Tychonoff.

We write $q(t)$ for the projection of $Q$ to $Q(t)$ and define a function $r$ from $P$ to $Q$ by

$$
\begin{equation*}
q(t) \circ r \circ \pi=\nu(t) \tag{4.1}
\end{equation*}
$$

for all $t \in T$. Let the range of $r$ be $Q_{0}$ so that $Q_{0}$ is Tychonoff since $Q_{0} \subset Q . r$ is
injective since if $h_{1}, h_{2} \in H$ are such that $\pi\left(h_{1}\right) \neq \pi\left(h_{2}\right)$ then there is a $t \in T$ such that $\nu(t)\left(h_{1}\right) \neq \nu(t)\left(h_{2}\right)$. Hence, we find

$$
(q(t) \circ r)\left(\pi\left(h_{1}\right)\right) \neq(q(t) \circ r)\left(\pi\left(h_{2}\right)\right)
$$

so that $r\left(\pi\left(h_{1}\right)\right) \neq r\left(\pi\left(h_{2}\right)\right)$.
Since $r$ is a bijection from $P$ to the Tychonoff space $Q_{0}$, it will follow that $P$ is Tychonoff when we have shown that $r$ is, in fact, a homeomorphism. This follows since for $t \in T, \nu(t) \circ \pi^{-1}=q(t) \circ r$ is continuous and thus $r$ is continuous. Moreover, by the definition of the weak topology on $P, r$ is open. The proof of Lemma 4.1 is now complete. Let $\beta(P)$ denote the Stone-Čech compactification of $P$. We extend in the usual way functions in $C(P)$ to functions in $C(\beta(P))$ and the extension of $x \in C(P)$ will be denoted (in this section only) by $\hat{x}$ or by $x^{\text {. }}$. In particular, for $t \in T,\left(\mu(t) \circ \pi^{-1}\right)^{-}$ is defined. Let $V$ denote the set of real-valued functions on $T$ such that all but a finite number of members of $T$ map to zero. We now consider the set $D$ defined as the closure in $C(\beta(P))$ of functions of the form

$$
\begin{equation*}
\sum_{t \in T} c(t)\left(\mu(t) \circ \pi^{-1}\right)^{\hat{n}} \tag{4.2}
\end{equation*}
$$

where $c \in V$ and the summation is defined in the obvious way as the sum of the nonzero terms. We have the following result.

Lemma 4.2. $D=C(\beta(P))$.
Proof. We first show that $D$ is an algebra. It will be sufficient to show that if $c_{1}, c_{2} \in V$ then there is a $c \in V$ such that

$$
\left(\sum_{t \in T} c_{1}(t) \mu(t)\right) \cdot\left(\sum_{t \in T} c_{2}(t) \mu(t)\right)=\sum_{t \in T} c(t) \mu(t) .
$$

However, this follows by noting that if $t_{1}, t_{2} \in T$ then there is a $t \in T$ such that $\mu(t)=\mu\left(t_{1}\right) \mu\left(t_{2}\right)$ and this is clear because $\mu\left(t_{1}\right), \mu\left(t_{2}\right)$ can be written as products of factors of the form $\nu\left(t^{\prime}\right)$.

We next note that $D$ separates points of $P$ since if $h_{1}, h_{2} \in H$ are such that $\pi\left(h_{1}\right) \neq$ $\pi\left(h_{2}\right)$ then there is a $t \in T$ such that $\mu(t)\left(h_{1}\right) \neq \mu(t)\left(h_{2}\right)$. Thus $\left(\mu(t) \circ \pi^{-1}\right)^{\wedge} \in D$ takes on different values at $\pi\left(h_{1}\right)$ and $\pi\left(h_{2}\right)$.

We note also that $\beta(P)$ is Hausdorff and compact and that $D$ contains the unit (the unit is in fact $\left.\left(\mu(\tau) \circ \pi^{-1}\right)^{\wedge}\right)$. We can now make use of the Stone-Weierstrass theorem to obtain the result of the lemma.

We now define $C_{a}(H)$ as the subset of $B(H)$ such that its elements are of the form $(x \mid P) \circ \pi$ where $x \in C(\beta(P))$. [The notation $(x \mid P)$ denotes the restriction of $x$ to $P$.] Using the previous lemma we see that elements of $C_{a}(H)$ can be approximated in the sense of the norm-topology by functions of the form

$$
\begin{equation*}
\sum_{t \in T} c(t) \mu(t) \tag{4.3}
\end{equation*}
$$

where $c \in V$. We now have a basic result.
Lemma 4.3. If $x \in C_{a}(H)$ then $a(x) \in C_{a}(H)$.
Proof. We must show that for each $x \in C_{a}(H)$ and positive number $\epsilon_{2}$, there exists $x_{2}$ of the form (4.3) such that $\left\|a(x)-x_{2}\right\|<\epsilon_{2}$. Let $\epsilon_{1}>0$ be such that $\epsilon_{1}\|a\|<$ $\epsilon_{2}$ and let

$$
x_{1}=\sum_{t \in \boldsymbol{T}} c_{1}(t) \mu(t),
$$

where $c_{1} \in V$, be such that $\left\|x_{1}-x\right\|<\epsilon_{1}$. We now define

$$
\begin{aligned}
x_{2} & =\sum_{t \in T} c_{1}(t) a(\mu(t))=\sum_{t \in T} c_{1}(t) \mu(\varphi(t)) \\
& =\sum_{t \in U}\left(c_{1} \circ \varphi^{-1}\right)(t) \mu(t)
\end{aligned}
$$

Let $c_{2} \in V$ be defined by $c_{2} \mid U=c_{1} \circ \varphi^{-1}$ and $c_{2}(t)=0$ if $t \notin U$. We now see that

$$
x_{2}=\sum_{t \in T} c_{2}(t) \mu(t)
$$

and, since $x_{2}=a\left(x_{1}\right)$ we have

$$
\left\|x_{2}-a(x)\right\| \leqq\|a\| \cdot\left\|x_{1}-x\right\| \leqq\|a\| \epsilon_{1}<\epsilon_{2}
$$

5. The Main Results. We now return to Eq. (3.1) with the assumption that $X$ is finite dimensional. This ensures that $X$ is reflexive as was assumed in Section 3 and also ensures that the weak and strong topologies for $X$ are identical. As was stated in Section 3, we suppose that $f$ is Lipschitz-continuous. (Throughout the rest of this paper, $X, f$ and $\eta_{0}$ will have the significance they had in Section 3.) Suppose the Lipschitz constant is $L$ where $L\|a\|<1$. We may now regard the space $Y$ as a metric space with metric $d$ given by

$$
d(y, z)=\sup _{h \in H}\|y(h)-z(h)\|
$$

so that the mapping which takes $y \in Y$ to $\eta_{0} e+a \times(f \circ y)$ is a contraction.
We write $C_{a}(H, X)$ for the subset of $Y$ such that $y \in C_{a}(H, X)$ if and only if for all $p \in X^{*}, p \circ y \in C_{a}(H)$. Our first main result will be stated as Theorem 5.4 but we precede this by three lemmas necessary for its proof.

Lemma 5.1. $\eta_{0} e \in C_{a}(H, X)$.
Proof. For $p \in X^{*}$ we have $p \circ\left(\eta_{0} e\right)=p\left(\eta_{0}\right) e \in C_{a}(H)$.
Lemma 5.2. If $y \in C_{a}(H, X)$ then $f \circ y \in C_{a}(H, X)$.
Proof. Since $f$ satisfies a Lipschitz condition, $p \circ f$ is continuous in the norm topology. Hence, it is continuous in the weak topology. Hence, $p \circ f \circ y$ is continuous in the topology of $H$. Hence $f \circ y \in C_{a}(H, X)$.

Lemma 5.3. If $y \in C_{a}(H, X)$, then $a \times y \in C_{a}(H, X)$.
Proof. $p \circ(a \times y)=a(p \circ y)=a(x)$ where $x \in C_{a}(H)$. But, by Lemma 4.3, $a(x) \in C_{a}(H)$. Hence, $a \times y \in C_{a}(H, X)$.

Theorem 5.4. y satisfying (3.1) is a member of $C_{a}(H, X)$.
Proof. If $y$ is evaluated as the limit of the sequence $\left(y_{0}, y_{1}, y_{2}, \cdots\right)$ where $y_{0}=\eta_{0} e$ and, for $n=1,2, \cdots$,

$$
\begin{equation*}
y_{n}=\eta_{0} e+a \times\left(f \circ y_{n-1}\right) \tag{5.1}
\end{equation*}
$$

then, using Lemmas $5.1,5.2,5.3$, we see that $y_{0} \in C_{a}(H, X)$ and that, if $y_{n-1} \in$ $C_{a}(H, X)$ then $y_{n} \in C_{a}(H, X)$. Hence, by induction, for all $n, y_{n} \in C_{a}(H, X)$. Since $C_{a}(H, X)$ is a closed subspace of $Y$, it follows that $y \in C_{a}(H, X)$.

As a consequence of this theorem, we have
Corollary 5.5. If y satisfies (3.1) and if $h_{1}, h_{2} \in H$ are such that $\pi_{a}\left(h_{1}\right)=\pi_{a}\left(h_{2}\right)$ then $y\left(h_{1}\right)=y\left(h_{2}\right)$,

Proof. Since $y \in C_{a}(H, X)$, then, for all $p \in X^{*}, p \circ y \in C_{a}(H)$. Let $x \in C_{a}\left(\beta\left(P_{a}\right)\right)$
be such that $p \circ y=\left(x \mid P_{a}\right) \circ \pi_{a}$ then $(p \circ y)\left(h_{1}\right)-(p \circ y)\left(h_{2}\right)=\left(x \mid P_{a}\right)\left(\pi_{a}\left(h_{1}\right)\right)-$ $\left(x \mid P_{a}\right)\left(\pi_{a}\left(h_{2}\right)\right)=0$. Hence, $p\left(y\left(h_{1}\right)-y\left(h_{2}\right)\right)=0$ for all $p \in X^{*}$ and the result follows.

The theorem which next follows expresses our second main result that in a certain sense a method $\left(H, a, h_{1}\right)$ is characterised by the function which takes $t \in T$ to $\nu_{o}(t)\left(h_{1}\right)$.

Theorem 5.6. Let $\left(H, a, h_{1}\right),\left(K, b, k_{1}\right)$ be two methods such that for all $t \in T$, $\nu_{a}(t)\left(h_{1}\right)=\nu_{b}(t)\left(k_{1}\right)$. Let $y, z$ be the solutions of

$$
\begin{align*}
& y=\eta_{0} e_{H}+a \times(f \circ y),  \tag{5.2}\\
& z=\eta_{0} e_{K}+b \times(f \circ z), \tag{5.3}
\end{align*}
$$

where $f$ is Lipschitz-continuous with constant $L$ such that $L\|a\|<1, L\|b\|<1$ and $e_{H}, e_{K}$ denote the units in $B(H), B(K)$ respectively. Then $y\left(h_{1}\right)=z\left(k_{1}\right)$.

Proof. Without loss of generality, we suppose that $H \cap K=\varnothing$. We now consider $c \in A(H \cup K)$ (we recall that $A(H)$ is defined for a set $H$ in Section 3) defined by

$$
\begin{align*}
& c(v)(h)=a(v \mid H)(h),  \tag{5.4}\\
& c(v)(k)=b(v \mid K)(k), \tag{5.5}
\end{align*}
$$

for $v \in B(H \cup K), h \in H$ and $k \in K$. It is found that

$$
\begin{aligned}
\|c(v)\| & =\max (\|a(v \mid H)\|,\|b(v \mid K)\|) \\
& \leqq \max (\|a\|,\|b\|) \cdot\|v\|
\end{aligned}
$$

so that $\|c\| \leqq \max (\|a\|,\|b\|)$. (In fact, it is easy to see that $\|c\|=\max (\|a\|,\|b\|)$.) Hence, $L\|c\|<1$.

If $w$ is defined as the bounded function on $H \cup K$ to $X$ satisfying

$$
\begin{equation*}
w=\eta_{0} e_{H \cup K}+c \times(f \circ w), \tag{5.6}
\end{equation*}
$$

then $w$ is given by

$$
\begin{equation*}
w|H=y, \quad w| K=z \tag{5.7}
\end{equation*}
$$

To see this, we evaluate the right-hand side of (5.6) at $h \in H$ with $w$ satisfying $(5,7)$ (the similar calculation for $k \in K$ will be omitted). We have

$$
\begin{aligned}
\left(\eta_{0} e_{H \cup K}+c \times(f \circ w)\right)(h) & =\eta_{0}+(a \times((f \circ w) \mid H))(h) \\
& =\eta_{0}+(a \times(f \circ(w \mid H)))(h) \\
& =\left(\eta_{0} e_{H}+a \times(f \circ y)\right)(h) .
\end{aligned}
$$

This shows in particular that $w\left(h_{1}\right)=y\left(h_{1}\right)$ and, similarly, that $w\left(k_{1}\right)=z\left(k_{1}\right)$.
We also see that if $t \in T, h \in H, k \in K$ then $\mu_{c}(t)(h)=\mu_{a}(t)(h)$ and $\mu_{c}(t)(k)=$ $\mu_{b}(t)(k)$. We verify this in the case of $h \in H$ using the induction principle for $T$. Thus, for $t=\tau$ we have $\mu_{c}(t)(h)=e_{H \cup K}(h)=1=e_{H}(h)=\mu_{a}(t)(h)$. To complete the induction argument, we write $t=u v$ and assume that $\mu_{c}(u)(h)=\mu_{a}(u)(h), \mu_{c}(v)(h)=$ $\mu_{a}(v)(h)$ for all $h \in H$. We now have

$$
\begin{aligned}
\mu_{c}(t)(h) & =\mu_{c}(u v)(h)=\left(\mu_{c}(u) c\left(\mu_{c}(v)\right)\right)(h) \\
& =\mu_{a}(u)(h) \cdot a\left(\mu_{c}(v) \mid H\right)(h)=\mu_{a}(u)(h) \cdot a\left(\mu_{a}(v)\right)(h) \\
& =\mu_{a}(t)(h) .
\end{aligned}
$$

It now follows that for all $t \in T, \nu_{c}(t)\left(h_{1}\right)=\nu_{a}(t)\left(h_{1}\right)=\nu_{b}(t)\left(k_{1}\right)=\nu_{c}(t)\left(k_{1}\right)$ so that $\pi_{c}\left(h_{1}\right)=\pi_{c}\left(k_{1}\right)$. Hence, by Corollary 5.5, $w\left(h_{1}\right)=w\left(k_{1}\right)$. That is, $y\left(h_{1}\right)=z\left(k_{1}\right)$.

Definition 5.7. Methods such as ( $H, a, h_{1}$ ), ( $K, b, k_{1}$ ) related as in Theorem 5.6 will be called equivalent methods.

If the methods $\left(H, a, h_{1}\right),\left(K, b, k_{1}\right)$ are so related that there exists a bijection $\varphi: H \rightarrow K$ such that $\varphi\left(h_{1}\right)=k_{1}$ and such that whenever $x \in B(K), k \in K$ then $a(x \circ \varphi)\left(\varphi^{-1}(k)\right)=b(x)(k)$, then it is easy to verify that the two methods are equivalent. However, this special type of relationship is a much closer one than equivalence. For example, it can be verified that $\varphi$ is a homeomorphism between $H$ and $K$ when the topologies of Section 4 are imposed and we have for all $h \in H, t \in T$ the relationship $\mu_{a}(t)(h)=\mu_{b}(t)(\varphi(h))$. Methods related in this special way will be regarded as identical. This means that whenever we have two different methods $\left(H, a, h_{1}\right),\left(H^{\prime}\right.$, $a^{\prime}, h_{1}^{\prime}$ ) to consider together for any purpose we can suppose (by replacing one of them by an identical method) that $H$ and $H^{\prime}$ are disjoint.
6. The Group of Integration Methods. For two integration methods ( $H, a, h_{1}$ ), ( $K, b, k_{1}$ ), where $H \cap K=\varnothing$, we define the product as the method ( $H \cup K, c, k_{1}$ ) where $c \in A(H \cup K)$ is defined for $x \in B(H \cup K), h \in H$ and $k \in K$ by

$$
\begin{equation*}
c(x)(h)=a(x \mid H)(h), \quad c(x)(k)=a(x \mid H)\left(h_{1}\right)+b(x \mid K)(k) \tag{6.1}
\end{equation*}
$$

that is, by

$$
\begin{equation*}
c(x)|H=a(x \mid H), \quad c(x)| K=a(x \mid H)\left(h_{1}\right) e_{K}+b(x \mid K) \tag{6.2}
\end{equation*}
$$

The significance of this definition is that the product of the two methods is the method that results by applying each of the given methods in order. This result is stated more precisely in the following.

Theorem 6.1. Let $\left(H \cup K, c, k_{1}\right)$ be the product of $\left(H, a, h_{1}\right)$ and $\left(K, b, k_{1}\right)$ and let $f$ satisfy a Lipschitz condition with constant $L$ such that $L\|c\|<1$. If $\eta_{1}, \eta_{2}, \eta_{3}$ are defined by

$$
\begin{array}{ll}
y_{1}=\eta_{0} e_{H}+a \times\left(f \circ y_{1}\right), & \eta_{1}=y_{1}\left(h_{1}\right) \\
y_{2}=\eta_{1} e_{K}+b \times\left(f \circ y_{2}\right), & \eta_{2}=y_{2}\left(k_{1}\right) \\
y_{3}=\eta_{0} e_{H \cup K}+c \times\left(f \circ y_{3}\right), & \eta_{3}=y_{3}\left(k_{1}\right) \tag{6.5}
\end{array}
$$

then $\eta_{2}=\eta_{3}$.
Proof. We first show that $\|a\| \leqq\|c\|$ and $\|b\| \leqq\|c\|$ so that $y_{1}, y_{2}$ are well defined. If $x \in B(H)$ and $x^{\prime}$ is an extension of $x$ to $H \cup K$ such that $\left\|x^{\prime}\right\|=\|x\|$ then $\|a(x)\| \leqq$ $\left\|c\left(x^{\prime}\right)\right\| \leqq\|c\| \cdot\|x\|$. Hence, $\|a\| \leqq\|c\|$. Now, let $x \in B(K)$ and let $x^{\prime}$ be the extension of $x$ such that $x^{\prime} \mid H$ is the zero function. We have $c\left(x^{\prime}\right) \mid K=b(x)$ so that $\|b(x)\| \leqq$ $\left\|c\left(x^{\prime}\right)\right\| \leqq\|c\| \cdot\|x\|$ and $\|b\| \leqq\|c\|$.

We now show that if $y_{3}$ satisfies (6.5), then $y_{3} \mid H=y_{1}$ and $y_{3} \mid K=y_{2}$ are the solutions of (6.3) and (6.4) respectively. We have

$$
\begin{aligned}
y_{3} \mid H & =\eta_{0} e_{H}+\left(c \times\left(f \circ y_{3}\right)\right) \mid H \\
& =\eta_{0} e_{H}+a \times\left(\left(f \circ y_{3}\right) \mid H\right) \\
& =\eta_{0} e_{H}+a \times\left(f \circ\left(y_{3} \mid H\right)\right) .
\end{aligned}
$$

Note that in this computation we made use of the result that $(c \times y) \mid H=a \times(y \mid H)$ where $y$ is a bounded function on $H \cup K$ to $X$. This is easily verified by noting that, for any $p \in X^{*}$,

$$
\begin{aligned}
p \circ((c \times y) \mid H) & =(c(p \circ y)) \mid H=a(p \circ y \mid H) \\
& =a(p \circ(y \mid H))=p \circ(a \times(y \mid H))
\end{aligned}
$$

We now perform a similar calculation with $K$ taking the place of $H$,

$$
\begin{aligned}
y_{3} \mid K & =\eta_{0} e_{K}+\left(c \times\left(f \circ y_{3}\right)\right) \mid K \\
& =\eta_{0} e_{K}+a \times\left(f \circ y_{3} \mid H\right)\left(h_{1}\right) e_{K}+b \times\left(\left(f \circ y_{3}\right) \mid K\right) \\
& =\eta_{1} e_{K}+b \times\left(f \circ\left(y_{3} \mid K\right)\right),
\end{aligned}
$$

where we have used that fact that $y_{3} \mid H$ and $y_{1}$ are identical.
For a method $m=\left(H, a, h_{1}\right)$, we may form the numbers $\nu_{a}(t)\left(h_{1}\right)$ for $t \in T$.
Definition 6.2. We define a function $w$ on the class of all integration methods to the set $G$ defined in Section 2 by $\nu_{a}(t)\left(h_{1}\right)=w(m)(t)$ for all $t \in T$. We denote by $G_{0}$, the range of $w$.

Theorem 6.3. If $m_{1}, m_{2}$ are integration methods, then

$$
\begin{equation*}
w\left(m_{1} m_{2}\right)=w\left(m_{1}\right) w\left(m_{2}\right) . \tag{6.6}
\end{equation*}
$$

Proof. The multiplication on the right of (6.6) is, of course, that of the group $G$ in Section 2. We will write $m_{1}=\left(H, a, h_{1}\right), m_{2}=\left(K, b, k_{1}\right)$ so that $m_{1} m_{2}=(H \cup K$, $c, k_{1}$ ) where $c$ is given by (6.1) or (6.2).

The proof will require a use of the function $\lambda$ defined in Section 2. We recall that for a given $g \in G$ and $t \in T, \lambda(g, t)$ is a member of $G^{*}$. That is, $\lambda(g, t)=$ $\sum_{t^{\prime} \in T} l_{t},\left(t^{\prime}\right)^{\wedge}$ where for each $t$, only a finite number of $t^{\prime} \in T$ are such that $l_{t}, \neq 0$. We write $g=w\left(m_{1}\right)$ so that the numbers $l_{t t}$, for $t, t^{\prime} \in T$ take on definite values.

We will now show by induction that, for all $t \in T$,

$$
\begin{align*}
& \mu_{c}(t) \mid H=\mu_{a}(t)  \tag{6.7}\\
& \mu_{c}(t) \mid K=\sum_{t^{\prime}} l_{t^{\prime}, \mu_{b}\left(t^{\prime}\right)}  \tag{6.8}\\
& \nu_{c}(t) \mid H=\nu_{a}(t)  \tag{6.9}\\
& \nu_{c}(t) \mid K=\sum_{t^{\prime}} l_{t t^{\prime}} \nu_{b}\left(t^{\prime}\right)+w\left(m_{1}\right)(t) e_{K} . \tag{6.10}
\end{align*}
$$

For $t=\tau$, we have by (3.4), $\mu_{c}(t)=e_{H \cup K}, \mu_{a}(t)=e_{H}, \mu_{b}(t)=e_{K}$ and from (2.17) we see that $l_{t t^{\prime}}=1$ if $t^{\prime}=\tau$ and $l_{t t^{\prime}}=0$ otherwise. Thus, (6.7), (6.8) are easy to verify in this case.

For any $t \in T$, if (6.7), (6.8) are known to hold then (6.9) and (6.10) follow by a calculation (which we will omit) based on (3.6) which gives in this case $\nu_{c}(t)=$ $c\left(\mu_{c}(t)\right)$.

To complete the proof of (6.7), (6.8), (6.9), (6.10) we will prove (6.7), (6.8) for $t=u v$ on the assumption that the four equations hold for $t=u$ and for $t=v$ where $u, v \in T$. From (3.5) we have $\mu_{c}(u v)=\mu_{c}(u) \nu_{c}(v)$ (with pointwise multiplication on the right-hand side) so that

$$
\begin{aligned}
\mu_{c}(u v) \mid H & =\left(\mu_{c}(u) \nu_{c}(v)\right) \mid H=\left(\mu_{c}(u) \mid H\right)\left(\nu_{c}(v) \mid H\right) \\
& =\mu_{a}(u) \nu_{a}(v)=\mu_{a}(u v)
\end{aligned}
$$

which is (6.7).
For $k \in K$ and for $g^{\prime} \in G$ defined by $g^{\prime}\left(t^{\prime}\right)=\mu_{b}\left(t^{\prime}\right)(k)$ for all $t^{\prime} \in T$, we also have

$$
\begin{aligned}
\mu_{c}(u v)(k) & =\left(\sum_{u^{\prime}} l_{u u^{\prime}} \mu_{b}\left(u^{\prime}\right)\left(\sum_{v^{\prime}} l_{v v} \nu_{b}\left(v^{\prime}\right)+w\left(m_{1}\right)(v) e_{K}\right)\right)(k) \\
& =\left(\sum_{u^{\prime}, v^{\prime}} l_{u u^{\prime}} l_{v v^{\prime}} \mu_{b}\left(u^{\prime} v^{\prime}\right)+w\left(m_{1}\right)(v) \sum_{u^{\prime}} l_{u u^{\prime}} \mu_{b}\left(u^{\prime}\right)\right)(k) \\
& =\left(\lambda\left(w\left(m_{1}\right), u\right) \lambda\left(w\left(m_{1}\right), v\right)+w\left(m_{1}\right)(v) \lambda\left(w\left(m_{1}\right), u\right)\right)\left(g^{\prime}\right) \\
& =\lambda\left(w\left(m_{1}\right), u v\right)\left(g^{\prime}\right) \\
& =\sum_{t^{\prime}} l_{u v, t^{\prime}} \mu_{b}\left(t^{\prime}\right)(k),
\end{aligned}
$$

so that (6.8) follows. Now, evaluating (6.10) at $k_{1}$ we have

$$
w\left(m_{1} m_{2}\right)(t)=w\left(m_{1}\right)(t)+\lambda\left(w\left(m_{1}\right), t\right)\left(w\left(m_{2}\right)\right)=\left(w\left(m_{1}\right) w\left(m_{2}\right)\right)(t)
$$

by (2.16).
Since $w\left(m_{1} m_{2}\right)$ can be determined just from $w\left(m_{1}\right)$ and $w\left(m_{2}\right)$ and making use of no other property of $m_{1}$ or $m_{2}$, we have the following theorem which we state without detailed proof.

Theorem 6.4. If $M_{1}, M_{2}$ are equivalent classes of integration methods then $\left\{m_{1} m_{2}\right.$ : $\left.m_{1} \in M_{1}, m_{2} \in M_{2}\right\}$ is an equivalent class of integration methods.

We now see fairly precisely the sense in which $w(m)$ characterises $m$. If $w(m)$ is given, $m$ is specified up to an equivalence class. Furthermore, the numerical result produced by a method is determined by $w(m)$ (for $f$ with small enough Lipschitz constant). We also have the convenient results of Theorem 6.3 that the products in $G_{0}$ correspond to successive applications of different numerical methods.

We might ask the questions: are there methods corresponding to the identity elements of $G_{0}$ and to $w(m)^{-1}$ where $m$ is a given method? The answers to both these questions are in the affirmative as shown by Lemma 6.5 below.

Suppose $m=\left(H, a, h_{1}\right)$ and $h_{0} \notin H$ then we will be concerned with two methods. The first is

$$
\begin{equation*}
m_{0}=(\{0\}, b, 0) \tag{6.11}
\end{equation*}
$$

where $b \in A(\{0\})$ is defined for $x \in B(\{0\})$ by $b(x)(0)=0$ and the second is

$$
\begin{equation*}
m_{1}=\left(H \cup\left\{h_{0}\right\}, c, h_{0}\right) \tag{6.12}
\end{equation*}
$$

where $c \in A\left(H \cup\left\{h_{0}\right\}\right)$ is defined for $x \in B\left(H \cup\left\{h_{0}\right\}\right)$ by

$$
\begin{aligned}
c(x)\left(h_{0}\right) & =-a(x \mid H)\left(h_{1}\right) \\
c(x) \mid H & =a(x \mid H)-a(x \mid H)\left(h_{1}\right) e_{H}
\end{aligned}
$$

Lemma 6.5. If $m_{0}$ is given by (6.11), then $w\left(m_{0}\right)$ is the identity of $G$. If $m_{1}$ is given by (6.12) then $w\left(m_{1}\right)=w(m)^{-1}$.

Proof. It is trivial that $w\left(m_{0}\right)(t)=0$ for all $t \in T$. Hence, $w\left(m_{0}\right)$ is the identity of $G$.
We now form the product of $m$ amd $m_{1}$. However, it is necessary to replace $m_{1}$ by an identical method, say ( $H^{\prime} \cup\left\{h_{0}\right\}, c, h_{0}$ ), where $H^{\prime}$ is a copy of $H$ which we regard as disjoint from it. If $h \in H$, we write $h^{\prime} \in H^{\prime}$ for its partner in a bijection between $H$ and $H^{\prime}$. The product method is $\left(H \cup H^{\prime} \cup\left\{h_{0}\right\}, d, h_{0}\right)$ where $d$ is defined for $x \in B\left(H \cup H^{\prime} \cup\left\{h_{0}\right\}\right)$ by

$$
\begin{aligned}
d(x) \mid H & =a(x \mid H) \\
d(x) \mid H^{\prime} & =a(x \mid H)\left(h_{1}\right) e_{H^{\prime}}+a^{\prime}\left(x \mid H^{\prime}\right)-a^{\prime}\left(x \mid H^{\prime}\right)\left(h_{1}^{\prime}\right) e_{H^{\prime}}, \\
d(x)\left(h_{0}\right) & =a(x \mid H)\left(h_{1}\right)-a^{\prime}\left(x \mid H^{\prime}\right)\left(h_{1}^{\prime}\right)
\end{aligned}
$$

and $a^{\prime}$ corresponds to $a$ in the identification of ( $H, a, h_{1}$ ) and ( $H^{\prime}, a^{\prime}, h_{1}^{\prime}$ ).
Let $B^{\prime} \subseteq B\left(H \cup H^{\prime} \cup\left\{h_{0}\right\}\right)$ and $B^{\prime \prime} \subseteq B^{\prime}$ be the sets defined so that if $x \in B^{\prime}$ then for all $h \in H, x(h)=x\left(h^{\prime}\right)$ and if $x \in B^{\prime \prime}$ we have the additional property that $x\left(h_{0}\right)=0$, then it is easy to verify that if $x \in B^{\prime}$ then $d(x) \in B^{\prime \prime}$. We also see that $B^{\prime}$ is a subalgebra of $B\left(H \cup H^{\prime} \cup\left\{h_{0}\right\}\right)$ (regarded as a Banach algebra with pointwise multiplication) and $B^{\prime \prime}$ is an ideal of $B^{\prime}$. We shall prove by induction that $\mu_{d}(t) \in$ $B^{\prime}$ and $\nu_{d}(t) \in B^{\prime \prime}$ for all $t \in T$. In fact, since $e_{H \cup H \prime \cup\left(h_{0}\right)} \in B^{\prime}$ we have that $\mu_{d}(\tau) \in B^{\prime}$. Hence, $\nu_{d}(\tau)=d\left(\mu_{d}(\tau)\right) \in B^{\prime \prime}$. If $t=u v$, where $\mu_{d}(u) \in B^{\prime}, \nu_{d}(v) \in B^{\prime \prime}$ then $\mu_{d}(u v)=$ $\mu_{d}(u) \nu_{d}(v) \in B^{\prime}$ and hence, $\nu_{d}(u v) \in B^{\prime \prime}$. Evaluating $\nu_{d}(t)$ at $h_{0}$ for any $t \in T$ gives the result 0 . Hence, $w\left(m m_{1}\right)=i$ (the identity for $G$ ). Hence, $w\left(m_{1}\right)=w(m)^{-1}$.

The following theorem will be stated without proof. It is a corollary to Lemma 6.5. Theorem 6.6. $G_{0}$ is a subgroup of $G$.
It is clear that $G_{0}$ is not identical with $G$ since an element of $G_{0}$ satisfies the condition of Theorem 6.8 below. In this result and in Lemma 6.7, $r$ will denote the function on $T$ which assigns to $t \in T$ the order of $t$. Thus, $r$ will satisfy the recursive definition

$$
\begin{align*}
r(\tau) & =1 \\
r(u v) & =r(u)+r(v), \quad \text { for all } u, v \in T \tag{6.13}
\end{align*}
$$

Lemma 6.7. If $m=\left(H, a, h_{1}\right)$ then for all $t \in T$,

$$
\begin{equation*}
|w(m)(t)| \leqq\|a\|^{r(t)} \tag{6.14}
\end{equation*}
$$

Proof. In fact, we will prove that, for all $t \in T$,
where, if $\|a\|=0$ and $t=\tau$, then the right-hand side of (6.15) is replaced by 1 .
For $t=\tau$, (6.15) is trivial since each side equals 1 ; for all $t \in T$, (6.16) follows from (6.15) since

$$
\left\|\nu_{a}(t)\right\|=\left\|a\left(\mu_{a}(t)\right)\right\| \leqq\|a\| \cdot\left\|\mu_{a}(t)\right\| .
$$

Furthermore, if $t=u v$ then

$$
\left\|\mu_{a}(u v)\right\|=\left\|\mu_{a}(u) \nu_{a}(v)\right\| \leqq\left\|\mu_{a}(u)\right\| \cdot\left\|\nu_{a}(v)\right\|
$$

and an inductive argument completes the proof of (6.15) and (6.16). Equation (6.14) now follows since $w(m)(t)=\nu_{a}(t)\left(h_{1}\right)$.

The theorem which now follows restates this result without making explicit reference to $m$.

Theorem 6.8. If $\alpha \in G_{0}$, then the set $\left\{|\alpha(t)|^{1 / r(t)}: t \in T\right\}$ is bounded.
Proof. Write $\alpha=w(m)$ where $m$ is as in Lemma 6.7. The result holds with an upper bound $\|a\|$.

Although $G_{0} \neq G$, there is a sense in which elements of $G$ can be approximated by elements of $G_{0}$. The result is as follows.

Theorem 6.9. If $\alpha \in G$ and $T_{0}$ is any finite subset of $T$, then there is $\beta \in G_{0}$ such that $\alpha\left|T_{0}=\beta\right| T_{0}$.

Proof. It is sufficient to prove the result in each of the cases $T_{0}=T_{1}, T_{2}, \cdots$ where for $n=1,2, \cdots, T_{n}$ is the set of trees of order $\leqq n$. We will prove this result (which is trivial for $n=1$ ) by induction on $n$.

We first define a function $\delta$ on $T \times T$ by the formulae

$$
\begin{array}{lll}
\delta(\tau, t)=0, & & \text { for all } t \in T \\
\delta(u v, t) & =\delta(u, t), &  \tag{6.17}\\
\text { where } u, v, t \in T \text { and } v \neq t \\
\delta(u v, t)=\delta(u, t)+1, & & \text { where } u, v, t \in T \text { and } v=t .
\end{array}
$$

If $r(t) \leqq r(u)$, we see that $\delta(t, u)=0$ and if for all $u \in T, \delta\left(t_{1}, u\right)=\delta\left(t_{2}, u\right)$, then $t_{1}=t_{2}$. Also, it can be verified that, for a method $m_{1}=\left(H, a_{1}, h_{1}\right)$, if $t \in T_{n}$ (where $n>1$ ), then

$$
\begin{equation*}
\mu_{a_{1}}(t)=\prod_{u \in T_{n-1}} \nu_{a_{1}}(u)^{\delta(t, u)} \tag{6.18}
\end{equation*}
$$

From the method $m_{1}$, we can form a new method $m_{2}=\left(H \cup\left\{h_{2}\right\}, a_{2}, h_{2}\right)$ where $h_{2} \notin H$ and $a_{2}$ is defined by

$$
a_{2}(x) \mid H=a_{1}(x \mid H), \quad a_{2}(x)\left(h_{2}\right)=x\left(h_{1}\right)
$$

It can be verified that for all $t \in T$ we have

$$
\nu_{a_{2}}(t) \mid H=\nu_{a_{1}}(t), \quad \nu_{a_{2}}(t)\left(h_{2}\right)=\mu_{a_{1}}(t)\left(h_{1}\right)
$$

Hence, using (6.18) we see that, for $t \in T_{n}$,

$$
\begin{equation*}
w\left(m_{2}\right)(t)=\prod_{u \in T_{n-1}}\left(w\left(m_{1}\right)(u)\right)^{\delta(t, u)} \tag{6.19}
\end{equation*}
$$

By the induction hypothesis, a method $m_{1}$ exists so that $w\left(m_{1}\right)$ takes on any required values on $T_{n-1}$. Let $S_{1}$ denote the set of all real-valued functions on $T_{n}$ which satisfy (6.19) for some $w\left(m_{1}\right) \mid T_{n-1}$. Also let $S_{2}$ denote the linear span of $S_{1}$, let $S_{3}$ denote the set of all real-valued functions which are restrictions to $T_{n}$ of members of $G_{0}$ and let $S_{4}$ denote the set of all real-valued functions on $T_{n}$. By definition, $S_{3} \subseteq S_{4}$ and we have already shown that $S_{1} \subseteq S_{3}$.

We now see that $S_{3}$ is a linear subspace of $S_{4}$. This follows by observing that if $c \in R, \alpha_{1}=w\left(m_{1}\right), \alpha_{2}=w\left(m_{2}\right), m_{1}=\left(H_{1}, a_{1}, h_{1}\right)$ and $m_{2}=\left(H_{2}, a_{2}, h_{2}\right)$ where $H_{1} \cap$ $H_{2}=\varnothing$ then $c \alpha_{1}=w\left(m_{3}\right)$ and $\alpha_{1}+\alpha_{2}=w\left(m_{4}\right)$ where

$$
m_{3}=\left(H_{1} \cup\left\{h_{3}\right\}, a_{3}, h_{3}\right), \quad m_{4}=\left(H_{1} \cup H_{2} \cup\left\{h_{4}\right\}, a_{4}, h_{4}\right),
$$

$h_{3} \notin H_{1}, h_{4} \notin H_{1} \cup H_{2}$, and $a_{3}, a_{4}$ are defined by

$$
a_{3}(x) \mid H_{1}=a_{1}\left(x \mid H_{1}\right), \quad a_{3}(x)\left(h_{3}\right)=c a_{1}\left(x \mid H_{1}\right)\left(h_{1}\right),
$$

$$
\begin{aligned}
a_{4}(x) \mid H_{1} & =a_{1}\left(x \mid H_{1}\right), \quad a_{4}(x) \mid H_{4}=a_{2}\left(x \mid H_{2}\right), \\
a_{4}(x)\left(h_{4}\right) & =a_{1}\left(x \mid H_{1}\right)\left(h_{1}\right)+a_{2}\left(x \mid H_{2}\right)\left(h_{2}\right) .
\end{aligned}
$$

Thus, we see that $S_{2} \subseteq S_{3} \subseteq S_{4}$ so that $S_{3}=S_{4}$ will follow when we have shown that $S_{2}=S_{4}$.

We first show that $S_{2}$ is a subalgebra of $S_{4}$ in the sense of pointwise multiplication. Let $\xi, \eta \in S_{2}$ be defined by

$$
\begin{align*}
& \xi(t)=\sum_{i \in I} c_{i} \prod_{u \in T_{n-1}} C_{i}(u)^{\delta(t, u)}  \tag{6.20}\\
& \eta(t)=\sum_{i \in J} d_{i} \prod_{u \in T_{n-1}} D_{i}(u)^{\delta(t, u)}
\end{align*}
$$

for all $t \in T_{n}$ where $I, J$ are finite index sets and for $i \in I$ and $j \in J, c_{i}$ and $d_{i}$ are real coefficients and $C_{i}, D_{i}$ are real-valued functions on $T_{n-1}$.

The (pointwise) product $\xi \eta$ is now defined by

$$
(\xi \eta)(t)=\xi(t) \eta(t)=\sum_{(i, i) \in I \times J} c_{i} d_{i} \prod_{u \in T_{n-1}}\left(C_{i}(u) D_{i}(u)\right)^{\delta(t, u)}
$$

so that $\xi \eta \in S_{2}$.
We next observe that the subalgebra $S_{2}$ distinguishes points of $T_{n}$. That is, for $t_{1}, t_{2} \in T_{n}$ with $t_{1} \neq t_{2}$ there is $\xi \in S_{2}$ such that $\xi\left(t_{1}\right) \neq \xi\left(t_{2}\right)$. If $u \in T_{n-1}$ is such that $\delta\left(t_{1}, u\right) \neq \delta\left(t_{2}, u\right)$, then we can, for example, define $\xi$ by (6.20) with $I=\{1\}$, $c_{1}=1, C_{1}(v)=1$ for $v \neq u$ and $C_{1}(u)=2$. Furthermore, $S_{2}$ contains the unit function since this is just $\xi$ defined in (6.20) with $I=\{1\}, c_{1}=1$ and $C_{1}(u)=1$ for all $u$.

Hence, $S_{2}=S_{4}$ and the proof of Theorem 6.9 is complete.
7. Some Properties of $G_{0}$ and $G$. In this section, we will develop some propties of some special elements of $G_{0}$, some of its subgroups that have a numerical significance and factor groups corresponding to some important normal subgroups. Since $G$ is, in some ways, simpler than $G_{0}$ it will be convenient to study properties of $G_{0}$ in terms of the corresponding properties of $G$.

For example, if $K$ is an invariant subgroup of $G$ such that every element of $G$ can be written as the product of an element in $K$ and an element in $G_{0}$ then the factor group $G_{0} / K \cap G_{0}$ is isomorphic to $G / K$. Thus, we can avoid explicit consideration of the group $G_{0} / K \cap G_{0}$ by considering $G / K$ instead.

It was mentioned in Section 3 that if $H$ is a finite set then the integration method $m=\left(H, a, h_{1}\right)$ is simply a Runge-Kutta method. Write $H=\{1,2, \cdots, n+1\}$ and define $a$ as the linear operator represented by the matrix

$$
\left[\begin{array}{ccccc}
a_{11} & a_{12} & \cdots & a_{1 n} & 0 \\
a_{21} & a_{22} & \cdots & a_{2 n} & 0 \\
\vdots & \vdots & & \vdots & \vdots \\
a_{n 1} & a_{n 2} & \cdots & a_{n n} & 0 \\
b_{1} & b_{2} & \cdots & b_{n} & 0
\end{array}\right]
$$

and write $h_{1}=n+1$. This leads to the integration method represented by the equations

$$
\begin{aligned}
& y(1)=\eta_{0}+\sum_{i} a_{1 i} f(y(j)), \\
& y(2)=\eta_{0}+\sum_{i} a_{2 i} f(y(j)), \\
& \vdots \\
& y(n)=\eta_{0}+\sum_{i} a_{n i} f(y(j)), \\
& \eta_{1}=\eta_{0}+\sum_{i} b_{i} f(y(j)) .
\end{aligned}
$$

It is easy to see that $w(m)$ maps a given tree $t$ onto the corresponding polynomial $\Phi$ as defined and used in [1], [2].

The special method $m=([0,1], a, 1)$, where $a$ is given by (3.2), will be referred to as the Picard method and we will write $p=w(m)$ in this case. Let $r: T \rightarrow R$ be defined by (6.13) so that $r(t)$ is the order of $t$. We then see that, according to [1] in which for a given $t, p(t)$ was written as $1 / \gamma, p$ would satisfy the recursion given by (7.3), (7.4) below. More detailed information is given in the following theorem.

Theorem 7.1. For the Picard method we have

$$
\begin{align*}
\mu_{a}(t)(h) & =p(t) r(t) h^{r(t)-1},  \tag{7.1}\\
\nu_{a}(t)(h) & =p(t) h^{r(t)}, \tag{7.2}
\end{align*}
$$

for $h \in[0,1], t \in T$ where $p$ is defined by

$$
\begin{align*}
p(\tau) & =1  \tag{7.3}\\
p(u v) & =p(u) p(v) r(u) / r(u v) \quad \text { for } u, v \in T \tag{7.4}
\end{align*}
$$

Proof. We shall verify (7.1) and (7.2) when $t=\tau$ and when $t=u 0$ given that they hold for $t=u$ and $t=0$. For any $t$, (7.2) follows from (7.1) by the formula

$$
\nu_{a}(t)(h)=\int_{0}^{h} \mu_{a}(t)(k) d k
$$

For $t=\tau$, we have $\mu_{a}(t)(h)=1$ which is the correct result. We also have, for $t=u v$,

$$
\begin{aligned}
\mu_{a}(u v)(h) & =\mu_{a}(u)(h) \nu_{a}(v)(h) \\
& =p(u) r(u) h^{r(u)-1} p(v) h^{r(v)} \\
& =p(u) p(v) r(u) r(u v) h^{r(u v)-1} / r(u v) \\
& =p(u v) r(u v) h^{r(u v)-1}
\end{aligned}
$$

so that (7.1) is proved in this case.
For a real number $A$, we define, for $\alpha \in G$, the element $\alpha^{(\Lambda)} \in G$ such that, for $t \in T$,

$$
\begin{equation*}
\alpha^{(A)}(t)=A^{r(t)} \alpha(t) . \tag{7.5}
\end{equation*}
$$

This definition is motivated by the observation that if $\alpha=w(m)$ for $m=\left(H, a, h_{1}\right)$ then $\alpha^{(\Delta)}=w\left(m^{\prime}\right)$ where $m^{\prime}$ is the same as $m$ except that $a$ is replaced by $A a$. This corresponds to a change in step size by a factor $A$.

The next result is concerned with a special property possessed by $p$.
Throrem 7.2. For $A, B \in R$, we have

$$
\begin{equation*}
p^{(A)} p^{(B)}=p^{(A+B)} \tag{7.6}
\end{equation*}
$$

Proof.

$$
p^{(A)}=w(([0,1], a, 1)), \quad p^{(B)}=w(([2,3], b, 3))
$$

where for $x \in B[0,1], y \in B[2,3], a$ and $b$ are defined by

$$
a(x)(h)=A \int_{0}^{h} x(\xi) d \xi, \quad b(y)(k)=B \int_{2}^{k} y(\xi) d \xi
$$

for $h \in[0,1], k \in[2,3]$.
Hence, $p^{(\Lambda)} p^{(B)}=w(([0,1] \cup[2,3], c, 3))$ where $c$ is defined for $x \in$ $B([0,1] \cup[2,3])$ by

$$
c(x)(h)=A \int_{0}^{h} x(\xi) d \xi, \quad c(x)(k)=A \int_{0}^{1} x(\xi) d \xi+B \int_{2}^{k} x(\xi) d \xi
$$

for $h \in[0,1]$ and $k \in[2,3]$.
As for the proof of Theorem 7.1, it can be verified that for all $t \in T, h \in[0,1]$ and $k \in[2,3]$ that

$$
\begin{align*}
& \mu_{c}(t)(h)=p(t) r(t)(A h)^{r(t)-1} \\
& \nu_{c}(t)(h)=p(t)(A h)^{r(t)} \\
& \mu_{c}(t)(k)=p(t) r(t)(A+B(k-2))^{r(t)-1} \\
& \nu_{c}(t)(k)=p(t)(A+B(k-2))^{r(t)} \tag{7.7}
\end{align*}
$$

Setting $k=3$ in (7.7) we obtain the result.
This result is characteristic of $p$. For example, we have
Theorem 7.3. If $\alpha \in G$ is such that $\alpha^{2}=\alpha^{(2)}$ and $\alpha(\tau)=1$, then $\alpha=p$.
Proof. The expression for $\alpha^{2}(t)$ is $2 \alpha(t)$ plus terms involving trees of lower order. On the other hand, $\alpha^{(2)}(t)$ equals $2^{r(t)} \alpha(t)$ so that for $r(t)>1$ there is only one possible value for $\alpha(t)$ so that the two expressions are equal. For $r(t)=1$, we have only $t=\tau$.

At this point, we introduce a certain type of subgroup of $G$.
Definition 7.4. For $n$ a positive integer, we define $K_{n}$ as the subset of $G$ such that for $\alpha \in K_{n}$ and any $t \in T$ such that $r(t) \leqq n$ we have $\alpha(t)=0$.

Theorem 7.5. $K_{n}$ is an invariant subgroup of $G$.
Proof. If $\alpha \in K_{n}$ we have, for all $t$ such that $r(t) \leqq n, \lambda(\alpha, t)=\hat{t}$. This follows since $\lambda(\alpha, \tau)=\hat{\tau}$ and if $r(u), r(v) \leqq n$ then

$$
\lambda(\alpha, u v)=\alpha(v) \lambda(\alpha, u)+\lambda(\alpha, u) \lambda(\alpha, v)=(u v)
$$

(Note that $\lambda(\alpha, t)=\hat{t}$ also for $r(t)=n+1$.)
Hence, if $\alpha \in K_{n}$ and $r(t) \leqq n$ then

$$
(\alpha \beta)(t)=\alpha(t)+\hat{t}(\beta)=\beta(t)
$$

so that $(\alpha \beta)(t)=0$ if and only if $\beta(t)=0$. Thus, if $\alpha \in K_{n}$ then $\alpha \beta \in K_{n}$ if and only if $\beta \in K_{n}$. Thus $K_{n}$ is a subgroup.

To show that $K_{n}$ is a normal subgroup, we will prove that if $\beta \in G$ then the sets

$$
\left\{\alpha \beta: \alpha \in K_{n}\right\} \quad \text { and } \quad\left\{\beta \alpha: \alpha \in K_{n}\right\}
$$

are identical. We shall show, in fact, that each of these sets is the subset of $G$ containing $\gamma \in G$ if and only if for all $t \in T$ such that $r(t) \leqq n$ we have $\gamma(t)=\beta(t)$. Call this set $S$. We must show that
(i) $\alpha \in K_{n} \Rightarrow \alpha \beta \in S$,
(ii) $\alpha \in K_{n} \Rightarrow \beta \alpha \in S$,
(iii) $\gamma \in S \Rightarrow\left(\exists \alpha \in K_{n}\right)(\alpha \beta=\gamma)$,
(iv) $\gamma \in S \Rightarrow\left(\exists \alpha \in K_{n}\right)(\beta \alpha=\gamma)$.

Proof of (i). We have already seen that if $\alpha \in K_{n}$ then $(\alpha \beta)(t)=\beta(t)$.
Proof of (ii). For fixed $\beta$, we have

$$
\lambda(\beta, t)=\sum_{r\left(t^{\prime}\right)<r(t)} c\left(t^{\prime}\right)\left(t^{\prime}\right) \hat{\imath}+\hat{i}
$$

where $c$ is some real-valued function on $T$. Hence, if $\alpha \in K_{n}$ and $r(t) \leqq n$ then $\lambda(\beta, t)(\alpha)=0$. Hence, $(\beta \alpha)(t)=\beta(t)$.

Proof of (iii). This is equivalent to (i) since $\beta=\alpha^{-1} \gamma$ and $\alpha^{-1} \in K_{n}$.
Proof of (iv). Similarly, this is equivalent to (ii).
The factor group $G / K_{n}\left(=G_{0} /\left(K_{n} \cap G_{0}\right)\right)$ will be denoted by $G_{n}$. We use the usual group-theoretic convention of writing $\alpha K_{n}$ where $\alpha \in G$ for the coset containing $\alpha$.

Definition 7.6. If $\alpha \in p K_{n}$, then we say that $\alpha$ is of order $n$. The greatest number $n$ such that $\alpha$ is of order $n$ is called the order of $\alpha$.

Note that we will use the word "order" as applying to a method $m$ such that $w(m)=\alpha$ in the same way as it would apply to $\alpha$. Sometimes it is convenient to generalise the concept of order and we say that a method $m$ or $\alpha=w(m)$ is of generalised order $n$ if $\alpha \in p^{(\alpha(\tau))} K_{n}$ and we call the greatest such $n$ the generalised order of $\alpha$ (or of $m$ ).

Definition 7.7. Let $J_{n}$ denote the set of elements of $G$ of generalised order $n$. That is

$$
J_{n}=\bigcup\left\{p^{(\theta)} K_{n}: \theta \in R\right\} .
$$

We have the following property of generalised order, which is not possessed by order.
Theorem 7.8. $J_{n}$ is a subgroup of $G$.
Proof. Suppose $\alpha \in p^{(A)} K_{n}, \beta \in p^{(B)} K_{n}=K_{n} p^{(B)}$ for real numbers $A$ and $B$. Then $\alpha \beta \in p^{(A)} K_{n} K_{n} p^{(B)}$. Since $K_{n} K_{n}=K_{n}$ and $K_{n} p^{(B)}=p^{(B)} K_{n}$ we have

$$
\alpha \beta \in p^{(A)} p^{(B)} K_{n} .
$$

However, by Theorem 7.2, $p^{(A)} p^{(B)}=p^{(A+B)}$ and it follows that $\alpha \beta \in J_{n}$. To prove that $\alpha^{-1} \in J_{n}$ where $\alpha \in p^{(A)} K_{n}$ we will show that $\alpha^{-1} \in p^{(-1)} K_{n}$. Let $\alpha=p^{(A)} k$ where $k \in K_{n}$ then

$$
\begin{aligned}
\alpha p^{(-\Lambda)} K_{n} & =p^{(\Lambda)} k p^{(-\Lambda)} K_{n}=p^{(\Lambda)} k K_{n} p^{(-\Lambda)} \\
& =p^{(\Lambda)} K_{n} p^{(-\Lambda)}=p^{(\Lambda)} p^{(-\Lambda)} K_{n}=K_{n}
\end{aligned}
$$

and this last set contains the identity of $G$.

We now introduce some other subgroups of $G$ that will be useful in the search for special numerical methods. First, however, we introduce some notational devices. For $t \in T$, we will, in this section, define $t_{0}, t_{1}, t_{2}, \cdots$ by

$$
\begin{align*}
& t_{0}=t  \tag{7.8}\\
& t_{n}=t_{n-1} \tau \quad(n=1,2, \cdots)
\end{align*}
$$

We recall that $\lambda: G \times T \rightarrow G^{*}$ was defined in Section 2. We now define $\Lambda: G \times$ $G^{*} \rightarrow G^{*}$ in such a way that $\Lambda$ is linear over $G^{*}$ and such that $\Lambda(\alpha, \hat{i})=\lambda(\alpha, t)$ whenever $t \in T$.

Definition 7.9. The function $\Lambda: G \times G^{*} \rightarrow G^{*}$ is defined by

$$
\begin{equation*}
\Lambda\left(\alpha, \sum_{t \in T} c(t) \hat{t}\right)=\sum_{t \in T} c(t) \lambda(\alpha, t) \tag{7.9}
\end{equation*}
$$

where $\alpha \in G$ and $c: T \rightarrow R$ takes all but a finite subset of $T$ to zero.
The following properties of $\Lambda$ that will be stated without proof are easily verified by substituting $\Lambda$ in terms of $\lambda$ in each formula.

Theorem 7.10. If $U \in G^{*}$ and $\alpha, \beta \in G$ then

$$
\begin{equation*}
U(\alpha \beta)=U(\alpha)+\Lambda(\alpha, U)(\beta) . \tag{7.10}
\end{equation*}
$$

Furthermore, if $U, V \in G^{*}, \alpha \in G$ and $c \in R$ then

$$
\begin{equation*}
\Lambda(\alpha, \hat{\tau})=\hat{\tau} \tag{7.11}
\end{equation*}
$$

$$
\begin{align*}
\Lambda(\alpha, U+V) & =\Lambda(\alpha, U)+\Lambda(\alpha, V)  \tag{7.12}\\
\Lambda(\alpha, c U) & =c \Lambda(\alpha, U)  \tag{7.13}\\
\Lambda(\alpha, U V) & =\Lambda(\alpha, U) \Lambda(\alpha, V)+V(\alpha) \Lambda(\alpha, U) \tag{7.14}
\end{align*}
$$

Let $L_{n}(n=1,2, \cdots)$ denote the subset of $G$ defined as that $\alpha \in L_{n}$ if and only if, for $m=0,1, \cdots, n-1$, it holds that

$$
\begin{equation*}
\alpha\left(\tau_{m}\right)=\alpha(\tau)^{m+1} /(m+1) \tag{7.15}
\end{equation*}
$$

Since $\tau_{0}=\tau, L_{1}$ is identical to $G$.
Definition 7.11. A subgroup $H$ of $G$ is said to be homogeneous if $\alpha \in H$ and $A \in R$ imply that $\alpha^{(A)} \in H$.

It is easy to see that $G_{0}, K_{n}(n=1,2, \cdots)$ and $J_{n}(n=1,2, \cdots)$ are all homogeneous.

We also have
Theorem 7.12. For $n=1,2, \cdots, L_{n}$ is a homogeneous subgroup of $G$ and $p \in L_{n}$.
Proof. From its definition, it is easy to see that $L_{n}$ is homogeneous and, from the definition of $p$ we see that $p$ satisfies (7.15). It remains to prove that $L_{n}$ is a subgroup.

We first show that, for $m=0,1,2, \cdots$ and $\alpha \in G$,

$$
\begin{equation*}
\lambda\left(\alpha, \tau_{m}\right)=\sum_{i=0}^{m}\binom{m}{i} \alpha(\tau)^{i} \hat{\tau}_{m-i} \tag{7.16}
\end{equation*}
$$

For $m=0$, both sides of (7.16) reduce to $\hat{\tau}$. To complete the inductive argument we assume the result for numbers lower than $m$. We have

$$
\begin{aligned}
\lambda\left(\alpha, \tau_{m}\right) & =\lambda\left(\alpha, \tau_{m-1} \tau\right)=\lambda\left(\alpha, \tau_{m-1}\right) \lambda(\alpha, \tau)+\alpha(\tau) \lambda\left(\alpha, \tau_{m-1}\right) \\
& =\sum_{i=0}^{m-1}\binom{m-1}{i} \alpha(\tau)^{i} \hat{\tau}_{m-1-i} \hat{\tau}+\alpha(\tau) \sum_{i=0}^{m-1}\binom{m-1}{i} \alpha(\tau)^{i} \hat{\tau}_{m-1-i} \\
& =\sum_{i=0}^{m}\binom{m}{i} \alpha(\tau)^{i} \hat{\tau}_{m-i} .
\end{aligned}
$$

Hence, if $\alpha, \beta \in G$,

$$
(\alpha \beta)\left(\tau_{m}\right)=\alpha\left(\tau_{m}\right)+\sum_{i=0}^{m}\binom{m}{i} \alpha(\tau)^{i} \beta\left(\tau_{m-i}\right)
$$

so that

$$
\begin{aligned}
(\alpha \beta)\left(\tau_{m}\right)- & \frac{1}{m+1}((\alpha \beta)(\tau))^{m+1} \\
& =\left(\alpha\left(\tau_{m}\right)-\frac{1}{m+1} \alpha(\tau)^{m+1}\right)+\sum_{i=0}^{m}\binom{m}{i} \alpha(\tau)^{i}\left(\beta\left(\tau_{m-i}\right)-\frac{\beta(\tau)^{m+1-i}}{m+1-i}\right)
\end{aligned}
$$

Hence, if $\beta \in L_{n}$ it follows that $\alpha \beta \in L_{n}$ if and only if $\alpha \in L_{n}$. Thus, $L_{n}$ is a subgroup.

The next subgroup to be defined is actually a subgroup of $L_{n}$. Let $U_{1}, U_{2}, \cdots$, $U_{n} \in G^{*}$ be defined by

$$
\begin{equation*}
U_{m}=m \hat{\tau}_{m-1}-\hat{\tau}_{m} \tag{7.17}
\end{equation*}
$$

so that a brief calculation gives for $\alpha \in G, m=1,2, \cdots, n$,

$$
\Lambda\left(\alpha, U_{m}\right)=\sum_{i=0}^{m-1}\binom{m}{i} \alpha(\tau)^{i} U_{m-i}+\left(m \alpha\left(\tau_{m-1}\right)-\alpha(\tau)^{m}\right) \hat{\tau}
$$

If $\alpha \in L_{n}$ we have

$$
\begin{equation*}
\Lambda\left(\alpha, U_{m}\right)=\sum_{i=0}^{m-1}\binom{m}{i} \alpha(\tau)^{i} U_{m-i} \tag{7.18}
\end{equation*}
$$

Definition 7.13. We define $S_{n}^{0}=\left\{U_{1}, U_{2}, \cdots, U_{n}\right\}$ and $S_{n}^{1}, S_{n}^{2}, \cdots$, recursively by the formula

$$
\begin{equation*}
S_{n}^{i}=\left\{\hat{t} V: t \in T, V \in S_{n}^{i-1}\right\} \cup\left\{V \hat{t}: t \in T, V \in S_{n}^{i-1}\right\} \tag{7.19}
\end{equation*}
$$

and we define $S_{n}$ as the linear subspace of $G^{*}$ spanned by the union of $S_{n}^{0}, S_{n}^{1}, \cdots$.
Finally, we define $L_{n, n}$ as the subset of $L_{n}$ containing $\alpha \in L_{n}$ if and only if $U(\alpha)=0$ for all $U \in S_{n}$, and $L_{m, n}=L_{m} \cap L_{n, n}$ for $m \geqq n$. The next two lemmas establish properties of $L_{n, n}$ that will enable us to prove it is a subgroup.

Lemma 7.14. If $\alpha \in L_{n, n}$ and $V \in S_{n}$ then

$$
\begin{equation*}
\Lambda(\alpha, V) \in S_{n} \tag{7.20}
\end{equation*}
$$

Proof. By the linearity of $\Lambda$, it is necessary to prove (7.20) only when $V \in S_{n}^{i}$ for some $i$. We prove this by induction on $i$. For $i=0$, the result follows from (7.18). For $i>0$, we assume the result has been proved for $V \in S_{n}^{i-1}$. Let $V=U W$ where $U \in S_{n}^{i-1}$ or $W \in S_{n}^{i-1}$. We have

$$
\begin{equation*}
\Lambda(\alpha, V)=\Lambda(\alpha, U) \Lambda(\alpha, W)+W(\alpha) \Lambda(\alpha, U) \tag{7.21}
\end{equation*}
$$

If $U \in S_{n}^{i-1}$, then both terms on the right of (7.21) are members of $S_{n}$. If $W \in S_{n}^{i-1}$, then $W(\alpha)=0$ so that again we see that $\Lambda(\alpha, V) \in S_{n}$.

Lemma 7.15. If $\alpha \in L_{n, n}$ and $\beta \in L_{n}$ are such that for all, $V \in S_{n}$,

$$
\begin{equation*}
\Lambda(\alpha, V)(\beta)=0, \tag{7.22}
\end{equation*}
$$

then $\beta \in L_{n, n}$.
Proof. By the construction of $\bigcup_{i} S_{n}^{i}$, its elements are of the form $m \hat{u}-\hat{0}$ where $u, v$ are of the same order (say $k$ ) and $m$ is an integer such that $m \leqq n$. Suppose that for all $m^{\prime} \hat{u}^{\prime}-\hat{v}^{\prime} \in \bigcup_{i} S_{n}^{i}$ such that $r\left(u^{\prime}\right)=r\left(v^{\prime}\right)<k$ it holds that $\left(m^{\prime} \hat{u}^{\prime}-\hat{v}^{\prime}\right)(\beta)=0$. Then, we will prove that $(m \hat{u}-\hat{v})(\beta)=0$.

Let $\beta_{0}$ be defined so that $\beta_{0}(t)=\beta(t)$ when $r(t)<k$ and $\beta_{0}(t)=0$ when $r(t) \geqq k$. Thus, $\beta_{0} \in L_{n, n}$ so that, by Lemma 7.14,

$$
\Lambda(\alpha, m \hat{u}-\hat{v})\left(\beta_{0}\right)=0=\Lambda(\alpha, m \hat{u}-\hat{v})(\beta)
$$

By writing $\lambda(\alpha, u)=\hat{u}+\sum_{r(t)<k} c(t) \hat{t}, \lambda(\alpha, v)=\hat{v}+\sum_{r(t)<k} d(t) \hat{t}$ and $e(t)=$ $m c(t)-d(t)$ we see that

$$
\Lambda(\alpha, m \hat{u}-\hat{v})=m \hat{u}-\hat{v}+\sum_{r(t)<k} e(t) \hat{t}
$$

so that

$$
\begin{aligned}
(m \hat{u}-\hat{v})(\beta) & =\Lambda(\alpha, m \hat{u}-\hat{v})(\beta)-\sum_{r(t)<k} e(t) \hat{t}(\beta) \\
& =\Lambda(\alpha, m \hat{u}-\hat{v})\left(\beta_{0}\right)-\sum_{r(t)<k} e(t) \beta_{0}(t) \\
& =(m \hat{u}-\hat{v})\left(\beta_{0}\right)=0
\end{aligned}
$$

Theorem 7.16. $L_{n, n}$ is a homogeneous subgroup of $L_{n}$ and $p \in L_{n, n}$.
Proof. If $\alpha \in L_{n, n}, U \in S_{n}$ and $\beta \in L_{n}$ we have

$$
U(\alpha \beta)=U(\alpha)+\Lambda(\alpha, U)(\beta)=\Lambda(\alpha, U)(\beta)
$$

By Lemmas 7.14, 7.15, $\alpha \beta \in L_{n, n}$ if and only if $\beta \in L_{n, n}$. Hence, $L_{n, n}$ is a subgroup.

The proof that $L_{n, n}$ is homogeneous is trivial and will be omitted. We will now show that $p \in L_{n, n}$. That is, for all $U \in S_{n}, U(p)=0$. This is readily verified for the case of $U_{1}, U_{2}, \cdots, U_{n}$ defined in the proof of Theorem 7.12. In fact, using (7.3), (7.4) and the fact that $r\left(\tau_{m}\right)=m+1$ for $m=0,1, \cdots$, we find that $p\left(\tau_{m}\right)=1 /(m+1)$, $p\left(\tau \tau_{m-1}\right)=1 / m(m+1)$ so that $U_{m}(p)=\left(m \hat{\tau}_{m-1}-\hat{\tau}_{m}\right)(p)=0$. Thus, we have proved that $U \in S_{n}^{i}$ implies $U(p)=0$ in the case $i=0$. We now proceed by induction on $i$. Suppose $m \hat{u}-\hat{v} \in S_{n}^{i-1}$ is such that $(m \hat{u}-\hat{v})(p)=0$ and we wish to prove the corresponding result in the case of $(m \hat{u}-\hat{v}) \hat{t}$ and $\hat{t}(m \hat{u}-\hat{v})$. Since $m p(u)=p(v)$, $r(u)=r(v)$ and $r(t u)=r(t v)=r(u t)=r(v t)$ we are in a position to evaluate $((m \hat{u}-\hat{v}) \hat{t})(p)$ and $(\hat{t}(m \hat{u}-\hat{v}))(p)$ making use of (7.4). In each case the result is zero.

In order to introduce our final subgroup, it is convenient to extend the notation given by (7.8) to elements of $G^{*}$. If $U \in G^{*}$ we will, for the remainder of this section, write $U_{0}=U, U_{1}=U_{0} \hat{\tau}, U_{2}=U_{1} \hat{r}, \cdots$.

Let $L^{n}$ be the subset of $G$ containing $\alpha$ if and only if for $m=1,2, \cdots, n-1$ and
for all $U \in G^{*}$ we have

$$
\begin{equation*}
\left(m \hat{\tau}_{m-1} U-\hat{\tau}(\alpha)^{m} U+U_{m}\right)(\alpha)=0 \tag{7.23}
\end{equation*}
$$

Note that $L^{n} \subseteq L_{n}$ since, if $U=\hat{\tau}$, (7.23) reduces to (7.15).
Theorem 7.17. $L^{n}$ is a homogeneous subgroup of $L_{n}$ and $p \in L^{n}$.
Proof. We will omit the trivial verification that $L^{n}$ is homogeneous. To prove that $p \in L^{n}$ it is sufficient to verify (7.23) with $\alpha=p$ (so that $\hat{\tau}(\alpha)=1$ ) and $U=\hat{\gamma}$ for all $t \in T$. It is found that

$$
U_{m}(\alpha)=r(t) p(t) /(m+r(t)), \quad\left(\hat{\tau}_{m-1} U\right)(\alpha)=p(t) /(m+r(t))
$$

so that (7.23) can easily be verified.
Since $U_{i}=U_{i-1} \hat{\imath}$, we can easily show, inductively, that

$$
\Lambda\left(\alpha, U_{m}\right)=\sum_{i=0}^{m}\binom{m}{i} \alpha(\tau)^{i} \Lambda(\alpha, U)_{m-i}
$$

so that

$$
\begin{aligned}
\Lambda\left(\alpha, \hat{\tau}_{m-1} U\right)= & \sum_{i=0}^{m-1}\binom{m-1}{i} \alpha(\tau)^{i} \hat{\tau}_{m-1-i} \Lambda(\alpha, U) \\
& +U(\alpha) \sum_{i=0}^{m-1}\binom{m-1}{i} \alpha(\tau)^{i} \hat{\tau}_{m-1-i}
\end{aligned}
$$

A straightforward calculation, making use of these results gives, for $m=1,2, \cdots$, $n-1$,

$$
\begin{aligned}
\left(m \hat{\tau}_{m-1}\right. & \left.U-\hat{\tau}(\alpha \beta)^{m} U+U_{m}\right)(\alpha \beta) \\
= & \left(m \hat{\tau}_{m-1} U-\hat{\tau}(\alpha)^{m} U+U_{m}\right)(\alpha) \\
& +\sum_{i=0}^{m}\binom{m}{i} \alpha(\tau)^{i}\left\{(m-i) \hat{\tau}_{m-i-1} \Lambda(\alpha, U)-\hat{\tau}(\beta)^{m-i} \Lambda(\alpha, U)+\Lambda(\alpha, U)_{m-i}\right\}(l) \\
& \quad+U(\alpha) \sum_{i=0}^{m-1}\binom{m}{i} \alpha(\tau)^{i}\left\{(m-i) \hat{\tau}_{m-1-i}(\beta)-\hat{\tau}(\beta)^{m-i}\right\} \\
= & \left(m \hat{\tau}_{m-1} U-\hat{\tau}(\alpha)^{m} U+U_{m}\right)(\alpha)
\end{aligned}
$$

if $\beta \in L^{n}$. Hence, $\alpha \beta \in L^{n}$ if and only if $\alpha \in L^{n}$. Thus, $L^{n}$ is a subgroup.
Definition 7.18. $L_{m, n}^{l}=L^{l} \cap L_{m, n}$.
Note that, as the intersection of homogeneous subgroups, $L_{m, n}^{l}$ is a homogeneous subgroup. Also $L_{m, n}^{l}$ contains $p$. We also see, since $L^{l} \subseteq L_{l}$, that we can assume without loss of generality that $m \geqq l$. In fact we can take $m$ as $l+n-1$ as the next theorem shows.

Theorem 7.19. If $l \geqq 2$, then $L_{m, n}^{l} \subseteq L_{l+n-1}$.
Proof. Replace $m$ by $i$ in (7.23) and write $U=\hat{\tau}_{i-1}$ where $j \leqq n$. We then have, for $i<l$,

$$
\begin{equation*}
\left(i \hat{\tau}_{i-1} \hat{\tau}_{i-1}-\hat{\tau}(\alpha)^{i} \hat{\tau}_{i-1}+\hat{\tau}_{i+i-1}\right)(\alpha)=0 \tag{7.24}
\end{equation*}
$$

for all $\alpha \in L_{m, n}^{l}$.
Since $\tau_{i-1} \tau_{i-1}=\left(\tau \tau_{i-1}\right)_{i-1}$ and since $\alpha\left(\left(\tau \tau_{i-1}\right)_{i-1}\right)=\alpha\left(\tau_{i+i-1}\right) / j$ because
$j\left(\hat{\tau} \hat{\tau}_{i-1}\right)_{i-1}-\left(\hat{\tau}_{i}\right)_{i-1} \in S_{n}$, we deduce from (7.24)

$$
\begin{equation*}
(i+j) \hat{\tau}_{i+i-1}(\alpha)=j \hat{\tau}(\alpha)^{i} \hat{\tau}_{i-1}(\alpha) \tag{7.25}
\end{equation*}
$$

Since (7.25) holds for $i<l$ and $j \leqq n$, it follows that $k \hat{\boldsymbol{\gamma}}_{k-1}(\alpha)=\hat{\tau}(\alpha)^{k}$ for all $k<$ $l+n$.
8. Applications and Extensions. In this paper, the concept of order has been introduced in a purely algebraic way although its true significance is analytical. It is the author's intention to present in a later paper an error analysis for Runge-Kutta methods that will fit very much into the spirit of the present paper and in which the concept of order will play an important part.

In another proposed contribution, a characterisation will be given for RungeKutta methods which, for appropriately smooth functions, have a global truncation error with an asymptotic expansion in even powers of the step size. Such methods will yield generalisations of the Romberg quadrature method. In particular, the implicit methods based on Gauss-Legendre quadrature have this property.

Again, the author intends to introduce a type of explicit method which, though of low order, may be made use of in a way which makes it effectively of higher order. In particular, a five stage method will be presented which can be used in a special way to yield answers that have the properties of answers produced by a fifth order method.

Finally, it is the author's hope that the results in Section 5 can be generalised by relaxing the assumption that $X$ is finite dimensional.
9. Appendix. In Table 9.1 are listed trees of order up to 5 . Also for given $t \in T$, the corresponding values of $r(t)$ and $\lambda(\alpha, t)$ for $\alpha \in G$ are given.

In Table 9.2, the expressions for $(\alpha \beta)(t)$ are given for $t \in T$ of order up to 5. To simplify the notation, the various $t$ are numbered serially $t_{0}=\tau, t_{1}=\tau \tau, t_{2}$ $=\tau \tau \cdot \tau, \cdots$ and $(\alpha \beta)_{i}, \alpha_{i}, \beta_{i}$ denote $(\alpha \beta)\left(t_{i}\right), \alpha\left(t_{i}\right), \beta\left(t_{i}\right)$ respectively for $i=0,1, \cdots$, 16. For convenience, $r_{i}=r\left(t_{i}\right)$ is also listed.

In Table 9.3 the values of $p\left(t_{i}\right)^{-1}$ are given where $t_{0}, t_{1}, \cdots, t_{16}$ are the trees of order up to 5 numbered in the same way as in Table 9.2.

Finally, we present Table 9.4 to enable certain computations to be conveniently performed in $L_{3,2}^{2}$. Let $t_{0}, t_{1}, \cdots t_{18}$ be as shown in Table 9.2 and let $\alpha$ be a member of $L_{3,2}^{2}$. Because of the relations which define this group, it can be seen that $\alpha_{i}$ for $i \neq 0,4,8$ or 10 can be written in terms of $\alpha_{0}, \alpha_{4}, \alpha_{8}, \alpha_{10}$. Where there are entries under the heading $\alpha_{i}$, expressions for these quantities are given. Also in this table are expressions for $(\alpha \beta)_{i}$ where $\beta$ is also a member of $L_{3,2}^{2}$ and $i=0,4,8,10$.

Mathematics Department
The University of Auckland
Auckland, New Zealand

[^0]Table 9.1

| $t$ | $r(t)$ | $\lambda(\alpha, t)$ |
| :---: | :---: | :---: |
|  | 1 | * |
| $\tau \tau$ | 2 | ${ }_{\boldsymbol{\tau}}^{\boldsymbol{\tau}} \boldsymbol{\tau}+\alpha(\tau) \hat{\tau}$ |
| $\boldsymbol{\tau} \cdot \boldsymbol{\tau}$ | 3 | $\hat{\boldsymbol{\tau}} \cdot \underline{\hat{\tau}}+2 \alpha(\tau) \hat{\tau} \hat{\tau}+\alpha(\tau){ }^{2} \hat{\tau}$ |
| $\boldsymbol{\tau} \cdot \boldsymbol{\tau} \boldsymbol{\tau}$ | 3 | $\hat{\boldsymbol{f}} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau) \hat{\tau} \hat{\tau}+\alpha(\tau \tau) \hat{\tau}$ |
| $(\tau \tau \cdot \tau) \tau$ | 4 | $(\hat{\tau} \hat{\tau} \cdot \hat{\tau}) \hat{\tau}+3 \alpha(\tau) \hat{\tau} \hat{\tau} \cdot \hat{\tau}+3 \alpha(\tau)^{2} \hat{\tau} \hat{\tau}+\alpha(\tau)^{3} \hat{\tau}$ |
| $\boldsymbol{\tau} \boldsymbol{\tau} \cdot \boldsymbol{\tau} \boldsymbol{\tau}$ | 4 | $\hat{\boldsymbol{\tau}} \hat{\boldsymbol{\tau}} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau) \hat{\tau} \hat{\tau} \cdot \hat{\tau}+\alpha(\tau) \hat{\tau} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau)^{2} \hat{\tau} \hat{\tau}+\alpha(\tau \tau) \dot{\tau} \hat{\tau}+\alpha(\tau \tau) \alpha(\tau) \hat{\tau}$ |
| $\tau(\tau \tau \cdot \tau)$ | 4 | $\hat{\boldsymbol{\tau}}(\boldsymbol{\tau} \boldsymbol{\tau} \cdot \hat{\tau})+2 \alpha(\tau) \hat{\tau} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau)^{2} \hat{\tau} \hat{\tau}+\alpha(\tau \tau \cdot \tau) \hat{\tau}$ |
| $\tau(\tau \cdot \tau \tau)$ | 4 | $\hat{\tau}(\hat{\tau} \cdot \hat{\tau} \hat{\tau})+\alpha(\tau) \hat{\tau} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau \tau) \hat{\tau} \hat{\tau}+\alpha(\tau \cdot \tau \tau) \hat{\tau}$ |
| $(\tau \tau \cdot \tau) \tau \cdot \tau$ | 5 | $\left.(\hat{\tau} \hat{\tau} \cdot \hat{\tau}) \hat{\tau} \cdot \hat{\tau}+4 \alpha(\tau)(\hat{\tau} \hat{\tau} \cdot \hat{\tau}) \hat{\tau}+6 \alpha(\tau){ }^{2} \hat{\tau} \hat{\tau} \cdot \hat{\tau}+4 \alpha(\tau)^{3} \hat{\tau}+\alpha{ }^{(\tau)}\right)^{4} \hat{\tau}$ |
| $(\tau \tau \cdot \tau) \cdot \tau \tau$ | 5 |  |
| $\boldsymbol{\tau \tau} \cdot(\tau \tau \cdot \tau)$ | 5 |  |
| $\tau \tau \cdot(\tau \cdot \tau \tau)$ | 5 | $\hat{\boldsymbol{\tau}} \hat{\boldsymbol{\tau}} \cdot(\hat{\tau} \cdot \hat{\tau} \hat{\tau})+\alpha(\tau) \hat{\tau}(\hat{\tau} \cdot \hat{\tau})+\alpha(\tau) \hat{\tau} \cdot \hat{\tau} \hat{\tau} \hat{\tau}+\alpha(\tau)^{2} \hat{\tau} \cdot \boldsymbol{\tau} \hat{\tau}+\alpha(\tau \tau) \hat{\tau} \cdot \hat{\tau}+\alpha(\tau) \alpha(\tau \tau) \hat{\tau} \hat{\gamma}+\alpha(\tau \cdot \tau \tau) \hat{\tau}+\alpha+\alpha(\tau) \alpha(\tau \cdot \tau \tau) \hat{\tau}$ |
| $(\tau \cdot \tau \tau) \cdot \tau \tau$ | 5 | $\left(\hat{\tau} \cdot \hat{\tau} \hat{\tau} \cdot \hat{\tau} \hat{\tau}+2 \alpha(\tau) \hat{\tau} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau)^{2} \hat{\tau} \hat{\tau} \cdot \hat{\tau}+2 \alpha(\tau \tau) \dot{\tau} \cdot \hat{\tau} \hat{\tau}+2 \alpha(\tau) \alpha(\tau \tau)+\boldsymbol{\tau}+\alpha(\tau \tau)^{2} \hat{\tau}\right.$ |
| $\tau \cdot(\tau \tau \cdot \tau) \tau$ | 5 |  |
| $\tau(\tau \tau \cdot \tau \tau)$ | 5 |  |
| $\tau \cdot \tau(\tau \tau \cdot \tau)$ | 5 | $\hat{f} \cdot \hat{f}(\hat{\tau} \hat{f} \cdot \hat{\tau})+2 \alpha(\tau) \hat{\tau}(\hat{\tau} \cdot \hat{\tau} \hat{\tau})+\alpha(\tau)^{2} \hat{\tau} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau \tau \cdot \tau) \hat{\tau} \hat{\tau}+\alpha(\tau(\tau \tau \cdot \tau)) \hat{t}$ |
| $\tau \cdot \tau(\tau \cdot \tau \tau)$ | 5 | $\hat{\tau} \cdot \hat{\tau}(\hat{\tau} \cdot \hat{\tau})+\alpha(\tau) \hat{\tau}(\hat{\tau} \cdot \hat{\tau} \hat{\tau})+\alpha(\tau \tau) \hat{\tau} \cdot \hat{\tau} \hat{\tau}+\alpha(\tau \cdot \tau \tau) \hat{\tau} \hat{\tau}+\alpha(\tau(\tau \tau \cdot \tau)) \hat{t}$ |

Table 9.2

| $i$ | $t_{i}$ | $r_{i}$ | $(\alpha \beta)_{i}$ |
| :--- | :--- | :--- | :--- |
| 0 | $\tau$ | 1 | $\alpha_{0}+\beta_{0}$ |
| 1 | $\tau \tau$ | 2 | $\alpha_{1}+\beta_{1}+\alpha_{0} \beta_{0}$ |
| 2 | $\tau \tau \cdot \tau$ | 3 | $\alpha_{2}+\beta_{2}+2 \alpha_{0} \beta_{1}+\alpha_{0}^{2} \beta_{0}$ |
| 3 | $\tau \cdot \tau \tau$ | 3 | $\alpha_{3}+\beta_{3}+\alpha_{0} \beta_{1}+\alpha_{1} \beta_{0}$ |
| 4 | $(\tau \tau \cdot \tau) \tau$ | 4 | $\alpha_{4}+\beta_{4}+3 \alpha_{0} \beta_{2}+3 \alpha_{0}^{2} \beta_{1}+\alpha_{0}^{3} \beta_{0}$ |
| 5 | $\tau \tau \cdot \tau \tau$ | 4 | $\alpha_{5}+\beta_{5}+\alpha_{0} \beta_{2}+\alpha_{0} \beta_{3}+\alpha_{0}^{2} \beta_{1}+\alpha_{1} \beta_{1}+\alpha_{0} \alpha_{1} \beta_{0}$ |
| 6 | $\tau(\tau \tau \cdot \tau)$ | 4 | $\alpha_{6}+\beta_{6}+2 \alpha_{0} \beta_{3}+\alpha_{0}^{2} \beta_{1}+\alpha_{2} \beta_{0}$ |
| 7 | $\tau(\tau \cdot \tau \tau)$ | 4 | $\alpha_{7}+\beta_{7}+\alpha_{0} \beta_{3}+\alpha_{1} \beta_{1}+\alpha_{3} \beta_{0}$ |
| 8 | $(\tau \tau \cdot \tau) \tau \cdot \tau$ | 5 | $\alpha_{8}+\beta_{8}+4 \alpha_{0} \beta_{4}+6 \alpha_{0}^{2} \beta_{2}+4 \alpha_{0}^{3} \beta_{1}+\alpha_{0}^{4} \beta_{0}$ |
| 9 | $(\tau \tau \cdot \tau) \cdot \tau \tau$ | 5 | $\alpha_{9}+\beta_{9}+\alpha_{0} \beta_{4}+2 \alpha_{0} \beta_{5}+2 \alpha_{0}^{2} \beta_{2}+\alpha_{0}^{2} \beta_{3}+\alpha_{0}^{3} \beta_{1}+\alpha_{1} \beta_{2}+2 \alpha_{0} \alpha_{1} \beta_{1}$ |
|  |  |  |  |
| 10 | $\tau \tau \cdot(\tau \tau \cdot \tau)$ | 5 | $\alpha_{10}+\beta_{10}+\alpha_{0} \beta_{6}+2 \alpha_{0} \beta_{5}+\alpha_{0}^{2} \beta_{2}+2 \alpha_{0}^{2} \beta_{3}+\alpha_{0}^{3} \beta_{1}+\alpha_{2} \beta_{1}+\alpha_{0} \alpha_{2} \beta_{0}$ |
| 11 | $\tau \tau \cdot(\tau \cdot \tau \tau)$ | 5 | $\alpha_{11}+\beta_{11}+\alpha_{0} \beta_{7}+\alpha_{0} \beta_{5}+\alpha_{0}^{2} \beta_{3}+\alpha_{1} \beta_{2}+\alpha_{0} \alpha_{1} \beta_{1}+\alpha_{3} \beta_{1}+\alpha_{0} \alpha_{2} \beta_{0}$ |
| 12 | $(\tau \cdot \tau \tau) \cdot \tau \tau$ | 5 | $\alpha_{12}+\beta_{12}+2 \alpha_{0} \beta_{5}+\alpha_{0}^{2} \beta_{2}+2 \alpha_{1} \beta_{3}+2 \alpha_{0} \alpha_{1} \beta_{1}+\alpha_{1}^{2} \beta_{0}$ |
| 13 | $\tau \cdot(\tau \tau \cdot \tau) \tau$ | 5 | $\alpha_{13}+\beta_{13}+3 \alpha_{0} \beta_{6}+3 \alpha_{0}^{2} \beta_{3}+\alpha_{0}^{3} \beta_{1}+\alpha_{4} \beta_{0}$ |
| 14 | $\tau(\tau \tau \cdot \tau \tau)$ | 5 | $\alpha_{14}+\beta_{14}+\alpha_{0} \beta_{6}+\alpha_{0} \beta_{7}+\alpha_{0}^{2} \beta_{3}+\alpha_{1} \beta_{3}+\alpha_{0} \alpha_{1} \beta_{1}+\alpha_{5} \beta_{0}$ |
| 15 | $\tau \cdot \tau(\tau \tau \cdot \tau)$ | 5 | $\alpha_{15}+\beta_{15}+2 \alpha_{0} \beta_{7}+\alpha_{0}^{2} \beta_{3}+\alpha_{2} \beta_{1}+\alpha_{6} \beta_{0}$ |
| 16 | $\tau \cdot \tau(\tau \cdot \tau \tau)$ | 5 | $\alpha_{16}+\beta_{16}+\alpha_{0} \beta_{7}+\alpha_{1} \beta_{3}+\alpha_{3} \beta_{1}+\alpha_{7} \beta_{0}$ |

Table 9.3

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| $p\left(t_{i}\right)^{-1}$ | 1 | 2 | 3 | 6 | 4 | 8 | 12 | 24 | 5 | 10 | 15 | 30 | 20 | 20 | 40 | 60 | 120 |

Table 9.4

| $i$ |  | $\alpha_{i}$ |
| :--- | :--- | :--- |
| 0 |  | $(\alpha \beta)_{i}$ |
| 1 | $\frac{1}{2} \alpha_{0}^{2}$ | $\alpha_{0}+\beta_{0}$ |
| 2 | $\frac{1}{3} \alpha_{0}^{3}$ |  |
| 3 | $\frac{1}{6} \alpha_{0}^{3}$ |  |
| 4 |  | $\alpha_{4}+\beta_{4}+\alpha_{0} \beta_{0}^{3}+\frac{3}{2} \alpha_{0}^{2} \beta_{0}^{2}+\alpha_{0}^{8} \beta_{0}$ |
| 5 | $\frac{1}{2} \alpha_{4}$ |  |
| 6 | $\frac{1}{3} \alpha_{0}^{4}-\alpha_{4}$ |  |
| 7 | $\frac{1}{6} \alpha_{0}^{4}-\frac{1}{2} \alpha_{4}$ | $\alpha_{8}+\beta_{8}+4 \alpha_{0} \beta_{4}+2 \alpha_{0}^{2} \beta_{0}^{3}+2 \alpha_{0}^{3} \beta_{0}^{2}+\alpha_{0}^{4} \beta_{0}$ |
| 8 |  |  |
| 9 | $\frac{1}{2} \alpha_{8}$ | $\alpha_{10}+\beta_{10}+\frac{1}{3} \alpha_{0} \beta_{0}^{4}+\frac{2}{3} \alpha_{0}^{2} \beta_{0}^{3}+\frac{2}{3} \alpha_{0}^{3} \beta_{0}^{2}+\frac{1}{3} \alpha_{0}^{4} \beta_{0}$ |
| 10 |  |  |
| 11 | $\frac{1}{2} \alpha_{10}$ |  |
| 12 | $\frac{1}{4} \alpha_{8}$ |  |
| 13 | $\alpha_{0} \alpha_{4}-\alpha_{8}$ |  |
| 14 | $\frac{1}{2} \alpha_{0} \alpha_{4}-\frac{1}{2} \alpha_{8}$ |  |
| 15 | $\frac{1}{3} \alpha_{0}^{5}-\alpha_{0} \alpha_{4}-\alpha_{10}$ |  |
| 16 | $\frac{1}{6} \alpha_{0}^{5}-\frac{1}{2} \alpha_{0} \alpha_{4}-\frac{1}{2} \alpha_{10}$ |  |


[^0]:    1. J. C. Butcher, "Coefficients for the study of Runge-Kutta integration processes," J. Austral. Math. Soc., v. 3, 1963, pp. 185-201. MR 27 \#2109.
    2. J. C. Butcher, "On the attainable order of Runge-Kutta methods," Math. Comp., v. 19, 1965, pp. 408-417. MR 31 \#4180.
