

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 σ and a binary relation ρ on σ . The pair (σ, ρ) will be called a labelled graph and the elements of σ will be known as labels. (σ, ρ) and (σ', ρ') will be said to be isomorphic if there is a bijection $\varphi: \sigma \rightarrow \sigma'$ such that ρ' contains $(\varphi(x), \varphi(y))$ if and only if ρ 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

σ	ρ		
{0}	{ }	0.	τ
{0, 1}	{(0, 1)}		$\tau\tau$
{0, 1, 2}	{(0, 1), (1, 2)}		$\tau \cdot \tau\tau$
{0, 1, 2}	{(0, 1), (0, 2)}		$\tau\tau \cdot \tau$
{0, 1, 2, 3}	{(0, 1), (1, 2), (2, 3)}		$\tau(\tau \cdot \tau\tau)$
{0, 1, 2, 3}	{(0, 1), (0, 2), (2, 3)}		$\tau\tau \cdot \tau\tau = (\tau \cdot \tau\tau)\tau$
{0, 1, 2, 3}	{(0, 1), (1, 2), (1, 3)}		$\tau(\tau\tau \cdot \tau)$
{0, 1, 2, 3}	{(0, 1), (0, 2), (0, 3)}		$(\tau\tau \cdot \tau)\tau$

If $\gamma_1 = (\sigma_1, \rho_1)$ and $\gamma_2 = (\sigma_2, \rho_2)$ are graphs such that $\sigma_1 \cap \sigma_2 = \emptyset$ 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 $\{\gamma_1, \gamma_2\}$ a disconnection of γ . If γ has no disconnection, then we say that γ 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 σ (r will be called the root of γ) such that for each $y \in \sigma \setminus \{r\}$ 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 σ . 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 σ are represented by points and the elements of ρ 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 (σ_1, ρ_1) and (σ_2, ρ_2) be labellings of two trees γ_1 and γ_2 such that σ_1 and σ_2 are disjoint and let r_1 and r_2 be the roots of (σ_1, ρ_1) and (σ_2, ρ_2) . We define the product $\gamma_1 \cdot \gamma_2$ as the tree which may be labelled as

$$(\sigma_1 \cup \sigma_2, \rho_1 \cup \rho_2 \cup \{(r_1, r_2)\}).$$

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 σ for a labelling (σ, ρ) , 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

$$(2.1) \quad (\gamma_1 \cdot \gamma_2) \cdot \gamma_3 = (\gamma_1 \cdot \gamma_3) \cdot \gamma_2.$$

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 (σ, ρ) to be a labelling of t and we shall prove that there are (σ_1, ρ_1) and (σ_2, ρ_2) such that $\sigma_1 \cap \sigma_2 = \emptyset$, $\sigma_1 \cup \sigma_2 = \sigma$, $\rho = \rho_1 \cup \rho_2 \cup \{(r_1, r_2)\}$, where r_1, r_2 are the roots of (σ_1, ρ_1) and (σ_2, ρ_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 $(\sigma_1, \rho_1) = (\{0\}, \{\})$ and $(\sigma_2, \rho_2) = (\{1\}, \{\})$. 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 ρ^{-1} is a function on $\sigma \setminus \{r\}$ to σ , where r is the root of (σ, ρ) , and since σ has 1 more element than $\sigma \setminus \{r\}$, ρ^{-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 σ such that $(z, x) \in \rho$. We define $(\sigma', \rho') = (\sigma \setminus \{x\}, \rho \setminus \{(z, x)\})$. It is easy to see that (σ', ρ') is a tree with root r and order $n - 1$. Let $(\sigma'_1, \rho'_1), (\sigma'_2, \rho'_2)$ be formed corresponding to a factorisation of (σ', ρ') . We distinguish two cases:

Case I. $z \in \sigma'_1$. In this case we define $(\sigma_1, \rho_1) = (\sigma'_1 \cup \{x\}, \rho'_1 \cup \{(z, x)\}), (\sigma_2, \rho_2) = (\sigma'_2, \rho'_2)$.

Case II. $z \in \sigma'_2$. We define $(\sigma_1, \rho_1) = (\sigma'_1, \rho'_1), (\sigma_2, \rho_2) = (\sigma'_2 \cup \{x\}, \rho'_2 \cup \{(z, x)\})$.

In each case, we see that (σ_1, ρ_1) and (σ_2, ρ_2) are trees satisfying the required conditions.

Let τ denote the unique tree with order 1. Then, using Lemma 2.1, we see that all trees are generated using τ 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 τ 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 τ 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 $t = u \cdot v$. 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 \cdot 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, $(t_1 t_2 \cdot t_3 t_4) t_5 \cdot t_6 t_7$ will denote the same tree as $((t_1 \cdot t_2) \cdot (t_3 \cdot t_4)) \cdot t_5 \cdot (t_6 \cdot t_7)$. 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 τ . 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 $n = 1, 2, 3, \dots$, of the sets of n -tuples of positive integers. That is,

$$S = \{(1), (2), \dots, (1, 1), (1, 2), \dots, (2, 1), \dots, (1, 1, 1), (1, 1, 2), \dots\}.$$

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 $x^- = () \in S$ 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 σ the labelled graph (σ, ρ) where

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

We define a labelled forest as a labelled graph (σ, ρ) such that (σ, ρ) can be disconnected into $(\sigma_1, \rho_1), \dots, (\sigma_m, \rho_m)$ (with m possibly equal to 1) where each of $(\sigma_1, \rho_1), \dots, (\sigma_m, \rho_m)$ is a labelled tree or such that $(\sigma, \rho) = (\emptyset, \emptyset)$. 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 \leq on F by writing $\sigma_1 \leq \sigma_2$ 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

$$(2.2) \quad (\alpha\beta)(\sigma) = \sum_{\sigma' \leq \sigma} \alpha(\sigma \setminus \sigma')\beta(\sigma')$$

for all $\sigma \in F$ and the summation is over all members σ' of F such that $\sigma' \leq \sigma$.

If we write $\sigma_1 + \sigma_2 + \dots = \sigma$ to indicate that $\sigma_1, \sigma_2, \dots$ are disjoint sets whose union is σ then we may rewrite (2.2) as

$$(\alpha\beta)(\sigma) = \sum_{\sigma_1 + \sigma_2 = \sigma : \sigma_2 \leq \sigma} \alpha(\sigma_1)\beta(\sigma_2).$$

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 $\sigma \in F$,

$$(2.3) \quad ((\alpha\beta)\gamma)(\sigma) = \sum_{\sigma_1 + \sigma_2 + \sigma_3 = \sigma; \sigma_1 \leq \sigma_2 \leq \sigma_1 \cup \sigma_2} \alpha(\sigma_1)\beta(\sigma_2)\gamma(\sigma_3),$$

$$(2.4) \quad (\alpha(\beta\gamma))(\sigma) = \sum_{\sigma_1 + \sigma_2 + \sigma_3 = \sigma; \sigma_1 \cup \sigma_2 \leq \sigma_3 \leq \sigma_2 \cup \sigma_3} \alpha(\sigma_1)\beta(\sigma_2)\gamma(\sigma_3).$$

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

$$(2.5) \quad \sigma_3 \leq \sigma_1 \cup \sigma_2 \cup \sigma_3$$

and

$$(2.6) \quad \sigma_2 \leq \sigma_1 \cup \sigma_2$$

if and only if

$$(2.7) \quad \sigma_2 \cup \sigma_3 \leq \sigma_1 \cup \sigma_2 \cup \sigma_3$$

and

$$(2.8) \quad \sigma_3 \leq \sigma_2 \cup \sigma_3.$$

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

$$(2.9) \quad x \in \sigma_3 \Rightarrow x^- \notin \sigma_1 \cup \sigma_2,$$

$$(2.10) \quad x \in \sigma_2 \Rightarrow x^- \notin \sigma_1,$$

$$(2.11) \quad x \in \sigma_2 \cup \sigma_3 \Rightarrow x^- \notin \sigma_1,$$

$$(2.12) \quad x \in \sigma_3 \Rightarrow x^- \notin \sigma_2.$$

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 $e(\emptyset) = 1, e(\sigma) = 0$ if $\sigma \in F \setminus \{\emptyset\}$. The subset $M_0 = \{\alpha \in M, \alpha(\emptyset) \neq 0\}$ with the operation is a group. To show this, we see that if $\alpha, \beta \in M_0$ then $(\alpha\beta)(\emptyset) = \alpha(\emptyset)\beta(\emptyset) \neq 0$ 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 σ . We have

$$(2.13) \quad \alpha^{-1}(\emptyset) = \alpha(\emptyset)^{-1},$$

$$(2.14) \quad \alpha^{-1}(\sigma) = -\left(\sum_{\sigma' \leq \sigma; \sigma' \neq \sigma} \alpha(\sigma \setminus \sigma')\alpha^{-1}(\sigma') \right) / \alpha(\emptyset),$$

for $\sigma \in F \setminus \{\emptyset\}$.

We consider further the subset M_1 of M_0 defined so that $\alpha \in M_1$ if and only if, whenever σ_1 and σ_2 correspond to the same forest, then $\alpha(\sigma_1) = \alpha(\sigma_2)$. 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 $\alpha(\emptyset) = 1$ and if, whenever

(σ_1, σ_2) is a disconnection of σ (that is $\sigma_1 + \sigma_2 = \sigma$, $\sigma_1 \leq \sigma$, $\sigma_2 \leq \sigma$), then $\alpha(\sigma) = \alpha(\sigma_1)\alpha(\sigma_2)$. 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 α^{-1} are multiplicative.*

Proof. Suppose (σ_1, σ_2) is a disconnection of σ then there is a bijection between the set of pairs (σ'_1, σ'_2) such that $\sigma'_1 \leq \sigma_1$, $\sigma'_2 \leq \sigma_2$ and the set $\{\sigma' \in F: \sigma' \leq \sigma\}$. In fact, the bijection is such that $\sigma' = \sigma'_1 \cup \sigma'_2$ corresponds to (σ'_1, σ'_2) . In the inverse relation, $(\sigma'_1, \sigma'_2) = (\sigma_1 \cap \sigma', \sigma_2 \cap \sigma')$ is a disconnection of the corresponding σ' and $(\sigma_1 \setminus \sigma'_1, \sigma_2 \setminus \sigma'_2)$ is a disconnection of $\sigma \setminus \sigma'$.

We now have

$$\begin{aligned} (\alpha\beta)(\sigma) &= \sum_{\sigma' \leq \sigma} \alpha(\sigma \setminus \sigma')\beta(\sigma') \\ &= \sum_{\sigma'_1 \leq \sigma_1} \sum_{\sigma'_2 \leq \sigma_2} \alpha(\sigma_1 \setminus \sigma'_1)\alpha(\sigma_2 \setminus \sigma'_2)\beta(\sigma'_1)\beta(\sigma'_2) \\ &= (\alpha\beta)(\sigma_1) \cdot (\alpha\beta)(\sigma_2) \end{aligned}$$

so that $\alpha\beta$ is multiplicative.

To prove α^{-1} is multiplicative, we proceed by induction. Thus, we suppose $\alpha^{-1}(\sigma') = \alpha^{-1}(\sigma'_1)\alpha^{-1}(\sigma'_2)$ whenever (σ'_1, σ'_2) is a disconnection of the forest σ' with fewer elements than σ . We now have

$$\begin{aligned} \alpha^{-1}(\sigma) &= - \sum_{\sigma' \leq \sigma, \sigma' \neq \sigma} \alpha(\sigma \setminus \sigma')\alpha^{-1}(\sigma') \\ &= \alpha^{-1}(\sigma_1)\alpha^{-1}(\sigma_2) - \sum_{\sigma'_1 \leq \sigma_1} \sum_{\sigma'_2 \leq \sigma_2} \alpha(\sigma_1 \setminus \sigma'_1)\alpha(\sigma_2 \setminus \sigma'_2)\alpha^{-1}(\sigma'_1)\alpha^{-1}(\sigma'_2) \\ &= \alpha^{-1}(\sigma_1)\alpha^{-1}(\sigma_2). \end{aligned}$$

We now observe that if $\sigma \in F$ is a labelled tree and $\sigma' \leq \sigma$, then either σ' is a labelled tree or $\sigma' \neq \emptyset$. Hence, in this case, (2.2) takes the form

$$\begin{aligned} (\alpha\beta)(\sigma) &= \sum_{\sigma' \leq \sigma} \alpha(\sigma \setminus \sigma')\beta(\sigma') \\ (2.15) \quad &= \alpha(\sigma) + \sum_{\sigma' \leq \sigma, \sigma' \neq \emptyset} \alpha(\sigma \setminus \sigma')\beta(\sigma') \end{aligned}$$

where the summation involves values of β evaluated only at trees.

Let a, b be functions on T to R corresponding to $\alpha, \beta \in M_2$. That is, if φ is a function on the nonempty connected members of F to T defined so that for each σ , $\varphi(\sigma)$ is the tree corresponding to σ , 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

$$(2.16) \quad (ab)(t) = a(t) + \lambda(a, t)(b)$$

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 λ 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 t^\wedge) 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 $\{t \in T: c_p(t) \neq 0\}$ is finite and so that for all $a \in G$,

$$p(a) = \sum_{t \in T} c_p(t)a(t).$$

For $\hat{i}, \hat{u} \in \hat{T}$ we define the product $\hat{i} \cdot \hat{u}$ by $\hat{i} \cdot \hat{u} = (tu)^\wedge$ 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 pq as

$$pq = \sum_{t, u \in T} c_p(t)c_q(u)\hat{t}\hat{u}.$$

We note that λ is a mapping on $G \times T$ to G^* . The following result enables us to determine λ 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

$$(2.17) \quad \lambda(a, \tau) = \hat{\tau},$$

$$(2.18) \quad \lambda(a, uv) = \lambda(a, u)\lambda(a, v) + a(v)\lambda(a, u).$$

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 $\sigma' \leq \sigma$ then $\sigma' = \emptyset$ or $\sigma' = \sigma$.

The proof of (2.18) is more involved. Let $\sigma \in F$ correspond to $t = uv$ and let σ_1, σ_2 correspond to u and v , respectively. Without loss of generality, we may suppose that the root of σ is (1) and the root of σ_2 is $(1, 1)$. The only connection between σ_1 and σ_2 is $((1), (1, 1))$ so any subset of σ_1 not containing (1) is not connected to any subset of σ_2 and any subset of σ_2 not containing $(1, 1)$ is not connected to any subset of σ_1 . Furthermore, if $\sigma'_1 \leq \sigma_1, \sigma'_2 \leq \sigma_2$ and $\sigma'_1 = \emptyset$ only if $\sigma'_2 = \emptyset$ then $\sigma'_1 \cup \sigma'_2 \leq \sigma$. Also we observe that the relationship between such pairs (σ'_1, σ'_2) as have these properties and $\sigma' = \sigma'_1 \cup \sigma'_2$ is a bijection. Hence, we have

$$\begin{aligned} (\alpha\beta)(\sigma) &= \sum_{\sigma' \leq \sigma} \alpha(\sigma \setminus \sigma')\beta(\sigma') \\ &= \alpha(\sigma) + \sum_{\emptyset \neq \sigma'_1 \leq \sigma_1} \alpha(\sigma_1 \setminus \sigma'_1)\alpha(\sigma_2)\beta(\sigma'_1) \\ &\quad + \sum_{\emptyset \neq \sigma'_1 \leq \sigma_1; \emptyset \neq \sigma'_2 \leq \sigma_2} \alpha(\sigma_1 \setminus \sigma'_1)\alpha(\sigma_2 \setminus \sigma'_2)\beta(\sigma'_1 \cup \sigma'_2) \end{aligned}$$

and this is equivalent to (2.18).

A final property of λ we will consider is

LEMMA 2.7. If $a, b, c \in G$ and $t \in T$ then

$$(2.19) \quad \lambda(ab, t)(c) + \lambda(a, t)(b) = \lambda(a, t)(bc).$$

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

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

$$(a(bc))(t) = \lambda(a, t)(bc) + a(t),$$

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 $\eta_1 = y(h_1)$ 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

$$(3.1) \quad y = \eta_0 e + a \times (f \circ y).$$

Here, e denotes the unit in $B(H)$ so that $\eta_0 e$ will be defined by $(\eta_0 e)(h) = \eta_0$ 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 η_1 is just that of a Runge-Kutta method. On the other hand, if $H = [0, 1]$ and a is defined by

$$(3.2) \quad a(x)(h) = \int_0^h x(t) dt$$

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

$$(3.3) \quad \frac{dy}{dt}(t) = f(y(t)), \quad y(0) = \eta_0.$$

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 $e, a(e), a(a(e)), a(e)^2, a(e) \cdot a(a(e)), \dots$. 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 μ_a now to be defined.

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

$$(3.4) \quad \mu_a(\tau) = e,$$

$$(3.5) \quad \mu_a(t_1 t_2) = \mu_a(t_1) a(\mu_a(t_2)),$$

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

$$\mu_a(t_1 t_2 \cdot t_3) = \mu_a(t_1) \cdot a(\mu_a(t_2)) \cdot a(\mu_a(t_3)) = \mu_a(t_1 t_3 \cdot t_2).$$

It is also convenient to introduce a mapping $\nu_a = \mu_a \circ \varphi$ where φ is the mapping on T to T defined by $\varphi(t) = \tau t$. Thus,

$$(3.6) \quad \nu_a(t) = a(\mu_a(t))$$

and the elements in the range of ν_a can be regarded as generators of the range of μ_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 μ_a and ν_a in each case. The range of φ will be

TABLE 3.1

t	$\mu_a(t)$	$\nu_a(t)$
τ	e	$a(e)$
$\tau\tau$	$a(e)$	$a(a(e))$
$\tau \cdot \tau\tau$	$a(a(e))$	$a(a(a(e)))$
$\tau\tau \cdot \tau$	$a(e)^2$	$a(a(e)^2)$
$\tau(\tau \cdot \tau\tau)$	$a(a(a(e)))$	$a(a(a(a(e))))$
$\tau\tau \cdot \tau\tau$	$a(e)a(a(e))$	$a(a(e)a(a(e)))$
$\tau(\tau\tau \cdot \tau)$	$a(a(e)^2)$	$a(a(a(e)^2))$
$(\tau\tau \cdot \tau)\tau$	$a(e)^3$	$a(a(e)^3)$

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 $t \in T$, $\mu_a(t)(h_1) = \mu_a(t)(h_2)$. It is clear that in this definition we could replace μ by ν 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 π_a mapping H onto P_a such that $\pi_a(h_1) = \pi_a(h_2)$ 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 μ_a, ν_a, π_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 $\{\nu(t) \circ \pi^{-1} : t \in T\}$. 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

$$(4.1) \quad q(t) \circ r \circ \pi = \nu(t)$$

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 $\pi(h_1) \neq \pi(h_2)$ then there is a $t \in T$ such that $\nu(t)(h_1) \neq \nu(t)(h_2)$. Hence, we find

$$(q(t) \circ r)(\pi(h_1)) \neq (q(t) \circ r)(\pi(h_2))$$

so that $r(\pi(h_1)) \neq r(\pi(h_2))$.

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-Ćech 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^\wedge . In particular, for $t \in T$, $(\mu(t) \circ \pi^{-1})^\wedge$ 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

$$(4.2) \quad \sum_{t \in T} c(t)(\mu(t) \circ \pi^{-1})^\wedge,$$

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 $\mu(t) = \mu(t_1)\mu(t_2)$ and this is clear because $\mu(t_1), \mu(t_2)$ can be written as products of factors of the form $\nu(t')$.

We next note that D separates points of P since if $h_1, h_2 \in H$ are such that $\pi(h_1) \neq \pi(h_2)$ then there is a $t \in T$ such that $\mu(t)(h_1) \neq \mu(t)(h_2)$. Thus $(\mu(t) \circ \pi^{-1})^\wedge \in D$ takes on different values at $\pi(h_1)$ and $\pi(h_2)$.

We note also that $\beta(P)$ is Hausdorff and compact and that D contains the unit (the unit is in fact $(\mu(\tau) \circ \pi^{-1})^\wedge$). 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 $(x | P) \circ \pi$ where $x \in C(\beta(P))$. [The notation $(x | P)$ 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

$$(4.3) \quad \sum_{t \in T} c(t)\mu(t)$$

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 ϵ_2 , there exists x_2 of the form (4.3) such that $\|a(x) - x_2\| < \epsilon_2$. Let $\epsilon_1 > 0$ be such that $\epsilon_1 \|a\| < \epsilon_2$ and let

$$x_1 = \sum_{t \in T} c_1(t)\mu(t),$$

where $c_1 \in V$, be such that $\|x_1 - x\| < \epsilon_1$. We now define

$$\begin{aligned} x_2 &= \sum_{t \in T} c_1(t)a(\mu(t)) = \sum_{t \in T} c_1(t)\mu(\varphi(t)) \\ &= \sum_{t \in U} (c_1 \circ \varphi^{-1})(t)\mu(t). \end{aligned}$$

Let $c_2 \in V$ be defined by $c_2 \upharpoonright U = c_1 \circ \varphi^{-1}$ 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 $x_2 = a(x_1)$ we have

$$\|x_2 - a(x)\| \leq \|a\| \cdot \|x_1 - x\| \leq \|a\| \epsilon_1 < \epsilon_2.$$

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 η_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 $p \circ (\eta_0 e) = p(\eta_0)e \in C_a(H)$.

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 (y_0, y_1, y_2, \dots) where $y_0 = \eta_0 e$ and, for $n = 1, 2, \dots$,

$$(5.1) \quad y_n = \eta_0 e + a \times (f \circ y_{n-1})$$

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 $\pi_a(h_1) = \pi_a(h_2)$ then $y(h_1) = y(h_2)$,

Proof. Since $y \in C_a(H, X)$, then, for all $p \in X^*$, $p \circ y \in C_a(H)$. Let $x \in C_a(\beta(P_a))$

be such that $p \circ y = (x \mid P_a) \circ \pi_a$ then $(p \circ y)(h_1) - (p \circ y)(h_2) = (x \mid P_a)(\pi_a(h_1)) - (x \mid P_a)(\pi_a(h_2)) = 0$. Hence, $p(y(h_1) - y(h_2)) = 0$ 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 (H, a, h_1) is characterised by the function which takes $t \in T$ to $\nu_a(t)(h_1)$.

THEOREM 5.6. *Let $(H, a, h_1), (K, b, k_1)$ be two methods such that for all $t \in T$, $\nu_a(t)(h_1) = \nu_b(t)(k_1)$. Let y, z be the solutions of*

$$(5.2) \quad y = \eta_0 e_H + a \times (f \circ y),$$

$$(5.3) \quad z = \eta_0 e_K + b \times (f \circ z),$$

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 $y(h_1) = z(k_1)$.

Proof. Without loss of generality, we suppose that $H \cap K = \emptyset$. 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

$$(5.4) \quad c(v)(h) = a(v \mid H)(h),$$

$$(5.5) \quad c(v)(k) = b(v \mid K)(k),$$

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

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

so that $\|c\| \leq \max(\|a\|, \|b\|)$. (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

$$(5.6) \quad w = \eta_0 e_{H \cup K} + c \times (f \circ w),$$

then w is given by

$$(5.7) \quad w \mid H = y, \quad w \mid K = z.$$

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

$$\begin{aligned} (\eta_0 e_{H \cup K} + c \times (f \circ w))(h) &= \eta_0 + (a \times ((f \circ w) \mid H))(h) \\ &= \eta_0 + (a \times (f \circ (w \mid H)))(h) \\ &= (\eta_0 e_H + a \times (f \circ y))(h). \end{aligned}$$

This shows in particular that $w(h_1) = y(h_1)$ and, similarly, that $w(k_1) = z(k_1)$.

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 $\mu_c(\tau)(h) = e_{H \cup K}(h) = 1 = e_H(h) = \mu_a(\tau)(h)$. To complete the induction argument, we write $t = uv$ 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

$$\begin{aligned} \mu_c(t)(h) &= \mu_c(uv)(h) = (\mu_c(u)c(\mu_c(v)))(h) \\ &= \mu_a(u)(h) \cdot a(\mu_c(v) \mid H)(h) = \mu_a(u)(h) \cdot a(\mu_a(v))(h) \\ &= \mu_a(t)(h). \end{aligned}$$

It now follows that for all $t \in T$, $\nu_c(t)(h_1) = \nu_a(t)(h_1) = \nu_b(t)(k_1) = \nu_c(t)(k_1)$ so that $\pi_c(h_1) = \pi_c(k_1)$. Hence, by Corollary 5.5, $w(h_1) = w(k_1)$. That is, $y(h_1) = z(k_1)$.

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 (H, a, h_1) , (K, b, k_1) are so related that there exists a bijection $\varphi: H \rightarrow K$ such that $\varphi(h_1) = k_1$ and such that whenever $x \in B(K)$, $k \in K$ then $a(x \circ \varphi)(\varphi^{-1}(k)) = b(x)(k)$, 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 φ 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 (H, a, h_1) , (H', a', h'_1) to consider together for any purpose we can suppose (by replacing one of them by an identical method) that H and H' are disjoint.

6. The Group of Integration Methods. For two integration methods (H, a, h_1) , (K, b, k_1) , where $H \cap K = \emptyset$, 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

$$(6.1) \quad c(x)(h) = a(x | H)(h), \quad c(x)(k) = a(x | H)(h_1) + b(x | K)(k),$$

that is, by

$$(6.2) \quad c(x) | H = a(x | H), \quad c(x) | K = a(x | H)(h_1)e_K + b(x | K).$$

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 $(H \cup K, c, k_1)$ be the product of (H, a, h_1) and (K, b, k_1) and let f satisfy a Lipschitz condition with constant L such that $L||c|| < 1$. If η_1, η_2, η_3 are defined by*

$$(6.3) \quad y_1 = \eta_0 e_H + a \times (f \circ y_1), \quad \eta_1 = y_1(h_1),$$

$$(6.4) \quad y_2 = \eta_1 e_K + b \times (f \circ y_2), \quad \eta_2 = y_2(k_1),$$

$$(6.5) \quad y_3 = \eta_0 e_{H \cup K} + c \times (f \circ y_3), \quad \eta_3 = y_3(k_1),$$

then $\eta_2 = \eta_3$.

Proof. We first show that $||a|| \leq ||c||$ and $||b|| \leq ||c||$ so that y_1, y_2 are well defined. If $x \in B(H)$ and x' is an extension of x to $H \cup K$ such that $||x'|| = ||x||$ then $||a(x)|| \leq ||c(x')|| \leq ||c|| \cdot ||x||$. Hence, $||a|| \leq ||c||$. Now, let $x \in B(K)$ and let x' be the extension of x such that $x' | H$ is the zero function. We have $c(x') | K = b(x)$ so that $||b(x)|| \leq ||c(x')|| \leq ||c|| \cdot ||x||$ and $||b|| \leq ||c||$.

We now show that if y_3 satisfies (6.5), then $y_3 | H = y_1$ and $y_3 | K = y_2$ are the solutions of (6.3) and (6.4) respectively. We have

$$\begin{aligned} y_3 | H &= \eta_0 e_H + (c \times (f \circ y_3)) | H \\ &= \eta_0 e_H + a \times ((f \circ y_3) | H) \\ &= \eta_0 e_H + a \times (f \circ (y_3 | H)). \end{aligned}$$

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

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

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

$$\begin{aligned} y_3 | K &= \eta_0 e_K + (c \times (f \circ y_3)) | K \\ &= \eta_0 e_K + a \times (f \circ y_3 | H)(h_1) e_K + b \times ((f \circ y_3) | K) \\ &= \eta_1 e_K + b \times (f \circ (y_3 | K)), \end{aligned}$$

where we have used that fact that $y_3 | H$ and y_1 are identical.

For a method $m = (H, a, h_1)$, we may form the numbers $\nu_a(t)(h_1)$ 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 $\nu_a(t)(h_1) = w(m)(t)$ 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*

$$(6.6) \quad w(m_1 m_2) = w(m_1) w(m_2).$$

Proof. The multiplication on the right of (6.6) is, of course, that of the group G in Section 2. We will write $m_1 = (H, a, h_1), m_2 = (K, b, k_1)$ 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 λ 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, $\lambda(g, t) = \sum_{t' \in T} l_{t,t'} \hat{(t')}$ where for each t , only a finite number of $t' \in T$ are such that $l_{t,t'} \neq 0$. We write $g = w(m_1)$ so that the numbers $l_{t,t'}$ for $t, t' \in T$ take on definite values.

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

$$(6.7) \quad \mu_c(t) | H = \mu_a(t),$$

$$(6.8) \quad \mu_c(t) | K = \sum_{t'} l_{t,t'} \mu_b(t'),$$

$$(6.9) \quad \nu_c(t) | H = \nu_a(t),$$

$$(6.10) \quad \nu_c(t) | K = \sum_{t'} l_{t,t'} \nu_b(t') + w(m_1)(t) e_K.$$

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 $l_{t,t'} = 1$ if $t' = \tau$ and $l_{t,t'} = 0$ 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 $\nu_c(t) = c(\mu_c(t))$.

To complete the proof of (6.7), (6.8), (6.9), (6.10) we will prove (6.7), (6.8) for $t = uv$ 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 $\mu_c(uv) = \mu_c(u) \nu_c(v)$ (with pointwise multiplication on the right-hand side) so that

$$\begin{aligned} \mu_c(uv) | H &= (\mu_c(u)\nu_c(v)) | H = (\mu_c(u) | H)(\nu_c(v) | H) \\ &= \mu_a(u)\nu_a(v) = \mu_a(uv) \end{aligned}$$

which is (6.7).

For $k \in K$ and for $g' \in G$ defined by $g'(t') = \mu_b(t')(k)$ for all $t' \in T$, we also have

$$\begin{aligned} \mu_c(uv)(k) &= \left(\sum_{u'} l_{uu'} \mu_b(u') \left(\sum_{v'} l_{vv'} \nu_b(v') + w(m_1)(v) e_K \right) \right) (k) \\ &= \left(\sum_{u', v'} l_{uu'} l_{vv'} \mu_b(u'v') + w(m_1)(v) \sum_{u'} l_{uu'} \mu_b(u') \right) (k) \\ &= (\lambda(w(m_1), u) \lambda(w(m_1), v) + w(m_1)(v) \lambda(w(m_1), u))(g') \\ &= \lambda(w(m_1), uv)(g') \\ &= \sum_{t'} l_{uv, t'} \mu_b(t')(k), \end{aligned}$$

so that (6.8) follows. Now, evaluating (6.10) at k_1 we have

$$w(m_1 m_2)(t) = w(m_1)(t) + \lambda(w(m_1), t)(w(m_2)) = (w(m_1)w(m_2))(t)$$

by (2.16).

Since $w(m_1 m_2)$ can be determined just from $w(m_1)$ and $w(m_2)$ 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 $\{m_1, m_2; m_1 \in M_1, m_2 \in M_2\}$ 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 $m = (H, a, h_1)$ and $h_0 \notin H$ then we will be concerned with two methods. The first is

$$(6.11) \quad m_0 = (\{0\}, b, 0)$$

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

$$(6.12) \quad m_1 = (H \cup \{h_0\}, c, h_0)$$

where $c \in A(H \cup \{h_0\})$ is defined for $x \in B(H \cup \{h_0\})$ by

$$\begin{aligned} c(x)(h_0) &= -a(x | H)(h_1), \\ c(x) | H &= a(x | H) - a(x | H)(h_1)e_H. \end{aligned}$$

LEMMA 6.5. *If m_0 is given by (6.11), then $w(m_0)$ is the identity of G . If m_1 is given by (6.12) then $w(m_1) = w(m)^{-1}$.*

Proof. It is trivial that $w(m_0)(t) = 0$ for all $t \in T$. Hence, $w(m_0)$ is the identity of G .

We now form the product of m and m_1 . However, it is necessary to replace m_1 by an identical method, say $(H' \cup \{h_0\}, c, h_0)$, where H' is a copy of H which we regard as disjoint from it. If $h \in H$, we write $h' \in H'$ for its partner in a bijection between H and H' . The product method is $(H \cup H' \cup \{h_0\}, d, h_0)$ where d is defined for $x \in B(H \cup H' \cup \{h_0\})$ by

$$\begin{aligned} d(x) | H &= a(x | H), \\ d(x) | H' &= a(x | H)(h_1)e_{H'} + a'(x | H') - a'(x | H')(h'_1)e_{H'}, \\ d(x)(h_0) &= a(x | H)(h_1) - a'(x | H')(h'_1), \end{aligned}$$

and a' corresponds to a in the identification of (H, a, h_1) and (H', a', h'_1) .

Let $B' \subseteq B(H \cup H' \cup \{h_0\})$ and $B'' \subseteq B'$ be the sets defined so that if $x \in B'$ then for all $h \in H$, $x(h) = x(h')$ and if $x \in B''$ we have the additional property that $x(h_0) = 0$, then it is easy to verify that if $x \in B'$ then $d(x) \in B''$. We also see that B' is a subalgebra of $B(H \cup H' \cup \{h_0\})$ (regarded as a Banach algebra with point-wise multiplication) and B'' is an ideal of B' . We shall prove by induction that $\mu_d(t) \in B'$ and $\nu_d(t) \in B''$ for all $t \in T$. In fact, since $e_{H \cup H' \cup \{h_0\}} \in B'$ we have that $\mu_d(\tau) \in B'$. Hence, $\nu_d(\tau) = d(\mu_d(\tau)) \in B''$. If $t = uv$, where $\mu_d(u) \in B'$, $\nu_d(v) \in B''$ then $\mu_d(uv) = \mu_d(u)\nu_d(v) \in B'$ and hence, $\nu_d(uv) \in B''$. Evaluating $\nu_d(t)$ at h_0 for any $t \in T$ gives the result 0. Hence, $w(mm_1) = i$ (the identity for G). Hence, $w(m_1) = w(m)^{-1}$.

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

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

LEMMA 6.7. If $m = (H, a, h_1)$ then for all $t \in T$,

$$(6.14) \quad |w(m)(t)| \leq \|a\|^{r(t)}.$$

Proof. In fact, we will prove that, for all $t \in T$,

$$(6.15) \quad \|\mu_a(t)\| \leq \|a\|^{r(t)-1},$$

$$(6.16) \quad \|\nu_a(t)\| \leq \|a\|^{r(t)},$$

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

$$\|\nu_a(t)\| = \|a(\mu_a(t))\| \leq \|a\| \cdot \|\mu_a(t)\|.$$

Furthermore, if $t = uv$ then

$$\|\mu_a(uv)\| = \|\mu_a(u)\nu_a(v)\| \leq \|\mu_a(u)\| \cdot \|\nu_a(v)\|$$

and an inductive argument completes the proof of (6.15) and (6.16). Equation (6.14) now follows since $w(m)(t) = \nu_a(t)(h_1)$.

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 $\{|\alpha(t)|^{1/r(t)} : t \in T\}$ 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 $\alpha | T_0 = \beta | T_0$.*

Proof. It is sufficient to prove the result in each of the cases $T_0 = T_1, T_2, \dots$ where for $n = 1, 2, \dots, T_n$ is the set of trees of order $\leq n$. We will prove this result (which is trivial for $n = 1$) by induction on n .

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

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

If $r(t) \leq r(u)$, we see that $\delta(t, u) = 0$ and if for all $u \in T$, $\delta(t_1, u) = \delta(t_2, u)$, then $t_1 = t_2$. Also, it can be verified that, for a method $m_1 = (H, a_1, h_1)$, if $t \in T_n$ (where $n > 1$), then

$$(6.18) \quad \mu_{a_1}(t) = \prod_{u \in T_{n-1}} \nu_{a_1}(u)^{\delta(t, u)}.$$

From the method m_1 , we can form a new method $m_2 = (H \cup \{h_2\}, a_2, h_2)$ where $h_2 \notin H$ and a_2 is defined by

$$a_2(x) | H = a_1(x | H), \quad a_2(x)(h_2) = x(h_1).$$

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

$$\nu_{a_2}(t) | H = \nu_{a_1}(t), \quad \nu_{a_2}(t)(h_2) = \mu_{a_1}(t)(h_1).$$

Hence, using (6.18) we see that, for $t \in T_n$,

$$(6.19) \quad w(m_2)(t) = \prod_{u \in T_{n-1}} (w(m_1)(u))^{\delta(t, u)}.$$

By the induction hypothesis, a method m_1 exists so that $w(m_1)$ 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 $w(m_1) | T_{n-1}$. 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 $c \in R$, $\alpha_1 = w(m_1)$, $\alpha_2 = w(m_2)$, $m_1 = (H_1, a_1, h_1)$ and $m_2 = (H_2, a_2, h_2)$ where $H_1 \cap H_2 = \emptyset$ then $c\alpha_1 = w(m_3)$ and $\alpha_1 + \alpha_2 = w(m_4)$ where

$$m_3 = (H_1 \cup \{h_3\}, a_3, h_3), \quad m_4 = (H_1 \cup H_2 \cup \{h_4\}, a_4, h_4),$$

$h_3 \notin H_1$, $h_4 \notin H_1 \cup H_2$, and a_3, a_4 are defined by

$$a_3(x) | H_1 = a_1(x | H_1), \quad a_3(x)(h_3) = ca_1(x | H_1)(h_1),$$

$$a_4(x) | H_1 = a_1(x | H_1), \quad a_4(x) | H_4 = a_2(x | H_2),$$

$$a_4(x)(h_4) = a_1(x | H_1)(h_1) + a_2(x | H_2)(h_2).$$

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

$$(6.20) \quad \xi(t) = \sum_{i \in I} c_i \prod_{u \in T_{n-1}} C_i(u)^{\delta(t,u)},$$

$$\eta(t) = \sum_{j \in J} d_j \prod_{u \in T_{n-1}} D_j(u)^{\delta(t,u)},$$

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_j are real coefficients and C_i, D_j are real-valued functions on T_{n-1} .

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

$$(\xi\eta)(t) = \xi(t)\eta(t) = \sum_{(i,j) \in I \times J} c_i d_j \prod_{u \in T_{n-1}} (C_i(u)D_j(u))^{\delta(t,u)}$$

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 $\xi(t_1) \neq \xi(t_2)$. If $u \in T_{n-1}$ is such that $\delta(t_1, u) \neq \delta(t_2, u)$, then we can, for example, define ξ 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 ξ 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 properties 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 $m = (H, a, h_1)$ is simply a Runge-Kutta method. Write $H = \{1, 2, \dots, n + 1\}$ and define a as the linear operator represented by the matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \\ b_1 & b_2 & \cdots & b_n & 0 \end{bmatrix}$$

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 Φ 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*

$$(7.1) \quad \mu_a(t)(h) = p(t)r(t)h^{r(t)-1},$$

$$(7.2) \quad \nu_a(t)(h) = p(t)h^{r(t)},$$

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

$$(7.3) \quad p(\tau) = 1,$$

$$(7.4) \quad p(uv) = p(u)p(v)r(u)/r(uv) \quad \text{for } u, v \in T.$$

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

$$\nu_a(t)(h) = \int_0^h \mu_a(t)(k) dk.$$

For $t = \tau$, we have $\mu_a(t)(h) = 1$ which is the correct result. We also have, for $t = uv$,

$$\begin{aligned}
 \mu_a(uv)(h) &= \mu_a(u)(h)\nu_a(v)(h) \\
 &= p(u)r(u)h^{r(u)-1}p(v)h^{r(v)} \\
 &= p(u)p(v)r(u)r(uv)h^{r(uv)-1}/r(uv) \\
 &= p(uv)r(uv)h^{r(uv)-1}
 \end{aligned}$$

so that (7.1) is proved in this case.

For a real number A , we define, for $\alpha \in G$, the element $\alpha^{(A)} \in G$ such that, for $t \in T$,

$$(7.5) \quad \alpha^{(A)}(t) = A^{r(t)}\alpha(t).$$

This definition is motivated by the observation that if $\alpha = w(m)$ for $m = (H, a, h_1)$ then $\alpha^{(A)} = w(m')$ where m' is the same as m except that a is replaced by Aa . This corresponds to a change in step size by a factor A .

The next result is concerned with a special property possessed by p .

THEOREM 7.2. For $A, B \in R$, we have

$$(7.6) \quad p^{(A)} p^{(B)} = p^{(A+B)}.$$

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, $p^{(A)} p^{(B)} = w([0, 1] \cup [2, 3], c, 3)$ 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

$$(7.7) \quad \begin{aligned} \mu_c(t)(h) &= p(t)r(t)(Ah)^{r(t)-1}, \\ \nu_c(t)(h) &= p(t)(Ah)^{r(t)}, \\ \mu_c(t)(k) &= p(t)r(t)(A + B(k - 2))^{r(t)-1}, \\ \nu_c(t)(k) &= p(t)(A + B(k - 2))^{r(t)}. \end{aligned}$$

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 $r(t) \leq n$ 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 $r(t) \leq n$, $\lambda(\alpha, t) = \hat{i}$. This follows since $\lambda(\alpha, \tau) = \hat{\tau}$ and if $r(u), r(v) \leq n$ then

$$\lambda(\alpha, uv) = \alpha(v)\lambda(\alpha, u) + \lambda(\alpha, u)\lambda(\alpha, v) = (uv)^\wedge.$$

(Note that $\lambda(\alpha, t) = \hat{i}$ also for $r(t) = n + 1$.)

Hence, if $\alpha \in K_n$ and $r(t) \leq n$ then

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

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

$$\{\alpha\beta: \alpha \in K_n\} \quad \text{and} \quad \{\beta\alpha: \alpha \in K_n\}$$

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 $r(t) \leq n$ 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) $\gamma \in S \Rightarrow (\exists \alpha \in K_n)(\alpha\beta = \gamma)$,
- (iv) $\gamma \in S \Rightarrow (\exists \alpha \in K_n)(\beta\alpha = \gamma)$.

Proof of (i). We have already seen that if $\alpha \in K_n$ then $(\alpha\beta)(t) = \beta(t)$.

Proof of (ii). For fixed β , we have

$$\lambda(\beta, t) = \sum_{r(t') < r(t)} c(t')(t')^{\wedge} + i$$

where c is some real-valued function on T . Hence, if $\alpha \in K_n$ and $r(t) \leq n$ 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 $G/K_n (= G_0/(K_n \cap G_0))$ will be denoted by G_n . We use the usual group-theoretic convention of writing αK_n where $\alpha \in G$ for the coset containing α .

Definition 7.6. If $\alpha \in pK_n$, then we say that α is of order n . The greatest number n such that α is of order n is called the order of α .

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 α . 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 $\alpha \in p^{(\alpha(r))}K_n$ and we call the greatest such n the generalised order of α (or of m).

Definition 7.7. Let J_n denote the set of elements of G of generalised order n . That is

$$J_n = \bigcup \{p^{(\theta)} K_n: \theta \in R\}.$$

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 $\alpha^{-1} \in p^{(-A)}K_n$. Let $\alpha = p^{(A)}k$ where $k \in K_n$ then

$$\begin{aligned} \alpha p^{(-A)}K_n &= p^{(A)}k p^{(-A)}K_n = p^{(A)}k K_n p^{(-A)} \\ &= p^{(A)}K_n p^{(-A)} = p^{(A)}p^{(-A)}K_n = K_n \end{aligned}$$

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, \dots by

$$(7.8) \quad \begin{aligned} t_0 &= t, \\ t_n &= t_{n-1}\tau \quad (n = 1, 2, \dots). \end{aligned}$$

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 Λ is linear over G^* and such that $\Lambda(\alpha, \hat{t}) = \lambda(\alpha, t)$ whenever $t \in T$.

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

$$(7.9) \quad \Lambda\left(\alpha, \sum_{t \in T} c(t)\hat{t}\right) = \sum_{t \in T} c(t)\lambda(\alpha, t)$$

where $\alpha \in G$ and $c : T \rightarrow R$ takes all but a finite subset of T to zero.

The following properties of Λ that will be stated without proof are easily verified by substituting Λ in terms of λ in each formula.

THEOREM 7.10. *If $U \in G^*$ and $\alpha, \beta \in G$ then*

$$(7.10) \quad U(\alpha\beta) = U(\alpha) + \Lambda(\alpha, U)(\beta).$$

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

$$(7.11) \quad \Lambda(\alpha, \hat{\tau}) = \hat{\tau},$$

$$(7.12) \quad \Lambda(\alpha, U + V) = \Lambda(\alpha, U) + \Lambda(\alpha, V),$$

$$(7.13) \quad \Lambda(\alpha, cU) = c\Lambda(\alpha, U),$$

$$(7.14) \quad \Lambda(\alpha, UV) = \Lambda(\alpha, U)\Lambda(\alpha, V) + V(\alpha)\Lambda(\alpha, U).$$

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

$$(7.15) \quad \alpha(\tau_m) = \alpha(\tau)^{m+1}/(m + 1).$$

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 ($n = 1, 2, \dots$) and J_n ($n = 1, 2, \dots$) are all homogeneous.

We also have

THEOREM 7.12. *For $n = 1, 2, \dots$, 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 $m = 0, 1, 2, \dots$ and $\alpha \in G$,

$$(7.16) \quad \lambda(\alpha, \tau_m) = \sum_{i=0}^m \binom{m}{i} \alpha(\tau)^i \hat{\tau}_{m-i}.$$

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

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

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

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

so that

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

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 $U_1, U_2, \dots, U_n \in G^*$ be defined by

$$(7.17) \quad U_m = m\hat{\tau}\hat{\tau}_{m-1} - \hat{\tau}_m$$

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

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

If $\alpha \in L_n$ we have

$$(7.18) \quad \Lambda(\alpha, U_m) = \sum_{i=0}^{m-1} \binom{m}{i} \alpha(\tau)^i U_{m-i}.$$

Definition 7.13. We define $S_n^0 = \{U_1, U_2, \dots, U_n\}$ and S_n^1, S_n^2, \dots , recursively by the formula

$$(7.19) \quad S_n^i = \{\hat{t}V: t \in T, V \in S_n^{i-1}\} \cup \{V\hat{t}: t \in T, V \in S_n^{i-1}\}$$

and we define S_n as the linear subspace of G^* spanned by the union of S_n^0, S_n^1, \dots .

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 $m \geq n$. 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*

$$(7.20) \quad \Lambda(\alpha, V) \in S_n.$$

Proof. By the linearity of Λ , 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 $V = UW$ where $U \in S_n^{i-1}$ or $W \in S_n^{i-1}$. We have

$$(7.21) \quad \Lambda(\alpha, V) = \Lambda(\alpha, U)\Lambda(\alpha, W) + W(\alpha)\Lambda(\alpha, U).$$

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, $V \in S_n$,*

$$(7.22) \quad \Lambda(\alpha, V)(\beta) = 0,$$

then $\beta \in L_{n,n}$.

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

Let β_0 be defined so that $\beta_0(t) = \beta(t)$ when $r(t) < k$ and $\beta_0(t) = 0$ when $r(t) \geq k$. Thus, $\beta_0 \in L_{n,n}$ so that, by Lemma 7.14,

$$\Lambda(\alpha, m\hat{u} - \hat{v})(\beta_0) = 0 = \Lambda(\alpha, m\hat{u} - \hat{v})(\beta).$$

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

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

so that

$$\begin{aligned} (m\hat{u} - \hat{v})(\beta) &= \Lambda(\alpha, m\hat{u} - \hat{v})(\beta) - \sum_{r(t) < k} e(t)\hat{i}(\beta) \\ &= \Lambda(\alpha, m\hat{u} - \hat{v})(\beta_0) - \sum_{r(t) < k} e(t)\beta_0(t) \\ &= (m\hat{u} - \hat{v})(\beta_0) = 0. \end{aligned}$$

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, \dots, U_n defined in the proof of Theorem 7.12. In fact, using (7.3), (7.4) and the fact that $r(\tau_m) = m+1$ for $m=0, 1, \dots$, we find that $p(\tau_m) = 1/(m+1)$, $p(\tau\tau_{m-1}) = 1/m(m+1)$ so that $U_m(p) = (m\hat{\tau}\hat{\tau}_{m-1} - \hat{\tau}_m)(p) = 0$. 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 $(m\hat{u} - \hat{v})\hat{i}$ and $\hat{i}(m\hat{u} - \hat{v})$. Since $mp(u) = p(v)$, $r(u) = r(v)$ and $r(tu) = r(tv) = r(ut) = r(vt)$ we are in a position to evaluate $((m\hat{u} - \hat{v})\hat{i})(p)$ and $(\hat{i}(m\hat{u} - \hat{v}))(p)$ 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 $U_0 = U$, $U_1 = U_0\hat{\tau}$, $U_2 = U_1\hat{\tau}$, \dots .

Let L^n be the subset of G containing α if and only if for $m = 1, 2, \dots, n-1$ and

for all $U \in G^*$ we have

$$(7.23) \quad (m\hat{\tau}_{m-1}U - \hat{\tau}(\alpha)^m U + U_m)(\alpha) = 0.$$

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 $U = \hat{\tau}$ for all $t \in T$. It is found that

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

so that (7.23) can easily be verified.

Since $U_i = U_{i-1}\hat{\tau}$, we can easily show, inductively, that

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

so that

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

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

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

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 $m \geq l$. In fact we can take m as $l + n - 1$ as the next theorem shows.

THEOREM 7.19. If $l \geq 2$, 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 $j \leq n$. We then have, for $i < l$,

$$(7.24) \quad (i\hat{\tau}_{i-1}\hat{\tau}_{i-1} - \hat{\tau}(\alpha)^i \hat{\tau}_{i-1} + \hat{\tau}_{i+i-1})(\alpha) = 0$$

for all $\alpha \in L_{m,n}^l$.

Since $\tau_{i-1}\tau_{i-1} = (\tau\tau_{i-1})_{i-1}$ and since $\alpha((\tau\tau_{i-1})_{i-1}) = \alpha(\tau_{i+i-1})/j$ because

$j(\hat{\tau}\hat{\tau}_{j-1})_{i-1} - (\hat{\tau})_{i-1} \in S_n$, we deduce from (7.24)

$$(7.25) \quad (i + j)\hat{\tau}_{i+j-1}(\alpha) = j\hat{\tau}(\alpha)^i \hat{\tau}_{j-1}(\alpha).$$

Since (7.25) holds for $i < l$ and $j \leq n$, it follows that $k\hat{\tau}_{k-1}(\alpha) = \hat{\tau}(\alpha)^k$ 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 Runge-Kutta 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 $t_0 = \tau$, $t_1 = \tau\tau$, $t_2 = \tau\tau \cdot \tau$, \dots and $(\alpha\beta)_i$, α_i , β_i denote $(\alpha\beta)(t_i)$, $\alpha(t_i)$, $\beta(t_i)$ respectively for $i = 0, 1, \dots, 16$. For convenience, $r_i = r(t_i)$ is also listed.

In Table 9.3 the values of $p(t_i)^{-1}$ are given where t_0, t_1, \dots, 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, \dots, t_{16} be as shown in Table 9.2 and let α be a member of $L_{3,2}^2$. Because of the relations which define this group, it can be seen that α_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 α_i , expressions for these quantities are given. Also in this table are expressions for $(\alpha\beta)$, where β is also a member of $L_{3,2}^2$ and $i = 0, 4, 8, 10$.

Mathematics Department
The University of Auckland
Auckland, New Zealand

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.

TABLE 9.1

t	$r(t)$	$\lambda(\alpha, t)$
τ	1	\dagger
$\tau\tau$	2	$\dagger\dagger + \alpha(\tau)\dagger$
$\tau\tau\tau$	3	$\dagger\dagger\dagger + 2\alpha(\tau)\dagger\dagger + \alpha(\tau)^2\dagger$
$\tau\tau\tau\tau$	3	$\dagger\dagger\dagger\dagger + \alpha(\tau)\dagger\dagger\dagger + \alpha(\tau\tau)\dagger$
$(\tau\tau\tau\tau)\tau$	4	$(\dagger\dagger\dagger\dagger\dagger) + 3\alpha(\tau)\dagger\dagger\dagger\dagger + 3\alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau)^3\dagger$
$\tau\tau\tau\tau\tau$	4	$\dagger\dagger\dagger\dagger\dagger + \alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau\tau)\alpha(\tau)\dagger$
$\tau(\tau\tau\tau\tau)$	4	$\dagger(\dagger\dagger\dagger\dagger) + 2\alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau\tau\tau)\dagger$
$(\tau\tau\tau\tau)\tau$	4	$\dagger\dagger(\dagger\dagger\dagger) + \alpha(\tau)\dagger\dagger\dagger + \alpha(\tau\tau)\dagger\dagger + \alpha(\tau\tau\tau)\dagger$
$(\tau\tau\tau\tau)\tau\tau$	5	$(\dagger\dagger\dagger\dagger\dagger) + 4\alpha(\tau)(\dagger\dagger\dagger\dagger) + 6\alpha(\tau)^2\dagger\dagger\dagger + 4\alpha(\tau)^3\dagger\dagger + \alpha(\tau)^4\dagger$
$(\tau\tau\tau\tau)\tau\tau\tau$	5	$(\dagger\dagger\dagger\dagger\dagger)\dagger + \alpha(\tau)(\dagger\dagger\dagger\dagger\dagger) + 2\alpha(\tau)\dagger\dagger\dagger\dagger\dagger + 2\alpha(\tau)^2\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + \alpha(\tau\tau)\alpha(\tau\tau)\dagger\dagger + \alpha(\tau)^3\alpha(\tau\tau)\dagger$
$\tau\tau\tau\tau\tau$	5	$\dagger\dagger\dagger(\dagger\dagger\dagger) + \alpha(\tau)\dagger\dagger\dagger\dagger + 2\alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + 2\alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau\tau)\alpha(\tau\tau)\dagger$
$(\tau\tau\tau\tau)\tau\tau$	5	$\dagger\dagger(\dagger\dagger\dagger) + \alpha(\tau)\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau\tau)\alpha(\tau\tau)\dagger$
$\tau(\tau\tau\tau\tau)\tau$	5	$\dagger(\dagger\dagger\dagger\dagger) + 2\alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + 2\alpha(\tau\tau)\dagger\dagger\dagger + \alpha(\tau)^3\dagger\dagger + \alpha(\tau\tau\tau)\dagger$
$\tau\tau\tau\tau\tau\tau$	5	$\dagger\dagger\dagger(\dagger\dagger\dagger) + \alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger + \alpha(\tau\tau)\dagger\dagger\dagger + \alpha(\tau\tau\tau)\dagger$
$\tau(\tau\tau\tau\tau)\tau\tau$	5	$\dagger(\dagger\dagger\dagger\dagger) + 2\alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + 3\alpha(\tau)^2\dagger\dagger + \alpha(\tau\tau\tau)\dagger$
$\tau\tau\tau\tau\tau\tau\tau$	5	$\dagger\dagger\dagger(\dagger\dagger\dagger) + 2\alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger + \alpha(\tau\tau)\dagger\dagger + \alpha(\tau\tau\tau\tau)\dagger$
$\tau\tau\tau\tau\tau\tau\tau\tau$	5	$\dagger\dagger\dagger(\dagger\dagger\dagger) + \alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger + \alpha(\tau\tau)\dagger\dagger + \alpha(\tau\tau\tau\tau)\dagger$
$\tau\tau\tau\tau\tau\tau\tau\tau\tau$	5	$\dagger\dagger\dagger(\dagger\dagger\dagger) + \alpha(\tau)\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger\dagger\dagger + \alpha(\tau)^2\dagger\dagger + \alpha(\tau\tau)\dagger\dagger + \alpha(\tau\tau\tau\tau)\dagger$

TABLE 9.2

i	t_i	r_i	$(\alpha\beta)_i$
0	τ	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	$\alpha_9 + \beta_9 + \alpha_0\beta_4 + 2\alpha_0\beta_5 + 2\alpha_0^2\beta_2 + \alpha_0^2\beta_3 + \alpha_0^3\beta_1 + \alpha_1\beta_2 + 2\alpha_0\alpha_1\beta_1 + \alpha_0^2\alpha_1\beta_0$
10	$\tau\tau\cdot(\tau\tau\cdot\tau)$	5	$\alpha_{10} + \beta_{10} + \alpha_0\beta_5 + 2\alpha_0\beta_6 + \alpha_0^2\beta_2 + 2\alpha_0^2\beta_3 + \alpha_0^3\beta_1 + \alpha_2\beta_1 + \alpha_0\alpha_2\beta_0$
11	$\tau\tau\cdot(\tau\cdot\tau\tau)$	5	$\alpha_{11} + \beta_{11} + \alpha_0\beta_7 + \alpha_0\beta_6 + \alpha_0^2\beta_3 + \alpha_1\beta_2 + \alpha_0\alpha_1\beta_1 + \alpha_3\beta_1 + \alpha_0\alpha_3\beta_0$
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	$\alpha_{14} + \beta_{14} + \alpha_0\beta_6 + \alpha_0\beta_7 + \alpha_0^2\beta_3 + \alpha_1\beta_3 + \alpha_0\alpha_1\beta_1 + \alpha_6\beta_0$
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
$p(t_i)^{-1}$	1	2	3	6	4	8	12	24	5	10	15	30	20	20	40	60	120

TABLE 9.4

i	α_i	$(\alpha\beta)_i$
0		$\alpha_0 + \beta_0$
1	$\frac{1}{2}\alpha_0^2$	
2	$\frac{1}{3}\alpha_0^3$	
3	$\frac{1}{6}\alpha_0^3$	
4		$\alpha_4 + \beta_4 + \alpha_0\beta_0^3 + \frac{3}{2}\alpha_0^2\beta_0^2 + \alpha_0^3\beta_0$
5	$\frac{1}{2}\alpha_4$	
6	$\frac{1}{3}\alpha_0^4 - \alpha_4$	
7	$\frac{1}{6}\alpha_0^4 - \frac{1}{2}\alpha_4$	
8		$\alpha_8 + \beta_8 + 4\alpha_0\beta_4 + 2\alpha_0^2\beta_0^3 + 2\alpha_0^3\beta_0^2 + \alpha_0^4\beta_0$
9	$\frac{1}{2}\alpha_8$	
10		$\alpha_{10} + \beta_{10} + \frac{1}{3}\alpha_0\beta_0^4 + \frac{2}{3}\alpha_0^2\beta_0^3 + \frac{2}{3}\alpha_0^3\beta_0^2 + \frac{1}{3}\alpha_0^4\beta_0$
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	$\frac{1}{6}\alpha_0^5 - \frac{1}{2}\alpha_0\alpha_4 - \frac{1}{2}\alpha_{10}$	