Nonstandard Analysis Applied to
Advanced Undergraduate Mathematics
♢ I- II Infinitesimal Modeling. III Very Elementary Physics and Generalized Functions♢

Robert A. Herrmann

A (July - August) 1989, 1990, 1991 Instructional Development Project from the
Mathematics Department
United States Naval Academy
572C Holloway Road
Annapolis, Maryland 21402-5002

IMPORTANT NOTICE

*This publication process is now being instituted with certain results published as of January 1992, accepted for publication or in preparation. Further, some typographical errors have been corrected as of this 17 June 1997 version.

CONTENTS

Chapter 1.

INTRODUCTION

1.1 A Brief History.

Scientists who use mathematical analysis as a tool have traditionally relied upon a vague process called "infinitesimal reasoning" - a process that from the time of Archimedes until 1961 had no fixed rules nor consistent language. However, application of this intuitive process is the exact cause that has led to our great analytical successes both in scientific and engineering endeavors. Unfortunately, it also led to great controversy.

Beginning in about 1600 a schism developed between some mathematicians and the foremost appliers of this analytical tool. Leibniz approved entirely of the concept of the infinitely small or infinitely large numbers but stated that they should be treated as "ideal" elements rather than real numbers. He also believed that they should be governed by the same laws that then controlled the behavior of the ordinary numbers. He claimed, but could not justify the assertion, that all arguments involving such ideal numbers could be replaced by arguing in terms of objects that are large enough or small enough to make error as small as one wished. De l'Hospital [1715] when he wrote the first Calculus textbook used the terminology exclusively and utilized a formal "definition - axiom" process supposedly delineating the notion of the infinitesimal. Unfortunately, his first axiom is logically contradictory. D'Alembert insisted that the Leibniz concepts were without merit and only a process using a modified "limit" idea was appropriate. Euler contended in opposition to D'Alembert that the Leibniz approach was the best that could be achieved and fought diligently for the acceptance of these ideal numbers.

Due to what appeared to be logical inconsistencies within the methods, those mathematicians trained in classical logic began to demand that applied mathematicians produce "proofs" of their derivations. In answer to this criticism Kepler wrote, *"We could obtain absolute and in all respects perfect demonstrations from the books of Archimedes themselves, were we not repelled by the thorny reading thereof."* The successes of these vague methods and those scientists and mathematicians such as Leibniz, Euler and Gauss who championed their continued use quieted the "unbelievers." It should be noted that the concern of the critics was based upon the fact that they used the same vague processes and terminology in their assumed rigorous demonstrations.

The major difficulty was the fact that mathematicians had not as yet developed a precise language for general mathematical discourse, nor had they even decided upon accepted definitions for such things as the real numbers. Within their discussions they conjoined terms such as "infinitely small" with the term limit in the hopes of bringing some logical consistency to their discipline.

The situation changed abruptly in 1821. Cauchy, the foremost mathematician of this period, is believed by many to be the founder of the modern limit concept that was eventually formalized by Weierstrass in the 1870's. A reading of Cauchy's Cours d'Analyse (Analyse Algébrique)[1821] yields the fact, even to the causal observer, that he relied heavily upon this amalgamation of terms and in numerous cases utilized infinitesimal reasoning entirely for his "rigorous" demonstrations. He claimed to establish an important theorem using his methods - a theorem that Abel [1826] showed by a counterexample to be in error. No matter how mathematicians of that time period described their vague infinitesimal methods they failed to produce the appropriately altered theorem - a modified theorem that is essential to Fourier and Generalize Fourier Analysis.

Beginning in about 1870, all of the language and methods of infinitesimal reasoning were replaced in the mathematical textbooks by the somewhat nonintuitive approximation methods we

term the "$\delta - \epsilon$" approach. These previous difficulties are the direct causes that have led to the modern use of axiom techniques and the great linguistic precision exhibited throughout modern mathematical literature.

However, scientists and engineers continued to use the old incorrect infinitesimal terminology. As an example, Max Planck wrote in his books on theoretical mechanics that *"a finite change in Nature always occurs in a finite time, and hence resolves into a series of infinitely small changes which occur in successive infinitely small intervals of time."* He then attempts to instruct the student in how to obtain mathematical models from this general description. Unfortunately, at that time, such terms as "infinitely small" had no mathematical counterpart.

In many textbooks that claim to bridge the gape between abstract analysis and applications, students often receive the impression that there is no consistent and fixed method to obtain applied mathematical expressions and indeed it takes some very special type of "intuition" that they do not possess. In fact, Spiegel in his present day textbook "Applied Differential Equations" writes the following when he discusses how certain partial differential equations should be "derived." He states that rigorous methods should not be attempted by the student, but *"it makes much more sense, however, to use plausible reasoning, intuition, ingenuity, etc., to obtain such equations and then simply postulate the equations."*

In 1961, Abraham Robinson of Yale solved the infinitesimal problem of Leibniz and discovered how to correct the concept of the infinitesimal. This has enabled us to return to the more intuitive analytical approach of the originators of the Calculus. Keisler writes that this achievement *"will probably rank as one of the major mathematical advances of the twentieth century."* Robinson, who from 1944 - 1954 developed much of the present supersonic aerofoil theory, suggested that his discovery would be highly significant to the applied areas. Such applied applications began in 1966, but until 1981 were confined to such areas as Brownian motion, stochastic analysis, ultralogic cosmogonies, quantum field theory and numerous other areas beyond the traditional experience of the student.

In 1980, while teaching basic Differential Equations, this author was disturbed by the false impression given by Spiegel in the above quotation relative to the one dimensional wave equation. It was suggested that I apply my background in these new infinitesimal methods and find a more acceptable approach. The approach discovered not only gives the correct derivation for the n-dimensional general wave equation but actually solves the d'Alembert - Euler problem and gives a fixed derivation method to obtain the partial differential equations for mechanics, hydrodynamics and the like. These rigorous derivation methods will bridge the gape between a student's laboratory, and basic textbook descriptions for natural system behavior, and the formal analytical expressions that mirror such behavior. Indeed, slightly more refined procedures can even produce the relativistic alteration taught in modern physics. Moreover, pure nonstandard models are now being used to investigate the properties of a substratum world that is believed to directly or indirectly effort our standard universe. These include pure nonstandard models for the fractal behavior of a natural system, nonstandard quantum fields, a necessary and purely nonstandard model for a cosmogony (or pregeometry) that generates many different standard cosmologies as well as automatically yielding a theory of ultimate entities termed subparticles.

The major goal for writing this and subsequent manuals is to present to the faculty, and through them to the student, these rigorous alterations to the old infinitesimal terminology so that the student can once again benefit from the highly intuitive processes of infinitesimal reasoning - so that they can better grasp and understand exactly why infinitesimal models are or are not appropriate and

when appropriate why they predict natural system behavior. Except for the basic calculus and the more advanced areas, there are no textbooks nor any properly structured documentation available which presents this material at the undergraduate level. In my opinion it will be 15 to 20 years, if not much longer, before such material is available in the commercial market and instructors properly trained. An immediate solution to this problem will give your students a substantial advantage over their contemporaries at other institutions and place your institution in the forefront of what will become a major worldwide trend in mathematical modeling.

1.2 Manual Construction.

The basic construction of these manuals will be considerably different from the usual mathematical textbook. No proofs of any of the fundamental propositions will appear within the main body of these manuals. However, all propositions that do not require certain special models to establish are proved within the various appendices. A large amount of attention is paid to the original intuitive approaches as envisioned by the creators of the Calculus and how these are modified in order to achieve a rigorous mathematical theory.

Another difference lies in the statements of the basic analytical definitions. Many definitions are formulated in terms of the original infinitesimal concepts and not in terms of those classical approximations developed after 1870. Each of these definitions is shown, again in an appropriate appendix, to be equivalent to some well-known "$\delta-\epsilon$" expression. Moreover, since these manuals are intended for individuals who have a good grasp of either undergraduate analysis or its application to models of natural system behavior then, when appropriate, each concept is extended immediately to Euclidean n-spaces.

Nonstandard analysis is NOT a substitute for standard analysis, it is a necessary rigorous extension. Correct and efficient infinitesimal modeling requires knowledge of both standard and nonstandard concepts and procedures. Indeed, the nonstandard methods that are the most proficient utilize all of known theories within standard mathematics in order to obtain the basic properties of these nonstandard extensions. It is the inner play between such notions as the standard, internal and external objects that leads to a truly significant comprehension of how mathematical structures correlate to patterns of natural system behavior. Our basic approach employs simple techniques relative to abstract model theory in order to take full advantage of all aspects of standard mathematics. The introduction of these techniques is in accordance with this author's intent to present the simplest and direct approach to this subject.

Since it is assumed that all readers of these manuals are well-versed in undergraduate Calculus, then your author believes that is it unnecessary to follow the accepted ordering of a basic Calculus course; but, rather, he will, now and then, rearrange and add to the standard content. This will tend to bring the most noteworthy aspects of infinitesimal modeling to your attention at the earliest possible moment. I have this special remark for the mathematician. *These manuals are mostly intended for those who apply mathematics to other disciplines. For this reason, many definitions, proofs and discussions are presented in extended form. Many would not normally appear in a mathematicians book since they are common knowledge to his discipline. Some would even be considered as "trivial." Please be patient with my exposition.*

It has taken 300 years to solve what has been termed "The problem of Leibniz" and it should not be assumed that the solution is easily grasped or readily obtained. You will experience some startling new ideas and encounter procedures that may be foreign to you. Hopefully, experience, intuition and knowledge are not immutable. It is my firm belief that, though proper training, these

three all important aspects of scientific progress can be expanded in order to reveal the true, albeit considerably different, mathematical world that underlies all aspects of rigorous scientific modeling. It has been hoped for many years that individuals who have a vast and intuitive understanding of their respective disciplines would learn these concepts and correctly apply them to enhance their mathematical models. It is through your willingness to discard the older less productive, and even incorrect, modeling language that this goal will eventually be met.

Chapter 2.

INFINITESIMALS, LIMITED AND INFINITE NUMBERS

2.1 Some Notation and Definitions.

When `Robinson [1961]` first introduced his new concepts he used both abstract algebraic and logic notions. A few years later, for simplicity in exposition, the basic algebraic ideas where expanded and, indeed, many of these algebraic results appear here for the first time. Utilizing some fundamental facts about algebraic structures the general properties of the infinitesimals can be readily obtained. Your author has successfully used this approach while instructing a series of 400 level courses. This algebraic approach is based upon but two assumed requirements, requirements that are established from fundamental set theory in a later section. First, however, we need to recall certain elementary definitions and identify notation.

Throughout these manuals the symbol $\mathbb{R}$ denotes the real numbers which we assume is a complete ordered field.

(I) Assume that there exists another ordered field ${}^*\mathbb{R}$ with the properties that $\mathbb{R}$ is a ordered subfield of ${}^*\mathbb{R}$ (i.e. $\mathbb{R}$'s basic order and field properties are those of ${}^*\mathbb{R}$ but restricted to members of $\mathbb{R}$) **AND** $\mathbb{R} \neq {}^*\mathbb{R}$.

The set ${}^*\mathbb{R}$ is called by various names, **hyperreals**, **star-reals** or **extended reals**. Also let the symbol $\mathbb{N}$ denote the natural numbers (including zero).

Note that any ordered field $\mathcal{F}$ can be assumed to contain a copy of $\mathbb{N}$ formed by adjoining to the additive identity (the zero of $\mathcal{F}$) finite sums of the multiplicative identity (the 1 of $\mathcal{F}$). You can then simply consider $\mathbb{N} \subset \mathcal{F}$, in general. The concept **complete** is discussed in most elementary analysis books where it is almost always shown that *if $\mathcal{F}$ is a complete ordered field and $r \in \mathcal{F}$, then there exists some $n \in \mathbb{N}$ such that $|r| < n$.* This property for a field is called the **Archimedean** property for a field. As will be shown, an important and general algebraic result states that **whatever the hyperreals may be they cannot be Archimedean.**

(II) The second assumption is the following: **assume that there exists some nonzero $\epsilon \in {}^*\mathbb{R}$ such that for each positive $r \in \mathbb{R}$; it follows that**

$$0 < \epsilon < r.$$

Historically, it is unfortunate that abstract algebra was not investigated at a much earlier date since much of the difficulty Leibniz and Euler faced in having the infinitesimals accepted as genuine entities would have been erased.

Theorem 2.1.1. *Assumption (II) holds for ${}^*\mathbb{R}$ if and only if ${}^*\mathbb{R}$ is not Archimedean.*

What Theorem 2.1.1 means is that the more we know about ordered fields that are not Archimedean the better. Of course, such fields also are not complete. It is now possible to define explicitly the "infinitesimals." Note that we interpret the logical symbol $\wedge$ as the word "and," the symbol $\forall$ by any one of the expressions "for each, for all, for every" and the symbol $\exists$ by one of the expressions "there exists some, there exists one, there exists an." Since such logical notation will become significant later in this manual we might as well practice its use as soon as possible. Let $\mathbb{R}^+$ denote the positive real numbers.

Definition 2.1.1 (Infinitesimals). Let $\mu(0) = \{x | (x \in {}^*\mathbb{R}) \wedge \forall r(r \in \mathbb{R}^+ \rightarrow |x| < r\}$. Or in words, $\mu(0)$ is the set of all hyperreal numbers x such that if r is a positive real number, then the absolute value of x is less than r. The set $\mu(0)$ is called the set of **INFINITESIMALS**. In the literature the set $\mu(0)$ is also denoted by the sysmbols M_1 and o.

Definition 2.1.2. (Limited). Let $\mathcal{O} = \{x | (x \in {}^*\mathbb{R}) \wedge \exists r(r \in \mathbb{R}^+ \wedge |x| < r\}$. Or in words, $\mathcal{O}$ is the set of all hyperreal numbers x such that there exists some positive real number r such that $|x| < r$. The set $\mathcal{O}$ is called the set of **LIMITED** numbers. Robinson first called $\mathcal{O}$ the "finite" numbers and denoted them by the symbol M_0.

The term "limited" is relatively new in the literature and many nonstandard analysts still employ the term "finite." When this is done, there is some confusion when the term finite is used in the ordinary since of set theory. The concept of limited can also be interpreted as meaning that each of these hyperreal numbers is "limited by" or "bounded by" a real number. There are many reasons why the term "bounded" would be confused with the same term as employed in standard analysis and for this reason is not used for the concept of limited.

Definition 2.1.3. (Infinite). The set ${}^*\mathbb{R} - \mathcal{O} = {}^*\mathbb{R}_\infty$ is the set of **INFINITE** hyperreal numbers. Or in words, those hyperreal numbers that are not limited are the infinite. The infinite numbers are also called the unlimited numbers.

I point out that the set of infinite numbers can also be characterized analytically as follows: $x \in {}^*\mathbb{R}$ is infinite if and only if for every $r \in \mathbb{R}^+$, $|x| > r$. This characterization is often very useful. The next definition relates the infinitesimals to the original concept of when two numbers are "infinitely close." This is probably the most significant concept for infinitesimal modeling and deserves attention.

Definition 2.1.4. (Infinitely Close). Two hyperreal numbers x, y are **INFINITELY CLOSE** if $x - y \in \mu(0)$. Or in words, if there difference is an infinitesimal. The symbol used for infinitely close is $\approx$.

How do we extend these definitions for the case of the Euclidean n-spaces, $\mathbb{R}^n$ where we are using the Euclidean norm $\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}$?

Definition 2.1.5. (Euclidean Extensions). In Definitions 2.1.1 and 2.1.2 substitute for the absolute value symbol $|\cdot|$ the norm symbol $\|\cdot\|$, as it would be defined relative to ${}^*\mathbb{R}^n$. Of course, we should also substitute the term "vector" or the symbol $\mathbb{R}^n$ when appropriate.

In the next section, the basic algebraic properties for the infinitesimals, limited numbers etc. are explored along with the relationships between the above hyperreal concepts and those of the hyperreal n-spaces.

2.2 Basic Algebra.

You should expect that the basic properties for the above defined entities will be presented rapidly since the proofs only appear in the appendix. This has both advantages and disadvantages. These properties are couched in terms of some very well-known algebraic structures and we need a very brief refresher course relative to field theory.

The ordered fields $\mathbb{R}$, ${}^*\mathbb{R}$ have the weaker structural property of being a **ring.** As far as nonempty subsets of a field $\mathcal{F}$ are concerned *rings may be characterized as any nonempty* $A \subset \mathcal{F}$ *that are closed under the operations of subtraction and multiplication.* A ring is also closed under addition, has a zero and additive inverses. Rings share with the field itself the right and left distributive laws, and the usual associative and commutative properties used in parentheses manipulation. Rings can differ greatly from a field in that they need not contain an element that has a multiplicative inverse. In the case that $\mathcal{F}$ is ordered, then rings, in general, only share the basic simple order property for the field $\mathcal{F}$ when it is considered to be restricted to the ring. Our last general ring concept it that of the "ideal." Let the ring $\Re \subset \mathcal{F}$. Then nonempty $\Im \subset \Re$ is an **ideal** *of (or in)* $\Re$ *if* $\Im$ *is a subring (i.e. a ring with respect to the ring structure of* $\Re$ *and a subset of* $\Re$*) and for each* $r \in \Re$ *and each* $x \in \Im$ *it follows that* $xr \in \Im$. Thus an ideal "absorbs" the members of $\Re$ by multiplication. Before proceeding with our first list of properties notice that the infinitesimals $\mu(0) = \{x | x \in {}^*\mathbb{R} \wedge x \approx 0\}$.

Theorem 2.2.1. *The set of limited numbers,* $\mathcal{O}$*, is a subring of* ${}^*\mathbb{R}$*,* $\mathbb{R} \subset \mathcal{O}$ *and* $\mathcal{O}$ *is not a field, but if* $x \in \mathcal{O} - \mu(0)$*, then* $x^{-1} \in \mathcal{O}$*.*

Theorem 2.2.2. *The infinitesimals,* $\mu(0)$*, form a subring of* $\mathcal{O}$ *and* Γ *is infinite if and only if there is some nonzero* $\epsilon \in \mu(0)$ *such that* $\Gamma = 1/\epsilon$*.*

Theorem 2.2.3. *The set of infinitesimals,* $\mu(0)$*, is an ideal of* $\mathcal{O}$*.*

The fact that $\mu(0)$ forms an ideal of $\mathcal{O}$ is very significant for the correct theory of infinitesimals. Not only are the infinitesimals closed under finite addition and finite product, but absorb, under product, all of the real numbers. It's unfortunate that Leibniz and others could not establish such results rigorously since, if they could have, undoubtedly much of the criticism of their concepts would not have developed.

We now consider what happens when we take any real number and add to it the infinitesimals.

Definition 2.2.1. (Monad). Let $r \in \mathbb{R}$. Then a **MONAD OF (ABOUT)** r is the set $\mu(r) = \{x | x \in {}^*\mathbb{R} \wedge (x - r) \in \mu(0)\} = \{x | x \in {}^*\mathbb{R} \wedge x \approx r\}$.

Theorem 2.2.4. *The binary relation* $\approx$ *is an equivalence relation on* ${}^*\mathbb{R}$*.*

Theorem 2.2.5. *For each* $x, y \in \mathbb{R}$*,*
(i) $\mu(x) \cap \mu(y) = \emptyset$ *if and only if* $x \neq y$*,*
(ii) $\mathcal{O} = \bigcup\{\mu(r) | r \in \mathbb{R}\}$*.*

Corollary 2.2.5.1. *If* $x, y \in \mathbb{R},\ x < y,\ z \in \mu(x),\ w \in \mu(y)$, *then* $z < w$.

Corollary 2.2.5.2. *If* $x, y \in \mu(r),\ z \in {}^*\mathbb{R},\ x < z < y$, *then* $z \in \mu(r)$.

The important Theorem 2.2.5 can be expressed in words by stating that the set of monads forms an ordered partition of the limited numbers. It is beginning to appear as if the infinitesimals are indeed behaving in the manner first envisioned by the founders of the infinitesimal calculus. HOWEVER, in at least one respect the monads as well as the set of limited numbers do not share an important real number property. From the previous discussion, ${}^*\mathbb{R}$ is known not to be complete. But, are there significant sets that have upper bounds in ${}^*\mathbb{R}$ and do not have a least upper bound?

Theorem 2.2.6. *Each monad and the set of limited numbers are bounded above* [*resp. below*], *but do not possess a least upper bound* [*resp. greatest lower bound*].

2.3 Euclidean n-spaces.

The following are additional propositions that extend some of the above results to Euclidean n-spaces. Whenever possible the same notation is used and should be understood from the context. It should be obvious how the previous definitions extend to Euclidean n-spaces.

Theorem 2.3.1. *For any* $n \in \mathbb{N},\ n \geq 1$, *the set of limited vectors in* ${}^*\mathbb{R}^n$ *is equal to*

$$\mathcal{O}^n = \overbrace{\mathcal{O} \times \cdots \times \mathcal{O}}^{n \text{ factors}}.$$

Theorem 2.3.2. *For any* $n \in \mathbb{N},\ n \geq 1$, *then monad of* $\vec{v} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ *is equal to*

$$\mu(\vec{v}) = \overbrace{\mu(x_1) \times \cdots \times \mu(x_n)}^{n \text{ factors}}.$$

Theorem 2.3.3. *For any* $n \in \mathbb{N},\ n \geq 1$, *and for each* $\vec{v}, \vec{w} \in \mathbb{R}^n$,
(i) $\mu(\vec{v}) \cap \mu(\vec{w}) = \emptyset$ *if and only if* $\vec{v} \neq \vec{w}$,
(ii) $\mathcal{O}^n = \bigcup\{\mu(\vec{v}) | \vec{v} \in \mathbb{R}^n\}$.

2.4 The Standard Part Operator

In infinitesimal modeling, various methods exist that allow us to investigate what might be termed as the *micro-effects* that occur within a world called the *Nonstandard Physical World*, *(i.e. NSP-world)* or *Deductive World (i.e.D-world).* When these micro-effects are modeled by means of the infinitesimals, certain mathematical operators applied to such infinitesimals yield standard mathematical objects that are often perceived to measure the corresponding *natural world (i.e.N-world)* effects that govern phenomenological behavior. The next operator is probably the most significant of these standardizing processes.

Definition 2.4.1. (Standard Part). Let $x \in \mathcal{O}$. Define the function st: $\mathcal{O} \to \mathbb{R}$ as follows:

(i) let $\mathtt{st}(x) = r$, where $r \in \mathbb{R}$ is the unique real number such that $x \in \mu(r)$.

(ii) The function st is often extended to all of ${}^*\mathbb{R}$ by letting $\mathtt{st}(x) = +\infty$, when $x \in {}^*\mathbb{R} - \mathcal{O}$ and $x > 0$ or $\mathtt{st}(x) = -\infty$, when $x \in {}^*\mathbb{R} - \mathcal{O}$ and $x < 0$. The map st is called the **standard part operator**.

Since $\mathbb{R} \subset \mathcal{O}$ then, as will be seen, the range of st is $\mathbb{R}$ (i.e. st is a surjection.) Please note **the basic properties of the standard part operator are highly important in elementary infinitesimal analysis.**

Theorem 2.4.1. *Let $x, y \in \mathcal{O}$. Then*

(i) $x \approx y$ if and only if $\mathtt{st}(x) = \mathtt{st}(y)$,

(ii) $x \approx \mathtt{st}(x)$,

(iii) if $x \in \mathbb{R}$, then $\mathtt{st}(x) = x$,

(iv) if $x \leq y$, then $\mathtt{st}(x) \leq \mathtt{st}(y)$,

(v) if $\mathtt{st}(x) \leq \mathtt{st}(y)$, then either $x \leq y$ or $x - y \in \mu(0)$ with unknown order.

Part (iii) of Theorem 2.4.1 is what implies that the range of st $= \mathbb{R}$. The next theorem is established from the basic definitions and the fact that $\mu(0)$ is an ideal of $\mathcal{O}$. The map st is purely algebraic in character and its application is a remarkable indication of how pure abstract algebra can be utilized to obtain classical results. For algebraists, these pure algebraic characterizations are discussed following the next theorem.

Theorem 2.4.2. *Let $x, y \in \mathcal{O}$. Then*

(i) $\mathtt{st}(x \pm y) = \mathtt{st}(x) \pm \mathtt{st}(y)$,

(ii) $\mathtt{st}(xy) = (\mathtt{st}(x))(\mathtt{st}(y))$.

Corollary 2.4.2.1 *Let $x, y \in \mathcal{O}$. Then*

(i) if $\mathtt{st}(y) \neq 0$, then $\mathtt{st}(x/y) = \mathtt{st}(x)/\mathtt{st}(y)$.

(ii) if $y = \sqrt[n]{x}$, then $\mathtt{st}(y) = (\mathtt{st}(x))^{(1/n)}$, where it is always the case that if $x \in \mathcal{O}$, then $\sqrt[n]{x} \in \mathcal{O}$.

Theorem 2.4.3. *For $\vec{v} \in \mathcal{O}^n$, let $\mathtt{st}(\vec{v}) = (\mathtt{st}(x_1), \ldots, \mathtt{st}(x_n))$. The set $\mathcal{O}^n$ forms a vector space with respect to the ring $\mathcal{O}$ (i.e. a module) and as such the map* st *distributes over the vector space algebra as well as the component defined dot and for $n = 3$ the cross product operators.*

Notice that even though $\mu(0)$ is an ideal of $\mathcal{O}$ it is most definitely only a subring of ${}^*\mathbb{R}$. Indeed, products are somewhat indefinite. For example, let $0 \neq \epsilon \in \mu(0)$. Then $\epsilon^2 \in \mu(0)$. But $\epsilon(\epsilon^{-2}) \in {}^*\mathbb{R} - \mathcal{O}$, while $\epsilon(1/\epsilon) \in \mathcal{O}$ and $\epsilon^2(1/\epsilon) \in \mu(0)$.

[*This is for the algebraist. The above theorems show that the mapping st is a ring epimorphism with ker(st) = $\mu(0)$.*

Theorem 2.4.4. *The set $\mu(0)$ is a maximual ideal in $\mathcal{O}$ and the quotient ring $\mathcal{O}/\mu(0)$ is isomorphic to $\mathbb{R}$.*]

2.5 A Slight Quandary.

In 1961 when Robinson first published his new theory, it was restricted to subsets of ${}^*\mathbb{R}$ and functions defined on ${}^*\mathbb{R}$. Shortly after this `Luxemburg [1962]` expanded upon Robinson's paper but still retained the same general restrictions. It became obvious that any extensive applications of infinitesimal analysis to functional analysis would require an extension to other set-theoretic objects. `Robinson [1966]` did just this but at a great expense to elementary exposition. He used a purely logical approach and the simplified theory of types. Indeed, this author entered this field in 1968 and first used the type-theoretic approach. There is no doubt that the type-theoretic approach is beyond almost all mathematicians and appliers of mathematics that have not had specialized training. `Machover and Hirschfeld [1969]` introduced a simplification to the Robinson type-theoretic approach by restricting their nonstandard analysis to set theory itself. Unfortunately, their approach requires that individuals utilize two "set theories" - a pseudoset theory and the standard set theory. Your author wrote his doctoral dissertation in pseudoset theory. At approximately the same time `Robinson and Zakon [1969]` published a paper that further simplified Robinson's approach. They were able to show how set theory itself could be used for all the necessary formations of nonstandard analysis and that pseudoset theory was no longer necessary. Since 1969 there have been other attempts at simplifying the foundations of nonstandard analysis but it is this author's belief that they are not appropriate to those mathematicians who are not accustomed to special logical procedures.

The quandary I face is that there are no other appropriate simplifications of Robinson's theory except for the 1969 Robinson - Zakon approach. This approach definitely needs to be presented within this basic manual so that you can encounter the full power of infinitesimal analysis. The difficulty is my wish not to present numerous definitions and constructions from mathematical logic in the next chapter; but, rather, actually to do some infinitesimal analysis immediately. There is a procedure that does allow this if we postpone until a later chapter the reasons why this procedure functions.

Chapter 3.

SOME SET THEORY, CONVERGENCE AND LEIBNIZ'S PRINCIPLE

3.1 Some Set Theory.

Historically we have over 5000 years of standard mathematics available. It is certainly reasonable to utilize all of this standard mathematics in the development of what has become know as "nonstandard analysis". The term "nonstandard" should not be taken to mean that the properties to be discussed are not relative to the known properties for standard mathematics. Indeed, the opposite is the case. In the development of nonstandard mathematics, it has always been required that nonstandard structures have all the same ordinary properties as the standard mathematical structures. They are nonstandard in the sense that they have additional useful properties not possessed by the standard structures. The simplification procedures mentioned at the conclusion of the last chapter have, for the most part, eliminated certain model theoretic difficulties associated with Robinson's application of the simplified theory of types. Depending upon the type of set-theoretic entity used no special considerations need to be considered. Even though what is about to be presented seems very straightforward and is easy to grasp, it should not be assumed that these set-theoretic procedures hold in the exact form given for more extensive set-theoretic processes that are not specifically mentioned. [Note: The formal set-theort used is found in **Suppes [1960]**.]

Why do we need to consider basic set theory at all? Are not the properties of the infinitesimals, limited and infinite numbers sufficient for analysis? Didn't Leibniz stress the calculus of the infinitesimal and infinite numbers as the basis for analysis? Since mathematicians investigate the *relations between individual objects* and our simplification is set-theoretical then basic set theory is required. Indeed, even an order pair (a, b) is set-theoretically defined as $(a, b) = \{\{a, b\}, \{a\}\}$.

Theorem 3.1.1. *It is permissible to assume that*

(i) if $A \subset \mathbb{R}$, *then* $A \subset {}^*\mathbb{R}$,

(ii) if $A \subset \mathbb{R}^n$, *then* $A \subset {}^*\mathbb{R}^n$,

(iii) if $A \subset (\mathbb{R}^n) \times (\mathbb{R}^m)$, *then* $A \subset ({}^*\mathbb{R}^n) \times ({}^*\mathbb{R}^m)$.

Notice that the subsets considered in Theorem 3.1.1 exhaust almost all of the mathematical objects studied in an ordinary course in undergraduate analysis and much more. However, at present, the notion of a "constant" in a formal mathematical language needs to be discussed. From the viewpoint of a formal language, it is assumed that every subset of $\mathbb{R}$, $\mathbb{R}^n$ or $(\mathbb{R}^n) \times (\mathbb{R}^m)$ is associated with at least one constant symbol that "names" that subset. These constants are employed in two contexts. First, as language constants and nothing more, then as names for mathematical objects within our special set theory. Technically these are two different usages which are not usually mentioned in introductory texts. In infinitesimal analysis these technicalities are significant, however. In the next theorem, the notation ${}^*A$ also has two meanings. ${}^*A$ is a constant symbol in a mathematical language and, in most cases, represents a mathematical entity different from A.

Theorem 3.1.2. *If* $A \subset \mathbb{R}$ [*resp.* $\mathbb{R}^n$, $(\mathbb{R}^n) \times (\mathbb{R}^m)$], *then there exists* ${}^*A \subset {}^*\mathbb{R}$ [*resp.* ${}^*\mathbb{R}^n$, $({}^*\mathbb{R}^n) \times ({}^*\mathbb{R}^m)$] *such that*

(i) $A \subset {}^*A$ *and*

(ii) $A = {}^*A$ *if and only if* A *is finite.*

Thus Theorem 3.1.2. tells us that a collection of new sets exist - the "star-sets" - and obviously we need a certain amount of new terminology in order to discuss their properties effectively. The basic operators that define the ordered field $\mathbb{R}$ may also be considered as subsets of $\mathbb{R} \times \mathbb{R}$ or $\mathbb{R}^3$. Do we then consider these operators extended to ${}^*\mathbb{R}$ and thus "star" the operator such as writing $a \ {}^*{+}\ b$, where $a, b \in {}^*\mathbb{R}$? If one wants to be technical about the matter, then the answer would be yes. However, in practice these specific operators are not so denoted. On the contrary, one may consider the *-field operators as basic and that $\mathbb{R}$ is defined by restricting them to the set $\mathbb{R}$. In all cases, the particular defining field operator is determinable from the context.

In all that follows let $\mathcal{U} = \mathbb{R} \cup \mathbb{R}^n \cup (\mathbb{R}^n) \times (\mathbb{R}^m)$. Now a major consideration in the writing of these manuals relates to mathematical rigor. Even though all stated propositions are established in the appendixes, how explicit should the main text be when confronted with the basic construction of a nonstandard model for analysis? Analysis of the fundamental objects in $\mathcal{U}$ often requires collections of members from $\mathcal{U}$ that are themselves not members of $\mathcal{U}$. In order to satisfy the modern trend that rejects pronouncements that are not at least superficially established, a slight acquaintance with the actual construction of the standard model for this analysis seems appropriate. It is definitely not essential that you have any in depth knowledge of the fine details of the construction of this standard structure. But, be assured that every standard set-theoretic object that will ever be consider in this analysis is a member of the standard structure.

For any set W in our set theory, let $\mathcal{P}(W)$ denote the set of all subsets of W, where the operator $\mathcal{P}$ is often called the **power set operator.** Hence $A \in \mathcal{P}(W)$ if and only if $A \subset W$. Starting with $\mathbb{R} = X_0$ an object called a **superstructure** is constructed by induction as follows: assume that X_n has been defined for $n \in \mathbb{N}$. Then define $X_{n+1} = \mathcal{P}(\bigcup\{X_i | \leq i \leq n\}$. Then the standard structure - **the superstructure** - is the set $\mathcal{H} = \bigcup\{X_n | n \in \mathbb{N}\}$. Various properties relative to $\mathcal{H}$ are discussed in Appendix 3.

Definition 3.1.1. (Individual, Entity, Star-Sets). Each member of $\mathbb{R}$ is called an **INDIVIDUAL**, while each set in $\mathcal{H}$ is called an **ENTITY** and for each entity A, the set ${}^*A$ is called an **EXTENDED STANDARD** set.

It turns out that for any $W \in \mathcal{H}$ the set $\mathcal{P}(W) \in \mathcal{H}$ and there exists the set ${}^*\mathcal{P}(W)$. {What appears between these two braces is a refinement for what appeared in this place originally. It clears up a certain notational confusion. The star operator "*" behaves like a mapping from $\mathcal{P}(W)$ into ${}^*\mathcal{P}(W)$. As such, there is a slight confusion in symbols that "name" the images. The usual notation for the range would be ${}^*(\mathcal{P}(W))$. The same notation holds for $A \in \mathcal{P}(W)$, where ${}^*(A)$ is the image notation. In order to relate this notation to our original when "*" is considered as an mapping, we let ${}^*A$ be a name for the image ${}^*(A)$. That is ${}^*(A) = {}^*A$.}

The next few propositions reveal the fundamental behavior of * and the first shows that * at least preserves all finite set-theoretic operations.

Theorem 3.1.3. *Unless otherwise stated all constants represent individuals or entities.*

(i) $a \in A$ *if and only if* ${}^*a \in {}^*A$; $A \neq B$ *if and only if* ${}^*A \neq {}^*B$.

(ii) $A \subset B$ *if and only if* ${}^*A \subset {}^*B$.

(iii) ${}^*\{A_1, \ldots, A_k\} = \{{}^*A_1, \ldots, {}^*A_k\}$.

(iv) ${}^*(A_1, \ldots, A_k) = ({}^*A_1, \ldots, {}^*A_k)$.

(v) $(A_1, \ldots, A_k) \in A$ *if and only* $({}^*A_1, \ldots, {}^*A_k) \in {}^*A$.

(vi) Let $A, B \subset \mathcal{U}$. *Then* ${}^*(A \cup B) = {}^*A \cup {}^*B$, ${}^*(A \cap B) = {}^*A \cup {}^*B$, ${}^*(A - B) = {}^*A - {}^*B$, ${}^*(A \times B) = {}^*A \times {}^*B$, ${}^*\emptyset = \emptyset$.

(vii) If $A \in \mathbb{R}$, *then* ${}^*A = A$.

Recall that if R is any n-ary relation ($n > 1$), then the ***i*th projection** is the set $P_i(R) = \{x_i | (x_1, \ldots, x_i, \ldots, x_n) \in R\}$, where $1 \leq i \leq n$. Also since $R \subset A^n$ if and only if ${}^*R \subset ({}^*A)^n$ then ${}^*R$ is any n-ary relation in $({}^*A)^n$ and the set-theoretic projections are defined for ${}^*R$. If R is an binary relation, then P_1 is called the **domain** and P_2 the **range** of R.

Theorem 3.1.4. *Let* $\emptyset \neq R \subset A^n$, $(n > 1)$,

(i) ${}^*P_i(R) = P_i({}^*R)$.

(ii) If R *is a binary relation (i.e.* $n = 2$*), then* ${}^*R^{-1} = ({}^*R)^{-1}$.

The next theorem presents two technical results and is stated here for completeness.

Theorem 3.1.5. *Let* A *be an entity. Then*

(i) ${}^*\{(x,x) | x \in A\} = \{(x,x) | x \in {}^*A\}$.

(ii) ${}^*\{(x,y) | x \in y \in A\} = \{(z,w) | z \in w \in {}^*A\}$.

Theorems 3.1.3, 3.1.4 can be paraphrased by simply considering the * map to be a natural map which appears to distribute over finite "everything."

[*This is for the algebraist. The map * when restricted to* $\mathcal{P}(W)$ *is a homomorphism on the Boolean algebra* $\mathcal{P}(W)$. *The map * is a monomorphism and it is a significant fact that it is not an epimorphism.*]

3.2 Convergence.

Since the time of Zeno [350 BC] the concept of the non-finite has been somewhat controversial. In modern axiomatic set theory, except for considerations of consistency, such controversies have little meaning. However, in physical modeling, accepting the reality of the non-finite has numerous philosophical ramifications. Unfortunately, these manuals are not the appropriate form to discuss these side issues. We have mentioned the concept from an intuitive set-theoretic point of view and this will suffice since infinitesimal analysis eliminates, to a great degree, much of the more esoteric and philosophic concerns that plagued early mathematical discourse.

`Leibniz [1701]` discussed what he claimed to be a relation between his calculus of infinitesimal and infinite numbers and the "method of Archimedes" (the so-called method of exhaustion). He, of course, never established this claim and did not seem to use any definable approximation process such as the "$\delta - \epsilon$." He mentions such things as fractions with infinite numbers as denominators, infinite sequences and the like. In particular, Leibniz, de l'Hospital and even Cauchy specifically required, as axiomized by de l'Hospital, that "a curved line be considered an infinite assemblage of straight line segments each infinitely small." This concept should be distinguished from the completely different idea of a curve as composed of an infinite number of points, with no magnitude, - the *indivisibles* - as put forth by Cavalieri [1635]. Leibniz also instituted a notion he called the "sovereign principle." This principle served as his justification that the infinitesimals and his infinite numbers obey the

same rules as the ordinary real numbers. As will be seen, many of these early ideas were not correct from the rigorous view point and, moreover, Robinson's methods are very broad in context since they may be applied to all standard mathematical structures not merely to $\mathbb{R}^n$.

Even though such statements as *a sequence converges to a real number "as n goes to infinity or as n grows without bound or as n gets infinitely large"* and other such intuitive expressions may be eliminated entirely from our language of convergence and replaced with statements involving the behavior of extended standard functions that are property defined on ${}^*\mathbb{N}$, it is a misconception that the formal limit definition with its approximating techniques is no longer needed. From the view point of mathematical modeling, nonstandard and standard analysis complement each other. The effects that appear in the natural world are modeled by standard analysis. The limit approach, with all of its approximation techniques, is often considered to be a natural world approximation for events modeled within the NSP-world by nonstandard objects. Our basic definition for sequential convergence is stated in a language that parallels some of Leibniz thoughts; but, as established in appendix 3 it is equivalent to the Bolzano-Weierstrass concept. **In all that follows, let $\mathbb{N}_\infty$ denote the infinite natural numbers.**

Definition 3.2.1. (Sequential Convergence). Let the sequence $S\colon \mathbb{N} \to \mathbb{R}^n,\ (n \geq 1)$. Then S converges to $\vec{r} \in \mathbb{R}^n$ (notation $S \to \vec{r}$), if for each $\Gamma \in \mathbb{N}_\infty$ it follows that ${}^*S(\Gamma) \in \mu(\vec{r})$. Thus S converges to $\vec{r} \in \mathbb{R}^n$ if ${}^*S$ maps all of the infinite natural numbers into a single monad (i.e. ${}^*S[\mathbb{N}_\infty] \subset \mu(\vec{r})$). Recall that the value $S(n)$ of a sequence is often written as S_n.

Theorem 3.2.1. *Definition 3.2.1 (i.e. $S \to \vec{r}$) is equivalent to the limit definition for convergence of a sequence (i.e. $\lim_{n\to\infty} S(n) = \vec{r}$).*

The next result simply recasts definition 3.2.1 in terms of the algebraic standard part operator. However, this seemingly "trivial" fact yields immediately all of the basic "limit theorems" that tend to be a first stumbling block for the first year calculus student.

Theorem 3.2.2. *Let the sequence $S\colon \mathbb{N} \to \mathbb{R}^n,\ (n \geq 1)$. Then $S \to \vec{r} \in \mathbb{R}^n$ if and only if for each $\Gamma \in \mathbb{N}_\infty,\ \mathtt{st}({}^*S_\Gamma) = \vec{r}$.*

Notice that when definition 3.2.1 is stated without the symbolism then it closely parallels the conceptional process outlined by Leibniz. **A sequence converges to a real number r if its value for each infinite number is infinitely close to r.** Thus Leibniz and others explained the result that the sequence $S(n) = n^{-2}$ converges to 0 by asserting that if Γ is an infinite number, then Γ^2 is an infinite number. This implies that Γ^{-2} is an infinitesimal and thus ${}^*S(\Gamma)$ is infinitely close to 0. (They did not use the * idea, however. And, when no confusion results, the * is often omitted from the extended standard function notation.) Unfortunately, in order to apply such intuitive procedures one must know how the *-function behaves. In the next section, such an investigation is initiated.

3.3 Constants and More Constants.

In section 3.2, it was pointed out that Leibniz believed that the infinitesimal and infinite numbers "obey the same rules as the ordinary real numbers." It was necessary to develop abstract model

theory before Leibniz's vague comment could be rigorously justified. As previously discussed, every object in $\Re = \mathcal{U} \cup \mathcal{P}(\mathcal{U})$ is denoted by a "constant" symbol as is every object in $\mathcal{H}$. When theorems about the set $\mathcal{U} \cup \mathcal{P}(\mathcal{U})$ are written in terms of these constants, variables and the symbols ϵ and $=$, an intuitive "first-order language" is often used. Indeed, every theorem and definition that appears in this manual has been stated or can be reformulated in such a "language." Here are a few examples of how these intuitive expressions appear, where all the constants are assumed to represent members of $\mathcal{H}$.

Example 3.3.1. In our set theory you have the basic definition for the union of two sets $A, B \subset \Re$. This might be expressed as follows: for each $x \in \Re,\ x \in A \cup B = C$ if and only if $x \in A$ or $x \in B$. This can also be expressed in a formalized first-order statement as $\forall x(x \in \Re \rightarrow (x \in A \cup B = C \leftrightarrow (x \in A) \vee (x \in B)))$.

In the above first-order statement, it is required that the variable symbol x be restricted to members of $\Re$. *The absolute requirement is that all quantified language variables must be restricted to specific sets that are entities (i.e. to sets that are elements of various X_p) and that are represented by constants within our language.* Mostly, the variables may simply be restricted to the set $\Re$, with certain exceptions mention in future chapters. However, better comprehension is often produced when the language variables are restricted to sets of immediate interest.

Example 3.3.2. Suppose that you have a sequence that is strictly increasing. How might this be stated? For each x and y if $x, y \in \mathbb{N}$ and $x < y$, then $S(x) < S(y)$.

Example 3.3.3. Since the usual operator and relation symbols have been retained as constant symbols, it is a simple matter to give explicit functional statements. For each $n \in \mathbb{N},\ S(n) = \sin(n^2)$. Notice that the symbol n has now been used as a variable and that the symbol sin is a constant that represents the sine function.

Obviously, in order to formulate properly these first-order expressions it is essential that individuals gain facility with the basic language of mathematics. In particular, great precision must be maintained in both written and oral mathematical exposition. This is also one of the goals of the core mathematics program at the Academy.

Are there significant mathematical concepts that might not be expressible by such first-order statements? Suppose you wanted to express the general concept of mathematical induction as put forth by Peano - a concept considered to be one of the most basic in all of mathematics. *If Q is any property which may or may not hold for the natural numbers, and if (1) the natural number 0 satisfies property Q and (2) whenever a natural number n satisfies property Q, then $n + 1$ has property Q, then all natural numbers satisfy property Q.* The difficulty in expressing the induction axiom lies in the requirement that we express the phrase "Q is any property."

There are two ways of expressing a property within a first-order language. First, a property is expressed by other first-order statements or secondly, it may be claimed that a property can be represented by a member of $\mathbb{N}^m$. In the first case, we need a variable that, at least, represents

predicates in our language and this is not allowed in a first-order language. A third approach would be to replace the single induction axiom by infinitely many axioms where each axiom refers to one specific property expressed by one specific first-order expression. Unfortunately, it can be shown that you would not obtain all of the possible properties by this process.

For the second set-theoretic case, it turns out that all of the possible properties for the natural numbers still cannot be obtained since we do not have a set in our slightly restricted set theory that contains all of the objects that represent all of the properties that can be expressed by our first-order set-theoretic language. Further, using the basic language of set theory one can conceive of a property determined by the predicate P(x) = "$\emptyset \subset x$," where $\emptyset$ (the empty set) is the object that satisfies the set-theoretic expression: there exists a set x such that for all sets y, $y \notin x$. However, for set theory to be consistent it must be assumed usually that if there does exist something that satisfies this property, it is not to be called a set. Consequently, one must be very sure that an informally stated notion is indeed expressible by means of our simplistic first-order language, even though it may not be done within an informal discussion or argument.

Every object in $\mathcal{H}$ is denoted by a logical constant. Some of these constants are in the customary form such as $\mathbb{N}$, $\mathbb{R}$, 2, $\sqrt{2}$, $+$, sin and all the very well-known symbols used in standard analysis. As an abbreviation, let $C(\mathcal{H})$ denote the set all such constants. As previously indicated, in nonstandard analysis the set ${}^*\mathcal{H} = \bigcup\{{}^*X_n | n \in \mathbb{N}\}$ is constructed and it is assumed that each object in ${}^*\mathcal{H}$ is denoted by a constant. The set of all such constants is denoted by $C({}^*\mathcal{H})$. In particular, it follows that ${}^*\Re = {}^*\mathcal{U} \cup {}^*\mathcal{P}(\mathcal{U})$ and each object in ${}^*\Re$ is also denotable by a constant. (Note: ${}^*(\mathcal{P}(\mathcal{U}))$ is denoted by ${}^*\mathcal{P}(\mathcal{U})$ since the symbol $\mathcal{P}(\mathcal{U})$, in this form, is considered as one constant.) Now if $A \in C(\mathcal{H})$, it denotes an object in $\mathcal{H}$ and the notation ${}^*A$ is the constant that denotes the object in ${}^*\mathcal{H}$ obtained by means of the * map. **Also note that Theorems 3.1.2 and 3.1.3 indicate that some members of ${}^*\mathcal{H}$ are denotable by both a starred and unstarred constant and by a previous convention the basic field operators defined on ${}^*\mathbb{R}$ are not starred. In these cases, we tend to use only the unstarred notation.** Theorems 3.1.2 and 3.1.3 imply that there are infinitely many members of ${}^*\mathcal{H}$ that are not named by starred members of $C(\mathcal{H})$. But, nevertheless, they do have constant names in an extended language.

Example 3.3.4. We know that $\mathbb{N}_\infty \neq \emptyset$. Moreover, ${}^*\mathbb{N} \subset {}^*\mathcal{U}$ implies that each member of ${}^*\mathbb{N}$ is a member of ${}^*\Re$ and $\Re$ is a member of some X_p. We have used capital Greek letters as the names for some of the members of ${}^*\mathbb{N}_\infty$. By the * process we only obtain the stars of the natural numbers, where by convention we drop the *. Since the * process is one-to-one into ${}^*\mathcal{H}$ by Theorem 3.1.3 then no member of ${}^*\mathbb{N}_\infty$ is the * of any of the constants in $C(\mathcal{H})$.

3.4 The Leibniz' Principle of *-transfer.

Thus far it may not appear that the previous rules and conventions associated with infinitesimal analysis yield a simplification. However, I invite you to compare the second chapter of Robinson's 1966 book with above procedures and judge for yourself. But, why do we need to consider these specialized first-order languages in the first place? The reason lies within certain very powerful results from the subject area of mathematical logic. These fundamental propositions coupled with our conventions lead to a completely correct formulation of the Leibniz Principle, which now becomes a theorem. But, first, one final procedure needs to be discussed prior to formalizing this highly useful principle.

Definition 3.4.1. (*-transfer). Let Φ be an intuitive first-order sentence written with respect to the above rules and only containing constants from $C(\mathcal{H})$ and the basic set-theoretic binary operators $\in$, $=$, $\cup$, $\cap$, $\subset$, $\times$, $-$ etc. Then ${}^*\Phi$ is the same sentence as Φ except every constant that appears in Φ is now proceeded by a *. In this process, the conventions as to when the * may be dropped are also followed.

The sentence ${}^*\Phi$ is called the ***-transform** of the sentence Φ. The mathematical object ${}^*\mathcal{H}$ is actually obtained by application of the compactness theorem for a formal first-order language or by an algebraic construction called the ultraproduct construction [`Stroyan and Luxemburg [1976]`] and a process known as the Mostowski collapse [`Barwise [1977]`, `Herrmann [1986]`]. The Leibniz Principle is an immediate consequence of the fundamental definition of what it means to say that $\mathcal{H}$ and ${}^*\mathcal{H}$ are the universes for first-order models. You can learn about first-order models by referring to any good book in mathematical logic.

Theorem 3.4.1 (The Leibniz' Principle). *A sentence Φ holds true for members of $\mathcal{H}$ if and only if the sentence ${}^*\Phi$ holds true for members of ${}^*\mathcal{H}$.*

What Theorem 3.4.1 says is that if we let K be the set of all of those specially written sentences that hold true for members of $\mathcal{H}$, then the members of ${}^*\mathcal{H}$ along with the basic set-theoretic operators form a model for the set of all sentences obtained from K by *-transfer. *Probably much more significant is the fact that infinitely many other statements not obtainable by *-transfer hold true for members of ${}^*\mathcal{H}$.* Each member of ${}^*\mathcal{H}$ is called an **internal** entity (or internal individual if it is a member of ${}^*\mathbb{R}$). Observe that an object A is internal if and only if there exists some $p \in \mathbb{N}$ such that $A \in {}^*X_p$. Consequently, the extended standard sets are internal entities; but, there are many internal entities that are not extended standard sets and internal individuals that are not individuals. Shortly the additional terms *extended standard* and *internal* are more fully exploited. But, first a few examples.

Example 3.4.1. Let $S_1,\ S_2$ be two sequences. The definition of the addition of two such functions can be stated as follows: for every x if $x \in \mathbb{N}$, then the function $C = S_1 + S_2$ if and only if $C(x) = S_1(x) + S_2(x)$. The *-transform becomes: for every x if $x \in {}^*\mathbb{N}$, then the function ${}^*C = {}^*S_1 + {}^*S_2$ if and only if ${}^*C(x) = {}^*S_1(x) + {}^*S_2(x)$. Notice that such forms as ${}^*C(x)$ mean that the * is applied to the C only. Thus *- transfer extends to all of the ordinary definitions of this type.

Example 3.4.2. Let's look at a few of the parts of Theorem 3.1.3. (i) Using just constants themselves is legal. Thus if $a, A \in C(\mathcal{H})$, and the expression $a \in A$ holds for $\mathcal{H}$, then the *-transform is simply the expression ${}^*a \in {}^*A$. (iii) Notice that if $\{A_1, \ldots, A_n\}$ is a finite set of members of $\mathcal{H}$, then there is a constant C in $C(\mathcal{H})$ that represents this set. Clearly, it is unnecessary to mention continually that a symbol is a constant on one hand and a name for an object in $\mathcal{H}$ on the other. The context of a statement will usually serve to indicate a constants usage. This first-order statement is: for every $x,\ x \in C$ if and only if $x = A_1 \vee x = A_2 \vee \cdots \vee x = A_k$. This is a collection of finitely many symbols and is thus allowed. The *-transfer becomes: for every $x,\ x \in {}^*C$ if and only if $x = {}^*A_1 \vee x = {}^*A_2 \vee \cdots \vee x = {}^*A_k$. [Note: translate $\vee$ by the word "or."]

[Note: Our defined operators $+,\ \cdot,\ |\cdot|,\ \|\cdot\|$, etc. are usually considered but the *-transfer of the operators (the nonstandard extension to ${}^*\mathbb{R}$, ${}^*\mathbb{R}^n$) as they are defined on $\mathbb{R}$ and $\mathbb{R}^n$.]

Example 3.4.3. An argument. *-transfer is certainly important, but it is our ability to argue by using statements that contain symbols that are not obtainable by *-transfer which is the key to nonstandard analysis. Indeed, as will be established, there are three different categories of constant mathematical symbols used in the following argument. Let $\Gamma \in {}^*\mathbb{N} - \mathbb{N}$. [$\Gamma \in C({}^*\mathcal{H})$ and there is no symbol for ${}^*\mathbb{N} - \mathbb{N}$ in $C(\mathcal{H})$.] Since $\Gamma \notin \mathbb{N}$ [it turns out that $\mathbb{N} \notin {}^*\mathcal{H}$)] then for each $x \in \mathbb{N},\ x < \Gamma$ for if not, then there exists some $n \in \mathbb{N}$ such that $0 < \Gamma < n$ and for each $x \in \mathbb{N},\ x \neq \Gamma$. However, we also know that the set $\{x | 0 < x < n\}$ is a finite set and thus each member is an element of $\mathbb{N}$ (even under *-transfer by our conventions). This would yield a contradiction. Now since $\mathbb{R}$ [$\mathbb{R} \in C(\mathcal{H})$] is Archimedean then for each $r \in \mathbb{R}$ there exists some $n \in \mathbb{N}$ such that $|r| < n$. Hence, $|r| < \Gamma$ implies that $\Gamma \in \mathbb{N}_\infty$. Therefore, ${}^*\mathbb{N} - \mathbb{N} \subset \mathbb{N}_\infty$. Since it is obvious that $\mathbb{N}_\infty \subset {}^*\mathbb{N} - \mathbb{N}$ then this implies that $\mathbb{N}_\infty = {}^*\mathbb{N} - \mathbb{N}$.

Theorem 3.4.2. *The set of infinite natural numbers* $\mathbb{N}_\infty = {}^*\mathbb{N} - \mathbb{N}$.

Mathematicians are so practiced in arguments such as illustrated in example 3.4.3 that it is often not realized that there are at least five categories of constants as symbols being used. (1) There are the **standard (unstarred) constants** that appear in $C(\mathcal{H})$. (2) The **extended standard constants** that appear in $C({}^*\mathcal{H})$. (3) The **internal constants** that appear in $C({}^*\mathcal{H})$ but are not obtained by *-transfer. (4). Then **external constants**, such as $\mathbb{N}_\infty$ that represent mathematical objects but do not fit categories (1) (2) (3). Last, there are the constants the comprise that word forms of the "metalanguage" that discusses and establishes things about the mathematical objects, such as the phrase "there is not a symbol for." Theorem 3.1.2 and part (vii) of Theorem 3.1.3 indicate the only instances where two categories - (1) and (2) - overlap in the sense that there are definitely two distinct constants representing the same mathematical object. With respect to the next definition, these basic facts about the usage of the constants - especially their correspondence to various mathematical objects - should always be kept in mind.

Definition 3.4.2. (Internal, External.) Any member of ${}^*\mathcal{H}$ will be termed, in general, an **INTERNAL ENTITY** or **INTERNAL INDIVIDUAL**. Any $A \subset {}^*X_p$ that IS NOT INTERNAL is termed an **EXTERNAL ENTITY.**

Definition 3.4.2 now allows for a somewhat better but equivalent statement of the Leibniz Principle for objects such as $\Re$.

Theorem 3.4.3. (Leibniz Principle Restated). *A sentence* Φ *holds true for members of* $\Re = \mathcal{U} \cup \mathcal{P}(\mathcal{U})$ *if and only if the sentence* ${}^*\Phi$ *holds true for the members of* ${}^*\mathcal{U}$ *or internal subsets of* ${}^*\mathcal{U}$.

3.5 A Few Simple Applications.

Using infinitesimal procedures, there are simple and specifically describable methods for determining the limit of a sequence, where Theorem 3.2.1 allows us to use the expression "limit of a

sequence" as an abbreviation for the phrase "real number to which the a sequence converges." The next applications amply illustrate the most significant of these procedures.

Application 3.5.1. *This is an example of the direct application of the algebraic properties of the infinite and infinitesimal numbers.*

Let $p \in \mathbb{N},\ p > 0$ and assume that for each $0 < n \in \mathbb{N},\ S_n = (1/n)^p$. Then $S \to 0$.

Proof. Let $\Gamma \in {}^*\mathbb{N} - \mathbb{N}$. Then $(1/\Gamma) \in \mu(0)$ implies that $(1/\Gamma)^p \in \mu(0)$. Thus result follows from Definition 3.2.1.

Application 3.5.2. *A bounding method.*

Let $a \in \mathbb{R},\ 0 < |a| < 1$. For each $0 < n \in \mathbb{N}$, let $S_n = a^n$. Then $S \to 0$.

Proof. By induction it is not difficult to show that if $0 \leq b \in \mathbb{R}$, then for each $n \in \mathbb{N},\ (1+b)^n \geq (1 + nb)$. Let $b = (1/|a|) - 1$. Then $b > 0$ and $|a| = 1/(1 + b)$. Thus for $n \in \mathbb{N}$,

$$0 < (|a|)^n = \frac{1}{(1+b)^n} \leq \frac{1}{1+nb} < \frac{1}{bn},\ (n > 0).$$

Thus by *-transfer for each $n \in {}^*\mathbb{N}$

$$0 < (|a|)^n < (1/b)(n)^{-1},\ (n > 0).$$

In particular, for each $\Gamma \in {}^*\mathbb{N} - \mathbb{N}$,

$$0 < (|a|)^\Gamma < (1/b)(\Gamma)^{-1}.$$

The result now follows since $a^\Gamma \in \mu(0)$.

Application 3.5.3 *A bounding method, redefinition and the standard part operator.*

Let $1 < a \in \mathbb{R}$. For each $0 < n \in \mathbb{N}$, let $S_n = (a)^{1/n}$. Then $S \to 1$.

Proof. Define the sequence $Q_n = a^{1/n} - 1,\ 0 < n \in \mathbb{N}$. Then

$$a = (1 + Q_n)^n = 1 + nQ_n + \textit{other positive terms},$$

where $0 < n \in \mathbb{N}$. Hence $a > nQ_n,\ \forall n \in \mathbb{N},\ (n > 0)$. By *-transfer $a > nQ_n,\ \forall n \in {}^*\mathbb{N}\ (n > 0)$. In particular, $a > \Gamma\, {}^*Q_\Gamma,\ \forall \Gamma \in {}^*\mathbb{N} - \mathbb{N}$. Hence

$$\forall \Gamma \in {}^*\mathbb{N} - \mathbb{N},\ 0 < {}^*Q_\Gamma < a(1/\Gamma) \in \mu(0).$$

Thus $Q \to 0$ implies that $\mathtt{st}({}^*Q_\Gamma) = 0 = \mathtt{st}({}^*S_\Gamma) - \mathtt{st}(1) = \mathtt{st}({}^*S_\Gamma) - 1,\ \forall \Gamma \in {}^*\mathbb{N} - \mathbb{N}$. The result follows from Theorem 3.2.2.

Application 3.5.4. *A bounding method, redefinition and the standard part operator.*

Let $S_n = \sqrt[n]{n},\ \forall n \in \mathbb{N},\ (n > 0)$. Then $S \to 1$.

Proof. Let $Q_n = S_n - 1 = \sqrt[n]{n} - 1,\ \forall n \in \mathbb{N},\ (n > 0)$. Then

$$n = (1 + Q_n)^n \geq (n(n-1)/2)Q_n^2,\ \forall n \geq 2,$$

by the binomial expansion. Hence

$$0 \leq Q_n \leq \sqrt{\frac{2}{n-1}},\ \forall n \geq 2.$$

By *-transfer,

$$0 \leq Q_\Gamma \leq \sqrt{\frac{2}{\Gamma - 1}},\ \forall \Gamma \in {}^*\mathbb{N} - \mathbb{N}.$$

But, $\sqrt{\frac{2}{\Gamma-1}} \in \mu(0),\ \forall \Gamma \in {}^*\mathbb{N} - \mathbb{N}$ implies that $Q \to 0$; which implies that $S \to 1$.

The bounding techniques and algebraic manipulations illustrated by above examples are actually of the same type that were used originally with old style infinitesimal analysis to argue for these results. As previously mentioned the standard part operator establishes the classical limit theorems which, of course, can be applied always. However, the standard part operator can be applied directly without referring to the limit theorem at all.

Application 3.5.5 *Using the standard part operator.*

Find the limit of the sequence $S_n = ((1/n)^{10})\sqrt[n]{n},\ n > 0$. Let arbitrary $\Gamma \in {}^*\mathbb{N} - \mathbb{N}$. Then $\mathtt{st}(1/\Gamma) = 0$ implies that $\mathtt{st}((1/\Gamma)^{10}) = 0$. From application 3.5.4, $\mathtt{st}((\Gamma)^{1/\Gamma}) = 1$. Thus $\mathtt{st}({}^*S_\Gamma) = 0 \cdot 1 = 0$. Hence, $S \to 0$.

One of the most significance statements made by a researcher is "What if...?" The basic propositions of infinitesimal analysis tend to lend themselves to many such "What if...?" type questions. Consider, for example, Theorem 3.2.2. One is prone to ask; what if $\forall \Gamma \in \mathbb{N}_\infty,\ \mathtt{st}({}^*S_\Gamma) \in \mathcal{O}^n$? Or, what if there exists some $\Gamma \in \mathbb{N}_\infty$ and $\mathtt{st}({}^*S_\Gamma) \in \mathcal{O}^n$? We state a theorem that represents an interesting result relative to the last "What if ...?" - a result established in appendix 3 solely by means of the standard part operator.

Theorem 3.5.1. *Let $S\colon \mathbb{N} \to \mathbb{R}$. If S is an increasing [resp. decreasing] sequence and there exists some $\Gamma \in \mathbb{N}_\infty$ such that ${}^*S_\Gamma \in \mathcal{O}$, then $S \to \mathtt{st}({}^*S_\Gamma)$.)*

Chapter 4.

SOME MODELING WITH THE INFINITE NUMBERS

4.1 Historical Confusion.

First, I point out that the infinite numbers are completely different from the idea of extending the real numbers by adjoining the new objects $\pm\infty$ to $\mathbb{R}$ and impressing upon these objects certain topological and algebraic properties. This can be readily shown by considering the concept of when $\lim_{n\to\infty} = +\infty$. Let $\mathbb{R}^+_\infty$ denote the positive infinite hyperreal numbers.

Definition 4.1.1. **($S \to +\infty$).** Let $S\colon \mathbb{N} \to \mathbb{R}$. Then $S \to +\infty$ if for each $\Gamma \in \mathbb{N}_\infty$, ${}^*S(\Gamma) \in \mathbb{R}^+_\infty$.

Theorem 4.1.1. *Definition 4.1.1 is equivalent to the limit definition for a sequence $S \to +\infty$.*

Theorem 4.1.1 indicates that the values of ${}^*S$ are scattered throughout the set $\mathbb{R}^+_\infty$.

Theorem 4.1.2. *Let $r \in {}^*\mathbb{R}^+$, $S\colon \mathbb{N} \to \mathbb{R}$ and $S \to +\infty$. Then for each $\Omega \in \mathbb{N}_\infty$ there exists some $\Delta \in \mathbb{N}_\infty$ such that ${}^*S(\Omega) + r \leq {}^*S(\Delta)$.*

Thus the sequential property that $S \to +\infty$ does not correspond to the idea that a sequence might "converge" to an infinite number. Indeed, we have not even mentioned the idea of a monad about an infinite number even though it is possible to extend monad theory to cover such cases. This is particular significant for applied modeling when the infinite series is discussed relative to the sequence of partial sums each infinitely small.

Recall that except for modern times most mathematics was almost used exclusively for applied physical or geometric modeling and was not consider as a study of abstract entities. De l'Hospital apparently believed in the objective reality of both the infinitesimal and infinite quantities as well as the existence of sets that contained infinitely many members. To him, they existed in the natural world. When de l'Hospital wrote that it was a requirement that one must regard a curve as a totality of an infinity of straight line segments, each infinitely small: or *"(which is the same) as a polygon with an infinite number of sides, each infinitely small, which determine by the angle at which they meet, the curvature of the curve..."* he apparently meant that all of these intuitively expressed objects exist in reality. Leibniz did not approve of this interpretation. Over and over again, Leibniz proclaimed that such objects as the infinitesimal or infinite numbers were "ideal" or imaginary. He claimed that they are theoretically useful but that they did not correspond to real natural things. Moreover, Leibniz apparently accepted the concept of a potentially infinite set as well, rather than the possibility of an objectively real infinite set of objects. Almost all of the ancient controversy as to the reality of such notions apparently came about as a direct result of the basic philosophical and theological predilections of the investigators. Their general philosophic belief systems were reflected in both their scientific and mathematical views.

With respect to the physical possibility of there being objectively real objects that are characterized by infinitesimals or even infinite numbers, Robinson has replace these abstruse philosophical

considerations with the following observation. In the first fundamental paper delineating his theory, he wrote: "*For phenomena on a different scale, such as are considered in Modern Physics, the dimensions of a particle may not be observable directly. Accordingly, the question whether or not a scale of non-standard analysis is appropriate to the physical world really amounts to asking whether or not such a system provides a better explanation of certain observable phenomena than the standard system of real numbers. The possibility that this is the case should be borne in mind.*" [Fine Hall, Princeton University] `Robinson [1961]`

As previously mentioned Zeno proposed his famous physical paradox of Achilles' and the Tortoise in their never ending (potentially infinite time) foot race. It is claimed, that this paradox is resolved by modern mathematics through application of the infinite series. Such a series can supposedly model this physical foot race - a statement that is obviously false. Such a series would require the non-mythlogical object, the Tortoise, and the mythological Archilles to have variable "sandal" sizes - sizes that decreased to the point of being unmeasurably small.

De l'Hospital accepted a curve as being identifiably the same as a polygon with a fixed infinite number of line segments comprising its sides. On the other hand, Eudoxus [370 BC] devised the method of exhaustion which assumes the true existence of a finite sequence of inscribed and circumscribed polygons. In general, for a closed non-polygonal curve none of these Eudoxus polygons were considered to be the curve under investigation; but, rather, by the "continuity process" they would continually squeeze the curve between these two types of polygons and "exhaust" the space in between. By this process the length of a curved segment was conceived of as an intuitive sequence composed of portions of the polygon's perimeters. Thus developed the idea of a partial sum that represented the sum of the lengths of the sides of an n-gon - a finite sum that remained finite but acquired more and more terms. Those that employed this method often guessed at a specific formula then justified their guess by indirect and not direct argument. In modern times, Planck described the accepted procedure for modeling the behavior of a natural system when he wrote that: *"a finite change in Nature always occurs in a finite time, and hence resolves into a <u>series</u> of infinitely small changes which occur in successive infinitely small intervals of time."* What might the term "series" mean in Planck's statement?

Leaving aside the ontological question associated with the notion of the infinite, it will be demonstrated that Robinson's theory of the infinitesimal and infinite hyperreal numbers brings a concrete and rigorous language to the above vague methods of physical and geometric modeling. Unfortunately, in doing so, many of these previous ideas will require modification.

4.2. <u>The Internal Definition Principle.</u>

Clearly there is a need to acquire a better understanding of the relationship between the concepts of the infinite series, the sequence of partial sums, the internal process of partial summing with its arithmetic and the notion expounded by Planck. But, prior to examining these concepts, yet another technical procedure needs to be discussed.

In definition 3.4.1, our first-order language was extended to include the basic set-theoretic binary operators. As explained in the Appendix to Chapter 3 immediately after the proof of Theorem 3.1.3, our first-order language may also include the symbol $(\cdot, \ldots, \cdot)$ for n-tuple formation where the coordinates are either variables or constants. Under *-transfer these operators and the n-tuple formation symbol are not starred in any Φ that has been properly formulated with the variables restricted to entities in $\mathcal{H}$.

Example 4.2.1. Assume that you are given some relation R, a nonempty set A and the

next statement. For each x if $x \in \mathcal{P}(\mathcal{U})$, then $(x, A) \in R$ and $x \subset A$. Or, as a formal statement $\forall x(x \in \mathcal{P}(\mathcal{U}) \rightarrow (x, A) \in R \wedge x \subset A)$. Then the *-transfer would read: For each x if $x \in {}^*\mathcal{P}(\mathcal{U})$, then $(x, {}^*A) \in {}^*R$ and $x \subset {}^*A$.

Please note that we needed to star the symbol $\mathcal{P}(\mathcal{U})$. This would also be the case if a variable appeared where the $\mathcal{U}$ appears.

When a collection C of mathematical expressions are written they include "mathematical" variables and constants. Within our standard set theory the constants represent individuals or entities. From the view point of the mathematical structure, the entities represent such things as basic sets, operators, relations, specific n-tuples and other definable objects. In general, the variables either vary over every member of some set or represent some one unknown element. It is common practice for mathematicians to use different variable symbols to represent elements of distinct sets or distinct positions in n-tuples and the like. Within C certain of these variables may also be "quantified." What this means is that if v is one of the variables in C, than there also appears in C in the customary location the phrase "for all v" (i.e. $\forall v$) or "there exists some v" (i.e. $\exists v$). Any variable in C that is not associated with some quantifier is termed a **free variable.** Now in formal logic the concept of the free and quantifier bounded variable must be more carefully described since formally the same variable can appear both in a free and not free position. Since we are trying to be as non-technical as possible and since it is customary to use many different variable symbols in mathematical prose, this somewhat vague free variable definition should suffice. If it still is confusing, then most elementary logic books have a more formal presentation.

Example 4.2.2. When the definition of continuity is expressed one usually states the following: The function $f\colon A \rightarrow \mathbb{R}$ is continuous at $p \in A$ if for each $\epsilon \in \mathbb{R}^+$ there exists some $\delta \in \mathbb{R}^+$ such that whenever $0 \leq |x-p| < \delta$ and $x \in A$, then $0 \leq |f(x)-f(p)| < \epsilon$. The symbols $f,\ A,\ \mathbb{R},\ \mathbb{R}^+,\ p,\ -,\ |,\ \leq,\ <,\ 0$ are all considered as constants. The symbols $\epsilon,\ \delta,\ x$ are variables. This entire definition may be formally expressed by $\Phi = \forall x(x \in \mathbb{R}^+ \rightarrow \exists y(y \in \mathbb{R}^+ \wedge \forall z(z \in A \wedge 0 \leq |z - p| < \delta \rightarrow |f(z) - f(p)| < \epsilon)))$.

Suppose that you are interested in the subset of A in example 4.2.2 that comprises all of the points of continuity. In this case, the constant p is considered a variable and one writes such a set as $\{p|p \in A \wedge \Phi(p)\}$. Built into our set builder notation is the quantifier $\forall$. "The set <u>of all</u> $p \in A$ such that $\Phi(p)$ holds true." The expression $\Phi(p)$ is now considered a formula in one variable p. Note that p is a free variable. Of course, these set builder formulas should present no difficulties since these language constructions are the ordinary and customary ones used by the mathematical community. Two more examples should sufficiently illustrate this easily grasped relation between our first-order language and common mathematical usage.

Example 4.2.3. Most of the time when we write mathematical formula in variables they are considered as free. Such an expression as $y = 3x$ has been defined for many different structures. Even though to obtain the graph of such an expression one needs to know the domain and codomain, this is not necessary until the structure itself is considered. Thus the graph may be the $\{(x, y)|x \in [0, 3] \subset \mathbb{R} \wedge y \in \mathbb{R} \wedge y = 3x\}$. On the other hand, we might also have $\{(x, y)|x \in \mathbb{R}^7 \wedge y \in \mathbb{R}^7 \wedge y = 3x\}$. Thus $y = 3x$ is a mathematical formula, $\Phi(x, y)$, in two variables with a constant 3 and a constant operator (multiplication) that can be used to generate many different sets in our set theory.

Formulas that appear in set builder notation can also contain quantifiers as the next example indicates.

Example 4.2.4. Let the sets $A, B \in X_p$ and let B^A denote the set of all functions with domain A and codomain B. Then $B^A \in X_{p+3}$. One might want to consider a special subset of B^A defined by $\{x|x \in B^A \textit{ and there exists } y \in A \textit{ such that } x(y) > 3\}$ (i.e. $\{x|x \in B^A \wedge \exists y(y \in A \wedge x(y) > 3\})$ or $\{x|x \in B^A \wedge \exists y \exists z(y \in A \wedge z \in B \wedge (y, z) \in x \wedge z > 3\})$. This set exists by the axiom of comprehension and every though the defining expression contains more than one variable, all but the x are bounded by a quantifier that appears immediately to the left.

The basic reason for discussing and presenting these simple examples lies in two most interesting results that allow for the set-theoretic generation of all extended standard or internal sets and n-ary relations by means of such set builder formulas. Of course, n-ary relations are sets but are singled out specifically due to their obvious usefulness. Recall that $C(\mathcal{H})$ denotes the set of all constants that signify members of the set $\mathcal{H}$ while $C({}^*\mathcal{H})$ the names for members of ${}^*\mathcal{H}$. A formula in our first-order language is called standard [resp. internal] **bound** if each quantified variable is restricted to an object represented by a constant in $C(\mathcal{H})$ [resp. $C({}^*\mathcal{H})$]. What this means is that, for a $\mathcal{H}$ [resp. ${}^*\mathcal{H}$] interpretation of the formula, each bounded variable must be interpretation as varying over a set contained in $\mathcal{H}$ [resp. ${}^*\mathcal{H}$]. The formula in Example 4.2.2 is bound. A formula like $\forall x \exists y(x \in y)$ is not bound.

Theorem 4.2.1. (The Extended Standard Definition Principle)

(i) A set A in our set theory is an extended standard set (i.e. there exists some $B \in C(\mathcal{H})$ such that $A = {}^*B$) if and only if there exists some standard set D and a standard bound formula $\Phi(x)$ in one free variable where each constant in $\Phi(x)$ is a member of $C(\mathcal{H})$ and

$$A = \{x|x \in {}^*D \wedge {}^*\Phi(x)\}.$$

(ii) A set A is an extended standard n-ary relation ($n > 1$) if and only if there exist n standard sets $D_1, \ldots, D_n$ and a standard bound formula $\Phi(x_1, \ldots, x_n)$ in n free variables where each constant in $\Phi(x_1, \ldots, x_n)$ is a member of $C(\mathcal{H})$ and

$$A = \{(x_1, \ldots, x_n)|x_1 \in {}^*D_1 \wedge \cdots \wedge x_n \in {}^*D_n \wedge {}^*\Phi(x_1, \ldots, x_n)\}.$$

Theorem 4.2.2. (The Internal Definition Principle)

(i) A set A in our set theory is an internal set if and only if there exists some internal set D and a internal bound formula $\Phi(x)$ in one free variable where each constant in $\Phi(x)$ is a member of $C({}^*\mathcal{H})$ and

$$A = \{x|x \in D \wedge \Phi(x)\}.$$

(ii) A set A is an internal n-ary relation ($n > 1$) if and only if there exist n internal sets $D_1, \ldots, D_n$ and an internal bound formula $\Phi(x_1, \ldots, x_n)$ in n free variables where each constant in $\Phi(x_1, \ldots, x_n)$ is a member of $C({}^*\mathcal{H})$ and

$$A = \{(x_1, \ldots, x_n)|x_1 \in D_1 \wedge \cdots \wedge x_n \in D_n \wedge \Phi(x_1, \ldots, x_n)\}.$$

Internal objects are basic to nonstandard analysis since it is only internal properties that hold for the model ${}^*\mathcal{H}$. As far as physical modeling is concerned **internal objects represent NSP-world effects that directly or indirectly yield the observed natural world behavior being modeled approximately by a standard mathematical structure.**

4.3. Hyperfinite Summation.

In example 4.2.4 the set-theoretic notation B^A is utilized to denote the set of all functions with domain A and codomain B. As previously mentioned there are procedures that require objects in $\mathcal{H}$ that may not be members of $\Re$. Assume that $\mathcal{A},\ \mathcal{B}$ are sets of subsets and $\mathcal{A},\ \mathcal{B} \in X_p$. Now let the sets $A \in \mathcal{A},\ B \in \mathcal{B}$. Then $A, B \in X_{p-1},\ (p \geq 2),\ A \cup B \subset X_0 \cup X_{p-2}$ implies that $A \cup B \in X_{p-1}$ and if $a \in A,\ b \in B$, then $\{a\}, \{a, b\} \in X_{p-1}$. Hence $(a, b) \in X_p$. Thus if $f \in B^A$, then $f \in X_{p+1}$. Indeed, $A \times B \in X_{p+1}$. Therefore, $B^A \in X_{p+2}$. It is possible to consider the formation of each set B^A as an operator $F(x, y)$ where $x \in \mathcal{A}$ and $y \in \mathcal{B}$. The *-transfer process can be extended to this operator in the same manner as is done with the $\cup,\ \cap,\ \times,\ -,\ (\cdot, \ldots, \cdot)$ operators in the sense that is it not starred when written entirely in variable form. For specific members of $C({}^*\mathcal{H})$, if $D \in {}^*\mathcal{A}$ and $E \in {}^*\mathcal{B}$, then ${}^*F(D, E) \in {}^*X_{p+2}$ and ${}^*F(D, E)$ is the internal set of all internal functions with domain D and codomain E. This last fact comes from the *-transfer of a general characterization for the set-theoretic concept of x^y.

Example 4.3.1. Let $p \in {}^*\mathbb{N}$. Then the set $\{x | x \in {}^*\mathbb{N} \wedge 0 \leq x \leq p\} = [0, p]$ is an internal subset of ${}^*\mathbb{N}$. Notice that the symbol $[0, p]$ can be used as an abbreviation for this set's defining property (i.e. $x \in [0, p]$ if and only if $x \in {}^*\mathbb{N} \wedge 0 \leq x \leq p$) which can be restated in an appropriate first-order expression and substituted for the notation $[0, p]$. With this in mind it is clearly possible to now consider p as a variable.

Hence it follows that $BHF = \{y | y \in {}^*\mathcal{P}(\mathbb{N}) \wedge \exists p(p \in {}^*\mathbb{N} \wedge y = [0, p])\}$ is an internal (indeed, an extended standard set) of subsets of ${}^*\mathbb{N}$. The set BHF is called the set of all **basic hyperfinite subsets of** ${}^*\mathbb{N}$.

A nonempty $A \subset B$ is **finite** if there exists some $n \in \mathbb{N}$ and a function $f : [0, n] \to B$ such that the range of f (i.e. $P_2(f)) = A$. The intuitive idea of a finite set appears to be equivalent to this functional definition if you are willing to accept such things as the "finite" axiom of choice and elementary procedures of recognition. The intuitive idea of the finite is based upon the human recognition of a distinction between symbols written on paper as they are considered geometric forms and consequently it is related to the most basic aspects of concrete geometry. It is this accepted recognition of the differences between geometric forms that than allows one to give a concrete meaning to a correspondence between these forms and $\mathbb{N}$. One does not really establish that such a correspondence exists but its existence is accepted as part of the metamathematical methods.

Theorem 3.1.4 states that for any standard function $f,\ {}^*(P_i(f)) = P_i({}^*f),\ i = 1, 2$. The projections P_i can be considered as maps from the set of all nonempty subsets of $A_1 \times \cdots \times A_n$ into the sets A_i, where $1 \leq i \leq n$. A modification of proof of Theorem 3.1.4 yields

Theorem 4.3.1. *Let $n > 1$. Then for each internal $R \subset {}^*A_1 \times \cdots \times {}^*A_n$ and for each $i,\ 1 \leq i \leq n$ it follows that ${}^*P_i(R) = P_i(R)$, where $P_i(R)$ is an internal subset of ${}^*A_i$.*

Definition 4.3.1 (Hyperfinite). An internal subset A of a set ${}^*B$ is **hyperfinite** if it is empty or there exists some $[0,p] \in BHF$ and $f \in {}^*F([0,p], {}^*B)$ such that $P_2(f) = A$.

Theorem 4.3.2. *Let $F(B)$ be the set of all finite subsets of B. Then nonempty $A \subset {}^*B$ is hyperfinite if and only if $A \in ({}^*F)(B))$, where F is considered as an operator that generates all of the finite subsets of a set.*

Theorem 4.3.3. *Any nonempty finite set of internal individuals or entities is internal and hyperfinite.*

With respect to Definition 4.3.1 the maps in each ${}^*F([0,p], {}^*B)$ are internal and behave like internal (partial) sequences. The ordinary finite manipulation we do with finite sets of real numbers can be extended to the hyperreals by means of *-transfer and by describing these processes by means of (partial) sequences defined on various $[0,p]$. As far as the NSP-world is concerned hyperfinite sets have all the same set-theoretic first-order properties as the finite sets and we do not usually establish these basic hyperfinite properties each time they are first employed. Thus, not only is the union of finitely many hyperfinite subsets of a set ${}^*B$ a hyperfinite subset of ${}^*B$ but the union of a hyperfinite collection of hyperfinite subsets of ${}^*B$ is a hyperfinite subset of ${}^*B$. *However, from the external or metamathematical point of view most hyperfinite sets are not finite as the next result indicates.*

Theorem 4.3.4. *Let $A \in \Re$ and assume that A is infinite. Then there exists a hyperfinite set F such that $F \neq A$, $F \neq {}^*A$ and $A \subset F \subset {}^*A$.*

It is precisely the concept of the hyperfinite that leads not only to a clear understanding of the processes that underlie a convergent infinite series, but also leads to the basic notion of the integral and Planck's meaning of the term "series."

Finite summation of elements of $\mathbb{R}$ can be consider a function Σ defined on $\{\mathbb{R}^{[0,n]} | n \in \mathbb{N}\}$ with values in $\mathbb{R}$ and a corresponding function for the finite summation of members of $\mathbb{R}^n$. It is customary to express the values as $\sum_{i=0}^{n} a_i$. By *-transfer if $\Gamma \in \mathbb{N}_\infty$, then the **hyperfinite sum** ${}^*\sum_{i=0}^{\Gamma} a_i \in \mathbb{R}$. By convention, the symbol ${}^*\sum_{i=0}^{\Gamma}$ is written as $\sum_{i=0}^{\Gamma}$. It is a simple matter to translate Definition 3.2.1 for sequential convergence into the following theorem for convergence of a infinite series.

Theorem 4.3.5. *An infinite series, $\sum_{i=0}^{\infty} a_i$, converges to $r \in \mathbb{R}$ if and only if for each $\Gamma \in \mathbb{N}_\infty$ it follows that $\sum_{i=0}^{\Gamma} a_i \in \mu(r)$.*

My experience indicates that Theorem 4.3.5 is not the most paramount application of the concept of hyperfinite summation for either physical or geometric modeling and, indeed, does not correspond to Planck's description for a series of infinitely small changes. In the next section, examples are discussed that substantially indicate the true character of this concept when applied to geometry or natural system behavior.

4.4. Continuity and a Few Examples.

De l'Hospital's concept of what constitutes a **curve** is not satisfactory for Robinson's theory. Recall that one acceptable analytical definition for the notion of a curve in $\mathbb{R}^n$ is the following: a curve is a continuous map $c\colon [0,1] \to \mathbb{R}^n$. This is equivalent to considering c as determined by n continuous coordinate functions $x_i = f_i(t),\ 1 \leq i \leq n$ each defined on $[0,1] \subset \mathbb{R}$. Of course, the geometric curve C determined by these functions is usually considered as the set $\{(x_1, \ldots, x_n) | t \in [0,1]\}$. The *-transform of these defining functions leads to the functions $x_i = {}^*f_i(t),\ 1 \leq i \leq n$ each defined on ${}^*[0,1] \subset {}^*\mathbb{R}$ and they generate the "hypercurve" ${}^*C \subset {}^*\mathbb{R}^n$. The analytic geometry of ${}^*\mathbb{R}^n$ is similar to the customary geometry except that it must be considered non-Archimedean in character. From the viewpoint of the geometry of the NSP-world, if the C is not linear, then ${}^*C$ is not linear and this would entail a necessary rejection de l'Hospital's infinitesimal description. As to the definition of continuity the following captures the envisioned belief that continuous functions preserve the infinitely close.

Definition 4.4.1. (Continuity). For any nonzero $n, m \in \mathbb{N}$ and any nonempty $A \subset \mathbb{R}^n$ a function $f\colon A \to \mathbb{R}^m$ is **CONTINUOUS** at $p \in A$ if ${}^*f[\mu(p) \cap {}^*A] \subset \mu(f(p))$. Also, f would be **UNIFORMLY CONTINUOUS** on A if for each $p, q \in {}^*A$ such that $p \approx q$, then ${}^*f(p) \approx {}^*f(q)$.

Observe that uniform continuity seems to preserve the infinitely close in the most satisfactory manner. One the other hand, pointwise continuity is a monad preserving property.

Theorem 4.4.1. *Definition 4.4.1 for continuity and uniform continuity is equivalent to the classical $\delta - \epsilon$ definition.*

One of the most powerful ideas in elementary analysis is that of the compact subset [or subspace if you wish] of the space $\mathbb{R}^n$. Rather than dwell upon the many equivalent standard definitions for this notion a direct nonstandard assault is very enlightening since it reveals immediately the relationship between compactness and continuity as well as an intuitive comprehension of what is being compressed or compacted.

Definition 4.4.2. (Compactness). For any nonzero $n \in \mathbb{N}$ a nonempty set $A \subset \mathbb{R}^n$ is **COMPACT** if ${}^*A \subset \bigcup\{\mu(r) | r \in A\}$.

Theorem 4.4.2. *Definition 4.4.2 for compactness is equivalent to the standard definition utilizing open covers.*

The reason that some 17'th century geometers considered non-linear curves to be collections of infinitesimal line segments was in their desire to use infinitesimal analysis to measure a curves length by corresponding this measure to the polygons of Eudoxus. Even though de l'Hospital's original description is inadequate, a modification does secure the accepted analytical results.

Example 4.4.1.A. Let $\Gamma \in \mathbb{N}_\infty$. Then $F = \{t_i | t_i = i/\Gamma \wedge 0 \leq i \leq \Gamma\}$ is an internal and hyperfinite subset of ${}^*[0,1]$. By *-transfer, F behaves like an ordered partition of the interval $[0,1]$ as

defined in the standard sense. Such a set is termed a **fine partition** (i.e. hyperfinitely many members of $[0,1]$ generating subintervals that are infinitesimal in length). The internal set F generates the internal set of "points" $P = \{({}^*f_1(t_i), \ldots, {}^*f_n(t_i)) | t_i \in F\}$ that are members of the hypercurve ${}^*C$. Now for each $i = 0, \ldots, \Gamma - 1$, and each j, $0 \leq j \leq n$ let ${}^*f_j(t_{i+1}) - {}^*f_j(t_i) = d(j,i)$. (If c is continuous, then each $d(j,i) \in \mu(0)$.) For each $i \in {}^*\mathbb{N}$ such that $0 \leq i \leq \Gamma - 1$, the internal set $\ell_i = \{(x_1, \ldots, x_n) | \forall j \in {}^*\mathbb{N},\ 0 \leq j \leq n,\ x_j = {}^*f_j(t_i) + t(d(j,i)) \wedge t \in {}^*[0,1]\}$ is a hyperline segment connecting the two points $({}^*f_1(t_i), \ldots, {}^*f_n(t_i))$, $({}^*f_1(t_{i+1}), \ldots, {}^*f_n(t_{i+1}))$ on the curve ${}^*C$. From this one obtains the internal hyperpolygonal curve $\mathcal{P}_\Gamma = \bigcup\{\ell_i | 0 \leq i \leq \Gamma - 1\}$. As to the length of $\mathcal{P}_\Gamma$ simply extend the concept of length in the classical sense by defining for each $i = 0, \ldots, \Gamma - 1$ the vector $\vec{v}_i = (d(1,i), \ldots, d(n,i)) \in {}^*\mathbb{R}^n$. Then let the hyperfinite sum $\sum_{i=0}^{\Gamma-1} \|\vec{v}_i\| = |\mathcal{P}_\Gamma| \in {}^*\mathbb{R}$. Even though, in general, you would have a different hyperpolygon with a different hyperreal length for $\forall \Gamma \in \mathbb{N}_\infty$ `Robinson [1966, 84-86]` showed that if c is continuously differentiable, then for all $\Gamma \in \mathbb{N}_\infty$, $|\mathcal{P}_\Gamma| \in \mu(r)$ and the real number r was the length of the curve obtained in the classical sense by means of the integral.

Example 4.4.1.B. Under the same criterion as stated in example 4.4.1.A that c is continuously differentiable, the length of a curve is actually closer to the limit concept then it is to the de l'Hospital description. This is seen by simply following the same process but replacing Γ with an arbitrary nonzero $n \in \mathbb{N}$. As n increases this yields an increasing sequence $|\mathcal{P}_n|$. Application of Theorem 3.5.1 implies that if there exists but one $\Gamma \in \mathbb{N}_\infty$ such that $|{}^*\mathcal{P}_\Gamma| \in \mathcal{O}$, then from example 4.4.1.A this sequences converges to $\mathtt{st}(|{}^*\mathcal{P}_\Gamma|)$ and has the same standard part for all $\Gamma \in \mathbb{N}_\infty$.

The process used to obtain the length of a curve in example 4.4.1.A shows that for most ordinary curves there are infinity many hyperpolygons that have the same standard part generated length. Thus each can be used as a NSP-world representative for the curve itself, at least as far as length is concerned. The fact that there does not, in general, exist a unique hyperpolygon is a disadvantage from the viewpoint of the founders of the infinitesimal method and forces a rejection of the de l'Hospital description. However, if in the natural world a curve is envisioned to be a path of motion produced by physical processes, then the lack of uniqueness could be a advantage. Under this interpretation, it would indicate that different and possibly interesting NSP-world ultranatural processes yield the same effect when they are restricted to the natural world.

There has arisen recently a significant application of the seeming esoteric idea that there may exist a multitude of distinct NSP-effects that yield the same natural world effects. Next is an example of how this might occur.

Example 4.4.2. (Fractals) Science has become interested in order and design as reflected in what has become known as "fractal" behavior. For this example, let $c\colon [0,1] \to \mathbb{R}^n$ be a curve; but, assume that the sequence of polygon approximations, $|\mathcal{P}_i|$, discussed in example 4.4.1.B has the property that $|\mathcal{P}_i| \to +\infty$. This is apparently one of the salient features of a fractal curve. In `Herrmann [1989]` it is shown that for any nonempty compact $K \subset \mathbb{R}$ and for any continuous $c\colon K \to \mathbb{R}^n$ there exists an internal $G\colon {}^*K \to {}^*\mathbb{R}^n$ such that G is *- differentiable of any order $m \in {}^*\mathbb{N}$, G has a well-defined hyperreal length and $\mathtt{st}(G) = c$. From the viewpoint of the NSP-world of processes and paths of motion, this G represents the same effects as does the function c except that G has an associated length concept and is ultrasmooth. Moreover, the internal object

G is somewhat less arbitrary in character than are those in example 4.4.1 since it is selectable from a specific algebra of functions. This may be significant since when it has been rigorously shown that certain physical attributes are representable by fractal curves then the theory that models such physical attributes is usually associated with some specific algebra of functions. Thus a standard fractal curve may be replaced by the standard part of an ultrasmooth curve with a well-defined length.

The idea of hyperfinite summation as representing the geometric length of a curve is, of course, closely associated with the elementary integral. Indeed, when we quoted Planck's fundamental description for physical modeling a question was asked, "What might the term 'series' mean in Planck's statement?" Apparently, what Planck meant by this term is the type of hyperfinite summation that, as seen in the next chapter, produces the integral.

Chapter 5.

STANDARD RULES FOR INTEGRAL MODELING

5.1 The Riemann Styled Integral.

The mathematical concept variously termed "integration" (i.e. to bring together the parts or to make whole) was, until after the time of Cauchy, always considered to be a specifically defined summation process. In 1823 Cauchy wrote the following description, where $f\colon [a,b] \to \mathbb{R}$ is assumed to be continuous. "...if one divides $X - x_0$ into infinitesimally small elements $x_1 - x_0, x_2 - x_1, \ldots, X - x_{n-1}$ the sum

$$S = (x_1 - x_0)f(x_0) + (x_2 - x_1)f(x_1) + \cdots + (X - x_{n-1})f(x_{n-1})$$

converges to a limit represented by the definite integral $\int_{x_0}^{X} f(x)\,dx$." From our new nonstandard point of view such a sum is produced by a hyperfinite partition [see Example 4.4.1.A] generated by some $\Gamma \in \mathbb{N}_\infty$ and the internal hyperfinite sequence of values $f(x_i),\ 0 \le i \le \Gamma - 1$. Obviously, the Riemann sum, where f may be evaluated at any member of each subinterval $[x_i, x_{i+1}]$, is styled after this Cauchy definition. However, does the standard part of such a hyperfinite sum exist and is it independent of the partition? Moreover, can the concept be extended to bounded not necessarily continuous functions?

In the following very brief discussion, neither the Stieltjes nor Lebesgue generalization is considered, even though these have been extensively investigated by nonstandard means. Indeed, research indicates that the use of arbitrary partitions along with the Darboux concept of the upper and lower sums and the upper and lower integrals which he proved to be equivalent to the Riemann integral are note worthy in that they more easily yield the rigorous proofs that establish the properties of the highly applicable converging Riemann sum notion. Since our paramount concern is modeling with the Riemann styled integral, our stated results are in terms of such easily conceived hyperfinite sums. A simple nonstandard definition for an integral of bounded functions - the H-integral - is given in `Herrmann [1985]`. In this paper, it is shown that the H-integral is equivalent to the Darboux integral. [Note: in Herrmann [1985] the Darboux integral is called the Riemann integral. Further, Theorem 3.3 and Corollaries 3.3.1, 3.3.2, 3.3.3, 3.3.4 in this paper are incorrect as stated. However, our use of the results from this paper are not related to these few erroneous conclusions.] As is well-known the Darboux integral is equivalent to the Riemann integral conceived of as approximated by the Riemann Sums. It seems expedient, however, to consider all such integral concepts extended to reasonable subsets of $\mathbb{R}^n$ using the simplest possible procedures [`Apostal [1957]`, `Spivak [1965]`].

For an n-dimensional space $n \ge 1$, the closed set $R = [a_1, b_1] \times \cdots \times [a_n, b_n],\ a_i < b_i,\ 1 \le i \le n$ is called **a rectangle.** Of course, if $n = 1$, then a "rectangle" is but a closed interval. As usual, consider for each $[a_i, b_i]$ a **partition** P_i as a finite set of members of $[a_i, b_i]$ such that $a_i, b_i \in P_i$ and where P_i is considered as ordered. This is often explicitly written as $P_i = \{x_{i0}, \ldots, x_{ik}\},\ a_i = x_{i0} < x_{i1} < \cdots < x_{ik} = b_i$. This determines the closed one-dimensional subintervals $[x_{i(p-1)}, x_{ip}],\ 1 \le p \le k$. In brief, this process obtains a partition $P = P_1 \times \cdots \times P_n$ of R and a finite collection of closed n-dimensional subrectangles R_q obtained by considering $([x_{10}, x_{11}] \cup \cdots \cup [x_{1k-1}, x_{1k}]) \times \cdots \times ([x_{n0}, x_{n1}] \cup \cdots \cup [x_{nm-1}, x_{nm}])$. Each R_q has a measure, $v(R_q) \in \mathbb{R}$, assigned to it which is intuitively the product of the lengths of the sides. For simplicity of notion the definition of the measure $v(R_q)$ is left intuitively understood.

Probably the simplest partition to consider would be the one termed a **simple** partition. These are formed by selecting n nonzero natural numbers $m_1, \ldots, m_n$ and dividing each interval $[a_i, b_i]$ into an equal length partition by adding to each successive partition point the number $(b_i - a_i)/(m_i)$. This concept is extended to the nonstandard world by selecting n infinite natural numbers $\Gamma_1, \cdots, \Gamma_n$ and generating for each interval $[a_i, b_i]$ an internal hyperfinite partition, P_i, each subinterval of which has positive infinitesimal length $(b_i - a_i)/(\Gamma_i) = dx_i$. Then the partition $P = P_1 \times \cdots \times P_n$ is a **simple fine partition** of R. Such a partition yields an internal set of hyperrectangles R_i such that ${}^*v(R_i) = dx_1 \cdots dx_n = dX \in \mu(0)^+$. You could be much more general and consider the **fine** partitions which are internal collections of hyperfinitely many members of ${}^*[a_i, b_i]$ such that the length of any subinterval is an infinitesimal. In Cauchy's definition he evaluated a function at specific endpoints of each subinterval. For a bounded function $f\colon R \to \mathbb{R}$ this evaluation concept can also be applied in the case of the subrectangles into which R is partitioned by evaluating the function f at, say, the corner nearest to the origin. However, it has become customary to be somewhat more general and include the concept of the intermediate partition. Let P be a partition of the rectangle R and assume that P determines the set of subrectangles $\{R_q | 1 \leq q \leq m\}$. An **intermediate partition**, Q, is any finite sequence of vectors $\{\vec{v}_q\}$, where $\vec{v}_q \in R_q$ for each q such that $1 \leq q \leq m$.

Definition 5.1.1. (The Integral). Let $f\colon R \to \mathbb{R}$ be bounded and $\mathcal{P}$ the set of simple partitions of R. Then f is said to be **INTEGABLE** if there exists some $r \in \mathbb{R}$ and a simple fine partition, $P \in {}^*\mathcal{P}$ such that for each of its internal intermediate partitions $\{\vec{v}_q\}$, where $1 \leq q \leq \Gamma \in \mathbb{N}_\infty$,

$$\sum_{k=1}^{\Gamma} {}^*f(\vec{v}_q)\, {}^*v(R_q) \in \mu(r).$$

Theorem 5.1.1. *A bounded function $f\colon R \to \mathbb{R}$ is integrable if and only if it is integrable in the sense of Darboux and Riemann Sums.*

Theorem 5.1.2. *If bounded $f\colon R \to \mathbb{R}$ is integrable, then there exists a unique $r \in \mathbb{R}$ such that for every fine partition $P = \{\vec{x}_0, \ldots \vec{x}_\Omega\},\ \Omega \in \mathbb{N}_\infty$ and every internal intermediate partition $Q = \{\vec{v}_q\},\ 1 \leq q \leq \Gamma \in \mathbb{N}_\infty$ it follows that*

$$\sum_{k=1}^{\Gamma} {}^*f(\vec{v}_q)\, {}^*v(R_q) \in \mu(r).$$

The unique real number that exists by Theorem 5.1.2 is, from theorem 5.1.1, the classical value of the definite integral and hence if bounded $f\colon R \to \mathbb{R}$ is integrable in the sense of Definition 5.1.1, then we may write

$$\int \cdots \int_R f(\vec{x})\, dx_1 \cdots dx_n = \int_R f(\vec{x})\, dX = \mathtt{st}(\sum_{k=1}^{\Gamma} {}^*f(\vec{v}_q)\, {}^*v(R_q)).$$

Observe that Theorem 5.1.2 indicates that Cavalieri's notion of the indivisible line segment as being the foundation for the definite integral is untenable. Even though the above hyperfinite sums correlate directly to the intuitive concept of the definite integral, even for possibly discontinuous

functions, this fact alone does not lead to the appropriate selection of specific integrands that will produce meaningful geometric or physical measures. *Indeed, one of the most significant aspects of modern infinitesimal analysis is that there does exist describable modeling procedures that rigorously establish that a specific integrand does provide the requisite value for a specific geometric or physical quantity.*

5.2 The Infinite Sum Theorems.

Throughout applied mathematics numerous linear functionals are utilized to discuss and predict geometric or physical qualities. Such functionals are essential to the indirect verification of many physical theories for within the laboratory environment it is the predicted values displayed by elaborate machinery that often yield the only indications that unobserved events may be occurring. As indicated in section 5.1, there are now rigorous rules that lead to an immediate adoption of the integral as the appropriate modeling structure when one mentally conceives of such events and applies experience to determine the geometric or physical properties that might establish that it is likely that certain hypothesized behavior is actually occurring.

Prior to Robinson's discovery certain vaguely described rules did appear in the mathematical literature; but none was consistently defined in a rigorous language nor did they have any particular relation to the successfully applied intuitive notions we term infinitesimal reasoning. As an intermediate step in establishing a consistent and complete approach to this problem, it is now possible to describe explicitly sufficient infinitesimal conditions that establish the integral as the correct modeling structure. These intermediate rules have become known as *The Infinite Sum Theorems.* For any bounded function $f\colon[a,b] \to \mathbb{R}$ a generalizations of these rules can be found in Herrmann [1985]. In appendix 5, this rule is further generalized and applied to the case that bounded $f\colon R \to \mathbb{R}$. Obviously, a **subrectangle,** R_S, is a rectangle that is a subset of R. Let $\mathcal{C}$ be the set of all simple partitions of R, $\mathcal{S}(P)$ the set of all subrectangles generated by $P \in \mathcal{C}$ and $\mathcal{C}_{PSR} = \{R_S | \exists P(P \in \mathcal{C} \wedge R_S \in \mathcal{S}(P)\}$ the set of all subrectangles contained in any simple partition of R.

Definition 5.2.1. (Simply Additive). Let $\{R_q | 1 \leq q \leq m\}$ be any simple partition of R and let B be any map defined on the collection $\mathcal{C}_{PSR}$ and having real number values. Then B is said to be **SIMPLY ADDITIVE** if for each $\{R_q | 1 \leq q \leq m\} = \mathcal{S}(P),\ P \in \mathcal{C}$ it follows that

$$B(R) = \sum_{k=1}^{m} B(R_k).$$

Obviously, ${}^*B$ is defined on the set ${}^*\mathcal{C}_{PSR}$ which contains all of the infinitesimal subrectangles contained in any simple fine partition of R. I point out that our first infinitesimal sum theorem actually holds for a slightly more general partition than a simple partition of R - the special partition. However, a simple partition is a special partition and in most applied cases the simple partition suffices.

Theorem 5.2.1. (An Infinite Sum Theorem.) *Let bounded $f\colon R \to \mathbb{R}$ and simply additive $B\colon \mathcal{C}_{PSR} \to \mathbb{R}$. If there exists a simple fine partition $\{R_q | 1 \leq q \leq \Gamma\}$ and for each R_q there exists some $\vec{p} \in R_q$ such that*

$$ {}^*B(R_q)/dX \approx {}^*f(\vec{p}), \qquad (*)$$

then f is integrable and

$$B(R) = \int_R f(\vec{x})\, dX.$$

Two observations about Theorem 5.2.1. The infinitesimal dX need not be considered the finite product of coordinate measures but may also take on the character of such physical quantities as the finite product of infinitesimal momenta, the finite product of infinitesimal probabilities and even the finite product of infinitesimal charges or infinitesimal numbers of elementary particles if such things can be conceived of in objective reality. Further, it is somewhat unfortunate for applied mathematics that the converse of Theorem 5.2.1 does not hold. N.J. `Cutland [1986]` has supplied your author with an example of a function defined on $[0, 1]$ that is Darboux integrable (hence integrable) but if you define $B(R_S) = \int_{R_S} f(x)\, dx$, then ${}^*B$ does not satisfy property (*). On the other hand, if f is continuous on R, then (*) does hold for such integrally defined functionals and, indeed, a much stronger property called *supernearness* holds as well. Let $\mathcal{C}_{SR}$ be the set of all subrectanghles contained in R. **For simplicity of notation**, **throughout this manual**, **maps such as B are denoted as being**, **at least**, **defined on sets such as $\mathcal{C}_{PSR}$, $\mathcal{C}_{SR}$ etc. Simple additivity and additivity will greatly enlarge their domains of definition.**

Definition 5.2.2. (Supernearness).
Let $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n) \in {}^*R$ and $y_i - x_i \in \mu(0)^+$, $1 \leq i \leq n$; and let $R_S = \{(z_1, \ldots, z_n) | \forall i(\, 1 \leq i \leq n \rightarrow x_i \leq z_i \leq y_i) \wedge (z_i \in {}^*\mathbb{R})\}$ denote an infinitesimal subrectangle of R. A map $B{:}\mathcal{C}_{SR} \rightarrow \mathbb{R}$ is **SUPERNEAR** to bounded $f{:}\, R \rightarrow \mathbb{R}$ if for every infinitesimal subrectangle R_S of ${}^*R$ and for every $\vec{p} \in R_S$ it follows that

$$ {}^*B(R_S)/dX \approx {}^*f(\vec{p}), \qquad (**)$$

where $dX = \prod_{i=1}^n (y_i - x_i) \in \mu(0)^+$.

Theorem 5.2.2. *A bounded function $f{:}\, R \rightarrow \mathbb{R}$ is continuous if and only if there exists a map $B{:}\mathcal{C}_{SR} \rightarrow \mathbb{R}$ that is supernear to f.*

In the proof of Theorem 5.2.2 the following interesting integral property is established.

Corollary 5.2.2. *Suppose that $f{:}\, R \rightarrow \mathbb{R}$ is continuous. For each $R_S \in \mathcal{C}_{SR}$ define $B(R_S) = \int_{R_S} f(\vec{x})\, dX$. Then B is supernear to f.*

Theorem 5.2.3. *Let bounded $f{:}\, R \rightarrow \mathbb{R}$. If $B{:}\mathcal{C}_{SR} \rightarrow \mathbb{R}$ is supernear to f and simply additive on each simple partition of each R_S, then f is continuous on R and*

$$B(R_S) = \int_{R_S} f(\vec{x})\, dX$$

for each $R_S \in \mathcal{C}_{SR}$.

In applications of the integral to geometric and physical problems it is usually assumed that the map $B{:}\mathcal{C}_{SR} \rightarrow \mathbb{R}$ is, at least, **additive** on $\mathcal{C}_{SR}$. Recall that this means that if nonempty

$\{R_i \mid 1 \leq i \leq n\} \subset \mathcal{C}_{SR}$ is pairwise disjoint or pairwise has only boundary points in common, then $B(\cup\{R_i\}) = \sum_{i=1}^{n} B(R_i)$. Note that if B is additive on $\mathcal{C}_{SR}$, then B is simply additive on each member of $\mathcal{C}$ and for each simple partition of $R_S \in \mathcal{C}_{SR}$.

Corollary 5.2.3.1 *Let bounded $f: R \to \mathbb{R}$. If $B: \mathcal{C}_{SR} \to \mathbb{R}$ is supernear to f and additive on $\mathcal{C}_{SR}$, then f is continuous on R and*

$$B(R_S) = \int_{R_S} f(\vec{x})\, dX$$

for each $R_S \in \mathcal{C}_{SR}$.

Corollary 5.2.3.2 *Let bounded $f: R \to \mathbb{R}$. There exists one and only one map $B: \mathcal{C}_{SR} \to \mathbb{R}$ that is supernear to f and either simply additive on each simple partition of each R_S or additive on $\mathcal{C}_{SR}$.*

5.3 Extensions.

In general, the bounded real valued function f need not be defined on such a convenient set as R. If $f: D \to \mathbb{R}$ is defined on a bounded set D, then the most expedient procedure to follow is to define a function $\hat{f}: R \to \mathbb{R}$, where $D \subset R$, by $\hat{f}(\vec{x}) = f(\vec{x})$ for each $\vec{x} \in D$ and $\hat{f}(\vec{x}) = 0$ for each $\vec{x} \in R - D$. With this case then, as is customary, let $\int_D f(\vec{x})\, dX = \int_R \hat{f}(\vec{x}),\ dX$.

As far as a map such as $B: \mathcal{C}_{SR} \to \mathbb{R}$ is concerned, the additivity of B may be extended to all Jordan-measurable subsets of R. The fact that B may be additive on a lesser collection of subsets of R will suffice for the basic modeling rules described in the next section. These modeling rules are very specific in character and if the proper simplistic assumptions for B are utilized, then they lead directly to the appropriate infinite sum theorem and its associated integral equivalence.

⇒ IMPORTANT ⇒

⇒ In the following applications, the standard requirements are stated in terms of what we perceive to be global behavior of well-known ordinary functionals and their relation to standard characterizing properties. **Many of these observation are not obvious. The following applications are actually intended to foster an appreciation for the nonstandard modeling rules and procedures that appear in Chapter 6 - rules that lead more directly to the appropriate conclusions.** The reason we present the following applications is that this global approach is used in the more elementary textbooks. ⇐

5.4 Applications and the Standard Modeling Rules.

One of the unusual aspects of the Infinite Sum Theorem 5.2.1 is that it does not view the functional B directly but, rather, a "mean value" must be considered if the attention is directed toward the concept of being infinitely close. As the derivations in the following elementary applications indicate this problem is submerged within the derivation itself and does not usually occur when the properties of the basic functional are proposed. Our applications are mostly geometric and elementary in character, while the major applications to the physical sciences will appear in their respect manuals. *Further, we concentrate upon those applications that traditionally appear in the customary core calculus, elementary differential equations and physical science courses.*

Application 5.4.1 *The 2-dimensional area between two continuous curves.*

First assume that all of the following functions are continuous on their indicated domains. Give $h\colon [c,d] \to \mathbb{R}$ and $k\colon [c,d] \to \mathbb{R}$, where we denote by the symbol $h \leq k$ the condition that $h(x) \leq k(x)$ for each $x \in [c,d]$. Let's look at the original idea behind an area function $A([c,d],h,k)$ that measures the intuitive area between these two curves and over the interval $[c,d]$. Define a **basic region** R for any pair of functions $h,k,\ h \leq k$ and over any interval $[c,d]$ contained in their common domain by $R = \{(x,y)|x \in [c,d] \wedge h(x) \leq y \leq k(x)\}$. Assume that we have two fixed functions $f,g,\ f \leq g$ defined on $[a,b]$. The follow rules (axioms) appear to model our intuitive notion of an area function.

(i) The area function A is, at least, defined on all basic regions determined by $[c,d] \subset [a,b]$.

(ii) For the above two fixed functions f,g defined on $[a,b]$ the area function A is, at least, additive on the set $\mathcal{C}_{SR}$ of $[a,b]$.

(iii) If D,E are two basic regions in the domain of A and $D \subset E$, then $A(D) \leq A(E)$.

(iv) If two functions h,k are constant over any $[c,d] \subset [a,b]$, then $A([c,d],h,k) = (d-c)(k-h)$.

These four properties for the area function A are certainly reasonable and seem to model the intuitive notions from elementary plane geometry. We now formally establish that

$$A([a,b],f,g) = \int_a^b (g(x) - f(x))\,dx. \tag{I_1}$$

Derivation. Let $[x_1, x_1 + dx]$ be any hyperinterval generated by any simple fine partition of ${}^*[a,b]$. Note in this case $dx = (b-a)/\Gamma$, where $\Gamma \in \mathbb{N}_\infty$. By considering the *-transfer of the standard extreme value theorem for continuous functions defined on closed intervals it follows that ${}^*f$ and ${}^*g$ attain their maximum and minimum values ${}^*f_M,\ {}^*g_M$ and ${}^*f_m,\ {}^*g_m$ respectively on $[x_1, x_1 + dx]$. It is not difficult to model statements (i), (ii), (iii) (iv) set- theoretically and extend these properties to the nonstandard world. We need only consider statements (i), (ii), (iii), (iv) as intuitively *-transformed by changing the terminology to "hyper" or "*" terminology. When this is done statements (i), (iii) and (iv) yield the result that

$$({}^*g_m - {}^*f_M)dx \leq {}^*A([x_1, x_1 + dx],\, {}^*f,\, {}^*g]) \leq ({}^*g_M - {}^*f_m)dx \Rightarrow \tag{1}$$

$$({}^*g_m - {}^*f_M) \leq ({}^*A([x_1, x_1 + dx],\, {}^*f,\, {}^*g]))/dx \leq ({}^*g_M - {}^*f_m). \tag{2}$$

Since f and g are uniformly continuous on $[a,b]$ and ${}^*f, {}^*g$ attain their respective maximum and minimum value at members of $[x_1, x_1 + dx]$ then definition 4.4.1 implies that

$${}^*f_m \approx {}^*f(x_1) \approx {}^*f_M,\ {}^*g_m \approx {}^*g(x_1) \approx {}^*g_M. \tag{3}$$

Consequently,

$$({}^*g_m - {}^*f_M) \approx (g(x_1) - f(x_1)) \approx ({}^*g_M - {}^*f_m) \Rightarrow \tag{4}$$

$$(A([x_1, x_1 + dx],\, {}^*f,\, {}^*g]))/dx \approx (g(x_1) - f(x_1)) \tag{5}$$

from the fact that f,g are bounded and Corollary 2.2.5.2. Statement (i) allows application of the Infinite Sum Theorem and integral equation I_1 is the consequence.

Once equation I_1 is obtained then it may be checked against the standard area measures for the ordinary Euclidean plane figures in order to insure that it is indeed an extension. I point out that throughout many of these applications similar modeling rules such as (i), (ii), (iii) and (iv) are essential if one wishes to achieve a formal derivation. As will be illustrated there are notable

exceptions to this general approach where one of the standard conditions (i) - (iv) may fail. However, here are two more applications where the standard functional characterizations can be formulated

Application 5.4.2. *Volume obtained by a 2-dimensional integral.*

Assume, as in the previous application, that all functions are continuous on their indicated domains. Suppose that two functions h, k are defined on a rectangle $R_0 \subset \mathbb{R}^2$ and have the property that the $h \leq k$ on R_0. Generalizing the definition in application 5.4.1, define a basic region R_B in $\mathbb{R}^3$ for each a pair h, k by $R_B = \{(x, y, z)|(x, y) \in R_0 \wedge h(x, y) \leq z \leq k(x, y)\}$. As before, we attempt to model the concept of a volume function $V(R_0, h, k)$ over any rectangle $R_0 \subset \mathbb{R}^2$. Assume that we have two real valued fixed functions $f, g,\ f \leq g$ defined on a rectangle R.

(i) The volume function V is defined, at least, for all basic regions determined by rectangles that are subsets of R.

(ii) For the above two functions f, g, the function V is, at least, additive on the set $\mathcal{C}_{SR}$ of R.

(iii) If D, E are two basic regions in the domain of V and $D \subset E$, then $V(D) \leq V(E)$.

(iv) If the functions h, k are constant over any rectangle $R_0 \subset R$, then $V(R_0, h, k) = (\text{area } R)(k - h)$.

If V satisfies these rules, then

$$V(R, f, g) = \int_R (g(\vec{x}) - f(\vec{x}))\, dX.$$

Derivation. Except for a very slight modification, this is exactly the same as the derivation for application 5.4.1. Simply let $[x_1, x_1 + dx_1] \times [x_2, x_2 + dx_2]$ be a hyperrectangle generated by a simple fine partition of ${}^*R$. In this case, $dx_1 = (b - a)/\Gamma,\ dx_2 = (d - c)/\Omega,\ \Gamma, \Omega \in \mathbb{N}_\infty$. Since R is compact the remainder of this derivation is as in application 5.4.1.

Application 5.4.3. *Mass obtained by a 3-dimensional integral.*

Assume that $\rho(\vec{x})$ is a continuous point density function defined on a rectangle $R \subset \mathbb{R}^3$. General physical experience leads to the following characterizations for the elementary mass, $M(R, \rho)$, of R.

(i) The mass is defined for, at least, the set of all subrectangles, $\mathcal{C}_{SR}$, of R.

(ii) The mass is additive on the set $\mathcal{C}_{SR}$.

(iii) If R_S is a subrectangle of R and $\rho_1(\vec{x})$ is a continuous density function defined on R_S with the property that $\rho_1(\vec{x}) \leq \rho(\vec{x})$ for each $x \in R_S$, then $M(R_S, \rho_1) \leq M(R_S, \rho)$.

(iv) If the function ρ is constant over any rectangle $R_0 \subset R$, then $M(R_0, \rho) = \rho(v(R_0))$.

If M satisfies rules (i) -(iv), then

$$M(R, \rho) = \int_R \rho(\vec{x})\, dX.$$

Derivation. As in the previous cases, one selects a simple fine partition of ${}^*R$ and lets R_q be some hyperrectangle determined by such a simple fine partition. The above characterizations are extended by *-transfer to the NSP-world. From continuity, ${}^*\rho$ attains its minimum value ${}^*\rho_m$ and maximum value ${}^*\rho_M$ at members of R_q. From (i), (iii) and (iv) it follows that

$$({}^*\rho_m)dX \leq {}^*M(R_q, {}^*\rho) \leq ({}^*\rho_M)dX \Rightarrow \tag{1}$$

$${}^*\rho_m \leq ({}^*M(R_q, {}^*\rho))/dX \leq {}^*\rho_M. \tag{2}$$

Let x_1 be any member of R_q. From the uniform continuity of ρ it follows that

$${}^*\rho_m \approx {}^*\rho(x_1) \approx {}^*\rho_M. \tag{3}$$

Consequently,

$$({}^*M(R_q, {}^*\rho))/dX \approx {}^*\rho(x_1) \tag{4}$$

and the derivation follows from the Infinite Sum Theorem.

In the above applications, the rules (i) – (v) obviously depict these functionals from a standard point of view as described completely in terms of the standard world. Except within the formal derivations, there may appear to be no consuming need for any insight into the infinitesimal NSP-world. As is be amply illustrated in the next chapter, one or more of these rules may not be self-evident when the integral is applied for both geometric and physical modeling. When this occurs, then it is often the case that certain simplistic and local aspects of the standard world are axiomatically impressed upon the pure NSP-world. This then leads to rules that do include descriptions for pure NSP-world behavior.

5.5 Extensions of the Standard Rules.

In practice, the rectangle R is too restrictive to be of much significance in applications. Fortunately, there are techniques that will allow us to remove this restriction - techniques that lead to a straightforward modification of a few of the terms that appear in such standard rules as (i) – (iv).

The appropriate alteration of these rules begins with the concepts briefly mentioned in section 5.3.

(1) Assume that $f: D \to \mathbb{R}$ is continuous on D, where D is a compact Jordan-measurable subset of $\mathbb{R}^n$. [Apostal [1957], `De Lillo [1982]`] The Jordan-measurable subsets of R include those that commonly appear throughout basic applications.

(2) Let $\mathcal{C}_{JR}$ denote the set of all Jordan-measurable subsets of R. Clearly, $\mathcal{C}_{SR} \subset \mathcal{C}_{JR}$ as is well-known.

(3) Assume that $B: \mathcal{C}_{JR} \to \mathbb{R}$.

(4) Now extend f to $\hat{f}$ and assume that $\hat{f}$ is integrable on R. It is easy to show that the value of $\int_D f(\vec{x})\, dX = \int_R \hat{f}(\vec{x})\, dX$ is independent of the choice of R.

Our next task is to see how a simple modification of the standard rules leads to a refined derivation that establishes the same integral expression. **In the following applications**, **the rule modifications are written in italics and** $v(J)$ **denotes the real Jordan content for any Jordan-measurable set** $J \subset R$.

Application 5.5.1. *Volume obtained by a 2-dimensional integral.*

Assume, as in application 5.4.2, that all functions are integrable on their indicated *compact Jordan-measurable* domains. Suppose that two functions h, k are defined on a *compact and Jordan-measurable* $J_0 \subset R \subset \mathbb{R}^2$ and have the property that the $h \leq k$ on J_0. Generalizing the definition in application 5.4.1, define a basic region J_B in $\mathbb{R}^3$ for each a pair h, k by $J_B = \{(x, y, z) | (x, y) \in J_0 \wedge h(x, y) \leq z \leq k(x, y)\}$. As before, we attempt to model the concept of a volume function $V(J_0, h, k)$ over any $J_0 \in \mathcal{C}_{JR}$. Assume that we have two real valued fixed and continuous functions $f, g,\ f \leq g$ defined on *compact and Jordan-measurable* $J \subset R$.

(i) The volume function V is defined, at least, for all basic regions determined by all *compact and Jordan-measurable sets* that are subsets of R.

(ii) For the *integrable extensions* $\hat{f}, \hat{g}$, *of* the above two functions the function V is, at least, additive on the set $\mathcal{C}_{JR}$ of R.

(iii) If D, E are two basic regions in the domain of V and $D \subset E$, then $V(D) \leq V(E)$.

(iv) If the functions h, k are constant over any $J_0 \in \mathcal{C}_{JR}$, then $V(J_0, h, k) = (v(J_0))(k - h)$. If V satisfies these rules, then

$$V(J, f, g) = \int_R (\hat{g}(\vec{x}) - \hat{f}(\vec{x}))\, dX = \int_J (g(\vec{x}) - f(\vec{x}))\, dX.$$

Derivation. Let $I = [x_1, x_1 + dx_1] \times [x_2, x_2 + dx_2]$ be a hyperrectangle generated by a simple fine partition of ${}^*R$. In this case, $dx_1 = (b - a)/\Gamma,\ dx_2 = (d - c)/\Omega,\ \Gamma, \Omega \in \mathbb{N}_\infty$. By *-transfer, we transfer the general results concerning Jordan-measurable sets to the NSP-world. Assume that ${}^*J \cap I = K \neq \emptyset$. Since $I \in {}^*\mathcal{C}_{JR}$ then $K \in {}^*\mathcal{C}_{JR}$. Further since I is *-compact then K is *-compact. Noting that ${}^*f = {}^*\hat{f},\ {}^*g = {}^*\hat{g}$ on K, then the *-extreme value theorem implies that ${}^*\hat{f},\ {}^*\hat{g}$ attain their maximum and minimum values ${}^*\hat{f}_M,,\ {}^*\hat{g}_M$ and ${}^*\hat{f}_m,\ {}^*\hat{g}_m$ respectively on K. Now by application of *-additivity and (iv) and the fact that $K, I, I - K$ are *-Jordan-measurable (if $I - K = \emptyset$, then we still let it be measurable with content equal to zero) we have that ${}^*V(I, {}^*\hat{f}, {}^*\hat{g}) = {}^*V(K, {}^*\hat{f}, {}^*\hat{g}) + {}^*V(I - K, {}^*\hat{f}, {}^*\hat{g}) = {}^*V(K, {}^*\hat{f}, {}^*\hat{g})$.

Next we also apply (iii) and obtain

$$({}^*\hat{g}_m - {}^*\hat{f}_M)\, {}^*v(K) \leq ({}^*\hat{g}_m - {}^*\hat{f}_M) dX \leq {}^*V(I, {}^*\hat{f}, {}^*\hat{g}) =$$

$${}^*V(K, {}^*\hat{f}, {}^*\hat{g}) \leq ({}^*\hat{g}_M - {}^*\hat{f}_m)\, {}^*v(K) \leq ({}^*\hat{g}_M - {}^*\hat{f}_m) dX. \tag{1}$$

Since $f = \hat{f},\ g = \hat{g}$ are uniformly continuous on K and $K \neq \emptyset$ then for $x_1 \in K$ we have that ${}^*\hat{f}_m \approx {}^*\hat{f}(x_1) \approx {}^*\hat{f}_M,\ {}^*\hat{g}_m \approx {}^*\hat{g}(x_1) \approx {}^*\hat{g}_M$. Consequently,

$$({}^*\hat{g}_m - {}^*\hat{f}_M) \approx ({}^*\hat{g}(x_1) - {}^*\hat{f}(x_1)) \approx ({}^*\hat{g}_M - {}^*\hat{f}_m). \tag{2}$$

Application of expressions (1) and (2) yields

$${}^*V(I, {}^*\hat{f}, {}^*\hat{g})/dX \approx ({}^*\hat{g}(x_1) - {}^*\hat{f}(x_1)). \tag{3}$$

For the case that $K = \emptyset$ it is obvious that ${}^*V(I, {}^*\hat{f}, {}^*\hat{g}) = 0 = ({}^*\hat{g}(x_1) - {}^*\hat{f}(x_1)) dX\, x_1 \in I$. In this case we also have that expression (3) holds. Application of the Infinite Sum Theorem completes the derivation.

The above example suffices to show how all of the previous standard modeling rules and applications can be extended to the case of the Jordan-measurable subsets and integrable functions. However, for many applications of integral modeling to geometric theories and natural system behavior such listed axioms for the behavior of the conjectured functionals are often not evident. This is particularly so for standard axioms such as (iii) and (iv). To eradicate this difficulty, a direct appeal is made to the NSP-world, either to the infinitesimal terms of the hyperfinite sum that appears in definition 5.1.1 or to statement (*) of Theorem 5.2.1.

Chapter 6.

NONSTANDARD RULES FOR INTEGRAL MODELING

6.1 Historical Examples.

In 1855, `Maxwell [1890]` presented his fluid flow analogue model for Faraday's concept of both magnetic and electric lines of force. An analysis of Maxwell's imagery relative to our present understanding of the behavior of infinitesimal quantities is very enlightening. Maxwell considers "tubes" of moving points (not particles) of fluid and their paths of motion as a pure imaginary picture of what one might conceive of as line a force. Obviously, the idea was not to consider the concept of "force" as an independent entity but, rather, to first picture "something" - the points of fluid material and their paths of motion - as representing the effects of unknown forces. "*The direction of motion of the fluid will in general be different at different points of the space which it occupies, but since the direction is determined for every such point, we may conceive a line to begin at any point and to continue so that every element of the line indicates by it direction the direction of motion at that point in space. Lines drawn in such a manner that their direction always indicates the direction of fluid motion are called lines of fluid motion.*" [Maxwell, 1890:160]

Maxwell then imagines a closed curve on a surface - a surface that "cuts" the lines of fluid motion - and the lines of fluid motion that intersect this surface curve. These curve generated flow lines then produce, in his mind, a tubular surface which he calls *a tube of fluid motion.* He then fills the interior of these tubular surface with the flow lines that intersect that portion of the surface which would have the curve as its boundary. He also assumes that the fluid is incompressible. This yields another postulated property of these flow lines. "*The quantity of fluid which in a unit of time crosses any fixed section of the tube is the same at whatever part of the tube the section is taken....and no part runs through the sides of the tube, therefore the quantity which escapes from the second section is equal to that which enters through the first.*" [Maxwell 1890:161]

Maxwell then supplies a paramount nonstandard rule to the methods of infinitesimal model - a rule that has recently been called the concept of the infinitesimal microscope. "*An infinite number of lines would have to be drawn at indefinitely small intervals; but since the description of such a system of lines would involve continual reference to the theory of limits, it has been thought better to suppose the lines drawn at intervals depending on the assumed unit, and afterwards to assume the unit as small of we please by taking a small submultiple of the standard unit.*" [Maxwell 1890:161] Notice that Maxwell's statement about the necessity of limit theory is now known to be false. He may have made such a statement, as did Kepler before him, to placate those who might be more attuned to rigorous derivations. Further, following general scientific practices, Maxwell does not establish his limit theory conclusions but, rather, ascribes to an infinitesimalizing approach he claims is equivalent to the physical limit theory. More importantly we have our first vague nonstandard rule for infinitesimal modeling

---

VR1.

These infinitesimalizing ideas are equivalent to a type of infinite magnification of a infinitely small portion of the fluid - a magnification that yields finitely many lines in our field of view that appear to be drawn at real finite distances apart.

I mention that such geometric notions as expressed in VR1 can indeed be formalized within the geometric theory of ${}^*\mathbb{R}^3$.

Within Maxwell's research reports he states numerous times that the reasons for his derivations are "evident." Thus, he often gives no specific causes for his logical conclusions and leaves them axiomatic in character. It is, however, the vague methods of infinitesimalizing that continue to interests us - methods that are also often assumed to be "evident" to Maxwell's audience. Intuitively, as it will be illustrated, physical and geometric infinitesimal integral modeling also displays yet another vague nonstandard rule.

VR2.

Infinitesimal integral modeling often makes a direct appeal to the Infinite Sum Theorems as well as a simplified interpretation of the equation (Theorem 5.1.2)

$$\int \cdots \int_R f(\vec{x})\, dx_1 \cdots dx_n = \int_R f(\vec{x})\, dX = \mathtt{st}(\sum_{k=1}^{\Gamma} {}^*f(\vec{v}_q)\, {}^*v(R_q)).$$

In the infinite sum theorems the equivalence relation $\approx$ is replaced by an equality while the standard part operator is ignored and the integral is made equal to some type of summation process which is often conceived of as finite in character.

How does Maxwell apply his fluid motion analogy to assumed continuously varying magnetic properties? Surprising, he views them as discrete and constant with relation to his magnified portion the of fluid material. "*The quantity of magnetism in any section of a magnetic body is measured by the number of lines of magnetic force that pass through it.*" Maxwell [1890:182] This "number" is assumed to be a standard natural number. "*If i be the quantity of the magnetization at any point, or the number of lines of force passing through unit of area in the section of the solenoid, then the total quantity of magnetization in the circuit is the number of lines which pass through a section, $I = \sum i\, dy\, dz$, where $dy\, dz$ is the element of the section, and the summation is performed over the whole section.*" [Maxwell 1890:183] In the magnified view, the section is to be conceived of as a rectangle with actual real number area that is then made arbitrary small by Maxwell's small unit convention. The constant numbers i are not assumed to be altered as a physical quantity by the small unit convention but are fixed constants. The cardinality (i.e. intuitively the number of terms in the summation) is ignored and this "summation" is equated to the integral value I.

Maxwell appears to have arrived at his conclusions by considering the simplest of laboratory experiences for assumed constant quantities and geometric configurations. He has then impressed these experiences upon the infinitesimal nonstandard world. By assuming that the outcome is somehow additive in character this leads directly to the integral model.

With respect to pure geometric nonstandard modeling the same general process, with certain exceptions, is also applied. It was seen in Example 4.4.1.A on page 33 that the length of an n-dimensional curve is viewed globally as approximated by a hyperpolygonal curve and following this the length of the curve is defined to be the hyperfinite sum $\sum_{i=0}^{\Gamma-1} \|\vec{v}_i\| = |\mathcal{P}_\Gamma| \in {}^*\mathbb{R}$. However, for curves that are continuously differentiable Robinson formally showed that each term of this

summation could be replaced by the term

$$\sqrt{\sum_{j=1}^{n} {}^*f_j'(t_i)^2\, dt}$$

where this replacement term can be conceived of as the infinitesimal length of an infinitesimal line segment infinitely close (in the sense of the infinitesimal sum theorem) to the original line segments that comprise the hyperpolygonal curve. **A proof of this can be found in appendix 6.** On the other hand, one could proceed backwards, as some geometers have done, and discuss the vague notion of the "element of length," ds, while forcing the length of a curve to be the "sum" of such elements ds. Indeed, one of the greatest of all infinitesimal geometers does all of his analysis in terms of the still vague "element" concept and uses, what would be today, totally unacceptable non-rigorous derivations. "*Let us now examine the integral curvature of this triangle, which is equal to* $\int k\, d\sigma$, $d\sigma$ *denoting a surface element of the triangle. Wherefore, since this element is expressed by* $mdp \cdot dq$, *we must extend the integral* $\int\int mdp \cdot dq$ *over the whole surface of the triangle.*"[`Gauss 1827; Art 20`] Gauss in his papers also states that various quantities are equal when in reality they are but infinitely close. Fortunately, the algebraic manipulations of infinitesimal quantities as employed by Gauss and Maxwell were restricted to their ring properties. Unfortunately, Gauss' derivations are highly non-rigorous in character.

By comparing Robinson's techniques with many of significant mathematical models developed over the past few centuries by application of intuitive infinitesimal modeling, it is now possible to gain a little better insight into the methods used, to make these methods somewhat more acceptable in rigor and to improve considerably upon the notions outlined in the vague nonstandard rules VR1 and VR2.

6.2 The Monadic Environment.

Vague rule 1 can now be made essentially rigorous in character. When we model the most elementary geometric or simplistic physical behavior it is usually conceived of within bounded portions of $\mathbb{R}^n$. Even though one may later remove the boundedness concept, such modeling often begins with such a restriction. Furthermore, this holds true whether or not one is concerned with real or complex variables. What is conceived of as "elementary geometric" or "simplistic physical" behavior is most often fostered by individual experiences and a personal development of discipline intuition. Simplistic diagrams and sketching develop an intuition for geometric concepts, while basic laboratory experimentation tends to yield to the conscientious investigator an intuitive understanding of basic natural system behavior. It is this phenomenological approach to simplistic physical behavior that has led to the development of most of our present day intricate mathematical models that, even though they may tend to predict observed behavior, may not correspond in their entirety to physical reality.

IR1

Elementary geometric or simplistic physical behavior takes place within an m-dimensional monad, $\mu(\vec{p}) \subset {}^*\mathbb{R}^m$, where standard $\vec{p} \in \mathbb{R}^m$. Such behavior may be intuitively *-transferred to similar behavior within $\mu(\vec{p})$.

It might be argued that IR1 is too restrictive for many of our modern applied mathematical structures. It is interesting to note that many monadic properties generalize not only to general topological spaces (see many of papers published by Robert A. Herrmann from 1975 - 1984) but even to the more general pre and pseudo topological spaces [Herrmann 1980]. Can we identify for ${}^*\mathbb{R}^m$ what, at least partially, constitutes basic simplistic behavior? The first step in this process is to study in $\mathbb{R}^n$ or the laboratory geometric or physical behavior restricted to entities termed "elements." Further, one is often only concerned with specific functionals associated with such objects. As previous mentioned, in applied infinitesimal analysis, infinitesimals that are denoted by such symbols as dx need not correspond to geometric concept of length. Indeed, they can be negative or correspond to such notions as velocity, acceleration or even infinitesimal "numbers of molecules." Nevertheless, it is convenient to regard the basic elements as geometric in character and not to be concerned with their specific functional or physical interpretation.

IR2

Let internal $L\colon {}^*\mathbb{R}^n \to {}^*\mathbb{R}^m$ be a designated linear transformation and $S \subset {}^*\mathbb{R}^n$ an infinitesimal subrectangle of ${}^*R$, where rectangle $R \subset \mathbb{R}^n$. **An m-dimensional element is the configuration** $L[S]$. The **basic element** is the infinitesimal subrectangle itself.

For all such designated L considered in IR2 the elements can be characterized as m-dimensional parallelepipeds, where the 1-dimensional parallelepiped is to be conceived of as a line segment. Further, we have the following little theorem that shows the relationship between infinitesimal subrectangles and monads.

Theorem 6.2.1. *Let A be a compact subset of $\mathbb{R}^n$ and the infinitesimal subrectangle $R \subset {}^*A$. Then there exists some $p \in A$ such that $R \subset \mu(p)$.*

⇓ IMPORTANT ⇓

⇒ The hypotheses of many of the following rules and theorems are stated in terms of a Jordan-measurable set J. Such premises may be weakened by restricting the functional to $\mathcal{C}_{SR}$ and assuming that J is but a subrectangle. **I mention the important but well-known fact that all of the usual geometric configurations utilized in the basic calculus and differential equation undergraduate courses are Jordan-measurable. Further, the elementary prototype used in place of the general Jordan-measurable set is the subrectangle itself. Most experimental investigations do not go beyond the subrectangle. After the hypotheses that appear in the following rules and theorems are restricted to subrectangles and their conclusions are verified, then their consequences are often extended to J without further confirmation.** ⇑

In order to use the infinite sum theorem directly for the **basic elementary integral** it is assumed that the internal linear map L is the extended identity may ${}^*I$. Then the following are

gleaned from the basic definitions or the intuitive methods of the geometer as well as from simple experiments on natural system behavior whether it be objectively real or imaginary.

IR3 – Infinitesimal Max. and Min. Rule

(1) We wish to measure a quantity M for a compact Jordan-measurable set $J \subset R \subset \mathbb{R}^n$, where M is defined on and, at least, additive over members of the set $\{\mathcal{C}_{SR}, R - J, J\}$. Further, if subrectangle $S \subset R - J$, $M(S) = 0$, and $M(R - J) = 0$. Let $v(J)$ denote the Jordan content.

(2) There is a generating function $f(\vec{x})$ that is related to the functional M in the following manner:

(i) The functions f is continuous on J.

(ii) Let P be some simple fine partition, $S \in {}^*\mathcal{S}(P)$ and $K = {}^*J \cap S \neq \emptyset$. Then there exist $\vec{x}_m \in K$ and $\vec{x}_M \in K$ such that ${}^*f_m = {}^*f(\vec{x}_m) = {}^*\inf\{{}^*f(\vec{x})|\vec{x} \in K\} = \inf\{{}^*f(\vec{x})|\vec{x} \in K\}$ and ${}^*f_M = {}^*f(\vec{x}_M) = {}^*\sup\{{}^*f(\vec{x})|\vec{x} \in K\} = \sup\{{}^*f(\vec{x})|\vec{x} \in K\}$ and

(iii) $({}^*f_m)\ {}^*v(S) \leq {}^*M(S) \leq ({}^*f_M)\ {}^*v(S)$. [Note: this is the case where ${}^*L$ is the identity map.]

Obviously the rules in IR3 are closely related to those that appear in chapter 5, section 5.4. Except for 2(iii) these rules are usually tacitly assumed. It is useful to repeat the observation made in section 5.4. *Certain aspects of rule IR3 could be relaxed if statement 2(iii) is formulated in terms of an infinitesimal "mean value" property for the functional M. However, it is the customary linear form in 2(iii) that appears throughout all of the traditional definitions - a form that we are trained to seek and experimentally justify.* Notice that the terms of a Riemann sum are represented in the inequality in 2(iii) by either ${}^*f_m$ or ${}^*f_M$. The following theorem is almost obvious.

Theorem 6.2.2. *If IR3 holds, then*

$$M(J) = \int_J f(\vec{x})\, dX.$$

Theorem 6.2.3. *Let compact Jordan-measurable $J \subset R \subset \mathbb{R}^n$. If continuous $f\colon J \to \mathbb{R}$, then for any partition P of R and any $S \in \mathcal{S}(P)$, where $K = J \cap S \neq \emptyset$ there exist $\vec{x}_m \in K$ and $\vec{x}_M \in K$ such that $f_m = f(\vec{x}_m) = \inf\{f(\vec{x})|\vec{x} \in K\}$ and $f_M = f(\vec{x}_M) = \sup\{f(\vec{x})|\vec{x} \in K\}$.*

The title of this chapter is actually somewhat misleading. Even though all of the previous rules relative to the monadic environment are stated in terms of this monadic world, generally for the integral, individuals still rely upon standard world observations. It is the generation of differential equation models that utilize this infinitesimal locale exclusively. What mental or experimental procedures does a researcher employ in order to arrive at the conclusion expressed in IR3 part 2(iii)? **Evidently**, **these standard procedures must be closely related to the rule IR1 - IR3 even if they are expressed in the language of a standard mathematical structure.** In derivations written prior to 1981, we read that such conclusions are "self-evident" or "obvious." These self-evident features of informal infinitesimal modeling are, of course, some type of unmentioned infinitesimal reasoning process. Formally, the unmentioned infinitesimalizing procedure can now be characterized as formal *-transfer and what needs to be determined in the large scale or

macroscopic environment prior to such infinitesimalizing is contained in the premises of the next somewhat obvious proposition. We call a set, $\mathcal{P}$, of simple partitions of R **acceptable** if there exists a simple fine partition $P \in {}^*\mathcal{P}$.

Theorem 6.2.4. (Self-evident Max. and Min.) *Let the rectangle $R \subset \mathbb{R}^n$ and suppose that compact Jordan-measurable $J \subset R$. Let M be defined as in (1) of IR3, continuous $f\colon J \to \mathbb{R}$, $\mathcal{P}$ an acceptable set of partitions of R and any $P \in \mathcal{P}$. If for any $S \in \mathcal{S}(P)$ such that $J \cap S \neq \emptyset$ it follows that $(f_m)v(S) \leq M(S) \leq (f_M)v(S)$, then the infinitesimalizing process IR3 holds.*

In order to better apply the self-evident theorem an intuitive discussion of its content is in order. The values f_m and f_M are values for the original defining function restricted to $J \cap S$. Thus essentially for the boundary type subrectangles S (i.e. $S \not\subset J$ but $S \cap J \neq \emptyset$) these values have been extended to the entire rectangle S. What has been done should be viewed as an application of these bounding values of f to the extended configurations $\bigcup\{S | S \in \mathcal{S}(P) \wedge J \cap S \neq \emptyset\}$. [**Important. See the simplest rules IR5, IR6 in Appendix 6 where a continuity concept eliminates this difficulty.**]

6.3 Simple Applications.

Assuming in the macroscopic world a continuous distribution of matter throughout a space region Synge and Griffith define the moment of inertia as $I = \int r^2\, dm$ and then state that "dm is the mass of an infinitesimal element..." [`Synge and Griffith [1959:173]`]. For a uniform rod of total mass m and length $2a$ positioned on the x-axis, they state that $dm = (m\, dx)/(2a)$. It is often the case that this concept is extended to the case the mass is not uniform but is rather determined by the continuous point density function $\rho(x)$. In this particular case it is claimed that $dm = \rho(x)\, dx$. However, from the corrected notions of infinitesimal modeling this last statement is simply incorrect. Based upon Corollaries 5.2.3.1 and 5.2.3.2 and the definition of supernearness we can only be assured that for an infinitesimal subrectangle R and for any $x \in R$, $dm = {}^*M(R, {}^*\rho)$ is infinitely close to ${}^*\rho(x)\, dx$. Indeed, the notion of "infinitely close" in this context is not sufficient for a proper understanding of the relationship between dm and ${}^*\rho(x)\, dx$. As well be discussed later these two objects must, for a given dx, be "closer" then indicated by the general infinitely close concept. It is, therefore, necessary to establish various elementary applications of these infinitesimal rules in a manner distinctly different from the customary ones. I envision that many of our present day texts that claim to teach the infinitesimal modeling of natural system behavior will need to be completely revised if rigor is to be incorporated.

Application 6.3.1. *The quantity of an incompressible fluid passing through a 2-dimensional rectangle* [*resp. Jordan-measurable plane region J.*]

Suppose that we have for a macroscopic environment a function $V(x, y)$ that represents the point velocity, in a normal direction, of the incompressible fluid passing through a 2-dimensional rectangle R. Then the amount of fluid passing through R per unit time, $Q(R)$, is

$$\int_R v(x, y)\, dX.$$

Derivation. Before we start this derivation observe that it must be considered slightly less rigorous than those that appear in chapter 5. What does experience indicates about such a value $Q(R)$? First, since the fluid is incompressible then Q is additive on $\mathcal{C}_{SR}$. Observation also indicates

that for a subrectangle R_S of any simple partition $(V_m)\, v(R_S) \leq Q(R_S) \leq (V_M)\, v(R_S)$. If you agree to this experiential argument, then by Theorems 6.2.2 and 6.2.3 the result follows.

Application 6.3.2. *Moment of inertia of a circular disc of radius a about a line ℓ through the center perpendicular to the plane of the disc.*

[Point mass method.] Consider the disc J centered at the origin of our 2-dimensional Cartesian coordinate system. Let $\rho(x, y)$ represent a continuous point density function for J. The moment of inertia, I_1, for a finite system of point masses, m_i, located on the disc each at a distance of r_i from ℓ is $I_1 = \sum r_i^2\, m_i$. This may be rewritten as $I_1 = \sum (x_i^2 + y_i^2)\, m_i$. The first requirement for application of IR3 is that this idea be extended to a continuous density function, $\rho(x, y)$, of point masses. This implies that we consider the continuous moment of inertia generating function $f(x, y) = \rho(x, y)\, (x^2 + y^2)$. Using this assumed function it follows that

$$Q(J) = \int_J f(x, y)\, dX.$$

If ρ is a constant, then

$$Q(J) = \int_J \rho(x, y)\, (x^2 + y^2)\, dX = \int_0^a 2\pi\rho r^3\, dr = (\pi\rho a^4)/2) = (m/2)a^2.$$

Derivation. It's clear that we have selected the basic aspects of this application to force it to have the properties expressed in IR3. However, what experiences do we have with moments of inertia that will allow us to conclude that the inequality $(f_m)\, v(S) \leq I(S) \leq (f_M)\, v(S)$ holds? If you have such experiences, then the result is immediate. If you don't or you cannot adequately explain you intuition, then another derivation method would be required.

Application 6.3.2 and the like seem to have a very weak derivations. There are two notations that will aid in eradicating these derivation difficulties. Indeed Maxwell explicitly states one of these procedures and there is a statement within the premises of application 6.3.2 that is significant and often appears when extensions are considered. **This leads to two often used, significant and powerful infinitesimal modeling procedures. The first is the extension of constants values and the second for, point definable quantities, is the extension of the finite to the hyperfinite.**

6.4. The Method of Constants.

The method of constants makes a direct appeal to expression (*) of the Infinite Sum Theorem (5.2.1) as it appears on page 37. Individuals observe simple properties about a functional relative to the assumption that the generating function can have constant values. These observations and a simplified physical theory are then transferred to the NSP-world. They then assume that these simple properties hold for the NSP- world and then without justification state that for a continuous generating function the value of the functional is but a "sum " of the appropriate (infinitesimal) quantities. A rigorous procedure is now possible.

IR4– Method of Constants

(1) In what follows, let for any $A \subset \mathbb{R}^n$ "int" denote the interior of A. Let $\mathcal{A} = \{\mathcal{C}_{SR}, \{\text{int}(J \cap S) \neq \emptyset \mid S \in \mathcal{C}_{SR}\}\}$. We wish to measure a quantity M for a Jordan-measurable set $J \subset R \subset \mathbb{R}^n$, where M is, at least, defined on and additive over the members of the set $\{\mathcal{A}, R - J, J\}$ and for a

subrectangle $S,\ S \subset R - \text{int}(J)$ it follows that $M(S) = 0$ and $M(R - J) = 0$. Let $v(J)$ denote the Jordan content.

(2) There is a generating function $f(\vec{x})$ that is related to the functional M in the following manner:

(i) The function f is bounded on J.

(ii) Let P be any arbitrary simple fine partition, $S \in {}^*\mathcal{S}(P)$ an arbitrary infinitesimal subrectangle and $\emptyset \neq K = {}^*\text{int}({}^*J \cap S)$.

(iii) There exists some $\vec{x} \in K$ such that ${}^*M(S) = {}^*f(\vec{x})\,{}^*v(S)$ or ${}^*M(S)/{}^*v(S) \approx {}^*f(\vec{x})$.

Please note carefully where IR4 differs from IR3. First, f is only assumed to be bounded and J need not be compact. Also note that for $\emptyset \neq K,\ {}^*v(K) \neq 0$ since in the standard case nonempty and Jordan-measurable $\text{int}(J \cap S)$ contains a rectangle.

Theorem 6.4.1. *If IR4 holds, then*

$$M(J) = \int_J f(\vec{x})\, dX.$$

Once again we need some sort of infinitesimal reasoning process that leads to application of IR4. Within the laboratory or an imaginary mind experiment individuals often tacitly observe that the hypotheses of the next "self-evident" theorem hold and, without stating it, take an intuitive *- transform and obtain IR4. It is IR4, or something akin to it, that appears in the literature. However, certain aspects of the next result are necessary due to technical difficulties and one useful requirement is some what unexpected.

Theorem 6.4.2. (Self-evident Method of Constants) *Let the rectangle $R \subset \mathbb{R}^n$ and suppose that Jordan-measurable $J \subset R$. Let M be defined as in (1) of IR4, continuous $f\colon R \to \mathbb{R}$, $\mathcal{P}$ an acceptable set of partitions of R and any $P \in \mathcal{P}$. If for any $S \in \mathcal{S}(P)$ such that $\emptyset \neq K = \text{int}(J \cap S)$ there exists some $\vec{x} \in K$ and some $\vec{y} \in S$ such that (i) $M(K) = f(\vec{x})\, v(K)$ and (ii) $M(S) = f(\vec{y})\, v(S)$, then the infinitesimalizing process IR4 holds for f restricted to J.*

As with IR3 the intuitive notion behind infinitesimal reasoning Theorem 6.4.2 is the extension of the generating function values to the boundary type subrectangles S. Both of our self-evident theorems display an interesting phenomenon. The rules IR3 and IR4 require the existence of but one simple fine partition for application; but, the self-evident theorems need an infinite collection of such partitions with the requisite properties in order to guarantee that such a partition exists in the NSP-world. In practice the requirement that f be continuous on R may be relaxed to piecewise continuity on a set of J covering subrectangles of R that at most overlap only on their boundaries. **Of interest is the necessity for premise (ii). In applications it is simply assumed from observation that there is a partition of subrectangles**, S, **that are "small enough" in size that there exists such a constant** $f(\vec{x}),\ \vec{x} \in \text{int}(J \cap S)$ **such that (i) holds. The same methodology should also indicate that there exists a constant** $f(\vec{y}),\ \vec{y} \in S$ **such that (ii) holds. However**, **the (ii) premise is never mentioned as a requirement.**

Application 6.4.1. (Application 6.3.1 revisited.) *The quantity of an incompressible fluid passing through a 2-dimensional rectangle* [*resp. Jordan-measurable plane region J.*]

Suppose that we have for a macroscopic environment a continuous function $V(x, y)$ that represents the point velocity, in a normal direction, of the incompressible fluid passing through a 2-dimensional rectangle R. Let J be any Jordan- measurable subset of R. Then the amount of fluid passing through J per unit time, $Q(J)$, is

$$\int_J v(x, y)\, dX.$$

Derivation. It appears that the method of constants yields a more satisfactory derivation of this integral formula. Consider any simple partition P and any $S \in \mathcal{S}(P)$. Let $\emptyset \neq K = \text{int}(J \cap S)$. Then experience indicates that there is some $\vec{x} \in K$ such that $Q(K) = V(\vec{x})\, v(K)$. Indeed, this can be most easily argued by the intermediate value theorem if J is connected. For the same reason there is some $\vec{y} \in S$ such that $Q(S) = V(\vec{y})\, v(S)$. If you agree to these seeming innocuous statements relating scalar constant velocity normal to J, then the result follows from the Method of Constants.

Originally the basic infinitesimal reasoning behind Theorem 6.4.2 was simply that in the NSP-world ${}^*M(K) \approx {}^*M(S)$, where $\emptyset \neq {}^*\text{int}({}^*J \cap S)$ and S is a infinitesimal subrectangle of ${}^*R$. However, technically, there seems to be no way to avoid that fact that this idea is incorrect. It may be a good starting point but, infinitesimally, it is necessary that ${}^*M(K)/{}^*v(K) \approx {}^*M(S)/{}^*v(S)$.

Theorem 6.4.3. (Extended Self-evident Method of Constants) *Let the rectangle $R \subset \mathbb{R}^n$ and suppose that Jordan-measurable $J \subset R$. Let M be defined as in (1) of IR4, continuous $f\colon R \to \mathbb{R}$, continuous $g\colon R \to \mathbb{R}$, $\mathcal{P}$ an acceptable set of partitions of R and any $P \in \mathcal{P}$. If for any $S \in \mathcal{S}(P)$ and $\emptyset \neq K = \text{int}(J \cap S)$ there exists some $\vec{x}_1,\ \vec{x}_2 \in K$ and some $\vec{y}_1,\ \vec{y}_2 \in S$ such that (i) $M(K) = f(\vec{x}_1)\, g(\vec{x}_2)\, v(K)$ and (ii) $M(S) = f(\vec{y}_1)\, g(\vec{y}_2)\, v(S)$, then the infinitesimalizing process IR4 holds for fg restricted to J.*

Why do we need the Extended Self-evident theorem? First, many physical and geometric quantities are defined as the product of other previously defined generating functions. Moreover, it is often the case that these quantities are actually defined for the sole purpose of applying the differential and integral calculus. One example of this should suffice. Consider the following supposedly non-calculus approach to the moment of inertia of a circular disc of radius a about a line ℓ through the center perpendicular to the plane of the disc. In the book *Mechanics of Engineers* `[Morley [1942]]` the following approach is used.

Assume that the density is unity. . . , consider the disc divided into n "ring-shaped strips such as PQ, each of width a/n." Morley then takes the distance of the pth strip from the center O as $p \times a/n$. He then lets the area be the same as that of the rectangle $2\pi \times pa/n \times a/n = 2\pi p\, a^2/n^2$. He then states that the moment of inertia about ℓ would be this area times the distance of the outer edge $p \times a/n$ which yields

$$2\pi a^4 p^3/n^4. \tag{1}$$

If now we add these n strips this would yield, after simplification, the expression

$$\big(\pi a^4/2\big)\big(1 + \frac{2}{n} + \frac{1}{n^2}\big). \tag{2}$$

However, on the other hand, if we choose PO to be the distance of the strip from the line ℓ then assuming the everything else holds the expression for the sum all the n strips would be

$$\big(\pi a^4/2\big)\big(1-\frac{2}{n}+\frac{1}{n^2}\big). \tag{3}$$

Observe that for either (2) or (3) the limit as $n \to \infty$ is $\pi a^4/2$. Morley's definition requires that such a limit be taken. If these ideas are now applied to parallel lines interior to the strips, then the same conclusion would follow. I discuss aspects of this example more fully at the beginning of the next section.

Remark 6.4.1 Closer examination of the Self-evident Theorems 6.4.2 is useful. For a particular linear function M defined as in IR4, the relation between the values $M(S) = 0$, where the subrectangle $S \subset R - \text{int}(J)$, is of a special nature. These values, in general, are assumed to be unaffected by the function f. It is as if the function f has degenerated to the zero function when it is observed by subrectangles exterior to $\text{int}(J)$. In most practical cases this is exactly how it should be. Since we are not interested in the behavior of f on such exterior objects. Two things often happen in practice. The function $f\colon J \to (a,b) \subset \mathbb{R}$ is only assumed to be continuous on J. It does not matter whether or not f is the restriction of a function that is continuous on R. The other case is that the function f is explicitly given and it is indeed continuous on some bounding set R. If J is compact and $f\colon J \to (a,b)$ is continuous and not explicitly expressed, M is defined as in IR4 and the remaining hypotheses hold for any continuous extension $F\colon R \to (a,b)$ of f to R, then by the Teitze Extension Theorem $M(J) = \int_J f(\vec{x})\, dX$. The same modifications could be made, if necessary, to the Extended Self-evident Theorem 6.4.3.

Application 6.4.2.(Application 6.3.2 revisited.) *Moment of inertia of a circular disc of radius a about a line ℓ through the center perpendicular to the plane of the disc.*

Consider the disc J centered at the origin of our 2-dimensional Cartesian coordinate system. Let $\rho(x,y)$ represent a continuous point density function for J which is assumed to be continuous on some $R \subset \mathbb{R}^2$ such that $J \subset R$. Then the moment of inertia, $Q(J)$, of J about ℓ a line perpendicular to the disc at its center is

$$Q(J) = \int_J f(x,y)\, dX,$$

where $f(x,y) = \rho(x,y)\,(x^2+y^2)$. Using this assumed function it follows that if ρ is a constant, then

$$Q(J) = \int_J \rho(x,y)\,(x^2+y^2)\, dX = \int_0^a 2\pi\rho r^3\, dr = (\pi\rho a^4)/2) = (m/2)a^2.$$

Derivation. Kinetic energy experimentation with a flywheel with movable weights attached to its surface indicates that for actually physical regions J, there does exist a "small enough" simple partition rectangle S such that there are two points $\vec{x},\ \vec{y} = (y_1, y_2) \in \text{int}(J \cap S) \neq \emptyset$ and $Q(J \cap S) = \rho(\vec{x})\,(y_1^2 + y_2^2)$. For this particular subrectangle S it is also observed that their exist two such points in S with the same property for $Q(S)$. The same conclusion apparently would hold for all other simple partition rectangles with a "smaller diagonal length." [This generalization to "all" such rectangles appears reasonable.] Let $(m_1, m_2) \in \mathbb{N}^2$ be the generator of the simple partition P such that $S \in \mathcal{S}(P)$. Let the set of partitions $\mathcal{P}$ generated by the pair $\{(x,y) \mid (x,y) \in \mathbb{N}^2 \wedge (x \geq m_1) \wedge (y \geq m_2)\}$. Then $\mathcal{P}$ is an acceptable partition. Letting $f(x,y) = (x^2+y^2)$, which is continuous on R, then the result follows from the Method of Constant Theorem 6.4.3.

6.5. The Hyperfinite Method.

Except when the operational approach to physical quantities is used, most authors when discussing the properties of rigid body motion immediately force upon the reader the imaginary notion of the point "particle." Synge and Griffith, in the text mentioned, use exclusively this technique. Also, even though it may not be apparent from his example in the previous section, Morley motivates all of his derivations similarly. It is now possible to establish in a somewhat rigorous fashion that the technique of particles is adequate.

As a prototype, we again concentration upon a nonempty plane Jordan-measurable region J. Letting ρ be an appropriate density function and d and appropriate distance function, from line ℓ of rotation one establishes that the moment of inertia is

$$I(J) = \int_J \rho(\vec{x})\, d^2(\vec{x})\, dX. \tag{1}$$

By *-transfer of the hypotheses of Theorem 6.4.3 it follows that for a simple fine partition $P = \{S_1, \ldots, s_\Gamma\}$ of ${}^*R$ and two hyperfinite sequences, $Q_1' = \{\vec{x}_1, \ldots, \vec{x}_\Gamma\},\ Q_2' = \{\vec{y}_1, \ldots, \vec{y}_\Gamma\},\ \vec{x}_i, \vec{y}_i$ members of $\emptyset \neq K_i = {}^*\mathrm{int}({}^*J \cap S_i)$ such that ${}^*I(K_i) = {}^*\rho(\vec{x}_i)\ {}^*d(\vec{y}_i)\, {}^*v(K_i)$. Observe that properties of Jordan-measurable sets and the measure v imply that $\mathrm{int}({}^*J \cap S_i)$ is *-Jordan-measurable and that ${}^*v(K_i) = {}^*v({}^*J \cap S_i)$; which leads to ${}^*I({}^*J \cap S_i) = {}^*\rho(\vec{x}_i)\ {}^*d(\vec{y}_i)\, {}^*v(K_i) = {}^*\rho(\vec{x}_i)\ {}^*d(\vec{y}_i)\, {}^*v({}^*J \cap S_i) = {}^*I(J \cap S_i)$. There is a nonempty internal $\mathcal{A} \subset {}^*\mathbb{N}$ such that $\mathcal{A} = \{n \mid n \in {}^*\mathbb{N} \wedge 1 \leq n \leq \Gamma \wedge {}^*J \cap S_n \neq \emptyset\}$. From the additivity of I it follows that

$${}^*I(J) = \sum_{j \in \mathcal{A}}{}^*\rho(\vec{x}_j)\ {}^*d(\vec{y}_j)\, {}^*v({}^*J \cap S_j). \tag{2}$$

In the standard case for every $P \in \mathcal{P}$ and $\{S_1, \ldots, S_n\} = \mathcal{S}(P)$ there are, of course, two finite sequences $Q_1 = \{\vec{x}_1, \ldots, \vec{x}_n\},\ Q_2 = \{\vec{y}_1, \ldots, \vec{y}_n\},\ \vec{x}_i, \vec{y}_i$ members of $\emptyset \neq \mathrm{int}(J \cap S_i)$ such that

$$I(J) = \sum_{j=1}^{n} \rho(\vec{x}_j)\, d(\vec{y}_j) v(J \cap S_j) = \mathtt{st}({}^*I(J)). \tag{3}$$

It's equation (3) above that yields the concept of the particle point masses by defining $m_j = \rho(\vec{x}_j)\, v(J \cap S_j)$ and, hence,

$$I(J) = \sum_{j=1}^{n} m_j\, d(\vec{y}_j) v(J \cap S_j) \tag{4}$$

If now one assumes the particle point mass equation (4) holds for each member of $\mathcal{P}$ then *-transfer yields

$${}^*I(J) = \sum_{j \in \mathcal{A}}{}^*m_i\ {}^*d(\vec{y}_j)\, {}^*v({}^*J \cap S_j). \tag{5}$$

The process of introducing (i) the additivity of I, (ii) simplifying the terms of the sum, if possible, by reducing to other geometrical quantities, and (iii) extending to the hyperfinite, is called the **hyperfinite method**. However, this method is unnecessary since it is but a simple extension of the Method of Constants. Note that for non-discrete quantities it is, technically, not correct to say that we "extend the finite sum of quantities to an infinite sum of such quantities" or some similar expression as is often done by Maxwell and others. The sum is not, in general, an (external) infinite sum, but, is an internal hyperfinite sum. The conceptual and formal differences between these two concepts are considerable.

But looking at our definition 5.1.1 for the integral, we are lead to the, usually not appreciated, approximation methods. It follows that for any position real r there is a simple partition P and a finite set of subrectangles $\{S_1, \ldots, S_2\} = \mathcal{S}(P)$ such that for any set of intermediate partition points $Q = \{\vec{x}_1, \ldots \vec{x}_n\}$,

$$\Big| \sum_{i=1}^{n} \hat{\rho}(\vec{x}_i)\, \hat{d}^2(\vec{x}_i)\, v(S_i) - I(J) \Big| < r. \tag{6}$$

Now each S_i is of one of three types. (a) $S_i \subset R - \text{int}(J)$, (b) $K = \text{int}(J \cap S_i) = \text{int}(J) \cap \text{int}(S_i) \neq \emptyset$ and $\text{int}(S_i) \not\subset \text{int}(J)$, (c) $\text{int}(S_i) \subset \text{int}(J)$. Let $\partial A,\ A \subset \mathbb{R}^n$ denote the **boundary points** of A. In case (a) since $v(\partial J) = 0,\ v(J) = v(\text{int}(J)),\ \partial J \subset R - \text{int}(J)$ and $v(S) \neq 0$ then there exists some $\vec{x} \in S$ such that $\vec{x} \notin J$. In case (b)and (c) there is some $\vec{x} \in \text{int}(J) \cap S$. Consequently, there is a finite sequence of points $S = \{\vec{y}_1, \ldots, \vec{y}_n\}$ that contains a subsequence $S_0 = \{\vec{y}_{k(1)}, \ldots, \vec{y}_{k(m)}\}$ such that

$$\Big| \sum_{i=1}^{m} \rho(\vec{y}_{k(i)})\, d^2(\vec{y}_{k(i)})\, v(S_i) - I(J) \Big| < r, \tag{7}$$

where the S_i are all of type (b) or (c) and each $\vec{y}_{k(i)} \in \text{int}(J)$. Yet there exists a simple partition and another sequence $S' = \{\vec{z}_1, \ldots, \vec{z}_p\}$ that contains a subsequence $S'_0 = \{\vec{z}_{k(1)}, \ldots, \vec{z}_{k(q)}\}$ such that

$$\Big| \sum_{i=1}^{q} \rho(\vec{z}_{k(i)})\, d^2(\vec{z}_{k(i)})\, v(S_i) - I(J) \Big| < r, \tag{8}$$

and all of the rectangles S_i are interior rectangles of type (c) and each $\vec{z}_{k(i)} \in \text{int}(J)$.

Of course these special selections of members in $\text{int}(J)$ extended to the NSP-world and equation (8) reads as

$$\sum_{i=1}^{\Gamma} {}^*\rho(\vec{z}_{k(i)})\ {}^*d^2(\vec{y}_{k(i)})\ {}^*v(S_i) \approx I(J),$$

and all of the infinitesimal subrectangles S_i are interior subrectangles of type (c) and each $\vec{z}_{k(i)} \in {}^*\text{int}(J)$.

Thus there are many different NSP-world configurations and expressions that have the same N-world effects. These effects are obtain by application of the standard part operator. However, even though it is obvious that all of the procedures discussed in this section and the above numbered expressions give a very detailed and analytically correct approach to infinitesimal modeling - an approach that reveals much about the nature of the NSP-world model - an early introduction of these infinitesimal concepts into an elementary exposition for physical modeling is often unnecessary and, indeed, they will tend to submerge the new discoveries expressed by the self-evident theorems. On the other hand, under certain circumstances, it appears necessary to include these infinitesimal notions. In the next section, I diverge briefly into certain instructional aspects of these discoveries.

6.6. Instruction.

The basic methods of infinitesimal modeling revolve about the application of simplified physical or geometric theories to "simple configurations" within the N-world environment. For the physical sciences, natural system behavior is viewed locally with many of the requirements of the self-evident theorem assumed. Physical modeling is viewed as objective while geometric modeling is subjective. Physical modeling deals with observation and experience with the behavior of natural systems. Geometric modeling, while originally motivated by physical concerns, has become a subject of abstract

definitions. This is obvious from the applications made in chapter 5 for there, the properties of the geometric measures are paramount and are global properties obtained from mathematical experience. The local self-evident rules and even the infinitesimal rules in this chapter are more relevant to natural system behavior. However, both the geometric and physical may be more closely associated with IR3 and IR4 then first assumed.

Experience dictates that the simplest and first level of comprehension for physical concerns is obtained from the two theorems on the Self-evident Method of Constants. They would be the easiest to apply for the neophyte. There is a reason, however, why in certain cases the actual infinitesimal rules IR3 and IR4 need to be applied. This is especially the case for geometric measures. These rules would be a second level of comprehension and this entails a certain basic familiarity with infinitesimal concepts. Moreover, the generating functions for a perceived linear functional need not be the integrand utilized. Indeed, many of the rules and integral notions within chapter 5 and chapter 6 can be expressed by hyperfinite summation that yields internal functionals not just the standard extension of a standard concept. An example of this is our proof of the "length of a curve integrand" in appendix 6 on pages 209–210.

In the proof for the length of a continuously differentiable curve $c\colon [0,1] \to \mathbb{R}^n$ an internal functional is defined for the length of a, possibly broken, hyperpolygonal line determined by a hypercurve ${}^*c\colon {}^*[0,1] \to {}^*\mathbb{R}^n$ with coordinates functions ${}^*f_j\colon {}^*[0,1] \to {}^*\mathbb{R}$. This internal function may be consider defined as follows: let $\{t_i\}$ be any hyperfinite partition of ${}^*[0,1]$. Let T be any nonempty internal, hence hyperfinite, subset of $\{t_i\}$ containing two or more members. For each pair of elements in T generate the hyperline segments determined by the curve ${}^*c$ and consider $\|\vec{v}_i\|$ from Example 4.4.1.A. Finally, consider the internal hyperfinite sum operator over the internal set of the $\|\vec{v}_i\|$ determined by T. Let L denote this internal functional. Now for each internal subinterval S determined by T define $L(S) = L({}^*\text{int}(S))$. This is our *-additive functional defined over the set of all such internal functions c. Under the given hypotheses it is shown that for any internal subrectangle S, taken from a simple fine partition of ${}^*[0,1]$ there exists some $t_i \in S$ such that $\|\vec{v}_i\|/dt \approx \sqrt{\sum_{j=1}^n ({}^*f_j'(t_i))^2}$. The expression on the right is the standard extension of the standard function $\sqrt{\sum_{j=1}^n (f_j'(t_i))^2}$ while the expression on the left is $L(S)/dt$. Thus our basic definition for the length of a curve - one that with a slight generalization is also shown to be equivalent to rectifiability - almost satisfies IR4 with the exception that IR4 is written in terms of a standard additive functional not an internal *-additive nonstandard entity. It is obvious how IR3 and IR4 would be modified to include such internal functionals.

The fact that geometric definitions are subjective in character would allow us to define many such concepts entirely in terms of infinitesimal concepts. This would be an aid in developing many integral statements about n-dimensional geometric properties in terms of integrals defined on k-dimensional spaces, where $k < n$.

6.7 Realism.

I will not, at this point, dwell upon the philosophical modeling concept termed "realism." This concept assumes that all mathematical objects within the mathematical formalism correspond to objects within objective (physical) reality. There is, however, an interesting historical fact relative to the rejection of realism. Even though Bohr rejected complete realism for his concept of quantum mechanics in order to force a type of physical completeness upon this theory - a completeness that is now known to be in error [`Aerts [1984]`] - he was not the first to do so. In his 1909 treatise on his theory of electrons, Lorentz accepted the notion of an ether but rejected complete realism. "*I*

should add that, while thus denying the real existence of ether stresses, we can still avail ourselves of all the mathematical transformations by which application of the formula (43) may be made easier. We need not refrain from reducing the force to a surface-integral, and for convenience's sake we may continue to apply to the quantities occurring in this integral the name of stresses. Only we must be aware that they are only imaginary ones, nothing else than auxiliary mathematical quantities. Perhaps all this that has been said about the absolute immobility of the ether and the non-existence of the stresses, may seem somewhat startling." `Lorentz [1952:31]`

As far as geometric modeling is concerned the concept of realism is not as significant as when it is applied to physical problems. Infinitesimal modeling may be considered as a partial realism. The notion of a continuum model in an ultimately discrete world seems non-realistic. However, what is being modeled is macroscopic and large scale behavior, macroscopic and large scale effects upon observers and other objects. It is not the actual physical entities that are being modeled; but, rather, the effects these entities produce relative to a specific set of scenarios.

However, on another level, modern physical theories often deal with entities that are not directly observable, only their indirect effects emerge within the laboratory setting. Technically, these entities are speculations that may in time be replaced by yet other speculative objects. Indeed, using special techniques, the language that actually expresses physical theories is mathematically embedded into the natural numbers by a Gödel coding; which is then embedded into a nonstandard structure. The linguistics of the physical theory produces a new collection of statements that tells us about the behavior of a new world, called in general, the NSP-world. `[Herrmann 1987]` Within this world we are not speculating ad hoc about infinitesimal objects, it is the standard theory that generates their logical existence and even predicts some of their extraordinary properties. Whether or not such entities are accepted as "real" or not depends upon their usefulness. In case you may have missed it previous, I again quote the following from the first paper Robinson published relative to his formal theory of infinitesimals. "*For phenomena on a different scale, such as considered in Modern Physics, the dimensions of a particular body or process may not be observable directly. Accordingly the question whether or not a scale of non-standard analysis is appropriate to the physical world really amounts to asking whether or not such a system provides a better explanation of certain observable phenomena than the standard system of real numbers. The possibility that this is the case should be borne in mind."* Fine Hall, Princeton University. `[Robinson 1961]`

Chapter 7.

PURE INFINITESIMAL INTEGRAL MODELING

7.1 Brief Discussion.

This basic manual is intended to present brief accounts of various approaches to infinitesimal modeling so that individuals may select the method the is most appropriate for their discipline. In Chapter 5, the standard and global type rules for integral modeling are introduced in a piecemeal fashion. Each of the applications in that chapter depends upon an in depth knowledge of the properties of a specific functional that measures a specific geometric or physical quantity. Actually, as is clear from the derivations, it is more likely that the properties of the specific functional are selected in concert with the requirements of the Infinite Sum Theorem. As for Chapter 6, although the Self-evident theorems are more general in character, they still rely upon the Infinite Sum Theorem as the proofs of Theorems 6.2.2 and 6.4.1 indicate. However, in general, neither the approach of Chapter 5 nor the approach of Chapter 6 appears within the pre-1960 literature. Previous appliers of the concepts of infinitesimal reasoning relied heavily upon a pure infinitesimal approach that is highly discipline orientated and which makes a direct appeal to the Riemann styled sum and Definition 5.1.1. This pure infinitesimal approach relies upon expedient definitions for elusive geometric or physical entities within the NSP-world - entities called by the general term "elements." The collection for geometric modeling includes the basic elements of IR2 in Chapter 6 and various Euclidean compositions and decompositions of these configurations.

For each discipline, **the collection of elements is sequentially defined from the most basic through the more complex in a**, **hopefully**, **consistent manner.** It is almost always the case that the defined objects are taken from those in the N-world that behave in the simplest possible discipline defined manner and this simplistic behavior is then impressed by *-transfer upon the NSP-world. For geometry, the selection of an acceptable set of elements is somewhat more arbitrary than for physical disciplines. The experimental or observational disciplines rely upon the concept of what might be termed as simple or idealized behavior and well-grounded physical methods of approximation. However, after the infinitesimal elements have been selected then, in all cases of which I am aware, the rigorous derivations that these elements lead to an integral model are very similar. Thus, there may be an unconscious interplay between the appropriate element, as axiomatically selected, and the rigorous method. Examination of the literature leads to the following general observations.

(i) The infinitesimal elements are axiomatically selected, per discipline, and the concept of (hyperfinite) summation is applied. Whether or not this summation can be calculated by means of a Riemann styled integral is somewhat secondary.

(ii) In anticipation of an analytical approach, it has become customary to employ certain symbols and terms that tend to describe inaccurately the actual situation. As a prototype consider the symbol ds. In infinitesimal modeling this is a **general symbol** that represents any of the lengths of hyperline segments that comprise any of the hyperpolygonal curves that determine the length of a geometric curve. With respect to our notation $ds = \|\vec{v}_i\|$. Only when actual attempts are made to calculate the length of the entire curve are additional analytical constraints placed upon the representing functions. For modeling purposes, the geometric or physical elements are considered more basic in character in order to avoid, if possible, all of the well-known analytical difficulties.

(iii) Restriction to the monadic environment takes place only when an n-dimensional quantity

is being measured with respect to an n-dimensional partition. Otherwise, the infinitesimal elements span both the N-world and the NSP-world; portions are monadic and portions are not.

(iv) The selected infinitesimal elements are not unique, even in their general character.

In the next section, certain elements will be selected for problems in elementary infinitesimal geometry. Following this, analytical constraints will be applied so that elementary integral expressions can be rigorously derived. These rigorous derivations display common derivation features that should be apparent. Indeed, the common features already appear in the proof on page 127 that gives the length of a continuously differentiable curve relative to the hyperpolygonal approximating curves discussed in examples 4.4.1.A and 4.4.1.B.

7.2 Geometric Elements.

The first and most basic measure that appears in Euclid's geometry book is the length of a line segment. I repeat, with slight modifications, the information in example 4.4.1.A for the generalization of this to curves.

Recall that a curve is a continuous map $c\colon [0,1] \to \mathbb{R}^n$. This is equivalent to considering c as determined by n continuous coordinate functions $x_i = f_i(t),\ 1 \leq i \leq n$ each defined on $[0,1] \subset \mathbb{R}$. The geometric curve C determined by these functions is usually considered as the set $\{(x_1,\ldots,x_n)|t \in [0,1]\}$. The *-transform of these defining functions leads to the functions $x_i = {}^*f_i(t),\ 1 \leq i \leq n$, each defined on ${}^*[0,1] \subset {}^*\mathbb{R}$ and they generate the "hypercurve" ${}^*C \subset {}^*\mathbb{R}^n$.

Let Q be any fine partition of ${}^*[0,1]$. By *-transfer, Q behaves like an ordered partition of the interval $[0,1]$ as defined in the standard sense and we write $Q = \{t_i \mid 0 \leq i \leq \Gamma\}$. The internal hyperfinite set Q generates the internal set of "points" $P = \{({}^*f_1(t_i),\ldots,{}^*f_n(t_i))|t_i \in Q\}$ that are members of the hypercurve ${}^*C$. Now for each $i = 0,\ldots,\Gamma-1$, and each $j,\ 0 \leq j \leq n$ let ${}^*f_j(t_{i+1}) - {}^*f_j(t_i) = d(j,i)$. Since c is continuous then each $d(j,i) \in \mu(0)$. For each $i \in {}^*\mathbb{N}$ such that $0 \leq i \leq \Gamma-1$, the internal set $\ell_i = \{(x_1,\ldots,x_n)|\forall j \in {}^*\mathbb{N},\ 0 \leq j \leq n,\ x_j = {}^*f_j(t_i) + t(d(j,i)) \wedge t \in {}^*[0,1]\}$ is a hyperline segment connecting the two points $({}^*f_1(t_i),\ldots,{}^*f_n(t_i)),\ ({}^*f_1(t_{i+1}),\ldots,{}^*f_n(t_{i+1}))$ on the curve ${}^*C$. From this one obtains the internal hyperpolygonal curve $\mathcal{P}_Q = \bigcup\{\ell_i|0 \leq i \leq \Gamma-1\}$. As to the length of $\mathcal{P}_Q$ simply extend the concept of length in the classical sense by defining for each $i = 0,\ldots,\Gamma-1$ the vector $\vec{v}_i = (d(1,i),\ldots,d(n,i)) \in {}^*\mathbb{R}^n$. Then let the hyperfinite sum $\sum_{i=0}^{\Gamma-1} \|\vec{v}_i\| = |\mathcal{P}_\Gamma| \in {}^*\mathbb{R}$. **For the curve**, **the geometric element is the hyperline segment** ℓ_i. Even though, in general, you would have a different hyperpolygon with a different hyperreal length for $\forall \Gamma \in \mathbb{N}_\infty$ the following Proposition is proved in appendix 6.

Theorem 7.2.1. *Consider continuous* $c\colon [0,1] \to \mathbb{R}^n$. *Then* c *is rectifiable if and only if there exists some* $L \in \mathbb{R}$ *such that for every fine partition* Q *of* ${}^*[0,1]$

$$\mathtt{st}({}^*|\mathcal{P}_Q|) = L.$$

Application 7.2.1. *The 2-dimensional area between two continuous curves using the 1-dimensional integral.*

Give continuous $f\colon [a,b] \to \mathbb{R}$ and continuous $g\colon [a,b] \to \mathbb{R}$, where $f(x) \leq g(x)$ for each $x \in [a,b]$. Then the area between the curves f and g is

$$\int_a^b (g(x) - f(x))\,dx.$$

Derivation. From the NSP-world view point each curve is viewed as any hyperpolygonal line created by any simple fine partition of ${}^*[a,b]$. Let Q be such a simple fine partition. For a given $I_i = [x_i, x_{i+1}] \in {}^*\mathcal{S}(Q)$, let $\ell(f)_i$ be the corresponding hyperline segment determined by the curve f and $\ell(g)_i$ that determined by g for the interval I_i. Select as the infinitesimal element the **hypertrapezoid**, T_i, with sides $\ell(f)_i,\ \ell(g)_i$ and the hyperline segments with end points $\{(x_i,\ {}^*g(x_i),\ (x_i,\ {}^*f(x_i))\}$ and $\{(x_{i+1},\ {}^*g(x_{i+1}),\ (x_{i+1},\ {}^*f(x_{i+1}))\}$. Observe that this hypertrapezoid selection is consistent with the NSP-world view of a curve and the hypertrapezoid is composed of *-Euclidean composition or decomposition of our previously defined basic elements. By *-transfer the hypertrapezoid's hyperarea ${}^*A(T_i) = (1/2)(({}^*g(x_i) - {}^*f(x_i)) + ({}^*g(x_{i+1}) - {}^*f(x_{i+1})))\,dx$. As was done for the length of a curve, let $T(Q) = \bigcup\{T_i \mid 0 \le i \le \Gamma - 1\}$ and consider the hyperfinite sum

$$\sum_{I_i \in {}^*\mathcal{S}(Q)} {}^*A(T_i) = {}^*A(T(Q)). \tag{1}$$

From this definition it follows that

$$\mathtt{st}({}^*A(T(Q))) = \int_a^b (g(x) - f(x))\,dx. \tag{2}$$

$\triangle$ Rather than relegate the proof that there exists a unique $r \in \mathbb{R}$ such that for any simple fine partition Q of ${}^*[a,b]$, ${}^*A(T(Q)) \in \mu(r)$ to the appendix we present it as follows by investigating the behavior of the expression $(({}^*g(x_i) - {}^*f(x_i)) + ({}^*g(x_{i+1}) - {}^*f(x_{i+1})))$. From the uniform continuity of $g,\ f$ it follows that $({}^*g(x_i) - {}^*f(x_i)) \approx ({}^*g(x_{i+1}) - {}^*f(x_{i+1}))$. Hence, ${}^*A(T_i) = (({}^*g(x_i) - {}^*f(x_i)) + \delta_i)\,dx$, where $\delta_i \in \mu(0)$. Consequently,

$$\sum_{I_i \in {}^*\mathcal{S}(Q)} {}^*A(T_i) = \sum_{I_i \in {}^*\mathcal{S}(Q)} (({}^*g(x_i) - {}^*f(x_i))\,dx + \sum_{I_i \in {}^*\mathcal{S}(Q)} \delta_i\,dx. \tag{3}$$

By considering the finite case it follows that there exists some $\delta \in \{|\delta_i| \mid 0 \le i \le \Gamma - 1\}$ such that $\delta = \max\{|\delta_i| \mid 0 \le i \le \Gamma - 1\}$ and, hence, $\delta \in \mu(0)$. By *-transfer we have that

$$\Big|\sum_{I_i \in {}^*\mathcal{S}(Q)} \delta_i\,dx\Big| \le \sum_{I_i \in {}^*\mathcal{S}(Q)} |\delta_i|\,dx \le \delta \sum_{I_i \in {}^*\mathcal{S}(Q)} dx = \delta(b-a) \in \mu(0). \tag{4}$$

Thus

$$\sum_{I_i \in {}^*\mathcal{S}(Q)} \delta_i\,dx = \lambda \in \mu(0). \tag{5}$$

Therefore, (2) can now be written as

$$\sum_{I_i \in {}^*\mathcal{S}(Q)} {}^*A(T_i) = \sum_{I_i \in {}^*\mathcal{S}(Q)} (({}^*g(x_i) - {}^*f(x_i))\,dx + \lambda \tag{6}$$

Since the function $(g(x) - f(x))$ is bounded and integrable and the partition Q can also be considered as an internal immediate partition then Theorem 5.2.2 yields

$$\mathtt{st}(\sum_{I_i \in {}^*\mathcal{S}(Q)} {}^*A(T_i)) = \mathtt{st}(\sum_{I_i \in {}^*\mathcal{S}(Q)} (({}^*g(x_i) - {}^*f(x_i))\,dx) = \int_a^b (g(x) - f(x))\,dx. \tag{7}$$

and this completes the proof. $\triangle$

The above derivation certainly appears interesting except for the obvious fact that the **hypertrapezoid configuration is not a unique geometric element.** The key to this derivation is the process that begins with the statement that the values are infinitely close ($\approx$) and the expressions (3)(4)(5)(6)(7). The derivation would hold for hyperrectangles, hypertriangles and the like. Hyperrectangles yield the same 1-dimensional integral expressions. Hypertriangles yield 1/2 the indicated integral expression, and so forth. This non-uniqueness for elements is sometimes considered one of the basic difficulties with infinitesimal modeling both for the geometric and physical. However, since the concept of area obtained by integrals is supposed to be an extension of the Euclidean area measure then any element that does not yield such an extension can be rejected. *For elementary calculus, the use of the hypertrapezoid is very expedient when the usual 1-dimensional integral expressions for the basic geometric measures are considered. It will be our element of choice in this case.*

Application 7.2.2. *Volume of revolution using the 1-dimensional integral.*

Given continuous $f\colon [a,b] \to \mathbb{R}$ where $f(x) \geq 0$ for each $x \in [a,b]$. Then the volume obtained by rotating this curve about the x-axis is

$$\int_a^b \pi f^2(x)\, dx.$$

Derivation. Let Q be a simple fine partition of ${}^*[a,b]$. For a given $I_i = [x_i, x_{i+1}] \in {}^*\mathcal{S}(Q)$, let $\ell(f)_i$ be the corresponding hyperline segment determined by f. Consider the hypertrapezoid composed of hyperline segments $\ell(f)_i$, I_i and hyperline segments with end point $\{(x_i,0),(x_i,\, {}^*f(x_i)\}$, $\{(x_{i+1},0),(x_{i+1},\, {}^*f(x_{i+1})\}$. The rotation of this hypertrapezoid about the x-axis is the frustum of a right cone, F_i. By *-transfer, the volume of such a *-geometric configuration is ${}^*V(F_i)) = (1/3)\pi(f^2(x_i) + f(x_i)f(x_{i+1}) + f^2(x_{i+1}))$. Assuming the usual *-additivity, the volume of the entire configuration $F(Q) = \bigcup\{F_i \mid 0 \leq i \leq \Gamma - 1\}$ is

$$\sum_{I_i \in {}^*\mathcal{S}(Q)} {}^*V(F_i) = {}^*V(F(Q)). \tag{8}$$

From this definition it will follow that

$$\mathtt{st}({}^*V(F(Q))) = \int_a^b \pi\, f^2(x)\, dx. \tag{9}$$

Noting that $f^2(x_i) + f(x_i)f(x_{i+1}) + f^2(x_{i+1}) \approx f^2(x_i) + f(x_i)f(x_i) + f^2(x_i) = 3f^2(x_i)$ then (9) follows in the same manner as in the derivation for application 7.2.1.

Application 7.2.3. *The surface area of revolution using the 1-dimensional integral.*

Given continuous differentiable $f\colon [a,b] \to \mathbb{R}$ where $f(x) \geq 0$ for each $x \in [a,b]$. Then the surface area obtained by rotating this curve about the x-axis is

$$\int_a^b 2\pi f(x)\sqrt{1 + (f'(x))^2}\, dx.$$

Derivation. Let Q be a simple fine partition of ${}^*[a,b]$. For a given $I_i = [x_i, x_{i+1}] \in {}^*\mathcal{S}(Q)$, let $\ell(f)_i$ be the corresponding hyperline segment determined by f. Consider the hypertrapezoid composed of hyperline segments $\ell(f)_i$, I_i and hyperline segments with end point

$\{(x_i,0),(x_i,\ {}^*f(x_i)\},\ \{(x_{i+1},0),(x_{i+1},\ {}^*f(x_{i+1})\}$. The rotation of this hypertrapezoid about the x-axis is the frustum of a right cone, F_i. By *-transfer, the surface area of such a *-geometric configuration is $S(F_i)) = \pi({}^*f(x_i) + {}^*f(x_{i+1}))\,{}^*|\ell(f)_i|$, where ${}^*|\ell(f)_i|$ is the *-length of the hyperline segment. In appendix 6, ${}^*|\ell(f)_i| = \|\vec{v}_i\|$. Assuming the usual *-additivity, the surface area of the entire configuration $F(Q) = \bigcup\{F_i \mid 0 \le i \le \Gamma - 1\}$ is

$$\sum_{I_i \in\ {}^*\mathcal{S}(Q)} {}^*S(F_i) = {}^*S(F(Q)). \tag{10}$$

From this definition it will follow that

$$\mathtt{st}({}^*S(F(Q))) = \int_a^b 2\pi f(x)\sqrt{1+(f'(x))^2}\,dx. \tag{11}$$

Now the complete derivation uses the proof of the integral length formula that appears in appendix 6. There it is shown that $\|\vec{v}_i\| = \sqrt{1+({}^*f'(x_i))^2}\,dx + \delta_i\,dx$, where $\delta_i \in \mu(0)$. But, ${}^*f(x_i) = {}^*f(x_{i+1}) + \lambda_i,\ \lambda_i \in \mu(0)$. Therefore, ${}^*f(x_i) + {}^*f(x_{i+1}) = 2\,{}^*f(x_i) + \lambda_i$. Since f is bounded then ${}^*S(F_i)) = \pi({}^*f(x_i) + {}^*f(x_{i+1}))\,{}^*|\ell(f)_i| = 2\pi f(x_i)\sqrt{1+({}^*f'(x_i))^2}\,dx + \gamma_i\,dx,\ \gamma_i \in \mu(0)$. The derivation is completed by application of steps (3)(4)(5)(6)(7) as demonstrated in application 7.2.1.

⇓ IMPORTANT ⇓

⇒ Please note that for applications 7.2.1, 7.2.2, and 7.2.3 the actual defining geometric quantities are given by equations (1), (8) and (10), respectively. The geometric configurations, $T(Q)$ and $F(Q)$ composed of the hyperfinite union of the respective infinitesimal elements may be considered as members of the nonstandard extension of the set of all ordinary Euclidean configurations. Configurations $T(Q),\ F(Q)$ are internal subsets of appropriate *-Euclidean entities that do exhibit a standard area or volume measure. By *-transfer of the standard properties of geometric measures relative to subsets it follows that the standard part of each of these expressions exists as a real number. What is established in these applications under the analytical constraints given is that the standard part is expressible by the indicated integral. ⇐

Relative to the infinitesimal geometric elements, I have not altered the requirement that for the n-dimensional integral these elements be the infinitesimal subrectangles. In the latter sections of this manual, since it is written for individuals with a strong undergraduate mathematics background, the Jordan-measurable generalization for the basic rectangular region was used. As far as a Jordan-measurable J is concerned do we ever need to consider any subrectangle that is not in the interior of J?

Let Jordan-measurable $J \subset R \subset \mathbb{R}^n$ and let Q be any fine partition of R. It is sometimes useful to assume that J is a closed subset of R (hence, compact). For, if J is not closed, then recall that the Jordan-content of $J,\ v(J) = v(\overline{J})$. **Intuitively**, **think of J as any of the ordinary regions studied in elementary calculus with their intuitive boundary**, **inner and exterior portions.** Using these ideas of boundary, inner and exterior portions of J, the set of all subrectangles determined by Q, which is denoted by ${}^*\mathcal{S}(Q)$, may be separated formally into three disjoint hyperfinite collections. Let the set of **boundary subrectangles** be

$\partial(Q) = \{S \mid S \in {}^*\mathcal{S}(Q) \wedge S \cap {}^*J \neq \emptyset \wedge S \cap ({}^*\mathbb{R}^n - {}^*J) \neq \emptyset\}$. Observe that if $S \in \partial(Q)$, then $S \subset \mu(p)$ and $\mu(p) \cap {}^*J \neq \emptyset$ and $\mu(p) \cap ({}^*\mathbb{R}^n - {}^*J) \neq \emptyset$ imply that $p \in \partial(J)$. Now consider the set of **exterior subrectangles** $\text{ext}(Q) = \{S \mid S \in {}^*\mathcal{S}(Q) \wedge S \subset ({}^*\mathbb{R}^n - {}^*J)\}$. Finally, the set of **inner subrectangles** is $\text{inn}(Q) = \mathcal{S}(Q) - (\partial(Q) \cup \text{ext}(Q))$. Observe that the boundary subrectangles generate boundary points of J. However, by considering J to be a rectangle or the interior of a rectangle then it is clear that there may exist some $S \in \text{inn}(Q)$ or $S \in \text{ext}(Q)$, respectively, such that $S \subset \mu(p)$ and $p \in \partial(J)$. It is definitely the case, however, that $S \subset {}^*J$ if and only if $S \in \text{inn}(Q)$. Of course the sets $\partial(Q)$, $\text{inn}(Q)$ and $\text{ext}(Q)$ are mutually disjoint. This leads to the very useful

Theorem 7.2.2. *Let Jordan-measurable $J \subset R \subset \mathbb{R}^n$, bounded $f\colon J \to \mathbb{R}$, and Q be any fine partition of ${}^*R$. Let hyperfinite ${}^*\mathcal{S}(Q) = \{S_i \mid 0 \leq i \leq \Gamma - 1\}$. Assume that there exists a hyperfinite sequence U_i, where $U_i \in S_i$ for each $S_i \in \mathcal{S}(Q)$. Then*

$$\mathtt{st}(\sum_{S(i)\in \text{inn}(Q)} {}^*f(U_i)\,{}^*v(S_i)) = \mathtt{st}(\sum_{S(i)\in \mathcal{S}(Q)} {}^*\hat{f}(U_i)\,{}^*v(S_i)).$$

Theorem 7.2.2 indicates, as expected, that only the interior infinitesimal subrectangles are significant when modeling with respect to the Jordan-content of any set.

Application 7.2.4. *Volume obtained by 2-dimensional integral.*

Suppose that compact Jordan-measurable $J \subset R \subset \mathbb{R}^n$. Let continuous $f\colon J \to \mathbb{R}$ For each $\vec{x} \in J$ let $f(x) \geq 0$. Then the volume between the surface determined by f and the xy-plane is

$$\int_J f(x)\,dX.$$

Derivation. Let Q be a simple fine partition of R and consider some $S_i \in \mathcal{S}(Q)$. From the NSP-world viewpoint, the surface curves obtained by intersecting the surfaces with planes parallel to the coordinate planes are hyperpolygonal lines. Since the choice of the particular hyperline segments is arbitrary and $S_i = [x_i, x_{i+1}] \times [y_i, y_{i+1}]$ then consider the *-Euclidean configuration $TR(S_i)$ composed of the truncated hyperrectangular solid with S_i as its base and its top a parallelogram with adjacent sides the hyperline segments $\{(x_i, y_i, {}^*\hat{f}(x_i, y_i), (x_i, y_{i+1}, {}^*\hat{f}(x_i, y_{i+1})\}$ and $\{(x_i, y_i, {}^*\hat{f}(x_i, y_i)), (x_{i+1}, y_i, {}^*\hat{f}(x_{i+1}, y_i)\}$ Using the *-Euclidean measure for the volume ${}^*V(T(S_i)$ then once again *-additivity yields for the configuration $TR(Q) = \bigcup\{TR(S_i) \mid 0 \leq i \leq \Gamma - 1\}$

$$\sum_{S(i)\in \mathcal{S}(Q)} {}^*V(TR(S_i)) = {}^*V(TR(Q)). \tag{12}$$

Investigating the various configurations $TR(S_i)$, and using the *-Euclidean measure of these configurations then Theorem 7.2.2 yields that

$$\sum_{S(i)\in \text{inn}(Q)} {}^*V(T(S_i)) \approx \sum_{S(i)\in \mathcal{S}(Q)} {}^*V(S_i). \tag{13}$$

Since J is compact then

$${}^*f(x_i, y_i) \approx {}^*f(x_i, y_{i+1}) \approx {}^*f(x_{i+1}, y_i). \tag{14}$$

Applying the *-Euclidean measure formula yields that ${}^*V(T(S_i) = ({}^*f(x_i, y_i) + \delta_i)\,dX$ and the method of Application 7.2.1 yields

$$\mathtt{st}(\sum_{S(i)\in \text{inn}(Q)} {}^*V(T(S_i)) = \mathtt{st}(\sum_{S(i)\in \mathcal{S}(Q)} {}^*V(S_i)) = \int_J f(x)\,dX. \tag{15}$$

For integral modeling, physical infinitesimal elements are based upon the Method of Constants. It assumes that the physical quantity being considered has the same effect as if it were concentrated at some point within subrectangle or another similar object. The method is essentially outlined by Maxwell in his previously quoted descriptions. Clearly, such infinitesimal physical modeling is highly discipline orientated and is closely associated with simple mind experiments.

Application 7.2.5. *The value of the electric field vector at a point P exterior to the plane of a charged 2-dimensional closed and bounded Jordan-measurable region.*

Consider a Jordan-measurable region J, a point $P = (a, b, c),\ (c \neq 0)$ in space exterior to the plane. Let $\rho(x, y, 0)$ be a continuous charge density function defined on J. The scalar value of the electric field at P is

$$|\vec{E}(a, b, c)| = \int_J \frac{\rho(x, y, 0)}{(x-a)^2 + (y-b)^2 + c^2}\, dX.$$

Derivation. (The method of point charges.) For a single point charge q at a distance r from P, the definition of the scalar vaule of the electric field is $|\vec{E}| = q/r^2$. Let $J \subset R \subset \mathbb{R}^2$, $(x_i, y_i, 0) \in J$, and consider a simple fine partition Q of R. If one considers a point charge with value ${}^*\rho(x, y, 0)\ {}^*v(S_i);\ (x, y, 0) \in S_i \in \text{inn}(Q)$, then $|\vec{E}(a, b, c)| = ({}^*\rho(x', y', 0)\ {}^*v(S_i))/((x-a)^2 + (y-b)^2 + c^2)$. But the function $h(x, y) = \rho(x, y, 0)/((x-a)^2 + (y-b)^2 + c^2)$ is uniformly continuous on J. Thus

$${}^*\rho(x', y', 0)\ {}^*v(S_i) = ({}^*\rho(x_i, y_i, 0) + \delta_i)\ {}^*v(S_i) =$$

$${}^*\rho(x_i, y_i, 0)\ {}^*v(S_i) + \delta_i\ {}^*v(S_i),$$

where $\delta_i \in \mu(0)$. Assuming that the scalar value of the electric field for any nonempty finite set of point charges is the sum of the individual values and that within any 2-dimensional rectangle, S, with charge density $\rho(x, y, 0)$ there is some point where the entire charge can be considered as concentrated, or that for a constant charge density ρ the total charge is $\rho v(S)$, then *-transfer and the method used in applications 7.2.1 and 7.2.4 imply that the scalar value of the electric field at P due to the charge on J is

$$|\vec{E}_J(P)| = \mathtt{st}(|\,{}^*\vec{E}_J(P)|) =$$

$$\mathtt{st}\Big(\sum_{S(i)\in \text{inn}(Q)} ({}^*\rho(x', y', 0)\ {}^*v(S_i))/((x-a)^2 + (y-b)^2 + c^2)\Big) =$$

$$\int_J \frac{\rho(x, y, 0)}{(x-a)^2 + (y-b)^2 + c^2}\, dX.$$

Remark: In 7.2.5, it is not correct to simply consider the points $(x', y', 0) \in S_i$. It is always necessary that the chosen intermediate partition be internal. Further, extending the concept of the constant charge density is a major approach to infinitesimal modeling since if $\rho(x, y, 0)$ is continuous on compact J, then ρ behaves in a constant-like manner in that if infinitesimal $S \subset J$ and $p, q \in S$, then $p \approx q$ implies that $\mathtt{st}(\rho(p)) = \mathtt{st}(\rho(q))$. Using the physical element concept, this is called the **elemental method of constants.**

Chapter 8.

REFINEMENTS FOR INTEGRAL MODELING

8.1 A Very General Approach.

`Hurd and Loeb [1985]` construct a very general integral concept. For example, consider any hyperfinite $\{x_1, \ldots, x_\Gamma\} \subset X$, let B be any set of internal hyperreal-valued functions defined on X, and $\{a_1, \ldots, a_\Gamma\}$ a fixed set of hyperreal nonnegative numbers. They then consider the hyperfinite sum operator $\sum_\Gamma$ defined on each $f \in B$ by $\sum_\Gamma f = \sum_1^\Gamma a_i\, f(x_i)$. This and other examples are generalized and an entire theory of integration is developed that incorporates various classical generalizations of the Riemann integral, especially the Lebesque. They apply their theory to stochastic processes such as the Poisson process and Brownian motion. Since the background necessary to study their generalization is beyond the scope of these manuals, their theory will not be presented. Indeed, except for concept of hyperfinite summation, their general approach is probably unsuited for elementary modeling. Our goal in this last chapter on integral modeling is to examine more closely the specific contents of the hyperfinite sum as defined in definition 5.1.1 relative to modifications of its geometric or physical meaning.

Recall definition 5.1.1. Let $f\colon R \to \mathbb{R}$ be bounded and $\mathcal{P}$ the set of simple partitions of R. Then f is said to be integrable if there exists some $r \in \mathbb{R}$ and a simple fine partition, $P \in {}^*\mathcal{P}$ such that for each of its internal intermediate partitions $Q = \{\vec{v}_q\}$, where $1 \le q \le \Gamma \in \mathbb{N}_\infty$, it follows that $\sum_{k=1}^\Gamma {}^*f(\vec{v}_q)\, {}^*v(R_q) \in \mu(r)$.

Obviously, modifications can be made in the concept of the basic partition P or the intermediate partition Q. Such modifications are discussed at the conclusion of this chapter. It is important to stress at this point something that is not apparent about this definition. It is immediate from examination of the proofs in appendix 5, that, basically, there are two reasons why this definition works. First, ${}^*v(R_q) \in \mu(0)$. And, secondly, due to behavior of the measure v, the integral value r is independent of the fine partition chosen. This must be taken into consideration if the partition concepts are not to be altered. For the basic modifications that follow, we extend the partitioning requirement to any arbitrary fine partition and any arbitrary intermediate partition and modify the terms of the hyperfinite sum with a view towards applications as they appear in elementary calculus courses.

8.2. The Line Integral.

Let bounded $\phi\colon [a,b] \to \mathbb{R}$. Our first consideration is to modify the value of the *-measure ${}^*v([t_{i-1}i, t_i])$, where $[t_{i-1}, t_i] \subset {}^*[a,b]$. We know the importance of the "increment" in the calculus, hence, our first modification is to replace ${}^*v([t_{i-1}, t_i])$ with the standard extension of the increment operator. In place of ${}^*v([t_{i-1}, t_i])$ write ${}^*\phi(t_i) - {}^*\phi(t_{i-1})$. What happends if ϕ is continuous? Well, in that case, if $t_i - t_{i-1} \in \mu(0)$, then ${}^*\phi(t_i) - {}^*\phi(t_{i-1}) \in \mu(0)$ and one of the most basic requirements for that factor of the term of the hyperfinite sum is met.

Definition 8.2.1. (Riemann - Stieltjes Integral). Let bounded $f\colon [a,b] \to \mathbb{R}$ and bounded $\phi\colon [a,b] \to \mathbb{R}$. Then f is **RIEMANN STIELTJES Integrable with respect to** ϕ if there exists a real r such that for any fine partition $P = \{a = t_0, \ldots, t_\Gamma = b\}$ of $[a,b]$ and any internal intermediate

partition $Q = \{t'_1, \ldots, t'_\Gamma\}$ of P it follows that

$$\sum_{i=1}^{\Gamma} {}^*f(t'_i)\,({}^*\phi(t_i) - {}^*\phi(t_{i-1})) \in \mu(r).$$

Rather than investigate the Riemann-Stieltjes integral as defined in 8.2.1, I pass directly to the line integral of elementary calculus. Referring to example 4.4.1A, let bounded $c\colon [a,b] \to \mathbb{R}^n,\ c(t) = (c_1(t), \ldots, c_n(t))$ be considered a curve with graph C. Assume that $E \subset \mathbb{R}^n,\ C \subset E$ and bounded $F\colon E \to \mathbb{R}^n$. Write F as $F(\vec{x}) = (f_1(\vec{x}), \ldots, f_n(\vec{x}))$. Consider the composite function $(Fc)\colon [a,b] \to \mathbb{R}^n$ defined by $(Fc)(t) = (f_1(c(t)), \ldots, f_n(c(t))),\ t \in [a,b]$. Notice that if $t'_j \in [t_{j-1}, t_j] \subset {}^*[a,b]$, then, letting $\bullet$ denote the "dot" (inner) product,

$${}^*(Fc)(t'_j) \bullet {}^*\vec{v}_j = \sum_{i=1}^{n} {}^*f_i({}^*c(t'_j))\,({}^*c_i(t_{j-1}) - {}^*c_i(t_j)),$$

where ${}^*\vec{v}_j = {}^*c(t_j) - {}^*c(t_{j-1})$ represents the directed hyperline segment portion of some hyperpolygonal curve representation for C. Now taking a hyperfinite sum yields

$$\sum_{j=1}^{\Gamma} {}^*(Fc)(t'_j) \bullet {}^*\vec{v}_j = \sum_{j=1}^{\Gamma}\Big(\sum_{n=1}^{n} {}^*f_i({}^*c(t'_j))\,({}^*c_i(t_{j-1}) - {}^*c_i(t_j))\Big) =$$

$$\sum_{i=1}^{n}\Big(\sum_{j=1}^{\Gamma} {}^*f_i({}^*c(t'_j))\,({}^*c_i(t_{j-1}) - {}^*c_i(t_j))\Big).$$

Hence, if for each $i = 1, \ldots, n,\ \sum_{j=1}^{\Gamma} {}^*f_i({}^*c(t'_j))\,({}^*c_i(t_{j-1}) - {}^*c_i(t_j)) \in \mu(r_i)$, then $\sum_{j=1}^{\Gamma} {}^*(Fc)(t'_j) \bullet {}^*\vec{v}_j \in \mu(r_1 + \cdots + r_n)$. On the other hand, if $\sum_{j=1}^{\Gamma} {}^*(Fc)(t'_j) \bullet {}^*\vec{v}_j \in \mu(r)$ and for each $i = 1, \ldots, n,\ \sum_{j=1}^{\Gamma} {}^*f_i({}^*c(t'_j))\,({}^*c_i(t_{j-1}) - {}^*c_i(t_j)) \in \mathcal{O}$, then $\mathtt{st}(\sum_{j=1}^{\Gamma} {}^*f_i({}^*c(t'_j))\,({}^*c_i(t_{j-1}) - {}^*c_i(t_j)) = r_i$ implies that $r_1 + \cdots r_n = r$. This leads to the notion of the line integral as an extension of the Riemann-Stieltjes integral.

Definition 8.2.2. (Line Integral). Let bounded $c\colon [a,b] \to \mathbb{R}^n,\ c(t) = (c_1(t), \ldots, c_n(t))$ be considered a curve with graph C. Assume that bounded $F\colon C \to \mathbb{R}^n$, where $F(\vec{x}) = (f_1(\vec{x}), \ldots, f_n(\vec{x}))$. Consider the composite function $(Fc)\colon [a,b] \to \mathbb{R}^n$ defined by $(Fc)(t) = (f_1(c(t)), \ldots, f_n(c(t))),\ t \in [a,b]$. Then F is **LINE INTEGRABLE with respect to C** if there exists some real r such that for any fine partition $P = \{a = t_0, \ldots, t_\Gamma = b\}$ of $[a,b]$ and any internal intermediate partition $Q = \{t'_1, \ldots, t'_\Gamma\}$ of P it follows that

$$\sum_{j=1}^{\Gamma} {}^*(Fc)(t'_j) \bullet {}^*\vec{v}_j \in \mu(r).$$

In which case we write

$$r = \mathtt{st}\Big(\sum_{j=1}^{\Gamma} {}^*(Fc)(t'_j) \bullet {}^*\vec{v}_j\Big) = \int_C F \bullet d\vec{R}.$$

⇓ IMPORTANT ⇓

⇒ In the next derivation, the significant method of the maximum and minimum is applied to physical elements. This approach is different than that used in integral rule IR3. Moreover, a special relationship between the work done by a force field over a hyperline segment and points on a rectifiable curve is advanced. This relationship aids in our comprehension of energy related NSP-world properties.⇐

Application 8.2.1. *Energy expended within a force field while moving along a curve.*

Let $c\colon [a,b] \to \mathbb{R}^n$ be a continuous differentiable curve with graph C. Assume that continuous $F\colon E \to \mathbb{R}^n$, open $E \supset C$. The work done in moving through the force field on the path C is

$$W(C) = \int_C F \bullet d\vec{R} = \int_a^b \Big(\sum_{i=1}^n f_i(c_1(t),\ldots,c_n(t))c_i'(t)\Big)\, dt.$$

Derivation. In experimental physics, the concept of "work" (energy expended) is introduce. **All one needs to do is to establish its properties for a polygonal curve.** Suppose we have continuous force field $F\colon E \to \mathbb{R}^n$. Let $\mathcal{P}_k \subset \mathbb{R}^n$ be any finite polygonal curve, ℓ_j one of the line segment portions of $\mathcal{P}_k$ with $\vec{v}_j = (c_1(t_j) - c_1(t_{j-1}), \ldots, c_n(t_j) - c_n(t_{j-1}))$, denoting this line segment considered as a directed line segment in the direction of motion through the field F. If F is constant on ℓ_j, then the work done moving along ℓ_j is defined as $W(\ell_j) = F \bullet (\vec{v}_j/\|\vec{v}_j\|)\|\vec{v}_j\|$, where length of $\ell_j = \|\vec{v}_j\|$. What if the force field is not constant? Consider $\mathcal{P}_k$ as represented by a continuous $\ell\colon [a,b] \to \mathbb{R}^n$ and assume that F is defined on ℓ. Then for a given $\ell_j = \{(x_1(t),\ldots,x_n(t)) \mid t \in [t_{j-1},t_j]\}$ there exists some $t_m,\ t_M$ such that $W_m(\ell_j) = F(\ell(t_m)) \bullet \vec{v}_j \le W(\ell_j) = F(\ell(t)) \bullet \vec{v}_j = W_M(\ell_j) = F(\ell(t_M)) \bullet \vec{v}_j$ for each $t \in [t_{j-1},t_j]$. Let's make the one assumption that the actual amount of work expended moving along the line segment ℓ_j is $W(\ell_j)$ and that $W_m(\ell_j) \le W(\ell_j) \le W_M(\ell_j)$. Then from continuity there exists some $h_j' \in [t_{j-1},t_j]$ such that $W(\ell_j) = F(\ell(h_j')) \bullet \vec{v}_j$. The idea of the nonconstant force field over a line segment is embedded into the the NSP-world by *-transfer assuming that what has been established above holds for all such polygonal curves. Hence, let $\mathcal{P}_\Omega$ be a hyperpolygonal representation for the curve generated by a fine partition, ℓ_j an hyperline segment in $\mathcal{P}_\Omega$. Since F is continuous on E then ${}^*F$ is defined on $\mathcal{P}_\Omega$. It follows that ${}^*W(\ell_j) = {}^*F({}^*\ell_j({}^*h_j')) \bullet {}^*\vec{v}_j$. For polygonal curves, in general, the work done is an additive function. Thus for the hyperpolygonal curve $\mathcal{P}_\Omega$

$$ {}^*W(\mathcal{P}_\Omega) = \sum_{j=1}^{\Omega} {}^*F({}^*\ell_j({}^*h_j')) \bullet {}^*\vec{v}_j. \tag{1}$$

We now show that there exists a real number $W(C)$ such that for any fine partition $\mathcal{P}_\Gamma$ and <u>any</u> intermediate partition $Q = \{t_1',\ldots,t_\Gamma'\}$, $\mathtt{st}({}^*W(\mathcal{P}_\Gamma)) = W(C)$.

Consider

$$ {}^*F({}^*c(t_j')) \bullet {}^*\vec{v}_j = \sum_{i=1}^{n} {}^*f_i({}^*c_1(t_j'),\ldots,{}^*c_n(t_j'))\,({}^*c_i(t_j) - {}^*c_i(t_{j-1})). \tag{2}$$

The curve c being continuously differentiable on $[a,b]$ implies that ${}^*c_i(t_j) - {}^*c_i(t_{j-1}) = \big({}^*c_i'(t_j) + \delta_{ij}\big)\,(t_j - t_{j-1})$, where $\delta_{ij} \in \mu(0)$. Hence, ${}^*F({}^*c(t_j')) \bullet {}^*\vec{v}_j =$

$$ \sum_{i=1}^{n} {}^*f_i({}^*c_1(t_j'),\ldots,{}^*c_n(t_j'))\ {}^*c_i'(t_j)\,(t_j - t_{j-1}) + \left(\sum_{i=1}^{n} \delta_{ij}\right)\,(t_j - t_{j-1}). \tag{3}$$

Consequently, ${}^*F({}^*c(t'_j)) \bullet {}^*\vec{v}_j =$

$$\sum_{i=1}^{n} {}^*f_i({}^*c_1(t'_j), \ldots, {}^*c_n(t'_j))\ {}^*c'_i(t_j)\,(t_j - t_{j-1}) + \delta_j\,(t_j - t_{j-1}). \tag{4}$$

Uniform continuity of the f_i yields that ${}^*f_i({}^*c_1(t'_j), \ldots, {}^*c_n(t'_j)) = {}^*f_i({}^*c_1(t_j), \ldots, {}^*c_n(t_j)) + \lambda_{ij}$, where $\lambda_{ij} \in \mu(0)$. Once again this yields ${}^*F({}^*c(t'_j)) \bullet {}^*\vec{v}_j =$

$$\sum_{i=1}^{n} {}^*f_i({}^*c_1(t_j), \ldots, {}^*c_n(t_j))\ {}^*c'_i(t_j)\,(t_j - t_{j-1}) +$$

$$\lambda_j\,(t_j - t_{j-1}) + \delta_j\,(t_j - t_{j-1}) \tag{5}$$

Now continuing the basic **elemental derivation process** leads to $\sum_{j=1}^{\Gamma}\big({}^*F({}^*c(t'_j)) \bullet {}^*\vec{v}_j\big) =$

$$\sum_{j=1}^{\Gamma}\Big(\sum_{i=1}^{n} {}^*f_i({}^*c_1(t_j), \ldots, {}^*c_n(t_j))\ {}^*c'_i(t_j)\,(t_j - t_{j-1})\Big) + \delta,\ \delta \in \mu(0). \tag{6}$$

Therefore, from Theorem 5.1.2

$$W(C) = \mathtt{st}\Big(\sum_{j=1}^{\Gamma}\big({}^*F({}^*c(t'_j)) \bullet {}^*\vec{v}_j\big)\Big) =$$

$$\int_a^b \Big(\sum_{i=1}^{n} f_i(c_1(t), \ldots, c_n(t))c'_i(t)\Big)\,dt. \tag{7}$$

The fact that this is a line integral follows from Definition 8.2.1. Finally, it will almost always be the case that the special energy property ${}^*F(\ell_j(t'_j)) \bullet {}^*\vec{v}_j = {}^*F({}^*c(t'_j)) \bullet {}^*\vec{v}_j + \epsilon_j \|{}^*\vec{v}_j\|$, $\epsilon_j \in \mu(0)$, $t'_j = h'_j$ holds. [See note [2] on page 148.] Since C is rectifiable then $\sum_{j=1}^{\Gamma} {}^*F(\ell_j(t'_j)) \bullet {}^*\vec{v}_j \approx \sum_{j=1}^{\Gamma} {}^*F({}^*c(t'_j)) \bullet {}^*\vec{v}_j$. We are using the hyperpolygonal representations for C as the basic NSP-world entity to determine the N-world physical effects. It is clear that the appropriate measure for the work expended moving along the path C should be the unique value $\mathtt{st}(\sum_{j=1}^{\Gamma} {}^*F(\ell_j(t'_j)) \bullet {}^*\vec{v}_j)$ obtained in (7). This completes the derivation.

Obviously, derivation 8.2.1 also establishes the elementary method for the calculation of a line integral.

8.3. Order Ideals and Approximations.

One often reads in the literature that such and such an expression is a "first- order approximation" or some such phrase. These vague approximation concepts can be discussed from the infinitesimal viewpoint and, indeed, lead to the notions of the "microconstruction" and "microeffects."

Theorem 8.3.1. *For each $\epsilon \in \mu(0)$ the set $o(\epsilon) = \{\epsilon\, h \mid h \in \mu(0)\}$ is an ideal in $\mu(0)$.*

Definition 8.3.1. (Order Ideals). For a given $\epsilon \in \mu(0)$ the set $o(\epsilon)$ is called an **ORDER IDEAL (of infinitesimals)**.

We briefly investigate some of the basic properties of the order ideals. (I note that these order ideal properties appear for the first time in this manual.) First, it is obvious that $o(\epsilon) = o(-\epsilon)$. In all that follows, let $\mu(0)^+ = \{x \mid x \in \mu(0) \wedge x \geq 0\}$ be the set of all **nonnegative infinitesimals**. The next theorem seems to be one of the more significant ones relative to order ideals.

Theorem 8.3.2. *Let $\epsilon \in \mu(0)^+$. Suppose that $w \in {}^*\mathbb{R}$ and $0 \leq w \leq \epsilon\, h \in o(\epsilon)$. Then $w \in o(\epsilon)$.*

Theorem 8.3.3. *Let $\epsilon,\ \delta \in \mu(0)^+$. If $0 \leq \delta \leq \epsilon$, then $o(\delta) \subset o(\epsilon)$ and $o(\delta)$ is a ideal in $o(\epsilon)$.*

Are there order ideals such that $o(\delta) \subset o(\epsilon)$ and $o(\delta) \neq o(\epsilon)$?

(1) Let $0 < \delta \leq \epsilon,\ \delta, \epsilon \in \mu(0)^+$. Then $o(\delta\epsilon) \subset o(\epsilon)$ and $o(\delta\epsilon) \neq o(\epsilon)$.

(2) Let $n \in \mathbb{N}^+ = \mathbb{N} - \{0\},\ \epsilon \in \mu(0)^+$. Then $o(\epsilon^n) \subset o(\epsilon^{n-1}) \subset \cdots \subset o(\epsilon)$ and $o(\epsilon^i) \neq o(\epsilon^j);\ 1 \leq i, j \leq n;\ i \neq j$.

Suppose that you have the set $\{\epsilon_1, \ldots, \epsilon_n\}$. Then consider the chain $C_1 :\ o(\epsilon_1 \times \cdots \times \epsilon_n) \subset o(\epsilon_1 \times \cdots \times \epsilon_{n-1}) \subset \cdots \subset o(\epsilon_1)$. The chain C_1 is just one possible chain of order ideals leading from $o(\epsilon_1 \times \cdots \times \epsilon_n)$ to $o(\epsilon_1)$. Such chains are used for comparison purposes and, in this case, the ideal $o(\epsilon_1 \times \cdots \times \epsilon_n)$ is called an **n'th order ideal** where, in general, such ideals as $o(\epsilon)$ are called **first-order ideals.** Part of this chapter will deal with the relation between n'th order ideals and the infinitesimal concept of n'th order approximations. It is useful to consider other operational methods that might generate different order ideals rather than simply restricting their generation to products of infinitesimals.

Let $\{\epsilon_1, \ldots, \epsilon_k\} \subset \mu(0)$. Define $o(\epsilon_1, \ldots, \epsilon_k) = \{\epsilon_1\, h_1 + \cdots + \epsilon_k\, h_k \mid h_i \in \mu(0) \wedge 1 \leq i \leq k\}$.

Theorem 8.3.4. *Let $\epsilon = \max\{|\epsilon_1|, \ldots, |\epsilon_k|\}$. Then $o(\epsilon_1, \ldots, \epsilon_k) = o(\epsilon)$.*

Theorem 8.3.5. *Let $\epsilon = \max\{|\epsilon_1|, \ldots, |\epsilon_k|\}$. Then*

$$o(\epsilon_1, \ldots, \epsilon_k) = o(\sqrt{\epsilon_1^2 + \cdots + \epsilon_k^2}) = o(\epsilon).$$

Thus neither the difference of infinitesimals nor the *-Euclidean norm is a useful process for the generation of higher order ideals. As will be illustrated through out the remainder of this manual, the n'th order ideals are related to the notion of the n'th order approximation. Refer back to all of our derivations where the basic elemental derivation process is used. In each case, a single term of the required hyperfinite sum of elemental measures, say ${}^*M(\cdot)$, is investigated for an arbitrary simple partition of infinitesimal volume dX. The derivation shows that this term is equal to the value of a standard extension, say ${}^*F(\vec{x}_i)$, plus $h\, dX$, where $h \in \mu(0)$. Consequently,

$${}^*M(\cdot) - {}^*F(\vec{x}_i) \in o(dX).$$

This expression also yields an equivalence relation which is often denoted by the algebraic notation

$${}^*M(\cdot) \approx {}^*F(\vec{x}_i) \pmod{o(dX)}$$

and is expressed by stating that ${}^*M(\cdot)$ and ${}^*F(\vec{x}_i)$ are **infinitely close of order** dX**.** Hence, to apply this derivation process the values ${}^*M(\cdot)$ are not just infinitely close to ${}^*F(\vec{x}_i)$, in a general sense; but,

they are infinitely close of order dX. Noting that for an n-dimensional integral $dX = dx_1 \times \cdots \times x_n$ it follows that, from a comparative viewpoint, ${}^*M(\cdot)$ and ${}^*F(\vec{x}_i)$ can be considered as **infinitely close of order n.** In the literature, you will also find the less descriptive expression "${}^*M(\cdot)$ is infinitely close to ${}^*F(\vec{x}_i)$ compared to dX" as a synonym for "infinitely close of order dX." The reason it appears necessary that such quantities need to be infinitely close of order dX is that the steps in the derivation process that proceed from this step require a hyperfinite sum to be extracted and the results must remain infinitely close.

One little observation about the order ideals that are created by products. If $0 \neq r \in {}^*[-1,1] - \mu(0)$, then for any $\epsilon \in \mu(0)$, it follows that $o(\epsilon\, r) = o(\epsilon)$. To see this, Theorem 8.3.3 yields that $o(|\epsilon\, r|) = o(\epsilon\, r) \subset o(|\epsilon|) = o(\epsilon)$. Let $\epsilon\, h \in o(\epsilon)$. Since $h/r \in \mu(0)$, then $(\epsilon\, r)(h/r) \in o(r\epsilon)$ implies that $\epsilon\, h \in o(\epsilon\, r)$. Further, note that if $0 \neq \delta \in \mu(0)$, then $\epsilon\, r \in o(\epsilon) - o(\delta\, \epsilon.)$

8.4. nth Order Increments.

In the next section, we investigate exactly what one means by a tangent line to a curve c at a point p, where c is differentiable at p. First, however, it is useful for this and the future sections on modeling by means of the derivative or differential to formally consider the NSP-world view of the nth order increment (nth difference). For $n \in \mathbb{N}^+$ and bounded $f\colon [a, nb] \to \mathbb{R}$, recall that the **nth order increment**, $\Delta^n f(x,b)$, is defined by induction, where for $x \in [a, nb]$, by $\Delta f(x,b) = f(x+b) - f(x)$. This leads to the general expression

$$\Delta^n f(x,b) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f(x + (n-k)b) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f(x + kb).$$

Notice that as an operator the *-nth order increment, ${}^*(\Delta^n)\, {}^*f(x,b) = \Delta^n\, {}^*f(x,b)$.

Theorem 8.4.1. *Let $1 \leq n \in \mathbb{N}$. Suppose that $f^{(n-1)}\colon [a, nb] \to \mathbb{R}$ and that $f^{(n)}\colon (a, nb) \to \mathbb{R}$, where $f^{(k)}$ denotes the kth derivative of f. Then there exists some $t \in (a, nb)$ such that $\Delta^n f(a,b) = f^{(n)}(t)\, b^n$.*

Corollary 8.4.1.1 *Let $1 \leq n \in \mathbb{N}$. Suppose that $f^{(n-1)}\colon [a,b] \to \mathbb{R}$ and that $f^{(n)}\colon (a,b) \to \mathbb{R}$, then for each $dx \in \mu(0)^+$ and $c \in {}^*[a,b)$, there exists some $t \in (c, c+ndx)$ such that $\Delta^n\, {}^*f(c, c+dx) = {}^*f^{(n)}(t)\, (dx)^n$.*

Theorem 8.4.1 holds if the hypotheses are appropriately altered to $f^{(n-1)}\colon [a - na, b] \to \mathbb{R}$ and that $f^{(n)}\colon (a - na, b) \to \mathbb{R}$. In this case, Corollary 8.4.1.1 may be altered to $dx \in \mu(0),\ dx < 0,\ c \in {}^*(a,b],\ t \in (c + ndx, c)$ and $\Delta^n\, {}^*f(c + dx, c) = {}^*f^{(n)}(t)\, (dx)^n$.

Theorem 8.4.2. *Let $1 \leq n \in \mathbb{N}$. Suppose that $f^{(n-1)}\colon [a,b] \to \mathbb{R}$ and that $f^{(n)}\colon (a,b) \to \mathbb{R}$. If $c \in (a,b)$, then for each $dx \in \mu(0),\ dx \geq 0$ [resp. $dx < 0$]*

$$f^n(c)\, (dx)^n \approx \Delta^n\, {}^*f(c, c+dx), [\text{resp.} f(c+dx, c)] \pmod{o((dx)^n)}.$$

One important aspect of Theorem 8.4.2 is that the quantities $f^n(c)\, (dx)^n$ and $\Delta^n\, {}^*f(c, c+dx)$ are not simply infinitely close; but, rather, are infinitely closed of order n. For an infinitesimal, dx, Robinson and those that founded the infinitesimal calculus consider $\Delta^n\, {}^*f(c, c+dx)$ to be the

nth order differential of f at c. However, most other authors still retain the notion that $d^n f(c) = f^{(n)}(c)\,(dx)^n = f^{(n)}(c)\,dx^n$; which I shall retain as well. Under the hypotheses of Theorem 8.4.2 it follows for nonzero positive [resp. negative] infinitesimal, dx, that $d^n f(c)/dx^n \approx \Delta^n\, {}^*f(c, c + dx), [\text{resp.} f(c + dx, c)] = {}^*f^{(n)}(t)$ and, hence, $\mathtt{st}(d^n f(c)/dx^n) = \mathtt{st}(\Delta^n\, {}^*f(c, c + dx), [\text{resp.} f(c + dx, c)]) = f^{(n)}(c)$.

Finally, all that has been said about nth order ideals is extended to m-dimensional objects of the form $o^n(\epsilon_1, \ldots, \epsilon_m) = o(\epsilon_1) \times \cdots \times o(\epsilon_m)$.

8.5. Microgeometry - Tangents to Curves.

A tangent to a curve $c\colon [a, b] \to \mathbb{R}^p$ is usually defined as the intuitive limit of a set of secants. How might this be viewed within the NSP-world? In order to analytically answer this question, let c be differentiable at $t \in (a, b)$. Next, let any $r \in {}^*[-1, 1]$. Fix $\epsilon \in \mu(0)^+$ and let f_i be the coordinate functions of c. The Fundamental Theorem of differential calculus states, in infinitesimal form, that there exists some $h_i \in \mu(0)$ such that

$$ {}^*f_i(t + r\,\epsilon) = f_i(t) + f_i'(t)(r\,\epsilon) + h_i(r\,\epsilon). \tag{1}$$

For any $r_1 \in {}^*[-1, 1] - \mu(0)$, equation (1) then yields, since $o(r_1\,\epsilon) = o(\epsilon)$

$$ {}^*c(t + r_1\,\epsilon) = c(t) + c'(t)(r_1\,\epsilon) + (h_1, \ldots, h_p)(r\,\epsilon), \tag{2}$$

$$ {}^*c(t + r_1\,\epsilon) \approx c(t) + c'(t)(r_1\,\epsilon) \pmod{o^n(\epsilon)}. \tag{3}$$

Each component of $o^n(\epsilon)$ is a first-order ideal and what follows next is a direct result of this fact and the concept of the **resolving power of a microscope.** Expression (1) is first transformed into an external relation determined by

$$ g_i(t + r\,\epsilon) = f_i(t) + f_i'(t)(r\,\epsilon), \tag{4}$$

$$ \vec{g}(t + r\,\epsilon) = c(t) + c'(t)(r\,\epsilon). \tag{5}$$

Assume that we are "looking at" the geometric NSP-world situation with an infinite powered microscope with "first-order ϵ-resolving power." Physically, this corresponds to the idea that if the distance between two objects is a member of a first-order ideal $o(\epsilon)$, then the objects cannot be resolved (i.e. cannot be distinguished one from the other.) A simple proof shows that if $r \in {}^*[-1, 1]$, then the best we can say is that

$$ \| {}^*c(t + r\,\epsilon) - \vec{g}(t + r\,\epsilon)\| \in o(\epsilon). \tag{6}$$

Equation (6) does not mean that for selected $r\,\epsilon$ the value $\|c(t + r\,\epsilon) - \vec{g}(t + r\,\epsilon)\|$ may not be in an nth order ideal. Indeed, if $r = \epsilon^{n-1}$, then this would be the case. Thus (6) means that, in general, $o(\epsilon)$ is the "smallest" order ideal that can be guaranteed to contain this value in all possible cases. For an "ϵ-infinitesimal microscope" (abbr: ϵ-IM) with first-order ϵ-resolving power, (6) implies that within the field of view

$$ {}^*c(t + r\,\epsilon) = \vec{g}(t + r\,\epsilon) = c(t) + c'(t)(r\,\epsilon). \tag{7}$$

A microscope is suppose to magnify, however. How is this feat accomplished. The point, $c(t)$, is being considered as the center of view of the ϵ-IM and this point is translated to the origin.

This gives us the expression ${}^*c_1(t + r\,\epsilon) = {}^*c(t + r\,\epsilon) - c(t) = c'(t)(r\,\epsilon)$. The magnification and resolving power of ordinary microscopes are related. Hence, for consistency, consider for any $x_j \in {}^*\mathbb{R}$ and $\delta_j \in \mu(0)$ the general **infinite magnification operator** $m(x_1\,\delta_1, \ldots, x_p\delta_p) = (x_1, \ldots, x_p)$, which is also assumed to be linear. If you wish to specify a specific shaped field of view, say a p-dimensional closed sphere, you can also restrict this magnification to $\sum_{j=1}^{p} x_j^2 \leq 2$. Obviously, both the translation and magnification can be combined into one operator. Letting, as before, $r \in {}^*[-1, 1]$ the final view in the ϵ-IM only shows the hyperline segment $\{m(f_i'(t)r\,\epsilon, \ldots, f_p'(t)r\,\epsilon) | r \in {}^*[-1, 1]\} = \{(f_1'(t)r, \ldots, f_p'(t)) | r \in {}^*[-1, 1]\} = \{c'(t)r | r \in {}^*[-1, 1]\}$.

8.6 Microgeometry - Surface Elements.

The well-known difficulties of determining a single Euclidean configuration as an appropriate approximation for the surface of a 3-dimensional object will not be discussed in this manual. [See `Cesari [1956]`] Instead, I pass directly to analytical considerations with the appropriate constraints. Let open $G \subset \mathbb{R}^2$. Assume that $\vec{r}\colon G \to \mathbb{R}^3$ and that for $(u, v) \in G$ the continuous partial derivatives $\vec{r}_u(u, v)$, $\vec{r}_v(u, v)$ exist. Let $\vec{r}_u(u_0, v_0) = \vec{a}$, $\vec{r}_v(u_0, v_0) = \vec{b}$ and $|\vec{a} \times \vec{b}| \neq 0$. In order to be consistent with requirement IR2, define, in matrix notation on column vectors, the linear transformation $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ as follows:

$$L = \begin{pmatrix} r_1(u_0, v_0) & a_1 & b_1 \\ r_2(u_0, v_0) & a_2 & b_2 \\ r_3(u_0, v_0) & a_3 & b_3 \end{pmatrix}. \tag{1}$$

In the previous section, a hyperline segment portion of the tangent line was viewed within an IM. The same procedures are now applied to the tangent plane to the surface generated by $\vec{r}$. Let $(s, t) \in {}^*[-1, 1] \times {}^*[-1, 1] = I$; $\delta, \epsilon \in \mu(0)^+$. Then the set $R_S = \{(u, v) \mid (u, v) = (u_0 + s\,\delta, v_0 + t\,\epsilon) \wedge (s, t) \in I\}$ is an internal infinitesimal rectangle and $R_S \subset \mu((u_0, v_0)) \in {}^*G$. The linear transformation ${}^*L$ transforms the internal set of vectors $\{(1, u - u_0, v - v_0)^T \mid (u, v) \in R_S\} = D$ onto a configuration ${}^*L(D)$ which is a *-Euclidean hyperparallelogram containing $\vec{r}(u_0, v_0)$. When considered restricted to D, this linear transformation can be expressed by

$$\vec{k}(u, v) = \vec{r}(u_0, v_0) + (u - u_0)\vec{a} + (v - v_0)\vec{b}. \tag{2}$$

Using the *-Euclidean concepts, this hyperparallelogram is a NSP-world portion of the standard tangent plane to the surface at $\vec{r}(u_0, v_0)$. Two adjacent sides of this hyperparallelogram have end points $\vec{k}(A) = \vec{r}(u_0, v_0) + (-\delta)\vec{a} + (-\epsilon)\vec{b}$; $\vec{k}(B) = \vec{r}(u_0, v_0) + (-\delta)\vec{a} + (\epsilon)\vec{b}$; $\vec{k}(C) = \vec{r}(u_0, v_0) + (\delta)\vec{a} + (-\epsilon)\vec{b}$. This yields two hyperline segments of hyperlength $|\vec{AB}| = \|2\delta\,\vec{a}\|$ and $|\vec{AC}| = \|2\epsilon\,\vec{b}\|$, with the hyperarea being $\|(4\delta\,\epsilon)\vec{a} \times \vec{b}\| = 4\delta\,\epsilon\|\vec{a} \times \vec{b}\| = \|\vec{a} \times \vec{b}\|\,dX$.

Now to analyze the order ideal relationship between members of this hyperparallelogram and the surface itself, assume that $\vec{r}_u$ and $\vec{r}_v$ are continuous at (u_0, v_0). The Fundamental Theorem of Differential Calculus in infinitesimal form, implies that for any $s\,\delta$ and any $t\,\epsilon$, $(s, t) \in I$ there exists $\vec{\eta} \in \mu^3(\vec{0})$ such that

$${}^*\vec{r}(u_0 + s\,\delta, v_0 + t\,\epsilon) = \vec{r}(u_0, v_0) + s\,\delta\,\vec{a} + t\,\epsilon\,\vec{b} + \|(s\,\delta, t\,\epsilon)\|\,\vec{\eta}. \tag{3}$$

Considering any $(s, t) \in {}^*[-1, 1] \times {}^*[-1, 1]$ a simple proof yields that

$$\|{}^*\vec{r}(u_0 + s\,\delta, v_0 + t\,\epsilon) - {}^*\vec{k}(u_0 + s\,\delta, v_0 + t\,\epsilon)\| \in o(\max\{\delta, \epsilon\}) = o(\lambda). \tag{4}$$

The magnification operator for our (δ, ϵ)-IM is the mapping $m(x_1\,\delta + y_1\,\epsilon, \ldots, x_3\,\delta + y_3\,\epsilon) = m(x_1 + y_1, \ldots, x_3 + y_3)$, where the x's and y's are hyperreal numbers. Translating and magnifying

${}^*k(u_0+s\,\delta, v_0+t\,\epsilon)$ yields as the final view in the (δ,ϵ)-IM the hyperparallelogram $\{s\vec{a}+t\vec{b} \mid (s,t) \in I.\}$ **BUT,** these results are more significant than a simple exercise in analyzing the IM view of the tangent plane.

Thus far we have decided upon three types of geometric elements.

(A) Rectifiable curves $\Leftrightarrow$ **hyperline segments** and **hyperpolygonal curves.**

(B) For 1 - 3 dimensional geometric measures by means of the 1-dimensional integral $\Leftrightarrow$ **hypertrapezoids.**

(C) For n-dimensional integrals $\Leftrightarrow$ **infinitesimal rectangles.**

These elements are intuitive in character and only by means of a restrictive analytical description are physical quantities relative to them, such as mass, infinitesimal energy and the other applications given in this manual, actually calculable by means of the integral. However, as evident from all that has preceded, **most individuals consider the integral as but a hyperfinite sum of entities that are intuitively defined** and do not, generally, concerned themselves with the difficulties in calculation. I have been slightly restrictive in some of the basic definitions by requiring that functions that generate hyperfinite sums, at least, be bounded. Of course, in certain cases this restriction might be relaxed. Indeed, in the older literature, geometers utilized intuitive infinitesimal geometry and these notions were not expressed originally in terms of any such analytical constraints.

If $\mathcal{S}$ denotes the geometric point-set called a surface, then ${}^*\mathcal{S}$ is the hypersurface. If, intuitively, $T \subset \mathcal{S}$ is the set of surface points at which tangent planes $\mathcal{T}$ to $\mathcal{S}$ exist, then ${}^*T$ is the set of points in ${}^*\mathcal{S}$ at which the hypertangent planes ${}^*\mathcal{T}$ exist. For the surface integral the geometric element - **the surface element**, σ - is an infinitesimal parallelogram containing a point from the hypersurface. This surface element is considered to be contained in a hypertangent plane. As such this element has an infinitesimal area $d\sigma$. **From the viewpoint of infinitesimal modeling this description of the geometric surface element is adequate.** From the view point of surface integral calculation, since there are infinitely many geometric surface elements of different area, such a description is not sufficient.

The only question that remains is which collection of surface elements should be required for an analytical definition? For this elementary manual, the *-Euclidean area notion will be maintained along with a fixed set of vectors normal to the hypertangent planes. Let $E \subset \mathbb{R}^2$ and bounded $\vec{r}\colon E \to \mathbb{R}^3$. Suppose that $\mathcal{S} = \{\vec{r}(u,v) \mid (u,v) \in E\}$ and that for nonempty $T' \subset E$ the set of points $\vec{r}\,[T'] = T$ is called a set of **tangent points to the surface** $\mathcal{S}$. Further, there exists a mapping $\vec{\nu}\colon T \to (\mathbb{R}^3 - \{\vec{0}\})$, where each $\nu(\vec{t}\,)$ is called a **normal vector** to the surface $\mathcal{S}$. Then each $\vec{t} \in T$ defines a unique **tangent plane** $\mathcal{T}(\vec{t}, \vec{\nu}\,(\vec{t}\,)) = \{(x_1,\ldots,x_3) \mid (x_1,\ldots,x_3) \in \mathbb{R}^3 \wedge \vec{\nu}\,(\vec{t}\,) \bullet ((x_1,\ldots,x_3) - \vec{t}\,) = 0\} \subset \mathbb{R}^3$. Assume that $E \subset R \subset \mathbb{R}^2$. If P' is any partition of R, then there exists a nonempty finite set of subrectangles $R'_i \subset R$ such that $T' \cap R'_i \neq \emptyset,\ 1 \leq i \leq k$ and for each such R'_i there exist the intermediate partitions $Q' = \{t'_1,\ldots,t'_k\}$ such that $t'_i \in R'_i \cap T'$. Hence, for any fine partition P of ${}^*R$ there exists a hyperfinite set of infinitesimal subrectangles $\tau(P) = \{S \mid S \cap {}^*T' \neq \emptyset \wedge S \in P\}$ and a corresponding set of internal intermediate partitions $\eta(P)$. These ideas and notations are used in the next definition.

Definition 8.6.1. (Surface Integral.) Let the surface with its tangent planes be defined as in the above paragraph and assume that bounded $F: \mathcal{S} \to \mathbb{R}$. Let $E \subset R \subset \mathbb{R}^2$. Then F is said to be **SURFACE INTEGRABLE with respect to** $\mathcal{S}$ if there exists some simple fine partition P of R and some $r \in \mathbb{R}$ such that for each intermediate partition $Q \in \eta(P)$

$$\sum_{S \in \tau(P),\ t' \in S \cap Q} {}^*F({}^*\vec{r}(t'))\| {}^*\vec{\nu}({}^*\vec{r}(t'))\| {}^*v(S) \in \mu(r).$$

[**Remark.** In the definition of the line integral, it is required that all fine partitions be considered. This was done so that consideration could be given to rectifiable curves that need not be continuously differentiable. It is obvious that Definition 8.6.1. is styled solely for the integral as defined by 5.1.1.]

Since it is clear from Theorem 7.2.2 and Definition 5.1.1 that for F to be integrable with respect to $\mathcal{S}$ the values ${}^*F({}^*\vec{r}(t'))\| {}^*\vec{\nu}({}^*\vec{r}(t'))\|$ must be obtained from an integrable function defined on an appropriate Jordan-measurable $J \subset R \subset \mathbb{R}^2$, I see no need to state the various well-known functions that lead to this conclusion - with one exception. Let compact $J \subset R \subset \mathbb{R}^2$ and bounded $\vec{r}: J \to \mathbb{R}^3$ generate a surface $\mathcal{S}$. How do we generalize the concept of Euclidean area to such a surface?

Suppose that $\vec{r}(u,v) = \vec{p} + u\,\vec{a} + v\,\vec{b}$, where $\vec{p},\ \vec{a} \neq \vec{0},\ \vec{b} \neq \vec{0}$ are fixed members of $\mathbb{R}^3$ and $(u,v) \in [a,b] \times [c,d] = J$. Then $\vec{r}[J]$ is a paralelogram with area equal to $(d-c)(b-a)\|\vec{a} \times \vec{b}\|$. But, $\vec{r}_u = \vec{a}$ and $\vec{r}_v = \vec{b}$. Assume that $K = \text{int}(J) - W$, where W is a set of Jordan-content zero. Then taking other simple Euclidean surfaces generated by corresponding simple sets of surface equations leads to the conclusion that it should be required for functions $\vec{r}: J \to \mathbb{R}^3$ that $\vec{r}_u(u,v)$ and $\vec{r}_v(u,v)$ exist and $\|\vec{r}_u(u,v) \times \vec{r}_v(u,v)\| \neq 0$ for $(u,v) \in K$, if area preserving surface integration is desired. This is precisely the starting point for elementary analysis. Notice that the linear transformations that correspond to IR2 are pointwise defined and each is considered to generate only the specific hyperparallelogram associated with a specific term in the hyperfinite sums used in Definition 8.6.1. I leave it to the reader to combine the elemental method of constants on surface elements with Definition 8.6.1 to obtain the concept of flux of a vector field across a surface.

8.7 Microgeometry - Other Stuff.

What about the geometry of the coordinate transformations? In general, it appears necessary to consider stronger constraints then for surface integration. If a map $f: R \to \mathbb{R}^m$, where $R \subset \mathbb{R}^m$, is to be considered a coordinate transformation, then it needs to be considered locally infinitely close to a linear transformation determined by a nonzero Jocobian. What this signifies is that each m-dimensional infinitesimal subrectangle is mapped onto an infinitesimal parallelepiped - a hyperparallelepiped. Thus the infinitesimally partitioned space is distorted into a space of hyperparallelepipeds. All the details as to why this appears necessary can be found in `Stroyan and Luxemburg` beginning in section 5.8 [110].

In the above mentioned section of the book by Stroyan and Luxemburg, you will also find in section 5.8 the infinitesimal approach to orientated partitions via differential forms. In section 5.9, the authors investigate the infinitesimal calculus on manifolds. In all of these sections, the authors extend the integral concept to internal functions that usually preserve monads. Thus their integral concept, although it is the same as has been presented here for nonstandard extensions of standard functions, is actually defined on a wider class of objects.

8.8 Gauge Integrals

Recently, additional emphases has been given to a generalization of the Riemann sum called the **gauge integral** [`Swartz and Thomson [1988]`]. This integral concept is also called the Riemann-complete integral, [`Henstock [1961]`]; or S-integral, [`Mawhin [1985]`]; among other names. The gauge integral can also be generalized to the `Jarnik`, `Kurzweil`, `Schwabik [1983]` integral (the **M-integral**, Mawhin [1985].) Indeed, a Riemann sum styled generalization leads to the classical Lebesgue integral [`McShane [1973]`]. From the infinitesimal viewpoint, all of these generalizations are very similar.

For the S and M-integrals, a positive real valued function, δ, called a **gauge** is utilized to select not only a special partition but a special intermediate partition as well.

Within the NSP-world a set of gauges determines an internal mapping $\hat{\delta}$ from ${}^*R \subset {}^*\mathbb{R}^m$ into $\mu(0)$ called a **microguage.** However, associated with every gauge, δ, is an object $(Q_\delta, P_\delta) = \{(t_1, S_1), \ldots, (t_k, S_k)\}$ where $P_\delta = \{S_1, \ldots, S_k\}$ is a subdivision of R by nonoverlapping subsets and $Q_\delta = \{t_1, \ldots, t_k\}$ is a partial sequence where each $t_i \in R$. Such an object (Q_δ, P_δ) called a **δ-fine partition**. Thus for a microguage $\hat{\delta}$ there exists a hyperfinite subdivision $P_{\hat{\delta}} = \{S_1, \ldots, S_\Gamma\}$ of ${}^*R$ and an hyperfinite intermediate partition $Q_{\hat{\delta}} = \{t_1, \ldots, t_\Gamma\}$. The pair $(Q_{\hat{\delta}}, P_{\hat{\delta}})$ is called a **micropartition.** I point out that each member of $P_{\hat{\delta}}$ is an infinitesimal subrectangle in this case. For a function $f\colon R \to \mathbb{R}^k$, Mawhin [1986] shows that *f is S-integrable if and only if there exists an $\vec{r} \in \mathbb{R}^k$ such that for each micropartition $(Q_{\hat{\delta}}, P_{\hat{\delta}})$ of ${}^*R$*

$$\sum_{i=1}^{\Gamma} {}^*f(t_i)\ {}^*v(S_i) \in \mu(\vec{r}).$$

Mawhin gives a completely similar characterization for the M-integral but replaces the general micropartition with the more specialized regular micropartition.

Looking at another type gauge, which I shall call the **L-gauge**, λ, and a corresponding partition pair (Q_λ, P_λ), where it is <u>not</u> assumed that if $t_i \in Q_\lambda$, then $t_i \in S_i \in P_\lambda$, McShane [1973] defines his integral by the same Riemann sum technique. The **McShane integral** is equivalent to the Lebesgue integral. Using the same technique employed by Mawhin, the following is established in appendix 8.

Theorem 8.8.1. *A bounded function $f\colon R \to \mathbb{R}$ is Lebesgue integrable with value $r \in \mathbb{R}$ if and only if for each L-micropartition $(Q_{\hat{\lambda}}, P_{\hat{\lambda}})$ of ${}^*R$*

$$\sum_{i=1}^{\Gamma} {}^*f(t_i)\ {}^*v(S_i) \in \mu(r).$$

Thus from the infinitesimal viewpoint there is no basic difference between any of these integral concepts discussed in this section. **They are all hyperfinite sums**, **where each term is the scalar product of a limited object (the value of a bounded function) by the volume of an infinitesimal subrectangle.**