A Formal Proof Of The Riesz Representation Theorem

Anthony Narkawicz
NASA Langley Research Center, Hampton, VA 23681-2199

This paper presents a formal proof of the Riesz representation theorem in the PVS theorem prover. The Riemann Stieltjes integral was defined in PVS, and the theorem relies on this integral. In order to prove the Riesz representation theorem, it was necessary to prove that continuous functions on a closed interval are Riemann Stieltjes integrable with respect to any function of bounded variation. This result follows from the equivalence of the Riemann Stieltjes and Darboux Stieltjes integrals, which would have been a lengthy result to prove in PVS, so a simpler lemma was proved that captures the underlying concept of this integral equivalence. In order to prove the Riesz theorem, the Hahn Banach theorem was proved in the case where the normed linear spaces are the continuous and bounded functions on a closed interval. The proof of the Riesz theorem follows the proof in Haaser and Sullivan's book *Real Analysis*. The formal proof of this result in PVS revealed an error in textbook's proof. Indeed, the proof of the Riesz representation theorem is constructive, and the function constructed in the textbook does not satisfy a key property. This error illustrates the ability of formal verification to find logical errors. A specific counterexample is given to the proof in the textbook. Finally, a corrected proof of the Riesz representation theorem is presented.

1. INTRODUCTION

Recently, the formal methods research group at NASA Langley has undertaken the challenge to formalize probabilistic models of aircraft behavior in the airspace. This has resulted in a project to develop a probability theory library in PVS. The beginning of such a library was started by David Lester, who has formalized the theory of Lebesgue integration. One challenging factor is that the formal theories should be capable of proving both (1) results with actual numbers and (2) theorems in general probability theory. A measure theory approach to probability lends itself more to proving general theorems, while the Riemann integral lends itself more to proving results with actual numbers. Thus, an alternative that is being explored is the Riemann-Stieljtes integral, which allows for both discrete and continuous distributions to be defined and makes numerical approximations of probabilities at least conceivable. Recently, the author has defined and proved basic properties of the Riemann-Stieltjes integral. This paper presents a result involving the Riemann-Stieltjes integral, which is not basic, although it is very well-known by any graduate student in functional analysis.

The Riesz representation theorem is a fundamental theorem in functional analysis. J.D. Gray's gripping exposé on the history of the theorem begins with the following remark:

> "Only rarely does the mathematical community pay a theorem the accolade of transforming it into a tautology. The Riesz representation theorem has received this accolade." [Gra84]

By this, Gray means that the theorem has become so embedded in the minds of todays mathematicians that they often implicitly assume it and do not bother to note its application. This theorem is useful in many areas of analysis, including probability theory, where it is used to prove the existence of solutions to moment problems. This paper presents a formal proof of the theorem, as originally stated in 1909 by Riesz [Rie09], in the PVS theorem prover. The PVS development can be downloaded at http://shemesh.larc.nasa.gov/people/ajn/pvs_development. The proof uses standard techniques found in many textbooks, but most of the construction is taken from Haaser and Sullivan's *Real Analysis* [HS71].

The Riesz representation theorem (henceforth called the Riesz theorem) classifies the bounded linear functionals on the space $C[a, b]$, of continuous functions on the closed, bounded interval $[a, b]$. A linear functional on $C[a, b]$ is a linear transformation $L\colon C[a, b] \to \mathbb{R}$, and it therefore satisfies the following two properties.

$$\text{(1)} \quad \text{For all } \ f, g \ \in C[a, b]\colon \ L(f + g) = L(f) + L(g)$$
$$\text{(2)} \quad \text{For all } \ c \ \in \mathbb{R}, \ f \in C[a, b]\colon \ L(c \cdot f) = c \cdot L(f)$$

The linear functional L is bounded if there exists a nonnegative real number M such that for all $f \in C[a, b]$,

$$|L(f)| \leq M \cdot \|f\|_{sup}, \tag{1}$$

where the (supremum) norm $\|g\|_{sup}$ of an arbitrary bounded function g on $[a, b]$ is defined by

$$\|g\|_{sup} \equiv \sup_{x \in [a,b]} |g(x)|. \tag{2}$$

This norm is defined for functions $f \in C[a, b]$ because any continuous function on a closed interval is necessarily bounded [HS71]. If L is a bounded linear functional on $C[a, b]$, then the operator norm $\|L\|_{C[a,b]}$ of L is the smallest real number M that satisfies Equation (1). It is easy to see that for any such L and any continuous function f on $[a, b]$, $|L(f)| \leq \|L\|_{C[a,b]} \cdot \|f\|_{sup}$.

The Riesz theorem classifies the bounded linear functionals on $C[a, b]$ in terms of the set $BV[a, b]$ of functions of bounded variation on $[a, b]$. A function $g\colon [a, b] \to \mathbb{R}$ is of bounded variation if there exists a real number K such that for every $n > 0$ and every partition $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$, $\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| \leq K$. The total variation $V_a^b(g)$ of a function $g \in BV[a, b]$ is the minimum of all such real numbers K. If $g \in BV[a, b]$, then there is an associated bounded linear functional I_g on $C[a, b]$ given by

$$I_g(f) \equiv \int_a^b f dg, \tag{3}$$

which is the Riemann-Stieltjes integral of f with respect to g (Section 2). In the construction of this integral and the linear functional in PVS, the original intent was to simply generalize the construction of the Riemann integral in the NASA PVS libraries that was created by Ricky Butler [But09]. Ideally, it would have been possible to generalize Butler's proof that a continuous function is Riemann integrable to a proof that the functional I_g is indeed well-defined. However, a new proof had to be developed, because Butler's proofs were based on step functions, which are not necessarily Riemann-Stieltjes integrable (Section 2).

A result needed in Haaser and Sullivan's proof of the Riesz theorem is the equivalence of the Riemann-Stieltjes and Darboux-Stieltjes integrals. This result is used to prove the lemma that a continuous function is Riemann-Stieltjes integrable with respect to any function of bounded variation. In the proof of the Riesz theorem in PVS, a simpler method was used to prove the integrability of continuous functions. This method is to prove one theorem that extracts the main idea of this integral equivalence and can be used to prove the integrability of continuous functions. The proof of this theorem is somewhat tricky due to its use of refinements of partitions and inequalities between finite sums.

The Riesz theorem says that every bounded linear functional on $C[a, b]$ is given by I_g for some $g \in BV[a, b]$.

Theorem 1.1. (The Riez Representation Theorem; Theorem 6.1 in [HS71]) If L is a bounded linear functional on $C[a, b]$, then there exists a function $g \in BV[a, b]$ such that $L = I_g$ and $\|L\|_{C[a,b]} = V_a^b(g)$.

The proof of the theorem is constructive. In fact, the function g such that $L = I_g$ is given by

$$g(x) = \begin{cases} 0 & \text{if } x = a \\ \bar{L}(\chi_{[a,x)}) & \text{if } a < x < b \\ \bar{L}(\chi_{[a,x]}) & \text{if } x = b \end{cases} \tag{4}$$

where $\bar{L}$ is an extension of L to the space $B[a, b]$ of bounded functions on $[a, b]$, and where, for a given interval Int, χ_{Int} denotes the characteristic function of Int.

Typically, constructing a formal proof of a difficult result in a mathematics textbook will reveal some special cases that were overlooked by the author. In the case of the Riesz theorem, the distinction between the formal PVS proof and the informal textbook proof was even greater than this, because the textbook proof in question [HS71] constructs the function g incorrectly. The definition given there of the function g is

$$g_{inc}(x) = \begin{cases} 0 & \text{if } x = a \\ \bar{L}(\chi_{[a,x)}) & \text{if } a < x \leq b \end{cases} \tag{5}$$

While the definitions of the functions g and g_{inc} are almost identical, the key property of g that is needed in the proof is not satisfied by the function g_{inc} (Section 5.1). A person trained in Lebesgue integration theory may look at the functions g and g_{inc} and assume that they are essentially equivalent, since the only difference occurs at the single point b. However, the Riemann-Stieltjes integral is different from the Lebesgue integral in that it is not true that two functions which differ only at one point have the same Riemann-Stieltjes integral, even if they are both integrable.

One difficult part of the proof of the Riesz theorem is verifying that any bounded linear functional L on $C[a, b]$ can be extended to a bounded linear functional $\bar{L}$ on $B[a, b]$ such that $\|L\|_{C[a,b]} = \|\bar{L}\|_{B[a,b]}$. This statement is an instantiation of the Hahn-Banach theorem, another fundamental theorem in functional analysis that is well known to graduate students in analysis. The Hahn-Banach theorem says that if X and Y are real normed linear spaces such that $X \subset Y$, and if $H : X \to \mathbb{R}$ is a bounded linear functional on X, then there exists an extension, $\bar{H} : Y \to \mathbb{R}$, of

H such that $\|H\|_X = \|\bar{H}\|_Y$. The following case of the Hahn-Banach theorem has been formally proved in PVS.

Theorem 1.2. (Special Case of Hahn-Banach) If L is a bounded linear functional on $C[a,b]$, then there exists an extension, $\bar{L}\colon B[a,b] \to \mathbb{R}$, of L such that $\|L\|_{C[a,b]} = \|\bar{L}\|_{B[a,b]}$.

Formal proofs of the complete Hahn-Banach theorem have been completed before [NT93, BW00, Bau01]. As in the case of the Riesz theorem, the proof follows the informal textbook proof in [HS71] (Theorem 5.6 and Corollary 5.7).

There is a critical step in the proof of the Hahn-Banach theorem in which one must show that if H is a bounded linear functional on a subspace S of $B[a,b]$, and if a function $f \in B[a,b]$ is not in S, then there exists an extension $\bar{H}$ of H to the subspace $\bar{S} = \{s + c \cdot f \mid s \in S,\ c \in \mathbb{R}\}$ of S. The extension $\bar{H}$ is defined by $\bar{H}(s+c\cdot f) = H(s)+c\cdot\beta$, where β is a fixed real number satisfying a certain property. Defining $\bar{H}$ directly in this way is not possible in PVS, so specific functions had to be defined that, given an element $r = s+c\cdot f$ of $\bar{S}$, compute the function s and the constant c, and the linearity and uniqueness of these functions had to be verified.

As usual, the formal PVS proofs of the Riesz and Hahn-Banach theorems were much longer and more detailed than the informal textbook proofs. The only exception was in the proof that a continuous function is Riemann-Stieltjes integrable with respect to any function of bounded variation. As mentioned above, Haaser and Sullivan's textbook proves this by first constructing the Darboux-Stieltjes integral and proving its equivalence to the Riemann-Stieltjes integral. The formal proof in PVS takes a simpler approach by proving one theorem that captures the needed underlying concept from this integral inequality.

1.1 Acknowledgements

The author would like to thank César A. Muñoz and Ricky Butler for their helpful comments and corrections to this paper.

2. RIEMANN-STIELTJES INTEGRABILITY OF CONTINUOUS FUNCTIONS

This section presents the formalization of the Riemann-Stieltjes (RS) integral and the proof that continuous functions on $[a,b]$ are integrable with respect to any function of bounded variation on $[a,b]$.

2.1 Functions on $[a,b]$

There are three specific spaces, each consisting of real-valued functions on $[a,b]$, whose properties are used in the Riesz representation theorem and its proof. The spaces are $C[a,b]$, which consists of all continuous functions on $[a,b]$, $B[a,b]$, which consists of all bounded functions on $[a,b]$, and $BV[a,b]$, which consists of all functions of bounded variation on $[a,b]$. The PVS theories in which the Riemann-Stieltjes integral is defined and its basic properties are proved all have a parameter T, which is a subtype of the real numbers. The type T must be connected and contain more than one element, and it is therefore an interval which can be infinite in length and may or may not contain either endpoint. In the PVS theories in which the Hahn-Banach and Riesz-Representation theorems are proved, the real numbers a and b are parameters of the theories, and a and b are restricted by their types to satisfy $a < b$. In these theories that have a and b as parameters, `INTab` refers to the type consisting of all real numbers in the interval $[a, b]$, and `Intab` (with lower-case n and t) refers to the set $[a, b]$.

2.2 Partitions on $[a, b]$

The formalization of the RS integral defines a partition using the finite sequence theories in the NASA PVS libraries. In these theories, a finite sequence is a record type with two fields, a natural number `length` and an infinite sequence `seq` of real numbers that evaluates to a default value for all indices greater than `length-1`. In PVS, elements of this finite sequence are written using the form (# `length := n, seq := f` #), where `n` is a natural number and `f` is a function from natural numbers into real numbers.

The formal definition of a partition in the RS theories is given as follows. It is defined by a predicate `partition_pred?` on finite sequences that tests whether a given finite sequence is a partition of the interval $[a, b]$.

```
partition_pred?(a:T,b:{x:T|a<x})(fs:fseq): MACRO bool =
                (Let N = fs`length, xx = fs`seq IN
                N > 1 AND xx(0) = a AND xx(N-1) = b AND
                increasing?(fs) AND
                (FORALL (i:below(N)): a <= xx(i) AND xx(i) <= b))

partition(a:T,b:{x:T|a<x}): TYPE = (partition_pred?(a,b))
```

In this definition, `T` is an interval of real numbers with positive length, and the notation `b:{x:T|a<x}` indicates that `b` has type `{x:T|a<x}`. This type declaration is possible in PVS because PVS allows predicate subtyping. The notation ``fs`length`` indicates record access. Here, `fs` is a finite sequence, and either of its fields, `length` and `seq` can be accessed by the notation ``f`length`` and ``f`seq``, respectively. The final condition in this definition, which states that every element of a partition is an element of the interval $[a, b]$, is implied by the other conditions in the definition. However, this condition was added to the partition type because many of the proofs required this statement to be verified, and adding it here allowed the type-checker in PVS to verify this condition automatically in the proofs.

Lemmas in PVS for partitions of the interval $[a, b]$ have been available for some time in the NASA PVS libraries and were developed by Ricky Butler [But09]. When formalizing the RS integral, many of the results in those theories could be used with only slight modifications. However, some new constructions were needed for the development of the RS integral, most notably the union of two partitions, which itself is a partition.

```
partition_union(a,(b|a<b))(P,Q: partition(a,b)):
                {PQ: partition(a,b) |
                (FORALL (x:T): member(x,PQ)
                IFF (member(x,P) OR member(x,Q))) AND
                strictly_increasing?(PQ)}
```

The union of two partitions in PVS is an uninterpreted function, and it was proved in PVS that this function exists and is unique. Several other key properties of these unions were proved, including that the width of the union of two partitions is bounded above by the widths of the individual partitions.

```
width(a:T, b:{x:T|a<x}, P: partition(a,b)): posreal =
                LET xx = P`seq, N = P`length IN
                max({ l: real | EXISTS (ii: below(N-1)):
                l = xx(ii+1) - xx(ii)})

partition_union_width: LEMMA a<b IMPLIES
                FORALL (P,Q:partition(a,b)):
                width(a,b,partition_union(a,b)(P,Q)) <=
                min(width(a,b,P),width(a,b,Q))
```

2.3 Summation in PVS

There are multiple places in the construction of the Riemann-Stieltjes integral and the proof of the Riesz representation theorem where summations are used, including the definitions for Riemann-Stieltjes sums and functions of bounded variation. There is extensive support for manipulating summations in the NASA PVS libraries. These PVS theories were developed, in part, for the purpose of defining the standard Riemann integral [But09]. Summation is defined in the NASA libraries with the function `sigma`:

```
sigma(low, high, F): RECURSIVE real
   = IF low > high THEN 0
     ELSE  F(high) + sigma(low, high-1, F)
     ENDIF
     MEASURE (LAMBDA low, high, F: abs(high+1-low))
```

In this definition, `F` is a real valued function whose domain is a subtype of the integers. Most of the standard results involving summation are available in the NASA libraries, such as the triangle inequality.

```
sigma_abs : THEOREM abs(sigma(low, high, F)) <=
     sigma(low, high, LAMBDA (n: T): abs(F(n)))
```

There are also more advanced lemmas that facilitate manipulations of summations, such as the fact that summation signs can be swapped.

```
sigma_swap: LEMMA
     sigma(low1,high1,LAMBDA (i:T):
        sigma(low2,high2,LAMBDA (j:T): F(i,j)))
     = sigma(low2,high2,LAMBDA (j:T):
        sigma(low1,high1,LAMBDA (i:T): F(i,j)))
```

2.4 Definition of the Riemann-Stieltjes integral

The Riemann-Stieltjes integral is a generalization of the Riemann integral, which has already been formalized in PVS [But09]. If f and g are functions $[a, b] \to \mathbb{R}$, then f is said to be *Riemann-Stieltjes integrable* with respect to g on $[a, b]$ if there exists a real number S with the following property. For every $\epsilon > 0$, there exists $\delta > 0$ with the property that for any partition $P:\ a = p_0 \le p_1 \le \cdots \le p_n = b$

with width less than δ, any choice of points $x_i \in [p_i, p_{i+1}]$ for $i < n$ satisfies

$$|S \; - \; \sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1}))| \; < \; \epsilon. \tag{6}$$

If f is RS integrable with respect to g, then the real number S is denoted $\int_a^b f \; dg$. The function f is called the "integrand", and g the "integrator". It is easy to see that if $g(x) = x$ for all $x \in [a, b]$ then this is the Riemann integral of f.

A popular alternative way to define the Riemann-Stieltjes integral is to require that for every $\epsilon > 0$, there exists a partition P, such that any refinement of P satisfies the conditions above. The approach taken in the formalization presented here is also taken by Haaser and Sullivan [HS71]. In fact, the definition chosen in this paper for the Riemann-Stieltjes integral is less general than this second definition, and it is easy to construct an example that is RS integrable for the second definition but not for the first. The author's main motivation for choosing this definition is that the NASA PVS libraries contain a complete set of theories on the Riemann integral. The Riemann integral is a special case of the Riemann-Stieltjes integral, so many of the definitions and proofs in the PVS theories for the Riemann integral were altered to give new definitions and proofs in the PVS theories for the Riemann-Stieltjes integral. However, not all of the results for the Riemann integral could be generalized to the Riemann-Stieltjes integral. This is because Butler's constructions for the Riemann integral [But09] were based on step functions, which are not necessarily Riemann-Stieltjes integrable. The PVS theories for these two integrals are separate and independent. The NASA PVS libraries also contain theories for Lebesgue integration, and the equivalence of Lebesgue and standard Riemann integration has been proved by David Lester.

A change was made (to Butler's constructions) in the definition of a set of points $\{x_0, \ldots, x_{n-1}\}$ such that $x_i \in [p_i, p_{i+1}]$, where the points p_j are in a partition P. Such a set $\{x_0, \ldots, x_{n-1}\}$ is said to *satisfy* the partition P. For the Riemann integral, such sets of points were simply defined as maps from `below(length(P)-1)` to `closed_interval(a,b)`. In the definition of the RS integral, this type, denoted `xis?(a,b,P)`, is also defined in terms of the theories on finite sequences from the NASA PVS libraries. It is defined using a predicate `xis_pred?` on finite sequences that tests whether a sequence satisfies a given partition `P`.

```
xis_pred?(a:T,b:{x:T|a<x},P:partition(a,b))(fs:fseq): MACRO bool =
    (fs`length = P`length-1 AND (FORALL (ii: below(P`length-1)):
    P`seq(ii) <= fs`seq(ii) AND fs`seq(ii) <= P`seq(ii+1)))

xis?(a:T,b:{x:T|a<x},P:partition(a,b)): TYPE = (xis_pred?(a,b,P))
```

As in the case of partitions, this change was not difficult to accommodate in the formal definitions and proofs.

The remainder of the construction of the Riemann-Stieltjes integral is identical to that for the Riemann integral [But09]. The following functions and predicates have been defined in PVS, and the return type is indicated in each case.

```
Rie_sum(a:T,b:{x:T|a<x},g:[T->real],P:partition(a,b),
                xis: xis?(a,b,P),f:[T->real]): real
```

```
Riemann_sum?(a:T,b:{x:T|a<x},P:partition(a,b),g,f:[T->real])
  (S:real): bool =
      (EXISTS (xis: xis?(a,b,P)): LET N = P`length-1 IN
                                 S = Rie_sum(a,b,g,P,xis,f))

integral?(a:T,b:{x:T|a<x},g,f:[T->real],S:real): bool

integrable?(a:T,b:{x:T|a<x},g,f:[T->real]): bool =
                  (EXISTS (S:real): integral?(a,b,g,f,S))

integral(a:T,b:{x:T|a<x},
              gg:[T->real],ff:{f | integrable?(a,b,gg,f)}):
                         {S:real | integral?(a,b,gg,ff,S)}
```

2.5 Functions of Bounded Variation

The Riesz representation theorem classifies the bounded linear functionals on $C[a,b]$ in terms of the set $BV[a,b]$ of functions of bounded variation on $[a,b]$. As noted in the introduction, a function $g\colon [a,b] \to \mathbb{R}$ is of bounded variation if there exists a real number K such that for every partition $a = p_0 \leq p_1 \leq \cdots \leq p_n = b$, $\sum_{i=1}^{n} |g(p_i) - g(p_{i-1})| \leq K$. The total variation $V_a^b(g)$ of a function $g \in BV[a,b]$ is the minimum of all such real numbers K.

The set of functions on $[a,b]$ of bounded variation is defined using a predicate `bounded_variation?` in PVS.

```
variation_on(a,b:{x:T|a<x},P:partition(a,b))(f) : nnreal =
   sigma[below(P`length-1)](0,P`length-2,
                           LAMBDA (n:below(P`length-1)):
                                abs(f(P`seq(n+1))-f(P`seq(n))))

bounded_variation?(a,b:{x:T|a<x})(f): bool = (EXISTS (M: nnreal):
   (FORALL (P:partition(a,b)): variation_on(a,b,P)(f) <= M))

total_variation(a,b:{x:T|a<x},f:(bounded_variation?(a,b)))
                                        (x: (closed_intv(a,b))):
   {M: nnreal|(x=a IMPLIES M=0) AND (x>a IMPLIES
   FORALL (P:partition(a,x)): variation_on(a,x,P)(f)<=M) AND
   (FORALL (M1: nnreal): M1<M IMPLIES EXISTS (P:partition(a,x)):
                                  variation_on(a,x,P)(f) > M1)}
```

The general strategy used in the PVS development, when proving that a property holds for functions of bounded variation, is to first prove that the property holds for functions that are increasing on $[a,b]$, then prove that the property is linear, and finally apply the following lemma.

Lemma 2.1. A function $g\colon [a,b] \to \mathbb{R}$ is of bounded variation on $[a,b]$ if and only if there are increasing functions h and r on $[a,b]$ such that $g = h - r$.

Proof (Sketch). If h is a function that is increasing on $[a, b]$, it is easy to see that it is of bounded variation, since for any partition $P:\ a = p_0 \leq p_1 \leq \cdots \leq p_n = b$, $\sum_{i=1}^{n} |(h(p_i) - h(p_{i-1})| = \sum_i (h(p_i) - h(p_{i-1})) = b - a$. Thus, the only tricky part of the proof is the forward implication, the proof of which is constructive. Indeed, given a function g of bounded variation, the functions h and r are given by $h(x) = V_a^x(g)$ and $r(x) = V_a^x(g) - f(x)$ for $x \in [a, b]$. □

This lemma has been proved in PVS, and it simplifies many proofs about functions of bounded variation. It is stated in PVS as follows.

```
BV_decomposition: LEMMA a<b IMPLIES
     (bounded_variation?(a,b)(f) IFF
     LET CI = closed_intv(a,b) IN
     EXISTS (g,h): f = (LAMBDA (x:T): g(x)-h(x)) AND
                  increasing?[(CI)](g) AND increasing?[(CI)](h))
```

2.6 The Darboux-Stieltjes Integral

A result used in Haaser and Sullivan's proof [HS71] of the Riesz Representation theorem is the equivalence of the Riemann-Stieltjes and Darboux-Stieltjes integrals. This result is used to prove the lemma that a continuous function is Riemann-Stieltjes integrable with respect to any function of bounded variation. In the proof of the Riesz Representation theorem in PVS, a simpler method was used to prove this result. This method is to prove one theorem that extracts the main idea of this integral equivalence and can be used to prove the integrability of continuous functions. The proof of this theorem is somewhat tricky due to its use of refinements of partitions and inequalities between finite sums.

The definition of the Darboux-Stieltjes integral of a bounded function f with respect to a function g requires that g be increasing on $[a, b]$. The basic quantities are defined as follows.

$$\begin{aligned} m_i(f) &= \inf\{f(x) \mid x \in [p_{i-1}, p_i]\} \\ M_i(f) &= \sup\{f(x) \mid x \in [p_{i-1}, p_i]\} \\ L(f, g, P) &= \sum_{i=1}^{n} m_i(f)(g(p_i) - g(p_{i-1})) \\ U(f, g, P) &= \sum_{i=1}^{n} M_i(f)(g(p_i) - g(p_{i-1})) \end{aligned} \tag{7}$$

The function g must be increasing to ensure that $L(f, g, P) \leq U(f, g, P)$. The upper and lower integrals, respectively, are defined as follows.

$$\begin{aligned} \overline{\int_a^b} f\ dg &= \ \inf\{U(f, g, P) \mid P \text{ is any partition on } [a, b]\} \\ \underline{\int_a^b} f\ dg &= \ \sup\{L(f, g, P) \mid P \text{ is any partition on } [a, b]\} \end{aligned} \tag{8}$$

If the upper and lower integrals are equal, then f is Darboux-Stieltjes integrable with respect to g.

The theorem that captures a key idea in the Riemann and Darboux Stieltjes integral equivalence and simplifies the proof of the Riesz theorem in PVS is stated below.

Theorem 2.2. If g is increasing on $[a, b]$, then f is integrable with respect to g if and only if for every $\epsilon > 0$, there exists $\delta > 0$ with the property that for every partition $P : \ a = p_0 \leq \cdots \leq p_n = b$ with width less than δ, and for any two sequences $\{x_0, \ldots, x_{n-1}\}$ and $\{y_0, \ldots, y_{n-1}\}$ that satisfy P,

$$|\sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})) \ - \ \sum_{i=1}^{n} f(y_{i-1})(g(p_i) - g(p_{i-1}))| < \epsilon. \tag{9}$$

It should be noted that this theorem is not the same as the true statement that f is integrable if and only if for all ϵ, there exists δ such that for all pairs of partitions P and Q with widths less than δ, every RS sum of f on P is less than ϵ away (in absolute value) from every RS sum of f on Q. This latter result is trivial since a sequence of real numbers converges if and only if it is Cauchy. Indeed, Theorem 2.2 is a stronger statement because it implies that the P and Q can be assumed to be equal. The proof relies on the following lemma, the proof of which contained the most challenging parts of the formalization of Theorem 2.2.

Lemma 2.3. If g is increasing on $[a, b]$, then for any partitions $P : \{p_0, \ldots, p_n\}$ and $Q : \{q_0, \ldots, q_m\}$ of $[a, b]$ and any sequences $\{x_0, \ldots, x_{n-1}\}$ and $\{y_0, \ldots, y_{r-1}\}$ that satisfy the partitions P and $P \cup Q$, respectively, there exists a sequence $\{z_0, \ldots, z_{n-1}\}$ that satisfies P such that

$$\begin{aligned} &|\sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})) \ - \ \sum_{j=1}^{r} f(y_{j-1})(g(pq_j) - g(pq_{j-1}))| \\ \leq \ &|\sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})) \ - \ \sum_{i=1}^{n} f(z_{i-1})(g(p_i) - g(p_{i-1}))|, \end{aligned} \tag{10}$$

where $P \cup Q = \{pq_0, \ldots, pq_r\}$. That is, $pq_0, \ldots, pq_r$ are the (increasing) elements of the partition $P \cup Q$.

Proof Of Lemma 2.3. The key to the proof is that $\{z_o, \ldots, z_{n-1}\}$ is a subset of $\{y_0, \ldots, y_{r-1}\}$. Suppose that $\sum_i f(x_{i-1})(g(p_i) - g(p_{i-1})) \leq \sum_j f(y_{j-1})(g(pq_j) - g(pq_{j-1}))$. The proof in the case where this inequality does not hold is nearly identical. It is trivial to see that it therefore suffices to find a sequence $\{z_0, \ldots, z_{n-1}\}$ that satisfies P such that $\sum_j f(y_{j-1})(g(pq_j) - g(pq_{j-1})) \geq \sum_i f(z_{i-1})(g(p_i) - g(p_{i-1}))$. A surjective increasing function $\sigma \colon \{0, \ldots, r\} \to \{0, \ldots, n\}$ can be defined such that $[pq_j, pq_{j+1}] \subset [p_{\sigma(j)}, p_{\sigma(j)+1}]$ for all $j < r$. Given $i \leq n-1$, define z_i to be any element of the set $\{y_j \mid \sigma(j) = i\}$ such that $f(z_i) \leq f(y_j)$ for all $j \leq r-1$ such that $\sigma(j) = i$. This is possible because the set in question is finite. Then, because

g is increasing and hence $g(pq_j) \geq g(pq_{j-1})$ for all j,

$$\begin{aligned}\sum_{j=1}^{r} f(y_{j-1})(g(pq_j) - g(pq_{j-1})) &\geq \sum_{j=1}^{r} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1})) \\ &= \sum_{i=1}^{n} \sum_{\{j|\sigma(j-1)=i-1\}} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1})) \\ &= \sum_{i=1}^{n} f(z_{i-1}) \sum_{\{j|\sigma(j-1)=i-1\}} (g(pq_j) - g(pq_{j-1})) \\ &= \sum_{i=1}^{n} f(z_{i-1})(g(p_i) - g(p_{i-1})).\end{aligned} \tag{11}$$

The last equality follows from the fact that

$$[p_{i-1}, p_i] = \bigcup_{\{j|\sigma(j-1)=i-1\}} [pq_j, pq_{j-1}].$$

The formalization of this proof is discussed in Section 2.7. □

Proof Of Theorem 2.2. If f is integrable with respect to g, then the result is trivial from definitions, so suppose that f is not integrable and choose any positive real number ϵ. Suppose that there exists $\delta > 0$ with the property that for every partition $P: \; a = p_0 \leq \cdots \leq p_n = b$ with width less than δ, any two sequences $\{x_0, \ldots, x_{n-1}\}$ and $\{y_0, \ldots, y_{n-1}\}$ such that $x_i, y_i \in [p_i, p_{i+1}]$ satisfy Equation (9) with ϵ replaced by $\epsilon/2$ in that equation. Let P and Q be any two partitions of $[a, b]$, with $n+1$ and $m+1$ elements, respectively, and suppose that the widths of P and Q are less than δ, and suppose that the sequences $\{s_0, \ldots, s_{n-1}\}$ and $\{t_0, \ldots, t_{m-1}\}$ satisfy P and Q, respectively. It suffices to show that

$$|\sum_{i=1}^{n} f(s_{i-1})(g(p_i) - g(p_{i-1})) \; - \; \sum_{j=1}^{m} f(t_{j-1})(g(q_j) - g(q_{j-1}))| < \epsilon. \tag{12}$$

Let $\{z_0, \ldots, z_{r-1}\}$ be any sequence that satisfies the partition $P \cup Q = \{pq_0, \ldots, pq_r\}$. Then by applying the hypothesis and Lemma 2.3 to both P and Q, it follows that there exist sequences $\{x_0, \ldots, x_{n-1}\}$ and $\{y_0, \ldots, y_{m-1}\}$ that satisfy P and Q, respectively, such that

$$
\begin{aligned}
&|\sum_{i=1}^{n} f(s_{i-1})(g(p_i) - g(p_{i-1})) \;-\; \sum_{j=1}^{m} f(t_{j-1})(g(q_j) - g(q_{j-1}))| \\
&\quad \leq |\sum_{i=1}^{n} f(s_{i-1})(g(p_i) - g(p_{i-1})) \;-\; \sum_{k=1}^{r} f(z_{k-1})(g(pq_k) - g(pq_{k-1}))| \\
&\quad\quad + \;|\sum_{j=1}^{m} f(t_{j-1})(g(q_j) - g(q_{j-1})) \;-\; \sum_{k=1}^{r} f(z_{k-1})(g(pq_k) - g(pq_{k-1}))| \\
&\quad \leq |\sum_{i=1}^{n} f(s_{i-1})(g(p_i) - g(p_{i-1})) \;-\; \sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1}))| \\
&\quad\quad + \;|\sum_{j=1}^{m} f(t_{j-1})(g(q_j) - g(q_{j-1})) \;-\; \sum_{j=1}^{m} f(y_{j-1})(g(q_j) - g(q_{j-1}))| \\
&\quad\quad < \;\frac{\epsilon}{2} + \frac{\epsilon}{2} \;=\; \epsilon.
\end{aligned} \tag{13}
$$

This completes the mathematical proof of Theorem 2.2. □

As discussed in Section 2.7, the formal development of the Riesz representation theorem in PVS uses Theorem 2.2 directly to prove that any function that is continuous on $[a, b]$ is integrable with respect to any function of bounded variation on $[a, b]$. This is a proof that was developed specifically for the formal verification, although it is quite simple. A shortened version is given here.

Theorem 2.4. For all functions $f \in C[a, b]$ and $g \in BV[a, b]$, f is Riemann-Stieltjes integrable with respect to g.

Proof. By Lemma 2.1, it suffices to assume that g is increasing. Choose $\epsilon > 0$. By lemma 2.2, it suffices to show that there exists $\delta > 0$ with the property that for every partition P with width less than δ, any two sequences $\{x_0, \ldots, x_{n-1}\}$ and $\{y_0, \ldots, y_{n-1}\}$ that satisfy P have Riemann-Stieltjes sums less than ϵ apart. Since f is uniformly continuous, choose $\delta > 0$ such that for all $s, t \in [a, b]$, $|s - t| < \delta$ implies $|f(s) - f(t)| < \epsilon/(b - a)$. Suppose that P is any partition of $[a, b]$ with $n + 1$ elements and width less than δ. Then

$$
\begin{aligned}
&|\sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})) \;-\; \sum_{i=1}^{n} f(y_{i-1})(g(p_i) - g(p_{i-1}))| \\
&\quad \leq \sum_{i=1}^{n} |f(x_{i-1}) - f(y_{i-1})|(g(p_i) - g(p_{i-1})) \\
&\quad < \sum_{i=1}^{n} \frac{\epsilon}{b - a}(g(p_i) - g(p_{i-1})) \;=\; \epsilon.
\end{aligned} \tag{14}
$$

The last equality follows from the fact that the sum in question is telescoping. □

2.7 The PVS Proof of Integrability of Continuous Functions

As noted above, the challenging part in developing formal proofs of Theorems 2.2 and 2.4 was the proof of Lemma 2.3. This lemma is stated in PVS as follows.

```
Rie_sum_diff_extend_union : LEMMA a<b IMPLIES
    LET CI = closed_intv(a,b) IN
        (increasing?[(CI)](g) IMPLIES FORALL (P,Q: partition(a,b)):
      LET PQ = partition_union(a,b)(P,Q) IN
        FORALL (xispq:xis?(a,b,PQ),xis:xis?(a,b,P)):
        EXISTS (xis2:xis?(a,b,P)):
        abs(Rie_sum(a,b,g,P,xis,f) - Rie_sum(a,b,g,PQ,xispq,f)) <=
        abs(Rie_sum(a,b,g,P,xis,f)-Rie_sum(a,b,g,P,xis2,f)))
```

In Section 2.6, there were several statements made in the informal proof of this lemma that are easy to see on paper but somewhat difficult to prove in PVS. Two of these statements are discussed here.

Statement 1: "Given $i \leq n-1$, define z_i to be any element of the set $\{y_j \mid \sigma(j) = i\}$ such that $f(z_i) \leq f(y_j)$ for all $j \leq r-1$ such that $\sigma(j) = i$."

This is a perfectly valid, well defined mathematical statement, and it is easy to see that such a definition is possible. In the PVS development, the function σ is named `partition_union_map_inv`, and there is (right) inverse function α, which is called `partition_union_map` in PVS. The function α takes an element of $\{0, \ldots, n\}$ and returns an element of $\{0, \ldots, r\}$ such that $p_i = pq_{\alpha(i)}$ for all i. It has been proved in PVS that the function α is a right inverse to σ. In fact, the following, stronger property holds: For all $i \leq n$ and $j \leq r$, if $\alpha(i) \leq j$ and either $i = n$ or $j < \alpha(i+1)$, then $\sigma(j) = i$. The formal statement of this result in PVS is given below.

```
partition_union_map_inv_def: LEMMA a<b IMPLIES
 FORALL (P,Q:partition(a,b)):
 LET pum = partition_union_map(a,b,P,Q),
   puminv = partition_union_map_inv(a,b,P,Q)
 IN
   FORALL (j:below(partition_union(a,b)(P,Q)`length),
                                    i:below(P`length)):
   (pum(i) <= j AND (i<P`length-1 IMPLIES j < pum(i+1)))
   IMPLIES  puminv(j) = i
```

Formally, the element z_i is defined to be y_j, where $j < r$, $\alpha(i) \leq j$, $j < \alpha(i+1)$, and for all $s < r$, $\alpha(i) \leq s$ and $s < \alpha(i+1)$ implies $f(y_t) \leq f(y_s)$. For any such j, $\sigma(j) = i$. The proof that j exists is elementary but time consuming for such a trivial result. In the PVS development, it is actually proved that for all $w \leq r$, this property is satisfied by some j if r is replaced by w, and the proof of this statement is by induction on w. The definition of the sequence $\{z_0, \ldots, z_{n-1}\}$ in PVS is given as follows.

```
(# length := P`length-1,
   seq := (LAMBDA (ii:nat): IF ii<P`length-1 THEN
    LET jj = choose({jj:below(PQ`length-1) |
     (pum(ii) <= jj AND jj < pum(ii + 1) AND
     (FORALL (zz: below(PQ`length - 1)):
       pum(ii) <= zz AND zz < pum(ii + 1)
     IMPLIES f(xispq`seq(zz)) >= f(xispq`seq(jj))))})
```

```
    IN xispq‘seq(jj) ELSE default[T] ENDIF) #)
```

In this formal definition in PVS, PQ $= P \cup Q$, xispq $= \{y_0, \ldots, y_{r-1}\}$, and pum $=$ partition_union_map(a,b,P,Q), which is referred to above by α.

Statement 2:

$$\sum_{j=1}^{r} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1})) = \sum_{i=1}^{n} \sum_{\{j|\sigma(j-1)=i-1\}} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1}))$$

This statement is trivial and takes up one line in an informal proof. However, it is tricky to prove in PVS. In the PVS proofs, the sum $\sum_{\{j|\sigma(j-1)=i-1\}} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1}))$ is written as a sum from one natural number to another. The proof is based on the fact that

$$\{j \mid \sigma(j-1) = i-1\} \; = \; \{j \mid \alpha(i-1) \leq j \leq \alpha(i) - 1\},$$

where α is the function partition_union_map, discussed above. Since this step is vital to the proof of Lemma 2.3, the formal proof of this lemma in PVS required that the function partition_union_map be defined and that its basic properties be verified. Using this function, the sum in question can be rewritten as follows.

$$\sum_{\{j|\sigma(j-1)=i-1\}} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1})) = \sum_{j=\alpha(i-1)}^{\alpha(i)-1} f(z_{\sigma(j-1)})(g(pq_j) - g(pq_{j-1})). \tag{15}$$

While this is an informal description of the solution to this problem, the formal statement is given in PVS below. It is found inside the proof of the Lemma called Rie_sum_extend_union in the PVS development.

```
FORALL (yy:below(PQ‘length - 1)):
   pum(1+nn)<=yy AND yy<=pum(2+nn)-1
   IMPLIES
   sigma[below(PQ‘length-1)](pum(1+nn),yy,
      (LAMBDA (n:below(PQ‘length-1)):
        f(xis‘seq(sig(n))) * g(PQ‘seq(1+n))
        - f(xis‘seq(sig(n))) * g(PQ‘seq(n))))
   = f(xis‘seq(1+nn))*(g(PQ‘seq(1 + yy))-g(PQ‘seq(pum(1+nn))))
```

Here, PQ $= P \cup Q$, xis $= \{z_0, \ldots, z_{n-1}\}$, pum $=$ partition_union_map(a,b,P,Q), and sig $=$ partition_union_map_inv(a,b,P,Q), which is referred to above by σ. This case statement is proved in PVS by induction on yy.

3. LINEAR SPACES AND FUNCTIONALS

The Riesz theorem classifies the bounded linear functionals on $C[a, b]$ in terms of functions of bounded variation on $[a, b]$. The proof relies on a special case of the Hahn-Banach theorem (Section 4), which requires a definition of arbitrary linear subspaces of $B[a, b]$. Bounded linear subspaces are defined in PVS as subsets of $B[a, b]$ that contain $C[a, b]$ and are closed under sum and scalar multiplication. Such sets are defined using the predicate called bounded_linear_subspace? in PVS.

```
Fnz: VAR set[[INTab->real]]

   funs_sum_closed?(Fnz) : bool= FORALL (f,g:(Fnz)): Fnz(f+g)
   funs_const_closed?(Fnz): bool= FORALL (f:(Fnz),c:real): Fnz(c*f)
   funs_bounded?(Fnz): bool= FORALL (f:(Fnz)): bounded_on_int?(f)
   funs_contain_constants?(Fnz): bool= FORALL (c:real):
                                          Fnz(LAMBDA (y:INTab):c)
   funs_contain_continuous?(Fnz): bool= FORALL (f:[INTab->real]):
                              continuous_on_int?(f) IMPLIES Fnz(f)

   bounded_linear_subspace?(Fnz): bool =
                                 funs_sum_closed?(Fnz) AND
                                 funs_const_closed?(Fnz) AND
                                 funs_bounded?(Fnz) AND
                                 funs_contain_constants?(Fnz) AND
                                 funs_contain_continuous?(Fnz)
```

Using these predicates, linear functionals are defined on any set of functions from $[a, b]$ to $\mathbb{R}$ that satisfy the predicate `bounded_linear_subspace?`.

```
   Operator: TYPE= [Funs->real]

   L: VAR Operator

   additive_op?(L): bool= (FORALL (f,g): L(f+g) = L(f)+L(g))
   const_inv_op?(L): bool= (FORALL (f:Funs,c:real): L(c*f) = c*L(f))
   linear_op?(L): bool= additive_op?(L) AND const_inv_op?(L)
```

The Hahn-Banach theorem is a statement about normed linear spaces. The space $B[a, b]$ is a normed linear space with the (supremum) norm

$$\|g\|_{sup} \equiv \sup_{x\in[a,b]} |g(x)|. \tag{16}$$

The linear functional L on a subspace M of $B[a, b]$ is bounded if there exists a nonnegative real number K such that for all $f \in M$,

$$|L(f)| \le K \cdot \|f\|_{sup}. \tag{17}$$

The operator norm $\|L\|_M$ of L is the smallest real number K that satisfies Equation (17).

Bounded linear functionals are defined in PVS, along with the function that computes the operator norm of any such functional:

```
   bounded_op?(L): bool= (EXISTS (M:nnreal):
       FORALL (f): abs(L(f)) <= M*fun_norm(f))

   op_norm(L: (bounded_op?)): {M:nnreal |
       (FORALL (f:Funs): abs(L(f))<=M*fun_norm(f))
       AND (FORALL (M1:real): M1 < M IMPLIES
          EXISTS (f:Funs): abs(L(f)) > M1*fun_norm(f))}
```

```
op_norm_bound: LEMMA FORALL (L:(bounded_op?),f):
    abs(L(f)) <= op_norm(L)*fun_norm(f)

bounded_linear_operator?(L): bool=
    bounded_op?(L) AND linear_op?(L)
```

4. THE HAHN BANACH THEOREM

The standard proof of the Riesz representation theorem relies on the Hahn Banach theorem.

Theorem 4.1. (Complete Hahn-Banach) If M is a normed linear space, N is a linear subspace of M, and L is a bounded linear functional on N, then there is an extension $\bar{L}$ of L to M such that $\|\bar{L}\|_M = \|L\|_N$.

The following special case of the Hahn-Banach theorem was proved in PVS.

Theorem 4.2. (Special Case of Hahn-Banach) If L is a bounded linear functional on $C[a,b]$, then there exists an extension, $\bar{L}\colon B[a,b] \to \mathbb{R}$, of L such that $\|L\|_{C[a,b]} = \|\bar{L}\|_{B[a,b]}$.

As noted in the introduction, there have been several formal proofs of the complete Hahn-Banach theorem [NT93, BW00, Bau01]. In this development, most of Haaser and Sullivan's proof [HS71] of the Hahn-Banach theorem easily translated into PVS, except Lemma 4.4 in Section 4.2, which required some additional constructions that are described in that section.

4.1 Zorn's Lemma and a Partial Ordering on Extensions

The proof of the Hahn-Banach theorem follows from Zorn's lemma, and this appears to be the first application of Zorn's lemma, in the NASA PVS libraries, to a result from analysis. Recall the type-theoretic version of Zorn's lemma.

Lemma 4.3. (Zorn's Lemma) Let T be a type and $\leq$ a partial ordering on T. Suppose that for every chain A (a totally ordered subset of T), there exists an element T which is an upper bound for A. Then T contains an element that is maximal.

Zorn's lemma has been in the NASA libraries for some time, and was proved by Jerry James.

```
zorn: THEOREM
  (FORALL (ch: chain[T, <=]): bounded_above?[T](ch, <=)) IMPLIES
   (EXISTS t: maximal?[T](t, fullset[T], <=))
```

In the proof of the Hahn-Banach theorem, the partial ordering in question is defined as follows. Given a linear functional L on $C[a,b]$, consider the set EXT_L of all pairs (E, L_E), where E is a linear subspace of $B[a,b]$ containing $C[a,b]$, and L_E is an extension of L to E such that $\|L_E\|_E = \|L\|_{C[a,b]}$. There is a partial ordering $\leq$ on EXT_L, where $(E, L_E) \leq (W, L_W)$ means that $E \subset W$, L_W is an extension of L_E to W, and $\|L_E\|_E = \|L_W\|_W$.

4.2 An Important Supporting Lemma

In the formal PVS development, the proof of Theorem 4.2 relies on the following, simpler lemma, which makes use of the partial ordering defined on EXT_L in Section 4.1. It is evident in the proof of this lemma that some challenges should arise when formalizing it in PVS. The proof is extracted from the reasoning found in the proof of the Hahn-Banach theorem in [HS71].

Lemma 4.4. If (E, L_E) is an element of EXT_L, and if f is an element of $B[a,b]$ that is not contained in E, then there is an element (W, L_W) of EXT_L such that $(E, L_E) \leq (W, L_W)$ and $f \in W$.

Proof. Let $W = \{e + c \cdot f \mid e \in E \text{ and } c \in \mathbb{R}\}$. Choose any real number β such that for all $g, h \in E$,

$$-\|L_E\|_E \; \|h + f\|_{\sup} - L_E(h) \;\leq \beta \leq\; \|L_E\|_E \; \|g + f\|_{\sup} - L_E(g). \tag{18}$$

Define a linear functional L_W on W by $L_W(e + c \cdot f) = L_E(e) + c\beta$. If L_W is bounded, then since E is a subspace of W, it follows that $\|L_E\|_E \leq \|L_W\|_W$. Thus, it suffices to show that $|L_W(e + c \cdot f)| \leq \|L_E\|_E \cdot \|e + c \cdot f\|_{\sup}$ for all $e \in E$ and $c \in \mathbb{R}$. If $c = 0$, then this is trivial, since $e \in E$.

Suppose first that $L_W(e + c \cdot f) \geq 0$. If $c > 0$, then by Equation (18) with $g = e/c$, $L_W(e+c \cdot f) = L_E(e)+c\beta \leq L_E(e)+c(\|L_E\|_E \cdot \|e/c+f\|_{\sup} - L_E(e/c)) = \|L_E\|_E \cdot \|s + c \cdot f\|_{\sup}$. Similarly, if $c < 0$, then by Equation (18) with $h = e/c$, $L_W(e + c \cdot f) = L_E(e) + c\beta \leq L_E(e) + c(-\|L_E\|_E \cdot \|e/c + f\|_{\sup} - L_E(e/c)) = \|L_E\|_E \cdot \|s + c \cdot f\|_{\sup}$. The case where $L_W(e + c \cdot f) \leq 0$ is nearly identical. The formalization of this proof is discussed below. □

The most difficult part of the formal verification of the Hahn-Banach theorem in PVS was proving Lemma 4.4. In particular, the following statement from the proof above highlighted the difference between formal mathematics and standard mathematics. This is a well-defined, perfectly acceptable statement in an informal proof, but in PVS, it must be stated differently.

Statement from the Proof: "Define a linear functional L_W on W by $L_W(e + c \cdot f) = L_E(e) + c\beta$."

Some care is needed to state this definition in PVS, since the set W is defined in PVS not by $\{e + c \cdot f \mid e \in E \text{ and } c \in \mathbb{R}\}$ but rather by $\{g \mid g \in B[a,b] \text{ and } \exists e \in E, c \in \mathbb{R} : \; g = e + c \cdot f\}$. The set W is defined in PVS as follows.

```
{gg:[INTab->real] | bounded_on_int?[a,b](gg) AND
EXISTS (rr: [INTab->real],ffc:real):
             OE`space(rr) AND gg = rr + ffc*ff}
```

This definition complicates the definition of the functional L_W in PVS. In order to define L_W, it is necessary to have formally defined functions that, given an element g of W, compute the function $e \in E$ and the constant $c \in \mathbb{R}$ such that $g = e + c \cdot f$. These functions, defined in the formal proof of Lemma 4.4, are called `funchoose` and `constchoose`, respectively. They are defined in PVS as follows.

```
LAMBDA (gg:(W)): choose({rr:[INTab->real]|
  EXISTS (ffc:real):OE`space(rr) AND gg = rr + ffc * ff})
```

```
LAMBDA (g: (W)): choose({c: real |
  EXISTS (e: [INTab->real]): E`space(e) AND g= e+c*f})
```

The key properties of these functions are that they (1) compute a function $e \in E$ and a constant c such that $g = e + c \cdot f$, (2) are linear, and (3) are unique.

```
g = funchoose(g)+constchoose(g)*ff

funchoose(g+h) = funchoose(g)+funchoose(h) AND
   constchoose(g+h) = constchoose(g)+constchoose(h)

funchoose(c*g) = c*funchoose(g) AND
   constchoose(c*g) = c*constchoose(g)

g= e+c*f IMPLIES e= funchoose(g) AND c= constchoose(g)
```

With these definitions, the functional L_W is defined in PVS as follows.

```
LAMBDA (g:(W)): E`Lop(funchoose(g)) + constchoose(g)*beta
```

4.3 Proof of the Hahn-Banach Theorem

The following is an informal textbook-style proof of Theorem 4.2, which is stated at the beginning of Section 4. It is nearly identical to the proof found in [HS71]. As mentioned above, the most difficult part of the proof of this theorem in PVS was proving Lemma 4.4 (Section 4.2). The reason for including this informal proof here is to illustrate the need for Lemma 4.4.

Proof of Theorem 4.2 (The Hahn-Banach Theorem). Given a bounded linear functional L on $C[a, b]$, let the space EXT_L be defined as in Section 4.1. If there exists an element (E, L_E) of EXT_L that is maximal, then Lemma 4.4 implies that $E = B[a, b]$. By the definition of EXT_L, $(C[a, b], L) \leq (E, L_E)$, so the Hahn-Banach theorem holds with $\bar{L} = L_E$.

It therefore suffices to show that there exists an element of EXT_L that is maximal. By Zorn's Lemma (Lemma 4.3 in Section 4.1), this problem reduces to showing that every totally ordered subset of EXT_L is bounded above by an element of EXT_L. The proof of this fact follows from basic reasoning. □

5. THE FORMAL PROOF OF THE RIESZ REPRESENTATION THEOREM: FIXING AN INCORRECT TEXTBOOK PROOF

If $g\colon [a, b] \to \mathbb{R}$ is of bounded variation, then there is an associated linear functional I_g on $C[a, b]$ given by

$$I_g(f) \equiv \int_a^b f dg, \tag{19}$$

which is the Riemann-Stieltjes integral of f with respect to g (Section 2). Theorem 2.4 in Section 2.6 implies that any function $f \in C[a, b]$ is RS integrable with respect to g, so I_g is well defined. The next lemma is an exercise in Haaser and Sullivan's book [HS71].

Lemma 5.1. If $g \in BV[a,b]$ and $f \in C[a,b]$, then $|I_g(f)| \leq V_a^b(f)\|f\|_{\sup}$.

Proof. Given $\epsilon > 0$, choose a partition $P = \{p_0, \ldots, p_n\}$ of $[a,b]$ such that $|I_g(f) - \sum_{i=1}^n f(x_{i-1})(g(p_i) - g(p_{i-1}))| < \epsilon$. Then for any sequence $\{x_0, \ldots, x_{n-1}\}$ that satisfies P,

$$\begin{aligned} |I_g(f)| &\leq |\sum_{i=1}^n f(x_{i-1})(g(p_i) - g(p_{i-1}))| + |I_g(f) - \sum_{i=1}^n f(x_{i-1})(g(p_i) - g(p_{i-1}))| \\ &< \sum_{i=1}^n |f(x_{i-1})||g(p_i) - g(p_{i-1})| + \epsilon \\ &\leq V_a^b(g)\|f\|_{\sup} + \epsilon. \end{aligned} \tag{20}$$

Since ϵ was arbitrary, this completes the proof. □

The Riesz representation theorem says that every bounded linear functional L on $C[a,b]$ is given by I_g for some $g \in BV[a,b]$. Recall its statement from the introduction:

Theorem 5.2. (The Riez Representation Theorem; Theorem 6.1 in [HS71]) If L is a bounded linear functional on $C[a,b]$, then there exists a function $g \in BV[a,b]$ such that $L = I_g$ and $\|L\|_{C[a,b]} = V_a^b(g)$.

5.1 A Counterexample to the Textbook's Proof

The formal verification of the Riesz representation theorem in PVS revealed an error in the proof used in the textbook [HS71]. The error is subtle, and the proof looks valid at first glance. Only a small change was needed to the textbook proof, but it is not too hard to construct a counterexample to its proof.

The first step of the proof in the textbook is to extend the bounded linear functional L on $C[a,b]$ to a bounded linear functional $\bar{L}$ on $B[a,b]$ that has the same norm, via the Hahn-Banach theorem (Theorem 4.2 in Section 4). The function g is then defined by

$$g(a) = 0 \qquad \text{and} \qquad g(x) = \bar{L}(\chi_{[a,x)}) \quad \text{for} \ \ x \in (a,b]. \tag{21}$$

The error in this definition is that $g(b)$ should equal $\bar{L}(\chi_{[a,b]})$ (note the closed interval $[a,b]$). A counterexample is given as follows.

Let $L\colon C[a,b] \to \mathbb{R}$ be given by evaluation at b:

$$L(f) \equiv f(b).$$

It is easy to see that L is a bounded linear functional on $C[a,b]$ with norm equal to 1. Then L can be extended to the bounded linear functional $\bar{L}$ on $B[a,b]$ that is also given by evaluation at b, which also has norm equal to 1. Then the function g, defined in Equation (21), is identically zero on $[a,b]$, so the bounded linear functional $I_g\colon C[a,b] \to \mathbb{R}$ is identically zero as well. However, L is not identically zero, so $L \neq I_g$. This shows that the textbook proof has a (minor) error.

Typically, constructing a formal proof of a difficult result in a mathematics textbook will reveal some special cases that were overlooked by the author. In the case of the Riesz theorem, the distinction between the formal PVS proof and the informal textbook proof was even greater than this due to the error in the definition of the function $g \in BV[a,b]$.

5.2 The Updated Textbook

Haaser and Sullivan's textbook [HS71] was originally published in 1971 by Van Nostrand Rienhold Company, and it was this first version that was used as a reference in developing the PVS theories for the Riesz representation theorem. There is a newer version of the textbook that was published in 1991 by Dover Publications, Inc., New York. When the author learned that the new version existed, he was quite eager to see whether, in the twenty years since the first publication, the error in the proof of the Riesz theorem had been fixed. Indeed, the 1991 version says the following: "This Dover edition, first published in 1991, is a revised and corrected republication of the work originally published in 1971 by Van Nostrand Rienhold Company, New York as part of The University Series in Mathematics. Certain passages have been deleted or replaced with new material..." [HS91].

It seemed certain that an error in the proof of an important theorem such as the Riesz representation theorem would have been caught in the twenty years between these two printings. However, the following statement can still be found in the proof of the theorem in the 1991 printing of Haaser and Sullivan's book:

"Define a function g on $[a,b]$ by the rule:

$$g(a) = 0 \quad \text{and} \quad g(x) = F(\chi_{[a,x)}) \text{ for } x \in [a,b].\text{" [HS91]}$$

In their proof, F denotes the extension of the given linear functional on $C[a,b]$ to $B[a,b]$. Clearly, this error is subtle enough that standard mathematical proof techniques are unlikely to reveal the problem. This highlights the ability of formalized reasoning to find subtle errors in mathematical proofs.

5.3 A Correct Proof of the Riesz Representation Theorem

The proof of the theorem is constructive. In fact, the function g such that $L = I_g$ is given by

$$g(a) = 0, \qquad g(x) = \bar{L}(\chi_{[a,x)}) \text{ for } a < x < b, \qquad \text{and} \qquad g(b) = \bar{L}(\chi_{[a,b]}), \tag{22}$$

where $\bar{L}$ is an extension of L to the space $B[a,b]$ of bounded functions on $[a,b]$. The only difference between this definition and the incorrect textbook definition of g discussed in Section 5.1 is that $g(b)$ is equal to $\bar{L}(\chi_{[a,b]})$ instead of $\bar{L}(\chi_{[a,b)})$.

The key property of the function g, defined in Equation (22), is that for any $x, y \in [a,b]$ with $x < y$,

$$g(y) - g(x) = \bar{L}(\chi_{[x,y)}) \text{ if } y < b \quad \text{and} \quad g(y) - g(x) = \bar{L}(\chi_{[x,b]}) \text{ if } y = b. \tag{23}$$

The following version the proof of the Riesz representation theorem follows the proof in [HS71], with this error corrected.

Proof of Theorem 5.2 (Correct Version). By the Hahn-Banach theorem, the bounded linear functional L on $C[a,b]$ can be extended to a bounded linear functional $\bar{L}$ on $B[a,b]$ with the same norm as L. Let g be defined by Equation (22). If

$P\colon a = x_0 < \cdots < x_n = b$ is any strictly increasing partition, then

$$\begin{aligned}\sum_{i=1}^{n} |g(p_i) - g(p_{i-1})| &= \sum_{i=1}^{n} \pm(g(p_i) - g(p_{i-1})) \\ &= \bar{L}(\pm\chi_{[p_{n-1},b]} \ + \ \sum_{i=1}^{n-1} \pm\chi_{[p_{i-1},p_i)}) \\ &\leq \|\bar{L}\|_{B[a,b]} \ = \ \|L\|_{C[a,b]}.\end{aligned} \tag{24}$$

where the inequality follows from the fact that the argument to $\bar{L}$ is a function that always takes a value in $\{-1, 1\}$ and therefore has supremum norm equal to 1. This shows that g has bounded variation on $[a, b]$ and that

$$V_a^b(g) \leq \|L\|_{C[a,b]}. \tag{25}$$

Given any $f \in C[a, b]$ and $\epsilon > 0$, choose $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon/2$ for all $x, y \in [a, b]$. This is possible since f is uniformly continuous. Using the fact that f is Riemann-Stieltjes integrable with respect to g, choose any partition P of $[a, b]$ with width less than δ and any sequence $\{x_0, \ldots, x_{n-1}\}$ that satisfies P such that

$$|\sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})) \ - \ I_g(f)| \leq \frac{\epsilon\|L\|_{C[a,b]}}{2}.$$

Define a bounded function $s \in B[a, b]$ by

$$s = f(x_{n-1})\chi_{[p_{n-1},b]} \ + \ \sum_{i=1}^{n} f(x_{i-1})\chi_{[p_{i-1},p_i)}.$$

By Equation (23) above,

$$\bar{L}(s) = \sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})).$$

It is easy to see from the uniform continuity hypothesis on f that $\|f - s\|_{\sup} \leq \epsilon/2$. That is, $|f(y) - s(y)| < \epsilon/2$ for all $y \in [a, b]$. Thus,

$$\begin{aligned}|L(f) - I_g(f)| \ &\leq |L(f) - \bar{L}(s)| + |\sum_{i=1}^{n} f(x_{i-1})(g(p_i) - g(p_{i-1})) \ - \ I_g(f)| \\ &\leq |L(f) - \bar{L}(s)| + \frac{\epsilon\|L\|_{C[a,b]}}{2} \\ &= |\bar{L}(f - s)| + \frac{\epsilon\|L\|_{C[a,b]}}{2} \\ &\leq \|f - s\|_{\sup}\|L\|_{C[a,b]} + \frac{\epsilon\|L\|_{C[a,b]}}{2} \\ &\leq \epsilon\|L\|_{C[a,b]}.\end{aligned} \tag{26}$$

Since f and ϵ were arbitrarily chosen, it follows that $L = I_g$. It therefore follows from Lemma 5.1 that $\|L\|_{C[a,b]} \leq V_a^b(g)$. Thus, Equation (25) implies that $\|L\|_{C[a,b]} = V_a^b(g)$. □

5.4 The PVS Proof of the Riesz Representation Theorem

The Riesz representation theorem is stated in PVS as follows for an arbitrary bounded linear functional LC on $C[a,b]$.

```
fun_to_op(g): C_Bounded_Linear_Operator = (LAMBDA (f):
                                        integral(a,b,g,f))

riesz_representation: LEMMA EXISTS g:
   op_norm[a,b,Continuous_Function[a,b]](LC) =
                              total_variation(a,b,g)(b) AND
                              LC = fun_to_op(g)
```

The bounded function s, constructed in the proof in Section 5.3, is explicitly defined in PVS, where it is named `step_approx`. An important property of s, that a partition can be chosen so that $\|f - s\|_{\sup}$ is arbitrarily small, is also proved as a lemma.

```
char_fun(r1,r2)(x): nnreal =
      IF r1<=x AND x<r2 AND r2<b THEN 1 ELSIF r1<=x
      AND x<=r2 AND r2=b THEN 1 ELSE 0 ENDIF

step_approx(P:partition(a,b),xis:xis?(a,b,P))
           (fr:(bounded_on_int?[a,b])): (bounded_on_int?[a,b]) =
 (LAMBDA (x:INTab): sigma[below(P`length-1)](0,P`length-2,
    LAMBDA (n:below(P`length-1)):
               fr(xis`seq(n))*char_fun(P`seq(n),P`seq(n+1))(x)))

step_approx_def: LEMMA FORALL (eps:posreal):
   EXISTS (delta:posreal):
   FORALL (P:partition(a,b),xis:xis?(a,b,P)):
   strictly_increasing?(P) AND width(a,b,P)<delta IMPLIES
    (FORALL (x:INTab): abs(f(x)-step_approx(P,xis)(f)(x)) < eps)
```

The most important property of s, namely that

$$s = f(x_{n-1})\chi_{[p_{n-1},b]} \; + \; \sum_{i=1}^{n} f(x_{i-1})\chi_{[p_{i-1},p_i)}.$$

for any partition $P = \{p_0, \ldots, p_n\}$ of $[a,b]$ and any sequence $\{x_0, \ldots, x_{n-1}\}$ satisfying P, is proved inside the proof of the Riesz representation theorem. It is used to prove the theorem and is then proved separately.

```
LET gg = (LAMBDA (x:INTab):
            IF x = a THEN 0 ELSE LB(char_fun(a,x)) ENDIF)
IN
  FORALL (ff:Continuous_Function[a,b],P:partition(a,b),
     xis:xis?(a,b,P)): strictly_increasing?(P) IMPLIES
     LB(step_approx(P,xis)(ff)) = Rie_sum(a,b,gg,P,xis,ff)
```

6. COMPLEXITY OF THE FORMAL PROOF

There were many results that were formally proved during this project that are easy to see on paper but significantly more difficult to prove in PVS. Examples include most of the basic results on partitions. Many of the results in PVS, such as Theorem 2.2 (Section 2.6), used refinements (i.e. unions) of partitions. These were not previously part of the NASA libraries, because they are not used by the theories on the standard Riemann integral. Here is an example of a statement that is obvious on paper but somewhat tricky in PVS.

Statement 1: "The union of two partitions is a partition that is strictly increasing."

As stated in PVS:

```
partition_union(a,(b|a<b))(P,Q: partition(a,b)):
    {PQ: partition(a,b) | (FORALL (x:T): member(x,PQ) IFF
    (member(x,P) OR member(x,Q))) AND strictly_increasing?(PQ)}
```

There are 258 steps in the proof of the validity of this type declaration.

As discussed in Section 2.6, the approach taken in the textbook by Haaser and Sullivan for proving that continuous functions are Riemann-Stieltjes integrable with respect to any function of bounded variation is to prove that this integral is equivalent to the Darboux-Stieltjes integral. In the formal PVS development, a simpler method was used. Theorem 2.2 (Section 2.6) was proved, which captures the underlying concept behind this equivalence in a statement that is strong enough to prove the integrability of continuous functions. The section on this integral equivalence is about four pages long in Haaser and Sullivan's book. The formal proof of Theorem 2.2 in PVS requires 465 proof statements, and it relies on another lemma, specifically designed to help prove the theorem, whose proof is 282 steps long.

The proof of the Hahn Banach theorem was significantly longer in PVS than in the textbook. The proof in Haaser and Sullivan's book is about one page. However, in PVS, the proof of the Hahn Banach theorem (Theorem 4.2 in Section 4) is 851 proof steps, although it is not too complicated. The formal proof of Lemma 4.4 (Section 4.2), which directly supports the Hahn-Banach theorem, is half a page in the textbook but 1010 steps long in PVS.

Finally, the proof of the Riesz representation theorem is also about one page long in the textbook. In PVS, the proof of the theorem itself takes 773 steps, and including all of the supporting lemmas in the same PVS file, the proof takes 1343 steps.

7. CONCLUSION

The Riesz representation theorem, as originally stated by Riesz [Rie09], has been proved in the PVS theorem prover. The formalization of the theorem and the Riemann Stieltjes integral is based mostly on Haaser and Sullivan's book [HS71]. In order to prove the Riesz representation theorem, it was necessary to prove that continuous functions on a closed interval are Riemann Stieltjes integrable with respect to any function of bounded variation. Haaser and Sullivan's book uses the equivalence of the Riemann Stieltjes and Darboux Stieltjes integrals to prove this integrability result. A simpler method was used in the formal PVS development, where a theorem was proved that captures the the main concept of this integral equivalence that is useful in the proof of the integrability of continuous functions.

The Hahn Banach theorem was proved in the case where the normed linear spaces are the continuous and bounded functions on a closed interval. The proof of the Hahn-Banach theorem follows from Zorn's lemma, and this appears to be the first application of Zorn's lemma, in the NASA PVS libraries, to a result from analysis.

There is an error in Haaser and Sullivan's proof of the Riesz representation theorem. Indeed, the proof is constructive, and the constructed function does not satisfy a necessary property. This error illustrates the ability of formal verification to find logical errors. A specific counterexample was given to the proof in the textbook. Finally, a corrected proof of the Riesz representation theorem was presented.

The PVS development presented in this paper can be downloaded online at the link http://shemesh.larc.nasa.gov/people/ajn/pvs_development/. Included there are the PVS theories in which the Riemann Stieltjes integral is defined, as well as the theories that directly support the Riesz representation theorem.

References

[Bau01] Gertrud Bauer. The Hahn-Banach theorem for real vector spaces. In *PART OF THE ISABELLE99-2 DISTRIBUTION, LIBRARY/HOL/HOL-REAL/HAHNBANACH/DOCUMENT.PDF*, 2001.

[But09] Ricky Butler. Formalization of the integral calculus in the pvs theorem prover. *Journal of Formalized Reasoning*, 2(1), 2009.

[BW00] Gertrud Bauer and Markus Wenzel. Computer-assisted mathematics at work – the Hahn-Banach theorem in isabelle/isar. In *TYPES FOR PROOFS AND PROGRAMS: TYPES99, LNCS*, 2000.

[Gra84] J.D. Gray. The shaping of the riesz representation theorem: A chapter in the history of analysis. *Archive for History of Exact Sciences*, 31(2), 1984.

[HS71] Norman B. Haaser and Joseph A. Sullivan. *Real Analysis*. Van Nostrand Rienhold Company, 450 West 33rd Street, New York, NY 10001, 1971.

[HS91] Norman B. Haaser and Joseph A. Sullivan. *Real Analysis*. Dover Publications, Inc., New York, 31 East 2nd Street, Mineola, New York 11501, 1991.

[NT93] B. Nowak and A. Trybulec. The Hahn-Banach theorem. *Journal of Formalized Mathematics*, (5), 1993.

[Rie09] Frédéric Riesz. Sur les opérations fonctionelles linéaires. *Comptes rendus de l'Académie des sciences*, (149):974–977, 1909.