# Real Vector Spaces and the Cauchy-Schwarz Inequality in ACL2(r)

Carl Kwan    Mark R. Greenstreet

Department of Computer Science
University of British Columbia*
Vancouver, Canada
{carlkwan,mrg}@cs.ubc.ca

We present a mechanical proof of the Cauchy-Schwarz inequality in ACL2(r) and a formalisation of the necessary mathematics to undertake such a proof. This includes the formalisation of $\mathbb{R}^n$ as an inner product space. We also provide an application of Cauchy-Schwarz by formalising $\mathbb{R}^n$ as a metric space and exhibiting continuity for some simple functions $\mathbb{R}^n \to \mathbb{R}$.

The Cauchy-Schwarz inequality relates the magnitude of a vector to its projection (or inner product) with another:

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

with equality iff the vectors are linearly dependent. It finds frequent use in many branches of mathematics including linear algebra, real analysis, functional analysis, probability, etc. Indeed, the inequality is considered to be among "The Hundred Greatest Theorems" and is listed in the "Formalizing 100 Theorems" project. To the best of our knowledge, our formalisation is the first published proof using ACL2(r) or any other first-order theorem prover.

## 1 Introduction

The Cauchy-Schwarz inequality is considered to be one of the most important inequalities in mathematics. Indeed, it appears in functional analysis, real analysis, probability theory, linear algebra, and combinatorics to name a few. Cauchy-Schwarz even made an appearance on an online list of "The Hundred Greatest Theorems" and the subsequent formalised version of the list "Formalizing 100 Theorems" [8, 24]. Some of the systems used for the proof include the usual suspects HOL/Isabelle, Coq, Mizar, PVS, etc. Notably missing, however, from the list of formalisations of Cauchy-Schwarz is a proof in ACL2 or ACL2(r). We remedy this.

In this paper, we present a formal proof of the Cauchy-Schwarz inequality in ACL2(r) including both forms (squared and norm versions) and the conditions for equality. This is the first proof of Cauchy-Schwarz for real vector spaces of arbitrary dimension $n \in \mathbb{N}$ in ACL2(r); in fact, to the best of our knowledge, this is the first proof in any first-order theorem prover. Such a formalisation suggests ACL2(r) applications in the various areas of mathematics in which the inequality appears. Indeed, we use Cauchy-Schwarz to prove $\mathbb{R}^n$ is a metric space in this paper and to prove theorems involving convex functions in [10].

The proof of Cauchy-Schwarz requires a theory of real vectors in ACL2(r). ACL2(r) extends ACL2 with real numbers formalised via non-standard analysis [3], and it supports automated reasoning involving irrational real and complex numbers in addition to the respective rational subsets that are supported

[*] This work is supported in part by the National Science and Engineering Research Council of Canada (NSERC) Discovery Grant program and the Institute for Computing, Information and Cognitive Systems (ICICS) at UBC.

Shilpi Goel and Matt Kaufmann (Eds.):
ACL2 Theorem Prover and its Applications (ACL2-2018)
EPTCS 280, 2018, pp. 111–127, doi:10.4204/EPTCS.280.9

in the vanilla distribution of ACL2 [5]. The natural next step beyond $\mathbb{R}$ is $\mathbb{R}^n$ which is fundamental to many branches of mathematics. Indeed, as a geometric structure, we view $\mathbb{R}^n$ as a place in which to perform analysis; as an algebraic object, $\mathbb{R}^n$ is a direct sum of subspaces. Under different lenses we view $\mathbb{R}^n$ as different structures, and thus it inherits the properties of the related structures. It is a metric space, a Hilbert space, a topological space, a group, etc. The ubiquitous nature of $\mathbb{R}^n$ suggests such a formalisation will open opportunities for applications in the many areas of mathematics $\mathbb{R}^n$ appears.

Formalising $\mathbb{R}^n$ for arbitrary $n$ introduces technical difficulties in ACL2(r). The fundamental differences between the rational and irrational numbers induce a subtle schism between ACL2 and ACL2(r) wherein the notions formalised in ACL2 (which are bestowed the title of *classical*) are far more well-behaved than those unique to ACL2(r) (which are respectively referred to as *non-classical*). The arbitrariness of $n$ suggests the necessity of defining operations recursively – yet non-classical recursive functions are not permitted in ACL2(r).

For the purposes of this paper, we present $\mathbb{R}^n$ formalised from two perspectives. First, we consider $\mathbb{R}^n$ as an inner product space under the usual dot product. The second perspective formalises $\mathbb{R}^n$ as a metric space. In this case, we use the usual Euclidean metric. Metrics also provide the framework in which to perform analysis and we introduce notions of continuity in ACL2(r) for functions on $\mathbb{R}^n$ and provide some simple examples of such functions.

Verifying the metric space properties of $\mathbb{R}^n$ traditionally requires the Cauchy-Schwarz inequality, which is the highlight of this paper. Accordingly, a focus on the formalisation of Cauchy-Schwarz will be emphasised. For the sake of brevity, a selection of the remaining definitions and theorems will be showcased with most details omitted. In the remainder of the paper we instead discuss the dynamics between approaching formalisation via simple definitions and avoiding fundamental logical limitations in expressibility. For example, real vectors are unsurprisingly represented using lists of ACL2(r) real numbers. However, we were surprised to find that deciding whether all components of a vector were infinitesimals via a recursive function that applies a recognizer to each entry was impossible due to the prohibition of non-standard recursive functions. We provide our solutions to this issue and other challenges that arose in the course of the formalisation.

Most of the mathematics encountered in this paper can be found in standard texts such as [20, 21, 11, 7]. Our proof Cauchy-Schwarz is similar to the one in [19, Chapter 9] and further reading specific to relevant inequalities can be found in [22]. Non-standard analysis is less standard of a topic but some common references are [18, 3, 12].

## 2 Related Work

Theorem provers with prior formalisations of Cauchy-Schwarz include HOL Light, Isabelle, Coq, Mizar, Metamath, ProofPower, and PVS. However, it appears some of the statements do not mention the conditions for equality [13, 9].

While this paper focuses on $\mathbb{R}^n$ from two perspectives, most other formalisations in the literature outline only one view of $\mathbb{R}^n$ – usually either neglecting to address metrics or lacking in the development of vectors. Moreover, to the best of our knowledge, these formalisations of $\mathbb{R}^n$ are verified in a higher-order setting. For, example, a theory of complex vectors with a nod towards applications in physics was formalised in HOL Light but does not address metrics [2]. On the other end of the spectrum, there is a HOL formalisation of Euclidean metric spaces and abstract metric spaces in general that does not fully include a theory of real vectors [6, 14]. This observation extends to similar results in Coq [23].

Within ACL2(r), there has been formalisation for some special cases of $n$. The work that handled

$n = 1$ is indeed fundamental and indispensable [5]. We may view $\mathbb{C} \simeq \mathbb{R}^2$ as a vector space over $\mathbb{R}$ so $n = 2$ is immediate since ACL2(r) supports complex numbers. Moreover, extensions of $\mathbb{C}$ such as the quaternions $\mathbb{H}$ and the octonions $\mathbb{O}$ with far richer mathematical structure than typical vector spaces have recently been formalised, which addresses the cases of $n = 4$ and $n = 8$ [4].

# 3 Preliminaries

## 3.1 Inner Product Spaces

Inner product spaces are vector spaces equipped with an inner product which induces – among other notions – rigorous definitions of the *angle* between two vectors and the *size* of a vector. More formally, a vector space is a couple $(V,F)$ where $V$ is a set of vectors and $F$ a field equipped with scalar multiplication such that

$$v + (u + w) = (v + u) + w \qquad \textit{(associativity)} \qquad (1)$$

$$v + u = u + v \qquad \textit{(commutativity)} \qquad (2)$$

$$\text{there exists a } 0 \in V \text{ such that } v + 0 = v \qquad \textit{(additive identity)} \qquad (3)$$

$$\text{there exists a } -v \in V \text{ such that } v + (-v) = 0 \qquad \textit{(additive inverse)} \qquad (4)$$

$$a(bv) = (ab)v \qquad \textit{(compatibility)} \qquad (5)$$

$$1v = v \qquad \textit{(scalar identity)} \qquad (6)$$

$$a(v + u) = av + au \qquad \textit{(distributivity of vector addition)} \qquad (7)$$

$$(a + b)v = av + bv \qquad \textit{(distributivity of field addition)} \qquad (8)$$

for any $v,u,w \in V$ and any $a,b \in F$ [19, Chapter 1].

An inner product space is a triple $(V,F,\langle -,- \rangle)$ where $(V,F)$ is a vector space and $\langle -,- \rangle : V \times V \to F$ is a function called an inner product, i.e. a function satisfying

$$\langle au + v, w \rangle = a\langle u, w \rangle + \langle v, w \rangle \qquad \textit{(linearity in the first coordinate)} \qquad (9)$$

$$\langle u, v \rangle = \langle v, u \rangle \text{ when } F = \mathbb{R} \qquad \textit{(commutativity)} \qquad (10)$$

$$\langle u, u \rangle \geq 0 \text{ with equality iff } u = 0 \qquad \textit{(positive definiteness)} \qquad (11)$$

for any $u,v,w \in V$ and $a \in F$ [19, Chapter 9].

For $\mathbb{R}^n$, this means $(\mathbb{R}^n,\mathbb{R},\langle -,- \rangle)$ is our inner product space and we choose $\langle -,- \rangle$ to be the bilinear map

$$\langle (u_1,u_2,\dots,u_n),(v_1,v_2,\dots,v_n) \rangle = u_1v_1 + u_2v_2 + \cdots u_nv_n \qquad (12)$$

also known as the dot product.

## 3.2 The Cauchy-Schwarz Inequality

Let $\|\cdot\|$ be the norm induced by $\langle -,- \rangle$. The Cauchy-Schwarz inequality states

$$|\langle u, v \rangle|^2 \leq \langle u, u \rangle \cdot \langle v, v \rangle \qquad \text{(Cauchy-Schwarz I)}$$

or, equivalently,

$$|\langle u, v \rangle| \leq \|u\|\|v\| \qquad \text{(Cauchy-Schwarz II)}$$

for any vectors $u$, $v$ [19, Chapter 9]. Moreover, equality holds iff $u$, $v$ are linearly dependent.

*Proof (sketch).* If $v = 0$, then the claims are immediate. Suppose $v \neq 0$ and let $a$ be a field element. Observe

$$0 \leq \|u-av\|^2 = \langle u-av, u-av\rangle = \|u\|^2 - 2a\langle u,v\rangle + a^2\|v\|^2. \tag{13}$$

Setting $a = \|v\|^{-2}\langle u,v\rangle$ and rearranging produces (Cauchy-Schwarz I). Take square roots and we get (Cauchy-Schwarz II). Note that $0 \leq \|u-av\|^2$ is the only step with an inequality so there is equality iff $u = av$. □

More details will be provided as we discuss the formal proof – especially the steps that involve "rearranging".

### 3.3 Metric Spaces

Metric spaces are topological spaces equipped with a metric function which is a rigorous approach to defining the intuitive notion of *distance* between two vectors. Formally, a metric space is a couple $(M,d)$ where $M$ is a set and $d : M \times M \to \mathbb{R}$ a function satisfying

$$d(x,y) = d(y,x) \qquad \textit{(commutativity)} \tag{14}$$

$$d(x,y) \geq 0 \text{ with equality iff } x = y \qquad \textit{(positive definiteness)} \tag{15}$$

$$d(x,y) \leq d(x,z) + d(z,y) \qquad \textit{(triangle inequality)} \tag{16}$$

for any $x,y,z \in M$ [19, Chapter 9]. The function $d$ is called a *metric*.

In this case, we view $\mathbb{R}^n$ as $(\mathbb{R}^n, d_2)$ where $d_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ denotes the Euclidean metric

$$d_2(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2} = \left(\sum_{i=1}^{n}(x_i-y_i)^2\right)^{1/2}. \tag{17}$$

Proofs for the first two properties of the metric $d_2$ follow directly from the definition of the function. The triangle inequality follows from the Cauchy-Schwarz inequality [11, Chapter 15].

A metric provides sufficient tools for defining continuity in a manner similar to that in single variable calculus [20, Chapter 4]. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *continuous* everywhere if for any $x \in \mathbb{R}^n$ and $\varepsilon > 0$ there is a $\delta > 0$ such that for any $y \in \mathbb{R}^n$, if

$$d_2(x,y) < \delta, \tag{18}$$

then

$$|f(x) - f(y)| < \varepsilon. \tag{19}$$

This naturally leads to differentiability which is mentioned briefly in [10].

### 3.4 Non-standard Analysis and ACL2(r)

Classical real analysis is well known for its epsilon-delta approach to mathematical theory-building. For example, we say that function $f : \mathbb{R} \to \mathbb{R}$ is continuous at $x \in \mathbb{R}$ iff [20, Chapter 4]

$$\forall \varepsilon > 0,\ \exists \delta > 0 :\ \forall y \in \mathbb{R},\ |y-x| < \delta \implies |f(y) - f(x)| < \varepsilon. \tag{20}$$

This classical approach makes extensive use of nested quantifiers and support for quantifiers in ACL2 and ACL2(r) is limited. In fact, proofs involving terms with quantifiers often involve recursive witness functions that enumerate all possible values for the quantified term, e.g. see [1]. Of course, we cannot enumerate all of the real numbers. Instead of using epsilon-delta style reasoning, ACL2(r) is built on a formalisation of non-standard analysis – a more algebraic yet isomorphic approach to the theory of real analysis [3].

Non-standard analysis introduces an extension of $\mathbb{R}$ called the *hyperreals* $^*\mathbb{R} \supset \mathbb{R}$ which include numbers larger in magnitude than any finite real and the reciprocals of such numbers. These large hyperreals are aptly named *infinite* and their reciprocals are named *infinitesimal*. If $\omega$ is an infinite hyperreal, then it follows that $|1/\omega| < x$ for any positive finite real $x$. Also, 0 is an infinitesimal.

Any finite hyperreal is the sum of a real number and an infinitesimal. The real part of a hyperreal can be obtained through the *standard-part* function $\mathrm{st} : {}^*\mathbb{R} \to \mathbb{R}$.

To state a function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous in the language of non-standard analysis amounts to: if $d(x,y)$ is an infinitesimal for a standard $x$, then so is $|f(x) - f(y)|$.

In ACL2(r), lists of real numbers are recognized by `real-listp`. The recognizer for infinitesimals is the function `i-small` and the recognizer for infinite hyperreals is `i-large`. Reals that are not `i-large` are called `i-limited`.

# 4 $\mathbb{R}^n$ as an Inner Product Space in ACL2(r)

## 4.1 Vector Space Axioms

Most of the properties of real vector spaces pass with relative ease and minimal guidance by the user. Vector addition is `(vec-+ u v)` and scalar multiplication is `(scalar-* a v)` where $a$ is a field element and $u$, $v$ are vectors in $\mathbb{R}^n$. The zero vector is recognized by `zvecp`.

One slightly more challenging set of theorems involve vector subtraction, `(vec-- u v)`. Ideally, vector subtraction would be defined as a macro equivalent to `(vec-+ u (scalar-* -1 v))` to remain consistent with subtraction for reals in ACL2(r) and we indeed do so. However, it turns out that proving theorems regarding a function equivalent to `vec--`, which we call `vec--x`, and then proving the theorems for `vec--` via the equivalence is more amenable to verification in ACL2(r) than immediately proving theorems about `vec--`. In particular, the theorems involving closure, identity, inverses, anticommutativity, etc. are almost immediate using this approach. An example of this can be seen in Program 4.1. Upon verification of the desired properties for `vec--`, the theorem positing the equivalence of `vec--` to `vec--x` is disabled so as to not pollute the space of rules.

## 4.2 Inner Product Space Axioms

Like the vector space axioms, the majority of the relevant inner product space theorems passes by guiding ACL2(r) through textbook proofs. Aside from the usual suspects, one notable set of theorems are the bilinearity of the dot product. In particular, while the proof for the linearity of the first coordinate of the dot product executes via induction without any hints, the proof of linearity for the second coordinate does not pass so easily. Providing an induction scheme in the form of a hint would likely produce the desired proof. It was simpler, however, to apply commutativity of the dot product and use linearity of the first coordinate to exhibit the same result, ie. given

$$\langle au, v\rangle = a\langle u, v\rangle, \tag{21}$$

$$\langle u, v\rangle = \langle v, u\rangle, \tag{22}$$

**Program 4.1** Using the equivalence between `vector--` and `vec---x`.

```
(defthm vec---equivalence
 (implies (and (real-listp vec1)
               (real-listp vec2)
               (= (len vec1) (len vec2)))
          (equal (vec-- vec1 vec2) (vec--x vec1 vec2)))
 :hints (("GOAL" :in-theory (enable vec--x vec-+ scalar-*)
                 :induct (and (nth i vec1) (nth i vec2)))))
...
(defthm vec--x-anticommutativity
        (= (vec--x vec1 vec2) (scalar-* -1 (vec--x vec2 vec1)))
 :hints (("GOAL" :in-theory (enable vec--x scalar-*))))

(defthm vec---anticommutativity
 (implies (and (real-listp vec1) (real-listp vec2)
               (= (len vec2) (len vec1)))
          (= (vec-- vec1 vec2) (scalar-* -1 (vec-- vec2 vec1))))
 :hints (("GOAL" :use ((:instance vec---equivalence)
                       (:instance vec---equivalence (vec1 vec2) (vec2 vec1))
                       (:instance vec--x-anticommutativity)))))
```

we have

$$\langle u, av\rangle = \langle av, u\rangle = a\langle v, u\rangle = a\langle u, v\rangle. \tag{23}$$

Program 4.2 shows the ACL2(r) version of this proof.

Our reliance of proving theorems about algebraic structures via their algebraic properties instead of via induction is well exemplified here. This is especially important for the following formalisation of metric spaces. Because non-classical recursive functions are not permitted in ACL2(r), suppressing the definition of recursive functions on vectors, say `dot`, within a `define` facilitates the reasoning of infinitesimals in the space of $\mathbb{R}^n$. In particular, since $\langle -, - \rangle : \mathbb{R}^n \to \mathbb{R}$ is a real-valued function, we may connect the notion of infinitesimal values of $\langle -, - \rangle$ with the entries of the vectors on which $\langle -, - \rangle$ is evaluated without unravelling the recursive definition of `dot`. Details will be provided when we discuss the formalisation of metric spaces.

## 5 Formalising Cauchy-Schwarz

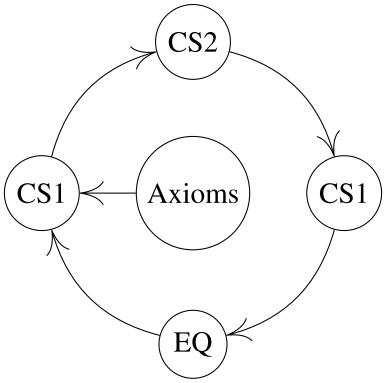
In this section we outline some of the key lemmas in ACL2(r) that result in the Cauchy-Schwarz inequality. Much of the proof is user guided via the algebraic properties of norms, the dot product, etc. Note that the lemma numbers correspond to those in the book `cauchy-schwarz.lisp`. By observing gaps in the numbering sequence in this presentation, the reader can infer where a handful of extra lemmas were needed to complete the proof in ACL2(r).

**Program 4.2** An example of using commutativity to prove bilinearity of the dot product.

```
(defthm dot-commutativity
 (implies (and (real-listp vec1) (real-listp vec2)
               (= (len vec2) (len vec1)))
          (= (dot vec1 vec2) (dot vec2 vec1)))
 :hints (("GOAL" :in-theory (enable dot))))

(defthm dot-linear-first-coordinate-1
 (implies (and (real-listp vec1) (real-listp vec2)
               (= (len vec2) (len vec1)) (realp a))
          (= (dot (scalar-* a vec1) vec2)
             (* a (dot vec1 vec2))))
 :hints (("GOAL" :in-theory (enable dot scalar-*))))
...
(defthm dot-linear-second-coordinate-1
 (implies (and (real-listp vec1) (real-listp vec2)
               (= (len vec2) (len vec1)) (realp a))
          (= (dot vec1 (scalar-* a vec2))
             (* a (dot vec1 vec2))))
 :hints (("GOAL" :do-not-induct t
                 :use ((:instance scalar-*-closure (vec vec2))
                       (:instance dot-commutativity (vec2 (scalar-* a vec2)))))))
```

Figure 1: Structure of the proof for Cauchy-Schwarz. CS1, CS2, and EQ denotes (Cauchy-Schwarz I), (Cauchy-Schwarz II), and the conditions for equality, respectively.



## 5.1 Axioms $\Longrightarrow$ (Cauchy-Schwarz I)

Suppose $v \neq 0$. First we prove

$$\|u-v\|^2 = \langle u,u\rangle - 2\langle u,v\rangle + \langle v,v\rangle \tag{24}$$

by applying multiple instances of the bilinearity of $\langle -, - \rangle$. This can be seen in Program 5.1. Since $\|u - v\| \geq 0$, replacing $v$ with $\frac{\langle u,v \rangle}{\langle v,v \rangle} v$ in Equation (24) above produces

$$0 \leq \left\| u - \frac{\langle u,v \rangle}{\langle v,v \rangle} v \right\|^2 = \langle u,u \rangle - 2 \left\langle u, \frac{\langle u,v \rangle}{\langle v,v \rangle} v \right\rangle + \left\langle \frac{\langle u,v \rangle}{\langle v,v \rangle} v, \frac{\langle u,v \rangle}{\langle v,v \rangle} v \right\rangle. \tag{25}$$

This can be seen in Program 5.2 and the inequality reduces to

$$0 \leq \langle u,u \rangle - 2 \frac{\langle u,v \rangle}{\langle v,v \rangle} \langle u,v \rangle + \frac{\langle u,v \rangle^2}{\langle v,v \rangle^2} \langle v,v \rangle = \langle u,u \rangle + \langle u,v \rangle \left( -2 \frac{\langle u,v \rangle}{\langle v,v \rangle} + \frac{\langle u,v \rangle}{\langle v,v \rangle} \right) = \langle u,u \rangle - \frac{\langle u,v \rangle}{\langle v,v \rangle} \tag{26}$$

Rearranging then produces (Cauchy-Schwarz I). Splitting into the cases $v \neq 0$ and $v = 0$ produces (Cauchy-Schwarz I) for arbitrary vectors $u$, $v$ since the case where $v$ is zero is trivial. This first version of Cauchy-Schwarz can be seen in Program 5.3.

**Program 5.1** Applying bilinearity of the dot product to prove a simple identity.

```
;; < u - v , u - v > = < u , u > - < u , v > - < v , u > + < v , v >
(defthm lemma-3
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (norm^2 (vec-- u v))
                 (+ (dot u u) (- (dot u v)) (- (dot v u)) (dot v v))))
 :hints (("GOAL" :use (...(:instance dot-linear-second-coordinate-2
                                     (vec1 v) (vec2 u)
                                     (vec3 (scalar-* -1 v)))
                          (:instance dot-linear-second-coordinate-2
                                     (vec1 u) (vec2 u)
                                     (vec3 (scalar-* -1 v))))))))

;; < u - v, u - v > = < u, u > - 2 < u , v > + < v, v >
(defthm lemma-4
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (norm^2 (vec-- u v))
                 (+ (dot u u) (- (* 2 (dot u v))) (dot v v))))
 :hints (("GOAL" :use ((:instance dot-commutativity (vec1 u) (vec2 v))))))
```

## 5.2 (Cauchy-Schwarz I) $\Longleftrightarrow$ (Cauchy-Schwarz II)

To see (Cauchy-Schwarz II) from (Cauchy-Schwarz I), we simply take square roots and show the equivalence between the dot products and the square of the norms. This part can be seen in Program 5.4. To see the other direction, we simply square both sides and rearrange.

## 5.3 (Cauchy-Schwarz I) $\Longleftrightarrow$ Conditions for Equality

Suppose $u, v \neq 0$ and

$$\langle u,v \rangle^2 = \langle u,u \rangle \langle v,v \rangle. \tag{27}$$

**Program 5.2** Substituting $v$ for $\langle v,v\rangle^{-1}\langle u,v\rangle v$.

```
;; 0 <= < u, u > - 2 < u , v > + < v, v >
(local (defthm lemma-6
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (<= 0 (norm^2 (vec-- u v)))
                 (<= 0 (+ (dot u u) (- (* 2 (dot u v))) (dot v v)))))))

;; let v = (scalar-* (* (/ (dot v v)) (dot u v)) v)
(local (defthm lemma-7
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (<= 0 (norm^2 (vec-- u (scalar-* (* (/ (dot v v)) (dot u v)) v))))
                 (<= 0 (+ (dot u u)
                          (- (* 2 (dot u (scalar-* (* (/ (dot v v)) (dot u v)) v))))
                          (dot (scalar-* (* (/ (dot v v)) (dot u v)) v)
                               (scalar-* (* (/ (dot v v)) (dot u v)) v))))))
 :hints (("GOAL" :use (...(:instance lemma-6 (v (scalar-* (* (/ (dot v v))
                                                             (dot u v)) v))))))))
```

**Program 5.3** Final form of (Cauchy-Schwarz I).

```
(defthm cauchy-schwarz-1
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
                 (<= (* (dot u v) (dot u v))
                     (* (dot u u) (dot v v))))
 :hints (("GOAL" ... :cases ((zvecp v) (not (zvecp v))))))
```

Then we simply reverse all the equalities used to prove (Cauchy-Schwarz I) from the axioms (see Inequality (25)) until we return to

$$0 = \left\| u - \frac{\langle u,v\rangle}{\langle v,v\rangle} v \right\|^2. \tag{28}$$

Since $\|\cdot\|^2$ is positive definite, we must have

$$u = \frac{\langle u,v\rangle}{\langle v,v\rangle} v. \tag{29}$$

This can be seen in Program 5.5.

To show linearity, we introduce a Skolem function so that the final form of the conditions for equality are in the greatest generality. We also attempted a proof where the value of the Skolem constant was explicitly computed – simply find the first non-zero element of $v$ and divide the corresponding element of $u$ by the $v$ element. The proof for the Skolem function approach was much simpler because the witness value comes already endowed with the properties we need for subsequent reasoning. In particular, invoking linear dependence to show (Cauchy-Schwarz I) simply amounts to applying algebraic rules to an arbitrary unknown witness which is simple (though occasionally tedious) from our formalisation. Otherwise, not using a Skolem function would necessitate the exhibition of a particular coefficient to complete the implication graph in Figure 1. The definition of the Skolem function is in Program 5.6. The cases for

**Program 5.4** Showing (Cauchy-Schwarz II).

```
(local (defthm lemma-16
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (and (equal (acl2-sqrt (* (dot u v) (dot u v))) (abs (dot u v)))
               (equal (acl2-sqrt (dot u u)) (eu-norm u))
               (equal (acl2-sqrt (dot v v)) (eu-norm v))
               (equal (acl2-sqrt (* (dot u u) (dot v v)))
                      (* (eu-norm u) (eu-norm v)))))
 :hints (("GOAL" :use ((:instance norm-inner-product-equivalence (vec u))
                       (:instance norm-inner-product-equivalence (vec v)))))))

(defthm cauchy-schwarz-2
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (<= (abs (dot u v))
              (* (eu-norm u) (eu-norm v))))
 :hints (("GOAL" :use ((:instance cauchy-schwarz-1)
                       (:instance norm-inner-product-equivalence (vec v))
                       (:instance norm-inner-product-equivalence (vec u)) ...) ...)))
```

$u = 0$ or $v = 0$ are immediate. The final result can be seen in Program 5.7. The conditions for equality for (Cauchy-Schwarz II) follows from its equivalence to (Cauchy-Schwarz I).

**Program 5.5** Linearity follows from equality.

```
(local (defthm lemma-19
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)) (not (zvecp v))
               (= (* (dot u v) (dot u v)) (* (dot u u) (dot v v))))
          (equal u (scalar-* (* (/ (dot v v)) (dot u v)) v))) ...))
```

**Program 5.6** Introducing a Skolem function for linearity.

```
(defun-sk linear-dependence-nz (u v)
 (exists a (equal u (scalar-* a v))))
```

## 5.4 Final Statement of the Cauchy-Schwarz Inequality

Program 5.8 displays our final form of Cauchy-Schwarz. The ACL2(r) theorems `cauchy-schwarz-1` and `cauchy-schwarz-2` are equivalent to (Cauchy-Schwarz I) and (Cauchy-Schwarz II), respectively. The theorems `cauchy-schwarz-3` and `cauchy-schwarz-4` correspond to the conditions for equality for (Cauchy-Schwarz I) and (Cauchy-Schwarz II), respectively.

**Program 5.7** The conditions for equality for (Cauchy-Schwarz I).

```
(defthm cauchy-schwarz-3
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (= (* (dot u v) (dot u v)) (* (dot u u) (dot v v)))
                 (or (zvecp u) (zvecp v) (linear-dependence-nz u v)))) ...)
```

**Program 5.8** The final form of Cauchy-Schwarz.

```
(defthm cauchy-schwarz-1
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
                 (<= (* (dot u v) (dot u v))
                     (* (dot u u) (dot v v)))) ...)

(defthm cauchy-schwarz-2
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (<= (abs (dot u v))
              (* (eu-norm u) (eu-norm v)))) ...)

(defthm cauchy-schwarz-3
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (= (* (dot u v) (dot u v))
                    (* (dot u u) (dot v v)))
                 (or (zvecp u) (zvecp v)
                     (linear-dependence-nz u v)))) ...)

(defthm cauchy-schwarz-4
 (implies (and (real-listp u) (real-listp v) (= (len u) (len v)))
          (equal (=  (abs (dot u v)) (* (eu-norm u) (eu-norm v)))
                 (or (zvecp u) (zvecp v)
                     (linear-dependence-nz u v)))) ...)
```

# 6 $\mathbb{R}^n$ as a Metric Space in ACL2(r)

## 6.1 Metric Space Axioms

Observe

$$d_2(x,y) = \|x-y\|_2 = \sqrt{\langle x-y, x-y\rangle}. \tag{30}$$

Proving theorems regarding metrics reduces to proving theorems about the norm from which the metric is induced. Likewise, proving theorems involving norms can be reduced to proving properties about the underlying inner product. The process of formalisation, then, should ideally define the metric via the norm, and the norm should be defined via the inner product. This is useful for proving properties such as positive definiteness of both the metric and the norm, eg. if

$$\langle u,u\rangle \geq 0 \tag{31}$$

with equality iff $u = 0$, then the same applies for

$$\|x\|_2 = \sqrt{\langle x,x\rangle} \geq 0 \tag{32}$$

and so

$$d_2(x,y) = \|x-y\|_2 = \sqrt{\langle x-y, x-y\rangle} \geq 0. \tag{33}$$

However, as exemplified by `vec--` and `vec--x`, the obvious sequence is not necessarily the easiest. Indeed, not only is it simpler to prove theorems on functions equivalent to the desired functions, we also prove properties of similar functions not equivalent to the desired functions but such that if the properties hold for the similar functions, then they also hold for the desired functions. For example, suppose we wish to prove commutativity for the Euclidean metric `eu-metric`. Recall `(vec-- x y)` is a macro for `(vec-+ x (scalar-* -1 y))` and, together with an instance of `acl2-sqrt`, formalising the proofs for commutativity require a non-trivial amount of hints and user guidance. We instead define recursively a function `metric^2`, which is the square of the norm of the difference of two vectors (which is equivalent to the square of the Euclidean metric, i.e. $\|x-y\|_2^2$). Moreover, proving those equivalences simply amounts to unwinding the definitions. Having established

$$\|x-y\|_2^2 = \|y-x\|_2^2, \tag{34}$$

it follows that

$$d_2(x,y) = \sqrt{\|x-y\|_2^2} = \sqrt{\|y-x\|_2^2} = d_2(y,x). \tag{35}$$

Hence, commutativity for `eu-metric` is proven.

**Program 6.1** The Euclidean norm, metric, and squared metric.

```
(defun eu-norm (u)
 (acl2-sqrt (dot u u)))
...
(defun eu-metric (u v)
 (eu-norm (vec-- u v)))
...
(define metric^2 (vec1 vec2) ...
 (norm^2 (vec-- vec1 vec2)) ...)
```

## 6.2 Continuity $\mathbb{R}^n \to \mathbb{R}$

To showcase continuity, let us begin with an enlightening example. Recall the non-standard analysis definition of continuity for a function $f : \mathbb{R}^n \to \mathbb{R}$ stated in the language of non-standard analysis: if $d_2(x,y)$ is an infinitesimal for a standard $x$, then so is $f(x) - f(y)$. Take $f(x) = \sum_{i=1}^n x_i$. It is clear that $f$ is continuous in our usual theory of classical real analysis. However, we must translate this into the language of infinitesimals. By hypothesis,

$$d_2(x,y) = \|x-y\|_2 = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2} \tag{36}$$

is an infinitesimal. We would like to show that

$$f(x) - f(y) = \sum_{i=1}^n x_i - \sum_{i=1}^n y_i = \sum_{i=1}^n (x_i - y_i) \tag{37}$$

is also an infinitesimal. Indeed, by Equation (36) we see that each $x_i - y_i$ must necessarily be infinitesimal since otherwise $d_2(x,y)$ wouldn't be an infinitesimal. Because the RHS of Equation (37) is a finite sum of infinitesimals, so is $f(x) - f(y)$ as desired. The two motivating questions are:

1. How do we make ACL2(r) recognize $x_i - y_i$ are infinitesimals from $d_2(x,y)$ being infinitesimal?
2. How do we state "all $x_i - y_i$ are infinitesimals"?

To answer the first question, observe for any vector $z$ and $i \leq n$,

$$\|z\|_2 = \sqrt{\sum_{i=1}^{n} z_i^2} \geq \max_i |z_i| \geq |z_i|. \tag{38}$$

Setting $z = x - y$, we see that if the norm is an infinitesimal, then so must each entry of the vector. By introducing an ACL2(r) function, say `max-abs-reals`, equivalent to $\max_i$ and reasoning over the arbitrariness of $i$ instead of over the length of $x$ and $y$, we may exhibit the infinitesimality of any entry in $x - y$ as seen in Program 6.2.

**Program 6.2** Showing arbitrary entries of a vector are infinitesimal via the maximum element.

```
(define max-abs-reals ((vec real-listp))
 ...
 (b* (((unless (consp vec)) 0)
       ((cons hd tl) vec)
       ((unless (realp hd)) 0))
      (max (abs hd) (max-abs-reals tl)))...)

(defthm eu-norm-i-small-implies-max-abs-reals-i-small
 (implies (and (real-listp vec) (i-small (eu-norm vec)))
          (i-small (max-abs-reals vec))))

(defthm eu-norm-i-small-implies-elements-i-small
 (implies (and (real-listp vec) (i-small (eu-norm vec)) (natp i) (< i (len vec)))
          (i-small (nth i vec))))
```

To address the second question, one could imagine a recognizer for vectors with infinitesimal entries – such a recognizer is depicted in Program 6.3. However, this recognizer would be recursive on the entries of the vector with each recursive step invoking the non-classical recognizer `i-small` for infinitesimal reals. Because non-classical recursive functions are forbidden, so is the suggested recognizer. A Skolem function was also considered as a possibility but to remain consistent with `eu-norm-i-small-implies-elements-i-small` a theorem positing the condition for an arbitrary index $i$ as seen in Program 6.4 was chosen instead.

**Program 6.3** A fantastical recognizer for vectors with infinitesimal entries that doesn't exist.

```
(defun i-small-vecp (x)
 (cond ((null x) t)
       ((not (real-listp x)) nil)
       (t (and (i-small (car x)) (i-small-vecp (cdr x))))))
```

To see `eu-metric-i-small-implies-difference-of-entries-i-small` in action, consider once again the example of $f(x) = \sum_{i=1}^{n} x_i$. If $n = 3$ and `sum` is the ACL2(r) function equivalent to $f$, then the following is the proof of continuity for `sum`. Other examples of functions with proofs of continuity

**Program 6.4** A theorem positing the infinitesimality of arbitrary entries in $x - y$.

```
(defthm eu-metric-i-small-implies-difference-of-entries-i-small
 (implies (and (real-listp x) (real-listp y) (= (len y) (len x))
               (natp i) (< i (len x)))
               (i-small (eu-metric x y))
          (i-small (- (nth i x) (nth i y)))) ...)
```

in ACL2(r) include the Euclidean norm, the dot product with one coordinate fixed, and the function $g(x,y) = xy$.

**Program 6.5** A theorem positing `sum` is continuous.

```
(defthm sum-is-continuous
 (implies (and (real-listp x) (real-listp y) (= (len x) 3) (= (len y) (len x))
               (i-small (eu-metric x y)))
          (i-small (- (sum x) (sum y))))...)
```

## 7 Conclusion

Firstly, we would like to note the choice of classical proof on which we base this formalisation. In particular, we would like to compare its flavour to other proofs of Cauchy-Schwarz. Indeed, there are geometric proofs, analytical proofs, combinatorial proofs, inductive proofs, etc. [25] whereas we followed a rather algebraic approach. Considering ACL2(r)'s strengths with regards to induction, the choice may seem odd. Indeed, there are several potential inductive candidates we considered at the onset of this endeavor before proceeding with the algebraic approach. However, most of the other candidates inducted over the dimension of $\mathbb{R}^n$ and required reasoning over the real entries of vectors. We suspect unwinding the vectors and guiding ACL2(r) through such a proof would be more onerous than the one outlined in this paper. Moreover, our formalisation of inner product spaces already provided the exact tools necessary for our preferred proof of Cauchy-Schwarz (eg. vectors, vector-vector operations, scalar-vector operations, inner products, etc.) without resorting to reasoning over individual reals. The precision of this approach, while arguably more elegant, also complements our approach to defining continuity. By reasoning over the properties of the vector itself instead of the individual components, we circumvent the soundness-motivated limitations of ACL2(r).

Secondly, this formalisation has two notable purposes. The first is the various applications that may be introduced as a result of the Cauchy-Schwarz inequality. The appearance of Cauchy-Schwarz in functional analysis, real analysis, probability theory, combinatorics, etc. speaks to its utility. A further application of Cauchy-Schwarz is in [10] where a set of theorems involving convex functions are formalised. Additionally, the various structures of $\mathbb{R}^n$ are very rich in mathematical theory and hold applications in various areas of science. In this paper, we presented a formalisation of the space from only two perspectives. However, the choice of perspectives is arguably among the most fundamental. It is the vector space structure of $\mathbb{R}^n$ that provides the necessary operations between its elements. Indeed, one would be hard-pressed to find any view of $\mathbb{R}^n$ that does not assume operations on the domain. Moreover, inner products are the path to calculus: inner products lead to norms; norms lead to metrics; metrics lead to real analysis. The formalisation of $\mathbb{R}^n$ as a metric space is the last step before multivariate calculus, which is in of itself highly applicable and left as future work. We also note the possibility of proving Cauchy-Schwarz for more abstract structures as future work.

During the course of formalisation, emphasis was placed on using algebraic methods to prove theorems that would have otherwise been proved via induction. However, algebraic approaches require significant guidance from the user. Instead, the properties of the inner product space axioms and the proof of Cauchy-Schwarz might be amenable to certification by a SMT solver via `smtlink` [16, 15]. The challenge here is that SMT solvers do not perform induction – we need to leave that for ACL2. On the other hand, we might be able to treat operations on vectors as uninterpreted functions with constraints corresponding to the requirements for a function to be an inner product, a norm, etc.

Finally, we discuss further formalisations of $\mathbb{R}^n$ under different lenses. In fact, there is still potential to further extend $\mathbb{R}^n$ as a metric space. The notions of continuity are independent of the metric used and $d_2$ may be replaced with any metric on $\mathbb{R}^n$. By way of encapsulation, pseudo-higher-order techniques may be employed to easily formalise various real metric spaces – especially if we consider the metric induced by other $p$-norms.

Among the extensions of $\mathbb{R}^n$ as a metric space is proving its completeness. Addressing Cauchy sequences traditionally follows from an application of Bolzano-Weierstrass which has yet to be formalised in ACL2 [20]. Stating completeness in terms infinitesimals and ACL2(r) is a farther but tantalizing prospect. Upon doing so, we would have a formalisation of $\mathbb{R}^n$ as a Hilbert space [17].

## References

[1] ACL2: *A Beginners Guide to Reasoning about Quantification in ACL2*. Available at `https://www.cs.utexas.edu/users/moore/acl2/current/manual/index.html?topic=ACL2____QUANTIFIER-TUTORIAL`. Originally written by Sandip Ray.

[2] Sanaz Khan Afshar, Vincent Aravantinos, Osman Hasan & Sofiène Tahar (2014): *Formalization of Complex Vectors in Higher-Order Logic*. In Stephen M. Watt, James H. Davenport, Alan P. Sexton, Petr Sojka & Josef Urban, editors: *Intelligent Computer Mathematics*, Springer International Publishing, Cham, pp. 123–137, doi:10.1023/A:1012692601098.

[3] Leif O. Arkeryd, Nigel J. Cutland & C. Ward Henson (Eds.) (1997): *Nonstandard Analysis: Theory and Applications*, 1st edition. *Nato Science Series C: 493* , Springer Netherlands, doi:10.1007/978-94-011-5544-1.

[4] John Cowles & Ruben Gamboa (2017): *The Cayley-Dickson Construction in ACL2*. In Anna Slobodova & Warren Hunt, Jr., editors: Proceedings 14th International Workshop on the *ACL2 Theorem Prover and its Applications*, Austin, Texas, USA, May 22-23, 2017, *Electronic Proceedings in Theoretical Computer Science* 249, Open Publishing Association, pp. 18–29, doi:10.4204/EPTCS.249.2.

[5] Ruben A. Gamboa & Matt Kaufmann (2001): *Nonstandard Analysis in ACL2*. *Journal of Automated Reasoning* 27(4), pp. 323–351, doi:10.1023/A:1011908113514.

[6] John Harrison (2005): *A HOL Theory of Euclidean Space*. In Joe Hurd & Tom Melham, editors: *Theorem Proving in Higher Order Logics*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 114–129, doi:10.1007/11541868_8.

[7] Nathan Jacobson (1985): *Basic Algebra I*, 2nd edition. Dover Publications.

[8] Nathan Kahl: *The Hundred Greatest Theorems*. Online. Available at `http://pirate.shu.edu/~kahlnath/Top100.html`. Originally published by Paul and Jack Abad (1999).

[9] Gerwin Klein: *The Top 100 Theorems in Isabelle*. Online. Available at `https://www.cse.unsw.edu.au/~kleing/top100/#78`.

[10] Carl Kwan & Mark R. Greenstreet (2018): *Convex Functions in ACL2(r)*. In: Proceedings 15th International Workshop on the *ACL2 Theorem Prover and its Applications*, Austin, Texas, USA, November 5-6, 2018, *Electronic Proceedings in Theoretical Computer Science* 280, Open Publishing Association, pp. 128–142, doi:10.4204/EPTCS.280.10.

[11] Serge Lang (2002): *Algebra*, 3rd edition. *Graduate Texts in Mathematics 211* , Springer-Verlag New York, doi:10.1007/978-1-4613-0041-0.

[12] Peter A. Loeb & Manfred P. H. Wolff (2015): *Nonstandard Analysis for the Working Mathematician*, 2nd edition. Springer Netherlands, doi:10.1007/978-94-017-7327-0.

[13] Jean-Marie Madiot: *Formalizing 100 theorems in Coq*. Online. Available at `https://madiot.fr/coq100/#78`.

[14] Marco Maggesi (2018): *A Formalization of Metric Spaces in HOL Light*. *Journal of Automated Reasoning* 60(2), pp. 237–254, doi:10.1007/s10817-017-9412-x.

[15] Yan Peng & Mark Greenstreet (2015): *Integrating SMT with Theorem Proving for Analog/Mixed-Signal Circuit Verification*. In Klaus Havelund, Gerard Holzmann & Rajeev Joshi, editors: *NASA Formal Methods*, Springer International Publishing, Cham, pp. 310–326, doi:10.1007/978-3-319-17524-9_22.

[16] Yan Peng & Mark R. Greenstreet (2015): *Extending ACL2 with SMT Solvers*. In: *Proceedings Thirteenth International Workshop on the ACL2 Theorem Prover and Its Applications, Austin, Texas, USA, 1-2 October 2015.*, pp. 61–77, doi:10.4204/EPTCS.192.6.

[17] Frigyes Riesz & Bela Sz.-Nagy (1990): *Functional Analysis*. Dover Publications.

[18] Abraham Robinson (1966): *Non-Standard Analysis*. North-Holland Publishing Company.

[19] Steven Roman (2008): *Advanced Linear Algebra*, 3rd edition. *Graduate Texts in Mathematics 135* , Springer-Verlag New York, doi:10.1007/978-0-387-72831-5.

[20] Walter Rudin (1976): *Principles of Mathematical Analysis*, 3rd edition. *International Series in Pure and Applied Mathematics* , McGraw-Hill.

[21] Georgi E. Shilov (1977): *Linear Algebra*. Dover Publications.

[22] J. Michael Steele (2004): *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press, doi:10.1017/CBO9780511817106.

[23] Jasper Stein (2001): *Documentation for the formalization of Linerar Agebra*. Online. Available at `http://www.cs.ru.nl/~jasper/`.

[24] Freek Wiedijk: *Formalizing 100 Theorems*. Online. Available at `http://www.cs.ru.nl/~freek/100`.

[25] Hui-Hua Wu & Shanhe Wu (2009): *Various proofs of the Cauchy-Schwarz inequality*. *Octogon Mathematical Magazine* 17(1), pp. 221–229.

## Appendix A Why are Non-classical Recursive Functions Prohibited?

A discussion with Ruben Gamboa[1] and members of the ACL2 Help Mailing List sheds light on why non-classical recursive functions are prohibited. In summary, the introduction of such functions will also introduce inconsistency into the logic of ACL2(r). It is possible to define a function that violates the rules of non-standard analysis.

For example, consider the hypothetical function defined in Program A.1. Suppose `f` terminates. If `n` is standard, then `f` returns it without issue. If `n` is infinite, then `f` returns the largest standard number. However, this is impossible since such a number does not exist and, if `n` is infinite, so is `(-1 n)` which means `f` should not have terminated anyways. Moreover, this applies if `n` is any non-standard hyperreal

[1] We would like to thank Ruben Gamboa for his insightful explanations on which most of this appendix is based.

**Program A.1** An impossible function in ACL2(r).

```
(defun f (n)
 (cond ((zp n) 0)
       ((standardp n) n)
       (t (f (-1 n)))))
```

since if $n = \mathrm{st}(n) + \varepsilon$ is equivalent to `n` and $\varepsilon > 0$ is an infinitesimal, then $n - 1 = \mathrm{st}(n-1) + \varepsilon$ is not standard either.

The essence of the issue with `f` is that its measure is non-standard. If the measure of a function can be proven to be standard, then a recursive non-classical function could be conceded. However, in practice, such a proof would likely be subtle and involved.