# Cauchy-Schwarz Inequality

Krish Nigam *

July, 2021

## Introduction

Why have I dedicated a whole handout to **one inequality**! This must be an *important* result, and I can assure you it is, as you will find out. But first, to approach this inequality most generally, formally and rigorously, we will have to define various ideas such as the **vector space** and **inner product**.

## Contents



*Email: kn221@ic.ac.uk

# §1 Vector Spaces

We begin by defining the notion of a **vector space**, however before that we must also be clear on the algebraic structures of **fields**.

## §1.1 Fields and a Vector Space

**Fields** are *algebraic structures* which come as an **extension** of the concept of **groups**, since these are structures that are characterised by **two** *operations*.[1]

**Definition 1.1** (Field) — A *field* is a set $\mathbb{F}$ with two operations

$$\begin{aligned} + &: \mathbb{F} \times \mathbb{F} \to \mathbb{F}, & (a, b) &\mapsto a + b, & &(\textit{addition}) \\ * &: \mathbb{F} \times \mathbb{F} \to \mathbb{F}, & (a, b) &\mapsto a * b, & &(\textit{multiplication}) \end{aligned}$$

that satisfy the following axioms:

(i) $(\mathbb{F}, +)$ is a **commutative group** (also known as an *Abelian group*). We call the **identity** element in this group *zero*, and write 0. We denote the **inverse** element of $a \in \mathbb{F}$ by $-a$.

(ii) $(\mathbb{F} \setminus \{0\}, *)$ is a **commutative group**.[a] We call the **identity** element in this group *unit*, and write 1. We denote the **inverse** element of $a \in \mathbb{F} \setminus \{0\}$ by $a^{-1}$.

(iii) The **distributive** laws hold, *i.e.* for all $a, b, c \in \mathbb{F}$ we have

$$\begin{aligned} a * (b + c) &= a * b + a * c \\ (a + b) * c &= a * c + b * c \end{aligned}$$

[a] Note that the set notation $\mathbb{F} \setminus \{0\}$ means 'the set of all elements of $\mathbb{F}$ *excluding* 0'.

If the reader is aware of **rings**, a **commutative** ring $R$ with unit is called a *field*, if $0 \neq 1$ and every $a \in R \setminus \{0\}$ is **invertible**. We may also say that every *field* is a *commutative ring* with **unit**, but the converse does not hold.

With these fundamental ideas done, we can now use the *field* in our definition of a **vector space**.

*Intuitively*, a *vector space* $\mathcal{V}$ over a *field* $\mathbb{F}$ (also known as an '$\mathbb{F}$-vector space') is a **space** with two **operations**:

- We can add two vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ to obtain $\mathbf{u} + \mathbf{v} \in \mathcal{V}$.
- We can multiply a scalar $\lambda \in \mathbb{F}$ with a vector $\mathbf{v} \in \mathcal{V}$ to obtain $\lambda \mathbf{v} \in \mathcal{V}$.

However, these two operations must satisfy certain axioms before we can call it a vector space, so let's formalise this now.

**Definition 1.2** (Vector Space) — An $\mathbb{F}$-*vector space* (or a vector space over $\mathbb{F}$) is an (additive) **commutative group** $(\mathcal{V}, +)$ together with a function called *scalar*

[1] In actual fact, *fields* are a special type of **ring**, but for the purposes of this document we do not need to be worried about rings, and a field can be defined without it.

*multiplication*:

$$\mathbb{F} \times \mathcal{V} \to \mathcal{V}$$
$$(\lambda, \mathbf{v}) \mapsto \lambda \cdot \mathbf{v}$$

satisfying the axioms:

(i) $\lambda(\mu\mathbf{v}) = \lambda\mu\mathbf{v}$ for all $\lambda, \mu \in \mathbb{F}$, $\mathbf{v} \in \mathcal{V}$. (*associativity*)

(ii) $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$ for all $\lambda \in \mathbb{F}$, $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. (*distributivity in* $\mathcal{V}$)

(iii) $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$ for all $\lambda, \mu \in \mathbb{F}$, $\mathbf{v} \in \mathcal{V}$. (*distributivity in* $\mathbb{F}$)

(iv) $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in \mathcal{V}$. (*identity*)

## §1.2 Inner Product

When generalised to all vector spaces, the Cauchy-Schwarz inequality involves use of the **inner product**.

The *inner product* is a **binary operation**[2], which associates each pair of vectors in the space with a **scalar quantity** known as the inner product of the vectors, often denoted with *angle brackets*.

**Definition 1.3** (Inner Product) — Let $\mathcal{V}$ be a *vector space* over *field* $\mathbb{F}$. An *inner product* $\langle \cdot, \cdot \rangle$ is a function $\mathcal{V} \times \mathcal{V} \to \mathbb{F}$ satisfying:

(i) for all $\mathbf{v} \in \mathcal{V}$, $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \iff \mathbf{v} = 0$.

(ii) for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$, $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

(iii) for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$ and for all $a, b \in \mathbb{F}$, $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{w} \rangle$

It follows from this definition that an **inner product space** is a *vector space* $\mathcal{V}$ over the *field* $\mathbb{F}$ together with an *inner product*, given by the map:

$$\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{F}.$$

## §1.3 Norm and Distance

In the Cauchy-Schwarz inequality we also have the **norm** at play.

The *norm* can be thought of as a notion of **length** of a vector.

**Definition 1.4** (Norm) — Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an *inner product space*. The **norm function** is a function $\mathcal{V} \to \mathbb{F}$ denoted as $\|\cdot\|$, and given by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

[2] An operation whose input(s) and output(s) are in the same set (so preserves *closure*).

**Example 1.5** (Euclidean Norm) The **Euclidean norm** in $\mathbb{R}^n$ is given by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{(v_1)^2 + (v_2)^2 + \cdots + (v_n)^2}$$

A *norm* in a *vector space* then induces the notion of **distance** between two vectors, which is defined by the **length** of their **difference**.

**Definition 1.6** (Distance) — Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an *inner product space*, and $\|\cdot\|$ be its associated *norm*. The **distance** between $\mathbf{u}$ and $\mathbf{v} \in \mathcal{V}$ is given by

$$\text{distance}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

**Example 1.7** (Euclidean Distance) The **Euclidean distance** between two vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ is given by

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$

## §1.4 Sets of Vectors

When we are dealing with **sets** of *vectors*, as we are, there are some important terminology to be aware of, specifically **linearly dependent** and **linearly independent** (the former more so than the latter in our case).

**Definition 1.8** (Linearly Dependent) — A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ from a vector space $\mathcal{V}$ is *linearly dependent* if there exist *scalars* $\lambda_1, \lambda_2, \ldots, \lambda_n$, not all zero, such that

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

where $\mathbf{0}$ denotes the zero vector.

Thus, a set of vectors is linearly dependent *if and only if* each vector in the set can be written as a **linear combination** of the others.

**Definition 1.9** (Linearly Independent) — A set of vectors is *linearly independent* if it is **not** *linearly dependent*, that is if the equation

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

can only be satisfied by $\lambda_i = 0$ for $i = 1, \ldots, n$.

# §2 Cauchy-Schwarz

**Theorem 2.1** (Cauchy-Schwarz Inequality) Let **u** and **v** be arbitrary vectors in an **inner product space** over the **scalar field** $\mathbb{F}$, where $\mathbb{F}$ is the field of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$. It follows that,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$

and equality holds *if and only if* **u** and **v** are **linearly dependent**.

**Remark 2.2** If we have this set of two vectors $\mathbf{u}, \mathbf{v}$, these vectors are linearly dependent *if and only if* they are **collinear** *i.e.* one is a **scalar multiple** of the other.

## §2.1 Proof I

*Proof* (I). We begin by definition of the *inner product* that $\langle \mathbf{v} + \lambda \mathbf{w}, \mathbf{v} + \lambda \mathbf{w} \rangle \geq 0$ for some scalar $\lambda$.

$$\langle \mathbf{u} + \lambda \mathbf{v}, \mathbf{u} + \lambda \mathbf{v} \rangle \geq 0$$

It then follows from the *linearity of the inner product* that,

$$\langle \mathbf{u}, \mathbf{u} \rangle + 2\lambda \langle \mathbf{u}, \mathbf{v} \rangle + \lambda^2 \langle \mathbf{v}, \mathbf{v} \rangle \geq 0$$

Here, we have a **quadratic** in $\lambda$, which is greater than or equal to 0. As a result, there is a repeated root or no $\mathbb{R}$ solutions for $\lambda$ and thus we have the discriminant $\leq 0$.

$$\therefore \quad 4\lambda^2 \langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\lambda^2 \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle \leq 0$$

$$\langle \mathbf{u}, \mathbf{v} \rangle^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{u}\|^2 \, \|\mathbf{v}\|^2$$

Hence, $\boxed{|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|}$. □

## §2.2 Proof II

*Proof* (II). Let's define a vector **z** as follows,

$$\mathbf{z} := \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$$

Now, let's consider $\langle \mathbf{z}, \mathbf{v} \rangle$. It follows from the *linearity of the inner product* that,

$$\langle \mathbf{z}, \mathbf{v} \rangle = \left\langle \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \right\rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{v} \rangle = 0$$

This shows that **z** is *orthogonal* to **v**, and hence **u** is written as the sum of two orthogonal vectors, so we can apply Pythagoras'.

$$\begin{aligned}\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v} + \mathbf{z} \implies \|\mathbf{u}\|^2 &= \left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\right)^2 \|\mathbf{v}\|^2 + \|\mathbf{z}\|^2 \\ &= \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{(\|\mathbf{v}\|)^2} \|\mathbf{v}\|^2 + \|\mathbf{z}\|^2 \\ &= \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{v}\|^2} + \|\mathbf{z}\|^2 \\ &\geq \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{v}\|^2}\end{aligned}$$

Hence, $\boxed{|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|}$ . □

## §2.3 Proof III

*Proof* (III). We can consider the expression $\frac{1}{\|\mathbf{v}\|^2} \left\| \|\mathbf{v}\|^2 \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v} \right\|^2$ for vectors $\mathbf{u}, \mathbf{v}$.

We can expand this expression by the definition of the *norm* and then the *linearity of the inner product*.

$$\begin{aligned}\frac{1}{\|\mathbf{v}\|^2} \left\| \|\mathbf{v}\|^2 \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v} \right\|^2 &= \frac{1}{\|\mathbf{v}\|^2} \left\langle \|\mathbf{v}\|^2 \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v},\ \|\mathbf{v}\|^2 \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v} \right\rangle \\ &= \frac{1}{\|\mathbf{v}\|^2} \left( \|\mathbf{v}\|^4 \langle \mathbf{u}, \mathbf{u} \rangle - 2 \|\mathbf{v}\|^2 \langle \mathbf{u}, \mathbf{v} \rangle \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle^2 \langle \mathbf{v}, \mathbf{v} \rangle \right) \\ &= \frac{1}{\|\mathbf{v}\|^2} \left( \|\mathbf{v}\|^4 \|\mathbf{u}\|^2 - 2 \|\mathbf{v}\|^2 \langle \mathbf{u}, \mathbf{v} \rangle^2 + \|\mathbf{v}\|^2 \langle \mathbf{u}, \mathbf{v} \rangle^2 \right) \\ &= \|\mathbf{u}\|^2 - \langle \mathbf{u}, \mathbf{v} \rangle^2\end{aligned}$$

Now, observe that the original expression $\frac{1}{\|\mathbf{v}\|^2} \left\| \|\mathbf{v}\|^2 \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v} \right\|^2 \geq 0$, and so

$$\|\mathbf{u}\|^2 - \langle \mathbf{u}, \mathbf{v} \rangle^2 \geq 0$$

Hence, $\boxed{|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|}$ . □

# §3 Other Results

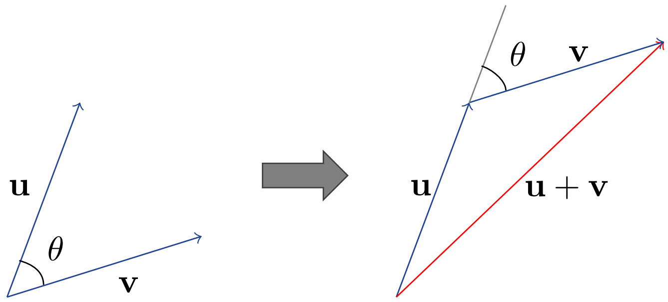
## §3.1 Triangle Inequality

**Corollary 3.1** (Triangle Inequality) Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let $\|\cdot\|$ denote the **norm** of a vector. Then,

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

**Remark 3.2** Notice that here also, equality holds *if and only if* $\mathbf{u}$ and $\mathbf{v}$ are *linearly dependent*, which corresponds to $\mathbf{u}, \mathbf{v}$ being **collinear**, and so the triangle is **degenerate**.

The following diagram illustrates the *triangle inequality*, for some two vectors $\mathbf{u}, \mathbf{v}$, which may seem like an 'obvious result'.




Clearly, if we reduce the angle $\theta$ to 0, then the vectors $\mathbf{u}, \mathbf{v}$ become collinear and this is where we obtain our degenerate triangle case.

## §3.2 Probability Theory

Let $X$ and $Y$ be **random variables**. We can define an *inner product* on the **set** of random variables using the **expectation** of their product.

$$\langle X, Y \rangle := \mathrm{E}(XY)$$

The *Cauchy-Schwarz inequality* then becomes:

**Theorem 3.3** (Cauchy-Schwarz for the expectation of random variables)

$$|\mathrm{E}(XY)|^2 \leq \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$

We can use this to prove the following corollary known as the **covariance inequality**.

**Corollary 3.4** (Covariance Inequality)

$$\mathrm{Cov}(X,Y)^2 \leq \mathrm{Var}(X)\,\mathrm{Var}(Y)$$

*Proof.*

$$\begin{aligned}\operatorname{Cov}(X,Y)^2 &= \mathrm{E}\left((X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right)^2 \\ &\leq \mathrm{E}((X-\mathrm{E}(X))^2)\cdot\mathrm{E}((Y-\mathrm{E}(Y))^2) \\ &= \operatorname{Var}(X)\cdot\operatorname{Var}(Y)\end{aligned}$$

□

## §3.3 Hölder's Inequality

It may be said that *Cauchy-Schwarz* is a **special case** of **Hölder's Inequality**, or it may be that Hölder's Inequality is a **generalisation** of *Cauchy-Schwarz*.

Either way, the two inequalities are inextricably linked.

**Theorem 3.5** (Hölder's Inequality) For sequences $a_i, b_i, \dots, z_i \in \mathbb{R}$, Hölder's Inequality states that

$$(a_1+a_2+\cdots+a_n)^{\lambda_a}\dots(z_1+z_2+\cdots+z_n)^{\lambda_z} \geq a_1^{\lambda_a}b_1^{\lambda_b}\dots z_1^{\lambda_z}+\cdots+a_n^{\lambda_a}b_n^{\lambda_b}\dots z_n^{\lambda_z}$$

for all $\lambda_a+\lambda_b+\cdots+\lambda_z=1$.

[*In the case of* $\lambda_a=\lambda_b=\frac{1}{2}$, *the inequality reduces to Cauchy-Schwarz.*]

**Remark 3.6** A more rigorous and generalised definition involves first defining the notion of a **measure space**, so for the purposes of this document I have only brushed the surface of Hölder's inequality.

# §4 Examples

## §4.1 Olympiads

We can convert our *Cauchy-Schwarz inequality* as expressed in theorem 2.1 into one which applies to **sequences** of **real numbers** by picking the *vector space* $\mathbb{R}^n$.

**Theorem 4.1** (Cauchy-Schwarz for Sequences) For all real numbers $a_i$ and $b_i$, we have

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) \geq \left(\sum_{i=1}^{n} a_i b_i\right)^2$$

with equality if and only if $a_i = kb_i$ for some constant $k \in \mathbb{R}^+$, for all $i = 1, \ldots, n$ which have $a_i b_i \neq 0$.

**Example 4.2** (IMO 1995) Let $a, b, c$ be positive real numbers such that $abc = 1$. Prove that

$$\frac{1}{a^3(b+c)} + \frac{1}{b^3(c+a)} + \frac{1}{c^3(a+b)} \geq \frac{3}{2}$$

*Solution.* We begin by making the substitution $x = \frac{1}{a}$, $y = \frac{1}{b}$, $z = \frac{1}{c}$. Then by the given condition, $xyz = 1$. Now, notice

$$\begin{aligned}
\frac{1}{a^3(b+c)} + \frac{1}{b^3(c+a)} + \frac{1}{c^3(a+b)} &= \frac{1}{\frac{1}{x^3}\left(\frac{1}{y}+\frac{1}{z}\right)} + \frac{1}{\frac{1}{y^3}\left(\frac{1}{z}+\frac{1}{x}\right)} + \frac{1}{\frac{1}{z^3}\left(\frac{1}{x}+\frac{1}{y}\right)} \\
&= \frac{x^2}{\frac{1}{x}\left(\frac{y+z}{yz}\right)} + \frac{y^2}{\frac{1}{y}\left(\frac{z+x}{zx}\right)} + \frac{z^2}{\frac{1}{z}\left(\frac{x+y}{xy}\right)} \\
&= \frac{x^2}{y+z} + \frac{y^2}{z+x} + \frac{z^2}{x+y}
\end{aligned}$$

But by *Cauchy-Schwarz*, if we go back to theorem 4.1, we can let $a_1 = \sqrt{x+y}, a_2 = \sqrt{y+z}, a_3 = \sqrt{z+x}$ and $b_1 = \frac{x}{\sqrt{y+z}}, b_2 = \frac{y}{\sqrt{z+x}}, b_3 = \frac{z}{\sqrt{x+y}}$. This then directly allows us to apply the inequality, giving

$$\begin{aligned}
[(x+y)+(y+z)+(z+x)] \cdot \left(\frac{x^2}{y+z} + \frac{y^2}{z+x} + \frac{z^2}{x+y}\right) &\geq (x+y+z)^2 \\
\implies \frac{x^2}{y+z} + \frac{y^2}{z+x} + \frac{z^2}{x+y} &\geq \frac{x+y+z}{2}
\end{aligned}$$

It then follows from the *AM-GM inequality* that $\frac{x+y+z}{2} \geq \frac{3}{2}\sqrt[3]{xyz} = \frac{3}{2}$.

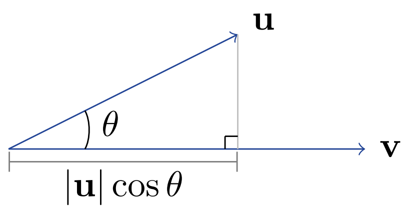
Equality holds *if and only if* $x = y = z = 1 \iff a = b = c = 1$. □

**Remark 4.3** If you are aware of **convex** functions and **Jensen's inequality**, we may also have proceeded from attempting to show $\frac{x^2}{y+z} + \frac{y^2}{z+x} + \frac{z^2}{x+y} \geq \frac{3}{2}$ by applying *Jensen's* to the function $\mathrm{f}(x, y) = \frac{x^2}{y}$.

## §4.2 Vectors

We can visualise the *Cauchy-Schwarz inequality* applied to two vectors in $\mathbb{R}^2$, as shown below.

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos\theta$$




In the above, we use the fact that the *inner product* in $\mathbb{R}^2$ is the **dot product** and this is given by the formula above. And, since $\cos\theta \leq 1$, $\mathbf{u} \cdot \mathbf{v} \leq |\mathbf{u}||\mathbf{v}|$.

**Example 4.4** Given that vectors $\mathbf{x}$ and $\mathbf{y}$ satisfy

$$\mathbf{x} + \mathbf{y}\,(\mathbf{x} \cdot \mathbf{y}) = \mathbf{a}$$

for a *fixed* vector $\mathbf{a}$, show that

$$|\mathbf{x}|\left(1 + |\mathbf{y}|^2\right) \geq |\mathbf{a}| \geq |\mathbf{x}|$$

and explain the circumstances under which equality is achieved, and describe the relation between $\mathbf{x}, \mathbf{y}$ and $\mathbf{a}$ in these circumstances.

*Solution.* Let's first consider $|\mathbf{a}|^2$.

$$\begin{aligned} |\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a} &= (\mathbf{x} + \mathbf{y}(\mathbf{x} \cdot \mathbf{y})) \cdot (\mathbf{x} + \mathbf{y}(\mathbf{x} \cdot \mathbf{y})) \\ &= \mathbf{x} \cdot \mathbf{x} + 2(\mathbf{x} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{y}) + \mathbf{y} \cdot \mathbf{y}(\mathbf{x} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{y}) \\ &= |\mathbf{x}|^2 + (\mathbf{x} \cdot \mathbf{y})^2\left(2 + |\mathbf{y}|^2\right) \end{aligned}$$

$$\therefore \;\; (\mathbf{x} \cdot \mathbf{y})^2 = \frac{|\mathbf{a}|^2 - |\mathbf{x}|^2}{2 + |\mathbf{y}|^2}$$

But, by *Cauchy-Schwarz*, $(\mathbf{x} \cdot \mathbf{y})^2 \leq |\mathbf{x}|^2|\mathbf{y}|^2$.

$$\begin{aligned} \implies \frac{|\mathbf{a}|^2 - |\mathbf{x}|^2}{2 + |\mathbf{y}|^2} &\leq |\mathbf{x}|^2|\mathbf{y}|^2 \\ |\mathbf{a}|^2 - |\mathbf{x}|^2 &\leq |\mathbf{x}|^2|\mathbf{y}|^4 + 2|\mathbf{x}|^2|\mathbf{y}|^2 \\ \therefore \;\; |\mathbf{a}|^2 &\leq |\mathbf{x}|^2\left(1 + |\mathbf{y}|^2\right)^2 \implies \underline{|\mathbf{a}| \leq |\mathbf{x}|\left(1 + |\mathbf{y}|^2\right)} \end{aligned}$$

We also know that $(\mathbf{x} \cdot \mathbf{y})^2 \geq 0 \implies \frac{|\mathbf{a}|^2 - |\mathbf{x}|^2}{2 + |\mathbf{y}|^2} \geq 0$. But since $2 + |\mathbf{y}|^2 > 0$, we must have $|\mathbf{a}| \geq |\mathbf{x}|$.

$$\therefore \;\; \boxed{|\mathbf{x}| \leq |\mathbf{a}| \leq |\mathbf{x}|\left(1 + |\mathbf{y}|^2\right)}$$

Equality is achieved when $|\mathbf{x}| = |\mathbf{a}|$ and $\mathbf{y} = 0$, provided $|\mathbf{x}| \neq 0$.

If $\mathbf{x} = \mathbf{0}$, for equality with $|\mathbf{a}|$, we require $\mathbf{a} = \mathbf{0}$.

If $\mathbf{x} \neq \mathbf{a}$, then $|\mathbf{a}| = |\mathbf{x}|\left(1 + |\mathbf{y}|^2\right)$. There exist infinite solutions $(|\mathbf{a}|, |\mathbf{x}|, |\mathbf{y}|)$.
*e.g.* $|\mathbf{a}| = 4, |\mathbf{x}| = 2, |\mathbf{y}| = 1$. □

## §4.3 Integrals

**Theorem 4.5** (Cauchy-Schwarz for Integrals) For 'well-behaved' functions $\mathrm{f}, \mathrm{g} : [a, b] \mapsto \mathbb{R}$,

$$\left(\int_a^b \mathrm{f}(x)\,\mathrm{g}(x)\,\mathrm{d}x\right)^2 \leq \left(\int_a^b (\mathrm{f}(x))^2\,\mathrm{d}x\right)\left(\int_a^b (\mathrm{g}(x))^2\,\mathrm{d}x\right)$$

**Example 4.6** Show that for $t > 0$,

$$\frac{e^t - 1}{e^t + 1} \leq \frac{t}{2}$$

*Solution.* We can use theorem 4.5, by choosing $\mathrm{f}(x), \mathrm{g}(x), a, b$ appropriately. Let's begin by setting $\mathrm{f}(x) = 1$, $\mathrm{g}(x) = e^x$.

$$\therefore \left(\int_a^b e^x\,\mathrm{d}x\right)^2 \leq \left(\int_a^b \mathrm{d}x\right)\left(\int_a^b e^{2x}\,\mathrm{d}x\right)$$
$$\implies \left(e^b - e^a\right)^2 \leq (b - a) \cdot \frac{1}{2}\left(e^{2b} - e^{2a}\right)$$

So, from this let's try $a = 0, b = t$.

$$\left(e^t - 1\right)^2 \leq \frac{t}{2}\left(e^{2t} - 1\right)$$

Now we can use the fact that we require $t > 0$, so $e^t - 1 > 0$, and by *difference of two squares* factorisation on the RHS, we get

$$e^t - 1 \leq \frac{t}{2}\left(e^t + 1\right) \implies \boxed{\frac{e^t - 1}{e^t + 1} \leq \frac{t}{2}}$$

□