How to Write Better Proofs

Dennis S. Bernstein

University of Michigan

Graduate Student Goals

- Learn new things
  - Take courses
  - Read books and papers
  - Attend seminars and conferences
  - Watch YouTube
  - Talk to people

- Build skills
  - Write code
  - Construct and run experiments (numerical or physical)
  - Make slides and give talks at conferences
  - Write and publish papers [possibly with **proofs**]
  - Write essays and proposals for applications
  - Teach, guide, mentor other students

Mathematics in Engineering

- Mathematical concepts are idealizations
- Often sharply defined
    - Continuous function, open set, orthogonal vectors
  - Idealizations are not physically verifiable
    - They are abstractions that guide our thinking about the real world

- Engineering “assumptions” are approximations
  - Water is incompressible; steel is rigid
  - All engineering models are approximations
  - The “correct” choice of model depends on how it will be used

- Approximation is a big part of mathematics
  - Metrics, norms, bounds
  - Asymptotic analysis and convergence

The art of engineering is to be able to bridge the subtle divide between mathematical idealizations and physical reality

Elements of Mathematics

- Definitions
- Theorems
- Proofs
- Examples

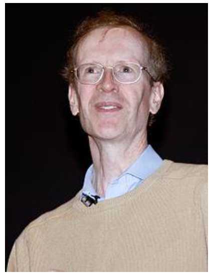
Proofs in Mathematics

- What is a proof?
  - A sequence of statements that verifies the correctness of a theorem
  - This is like saying “a poem is a bunch of words”

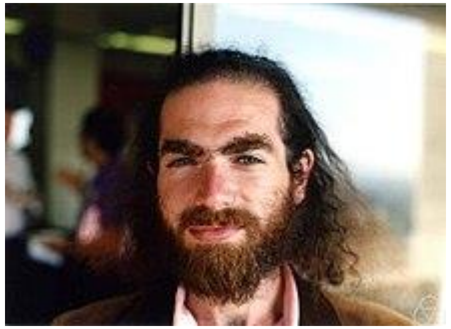
- Why do mathematicians like proofs?
  - Proofs provide the foundation of rigorous mathematics
  - Without proofs, nothing in mathematics would be known with certainty
    - Everything would be a conjecture
  - Proofs ***are*** mathematics
-

Famous Proofs

- Theorem (Fermat): If k, l, m, n are positive integers and $n \geq 3$, then $k^n + l^n \neq m^n$
  - $n = 2$: $3^2 + 4^2 = 5^2$
  - $n = 3$: $k^3 + l^3 \neq m^3$
  - Proof took 358 years
  - No prize

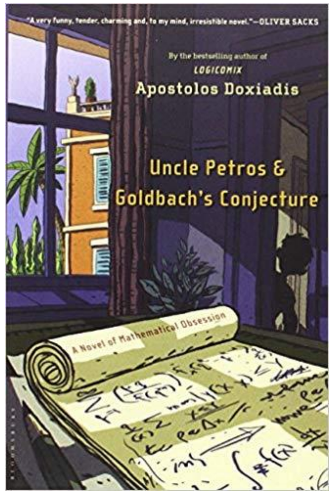


- Theorem (Poincare): Every simply connected closed 3-manifold is homeomorphic to the 3-sphere
  - Proof took 106 years
  - Won $1,000,000 prize
    - Not claimed



Famous Conjectures

- Conjecture 1 (Goldbach): Every even integer is the sum of two primes
  - Example: 34 = 11+23
  - Not proved for 276 years
  - Might be false but likely is true
  - No prize



"A very funny, tender, charming and, to my mind, irresistible novel."—OLIVER SACKS

By the bestselling author of LOGICOMIX

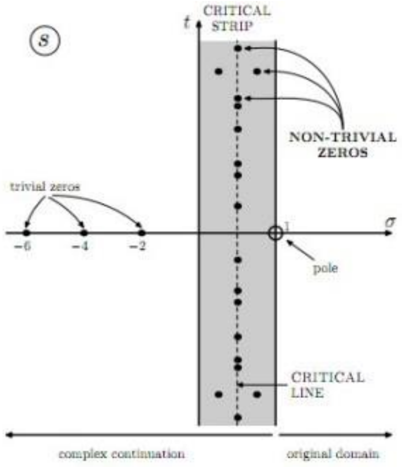
Apostolos Doxiadis

Uncle Petros & Goldbach's Conjecture

A Novel of Mathematical Obsession

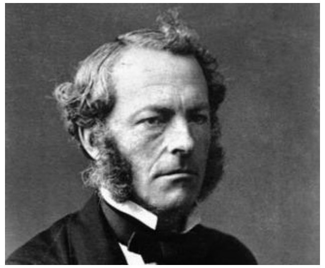
- Conjecture 2 (Riemann): All nontrivial zeros of the zeta function have real part $\frac{1}{2}$
  - Not proved for 159 years
  - The zeros are deeply related to the prime numbers
  - Might be false
  - $1,000,000 prize
  - Not likely to be "Starbucks-solvable"

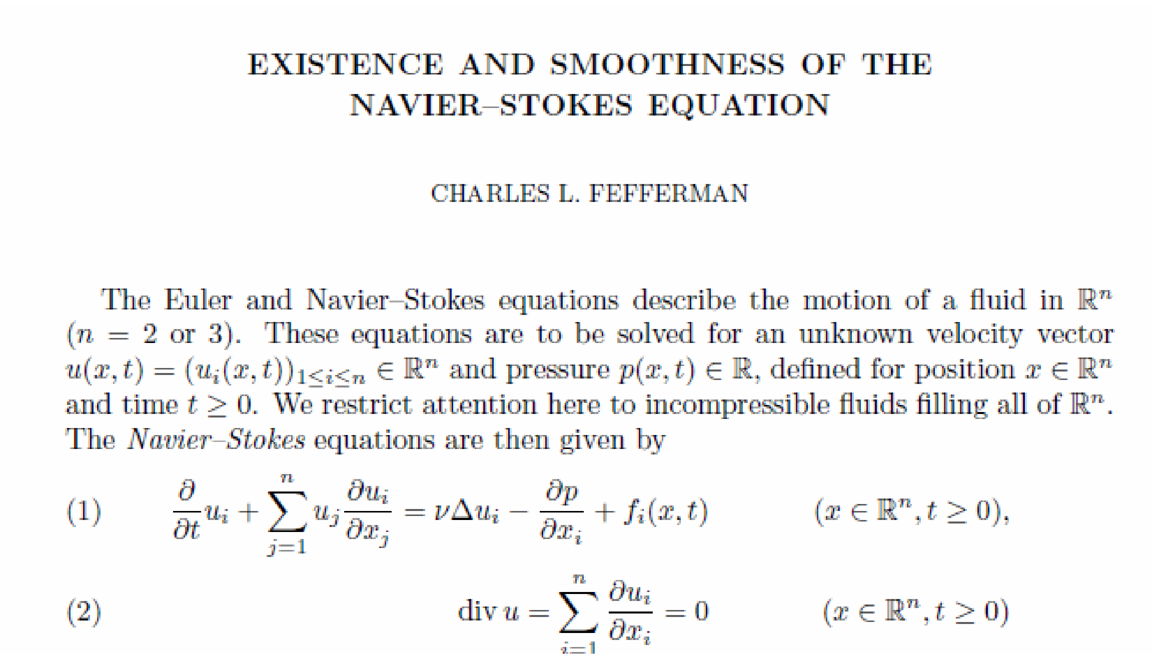
$$\zeta(z) \triangleq \sum_{i=1}^{\infty} \frac{1}{i^z}.$$



Famous Conjectures

- Conjecture 3: The Navier-Stokes equation has a solution
  - $1,000,000 prize
  - Relevant to aerospace engineering!





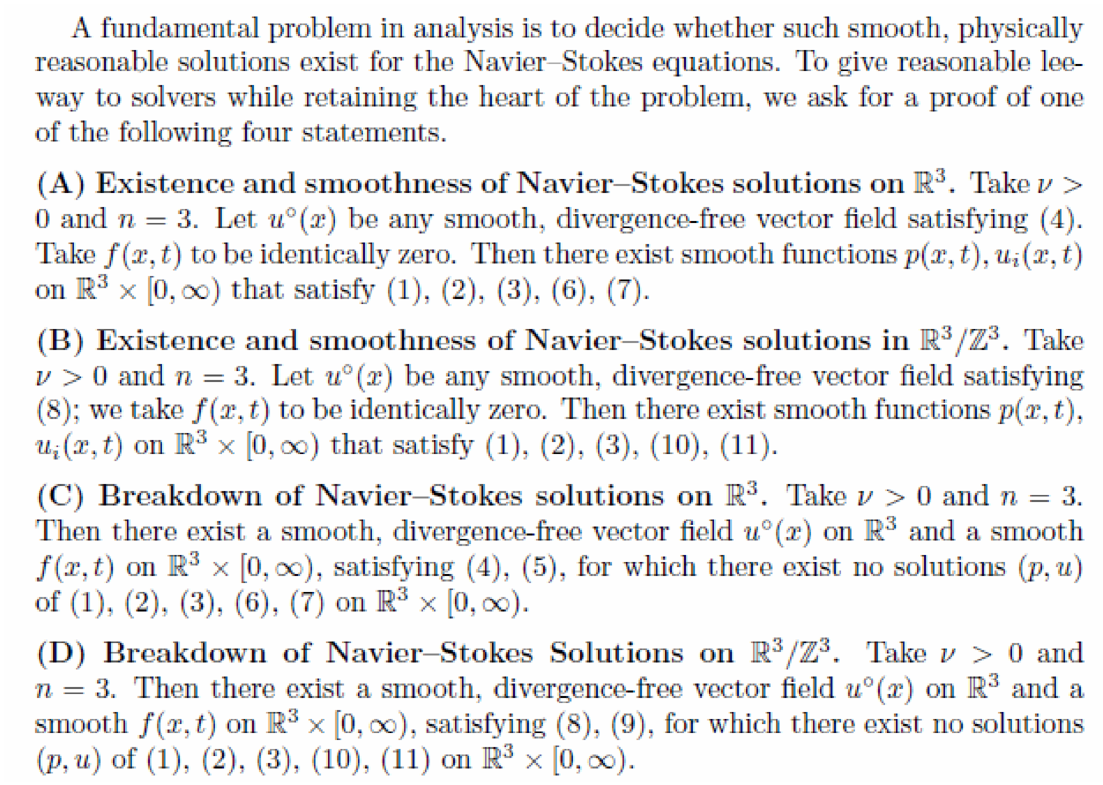
EXISTENCE AND SMOOTHNESS OF THE NAVIER–STOKES EQUATION

CHARLES L. FEFFERMAN

The Euler and Navier–Stokes equations describe the motion of a fluid in $\mathbb{R}^n$ ($n = 2$ or 3). These equations are to be solved for an unknown velocity vector $u(x,t) = (u_i(x,t))_{1\le i\le n} \in \mathbb{R}^n$ and pressure $p(x,t) \in \mathbb{R}$, defined for position $x \in \mathbb{R}^n$ and time $t \ge 0$. We restrict attention here to incompressible fluids filling all of $\mathbb{R}^n$. The *Navier–Stokes* equations are then given by

$$\text{(1)} \qquad \frac{\partial}{\partial t}u_i + \sum_{j=1}^{n} u_j \frac{\partial u_i}{\partial x_j} = \nu \Delta u_i - \frac{\partial p}{\partial x_i} + f_i(x,t) \qquad (x \in \mathbb{R}^n, t \ge 0),$$

$$\text{(2)} \qquad \operatorname{div} u = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i} = 0 \qquad (x \in \mathbb{R}^n, t \ge 0)$$



A fundamental problem in analysis is to decide whether such smooth, physically reasonable solutions exist for the Navier–Stokes equations. To give reasonable leeway to solvers while retaining the heart of the problem, we ask for a proof of one of the following four statements.

(A) Existence and smoothness of Navier–Stokes solutions on $\mathbb{R}^3$. Take $\nu > 0$ and $n = 3$. Let $u^\circ(x)$ be any smooth, divergence-free vector field satisfying (4). Take $f(x,t)$ to be identically zero. Then there exist smooth functions $p(x,t), u_i(x,t)$ on $\mathbb{R}^3 \times [0,\infty)$ that satisfy (1), (2), (3), (6), (7).

(B) Existence and smoothness of Navier–Stokes solutions in $\mathbb{R}^3/\mathbb{Z}^3$. Take $\nu > 0$ and $n = 3$. Let $u^\circ(x)$ be any smooth, divergence-free vector field satisfying (8); we take $f(x,t)$ to be identically zero. Then there exist smooth functions $p(x,t)$, $u_i(x,t)$ on $\mathbb{R}^3 \times [0,\infty)$ that satisfy (1), (2), (3), (10), (11).

(C) Breakdown of Navier–Stokes solutions on $\mathbb{R}^3$. Take $\nu > 0$ and $n = 3$. Then there exist a smooth, divergence-free vector field $u^\circ(x)$ on $\mathbb{R}^3$ and a smooth $f(x,t)$ on $\mathbb{R}^3 \times [0,\infty)$, satisfying (4), (5), for which there exist no solutions (p,u) of (1), (2), (3), (6), (7) on $\mathbb{R}^3 \times [0,\infty)$.

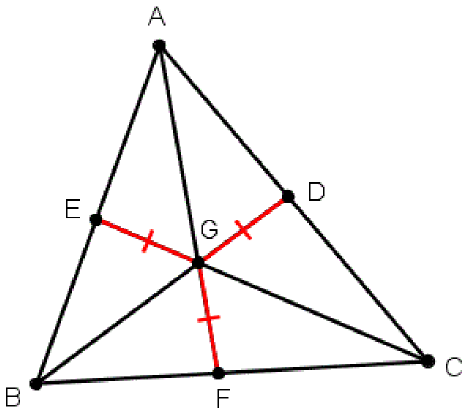
(D) Breakdown of Navier–Stokes Solutions on $\mathbb{R}^3/\mathbb{Z}^3$. Take $\nu > 0$ and $n = 3$. Then there exist a smooth, divergence-free vector field $u^\circ(x)$ on $\mathbb{R}^3$ and a smooth $f(x,t)$ on $\mathbb{R}^3 \times [0,\infty)$, satisfying (8), (9), for which there exist no solutions (p,u) of (1), (2), (3), (10), (11) on $\mathbb{R}^3 \times [0,\infty)$.

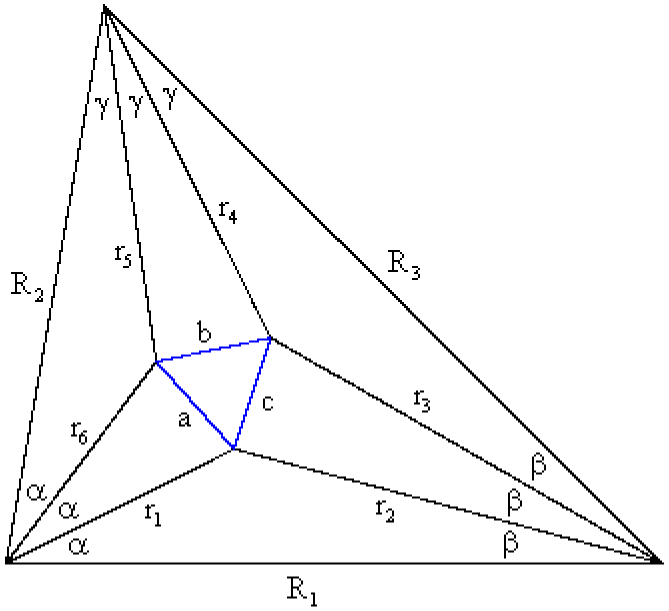
Types of Proofs

- Messy (typical of early proofs)
  - Lots of algebra; brute force; constructive
- Heavy machinery (typical of deep proofs)
  - Use lots of advanced mathematics
- Elegant
  - Elementary but clever mathematics
  - “Who would ever think of this?”

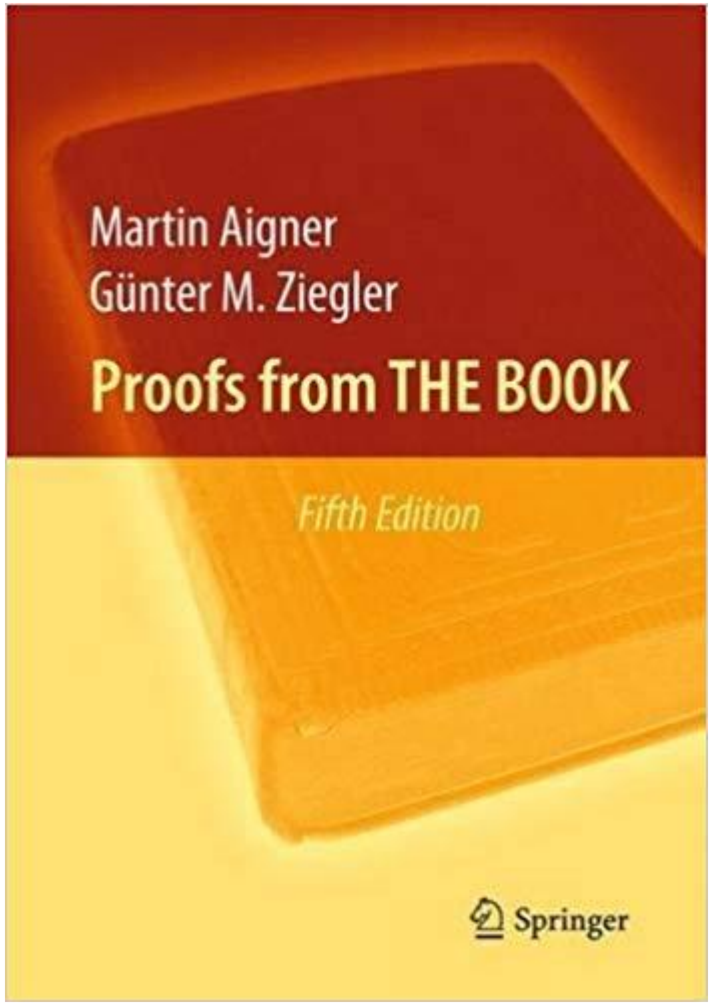
- Some results have dozens of proofs

- Mathematicians seek elegant proofs as an intellectual and aesthetic goal





Proofs as Art

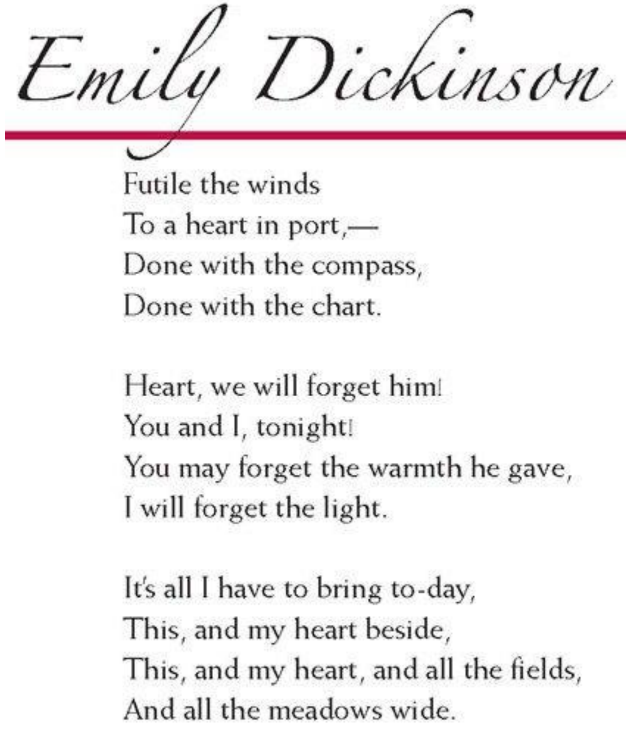


Martin Aigner
Günter M. Ziegler

Proofs from THE BOOK

Fifth Edition

Springer



Emily Dickinson

Futile the winds
To a heart in port,—
Done with the compass,
Done with the chart.

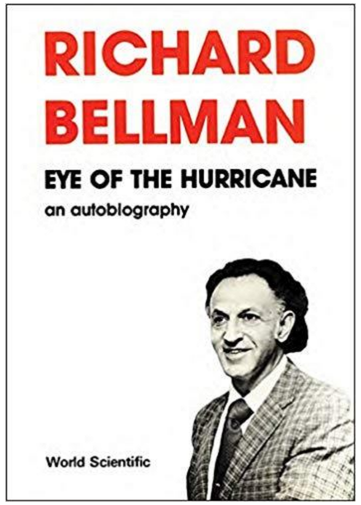
Heart, we will forget him!
You and I, tonight!
You may forget the warmth he gave,
I will forget the light.

It's all I have to bring to-day,
This, and my heart beside,
This, and my heart, and all the fields,
And all the meadows wide.

Engineering View of Proofs

- Engineering is about applications
  - We may need variations of a theorem
  - Slightly different hypotheses yield slightly different results

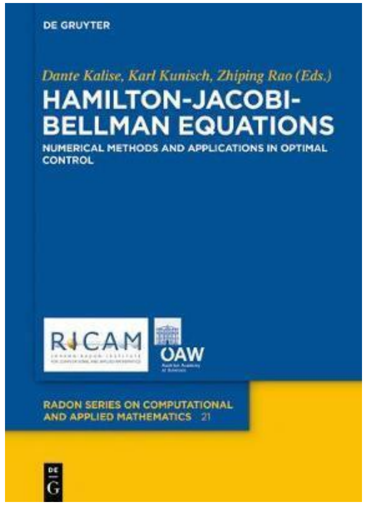
- Richard Bellman:
  - **It is desirable to prove a result in as many different ways as possible since different proofs generalize in different ways**


RICHARD BELLMAN

EYE OF THE HURRICANE

an autobiography

World Scientific


DE GRUYTER

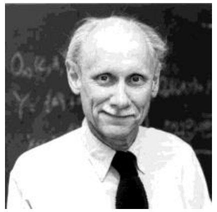
Dante Kalise, Karl Kunisch, Zhiping Rao (Eds.)

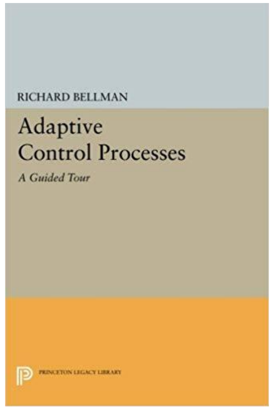
HAMILTON-JACOBI-BELLMAN EQUATIONS

NUMERICAL METHODS AND APPLICATIONS IN OPTIMAL CONTROL

RICAM ÖAW

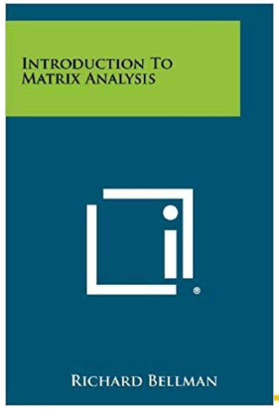
RADON SERIES ON COMPUTATIONAL AND APPLIED MATHEMATICS 21


1997

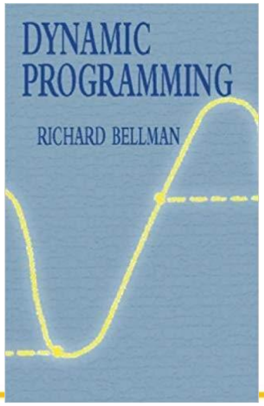

RICHARD BELLMAN

Adaptive Control Processes

A Guided Tour


INTRODUCTION TO MATRIX ANALYSIS

RICHARD BELLMAN


DYNAMIC PROGRAMMING

RICHARD BELLMAN

$$
\begin{aligned}
-\frac{\partial}{\partial t}V(x,t) &= \min_u \left[c(x,u) + \frac{\partial V}{\partial x}f(x,u)\right] \\
-x^\top \dot{P}(t)x &= \min_u \left[x^\top Qx + u^\top Ru + 2x^\top P(t)(Ax+Bu)\right] \\
0 &= \frac{\partial}{\partial u}\left[x^\top Qx + u^\top Ru + 2x^\top P(t)(Ax+Bu)\right] \\
&= 2u^\top R + 2x^\top P(t)B \\
u^* &= -R^{1}B^\top Px
\end{aligned}
$$

Proof and Non-Proof Journals

- Very few engineering disciplines write in a theorem-proof format
  - Control theory, signal processing, and information theory are exceptions
  - To publish in a “proof journal”, it is essential to know how to write good proofs

- Most other areas of science and engineering rarely write theorems and proofs
  - They could, but traditionally they do not
  - Most engineering journals are “non-proof journals”

Pros and Cons of Proofs in Engineering

- Pros of proofs:
  - Proof journals expect theorems and proofs
  - A proof tells us precise conditions under which something is true
    - These assumptions sharpen our thinking and understanding

- Cons of proofs:
  - Non-proof journals may be hostile to theorems and proofs
  - Tendency to focus on math for its own sake at the expense of what is useful
  - A theorem may not have practical value if it has
    - Unrealistic or restrictive assumptions
    - Unverifiable assumptions

What Is a “Good” Proof?

- The reader can understand the main ideas
  - The logical structure is clear
- The reader can verify the details
  - Every statement is supported
- Bonus: The reader can modify or generalize it
  - Amenable to variations

What Is a "Bad" Proof?

- Theorem statement is not clear
  - If you can't understand the theorem statement, the proof will probably not help

- Proof is hard to understand or verify
  - Obscure logical structure
  - Too much omission ("It can be shown…" or no reason given)
  - "It follows from [5]" where [5] is an 800 page book
    - Notation, terminology, and assumptions in [5] must be checked
  - Esoteric/specialized/advanced mathematics
  - Obscure references (not in English, not available)

Writing

- Technical writing
- Scientific writing
- Mathematical writing

Technical Writing

- Technical writing should be
  - Clear and friendly (understandable by humans)
  - Precise and unambiguous
  - Accurate (no errors)

- Use terminology consistently
  - Minimize variation in language
  - It’s not literature

Scientific Writing

- Scientific writing is a specialized version of technical writing
  - Objective and scholarly
  - Special vocabulary (words with accepted, unambiguous meaning)
  - Scientific writing style (objective, supported statements)

Mathematical Writing

- For definitions, theorems, proofs

- Mathematically rigorous
  - Logically rigorous
  - Unambiguous in notation and terminology
  - Clear and readable
  - Succinct

Nonsensical Mathematical Statements

- A mathematical statement can be nonsensical for many reasons
  - Ambiguous notation (Let $r(t)$ denote the position of B.)
  - Missing qualifiers (The solution of $\dot{x} = f(x)$ is $x(t) = \sin t$.)
  - Imprecise terminology (Assume that $x(t)$ does not change sign.)
  - Obscure assumptions (Theorem. $x(t)$ is bounded.)
  - Unclear logic
    - The equation has a unique, positive-semidefinite solution.
    - The equation has a unique positive-semidefinite solution.

- A sensical mathematical statement can be wrong (“2=3”)
- A nonsensical mathematical statement is “not even wrong”

Define, Let, Assume

- Define
  - Define $y \triangleq \pi^2$
  - We know what π is
  - We assign meaning to the symbol y
  - We define a symbol or word (open, closed, continuous, etc.)

- Let
  - Let x be a real number.
  - We evoke the existence of x
  - We give it the name “x”

- Assume
  - Let x be a real number, and assume that is not the square of an integer.
  - We evoke, name, and assign properties.

Definitions

- A definition must be meaningful
  - Let n be the largest integer.
  - Theorem: $n = 1$. Proof. Suppose $n \geq 2$. Then $n^2 > n$. Therefore $n = 1$.

- A definition must be useful
  - There are many ways to define the relative degree of a MIMO transfer function
  - Which one is useful? For what?

Definitions

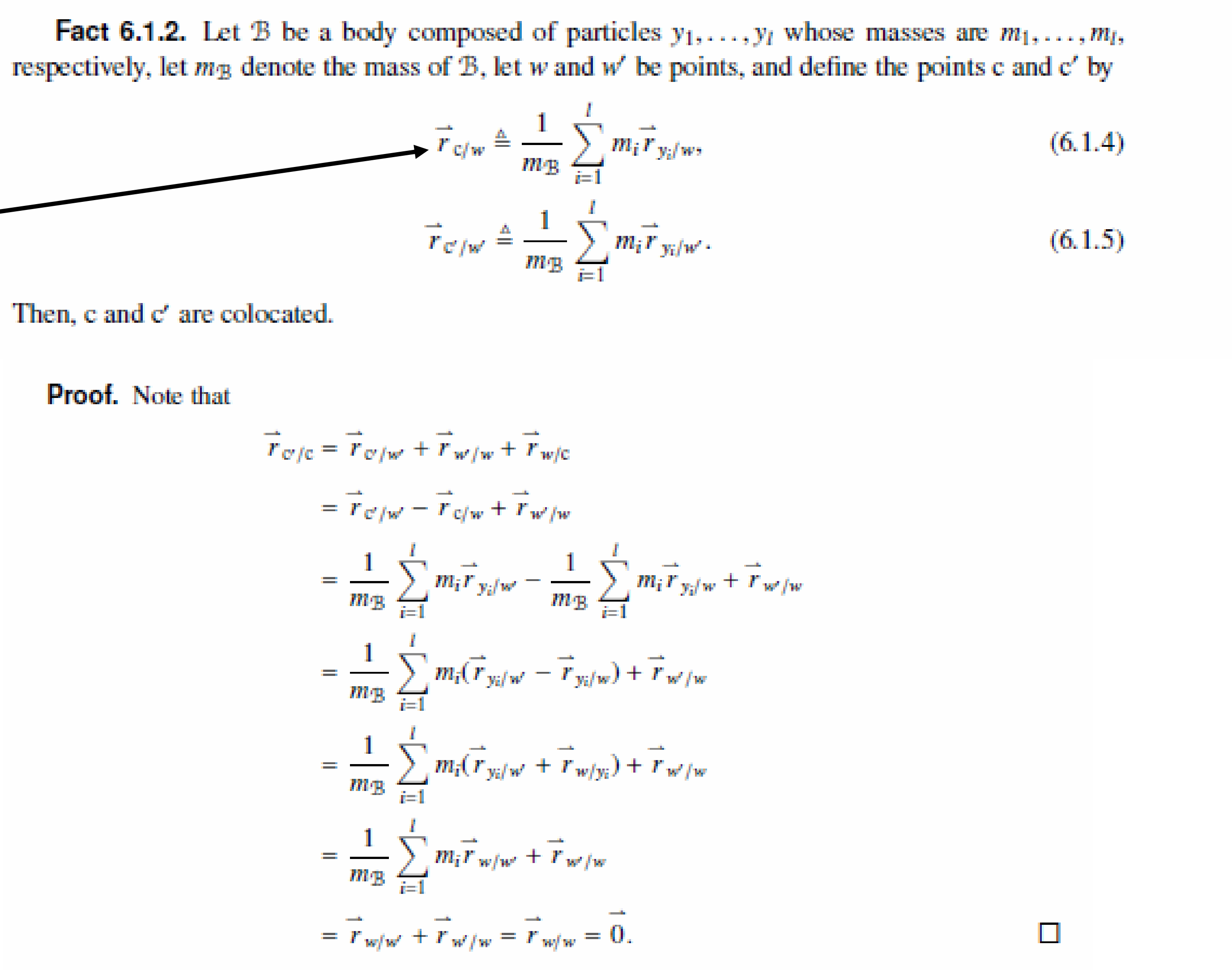
- Highlight the word or phase being defined
  - A set is *closed* if it contains its boundary.
  - The set X is *closed* if it contains its boundary.
  - The set X is *closed* if X contains its boundary.
  - Let X be a set. Then X is *closed* if X contains its boundary.

- “if” in a definition means “if and only if”----by convention
  - The integer n is *even* if it is divisible by 2.

Well-Definedness

- Definitions must be well-defined
  - This takes some care

Definition of center of mass c seems to depend on the reference point w



Fact 6.1.2. Let $\mathcal{B}$ be a body composed of particles $y_1, \ldots, y_l$ whose masses are $m_1, \ldots, m_l$, respectively, let $m_\mathcal{B}$ denote the mass of $\mathcal{B}$, let w and w' be points, and define the points c and c′ by

$$\vec{r}_{\mathrm{c}/w} \triangleq \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i \vec{r}_{y_i/w}, \tag{6.1.4}$$

$$\vec{r}_{\mathrm{c}'/w'} \triangleq \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i \vec{r}_{y_i/w'}. \tag{6.1.5}$$

Then, c and c′ are colocated.

Proof. Note that

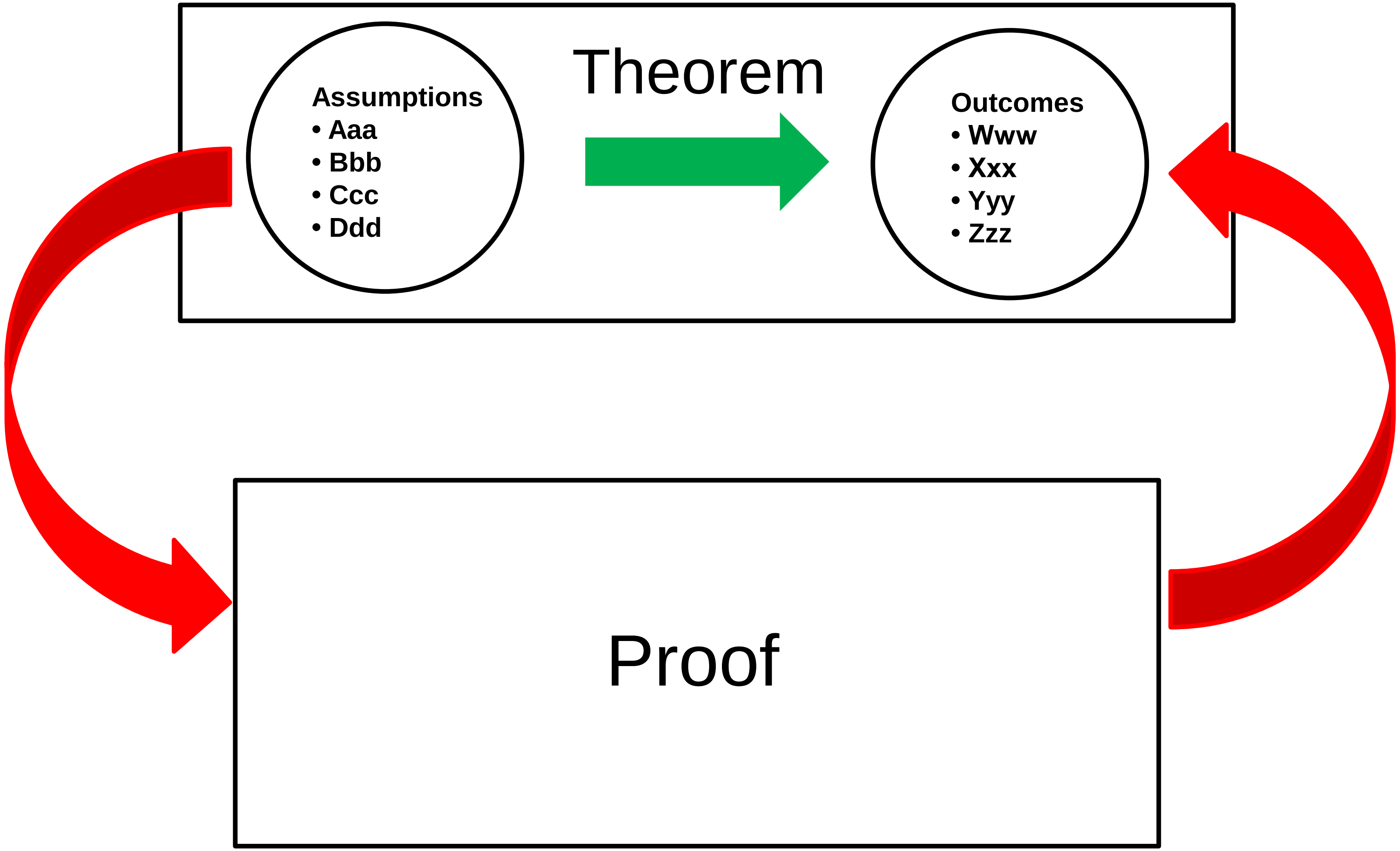
$$\begin{aligned}
\vec{r}_{\mathrm{c}'/\mathrm{c}} &= \vec{r}_{\mathrm{c}'/w'} + \vec{r}_{w'/w} + \vec{r}_{w/\mathrm{c}} \\
&= \vec{r}_{\mathrm{c}'/w'} - \vec{r}_{\mathrm{c}/w} + \vec{r}_{w'/w} \\
&= \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i \vec{r}_{y_i/w'} - \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i \vec{r}_{y_i/w} + \vec{r}_{w'/w} \\
&= \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i (\vec{r}_{y_i/w'} - \vec{r}_{y_i/w}) + \vec{r}_{w'/w} \\
&= \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i (\vec{r}_{y_i/w'} + \vec{r}_{w/y_i}) + \vec{r}_{w'/w} \\
&= \frac{1}{m_\mathcal{B}} \sum_{i=1}^{l} m_i \vec{r}_{w/w'} + \vec{r}_{w'/w} \\
&= \vec{r}_{w/w'} + \vec{r}_{w'/w} = \vec{r}_{w/w} = \vec{0}.
\end{aligned}$$

□

Theorems et al

- Theorem
  - Major result
- Proposition
  - Minor result but still important
- Lemma
  - Technical result mainly of interest to prove a theorem or proposition
- Corollary
  - Consequence of a theorem or proposition
-

Theorem and Proof



Theorem

Assumptions
- Aaa
- Bbb
- Ccc
- Ddd

Outcomes
- Www
- Xxx
- Yyy
- Zzz

Proof

What Is a Theorem?

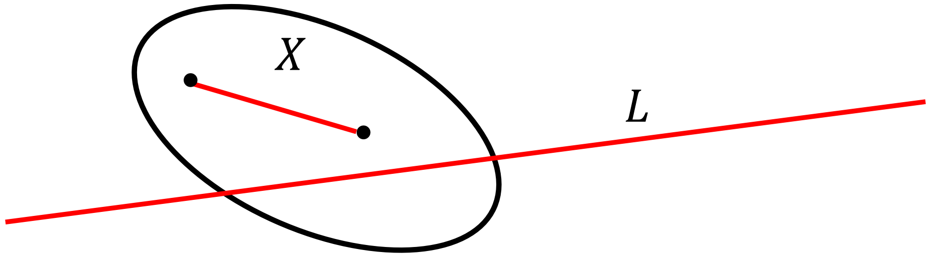
- Movie analogy
  - Introduce actors and their backstories
    - Let x and y be a positive numbers, and define $z \triangleq \log xy$
  - Plot unfolds (action)
    - Then $z = \log x + \log y$

- Setting up the action (with a little suspense)
  - Theorem. Let X, Y be open sets, let $f: X \to Y$ be continuous, and let $Y_0 \subset Y$ be an open set. Then $f^{-1}(Y_0)$ is open.

  - Theorem. Let X, Y be open sets, let $f: X \to Y$, <u>assume</u> that f is continuous, and let $Y_0 \subset Y$ be an open set. Then $f^{-1}(Y_0)$ is open.

  - Theorem. Let X, Y be open sets, let $f: X \to Y$, and let $Y_0 \subset Y$ be an open set. <u>If</u> f is continuous, then $f^{-1}(Y_0)$ is open.

Interchanging Definitions and Theorems



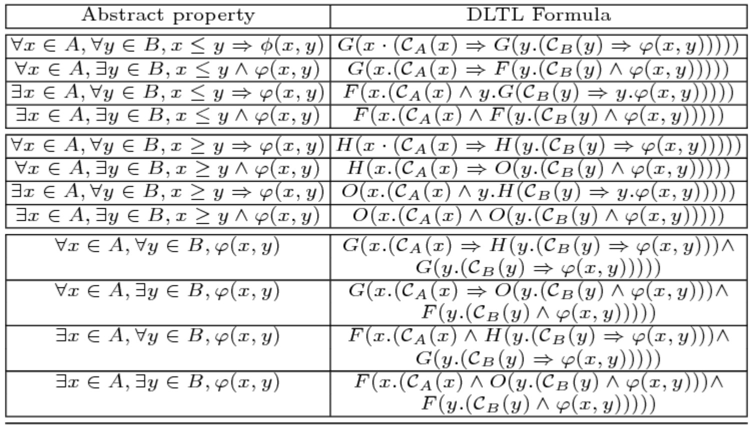
- The definition creates and classifies objects
- Theorems characterize properties of those objects

- Approach 1:
  - Definition: The set X is *convex* if, for all pairs of points x, y in X, the line segment connecting x and y is contained in X.
  - Theorem: The set X is convex if and only if, for every line L, the set $L \cap X$ is either empty or a line segment.

- Approach 2:
  - Definition: The set X is *convex* if, for every line L, the set $L \cap X$ is either empty or a line segment.
  - Theorem: The set X is convex if and only if, for all pairs of points x, y in X, the line segment connecting x and y is contained in X.

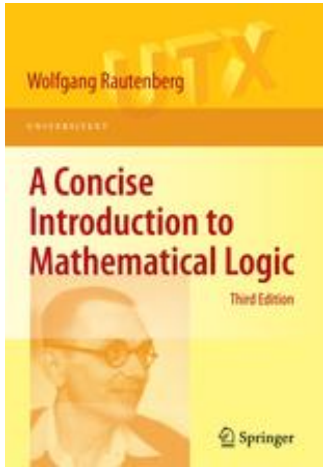
- Definitions and “iff” theorems are interchangeable!
  - Choice is arbitrary but based on tradition
  - Need “iff” theorems to do this

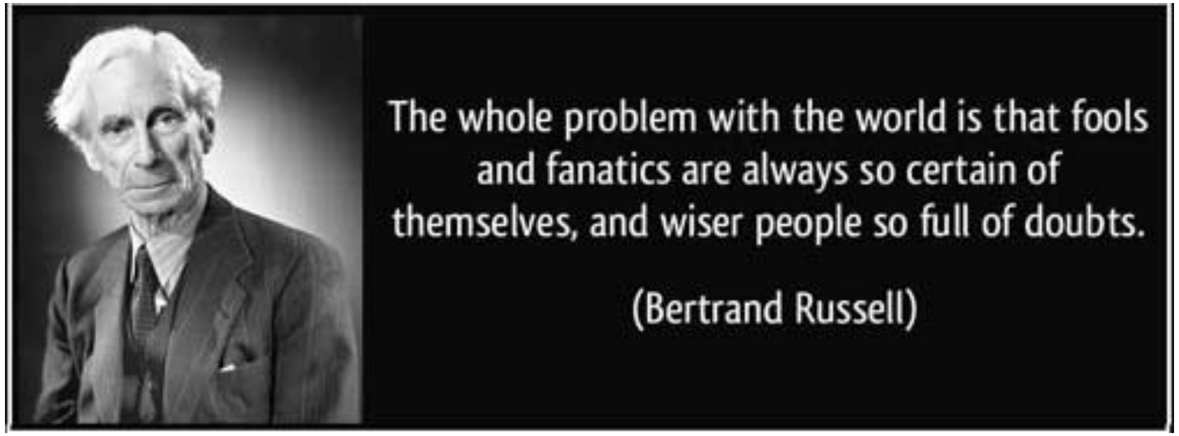
Logic



Abstract property	DLTL Formula
$\forall x \in A, \forall y \in B, x \leq y \Rightarrow \phi(x,y)$	$G(x \cdot (\mathcal{C}_A(x) \Rightarrow G(y.(\mathcal{C}_B(y) \Rightarrow \varphi(x,y)))))$
$\forall x \in A, \exists y \in B, x \leq y \wedge \varphi(x,y)$	$G(x.(\mathcal{C}_A(x) \Rightarrow F(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))))$
$\exists x \in A, \forall y \in B, x \leq y \Rightarrow \varphi(x,y)$	$F(x.(\mathcal{C}_A(x) \wedge y.G(\mathcal{C}_B(y) \Rightarrow y.\varphi(x,y)))))$
$\exists x \in A, \exists y \in B, x \leq y \wedge \varphi(x,y)$	$F(x.(\mathcal{C}_A(x) \wedge F(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))))$
$\forall x \in A, \forall y \in B, x \geq y \Rightarrow \varphi(x,y)$	$H(x \cdot (\mathcal{C}_A(x) \Rightarrow H(y.(\mathcal{C}_B(y) \Rightarrow \varphi(x,y)))))$
$\forall x \in A, \exists y \in B, x \geq y \wedge \varphi(x,y)$	$H(x.(\mathcal{C}_A(x) \Rightarrow O(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))))$
$\exists x \in A, \forall y \in B, x \geq y \Rightarrow \varphi(x,y)$	$O(x.(\mathcal{C}_A(x) \wedge y.H(\mathcal{C}_B(y) \Rightarrow y.\varphi(x,y)))))$
$\exists x \in A, \exists y \in B, x \geq y \wedge \varphi(x,y)$	$O(x.(\mathcal{C}_A(x) \wedge O(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))))$
$\forall x \in A, \forall y \in B, \varphi(x,y)$	$G(x.(\mathcal{C}_A(x) \Rightarrow H(y.(\mathcal{C}_B(y) \Rightarrow \varphi(x,y)))\wedge$ $G(y.(\mathcal{C}_B(y) \Rightarrow \varphi(x,y)))))$
$\forall x \in A, \exists y \in B, \varphi(x,y)$	$G(x.(\mathcal{C}_A(x) \Rightarrow O(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))\wedge$ $F(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))))$
$\exists x \in A, \forall y \in B, \varphi(x,y)$	$F(x.(\mathcal{C}_A(x) \wedge H(y.(\mathcal{C}_B(y) \Rightarrow \varphi(x,y)))\wedge$ $G(y.(\mathcal{C}_B(y) \Rightarrow \varphi(x,y)))))$
$\exists x \in A, \exists y \in B, \varphi(x,y)$	$F(x.(\mathcal{C}_A(x) \wedge O(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))\wedge$ $F(y.(\mathcal{C}_B(y) \wedge \varphi(x,y)))))$

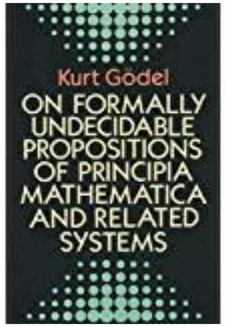
Table 1. DLTL formula schemes correlating two word positions



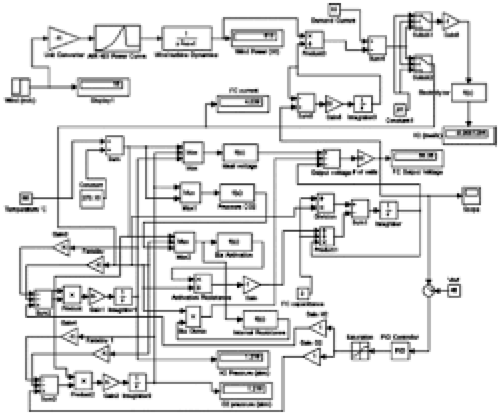


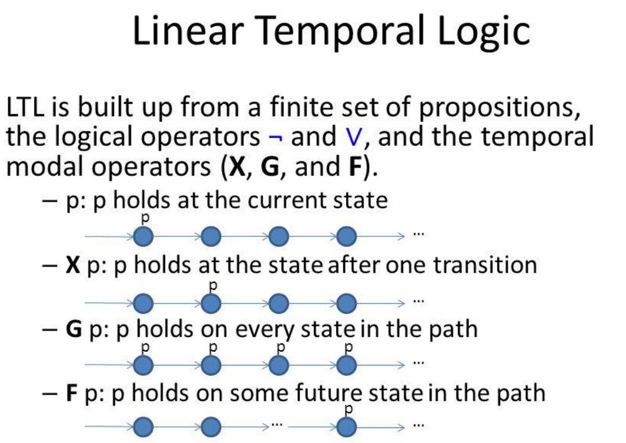
The whole problem with the world is that fools and fanatics are always so certain of themselves, and wiser people so full of doubts.

(Bertrand Russell)



Kurt Gödel
ON FORMALLY UNDECIDABLE PROPOSITIONS OF PRINCIPIA MATHEMATICA AND RELATED SYSTEMS

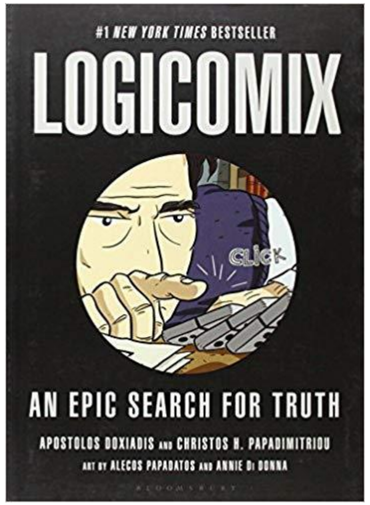




Linear Temporal Logic

LTL is built up from a finite set of propositions, the logical operators ¬ and V, and the temporal modal operators (**X**, **G**, and **F**).

- p: p holds at the current state
- **X** p: p holds at the state after one transition
- **G** p: p holds on every state in the path
- **F** p: p holds on some future state in the path



#1 NEW YORK TIMES BESTSELLER
LOGICOMIX
CLICK
AN EPIC SEARCH FOR TRUTH
APOSTOLOS DOXIADIS AND CHRISTOS H. PAPADIMITRIOU
ART BY ALECOS PAPADATOS AND ANNIE DI DONNA
BLOOMSBURY

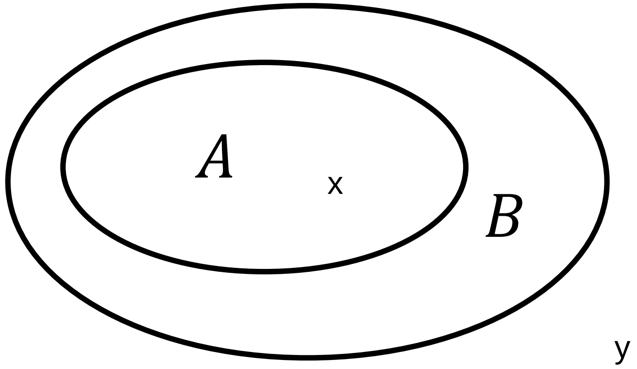
© 2002 by John Holden.

Logic Words

- Essential short words
  - If…then
  - And
  - Or
  - Not
  - Exists
  - All

Logic: If…, then…

- If A, then B
  - A implies B
  - $A \Longrightarrow B$
  - Venn diagram: $A \subseteq B$



- Contrapositive [equivalent]
  - If not B, then not A
  - not $B \Longrightarrow$ not A
  - Venn diagram: $B^{\sim} \subseteq A^{\sim}$

What Is a Corollary?

Theorem

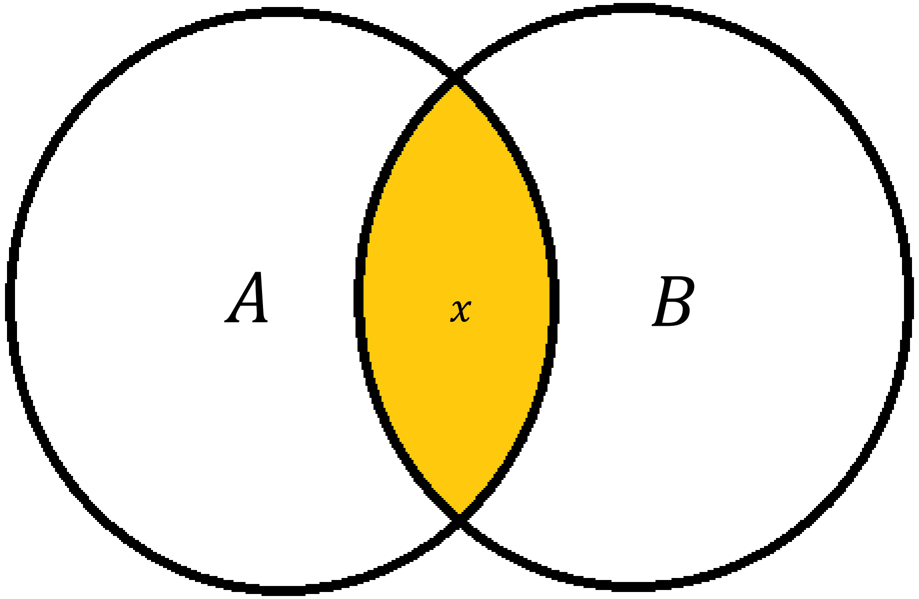
$$\begin{tikzcd}
A \arrow[r, Rightarrow] & B \arrow[d, Rightarrow] \\
C \arrow[u, Rightarrow] \arrow[r, Rightarrow] & D
\end{tikzcd}$$

Corollary

$\Rightarrow$ is transitive

And

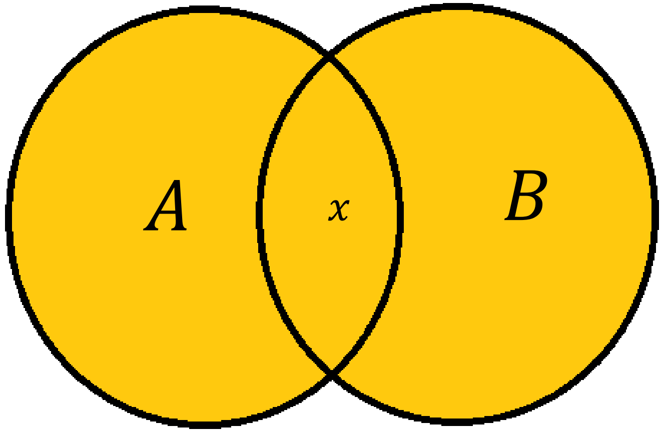
$x \in A$ and $x \in B$



$x \in A \cap B$

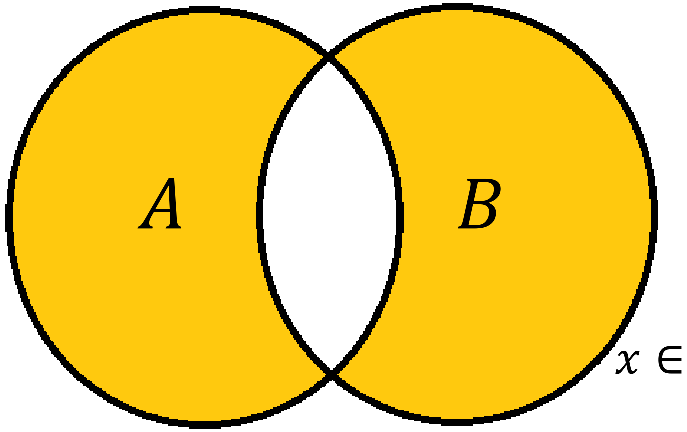
Or

- In mathematics
  - A “or” B is inclusive
    - x may be an element of $A \cap B$
  - “xor” means “exclusive or”
    - x cannot be an element of $A \cap B$
- In English
  - “or” is exclusive
  - Choose fries or a salad.
    - Obviously not both
  - Choose fries and/or a salad.
    - Possibly both
  - The model is valid if the set is closed and/or open.



$x \in A \cup B$

English “and/or” = mathematics “or”



$x \in A \cup B \sim (A \cap B)$

English “or” = mathematics “xor”

Parentheses versus Ambiguity

- Let X be bounded and open or convex
  - ???

- Let X be (bounded and open) or convex
- Let X be bounded and (open or convex)

- Let X be both bounded and either open or convex
- Let X be bounded and either open or convex

Not

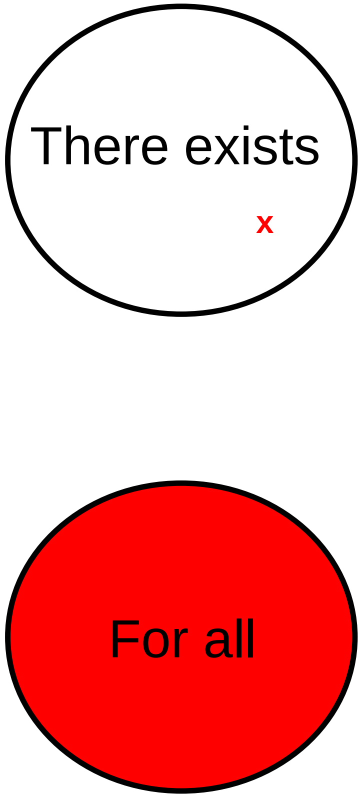
- The not of this:
  - For all $\varepsilon > 0$, there exists a $\delta > 0$ such that A is true.

- Is this:
  - Not(for all $\varepsilon > 0$, there exists a $\delta > 0$ such that A is true.)

- Which is actually this:
  - There exists $\varepsilon > 0$ such that, for all $\delta > 0$, A is not true.

Qualifiers: Exist, All

- There exists
  - “There exists at least one”
  - Existence + uniqueness: There exists exactly one
- For all
  - For all $\varepsilon > 0$, there exists $\delta > 0$...
  - For every $\varepsilon > 0$, there exists $\delta > 0$...
  - For each $\varepsilon > 0$, there exists $\delta > 0$...
  - Let $\varepsilon > 0$. Then there exists $\delta > 0$...
- Every variable in every mathematical statement must be explicitly and unambiguously qualified
  - No matter what



More on Qualifiers

- Order of qualifiers
  - For all $\varepsilon > 0$, there exists $\delta > 0$...
    - δ may depend on ε
  - There exists $\delta > 0$ such that, for all $\varepsilon > 0$,...
    - δ does not depend on ε

- The dependence of objects can be correctly inferred from the logic
  - No need to say “(not depending on ε)
  - Can mention for emphasis (“Note that…”)

Words that Are OK to Use in a Defn, Theorem, Proof

- Setup words: let, define, assume*
- Logic words: if*, then, and, or, not
  - if and assume are not ok in proofs
    - Say “suppose” in a proof by contradiction
  - Assumptions are made in the theorem, not in the proof
  - Say “in the case where”
- In the case where
- Since…it follows that
  - Reason before result
- Yields, implies
- However, moreover, in addition, furthermore, but, although, whereas
- Whose, with

- The allowable vocabulary used in a proof is severely limited to a small set of words that have clear and precise meaning

Words to Never Use in a Defn, Thm, Proof

- $\forall, \exists$
- When, whenever, always
  - Time is irrelevant
- Unless, even though, provided, necessarily, arbitrarily, sufficiently small, sufficiently large
  - OK for discussion but could be unnecessary, imprecise, or logically confusing
- Would, could, should
  - It should rain today
  - Dennis should get a Nobel prize
- Some (for some = there exists)
- Fixed, Given

Words that Clarify Logic

- Condition A is satisfied whether or not Condition B is satisfied
- If Condition A is satisfied, then it does not necessarily follow that Condition B is satisfied
- Can, may, might
  - Imprecise, but OK for discussion
  - Can: strong cause and effect; may is weak causality (takes place of might)

Theorem Statements

- If n is an even integer, then $n + 1$ is an odd integer

- The following statements are equivalent:
  - □ $i) \ \ |x| < 1$
  - □ $ii) \ \ x^2 < 1$
  - □ $iii) \ x^4 < 1$

- Exactly one of the following statements holds:
  - □ $i) \quad x < 1$
  - □ $ii) \quad x = 1$
  - □ $iii) \ \ x > 1$

Theorem Statements

- If at least two of the following statements holds, then so does the remaining statement:
  - $i)\quad x \leq 1$
  - $ii)\quad x = 1$
  - $iii)\ x \geq 1$

- Consider the following statements:
  - $i)\quad x < 1$
  - $ii)\quad x < 2$
  - $iii)\ x < 3$
  - Then $i) \Longrightarrow ii) \Longrightarrow iii)$

Proof Logic: 4 Types

- 1. Direct
  - $A \Longrightarrow B$

- 2. Contrapositive
  - not $B \Longrightarrow$ not A

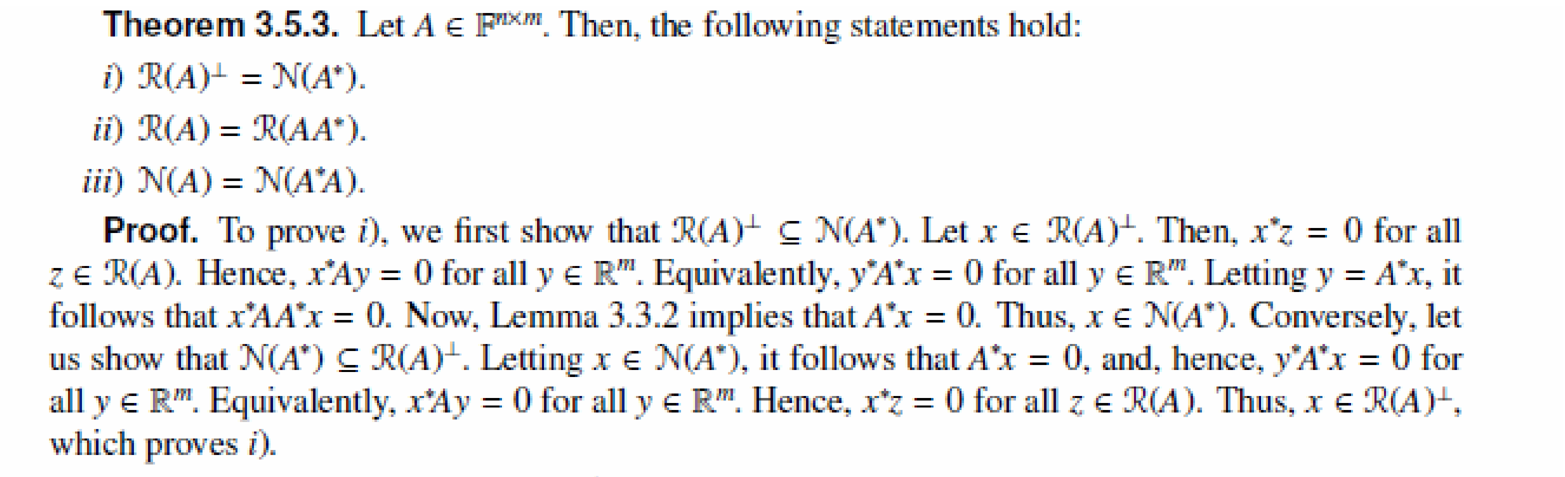
- 3. Contradiction
  - A and (not B) leads to a contradiction

- 4. Induction
  - Show $\sum_{i=1}^{n} i^3 = \frac{1}{4}n^2(n+1)^2$
  - Prove for $n = 1$. Assume true for $n \geq 2$.   Prove for $n + 1$.

Make the strategy clear to the reader

Avoid proving both directions of “iff” at the same time

Example: Direct Proof



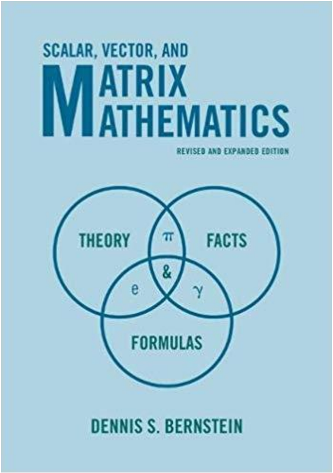
Theorem 3.5.3. Let $A \in \mathbb{F}^{n\times m}$. Then, the following statements hold:

i) $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^*)$.

ii) $\mathcal{R}(A) = \mathcal{R}(AA^*)$.

iii) $\mathcal{N}(A) = \mathcal{N}(A^*A)$.

Proof. To prove *i)*, we first show that $\mathcal{R}(A)^{\perp} \subseteq \mathcal{N}(A^*)$. Let $x \in \mathcal{R}(A)^{\perp}$. Then, $x^*z = 0$ for all $z \in \mathcal{R}(A)$. Hence, $x^*Ay = 0$ for all $y \in \mathbb{R}^m$. Equivalently, $y^*A^*x = 0$ for all $y \in \mathbb{R}^m$. Letting $y = A^*x$, it follows that $x^*AA^*x = 0$. Now, Lemma 3.3.2 implies that $A^*x = 0$. Thus, $x \in \mathcal{N}(A^*)$. Conversely, let us show that $\mathcal{N}(A^*) \subseteq \mathcal{R}(A)^{\perp}$. Letting $x \in \mathcal{N}(A^*)$, it follows that $A^*x = 0$, and, hence, $y^*A^*x = 0$ for all $y \in \mathbb{R}^m$. Equivalently, $x^*Ay = 0$ for all $y \in \mathbb{R}^m$. Hence, $x^*z = 0$ for all $z \in \mathcal{R}(A)$. Thus, $x \in \mathcal{R}(A)^{\perp}$, which proves *i)*.

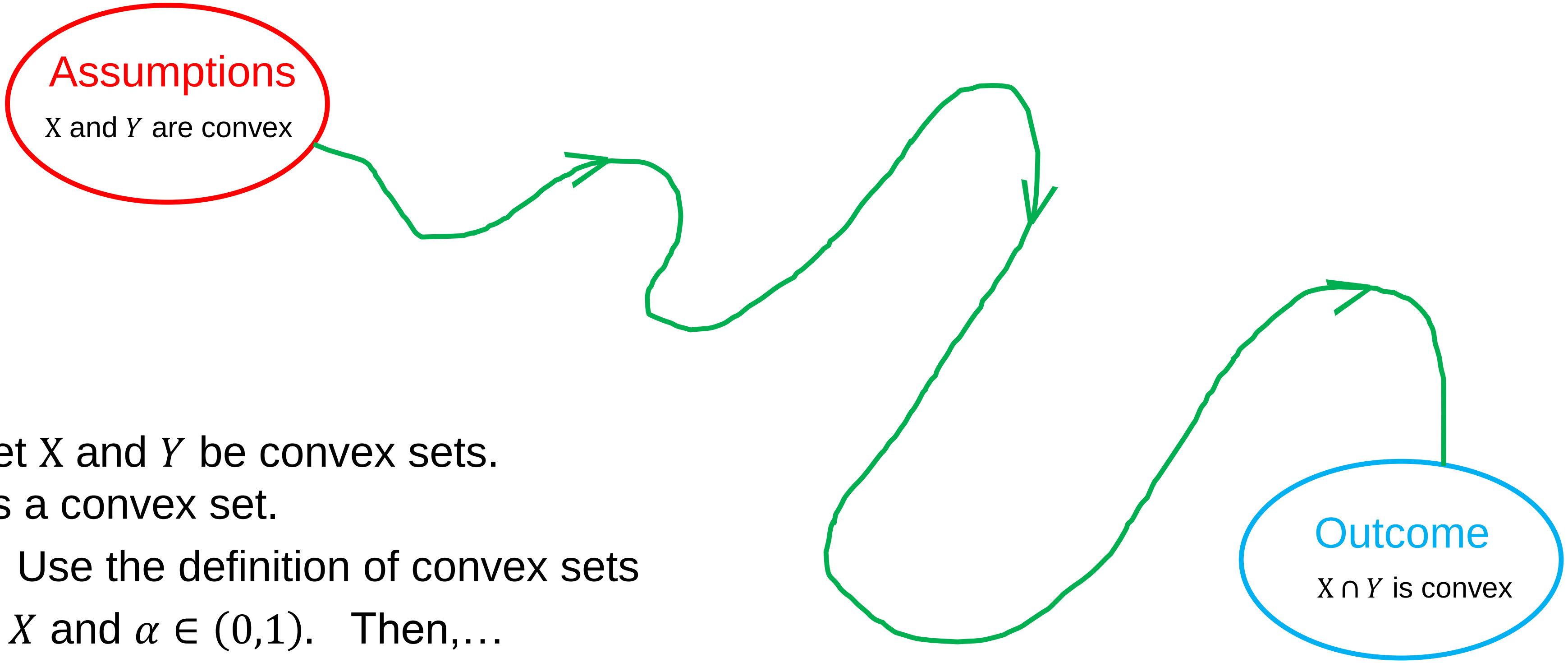


SCALAR, VECTOR, AND
MATRIX MATHEMATICS
REVISED AND EXPANDED EDITION

DENNIS S. BERNSTEIN

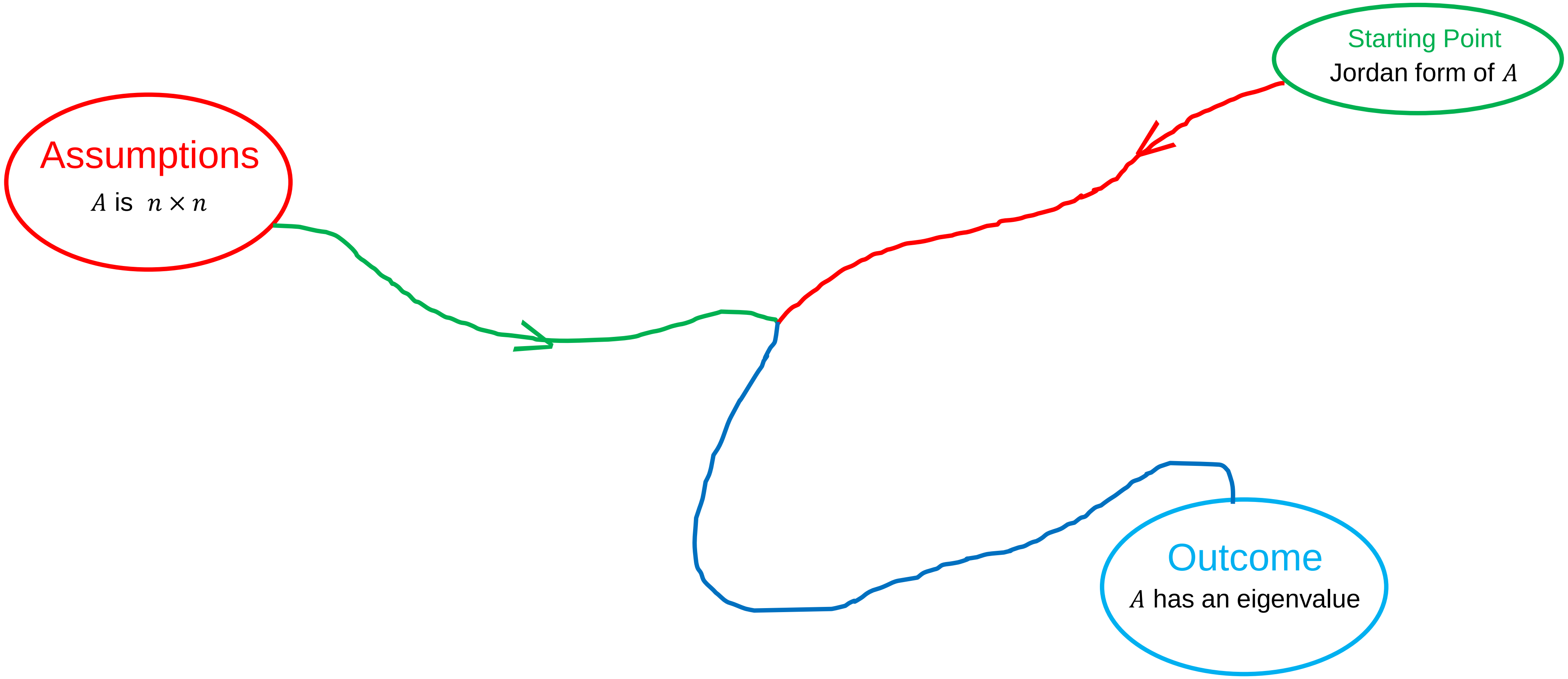
Starting a Proof: From the Assumptions

- A proof is a path from the assumptions to the result
  - It requires a strategy and structure



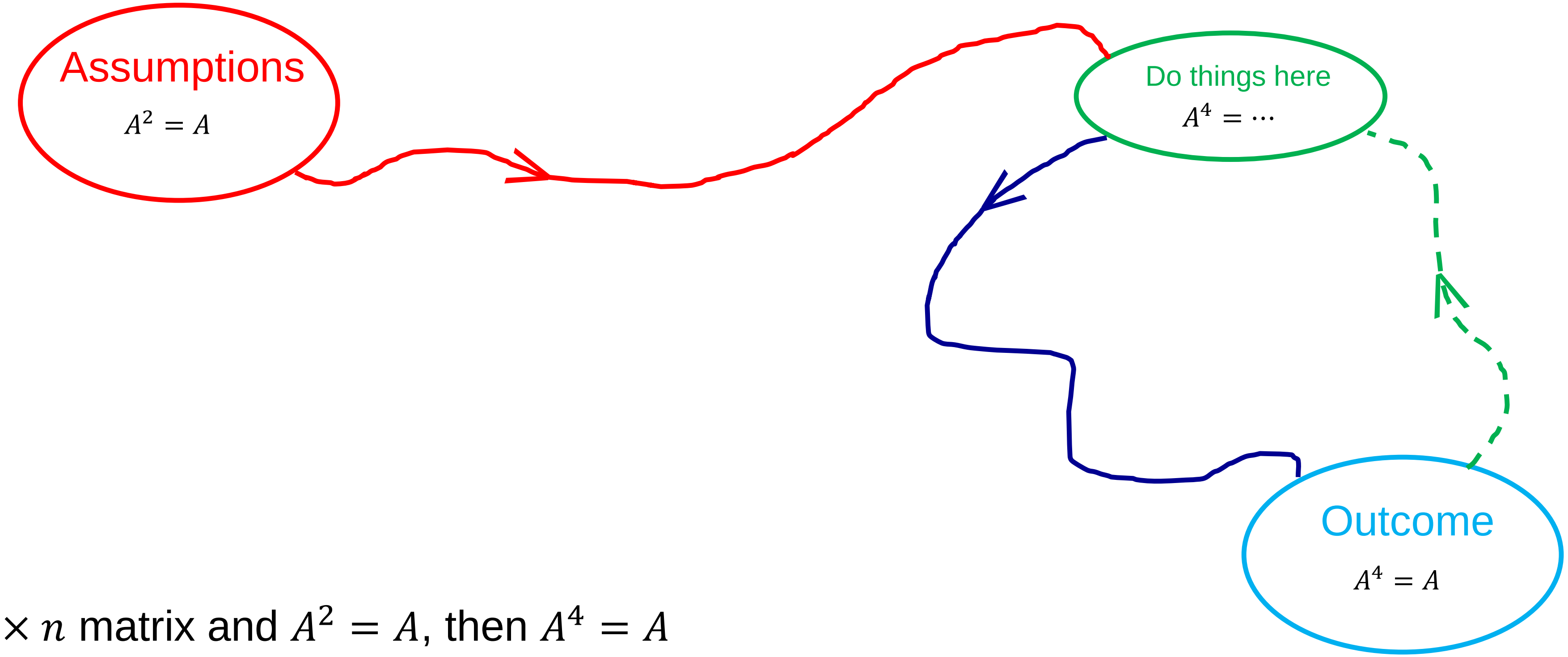
- Theorem: Let X and Y be convex sets.
  Then $X \cap Y$ is a convex set.
  - Strategy: Use the definition of convex sets
  - Let $x, y \in X$ and $\alpha \in (0,1)$.   Then,…

Starting a Proof: From Machinery



- If A is an $n \times n$ matrix, then A has at least one eigenvalue
  - Strategy: Construct nonsingular S such that $SAS^{-1}$is in Jordan form

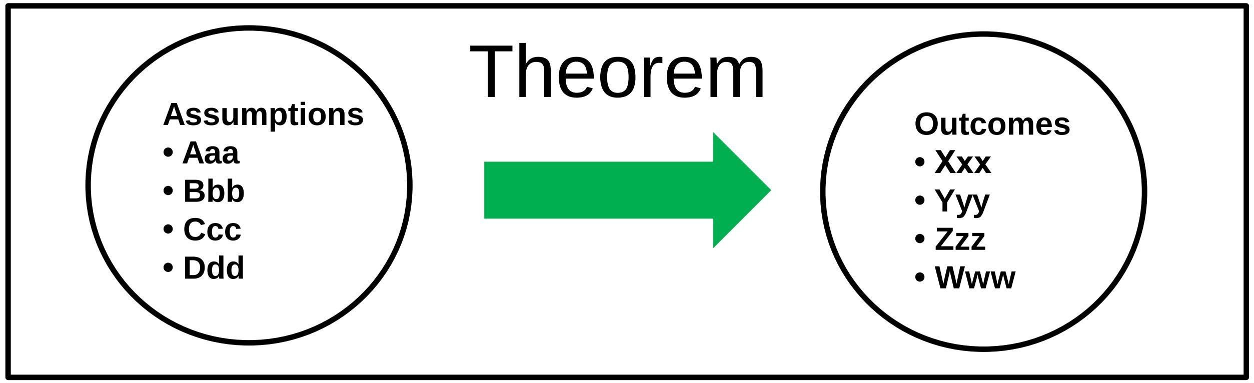
Starting a Proof: From the Outcome



If A is an $n \times n$ matrix and $A^2 = A$, then $A^4 = A$

Mechanics: Checklist

- All assumptions of a theorem must be used in the proof
- Help the reader by “checking off” each assumption
  - Since X is closed it follows that… Furthemore, since X is bounded, it follows that…
- Tell the reader where each assumption is used in the proof



Theorem

Assumptions
- Aaa
- Bbb
- Ccc
- Ddd

Outcomes
- Xxx
- Yyy
- Zzz
- Www

Mechanics: A Derivation Is Not a Proof!

- Theorem: $x(t) = \sin t^2$ is the solution to the ODE $\dot{x} = x^4 + \cos x^3$
  - Proof. Substitute and check
  - You have no obligation to explain how you found the solution

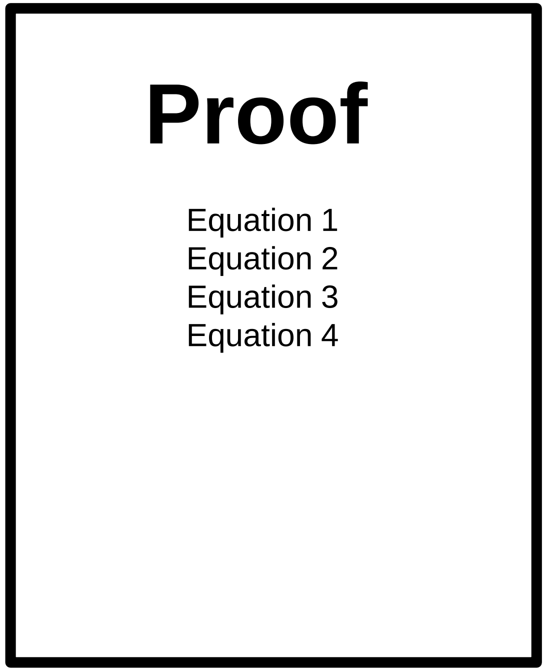
- Proof by “derivation” is logically suspect
  - A derivation begins by assuming existence
  - But existence cannot be assumed a priori---not justified

- The logical direction of a derivation might not be reversible
  - $2x < 8$. Therefore, $2x < 11$. Choose $x = 5$. Wrong.

Mechanics: Proof in a Box

- Think of a proof as a sealed box
  - A proof is a sequence of statements that form an argument
  - The status of each statement has meaning only within the context of the proof
  - Therefore, no statement (equation) should be accessed from outside the box

- Labeling equations
  - In a paper, it helps to label every equation whether or not it is referred to
    - Readers and reviewers may wish to refer to it
  - In a proof, ONLY equations that are referred to should be labeled
  - If an equation inside a proof is useful, then “bubble” it up to the statement of the theorem

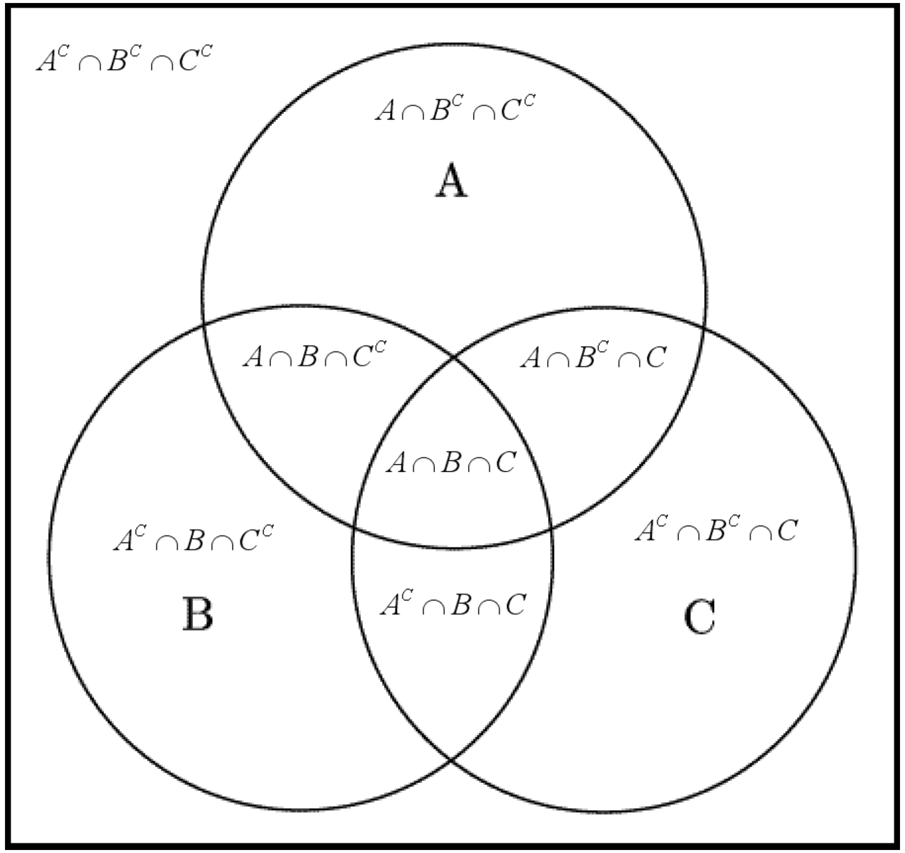


Proof

Equation 1
Equation 2
Equation 3
Equation 4

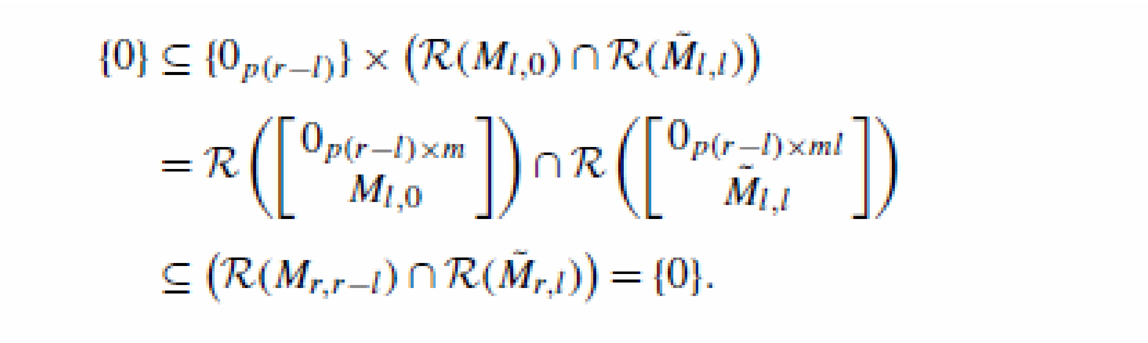
Techniques: Sets

- Sets are the language of mathematics
  - Define and use sets to describe objects and their relationship
  - Venn diagrams can clarify logic



Math Technique: Squeezing

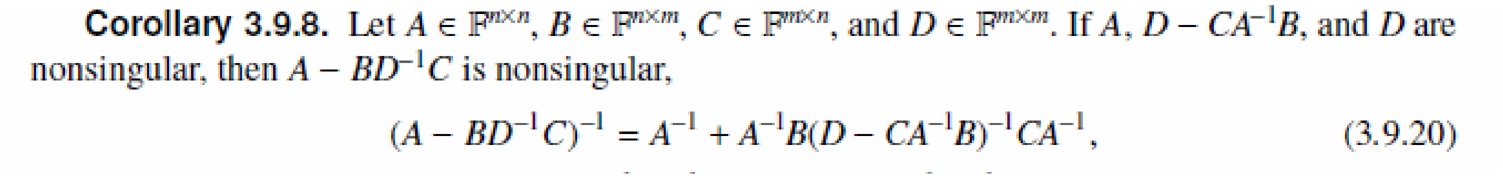
- Let A and B be sets. Prove that $A = B$.
  - Technique: Show that 1) $A \subseteq B$ and 2) $B \subseteq A$
  - $A \subseteq B \subseteq A \Leftrightarrow A = B$ Why does this work?
  - $x \leq y \leq x$ implies $x = y$
  - "$\subseteq$" and "$\leq$" are antisymmetric
  - All partial orderings are antisymmetric



$$\begin{aligned}
\{0\} &\subseteq \{0_{p(r-l)}\} \times \left(\mathcal{R}(M_{l,0}) \cap \mathcal{R}(\tilde{M}_{l,l})\right) \\
&= \mathcal{R}\left(\begin{bmatrix} 0_{p(r-l)\times m} \\ M_{l,0} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} 0_{p(r-l)\times ml} \\ \tilde{M}_{l,l} \end{bmatrix}\right) \\
&\subseteq \left(\mathcal{R}(M_{r,r-l}) \cap \mathcal{R}(\tilde{M}_{r,l})\right) = \{0\}.
\end{aligned}$$

Math Technique: Augmentation

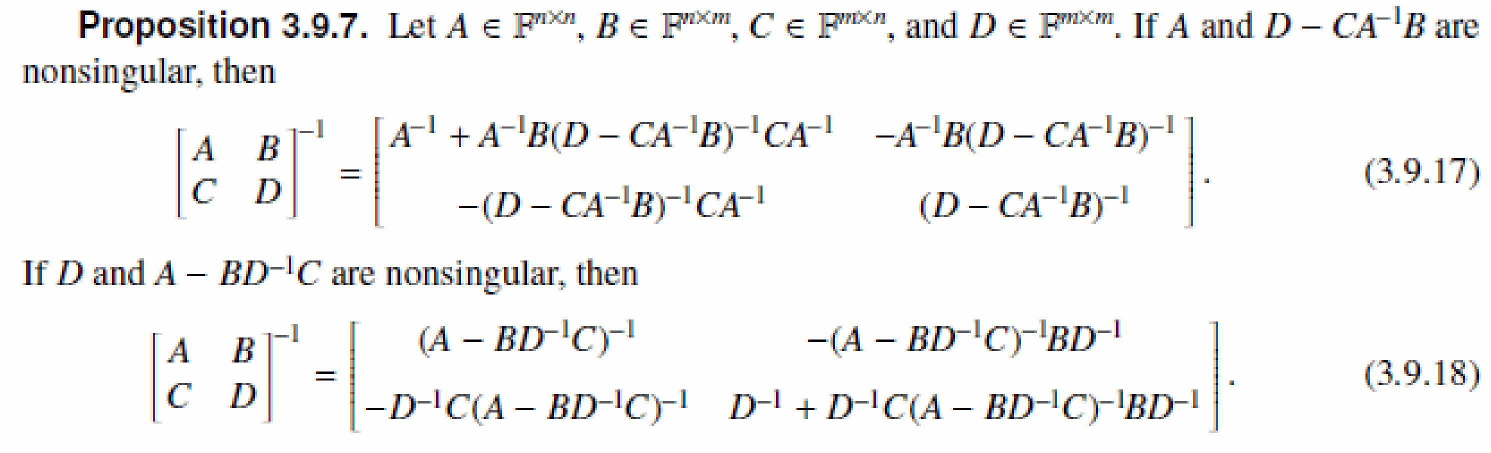
- Matrix inversion lemma



Corollary 3.9.8. Let $A \in \mathbb{F}^{n\times n}$, $B \in \mathbb{F}^{n\times m}$, $C \in \mathbb{F}^{m\times n}$, and $D \in \mathbb{F}^{m\times m}$. If A, $D - CA^{-1}B$, and D are nonsingular, then $A - BD^{-1}C$ is nonsingular,

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}, \tag{3.9.20}$$

- Proof



Proposition 3.9.7. Let $A \in \mathbb{F}^{n\times n}$, $B \in \mathbb{F}^{n\times m}$, $C \in \mathbb{F}^{m\times n}$, and $D \in \mathbb{F}^{m\times m}$. If A and $D - CA^{-1}B$ are nonsingular, then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}. \tag{3.9.17}$$

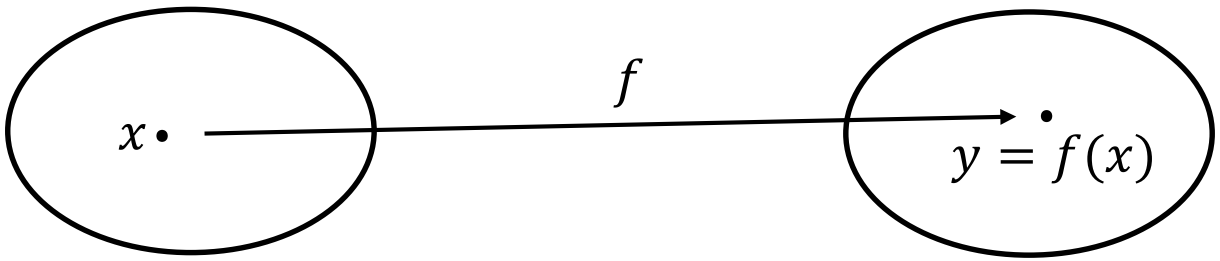
If D and $A - BD^{-1}C$ are nonsingular, then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}. \tag{3.9.18}$$

-

Math Technique: Inverses

- Consider the function $f: X \rightarrow Y$
- Properties of f
  - f is onto (surjective): For all y there exists x such that $y = f(x)$
  - f is one-to-one (injective): If $f(x_1) = f(x_2)$, then $x_1 = x_2$



x

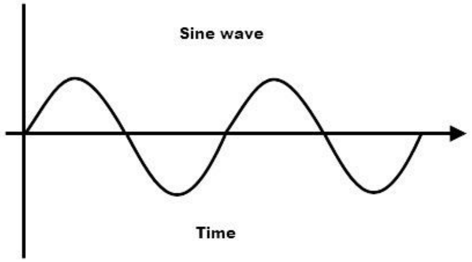
f

$y = f(x)$

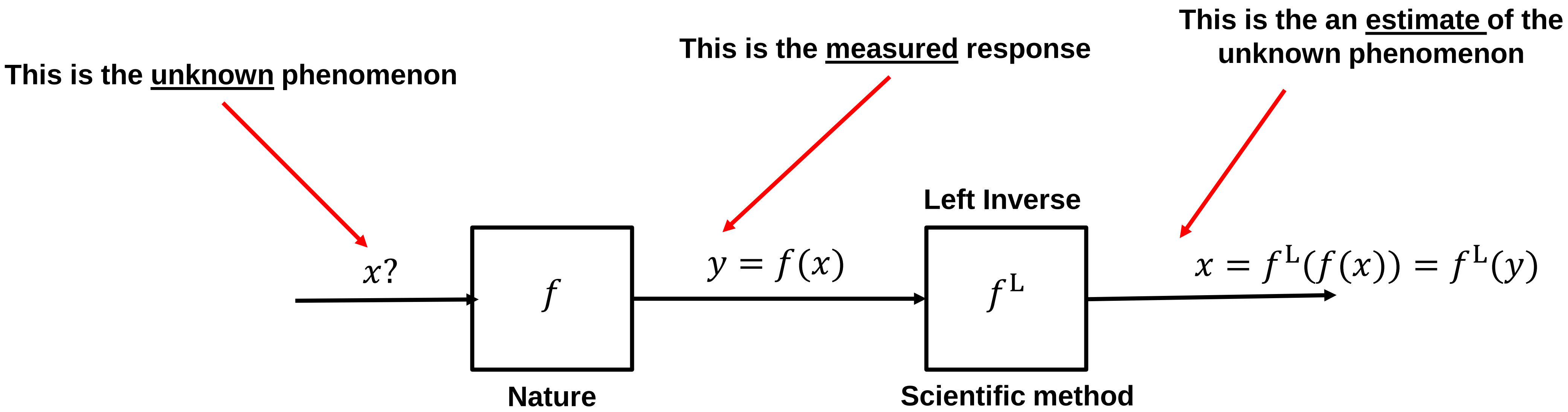
Domain X

Codomain Y

- f is invertible if it is one-to-one and onto
- If f is not onto, then we can create a “new” f by restricting its codomain
- If f is not one-to-one, then we can create a “new” f by restricting its domain
- Example: sin, cos, and tan are neither one-to-one nor onto
  - We artificially create arcsin, arcos, arctan by restricting their domain and codomain



Science as a Left Inverse Problem



This is the <u>unknown</u> phenomenon

This is the <u>measured</u> response

This is the an <u>estimate</u> of the unknown phenomenon

$x?$

f

Nature

$y = f(x)$

Left Inverse

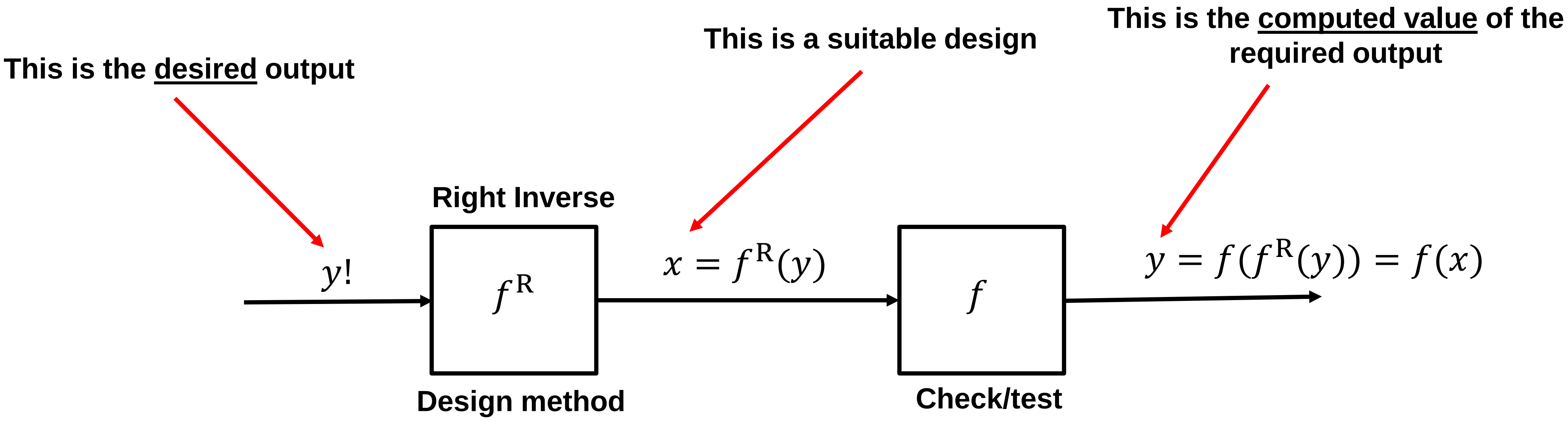
f^{L}

Scientific method

$x = f^{\mathrm{L}}(f(x)) = f^{\mathrm{L}}(y)$

This is a science problem.
Want f one-to-one.

Engineering as a Right Inverse Problem



This is the <u>desired</u> output

This is a suitable design

This is the <u>computed value</u> of the required output

Right Inverse

$y!$

f^{R}

$x = f^{\mathrm{R}}(y)$

f

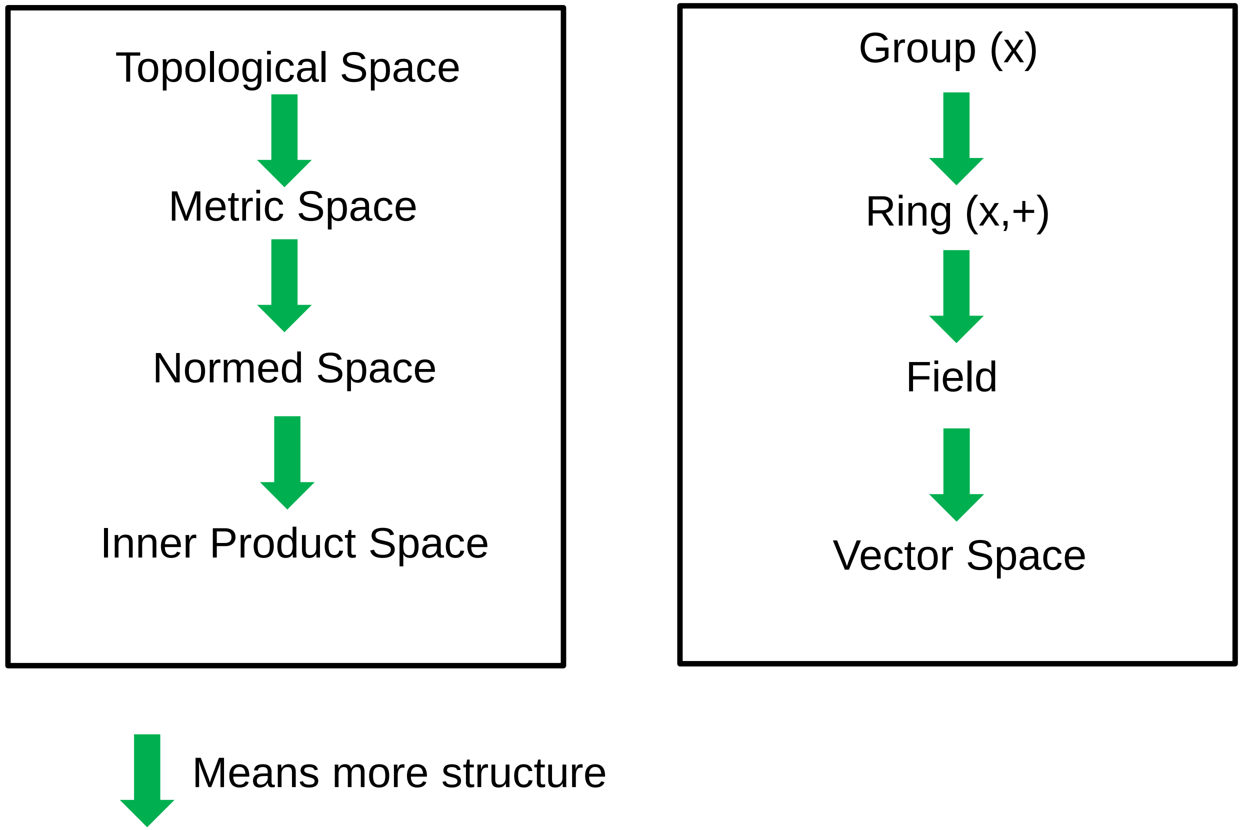
$y = f(f^{\mathrm{R}}(y)) = f(x)$

Design method

Check/test

This is an engineering problem.
Want f onto.

Mathematical Structures



Geometric Structures

Topological Space

↓

Metric Space

↓

Normed Space

↓

Inner Product Space

Algebraic Structures

Group (x)

↓

Ring (x,+)

↓

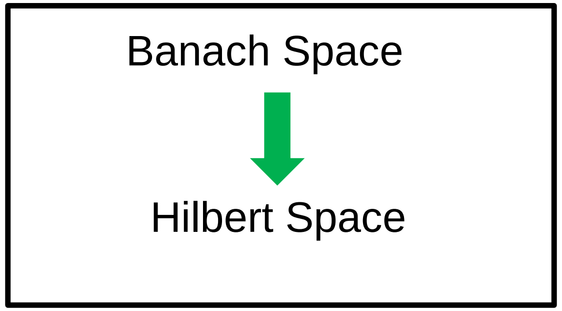
Field

↓

Vector Space

↓ Means more structure

Combined Structures



Banach Space

↓

Hilbert Space

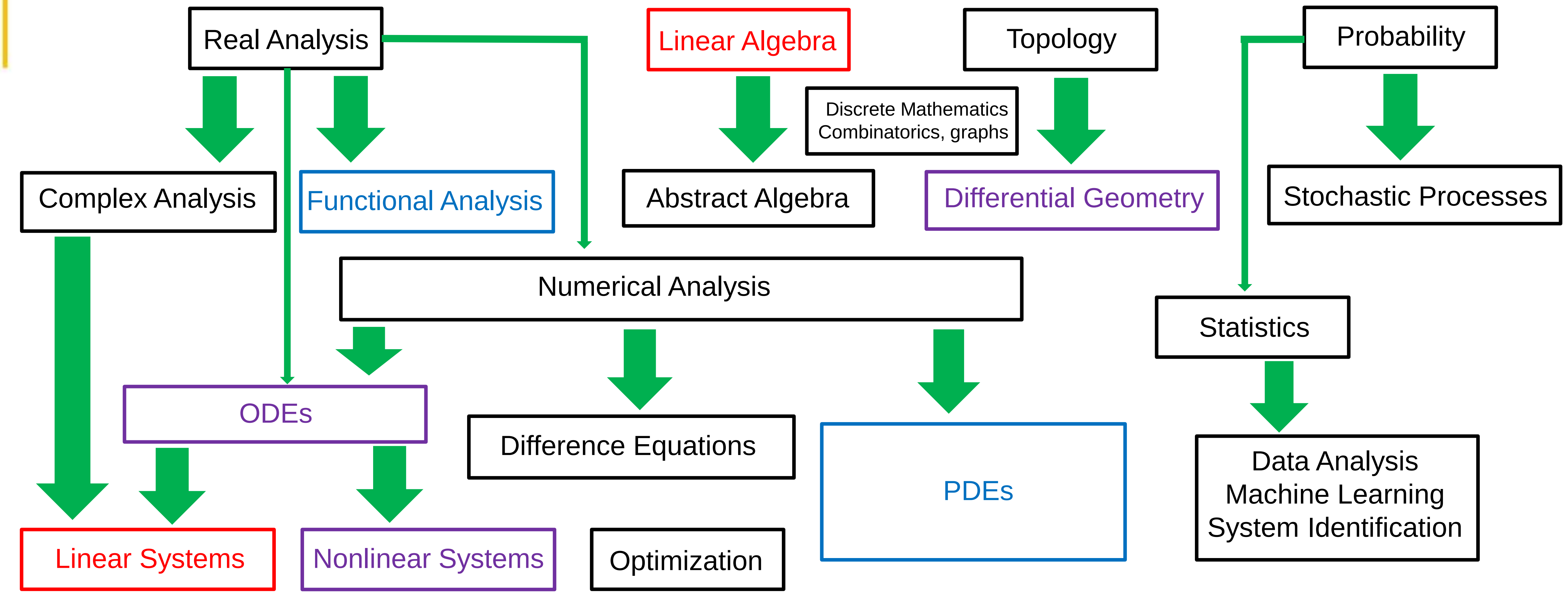
Euclidean space is a finite dimensional Hilbert space (linear algebra)

Function spaces are infinite-dimensional Banach or Hilbert spaces (functional analysis)

ODE's are infinite-dimensional in time but finite-dimensional in space.

PDE's are infinite-dimensional in space and time.

Math Beyond Calculus



Random Advice

How to Get Good at Writing Proofs?

- Master the mechanics of proofs
  - This talk discussed some of these

- Emulate good proofs
  - Find some well-written proofs
  - Study them
  - Mimic them
  - Develop your own style

- Take math courses that require lots of proofs
  - Practice, practice, practice

Research Advice

- Work on a feasible problem
  - You need an attack---idea or technique
- Alternate between general and special cases
  - Big picture versus little picture
- Isolate the key issues
- Experiment and tinker
  - Numerically or in the lab
  - Collect evidence and formulate conjectures
- Look for anomalies
  - “That’s funny” is a clue (do not expect eureka)
- Solve problems in two different ways if you can
  - Confirms that you are correct; can yield new identities
- Avoid trying to prove a theorem until you are very sure it is true
  - But: Attempts to prove a conjecture can help us understand why a conjecture may be false

Career Advice

- Be passionate about your research
- Be curious about everything
- Be creative
- Remember that research is not a performing art
  - You only need to get it right once

- Admit ignorance
  - Even professors don’t know everything (!)
  - Practice saying “I don’t know” without reservation or hesitation
  - “Not knowing is not failure, it's the first step to understanding” (TED Talk)

- Learn
  - Life is for learning
  - Learn how to learn
  - Learn how to listen
  - Context and need motivate effective learning

Extra on Writing

Adverbs

- Adverbs modify adjectives and verbs
  - Clearly seen
  - Very hot

- Adverbs add emphasis
  - They are usually imprecise
  - They are essentially hype

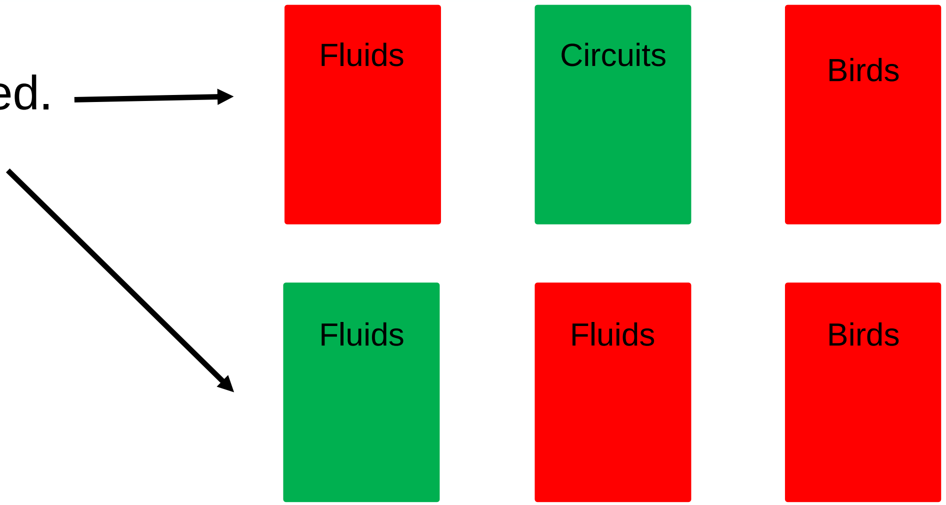
- Exception: The equation has exactly one solution.

Any

- Show that the sum of any two odd integers is even
  - 1+1 = 2
- Show that sum of ANY two odd integers is even
  - 2n+1 + 2m+1 = 2(n+m+1)
- Any seems to have two meanings
  - Some (there exists)
  - All (for all)
- What does “any” really mean?
  - I do not have a dog, and I do not have a cat.
  - I do not have a dog and a cat. (??)
  - I do not have a dog or a cat.
  - I do not have any pets.
- Halmos: Never use “any” in a math paper

Which versus That

- Please read the fluid dynamics book, which is red.
- Please read the fluid dynamics book that is red.



- Which
  - Not essential, always preceded by a comma

- That
  - Essential, never preceded by a comma
  - “that” should immediately follow what it modifies
  - We derive an adaptive control algorithm in discrete time that is globally convergent.
  - We derive a discrete-time adaptive control algorithm that is globally convergent.

- Correct usage adds clarity and precision

Style: Equations and Punctuation

- Equations are grammatically integral part of sentences
  - Not preceded by colons
  - Deserve punctuation

- Label equations
  - Label every equation except inside a proof

- Semicolons are nice for connecting short sentences
  - Some sentences are long for various reasons. Others are not.
  - Some sentences are long for various reasons; others are not.

- Colons are an interruption
  - The set X contains three elements: a,b,c.
  - The set X contains three elements, namely, a,b,c.

Style: Sentences and Paragraphs

- Sentences
  - Smooth, flowing, hooked together; no jumps, no non sequiturs

- Paragraph structure
  - Start by announcing the point or topic
  - Coherent discussion

- Paragraph style
  - Not too long (1/3rd of a page is a nice length)
  - Could be just one sentence
  - Announce each paragraph with indentation
  - Every theorem, proposition, lemma, corollary is exactly 1 paragraph
  - A proof can have many paragraphs.
    - Indicate end with a box □
  - An example can have many paragraphs.
    - Indicate end with a diamond ◊

Mechanics: Super String Calculations

$$
\begin{aligned}
Q_j(q,\dot{q},t) &= \vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \partial_{\dot{q}_j}\left(\vec{v}_{y_1/w/\mathrm{A}}\Big|_{\mathrm{A}}\right) - \vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \partial_{\dot{q}_j}\left(\vec{v}_{y_2/w/\mathrm{A}}\Big|_{\mathrm{A}}\right) \\
&= \vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \left[\partial_{\dot{q}_j}\left(\vec{v}_{y_1/w/\mathrm{A}}\Big|_{\mathrm{A}}\right) - \partial_{\dot{q}_j}\left(\vec{v}_{y_2/w/\mathrm{A}}\Big|_{\mathrm{A}}\right)\right] \\
&= \vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \left[\partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}} \times \vec{r}_{y_1/w}\Big|_{\mathrm{A}}\right) - \partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}} \times \vec{r}_{y_2/w}\Big|_{\mathrm{A}}\right)\right] \\
&= \vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \left[\partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}}\right) \times \left(\vec{r}_{y_1/w}\Big|_{\mathrm{A}} - \vec{r}_{y_2/w}\Big|_{\mathrm{A}}\right)\right] \\
&= \vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \left[\partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}}\right) \times \vec{r}_{y_1/y_2}\Big|_{\mathrm{A}}\right] \\
&= -\vec{f}\Big|_{\mathrm{A}}^{\mathrm{T}} \left[\vec{r}_{y_1/y_2}\Big|_{\mathrm{A}} \times \partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}}\right)\right] \\
&= -\left(\vec{f}\Big|_{\mathrm{A}} \times \vec{r}_{y_1/y_2}\Big|_{\mathrm{A}}\right)^{\mathrm{T}} \partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}}\right) \\
&= \left(\vec{r}_{y_1/y_2}\Big|_{\mathrm{A}} \times \vec{f}\Big|_{\mathrm{A}}\right)^{\mathrm{T}} \partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}}\right) \\
&= \vec{M}_k\Big|_{\mathrm{A}}^{\mathrm{T}} \partial_{\dot{q}_j}\left(\vec{\omega}_{\mathrm{B/A}}\Big|_{\mathrm{A}}\right).
\end{aligned}
$$