

Microsoft Research

Summit 2022

# The Digital Revolution in Mathematics

Jeremy Avigad, Carnegie Mellon University

# New technology for mathematics

Lean as a theorem prover: a platform for

- defining mathematical objects,
- stating theorems,
- and writing complex proofs.

Lean as a programming language:

- a performant functional programming language, with
- means for writing specifications and proving that programs satisfy them.

Combining the two brings

- computational methods to mathematical reasoning, and
- mathematical reasoning to computation.

## New technology for mathematics

The technology is based on *formal methods* in computer science, namely, logic-based computational methods for specifying and verifying software and hardware.

Until recently, very few mathematicians were using proof assistants.

In 2017, a number of mathematicians discovered Lean, and the Lean community was born.

# The Lean community

The screenshot shows a web browser window with the URL `leanprover-community.github.io`. The page layout includes a left-hand navigation menu, a main content area with a large logo, and a grid of three introductory boxes.

**Lean Community**

- Community**
  - Zulip chat
  - GitHub
  - Community information
  - Papers about Lean
  - Projects using Lean
- Installation**
  - Get started
  - Debian/Ubuntu installation
  - Generic Linux installation
  - MacOS installation
  - Windows installation
  - Online version (no installation)
  - Using leanproject
  - The Lean toolchain
- Documentation**
  - Learning resources (start here)
  - API documentation
  - Calc mode
  - Conv mode
  - Simplifier
  - Tactic writing tutorial
  - Well-founded recursion
  - About MWEs
- Library overviews**
  - Library overview
  - Undergraduate maths
  - Wiedijk's 100 theorems
- Theory docs**
  - Category theory
  - Linear algebra
  - Natural numbers
  - Sets and set-like objects
  - Topology

**LEAN**  
*Community*

## Lean and its Mathematical Library

The Lean [theorem prover](#) is a proof assistant developed principally by Leonardo de Moura at Microsoft Research.

The Lean mathematical library, *mathlib*, is a community-driven effort to build a unified library of mathematics formalized in the Lean proof assistant. The library also contains definitions useful for programming. This project is very active, with many regular contributors and daily activity.

The contents, design, and community organization of mathlib are described in the paper [The Lean mathematical library](#), which appeared at CPP 2020. You can get a bird's eye view of what is in the library by reading the [library overview](#). You can also have a look at our [repository statistics](#) to see how it grows and who contributes to it.

### Try it!

You can try Lean in your web browser, install it in an isolated folder, or go for the full install. Lean is free, open source software. It works on Linux, Windows, and MacOS.

[Try the online version of Lean](#)

[Installation instructions](#)

[Working on Lean projects](#)

### Learn to Lean!

You can learn by playing a game, following tutorials, or reading books.

[Learning resources](#)

[Theorem Proving in Lean \(an introduction\)](#)

[API documentation of mathlib](#)

### Meet the community!

Lean has very diverse and active community. It gathers mostly on a Zulip chat and on GitHub. You can get involved and join the fun!

[Meet us](#)

[How to contribute](#)

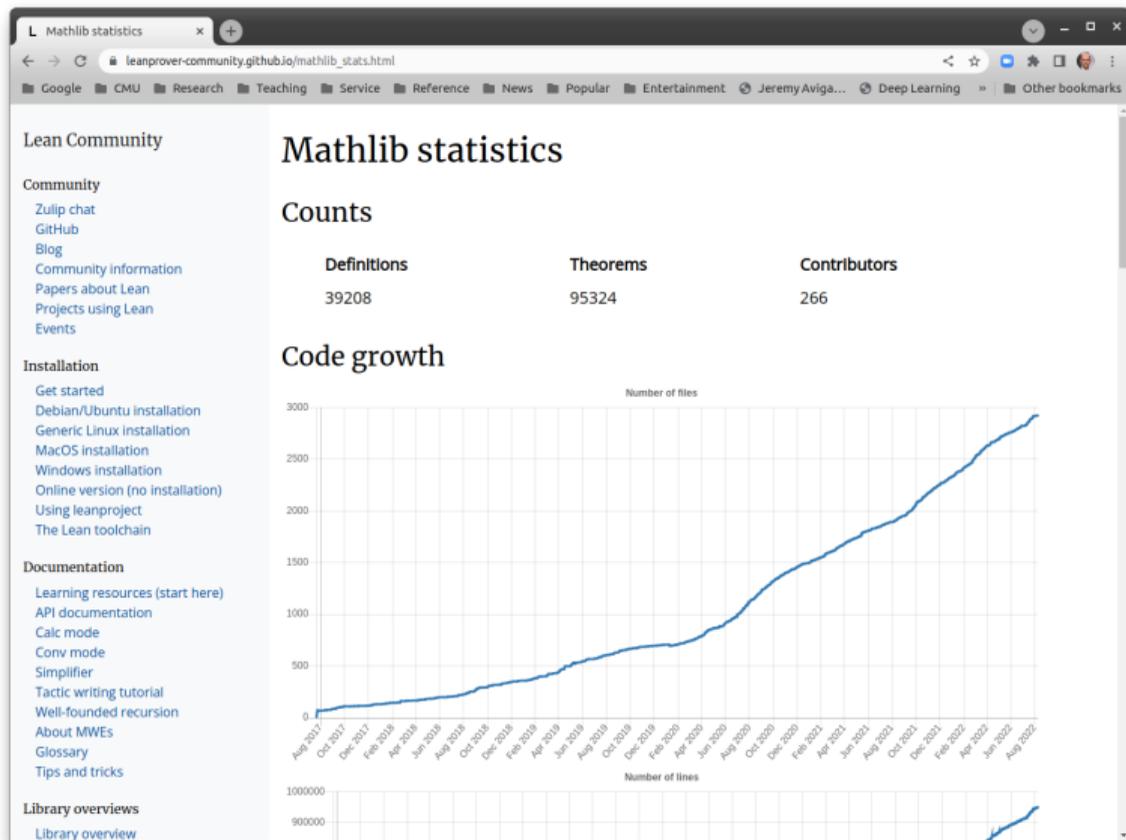
[Papers involving Lean](#)

# The Lean community

Where we are now:

- Hundreds of people have contributed to Lean's library, *mathlib*.
- The library has almost a million lines of formal definitions, theorems, and proofs.
- Lean's social media channel on *Zulip* gets hundreds of messages every day.
- There are a growing number of papers, conferences, and workshops dedicated to formalization of mathematics and Lean.

# Mathlib statistics



# The Lean Zulip channel

The screenshot shows a Zulip chat window for the 'Lean' channel. The browser address bar is 'leanprover.zulipchat.com/#all\_messages'. The chat interface includes a sidebar with navigation options like 'All messages', 'Recent topics', and 'STREAMS'. The main chat area displays several messages:

- Message 1:** A message in the 'general' stream titled 'Automatically checking function is polynomial time' from 'Today'. It contains Lean code defining `nat_polytime`, `eg_infer`, and `eg_multi_arg`. The code is as follows:

```
/- Every natural number is polynomial time (one way to make sense of this is that every constant is a 0)
instance nat_polytime (n : ℕ) : polytime n := sorry

-- Here the 'a' in 'id a' is ignored and only the relevant parameters are considered
instance eg_infer : polytime (λ n, nat.sq (id n)) := infer_instance

-- A bit more work is required here, but this complicated is still derivable from typeclass inference
instance eg_multi_arg : polytime (λ n, (nat.add (nat.sq n)) 5) := by {
  show polytime (λ n, ((λ x, nat.add (nat.sq x)) n) 5),
  apply_instance,
}
```

The message concludes with: "This is by no means a complete solution, but I think it addresses some parts of the question (such as ignoring implicit arguments). There may also be a better way to do this that avoids the `show` in the last example, but I am not aware of one. I suspect that the three typeclass instances above will be enough to handle a wide variety of cases thanks to currying."
- Message 2:** A message in the 'general' stream titled 'Namespaces for ring lemmas' from '1:52 PM'. It shows a user 'Vaik Dillies' mentioning `docs#sub_eq_add_neg`.
- Message 3:** A reply from 'Michael Stoll' at '1:54 PM' asking: "Thanks. Why is `neg_add_eq_sub` in the other direction?"
- Message 4:** A message in the 'is there code for X?' stream titled 'more linear algebra' from '2:11 PM'. It is from 'Junyan Xu' and includes a proof for the lemma `matrix.range_to_lin'`. The code is:

```
import analysis.inner_product_space.basic
import data.real.basic
import data.matrix.rank

lemma matrix.range_to_lin' (R) [comm_semiring R] (m n) [fintype n] [decidable_eq n]
  (M : matrix (m × m) R) (M' : matrix (m × n) R) (M'' : matrix (m × n) R) (M''' : matrix (m × n) R) :
  M' ∈ M.range_to_lin' M'' M''' := ...
```

A link to 'More...' is provided. The message ends with: "To show it spans, it should be easy from linear independence + `docs#span_eq_top_of_linear_independent_of_card_eq_firank`."
- Message 5:** A message in the 'maths' stream titled 'Missing facts about 'locally\_constant'' from '2:16 PM' by 'Anatole Dedecker'.

At the bottom of the chat, there is a message input field containing 'Message #1s there code for X? > more linear algebra' and buttons for 'New topic' and 'New private message'.

## Notable achievements

For example:

- Jesse Han and Floris van Doorn gave the first formal verification of the independence of the continuum hypothesis, an important result in set theory.
- Johan Commelin led the *Liquid Tensor Experiment*, in response to a challenge by Fields Medalist Peter Scholze.
- Bhavik Mehta and Thomas Bloom verified an important result in number theory.

Scholze: “I find it absolutely insane that interactive proof assistants are now at the level that within a very reasonable time span they can formally verify difficult original research.”

## Lean in the news

Lean has been getting good press:

- *Quanta*: “Building the mathematical library of the future”
- *Quanta*: “At the Math Olympiad, computers prepare to go for the gold”
- *Nature*: “Mathematicians welcome computer-assisted proof in ‘grand unification’ theory”
- *Quanta*: “Proof Assistant Makes Jump to Big-League Math”

Kevin Buzzard recently gave a talk, “The Rise of Formalism in Mathematics,” at the 2022 International Congress of Mathematicians.

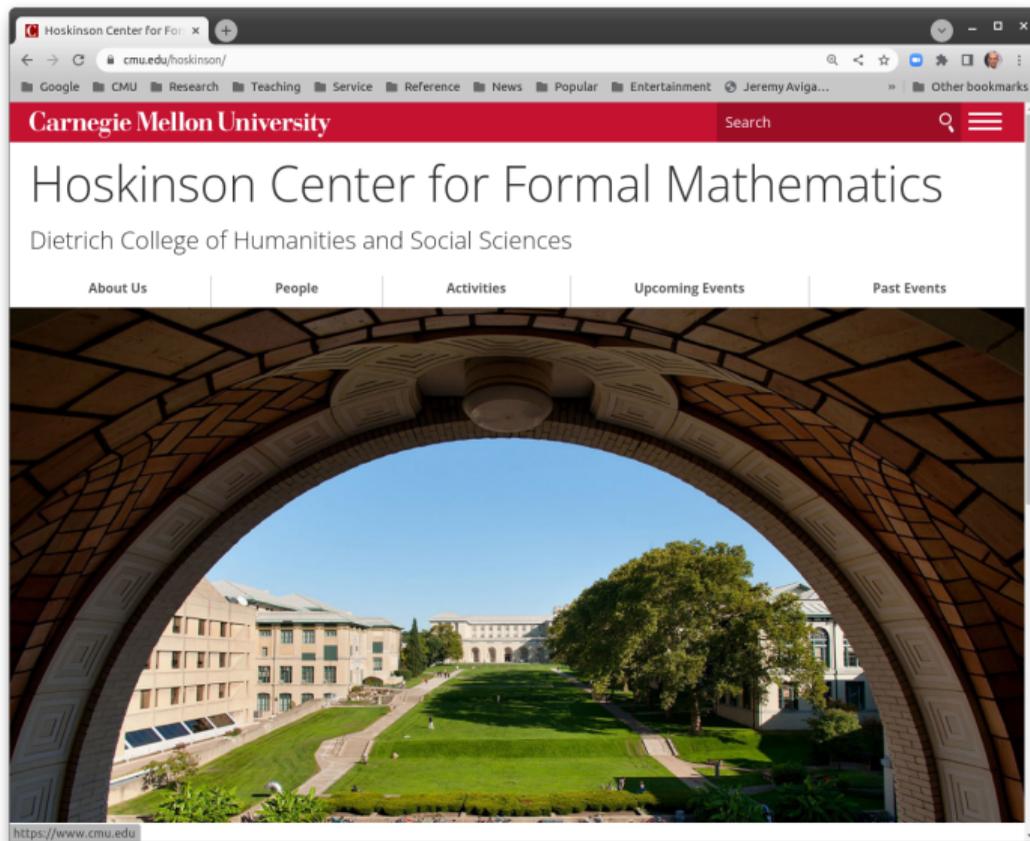
# The Hoskinson Center

In September of 2021, Carnegie Mellon launched the **Charles C. Hoskinson Center for Formal Mathematics**.

It is dedicated to the use of formal computational methods in mathematical research and education.

Most of the center's activities are based on Lean.

# The Hoskinson Center



# What's the big deal?

Some are calling this the start of a revolution in mathematics.

It's reasonable to ask: why all the excitement?

This talk:

- Lean and formal methods in mathematics
- the nature of mathematical revolutions
- the digital revolution in mathematics

# Revolutions in mathematics

Examples:

- the appearance of deductive reasoning in ancient Greece.
- the rise of algebraic methods
- the birth of calculus
- the inauguration of infinitary reasoning in the 19th century
- the advent of the computer and numeric computation

## The rise of algebraic methods

The roots of algebra can be found in Al-Khwarizmi (9th century), and even earlier in ancient Greece.

The turning point in the early 17th century:

- the development of better algebraic notation
- the mathematization of natural science
- the use of algebraic methods to solve problems in geometry and science

## The rise of algebraic methods

Cardano's solution  $x^3 + px = q$  in 1545:

“Cube the third part of the number of unknowns, to which you add the square of half the number of the equation, and take the root of the whole, that is, the square root, which you will use, in one case adding the half of the number which you just multiplied by itself, in the other case subtracting the same half, and you will have a binomial and apotome respectively; then subtract the cube root of the apotome from the cube root of the binomial, and the remainder from this is the value of the unknown.”

Today:

$$\sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

## The rise of algebraic methods

Galileo, *The Assayer*, 1623:

“Philosophy [i.e. natural philosophy] is written in this grand book — I mean the Universe — which stands continually open to our gaze, but it cannot be understood unless one first learns to comprehend the language and interpret the characters in which it is written. It is written in the language of mathematics. . .”

# The rise of algebraic methods

1637:

## The Geometry of René Descartes

### BOOK I

#### PROBLEMS THE CONSTRUCTION OF WHICH REQUIRES ONLY STRAIGHT LINES AND CIRCLES

ANY problem in geometry can easily be reduced to such terms that a knowledge of the lengths of certain straight lines is sufficient for its construction.<sup>[1]</sup> Just as arithmetic consists of only four or five operations, namely, addition, subtraction, multiplication, division and the extraction of roots, which may be considered a kind of division, so in geometry, to find required lines it is merely necessary to add or subtract other lines; or else, taking one line which I shall call unity in order to

## The rise of algebraic methods

Often it is not necessary thus to draw the lines on paper, but it is sufficient to designate each by a single letter. Thus, to add the lines BD and GH, I call one  $a$  and the other  $b$ , and write  $a + b$ . Then  $a - b$  will indicate that  $b$  is subtracted from  $a$ ;  $ab$  that  $a$  is multiplied by  $b$ ;  $\frac{a}{b}$  that  $a$  is divided by  $b$ ;  $aa$  or  $a^2$  that  $a$  is multiplied by itself;  $a^3$  that this result is multiplied by  $a$ , and so on, indefinitely.<sup>[9]</sup> Again, if I wish to extract the square root of  $a^2 + b^2$ , I write  $\sqrt{a^2 + b^2}$ ; if I wish to extract the cube root of  $a^3 - b^3 + ab^2$ , I write  $\sqrt[3]{a^3 - b^3 + ab^2}$ , and similarly for other roots.<sup>[10]</sup> Here it must be observed that by  $a^2$ ,  $b^3$ , and similar expressions, I ordinarily mean only simple lines, which, however, I name squares, cubes, etc., so that I may make use of the terms employed in algebra.<sup>[11]</sup>

# Mathematical revolutions

Mathematical revolutions don't happen all at once.

They are not revolutions in the sense of overthrowing the old order. Rather they incorporate the past and build on it.

They open up new capacities for thought:

- Things that were hard become easier.
- Problems that were out of reach become solvable.
- New questions and problems arise.

## What's important to mathematics

- Mathematics has practical applications.
- What we really care about in mathematical *understanding*.
- We value powerful intuitions, insights, and ideas.
- We need to communicate these ideas to one another in precise ways.
- The subject provides extraordinary means to come to consensus as to whether a proof is correct.
- The main challenge is complexity.

# The digital revolution

Remember the outline:

- Lean and formal methods in mathematics
- the nature of mathematical revolutions
- the digital revolution in mathematics

Let's think about the new technology in these terms.

## Verifying correctness

In early 2022, Thomas Bloom solved a problem posed by Paul Erdős and Ronald Graham.

The headline in Quanta read “**Math’s ‘Oldest Problem Ever’ Gets a New Answer.**”

Within in a few months, Bloom and Bhavik Mehta verified the correctness of the proof in Lean.

# Verifying correctness



**Timothy Gowers**

@wtgowers · Jun 13

Very excited that Thomas Bloom and Bhavik Mehta have done this. I think it's the first time that a serious contemporary result in "mainstream" mathematics doesn't have to be checked by a referee, because it has been checked formally. Maybe the sign of things to come ... 1/

 **Kevin Buzzard** @XenaProject · Jun 12

Happy to report that Bloom went on to learn Lean this year and, together with Bhavik Mehta, has now formalised his proof in Lean `b-mehta.github.io/unit-fractions/` (including formalising the Hardy-Littlewood circle method), finishing before he got a referee's report for the paper ;-)

[Show this thread](#)



## Exploring mathematics

Similarly, Peter Scholze wrote:

“I am excited to announce that the Experiment has verified the entire part of the argument that I was unsure about.”

But he went on:

“[H]alf a year ago, I did not understand why the argument worked. . . .”

The formalization helped him realize that

“the key thing happening is a reduction from a non-convex problem over the reals to a convex problem over the integers.”

# Collaboration

The liquid tensor experiment is also a model for digital collaboration.

- The formalization was in kept in a shared online repository.
- Participants followed an informal blueprint with links to the repository.
- Participants were in constant contact on Zulip.
- Lean made sure the pieces fit together.

# Collaboration

Blueprint for the Liquid Tensor Experiment

Introduction

1 First part

1.1 Breen–Deligne data

**1.2 Variants of normed groups**

1.3 Spaces of convergent power series

1.4 Some normed homological algebra

1.5 Completions of locally constant functions

1.6 Polyhedral lattices

1.7 Key technical result

2 Second part

3 Bibliography

Section 1 graph

Section 2 graph

## 1.2 Variants of normed groups

Normed groups are well-studied objects. In this text it will be helpful to work with the more general notion of *semi-normed group*. This drops the separation axiom  $\|x\| = 0 \iff x = 0$  but is otherwise the same as a normed group.

The main difference is that this includes “uglier” objects, but creates a “nicer” category: semi-normed groups need not be Hausdorff, but quotients by arbitrary (possibly non-closed) subgroups are naturally semi-normed groups.

Nevertheless, there is the occasional use for the more restrictive notion of normed group, when we come to polyhedral lattices below (see Section 1.6).

In this text, a morphism of (semi-)normed groups will always be bounded. If the morphism is supposed to be norm-nonincreasing, this will be mentioned explicitly.

**Definition 1.2.1** ✓

Let  $r > 0$  be a real number. An  $r$ -normed  $\mathbb{Z}[T^{\pm 1}]$ -module is a semi-normed group  $V$  endowed with an automorphism  $T: V \rightarrow V$  such that for all  $v \in V$  we have  $\|T(v)\| = r\|v\|$ .

The remainder of this subsection sets up some algebraic variants of semi-normed groups.

**Definition 1.2.2** ✓

A *pseudo-normed group* is an abelian group  $(M, +)$ , together with an increasing filtration  $M_c \subseteq M$  of subsets  $M_c$  indexed by  $\mathbb{R}_{\geq 0}$ , such that each  $M_c$  contains 0, is closed under negation, and  $M_{c_1} + M_{c_2} \subseteq M_{c_1+c_2}$ . An example would be  $M = \mathbb{R}$  or  $M = \mathbb{Q}_p$  with  $M_c := \{x : |x| \leq c\}$ .

A pseudo-normed group  $M$  is *exhaustive* if  $\bigcup_c M_c = M$ .

All pseudo-normed groups that we consider will have a topology on the filtration sets  $M_c$ . The most general variant is the following notion.

**Definition 1.2.3** ✓

A pseudo-normed group  $M$  is *CH-filtered* if each of the sets  $M_c$  is endowed with a topological space structure making it a compact Hausdorff space, such that following maps are all continuous:

- the inclusion  $M_{c_1} \rightarrow M_{c_2}$  (for  $c_1 \leq c_2$ );
- the negation  $M_c \rightarrow M_c$

# Teaching

An interactive proof assistant is a powerful tool for teaching mathematics.

It empowers students to explore mathematical reasoning on their own.

There have been workshops and conference sessions dedicated to learning how to use the technology effectively.

# Teaching

Learning Mathematics with LEAN

6 April 2022 10:30-16:15 Loughborough University campus - room to be confirmed

   LONDON MATHEMATICAL SOCIETY EST. 1865

This workshop is funded by the London Mathematical Society (LMS).

Join us at Loughborough University for a workshop about using the automated theorem prover Lean (<https://leanprover.github.io>) to teach first year pure mathematics. Speakers will talk about their experiences of using Lean for teaching and the resources they created. There will also be the opportunity for hands-on sessions to try some of resources designed for students. You can join the workshop here at Loughborough or online - but

**Department of Mathematics Education**

- About
- Research
- Teaching
- Staff
- MENU

**Department of Mathematics Education**

- About the Department
- Research
- Teaching
- Staff
- Postgraduate study
- News
- Events**
- Contact us

Contact us

- +44 (0)1509 222681
- Send email
- Department of Mathematics Education

## Other applications

This only scratches the surface.

Lean can also be used as a platform for numerical and symbolic computation, as well as automated reasoning and machine learning.

- It enables us to apply computational tools to precise mathematical formulations.
- It can be used to verify and interpret the computational results.

# The digital revolution in mathematics

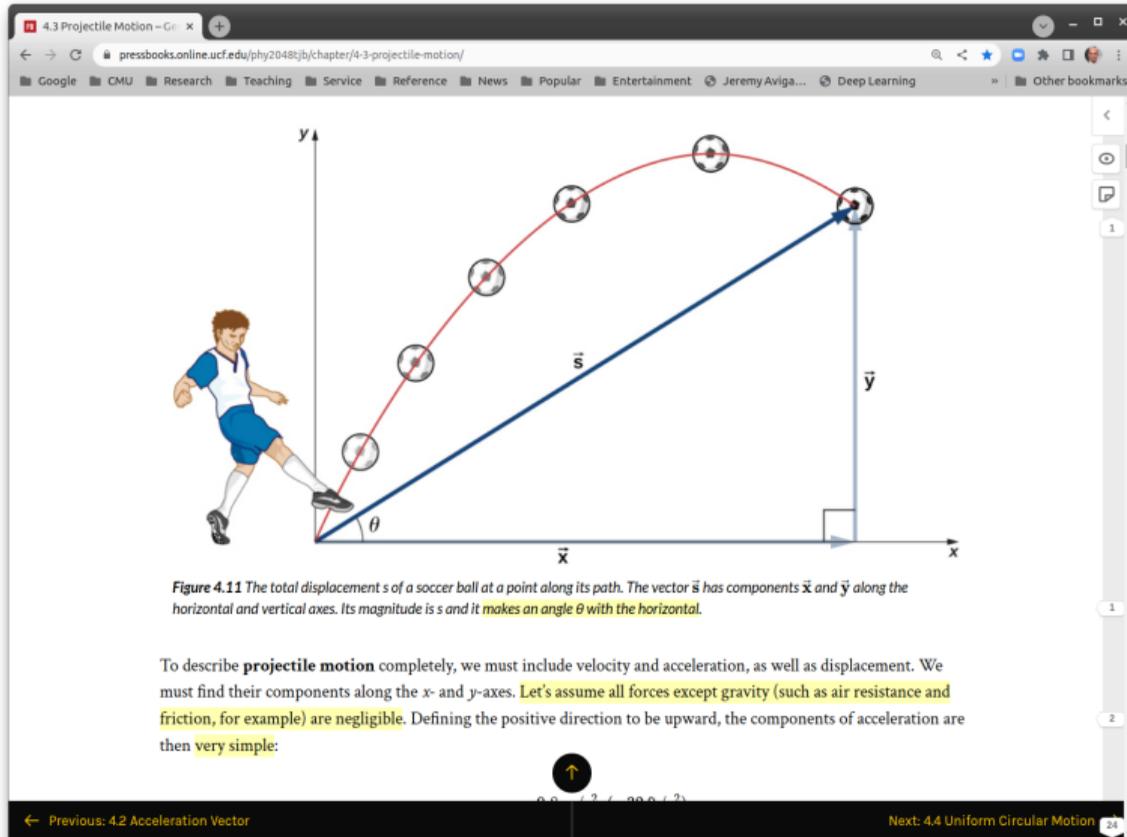
Formal technology can help us:

- build mathematical libraries,
- verify results,
- explore new concepts,
- collaborate,
- teach mathematics,
- carry out mathematical computation more rigorously, and
- use AI to discover new mathematics.

## Revolutions in mathematics

I sometimes wonder whether people *knew*, at the time, that they were in middle of a revolution.

# Projectile motion



4.3 Projectile Motion – Ge... x

pressbooks.online.ucf.edu/phy2048t/b/chapter/4-3-projectile-motion/

Google CMU Research Teaching Service Reference News Popular Entertainment Jeremy Aviga... Deep Learning Other bookmarks

$y$

$x$

$\vec{s}$

$\vec{x}$

$\vec{y}$

$\theta$

**Figure 4.11** The total displacement  $s$  of a soccer ball at a point along its path. The vector  $\vec{s}$  has components  $\vec{x}$  and  $\vec{y}$  along the horizontal and vertical axes. Its magnitude is  $s$  and it makes an angle  $\theta$  with the horizontal.

To describe **projectile motion** completely, we must include velocity and acceleration, as well as displacement. We must find their components along the  $x$ - and  $y$ -axes. Let's assume all forces except gravity (such as air resistance and friction, for example) are negligible. Defining the positive direction to be upward, the components of acceleration are then very simple:

← Previous: 4.2 Acceleration Vector

Next: 4.4 Uniform Circular Motion 25

# Projectile motion

4.3 Projectile Motion - Ge...

pressbooks.online.ucf.edu/phy2048tjb/chapter/4-3-projectile-motion/

Horizontal Motion

$$v_{0x} = v_x, x = x_0 + v_x t$$

Vertical Motion

$$y = y_0 + \frac{1}{2}(v_{0y} + v_y)t$$
$$v_y = v_{0y} - gt$$
$$y = y_0 + v_{0y}t - \frac{1}{2}gt^2$$
$$v_y^2 = v_{0y}^2 - 2g(y - y_0)$$

Using this set of equations, we can analyze projectile motion, keeping in mind some important points.

**Problem-Solving Strategy: Projectile Motion**

1. Resolve the motion into horizontal and vertical components along the  $x$ - and  $y$ -axes. The magnitudes of the components of displacement  $\vec{s}$  along these axes are  $x$  and  $y$ . The magnitudes of the components of velocity  $\vec{v}$  are  $v_x = v \cos \theta$  and  $v_y = v \sin \theta$ , where  $v$  is the magnitude of the velocity and  $\theta$  is its direction relative to the horizontal, as shown in [Figure](#).
2. Treat the motion as **two independent one-dimensional motions**: one horizontal and the other vertical. Use the kinematic equations for horizontal and vertical motion presented earlier.
3. Solve for the unknowns in the two separate motions: one horizontal and one vertical. Note that the only com...

← Previous: 4.2 Acceleration Vector

Next: 4.4 Uniform Circular Motion

22

# Projectile motion



## THIRD DAY

[190]

### CHANGE OF POSITION. [*De Motu Locali*]



Y purpose is to set forth a very new science dealing with a very ancient subject. There is, in nature, perhaps nothing older than motion, concerning which the books written by philosophers are neither few nor small; nevertheless I have discovered by experiment some properties of it which are worth knowing and which have not hitherto been either observed or demonstrated. Some superficial observations have been made, as, for instance, that the free motion [*naturalem motum*] of a heavy falling body is continuously accelerated; \* but to just what extent this acceleration occurs has not yet been announced; for so far as I know, no one has yet pointed out that the distances traversed, during equal intervals of time, by a body falling from rest, stand to one another in the same ratio as the odd numbers beginning with unity.†

It has been observed that missiles and projectiles describe a curved path of some sort; however no one has pointed out the fact that this path is a parabola. But this and other facts, not few in number or less worth knowing, I have succeeded in proving; and what I consider more important, there have been opened up to this vast and most excellent science, of which my

## Projectile motion

### THEOREM I, PROPOSITION I

A projectile which is carried by a uniform horizontal motion compounded with a naturally accelerated vertical motion describes a path which is a semi-parabola.

## Revolutions in mathematics

But these and other facts, not few in number or less worth knowing, I have succeeded in proving; and what I consider more important, there have been opened up to this vast and most excellent science, of which my work is merely the beginning, ways and means by which other minds more acute than mine will explore its remote corners.

## Revolutions in mathematics

So that we may say the door is now opened, for the first time, to a new method fraught with numerous and wonderful results which in future years will command the attention of other minds.

Microsoft Research

**Summit 2022**

**Thank you**

 <https://www.andrew.cmu.edu/user/avigad>