AMERICAN MATHEMATICAL SOCIETY

CURRENT EVENTS BULLETIN

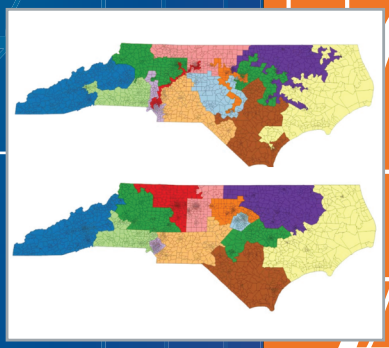
Friday, January 17, 2020, 1:00 PM to 4:45 PM

Room 205, Colorado Convention Center
Joint Mathematics Meeting, Denver, CO

1:00 PM | **Jordan S. Ellenberg**
University of Wisconsin-Madison

Geometry, inference, and democracy

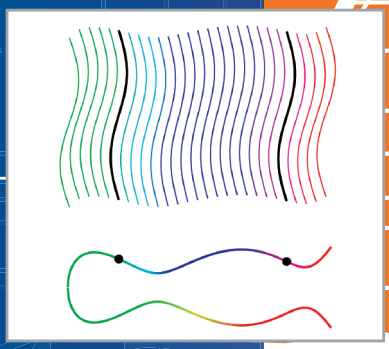
There is no democracy without representation. How can mathematics help to ensure it?



2:00 PM | **Bjorn Poonen**
Massachusetts Institute of Technology

A *p*-adic approach to rational points on curves

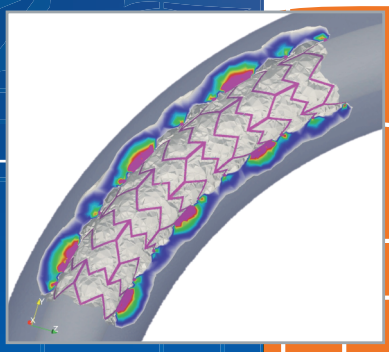
A new proof of a famous theorem!



3:00 PM | **Sunčica Čanić**
University of California, Berkeley

Recent progress on moving boundary problems

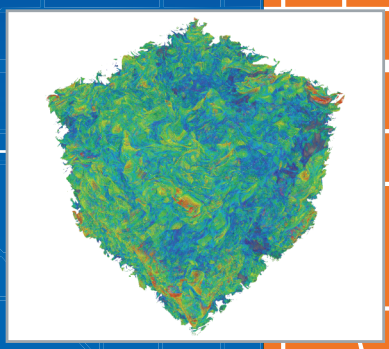
Cardiovascular interventions with the help of math (and trying to understand how we swim).



4:00 PM | **Vlad C. Vicol**
Courant Institute of Mathematical Sciences, New York University

Convex integration and fluid turbulence

Turbulence is one of the great abiding mysteries, but there is progress toward understanding.



AMS AMERICAN MATHEMATICAL SOCIETY
Advancing research. Creating connections.

Organized by David Eisenbud, University of California, Berkeley

Introduction to the Current Events Bulletin

Will the Riemann Hypothesis be proved this week? What is the Geometric Langlands Conjecture about? How could you best exploit a stream of data flowing by too fast to capture? I think we mathematicians are provoked to ask such questions by our sense that underneath the vastness of mathematics is a fundamental unity allowing us to look into many different corners -- though we couldn't possibly work in all of them. I love the idea of having an expert explain such things to me in a brief, accessible way. And I, like most of us, love common-room gossip.

The Current Events Bulletin Session at the Joint Mathematics Meetings, begun in 2003, is an event where the speakers do not report on their own work, but survey some of the most interesting current developments in mathematics, pure and applied. The wonderful tradition of the Bourbaki Seminar is an inspiration, but we aim for more accessible treatments and a wider range of subjects. I've been the organizer of these sessions since they started, but a varying, broadly constituted advisory committee helps select the topics and speakers. Excellence in exposition is a prime consideration.

A written exposition greatly increases the number of people who can enjoy the product of the sessions, so speakers are asked to do the hard work of producing such articles. These are made into a booklet distributed at the meeting. Speakers are then invited to submit papers based on them to the *Bulletin of the AMS*, and this has led to many fine publications.

I hope you'll enjoy the papers produced from these sessions, but there's nothing like being at the talks -- don't miss them!

David Eisenbud, Organizer
Mathematical Sciences Research Institute
de@msri.org

For PDF files of talks given in prior years, see
http://www.ams.org/ams/current-events-bulletin.html.
The list of speakers/titles from prior years may be found at the end of this booklet.

GEOMETRY, INFERENCE, AND DEMOCRACY

JORDAN S. ELLENBERG

Abstract. Decisions about how the population of the United States should be divided into legislative districts have powerful and not fully understood effects on the outcomes of elections. The problem of understanding what we might mean by "fair districting" intertwines mathematical, political, and legal reasoning; but only in recent years has the academic mathematical community gotten directly involved in the process. I'll report on recent progress in this area, how newly developed mathematical tools have affected real political decisions, and what remains to be done.

1. What does it mean to be represented?

All democratic states are founded on the principle that every citizen's views are to be represented in the conduct of the government. This principle, like most important principles, is easy to state, difficult to make precise, and almost impossible to implement in a fully satisfying way.

For one thing, democratic states are big. Even a modestly sized city is large enough that it would be impractical for every decision about zoning, school curriculum, public transport, and taxes to be put to a public plebiscite, let alone to arrive at a consensus. So modern governments typically operate under some form of representative democracy, where a small group of legislators are elected by the population to write laws and vote on their passage. But how to generate this group of popular representatives? There are a lot of different ways. In Israel, voters vote for their preferred political party, which is awarded a number of seats in the Knesset roughly in proportion to its share of the popular vote, and then the party chooses the occupants of those seats. For the Senate of the Phillippines, each voter casts a vote for as many as twelve candidates, and the top twelve vote-getters overall join the Senate. The most common means of choosing representatives, though, is the one used by the United States Congress and by the legislatures of most of the states; the population is divided up into *legislative districts*, and each district chooses a representative by plurality vote. Under a district system, every voter has a specific representative who, at least in theory, governs on their behalf.

In some systems, this partition of the electorate reflects natural (or at least historically settled) political divisions. Each U.S. state has two U.S. Senators, because, at least formally, each state is a semi-autonomous political entity with its own particular interests. (These are thus examples of *multi-member districts*, a variant of single-member districting in which each district selects not just one but several representatives in the legislature; multimember districts are also used for a a few US state legislatures.) The partition is almost always along geographic lines, though not always. In New Zealand, Māori people have their own electoral districts, which are superimposed on the general districts; Māori voters have the choice in each election whether to vote in the Māori or the general district containing their residence. Or the partition might not have any geographic component

at all. In Hong Kong, there's a seat in the Legislative Council only teachers and school administrators can vote for, one of 30 seats elected by so-called functional constituencies. The Centuriate Assembly of the Roman Republican had constituencies separated by wealth bracket. In the upper house of the Oireachtas in Ireland, there is a three-seat constituency consisting of students and graduates of Trinity College Dublin, and another for alumni of the National University of Ireland.

Electoral districts within U.S. states are a different story. They are patches of land without much meaning. Nobody in the 2nd Congressional District of Wisconsin, where I live, wears a WI-2 sweatshirt, or could recognize the district from its silhouette. As for my state legislative district, I had to look it up to be sure I had the number right. These districts, have to be determined somehow, despite not having robust pre-existing political identities; that is, someone has to select a partition of the population of the state chosen from the ensemble of all possible partitions, a set of unmanageably large size. This process, historically, has not been the subject of much public attention. That has now changed. That's because we now understand something we didn't fully grasp before, which is at least in part a mathematical statement; that the way the population is broken up into districts has an enormous effect on the makeup of a legislature.

To some extent this is obvious. If I am in complete control of the districting of Wisconsin, with the power to partition the population any way I wish, and there is a cabal of like-minded people I want to be in control of the state, I could simply make each one of those people their own district, and then create one more district consisting of everybody else. My hand-picked candidates vote for themselves and then rule the legislature with at most one potential voice of opposition.

That's not fair! Certainly the people of Wisconsin, with the exception of the cabal itself, would be right to feel they were unrepresented in the decision-making of the state.

In real life, no one tries to implement a scheme like this. For one thing, state governments are not allowed to create districts with radically different populations; though until the 1964 Supreme Court decision in *Reynolds v. Sims*, state governments in the United States could, and did, do exactly this. In the United Kingdom, so-called "rotten boroughs" with only a few dozen voters were common until the 19th century.

Nowadays, in the United States and many other representative democracies (though not Canada!) districts are fixed by law to be approximately equal in size. That prevents the kind of cabalization of the legislature I described above. But it is *not*, it turns out, sufficient to keep the choice of partition from having a dramatic influence on the outcome of the election. The manipulation of district boundaries in order to achieve a desired outcome (most commonly a majority or supermajority of seats for one's own party, or protection of incumbent legislators) is often called *gerrymandering*, after a 19th-century Massachusetts governor sometimes thought of as a pioneer of the practice. In most states, the power to determine legislative districts is held by the legislature itself, creating an obvious incentive and opportunity for a disciplined partisan majority to preserve itself by gerrymandering against the will of an unfriendly electorate.

We are faced with the following ensemble of questions. We write $\mathbf{\Pi} = \{\Pi_1, \ldots, \Pi_k\}$ for a partition of the state's population into k subsets. We want to know:

- What properties should $\mathbf{\Pi}$ have in order to be considered "fair"?
- Given a proposed $\mathbf{\Pi}$, are there quantitative measurements of unfairness which are robust, reliable, and simple enough to be used by judges and courts who have to decide whether $\mathbf{\Pi}$ is too unfair to use?
- In a U.S. context, what kind of constraints on $\mathbf{\Pi}$ do the US and state constitutions allow us to impose, and what kind of constraints do those constitutions *require* us to impose?

As you can see, these questions are not purely mathematical in nature. They have a legal, a political, and a philosophical strands as well, and the strands can't really be unwound from each other. If mathematicians work on these problems alone, ignoring the other strands, the results are not going to be very useful. ("Why don't we just draw a grid over the state and make each box a district....?") But when lawyers and politicians think about redistricting while neglecting the mathematical strand, the result of their work will be no better; and that, by and large, is exactly how these issues have been addressed through most of American history. In recent years, I am happy to report, there has been a flowering of truly interdisciplinary work, involving both serious mathematics and conscientious attention to political and legal realities, and we have begun to move toward a way of thinking about legislative districting which is sound from all the relevant points of view.

2. Measures of fairness

What do we mean when we say a districting is "fair" to the residents of a state? A good way to get a sense of the difficulties here is to contemplate a toy example. Imagine a state with a population of just 100 people, of whom 60 are members of the Purple Party and 40 vote for the Orange Party. The population of this state is partitioned into five legislative districts. Here are four ways the task could be done:

$\mathbf{\Pi}_1$:

Purple	Orange
15	5
15	5
15	5
7	13
8	12

$\mathbf{\Pi}_2$:

Purple	Orange
9	11
9	11
9	11
17	3
16	4

$\mathbf{\Pi}_3$:

Purple	Orange
14	6
14	6
13	7
11	9
8	12

$\mathbf{\Pi}_4$:

Purple	Orange
12	8
12	8
12	8
12	8
12	8

Each of these districtings obeys the constraint that districts be of equal size. But the legislatures they produce are very different. In $\mathbf{\Pi}_1$, the Purple Party holds three seats and the Orange Party two. In $\mathbf{\Pi}_2$, the Orange Party holds a legislative majority, with three out of the five seats. In $\mathbf{\Pi}_3$, Purple holds a 4-1 majority of seats. And in $\mathbf{\Pi}_4$, Purple holds all five seats and Orange is utterly shut out.

Which of these choices is the most fair? Which is the least?

With this toy case in mind, let's talk about the main existing flavors of quantitative measures of fairness.

2.1. **Proportional representation.** One of the most broadly popular and intuitively appealing measures of districting fairness is provided by the principle of *proportional representation*.

Definition 2.1. A districting satisfies *proportional representation* when the proportion of seats held by each party is equal to the proportion of votes won by that party.

Of the districtings above, only $\mathbf{\Pi}_1$ satisfies proportional representation; the Purple Party got 60 percent of the vote, and it holds 60 percent of the seats. Achieving proportional representation is often seen as a goal, or even *the* goal, of districting reform. The New York Times, in a 2018 feature story on gerrymandering, wrote of Pennsylvania's Congressional districts: "Republicans got 54 percent of U.S. House votes statewide, but won 13 of 18 seats," suggesting that this deviation of seat proportion from vote proportion is the problem districting reform is meant to solve. Supreme Court Justice Neal Gorsuch, in the oral arguments on the redistricting case *Rucho v. Common Cause*, also took this to be the standard at issue, pointedly asking, "[A]ren't we just back in the business of deciding what degree of tolerance we're willing to put up with from proportional representation?"

I hope it is clear at the very outset that this definition suffers from many practical problems. For one thing, it is impossible to satisfy exactly; the proportion of votes cast for a party need not be anywhere near a rational number whose denominator is the number of districts! This is most notable in states consisting of a single Congressional district; we accept, without hesitation, that whoever gets the most votes should occupy 100% of the seats, even though the proportion of votes that candidate received may be far from 1.

Would fairly drawn maps even be likely to yield proportional representation? It's unlikely. Look at the Wyoming State Senate, for instance. Wyoming is by

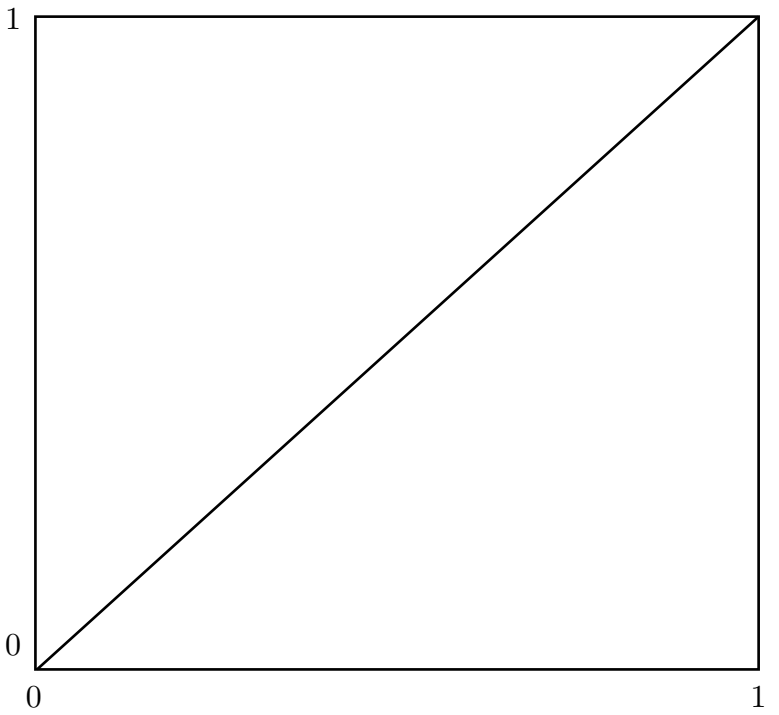
some measures the most strongly Republican state in America. Two-thirds of its voters picked Donald Trump in 2016, and the same proportion voted Republican in the governor's race in 2018. But the state senate isn't two-thirds Republican; there are 27 GOP Senators and only 3 Democrats. That shouldn't necessarily be seen as unfair! When a state's population is two-thirds Republican, the likelihood is that most geographic segments of the state are pretty Republican. In the extreme case of this, where the state is utterly homogeneous politically, *every* district would be represented by Republican Senator. This is the situation depicted in $\mathbf{\Pi}_4$. By the central limit theorem, these are the kinds of districts we'd get if we selected the districting entirely at random from the set of all possible partitions of the state's population into equal-cardinality pieces, with no attention paid to geography. Real-life states, even Wyoming, are *not* exactly homogeneous; but they also often don't look like $\mathbf{\Pi}_1$, in which there's not a single district that approximates the overall political distribution of the state. When districts are drawn geographically, an intermediate scenario like $\mathbf{\Pi}_3$ is more common; substantial variation from the statewide average, but with some concentration around that average.

The final problem with asking districtings to approximate a proportional representation system is that it's not *our* system. We have chosen to accept, for instance, that parties with small but geographically dispersed support don't get representation in the legislature. The proportion of Americans voting for Libertarian candidates for the House of Representatives consistently hovers around 1%; but there has never been a representative from that party, let alone the $3-5$ that strict proportional representation would recommend. (In Canada, whose elections are very similar to those in the U.S., the deviations are even starker; in the 2019 federal elections there, the New Democratic Party drew 16% of the vote against only 8% for the Bloc Québécois, but the Bloc, whose voters are concentrated in a single province, won substantially more seats in Parliament.)

This is not a matter of mathematical or purely philosophical fairness; it's a decision the United States made a long time ago, baked into the way our legal system views elections. No matter how many party-line voters there are in practice, our votes are formally for people, not parties.

2.2. **Partisan symmetry.** One visually effective way to think about measures of fairness of a districting $\mathbf{\Pi}$ is the *seats-votes curve*. This is just what it says on the box; the locus $\{(x,y)\} \in [0,1]\times[0,1]$ consisting of points where x is the proportion of overall votes going to a party and y is the proportion of legislative seats that party wins. There is a separate seats-votes curve for each party; in the present US-centric discussion, we are going to stick to cases where only two parties compete (sorry, Libertarians!) in which case the seats-votes curve for one party is the image of the seats-votes curve for the other by the transformation $(x,y)\mapsto(1-x,1-y)$.

Proportional representation is the requirement that the curve is just the line $x=y$



or maybe, the number of seats typically being pretty small, a step function approximating $x = y$;

I've hidden something important here! There are only finitely many elections held under a given district map, and so this curve is something we have only finitely many points on; what's more, it's certainly possible for two different elections to yield the same vote share but different seat shares, depending on the distribution of votes; what's *still* more, elections held in different years may be held under the same geographic district maps but don't represent *exactly* the same districting, some voters inevitably having moved out of the district, into the district, off this mortal coil, etc. So the seats-votes curve is probably best thought of as a cloud of points around an ideal curve, and we may use as a criterion of fairness that the ideal curve has certain properties. That it be $x = y$ on the nose is, we have argued, too much to ask.

Partisan symmetry is a much more modest request:

Definition 2.2. A districting satisfies *partisan symmetry* if the seats-votes curve is invariant under the symmetry $(x, y) \mapsto (1 - x, 1 - y)$.

This criterion seems very natural: if the Purple Party gets 4 seats with 60% of the vote, then the Orange Party should get 4 seats if *it* gets 60% of the vote. In particular, under the partisan symmetry constraint, a party that gets exactly half the votes gets exactly half the seats.

One challenge for this notion is that it asks us to test whether a curve satisfies a symmetry criterion when we have access only to a set of points on the curve (or, really, a set of points *near* the curve.) If we want to test whether the curve is $x = y$, that's no problem; we can use the difference between the measured x and the measured y as our measure on unfairness. In order to test symmetry, we would

have to probe the seats-votes curve further, inferring something about the results of elections that might have happened, but didn't. As a simple rule of thumb, for instance, we might imagine that partisan swings are roughly uniform across districts. So in $\mathbf{\Pi}_1$, if we moved the overall voteshare to $50-50$, we would similarly shift 2 votes from Purple to Orange in each of the five districts; then Purple still wins three seats, by the narrower margin of 13 to 7, and Orange still wins two, now in $14-6$ and $15-5$ blowouts. This means partisan symmetry has been violated, since Purple still holds a majority of seats while getting only half the votes. More generally, the approximate seat curve for $\mathbf{\Pi}_1$ would be given by the step function

$$y = 3\sigma(x+0.15) + \sigma(x-0.25) + \sigma(x-0.2)$$

where $\sigma(x)$ is 1 for $x > 0.5$ and 0 for $x \leq 0.5$. The reader can check that this seats-votes curve is not symmetric; indeed, its image under $(x, y) \mapsto (1-x, 1-y)$ is the seats-votes curve for $\mathbf{\Pi}_2$. The only one of the four districtings we showed which satisfies partisan symmetry in this sense is $\mathbf{\Pi}_4$.

Another criticism: partisan symmetry may reflect factors other than self-interested malfeasance. The districting $\mathbf{\Pi}_2$ awards a majority of seats to the Orange Party, even as they get thumped by the Purples in the popular vote. But what if the Purples of the state are packed into a couple of dark-Purple metro areas, set against the background of a countryside that leans orange? Isn't it possible you'd see results a lot like this, without any self-dealing? Is "organic partisan asymmetry" like this actually unfair? If we ask the state to vote on a ballot referendum, people who feel strongly about the issue don't get more votes than people who barely care. Some would apply the same reasoning to geographic regions: each patch of land gets one vote in the legislature, even though some patches may be strong supporters of one party and others more ambivalent.

2.3. **Efficiency gap.** In the last decade, law professor Nicholas Stephanopoulos and political scientist Eric McGhee introduced and popularized a new metric for unfairness, called the *efficiency gap*. [22] To see what motivates their definition, look back at our four example districtings. What makes $\mathbf{\Pi}_2$ such a good choice for Orange? It's that Orange voters are deployed with exquisite strategic precision, to exactly the districts where they're needed to ensure a narrow victory. Purple voters, by contrast, are in exactly the *wrong* places; almost half of them reside in the three districts where Purple loses narrowly, and thus contribute nothing to Purple's representation in the statehouse. One might say their votes were wasted. This leads us to a definition.

Definition 2.3. Suppose the candidates in a two-party election receive A and B votes, respectively, with $A \geq B$. Then the number of wasted votes for the losing candidate is B, and the number of wasted votes for the winning candidate is $A - (1/2)(A+B)$.[1]

This captures the notion that a vote is wasted just insofar as it fails to contribute to a candidate's victory. A districting drawn to favor one party does so by causing the other party to waste votes. That motivates the definition of efficiency gap.

[1]We are not going to worry in this space about the difference between half the votes and half the votes plus one.

Definition 2.4. Suppose the total number of wasted votes for the two parties in an election is w_1 and w_2 respectively, and the total number of votes cast in the election is N. Then the *efficiency gap* is $(1/N)(w_1 - w_2)$.

For example, in $\mathbf{\Pi}_1$, the Purple party wastes 5 votes in each of the three seats they win, and 7 and 8 votes respectively in the two seats they lose, for a total of 30. The Orange party, by contrast, also wastes 5 votes in the first three districts, but only 3 and 2 in the two districts where they win, totaling to 20. So the efficiency gap here is $10/100 = 0.1$, favoring the Orange party. In $\mathbf{\Pi}_2$, the efficiency gap is much larger, 0.3 favoring Orange. In $\mathbf{\Pi}_3$, the efficiency gap is 0.1 in favor of Purple, and in $\mathbf{\Pi}_4$ the efficiency gap is 0.3 in Purple's direction. This measure rates $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_3$ as the fairest districtings; so do most people who look at those four tables, which is a point in efficiency gap's favor.

Another bonus of efficiency gap is that, unlike partisan symmetry, it doesn't require any imputation of election results under conditions other than the real ones; it computes its measure directly from the election results that have already happened. Efficiency gap has this in common with proportional representation, and indeed the two measures are very similar in spirit; both specify exactly what they want the seats-votes curves to be.

Proposition 2.5. A two-party election (with all districts of equal size) in which one party receives proportion x of the votes and proportion y of the seats has efficiency gap $2(x - 1/2) - (y - 1/2)$. In particular, the efficiency cap is zero when $y = 2x - 1/2$.

The latter statement means efficiency gap can be thought of as demanding adherence to the seats-votes curve $y = 2x - 1/2$, instead of the curve $y = x$ required by proportional representation. In other words, those two criteria are not only different, they are incompatible!

Proof. For each district π_i, let x_i be the proportion of votes won by the first party. Let y_i be $1/2$ if the first party wins and $-1/2$ if the second party wins. So the average of y_i over all districts is $y - 1/2$, and the average of x_i over all districts is x. The proportion of votes in π_i which are wasted votes for the first party is then $x_i - (1/2)(y_i + 1/2)$, and the proportion of votes which are wasted votes for the second party is $(1 - x_i) - (1/2)(-y_i + 1/2)$, so the difference – the quantity whose average over all districts is the efficiency gap – is $2x_i - y_i - 1$. Thus the efficiency gap is $2x - 1 - (y - 1/2)$, as claimed. □

The efficiency gap was a huge step forward for attempts to bring mathematical reasoning to bear on the legal problem of districting. Prior to the efficiency gap, courts had typically declined to intervene in partisan gerrymandering cases, following the Supreme Court ruling in *Vieth v. Jubelirer* that there was no sufficiently clear standard for distinguishing a gerrymandered map from a fair one. The efficiency gap filled that, well, gap. It is easy to compute, it doesn't rely on any hypotheticals, it's clearly distinct from proportional representation, and it conforms well with our native intuition about what makes a districting unfair. It formed the centerpiece of *Whitford v. Gill*, the case against the state legislative districts drawn in Wisconsin after the 2010 census. The new maps created a sharp rise in efficiency gap which persisted in several consecutive elections following redistricting. That evidence (together with ample documentation that the map was drawn with the

specific intent to help the Republican party) convinced a three-judge federal panel to throw out those districts.

But the efficiency gap isn't the end of the story. For one thing, it is surely too strict. Like the proportional representation standard, it asks for adherence to a specific seats-votes curve. Do we really think one curve suits all states at all times? The curve $2(x-1/2) = (y-1/2)$ also has the problem that it's literally impossible for the outcome to match that curve if one party gets more than 75% of the vote (though admittedly this scenario is extremely rare, even in states like Wyoming where one party is much more popular than the other.) All these difficulties are, of course, superable. Legal arguments based on the efficiency gap don't propose a strict standard where a map with efficiency gap exceeding some threshhold are automatically declared unconstitutional; rather, a high value of the efficiency gap is to be used as a "red flag" providing evidence, but not dispositive evidence, that a map has been gerrymandered.

Another vulnerability of the efficiency gap standard arises directly from one of its strengths. The efficiency gap depends on real outcomes, not hypotheticals. But that makes the efficiency gap *discontinuous*. At its heart is the function y_i in Proposition 2.5, which can jump from $-1/2$ to $1/2$ with a tiny change in vote-share. A districting that fails an efficiency gap test might easily pass if a few close races switched their outcome. That feels wrong. (Though of course there are workarounds; one might, for instance, report the distribution of efficiency gap on a small ball around the actual outcome rather than relying on a single point)

The efficiency gap measure has another problem in the U.S. context. American courts have generally not been sympathetic to the idea that political parties with substantial support have a constitutionally guaranteed right to legislative representation. Claims made by individual voters that they've suffered harm to their ability to vote, or their first amendment right to express their politics, have been more successful. So a suit filed against a district map has to argue that individual voters have had their rights removed, or at least meaningfully shaved down. Which voters are these? It can't just be someone whose vote was "wasted," in the efficiency gap sense; after all, in every single district, half the votes cast are wasted, whether gerrymandering takes place or not! The design of the efficiency gap measure is purely global; it doesn't tell you, because it wasn't designed to tell you, which districts are the ones that were maliciously modified to help one party. A detailed critique of the efficiency gap from a mathematical perspective can be found in [1]; for further description and several more refined measurements in the spirit of the efficiency gap, see [23].

2.4. **Ensemble sampling.** I have discussed a lot of simple answers to a complex question, each of which has real merits, each of which is in some ways lacking. Now we turn to the part where deeper mathematics comes into play; I want to discuss the method of *ensemble sampling*, which in my view is the current state of the art for measuring gerrymandering.

Let's return to our basic question. What does a district map that *isn't* gerrymandered look like? Does it look more like $\mathbf{\Pi}_1, \mathbf{\Pi}_2, \mathbf{\Pi}_3$, or $\mathbf{\Pi}_4$? We have seen that attempts to assign a numerical unfairness score to a map based on a single election outcome all have problems, and not all agree about the relative fairness of the four districtings in our toy example.

And why *should* they agree? After all, fairness is not purely a matter of the numbers on the spreadsheet; some actual knowledge of the political landscape to be partitioned is required. If partisanship in a state is homogeneously distributed, the same in the east as in the west, in the north and in the south, then any geographically based map is going to look like $\mathbf{\Pi}_4$, with all seats having roughly the same partisan distribution as the whole state. In that case, one party will hold all or almost all the seats, even if their statewide share of the vote is only a modest minority. In a state like that, a seat distribution closer to proportional representation, or to efficiency gap 0 would be strong evidence *for* gerrymandering, not for its absence. Likewise, in a state where one party's support was concentrated in a geographic region, a districting like $\mathbf{\Pi}_4$ would be ironclad proof someone had their thumb pressed firmly on the scale.

The philosophy behind ensemble sampling is a simple one: the opposite of gerrymandering isn't proportional representation or adherence to any other strict numerical standard; the opposite of gerrymandering is not gerrymandering. If we want to know whether a map is fair, the right question is

> Does this district map tend to produce outcomes similar to a map that *would have been drawn* by an authority who wasn't aiming to privilege one party's interests over another?

That neatly solves some of the problems with the measures described above, but at the expense of introducing a new problem, which is now not really legal or philosophical but inferential: how can we assess what would have happened if the maps had been drawn without prejudice? The idea of studying gerrymandering through this lens was first popularized in an influential 2013 paper by the political scientists Jowei Chen and Jonathan Rodden [5]. They were troubled by the issues above, especially the phenomenon mentioned at the end of section 2.2: when Democrats are predominant in cities and present in moderate numbers throughout the state, while Republicans are concentrated in more rural districts and almost entirely absent from more densely populated areas, partisan symmetry can fail even when districts are drawn indifferently to partisan advantage. How do we distinguish an asymmetric districting like $\mathbf{\Pi}_2$ that arises from gerrymandering from one that reflects the disinclination of Oranges to live in Purpleopolis?[2]

Chen and Rodden write:

> To what extent is observed pro-Republican electoral bias a function of human geography rather than intentional gerrymandering? To what extent might pro-Republican bias persist in the absence of partisan and racial gerrymandering?
>
> The main contribution of this paper is to answer these questions by generating a large number of hypothetical alternative districting plans that are blind as to party and race, relying only on criteria of geographic contiguity and compactness. We achieve this through a series of automated districting simulations. The simulation results provide a useful benchmark against which to contrast observed districting plans.

[2]Proponents of proportional representation should pipe up here and ask, why do we need to make this distinction? Isn't partisan bias unfair whether or not it's enacted on purpose? Maybe so – but the legal and political barriers to moving away from one-geographic-district-one-seat are a lot higher than anything else described here.

Where do the "hypothetical alternative districting plans" come from? That's where ensemble sampling comes in. First of all, let's rephrase the basic fairness question along the lines of the Chen-Rodden approach:

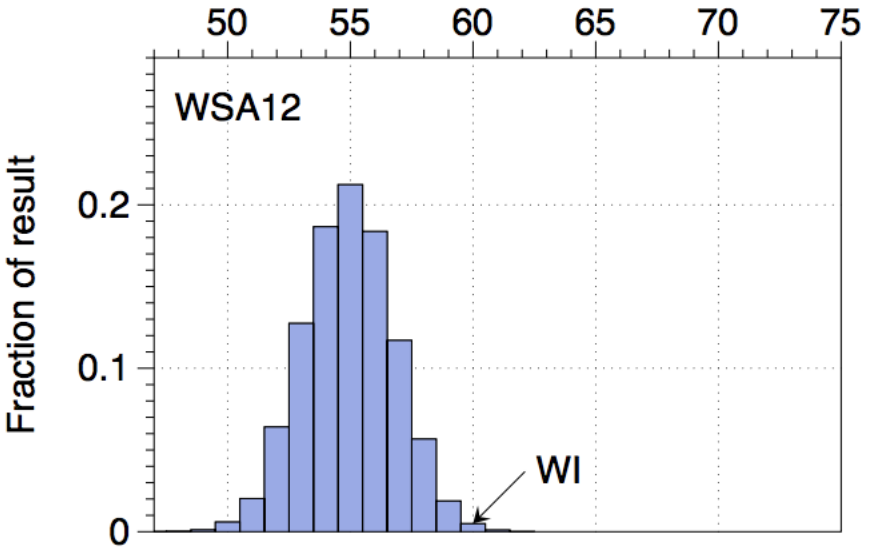
> Does this district map tend to produce outcomes similar to a map *randomly selected* from the set of all possible maps?

This suits our intuition; one might imagine, as a first approximation, that a map-drawer indifferent to who wins would consider any partition of the population into equally-sized geographic chunks to be as good as any other. Or one might very reasonably *not* assume that. Each state has its own constitutional and other legal constraints on what districts can look like – for example, in most states they need to be connected – and overlaid on these is the federal Voting Rights Act, which guarantees that among the Congressional districts there are, where possible, majority-minority "districts of opportunity."

It gets more complicated still. Mathematicians often think of the law as consisting of a series of hard and fast rules, like axioms, from which outcomes can be drawn. Law isn't really like that. Law probably *couldn't* be like that. In the case of districting, one quickly finds that much of the relevant law is not so much a set of constraints as a collection of preferences. In Wisconsin, for instance, legislative districts are not supposed to cross county lines, except that decades of precedent have established that they *can* cross county lines when other legal requirements make it necessary, but it's better not to do it too much. We also have requirement that state legislative districts be "as compact as practicable." What does that mean? It is certainly not a strict numerical bound on the pair (perimeter, area) in $\mathbb{R}^2$. We can capture all this by refining our question one more time:

> Does this district map tend to produce outcomes similar to a map *randomly selected* from the set of all legally permissible maps, on which we place a probability distribution that reflects this state's legal preferences between districts?

The standard is then that unfair maps are ones which are *extreme outliers* in that distribution – that is, those that look like this:



This diagram, from a 2017 paper of Herschlag, Ravier, and Mattingly [15], depicts the results of a simulation of the 2012 Wisconsin State Assembly election, under 19,184 alternative districting plans sampled from their ensemble. The outcomes form a reasonably normalish distribution centered on a modal outcome of 55 Republican seats out of 99. In our world, under the maps drawn by the Republican majority in the state legislature.after the 2010 census, the Republicans won 60 seats. We note in passing that Herschlag-Ravier-Mattingly's finding matches that of Chen and Rodden concerning "unintentional gerrymandering"; Wisconsin is a state where two large urbanized areas (Madison and Milwaukee) strongly favor Democrats and only one (Waukesha County) strongly favors Republicans, and indeed, in the 2012 election where the popular vote for Assembly candidates was very close to evenly split between Democrats and Republicans, a typical map drawn without prejudice gives Republicans a modest majority of seats, though much less than the majority supplied by the gerrymandered map in actual use.

As you've probably noticed, there's a major methodological issue we've been keeping silent about. *How* do we sample from the set of all permissible maps? This is no small question, and it's the hardest mathematical part of the problem. Wisconsin has $6,672$ voting wards. The number of ways to partition those wards into 99 Assembly districts is big – really big, massively uncomputably big. It is 99^{6672}, if you ignore all constraints on what makes an assembly map permissible.

Now just because a set is big doesn't mean you can't uniformly sample from it. It's easy to choose an integer uniformly at random from the interval $[0, 10^{100}]$. Or to choose a spanning tree from a graph on 1000 vertices; you can do this efficiently via Wilson's algorithm. Closer to the problem at hand, the set of *all* partitions of Wisconsin's wards very easy to sample uniformly – just hand each ward an integer uniformly and independently chosen from $\{1, \dots, 99\}$.

But that's not the sample we want. First of all, we need the districts to be of approximately equal size. But more than that, we need the districts to be *contiguous*, not built out of wards scattered all over the state from Turtle Lake

to Oconomowoc. We can think of the wards as forming a *weighted planar graph*, where each vertices is a ward and two wards are connected if they border each other. Then a collection of wards forms a contiguous district just when the set of vertices corresponding to that collection induces a connected subgraph. If Γ is a graph, we define a *connected k-partition* to be a partition of the vertices into subsets $V_1, \ldots, V_k$ such that the induced subgraph on each V_i is connected.

Now we're faced with a question in the theory of algorithms: given a weighted planar graph on N nodes, how do you randomly sample from the set of connected k-partitions whose constituents have roughly equal total weight? This is not so easy. Even when $k = 2$, this problem is NP-hard [20].

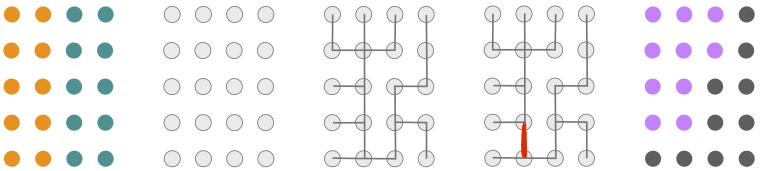
But lots of NP-hard problems have tractable approximations. If we want to sample from an unknown distribution on the vertices of a very large graph, one can often do very well in practice by means of a random walk on the graph. (In pure math, a well-known example of this technique is the *product replacement algorithm* for choosing a uniform random element of a large finite group; see [14].) The idea of using this kind of Markov chain in the context of redistricting originated with Mattingly-Vaughn [17] and Fifield et al [13]. It has become the dominant method of ensemble sampling among mathematicians working in this area.

The graph we're sampling from is not the graph Γ of wards described above, but a vastly huger one: we consider a graph P whose vertices are all connected k-partitions of Γ – or, better, all connected k-partitions corresponding to districtings compliant with state and federal law. I've told you the vertices; what are the edges? Here's where things get really interesting. There are different choices, and the corresponding random walks may certainly have different properties. The simplest and in some ways most natural case is that of the "flip graph," in which two vertices are adjacent in P just when the corresponding partitions differ with respect to only one vertex of Γ. That is a gigantic graph whose structure we know next to nothing about. But given a vertex $\mathbf{\Pi}$ of this graph, we *can* generate a list of its neighbors; each vertex adjacent to $\mathbf{\Pi}$ is obtained by a "flip": take a ward in district i which lies on the boundary with district j, and reassign it to district j, as long as this violates no legal constraint. It's easy to list and sample from all possible flips; so we can efficiently carry out a random walk on P, and thereby generate a large population of legally acceptable districtings. If you want to bias your walk towards districtings whose districts have more compact shapes, or shatter fewer counties, or whatever, you can weight your choice of moves in the random walk to promote those virtues. In other words: the distribution we want to sample from *isn't* the uniform distribution, because we prefer some maps to others, so we set up our walk to get the stationary distribution we desire. This is the walk we see used in the work of Herschlag-Ravier-Mattingly [15] on the Wisconsin Assembly districts. We can then identify the problematic maps as those whose electoral outcomes are extreme outliers in the sample distribution.

Another graph structure you can put on P is the *recombination* graph [9]. Suppose $\mathbf{\Pi}$ is a districting. Then a neighbor of $\mathbf{\Pi}$ is obtained as follows. Choose two adjacent districts. Merge them into one. Then split the new double-sized district into two connected pieces of roughly equal size. That's a ReCom move, and two districtings are adjacent if a ReCom move sends one to the other.

But didn't I *just say* the problem of sampling uniformly from the set of 2-partitions is hard? And isn't that just what choosing a ReCom move requires

A step in the ReCom process (Figure 4 in [9]).



us to do? Well, not quite. In the ReCom process, the splitting of the double-sized district D is *not* uniform over the set of all 2-partitions. Instead, we proceed as follows. Choose a spanning tree T uniformly from the set of all spanning trees of D; *this* we can do uniformly in polynomial time, by Wilson's algorithm referenced above.[3] Now choose an edge at random, subject to the constraint that deleting the edge splits the vertices of T into two roughly equal parts T_1, T_2; then the vertices of T_1 and T_2 constitute a partition of the vertices of D into two subsets. Now the restriction of P to the vertices of T_i for each i contains a tree, and so is connected; we have constructed the desired 2-partition of D.[4].

(I can't resist pointing out in passing that the set of spanning trees of a planar graph has an extremely interesting structure. By Kirchhoff's matrix-tree theorem, the number of spanning trees is the same as the order of the finite abelian group alternately called the sandpile group, the Pic^0 of the graph, or the tropical Jacobian. When the graph is planar, a beautiful categorification of this numerical identity holds: the set of spanning trees is canonically a *torsor* for the sandpile group. (See [4],[2].) Is there any chance that the tropical viewpoint on graphs is useful for computation?)

The non-uniformity of the choice of the 2-partition might at first seem artificial, but in fact it's more feature than bug! A uniformly chosen 2-partition typically has a long, snaky boundary, a sort of space-fillling curve. The partitions coming from the spanning tree method, by contrast, are strongly biased towards having short boundaries; in graph-theoretic terms, there tend to be few edges of D joining vertices of T_1 to vertices of T_2. For the flip walk, we need to builld a preference for compact districts into the weights of the random walk if we want to get decent-looking districts; ReCom, by contrast, tends to form compact districts without any extra infrastructure. What's more, the ReCom walk appears to converge to a stationary distribution more efficiently than does the flip walk, although it is more difficult to describe in explicit terms what the stationary distribution is. Gaining a better understanding of the stationary distributions attached to various random walks on k-partitions is one of the richest open problems in the subject, and is of interest not only in practical terms but as a question in pure stochastic processes.

Yet another appealing feature of the ReCom walk is that software to do it is open-source and publicly available [18]; I encourage you to mess with it yourself!

[3]The Kruskal-Karger algorithm is just as good in practice, and is sometimes used in ReCom implementations.

[4]There are variants of ReCom which choose the 2-partition in a different way, but for simplicity we'll stick to spanning tree recombination in these notes

Ensemble sampling has proven to provide a more effective, convincing, and intuitive quantitative measure of gerrymandering than those before it. It was the centerpiece of the gerrymandering cases presented to the Supreme Court in 2019; a bipartisan twin set, one (*Rucho v. Common Cause*) addressing a district map gerrymandered by Republicans in North Carolina, the other (*Lamone v. Benisek*) a map gerrymandered by Democrats in Maryland. I don't know if this is the most mathematical case the Supreme Court has ever addressed, but I believe it is the first time the Court received a "Mathematicians' Brief," an amicus brief signed by eleven mathematical scientists, including me, explaining the quantitative aspects of the case.

In *Rucho* and *Lamone*, I hurry to point out, the Supreme Court ruled that partisan gerrymandering was not justiciable; that is, it was not a matter where the federal courts had the right to intervene. To anyone who had been following the mathematical study of redistricting, the court's decision was puzzling. Justice Roberts, writing for the majority, says:

> Partisan gerrymandering claims invariably sound in[5] a desire for proportional representation. As Justice O'Connor put it, such claims are based on a conviction that the greater the departure from proportionality, the more suspect an apportionment plan becomes.

I am not weighing in on the legal merits of the decision when I say that this claim is wildly off the mark. As we have seen, proportional representation is not the aim of any modern measure of electoral fairness, or any claim brought before the court in the 2019 districting cases (It may indeed have been an issue 33 years ago, when Justice O'Connor wrote the words Roberts quotes in her concurrence in *Davis v. Bandemer*) Indeed, one measure arising in contemporary cases, the efficency gap, is *incompatible* with proportional representation, while the ensemble sampling methods (correctly, in my view) are orthogonal to proportional representation. In some case a map that yields proportional-representation outcomes would be licensed by those methods; in other cases, like U.S. congressional districts Massachusetts, a map yielding proportional representation would be flagged as an extreme outlier [11].

To be clear, the Supreme Court did not take seriously any claims that gerrymandering had not taken place, and indeed the majority decision endorsed the view that the practice of gerrymandering is "incompatible with democratic principles." The Court's decision rests on the fact that some things are incompatible with democratic principles but not incompatible with the Constitution. When this is held to be the case, the federal judiciary walks on by with eyes modestly averted. But the Court did portray partisan gerrymandering as a problem somebody – just not they themselves – ought to remedy.

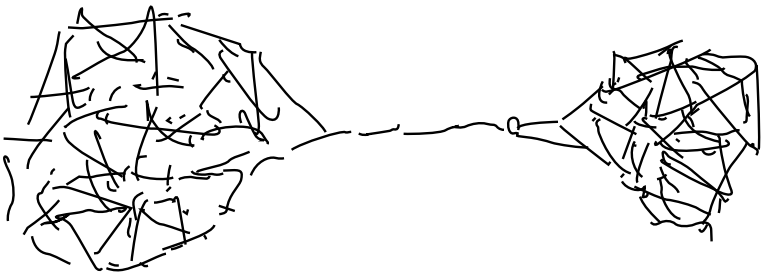
And indeed, the arguments derived from ensemble sampling have found more purchase elsewhere. In 2018, the Pennsylvania Supreme Court threw out the state's U.S. Congressional district map, relying heavily on expert testimony from Jowei Chen and from Wesley Pegden, a mathematician at Carnegie Mellon (more on Pegden's work just down the page!) identifying the map as an extreme partisan

[5]As best I can tell from my lawyer friends, "sound in" here means something in between "derives from" and "amounts to" – and people say *we* talk in impenetrable jargon!

outlier from a random walk ensemble. Pegden was subsequently appointed by Governor Tom Wolf to serve on the Pennsylvania Redistricting Reform Commission, along with the more traditional group of elected officials and community leaders you might expect to find on a political panel like this. Wolf also enlisted Moon Duchin to help evaluate replacement districting plans. In North Carolina, a state court similarly found that state legislative districts drawn by the majority in the state legislature violated both the Free Elections Clause and the Equal Protection Clause of the North Carolina Constitution. Their decision, like the one in Pennsylvania, is deeply rooted in the analysis of ensemble samples, citing testimony of Jonathan Mattingly as well as Pegden and Chen. (To get a sense of what kind of testimony mathematicians provide, you could look at Mattingly's expert report in that case [16].) One key point from a legal perspective is that ensemble methods can make local assessments, not just global ones as the efficiency gap does, identifying *individual districts* as outliers, which helps courts in several ways. It makes it easier identify harm to individual voters; it can provide courts with a remedy that throws out only part of the map instead of the whole thing; and, most importantly, it presents judges with a clearer picture of *how* gerrymandering is accomplishing its goals.

I don't want to leave the impression that ensemble sampling is an infallible gerrymandering detector. Many challenges, some purely mathematical, others intertwining math and law, remain.

A central open question concerns speed of mixing. A random walk on a connected graph converges to the stationary distribution. But how *quickly* it converges is a delicate matter, involving, among other things, the spectral gap of the adjacency matrix of the graph. For the graph P of k-partitions, we have no control over mixing time, and so no guarantee that the ensembles obtained by random walk are drawn from the stationary distribution. Indeed, if we impose on P the condition that the k components of the partition have roughly the same size, as far as I know there is no proof that P is even connected! Even if P is connected, it might for instance have a "long neck," like so:



If our random walk starts on the right-hand side of the graph, crossing the neck is a very low-probability event; so even if we run the random walk for a long time, we may be approximating a distribution which is far from the stationary one, but rather approximates a distribution supported on the right-hand side of the neck.

From a political and legal perspective, it's not clear this matters. Suppose a district map $\mathbf{\Pi}$ tends to give the Purple Party 60 out of 100 seats, and suppose this figure is an outlier in a sample of maps near $\mathbf{\Pi}$ in P, 99.9% of which give the Purple Party at most 55 seats. One cannot strictly rule out the existence of an "undiscovered country" of districtings in which 60 Purple seats are the norm. But

the ensemble sample still feels like very strong evidence that this map $\mathbf{\Pi}$, so unusual among its neighbors, was not picked out indifferently to its Purple-friendliness. This insight was formalized and made into a theorem by Chikina, Frieze, and Pegden ([6], see also [7]) who obtain a provable threshold for statistical significance without any need for a bound on mixing time. We'll state their result in the language of random walks on graphs, though it actually holds for any reversible Markov chain.[6] They prove: for any small $\epsilon > 0$, and any real-valued score function ω on the vertices of a graph G, the probability that a vertex v_0 of P chosen uniformly (the null hypothesis) has a higher score than $1 - \epsilon$ of the first k vertices in a random walk starting from v_0 is at most $\sqrt{2\epsilon}$.

Nonetheless, the question of what kind of mixing time one expects for random walks of various flavors on sets of k-partitions of planar graphs is a really interesting one. In practice, as we have mentioned, the ReCon walk seems to converge substantiallly more efficiently than does the flip walk. Why? And are there still more effiicient graph structures on P out there to be exploited? The study of mixing on these graphs has the flavor of statistical physics and the theory of self-avoiding random walks; see the recent paper of Najt, DeFord, and Solomon [20] for this connection, together with questions about the way different discretizations of the same planar landscape can yield surprisingly large differences in the behavior of the corresponding random walks.

Another challenge of ensemble sampling is that different elections are different. One certainly doesn't want to say that an individual voter is, once and for all, a Democrat or a Republican whose voting behavior is independent of time and the individual candidate before them. If we did say this, empirical data would contradict us. In practice, what this means is that a given map may be an extreme outlier in some elections and not in others. Take, for instance, the map of Wisconsin state assembly districts. In 2012, a year when the statewide vote for Assembly seats in Wisconsin was very close to 50-50, the result of 60 seats for Republicans was very different from the ensemble-modal value of 55 seats. But two years later, in 2014, the electorate leaned much more towards Republicans; and in this election, the 63 seats won by Republicans sit comfortably in the middle of the range of outcomes produced by the ensemble.

So is the Wisconsin district map an extreme outlying gerrymander, or is it not? The work of Herschlag, Ravier, and Mattingly [15] provides critical insight here. A district map, remember, is made without knowledge of exactly who is going to vote, or how. The dark art of gerrymandering has to be robust to this basic uncertainty. And there are tradeoffs: maximizing the extent to which the map helps your party under one set of circumstances may make the map less effective under other conditions, or even hurt your party if things go really sideways. Herschlag-Ravier-Mattingly find that the Wisconsin map is designed as a sort of "firewall." In electoral environments where the statewide vote leans Republican, the gerrymander doesn't do much work. But when the electorate is split evenly or even leaning slightly Democratic, it provides a powerful force towards maintaining a Republican majority in the assembly. The gerrymander works exactly when, at least according to the desires of its makers, it needs to work, locking in a Republican seat majority

[6]The flip walk, in its most commonly used form, is reversible, but the ReCom walk; see [3] for a reversible modification of ReCom.

over the whole range of statewide electoral conditions one might reasonably expect to encounter in an evenly split state like Wisconsin.

3. What's next?

The study of random walks applied to districting has developed very rapidly in the last five years, and provides opporunities for an extraordinarily direct interaction between advanced mathematics and public policy. But the story is far from over, and there's a lot of work still to do for mathematicians interested in these problems.

Part of the work is outreach. Moon Duchin and Jonathan Mattingly have both launched centers – respectively, the Metric Geometry and Gerrymandering Group at Tufts, and the Quantifying Gerrymandering group at Duke – which serve as clearinghouses for new research on districting and launchpads for both early-career and senior mathematical scientists interesting in getting involved.

There is also much more mathematics to do, besides the rich vein of questions indicated above about mixing times for Markov process on the k-partitions of planar graphs. One key question, which has been raised a lot but so far has not been extensively addressed, is: what to do about situations where there are more than two parties? Suppose I don't care for either Purple nor Orange, preferring the Ecru Party to either one. But perhaps, if Ecru is not an option, I like Orange better than Purple. In a first-past-the-post system like that in the U.S., my voting behavior may depend on the district I'm drawn into. If Ecru has no chance and Orange and Purple are close, I may vote Orange to stop Purple. But if Purple is safely ahead in my district, I'm more likely to vote my Ecru conscience. *Everything* about voting behavior is much harder to analyze when more than two parties have substantial support; the problem of districting is no exception. In a U.S. context, where parties other than Democrats and Republicans are very rarely competitive for legislative seats, one might this question can be ignored. But the U.S. is not the only country with geographic districts. What's more, many American jurisdictions, including the entire state of Maine, have abandoned first-past-the-post in favor of ranked-choice voting. To the extent RCV becomes mainstream, there will likely be more votes for candidates other than Democrats and Republicans; it seems prudent to have the mathematical machinery for analyzing districtings ready in advance.

What is the future for this interaction between mathematics and politics? The Supreme Court has, for now, put an end to the idea that political parties nationwide will be forbidden from extreme gerrymandering. In some states, like Pennsylvania and North Carolina, courts will throw out existing gerrymandered maps; in others, like Ohio and Michigan, legislation or ballot initiatives will delegate the process of redistricting to non-partisan panels. In both cases, though, there are still fundamental design questions: if the process of drawing the map isn't to be "party operatives in a smoke-filled room tweak and twist the districting until it delivers every possible advantage to their party," what *is* the right process?

One question that comes to many mathematicians' minds at this point is: if we can generate an ensemble of thousands and thousands of potential district maps, which are compliant with the Voting Rights Act and other legal constraints, which perform well on traditional districting criteria like compactness and county-splitting, and which are completely indifferent to which party does better, why don't we . . . just pick one of those maps at random and call it a day?

The reasons are political. For one thing, algorithmic maps will inevitably miss criteria specific to the case at hand that are legitimately important to constituents. As DeFord, Duchin, and Solomon write in their report on districting alternatives in Virginia [19], "We emphasize that these ensemble methods should not be used to select a plan for enactment because they are made without local and community-based considerations. Instead, ensemble methods give an effective means of verifying whether a newly proposed plan is an extreme outlier in the universe of valid plans." Even when used as a benchmark, a process like ensemble sampling is viewed with some suspicion by political actors. They can't see what's under the hood. To actually hand over map-drawing power to the algorithm is something neither elected officials nor their constituents are likely to swallow. To a lesser extent, proposals to delegate the power to an independent nonpartisan commission – say, retired judges or a panel of state residents – meet the same resistance (from legislators, if not from their constituents, who have recently voted for such plans by referendum in Colorado, Michigan, Missouri, and Utah.) Elected officials, as a rule, don't like relinquishing powers the existing framework affords them.

But how can there possibly be a fair protocol for district-drawing if the district-drawing is to be done by the legislature itself? In [21] Pegden, Procaccia, and Yu propose a really interesting idea, deriving from an entirely different area of math: the theory of fair-division games. That theory descends from a single algorithm: "I cut you choose." Two players, who perhaps don't trust or even like each other, want to divide a cake, and each wants to make sure the other doesn't get more than their share. (See the relevance?) Algorithm: one player cuts the cake in two pieces, and the other picks which piece they get. The cutter has an incentve to make the division as close to 50-50 as possible, and the chooser, if the cutter does their job properly, is indifferent to the choice they make. Crucially, the algorithm doesn't *enforce* fairness; but it leaves both parties to the decision feeling like they had a fair chance to affect the outcome.

The Pegden-Procaccia-Yu protocol is called "I cut you freeze." It's a game where Purple and Orange take turns. They start with the district map $\mathbf{\Pi} = \mathbf{\Pi}_1, \dots \mathbf{\Pi}_k$ as it currently exists. At each stage, some subset of the districts is *frozen* – its boundaries are fixed and it can no longer be modified. Each turn has two parts: you freeze one of the not-yet-frozen districts, and then you redistrict the unfrozen part of the map however you like, then you pass the new map over to the other player. The game ends when all the districts are frozen, each party having locked in the final form of half the districts. (This might be better called "I freeze and cut, then you freeze and cut," but I think the authors' choice to contract this for euphony was a wise one.) Note, crucially, that the district a player gets to freeze is one chosen from a map *created by the other party*; if the steps were reversed, so that each party redistricted the unfrozen part of the map first and then froze a district of their own making, it would be substantially easier for parties to create mischievous districts.

There are lots of possible game protocols one could use for redistricting, and the Pegden-Procaccia-Yu paper has inspired several competing proposals. Understanding which of these provides the most effective buffer against extreme gerrymandering while remaining palatable to elected officials and compliant with legal requirements is a rich and fundamentally interdisciplinary question which we've just started to penetrate.

So far, the role of mathematicians in districting has been a form of damage control – we come in to assess the outcome of decisions that have already been made, under a regime of rules already been set in place. The work of Pegden, Procaccia, and Yu is an example of a deeper interaction that's just starting to take shape, namely: which regimes of rules would lead to better outcomes? This is especially important at this time of flux, when the issue of districting is at its highest political salience in years and many states are launching brand-new commissions. Decisions about the design of the districting process and the rules governing it are going to have consequences for decades to come, and it seems like a really good idea for mathematicians to be in the room when those decisions are being made.[7] Considering different rules and figuring out what would follow from them is *kind of our thing.*

Ensemble sampling is well-suited for the task because it enables us to efficiently explore the space of what's possible under various collections of rules. People may once have thought that requiring districts to have equal population ensures fair representation; we now know that's not true. The picture associated with "gerrymandering" for most people is a bizarrely shaped branching snake of a district, drawn to its odd contours to achieve an electoral result; that happens, but it turns out requiring districts to be roughly round [8] is also not sufficient to prevent strong partisan gerrymandering. Ensemble methods have also shown that rules which are facially neutral to party, like maximization of the number of competitive districts, can introduce partisan bias [8]. What's more, the ensemble method speaks to much more than the number of seats each party wins – that's just one statistic attached to the maps in the ensemble. Which rules promote, or suppress, the power of minority voters? Which tend to lead to more competitive districts? What are the tradeoffs between properties of district maps we think of as virtues? The mathematics of redistriciting isn't just a gerrymandering detector; it is, at least potentially, a full-fledged tool for the exploration of the mysterious space of protocols for representative democracy.

Whether the future of districting in the United States is independent commissions put in place by popular ballot initiatives, intricate games of bipartisan cake-cutting, or the status quo of entrenched parties grimly maximizing their own interest, is a political question. But it's a political question shot through with mathematical content, and thanks to the work discussed here, the mathematical community has gotten engaged with this content to an extent rarely seen in American politics. I hope this engagement continues, and I hope some of the audience here will become part of it!

References

1. Bernstein, M. and Duchin, M., "A formula goes to court: Partisan gerrymandering and the efficiency gap." *Notices of the AMS* 64.9: 1020-1024 (2017)
2. Baker, M., and Wang, Y. The Bernardi process and torsor structures on spanning trees. *International Mathematics Research Notices*, 16), 5120-5147. (2018)

[7]Maybe we could have kept Missouri from enshrining in its state constitution the principle that "In general, compact districts are those which are square, rectangular, or hexagonal in shape." It's ironic that Missouri, a trapezoid, would be so hostile to general quadrilaterals.

[8]in whatever sense: see [12] for a suggestion of a discretized notion of compactness which seems better suited for modern applications than the many, many traditional methods in use.

3. Carter, D., Herschlag, G., Hunter, Z., and Mattingly, J. A Merge-Split Proposal for Reversible Monte Carlo Markov Chain Sampling of Redistricting Plans, arXiv:1911.01503. (2019)
4. Chan, M., Church, T., and Grochow, J. A. Rotor-routing and spanning trees on planar graphs. *International Mathematics Research Notices*, 2015(11), 3225-3244. (2015)
5. Chen, J., and Rodden, J.. "Unintentional gerrymandering: Political geography and electoral bias in legislatures." Quarterly Journal of Political Science, 8(3), 239-269. (2013)
6. Chikina, M. and Frieze A., and Pegden, W. "Assessing significance in a Markov chain without mixing." *Proceedings of the National Academy of Sciences* 114.11: 2860-2864 (2017).
7. Chikina, M., Frieze, A., Mattingly, J., and Pegden, W. "Practical tests for significance in Markov Chains." arXiv:1904.04052. (2019)
8. DeFord, D., Duchin, M., and Solomon, J., "A Computational Approach to Measuring Vote Elasticity and Competitiveness," (2019)
9. DeFord, D., Duchin, M., and Solomon, J., "Recombination: a family of Markov chains for redistricting" (2019, forthcoming)
10. Duchin, M., "Outlier analysis for Pennsylvania congressional redistricting," (2018)
11. Duchin, M., Gladkova, T., Henninger-Voss, E., Klingensmith, B., Newman, H., and Wheelen, H. "Locating the representational baseline: Republicans in Massachusetts" arXiv:1810.09051. (2018)
12. Duchin, M., and Tenner, B. E. "Discrete geometry for electoral geography." arXiv preprint arXiv:1808.05860. (2018)
13. Fifield, B., Higgins, M., Imai, K., and Tarr, A. A new automated redistricting simulator using markov chain monte carlo. Work. Pap., Princeton Univ., Princeton, NJ (2015).
14. Gamburd, A. and Pak, I., "Expansion of product replacement graphs," *Combinatorica* 26.4: 411-429 (2006)
15. Herschlag, G., Ravier, R., and Mattingly, J., "Evaluating Partisan Gerrymandering in Wisconsin", arXiv: 1709.01596. (2017)
16. Mattingly, J. "Expert Report on the North Carolina State Legislature," available at `https://sites.duke.edu/quantifyinggerrymandering/files/2019/09/Report.pdf`
17. Mattingly, J. C., and Vaughn, C. Redistricting and the Will of the People. arXiv preprint arXiv:1410.8796 (2014)
18. Metric Geometry and Gerrymandering Group. **GerryChain** available at `https://gerrychain.readthedocs.io/en/latest/`.
19. Metric Geometry and Gerrymandering Group. "Comparison of Districting Plans for the Virginia House of Delegates," (2018)
20. Najt, Lorenzo, Daryl DeFord, and Justin Solomon. "Complexity and Geometry of Sampling Connected Graph Partitions." arXiv: 1908.08881. (2019)
21. Pegden, W., and Procaccia, A. and Yu, D. "A partisan districting protocol with provably nonpartisan outcomes." arXiv:1710.08781 (2017)
22. Stephanopoulos, N.., and McGhee, E.. "Partisan gerrymandering and the efficiency gap." *U. Chi. L. Rev. 82*: 831 (2015)
23. Tapp, K. Measuring Political Gerrymandering. *The American Mathematical Monthly*, 126(7), 593-609. (2019)

University of Wisconsin-Madison
Email address: `ellenber@math.wisc.edu`

A p-ADIC APPROACH TO RATIONAL POINTS ON CURVES

BJORN POONEN

Abstract. In 1922, Mordell conjectured the striking statement that for a polynomial equation $f(x, y) = 0$, if the topology of the set of complex number solutions is complicated enough, then the set of rational number solutions is finite. This was proved by Faltings in 1983, and again by a different method by Vojta in 1991, but neither proof provided a way to provably find all the rational solutions, so the search for other proofs has continued. Recently, Lawrence and Venkatesh found a third proof, relying on variation in families of p-adic Galois representations; this is the subject of the present exposition.

1. The Mordell conjecture

1.1. **Rational points on curves.** The equation $x^2 + y^2 = z^2$ has infinitely many solutions in integers satisfying $\gcd(x, y, z) = 1$. Equivalently, the circle $x^2 + y^2 = 1$ has infinitely many *rational* points ($(3/5, 4/5)$, $(5/13, 12/13)$, etc.) This can be understood geometrically: each line through $(-1, 0)$ with rational slope intersects the circle at one other point, which must have rational coordinates since finding its coordinates amounts to solving a quadratic equation over $\mathbb{Q}$ for which one rational root is already known. The same argument shows that any nonsingular conic section defined by a polynomial with rational coefficients having one rational point has infinitely many.

In contrast, Fermat proved that the equation $x^4 + y^4 = z^4$ has no positive integer solutions. Equivalently, the set of rational points on the plane curve $x^4 + y^4 = 1$ is $\{(\pm 1, 0), (0, \pm 1)\}$. It turns out that the curve $x^4 + y^4 = 17$ again has only finitely many rational points: $(\pm 2, \pm 1)$ and $(\pm 1, \pm 2)$. More generally, for any fixed $d \geq 4$ and nonzero $a \in \mathbb{Q}$, the curve $x^d + y^d = a$ has only finitely many rational points. These are special cases of the Mordell conjecture, which states that a "complicated enough" curve has only finitely many rational points, if any at all.

In the previous paragraph, the condition $d \geq 4$ is what made the curve "complicated enough". To state the Mordell conjecture fully, however, we need to consider also curves

Date: December 3, 2019.

This article is associated with a lecture to be given January 17, 2020 in the Current Events Bulletin at the 2020 Joint Mathematics Meetings in Denver. The writing of this article was supported in part by National Science Foundation grant DMS-1601946 and Simons Foundation grants #402472 (to Bjorn Poonen) and #550033.

defined by several polynomials in higher-dimensional space and to introduce the notion of genus to measure their geometric complexity.

1.2. **Projective space.** Let k be a field and let $n \in \mathbb{Z}_{\geq 0}$. The set of k-points on n-dimensional affine space is $\mathbb{A}^n(k) := k^n$.

Define an equivalence relation $\sim$ on $k^{n+1} - \{\vec{0}\}$ by $(a_0, \ldots, a_n) \sim (\lambda a_0, \ldots, \lambda a_n)$ for all $\lambda \in k^\times$. Let $(a_0 : \ldots : a_n)$ denote the equivalence class of $(a_0, \ldots, a_n)$. The set of all such equivalence classes is the set

$$\mathbb{P}^n(k) := \frac{k^{n+1} - \{\vec{0}\}}{k^\times},$$

of k-points on n-dimensional projective space.

The points $(a_0 : \ldots : a_n) \in \mathbb{P}^n(k)$ with $a_0 \neq 0$ have a unique representative of the form $(1, a_1, \ldots, a_n)$, so they form a copy of $\mathbb{A}^n(k)$. For each i, the same holds for the points with $a_i \neq 0$. Moreover, $\mathbb{P}^n(k)$ is the union of these $n+1$ overlapping copies of $\mathbb{A}^n(k)$.

One advantage of projective space over affine space is that $\mathbb{P}^n(\mathbb{R})$ is compact for the topology coming from the Euclidean topology on each $\mathbb{R}^n$; similarly, $\mathbb{P}^n(\mathbb{C})$ is compact. Related to this is that intersection theory works better in projective space: for example, two distinct lines in $\mathbb{P}^2(k)$ always meet in exactly one point.

1.3. **Projective varieties.** A finite list of polynomials $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ defines an affine variety[1] $X \subset \mathbb{A}^n$ whose set of k-points is

$$X(k) := \{\vec{a} \in \mathbb{A}^n(k) : f_1(\vec{a}) = \cdots = f_m(\vec{a}) = 0\}.$$

But for a point $(a_0 : \ldots : a_n) \in \mathbb{P}^n(k)$, a polynomial condition $f(\vec{a}) = 0$ makes sense (is unchanged by scaling $\vec{a}$) only if f is homogeneous, a sum of monomials of the same total degree, such as $x_0^5 x_1^2 - x_0^4 x_1^3 + 9x_1^7$ of degree 7. A finite list of homogeneous polynomials $f_1, \ldots, f_m \in k[x_0, \ldots, x_n]$ defines a projective variety $X \subset \mathbb{P}^n$ whose set of k-points is

$$X(k) := \{(a_0 : \cdots : a_n) \in \mathbb{P}^n(k) : f_1(\vec{a}) = \cdots = f_m(\vec{a}) = 0\}.$$

The decomposition of $\mathbb{P}^n$ as a union of $n+1$ copies of $\mathbb{A}^n$ restricts to express X as a union of $n+1$ affine varieties called affine patches. For each i, dehomogenizing $f_1, \ldots, f_m$ by setting x_i equal to 1 gives polynomials cutting out the ith affine patch in $\mathbb{A}^n$.

1.4. **Smooth varieties.** If a variety $Y \subset \mathbb{A}^n$ is defined by $f_1, \ldots, f_{n-r}$ such that for every field extension $L \supset k$ and point $\vec{a} \in Y(L)$, the matrix $\left(\left(\frac{\partial f_i}{\partial x_j}\right)(\vec{a})\right) \in \mathrm{M}_{n-r,n}(L)$ has rank $n-r$, then call Y obviously smooth of dimension r; the rank condition should be familiar as the Jacobian criterion in the implicit function theorem. More generally, any affine or projective variety X is called smooth of dimension r if it can be covered by subvarieties isomorphic to obviously smooth varieties Y as above (in a sense we will not make precise).

[1]Some people require a variety to satisfy additional conditions, such as not being a union of two strictly smaller such varieties.

If X is smooth of dimension r over $\mathbb{R}$, then $X(\mathbb{R})$ is a smooth $\mathbb{R}$-manifold of dimension r. The same holds if $\mathbb{R}$ is replaced by $\mathbb{C}$ in all three places.

1.5. **Genus of a curve.** From now on, we consider a smooth projective curve X over $\mathbb{Q}$, that is, a projective variety over $\mathbb{Q}$ that is smooth of dimension 1. We assume, moreover, that X is geometrically connected, meaning that the variety defined by the same polynomials over an algebraically closed extension field (such as $\mathbb{C}$) is nonempty and not the disjoint union of two strictly smaller varieties. Then $X(\mathbb{C})$ is a compact connected 1-dimensional $\mathbb{C}$-manifold, that is, a compact Riemann surface. Forgetting the complex structure, we find that $X(\mathbb{C})$ is a compact connected oriented 2-dimensional real manifold; by the classification of such, $X(\mathbb{C})$ is homeomorphic to a g-holed torus for some $g \in \mathbb{Z}_{\geq 0}$. The integer g is called the genus of X. It measures the geometric complexity of X.

Remark 1.1. It turns out that g also equals the dimension of the space of holomorphic 1-forms on $X(\mathbb{C})$. One can also define g algebraically, either by using Kähler differentials in place of holomorphic forms, or by computing the dimension of a sheaf cohomology group $\mathrm{H}^1(X, \mathscr{O}_X)$.

Example 1.2 (The Riemann sphere). If $X = \mathbb{P}^1$, then the space $X(\mathbb{C}) = \mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ is homeomorphic to a sphere via (the inverse of) stereographic projection. Thus $g = 0$.

Example 1.3 (Plane curves). If $X \subset \mathbb{P}^2$ is a smooth projective curve defined by a degree d homogeneous polynomial, then it turns out that $g = (d-1)(d-2)/2$.

Example 1.4 (Conic sections). A nondegenerate conic section is a smooth curve of degree 2 in $\mathbb{P}^2$. By Example 1.3, it is of genus 0.

Example 1.5 (Elliptic curves). An elliptic curve is a smooth degree 3 curve $y^2z - x^3 - Axz^2 - Bz^3 = 0$ in $\mathbb{P}^2$ for some numbers $A, B \in \mathbb{Q}$ (dehomogenizing by setting $z = 1$ gives the more familiar equation $y^2 = x^3 + Ax + B$ for one affine patch). By Example 1.3, an elliptic curve is of genus 1.

Example 1.6 (Hyperelliptic curves). Let $f(x) \in \mathbb{Q}[x]$ be a nonconstant polynomial with no repeated factors. Then $y^2 = f(x)$ defines a smooth curve in $\mathbb{A}^2$. It is isomorphic to an affine patch of some smooth projective geometrically connected curve X. If f has degree $2g+1$ or $2g+2$, then the genus of X is g.

Remark 1.7. The problem of determining the rational points on a general curve can be reduced to the problem for a smooth projective geometrically connected curve. That is why it suffices to consider only the latter.

1.6. **The conjecture.**

Mordell conjecture ([Mor22], first proved in [Fal83]). *Let X be a smooth projective geometrically connected curve of genus g over $\mathbb{Q}$. If $g > 1$, then $X(\mathbb{Q})$ is finite.*

Remark 1.8. One can say qualitatively what happens for curves of genus 0 and 1 as well:

Genus g	$X(\mathbb{Q})$	Some examples
0	infinite, if nonempty	lines and conics[2]
1	can be finite or infinite	elliptic curves, ...
> 1	finite	plane curves of degree ≥ 4, ...

Several proofs of the Mordell conjecture are known, none of them easy:

- Faltings [Fal83] proved the conjecture in 1983 using methods from Arakelov theory, a kind of arithmetic intersection theory that combines number-theoretic data with complex-analytic data.
- Vojta [Voj91] gave a completely different proof based on diophantine approximation, a theory whose original goal was to quantify how closely irrational algebraic numbers such as $\sqrt[3]{2}$ could be approximated by rational numbers with denominator of at most a certain size. For a more elementary variant of Vojta's proof due to Bombieri, see [Bom90] or [HS00].
- Lawrence and Venkatesh [LV18] recently gave yet another proof. Their proof shares some ingredients with Faltings's but replaces the most difficult steps by arguments involving p-adic Hodge theory. The rest of this article is devoted to explaining some of the ideas underlying their proof.

Remark 1.9. All of these proofs generalize to the case of curves defined over number fields instead of just $\mathbb{Q}$. (A number field is a finite field extension over $\mathbb{Q}$, such as $\mathbb{Q}(\sqrt{5})$.)

Remark 1.10. Although the proof of Lawrence and Venkatesh is the first *complete* proof of the Mordell conjecture using p-adic methods, older p-adic approaches have given partial results. Chabauty [Cha41] gave a proof for X satisfying an additional hypothesis, namely $\operatorname{rank} J(\mathbb{Q}) < g$ for a certain projective group variety J associated to X, the Jacobian. More recently, Kim [Kim05, Kim09] proposed a sophisticated extension of Chabauty's ideas, using the nilpotent fundamental group of X as a substitute for J. He proved that his approach combined with well-known conjectures would imply the Mordell conjecture. Kim's approach has already led to the explicit determination of $X(\mathbb{Q})$ for some X outside the reach of previous methods [BDM+19], and it may be that Kim's approach succeeds for every X of genus > 1.

Remark 1.11. All the proofs so far are ineffective: they do not prove that there is an algorithm that takes as input the list of polynomials defining a curve X of genus > 1 and outputs the list of all rational points on X. At best they give a computable upper bound on $\#X(\mathbb{Q})$ in terms of X. See [Poo02] for more about the algorithmic problem.

2. Galois representations

The Lawrence–Venkatesh proof makes essential use of p-adic Galois representations. Therefore, in this section we define $\mathbb{Q}_p$, define the absolute Galois group of a field, and give examples and properties of $\mathbb{Q}_p$-representations of the absolute Galois group of $\mathbb{Q}$.

2.1. **The field of p-adic numbers.** Let p be a prime number. The ring of p-adic integers is the inverse limit $\mathbb{Z}_p := \varprojlim \mathbb{Z}/p^n\mathbb{Z}$. Thus an element of $\mathbb{Z}_p$ is a sequence $(a_1, a_2, \ldots)$ where the elements $a_n \in \mathbb{Z}/p^n\mathbb{Z}$ are compatible in the sense that the natural homomorphism $\mathbb{Z}/p^{n+1}\mathbb{Z} \twoheadrightarrow \mathbb{Z}/p^n\mathbb{Z}$ maps a_{n+1} to a_n for each n. For example,

$$(3 \bmod 5,\ 13 \bmod 5^2,\ 38 \bmod 5^3,\ \ldots) \in \mathbb{Z}_5.$$

As a ring, $\mathbb{Z}_p$ is a domain of characteristic 0. Its fraction field, denoted $\mathbb{Q}_p$, is called the field of p-adic numbers.

For each $n \geq 1$, the homomorphism $\pi_n \colon \mathbb{Z}_p \to \mathbb{Z}/p^n\mathbb{Z}$ sending $(a_1, a_2, \ldots)$ to a_n is surjective with kernel $p^n\mathbb{Z}_p$. The kernel of $\pi_1 \colon \mathbb{Z}_p \to \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ is the unique maximal ideal $p\mathbb{Z}_p$ of $\mathbb{Z}_p$. The collection of subsets $\pi_n^{-1}(a)$ for all $n \in \mathbb{Z}_{\geq 1}$ and $a \in \mathbb{Z}/p^n\mathbb{Z}$ is a basis of a topology on $\mathbb{Z}_p$. Equip $\mathbb{Q}_p$ with the unique topology making it a topological group having $\mathbb{Z}_p$ as an open subgroup.

Remark 2.1. Here we explain an alternative construction of $\mathbb{Z}_p$ and $\mathbb{Q}_p$ and their topologies, producing the same results. The p-adic absolute value on $\mathbb{Q}$ is characterized by $\left|p^n \frac{a}{b}\right|_p := p^{-n}$ whenever $a, b, n \in \mathbb{Z}$ and $p \nmid a, b$. Define $\mathbb{Q}_p$ as the completion of $\mathbb{Q}$ with respect to $|\ |_p$, just as $\mathbb{R}$ is the completion of $\mathbb{Q}$ with respect to the standard absolute value. Then $|\ |_p$ extends to an absolute value on $\mathbb{Q}_p$. Define $\mathbb{Z}_p$ as the closed unit disk $\{x \in \mathbb{Q}_p : |x|_p \leq 1\}$. Finally, $|\ |_p$ induces a metric on $\mathbb{Q}_p$, which defines a topology on $\mathbb{Z}_p$ and $\mathbb{Q}_p$.

Working with $\mathbb{Z}_p$ or $\mathbb{Q}_p$ amounts to working with infinitely many congruences at once, but passing to the limit has advantages. One is that one can work over a domain or field of characteristic 0. Another is that some ideas from analysis over $\mathbb{R}$ have analogues for $\mathbb{Q}_p$.

Whereas number fields such as $\mathbb{Q}$ are examples of what are called global fields, $\mathbb{Q}_p$ is an example of a local field. For a more detailed introduction to p-adic numbers, see [Kob84].

2.2. **The absolute Galois group of $\mathbb{Q}$.** A complex number is algebraic over $\mathbb{Q}$ if it is a zero of some nonzero polynomial in $\mathbb{Q}[x]$. The set of all algebraic numbers is a subfield $\overline{\mathbb{Q}}$ of $\mathbb{C}$, called an algebraic closure of $\mathbb{Q}$.

Now let K be a subfield of $\overline{\mathbb{Q}}$. Call $K \supset \mathbb{Q}$ a finite extension if $\dim_{\mathbb{Q}} K$ is finite. Call $K \supset \mathbb{Q}$ a Galois extension if it is generated by the set of *all* zeros of some collection of polynomials in

$\mathbb{Q}[x]$.[3] For example, $\mathbb{Q}(\sqrt[3]{2})$ is not a Galois extension of $\mathbb{Q}$, but $\mathbb{Q}(\sqrt[3]{2}, e^{2\pi i/3}\sqrt[3]{2}, e^{4\pi i/3}\sqrt[3]{2})$ is. The field $\overline{\mathbb{Q}}$ is the union of its finite Galois subextensions K.

For a Galois extension $K \supset \mathbb{Q}$, the Galois group $\mathrm{Gal}(K/\mathbb{Q})$ is the set of automorphisms of K that fix $\mathbb{Q}$ pointwise[4]. The absolute Galois group of $\mathbb{Q}$ is $G_{\mathbb{Q}} := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Each automorphism of $\overline{\mathbb{Q}}$ restricts to give an automorphism of each finite Galois subextension K, and any compatible collection of such automorphisms defines an automorphism of $\overline{\mathbb{Q}}$, so

$$G_{\mathbb{Q}} \simeq \varprojlim_{\text{finite Galois } K \supset \mathbb{Q}} \mathrm{Gal}(K/\mathbb{Q}).$$

Just as the inverse limit $\mathbb{Z}_p$ had a topology, the inverse limit $G_{\mathbb{Q}}$ has a topology.

Remark 2.2. More generally, for any field F, one can construct a field $\overline{F}$ and topological group G_F.

2.3. **Global p-adic Galois representations.** Let V be a finite-dimensional $\mathbb{Q}_p$-vector space. If $\dim V = r$, then $\mathrm{Aut}\, V \simeq \mathrm{GL}_r(\mathbb{Q}_p)$, which has a topology coming from the topology of $\mathbb{Q}_p$. Call a $\mathbb{Q}_p$-linear action of $G_{\mathbb{Q}}$ on V continuous if the homomorphism $\rho\colon G_{\mathbb{Q}} \to \mathrm{Aut}\, V$ defined by the action is continuous. By a $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}}$ we mean a finite-dimensional $\mathbb{Q}_p$-vector space V equipped with a continuous action of $G_{\mathbb{Q}}$. In the next few sections, we give examples of such representations arising in number theory and arithmetic geometry.

2.4. **The cyclotomic character.** Let m be a positive integer. Define

$$\mu_m := \{z \in \overline{\mathbb{Q}} : z^m = 1\},$$

which under multiplication is a cyclic group of order m. Thus μ_m is a free $\mathbb{Z}/m\mathbb{Z}$-module of rank 1. The group $G_{\mathbb{Q}}$ acts on the group μ_m.

Now fix a prime p, and let m range through the powers of p. Form the inverse limit

$$T := \varprojlim \mu_{p^n}$$

with respect to the homomorphisms $\mu_{p^{n+1}} \twoheadrightarrow \mu_{p^n}$ sending ζ to ζ^p. Then T is a free rank 1 module under the ring $\mathbb{Z}_p := \varprojlim \mathbb{Z}/p^n\mathbb{Z}$, and $G_{\mathbb{Q}}$ acts on T.

Next let

$$V := T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p.$$

Then V is a 1-dimensional $\mathbb{Q}_p$-vector space, and $G_{\mathbb{Q}}$ acts on V. It follows from the definitions that the action is continuous, so V is a 1-dimensional $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}}$, called the cyclotomic character.

[3]For a definition that works over an arbitrary ground field k instead of $\mathbb{Q}$, one should require each polynomial to have distinct zeros in $\overline{k}$.

[4]Fixing $\mathbb{Q}$ pointwise is automatic; this condition becomes relevant only over other ground fields.

2.5. **Galois representations associated to elliptic curves.** Let E be an elliptic curve over $\mathbb{Q}$. It turns out that E is a group variety; in particular, there is a map of varieties $E \times E \to E$ making $E(\overline{\mathbb{Q}})$ an abelian group. If $P \in E(\overline{\mathbb{Q}})$, we may use this group law to define $3P := P + P + P$ and so on. For each $m \geq 1$, it turns out that the m-torsion subgroup

$$E[m] := \{P \in E(\overline{\mathbb{Q}}) : mP = 0\}$$

is a free $\mathbb{Z}/m\mathbb{Z}$-module of rank 2. Therefore the inverse limit

$$T_p E := \varprojlim E[p^n]$$

(with respect to the homomorphisms $E[p^{n+1}] \to E[p^n]$ sending P to pP) is a free $\mathbb{Z}_p$-module of rank 2, called a Tate module. Next,

$$V_p E := T_p E \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$$

is a 2-dimensional $\mathbb{Q}_p$-vector space. The continuous action of $G_{\mathbb{Q}}$ on $E(\overline{\mathbb{Q}})$ induces continuous actions on $E[p^n]$, $T_p E$, and $V_p E$. Thus $V_p E$ is a 2-dimensional $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}}$.

2.6. **Galois representations associated to higher-genus curves.** Let X be a smooth projective geometrically connected curve of genus g over $\mathbb{Q}$. If $g \neq 1$, there is no group law $X \times X \to X$, but the Jacobian J of X does have a group law. The construction of $V_p E$ generalizes to produce a $2g$-dimensional $\mathbb{Q}_p$-representation $V_p J$ of $G_{\mathbb{Q}}$.

2.7. **Galois representations from étale cohomology.** If X is a smooth projective variety over $\mathbb{Q}$, and $i \in \mathbb{Z}_{\geq 0}$ then the étale cohomology group $\mathrm{H}^i(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_p)$ (which we will not attempt to define here) is a $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}}$.

Example 2.3. If E is an elliptic curve, then it turns out that $\mathrm{H}^1(E_{\overline{\mathbb{Q}}}, \mathbb{Q}_p)$ is the dual of the representation $V_p E$. If X and J are as in Section 2.6, then $\mathrm{H}^1(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_p)$ is the dual of $V_p J$.

2.8. **Semisimple representations.** Let V be a $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}}$. Call V irreducible if $V \neq 0$ and there is no $G_{\mathbb{Q}}$-invariant subspace W with $0 \subsetneq W \subsetneq V$. Call V semisimple if it is a direct sum of irreducible representations. Maschke's theorem [Ser77, §1.4, Theorem 2] states that any $\mathbb{C}$-representation of a finite group is semisimple, but this is not true for $\mathbb{F}_p$-representations of a finite group of order divisible by p, and $\mathbb{Q}_p$-representations of $G_{\mathbb{Q}}$ are more like the latter in this regard: they need not be semisimple.

Example 2.4. Let $\chi \colon G_{\mathbb{Q}} \to \mathbb{Q}_p^\times$ be the cyclotomic character. There is a homomorphism $\log_p \colon \mathbb{Q}_p^\times \to \mathbb{Q}_p$ from the multiplicative group to the additive group; see [Kob84, IV.2]. Let $V := \mathbb{Q}_p \oplus \mathbb{Q}_p$, viewed as a space of column vectors. Let each $g \in G_{\mathbb{Q}}$ act as $\begin{pmatrix} 1 & \log_p \chi(g) \\ 0 & 1 \end{pmatrix}$ on V. The only invariant subspace of V is $\mathbb{Q}_p \oplus 0$, so V is not semisimple.

2.9. **The absolute Galois group of $\mathbb{Q}_p$.** Let $\overline{\mathbb{Q}}_p$ denote an algebraic closure of $\mathbb{Q}_p$. The homomorphism $\mathbb{Z} \hookrightarrow \mathbb{Z}_p$ extends uniquely to $\mathbb{Q} \hookrightarrow \mathbb{Q}_p$ and non-uniquely to $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$; fix one such embedding. Define $G_{\mathbb{Q}_p} := \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. It turns out that $\overline{\mathbb{Q}}_p$ is generated by its subfields $\overline{\mathbb{Q}}$ and $\mathbb{Q}_p$, so the homomorphism $G_{\mathbb{Q}_p} \to G_{\mathbb{Q}}$ sending each σ to $\sigma|_{\overline{\mathbb{Q}}}$ is injective. Identify $G_{\mathbb{Q}_p}$ with its image, which is called a decomposition group of $G_{\mathbb{Q}}$.

The absolute value $|\ |_p$ on $\mathbb{Q}_p$ extends in a unique way to $\overline{\mathbb{Q}}_p$. Let $\overline{\mathbb{Z}}_p := \{x \in \overline{\mathbb{Q}}_p : |x|_p \leq 1\}$; it is a subring. The unique maximal ideal of $\overline{\mathbb{Z}}_p$ is $\mathfrak{m} := \{x \in \overline{\mathbb{Q}}_p : |x|_p < 1\}$, and the quotient $\overline{\mathbb{Z}}_p/\mathfrak{m}$ is an algebraic closure $\overline{\mathbb{F}}_p$ of $\mathbb{F}_p$. Each element of $G_{\mathbb{Q}_p}$ preserves $|\ |_p$ and hence induces an automorphism of $\overline{\mathbb{Z}}_p/\mathfrak{m}$. Thus we obtain a homomorphism $G_{\mathbb{Q}_p} \to G_{\mathbb{F}_p}$. It is surjective, and its kernel $I_p \subset G_{\mathbb{Q}_p} \subset G_{\mathbb{Q}}$ is called an inertia group. To summarize, we have a diagram

$$\begin{array}{ccccccc} 1 & \longrightarrow & I_p & \longrightarrow & G_{\mathbb{Q}_p} & \longrightarrow & G_{\mathbb{F}_p} \longrightarrow 1 \\ & & & & \big\downarrow & & \\ & & & & G_{\mathbb{Q}}. & & \end{array}$$

The Frobenius automorphism $\mathrm{Frob}_p \in G_{\mathbb{F}_p}$ is the automorphism $x \mapsto x^p$ of $\overline{\mathbb{F}}_p$; it generates a dense subgroup of $G_{\mathbb{F}_p}$ since it restricts to a generator of each finite quotient $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$. Write Frob_p also for any element of $G_{\mathbb{Q}_p}$ mapping to $\mathrm{Frob}_p \in G_{\mathbb{F}_p}$, or for the corresponding element of $G_{\mathbb{Q}}$.

2.10. **Local Galois representations.** Let p and q be primes. (Soon we will take $q = p$.) A $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}_q}$ is a finite-dimensional $\mathbb{Q}_p$-vector space V equipped with a continuous action of $G_{\mathbb{Q}_q}$. Call V unramified if I_q acts trivially on V; in that case the $G_{\mathbb{Q}_q}$-action can be described by one matrix, namely the automorphism $\mathrm{Frob}_q|_V \in \mathrm{GL}(V)$ given by the action of any $\mathrm{Frob}_q \in G_{\mathbb{Q}_q}$. Given $w \in \mathbb{Z}$, call such a V pure of weight w if the characteristic polynomial of $\mathrm{Frob}_q|_V$ is a polynomial in $\mathbb{Z}[x]$ whose complex zeros have absolute value $q^{w/2}$.

2.11. **Properties of representations coming from geometry.** Now return to a global representation V, a $\mathbb{Q}_p$-representation of $G_{\mathbb{Q}}$. For each prime q, restricting the $G_{\mathbb{Q}}$-action to the subgroup $G_{\mathbb{Q}_q}$ yields a local representation V_q. For a finite set S of primes, call V unramified outside S if V_q is unramified for every $q \notin S$. For $w \in \mathbb{Z}$, call such a V pure of weight w if V_q is pure of weight w for every $q \notin S$. These properties were introduced because they hold for representations "coming from geometry":

Theorem 2.5. *Each representation* $\mathrm{H}^i(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_p)$ *as in Section 2.7 is unramified outside a finite set* S *and is pure of weight* i *(cf.* [Del74, Théorème 1.6]*).*

Remark 2.6. One can say more about S. The variety X can be defined by polynomials with coefficients in $\mathbb{Z}$. Reducing all the coefficients of the polynomials modulo ℓ produces polynomials defining a variety over $\mathbb{F}_\ell$. For most ℓ, this variety is again smooth; more

precisely, this holds for all primes ℓ outside a finite set S_0. Then in Theorem 2.5 one may take $S = S_0 \cup \{p\}$.

2.12. **Faltings's finiteness theorem for global Galois representations.** Faltings cleverly combined a few classical facts from number theory (Hermite's finiteness theorem and the Chebotarev density theorem) to prove the following finiteness statement.

Theorem 2.7 (cf. [Fal83, proof of Satz 5])**.** *Fix a nonnegative integer d, a prime p, a finite set S of primes, and an integer w. Then the set of isomorphism classes of semisimple d-dimensional $\mathbb{Q}_p$-representations of $G_{\mathbb{Q}}$ that are unramified outside S and pure of weight w is finite.*

3. A family of curves

In this section, we construct an algebraic family of curves that plays a key role in the Lawrence–Venkatesh proof.

3.1. **Fundamental group of a punctured Riemann surface.** Let X be a compact Riemann surface of genus g. Because X is homeomorphic to a $4g$-gon with edges glued appropriately, the fundamental group of X (with respect to any base point) has a presentation

$$\pi_1(X) \simeq \left\langle a_1, b_1, \ldots, a_g, b_g \;\middle|\; [a_1, b_1] \cdots [a_g, b_g] \right\rangle,$$

where $[a, b] := aba^{-1}b^{-1}$; that is, $\pi_1(X)$ is the quotient of a free group on $2g$ generators by the smallest normal subgroup containing the indicated product of g commutators. More generally, if B is a finite subset of X, then

$$\pi_1(X - B) \simeq \left\langle a_1, b_1, \ldots, a_g, b_g, c_1, \ldots, c_r \;\middle|\; [a_1, b_1] \cdots [a_g, b_g] c_1 \cdots c_r \right\rangle.$$

3.2. **Analytic construction of a family of ramified covers.** Now fix X and a finite group G. Let $x \in X$. A surjective homomorphism $\pi_1(X - \{x\}) \overset{\alpha}{\twoheadrightarrow} G$ defines a finite covering space of $X - \{x\}$, and it can be completed to a finite *ramified* covering $Y_{x,\alpha} \to X$, with some branches possibly coming together above $x \in X$.

This covering depends on α, but there are only finitely many α since $\pi_1(X - \{x\})$ is finitely generated. To obtain a space not depending on a choice of any one α, define the finite disjoint union $Y_x := \coprod_\alpha Y_{x,\alpha}$, which is a disconnected ramified covering of X.[5] As x varies, the Y_x vary continuously in a family. The total space of this family is a *2-dimensional* compact complex manifold Y with a proper submersion $\pi\colon Y \to X$ such that $\pi^{-1}(x) = Y_x$ for each $x \in X$.

[5]Lawrence and Venkatesh use a variant in which G has trivial center and the disjoint union is over *conjugacy classes* of surjective homomorphisms α, which makes sense since the isomorphism type of $Y_{x,\alpha}$ depends only on the conjugacy class.

3.3. **An algebraic family of curves.** The constructions above can be made algebraic, in the following sense. Suppose that X is a smooth projective connected curve over $\mathbb{C}$. Then by the Riemann existence theorem, $Y_{x,\alpha} \to X$ arises from an algebraic morphism of algebraic curves. Moreover, there is a 2-dimensional variety Y with a morphism $\pi\colon Y \to X$ whose fibers are the disconnected curves $Y_x = \coprod_\alpha Y_{x,\alpha}$.

Even better, the construction is canonical enough that if X is defined over $\mathbb{Q}$, then $\pi\colon Y \to X$ can be defined over $\mathbb{Q}$. This is called a Kodaira–Parshin family; see [LV18, §7] for details.

Remark 3.1. The curve X is playing two roles: it is the base of the family $Y \to X$, but also each fiber Y_x is a ramified covering of X.

4. The Lawrence–Venkatesh proof of the Mordell conjecture

We now summarize the Lawrence–Venkatesh approach to the Mordell conjecture; of course, in doing so we gloss over many difficult arguments.

4.1. **From rational points to representations.** Let X be a smooth projective geometrically connected curve of genus > 1 over $\mathbb{Q}$. The goal is to prove that $X(\mathbb{Q})$ is finite.

All the claims below will be true for a suitable choice of finite group G. (Lawrence and Venkatesh take G to be the semidirect product $\mathbb{F}_q \rtimes \mathbb{F}_q^\times$ for a suitable large prime q.) Let $\pi\colon Y \to X$ be the Kodaira–Parshin family of curves over X defined using G. To each $x \in X(\mathbb{Q})$ one may associate the global Galois representation

$$V_x := \mathrm{H}^1_{\mathrm{et}}((Y_x)_{\overline{\mathbb{Q}}}, \mathbb{Q}_p).$$

Thus one obtains a map of sets[6]

$$\begin{aligned} X(\mathbb{Q}) &\xrightarrow{\mathscr{V}} \{\mathbb{Q}_p\text{-representations of } G_{\mathbb{Q}}\} \\ x &\longmapsto V_x. \end{aligned} \tag{1}$$

Now

- The representations V_x are all of the same dimension.
- They are semisimple.[7]
- They are unramified outside a set S that is independent of x, because one can choose a set S_0 as in Remark 2.6 that works for the whole family $Y \to X$.
- They are pure of weight 1.

[6]The set of representations should really be a set of *isomorphism classes* of representations (we abuse notation).

[7]The semisimplicity is actually a difficult theorem, proved by Faltings in his paper on the Mordell conjecture. Lawrence and Venkatesh would be "cheating" if they used this, so instead they give an independent argument using Hodge–Tate weights to prove that V_x is semisimple for all but finitely many $x \in X(\mathbb{Q})$; that is sufficient for their proof of the Mordell conjecture.

That is, the representations V_x satisfy all the conditions of Faltings's finiteness theorem (Theorem 2.7), so

$$\text{the map } \mathscr{V} \text{ in (1) has finite image!}$$

To finish the proof that $X(\mathbb{Q})$ itself is finite, one needs to show that every fiber of $\mathscr{V}$ is finite, i.e., that the V_x vary enough that there are only finitely many $x \in X(\mathbb{Q})$ mapping to any given isomorphism class of representations.

4.2. **Variation in a *p*-adic family of local Galois representations.** The plan is to show that the global representations V_x vary enough by showing that even their restrictions to $G_{\mathbb{Q}_p}$ vary enough.

To each $x \in X(\mathbb{Q}_p)$ one may associate the local Galois representation

$$V_x := \mathrm{H}^1_{\mathrm{et}}((Y_x)_{\overline{\mathbb{Q}}_p}, \mathbb{Q}_p).$$

This provides the map $\mathscr{V}_p$ in the following commutative diagram of sets:

$$\begin{array}{ccc} X(\mathbb{Q}) & \xrightarrow{\mathscr{V}} & \{\mathbb{Q}_p\text{-representations of } G_{\mathbb{Q}}\} \\ \downarrow & & \downarrow \text{restriction} \\ X(\mathbb{Q}_p) & \xrightarrow{\mathscr{V}_p} & \{\mathbb{Q}_p\text{-representations of } G_{\mathbb{Q}_p}\}. \end{array}$$

To prove that $\mathscr{V}$ has finite fibers, it suffices to prove that $\mathscr{V}_p$ has finite fibers. That is, loosely speaking, one must show that the local representation V_x varies enough as x ranges over $X(\mathbb{Q}_p)$. The rest of the proof proceeds as follows:

- Use p-adic Hodge theory to relate the variation of the étale cohomology groups V_x for $x \in X(\mathbb{Q}_p)$ to the variation of the Hodge filtration in the corresponding de Rham cohomology groups.
- The variation of the Hodge filtration is described by the Gauss–Manin connection, which in down-to-earth terms means that it is described by the solutions to a system of differential equations whose coefficients are algebraic functions on X over $\mathbb{Q}$.
- The same differential equations describe the variation of the Hodge filtration for the family $Y_{\mathbb{C}} \to X_{\mathbb{C}}$ of complex projective curves.
- A lower bound on that variation is given by the monodromy of the Kodaira–Parshin family over $\mathbb{C}$.
- Use an extensive calculation in topology (involving mapping class groups, Dehn twists, and the like) to prove that indeed the monodromy group is large enough.

This completes the proof of the Mordell conjecture.

Remark 4.1. Lawrence and Venkatesh show that their approach has applications beyond rational points on curves. In particular, using recent work of Bakker and Tsimerman [BT19], they prove that certain affine varieties F of higher dimension (moduli spaces of smooth

hypersurfaces in projective space) have few *integral* points, where "few" means that they are contained in a subvariety of F of lower dimension.

References

[BT19] Benjamin Bakker and Jacob Tsimerman, *The Ax-Schanuel conjecture for variations of Hodge structures*, Invent. Math. **217** (2019), no. 1, 77–94, DOI 10.1007/s00222-019-00863-8. MR3958791 ↑4.1

[BDM+19] Jennifer Balakrishnan, Netan Dogra, J. Steffen Müller, Jan Tuitman, and Jan Vonk, *Explicit Chabauty-Kim for the split Cartan modular curve of level 13*, Ann. of Math. (2) **189** (2019), no. 3, 885–944, DOI 10.4007/annals.2019.189.3.6. MR3961086 ↑1.10

[Bom90] Enrico Bombieri, *The Mordell conjecture revisited*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **17** (1990), no. 4, 615–640. MR1093712 (92a:11072) ↑1.6

[Cha41] Claude Chabauty, *Sur les points rationnels des courbes algébriques de genre supérieur à l'unité*, C. R. Acad. Sci. Paris **212** (1941), 882–885 (French). MR0004484 (3,14d) ↑1.10

[Del74] Pierre Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. **43** (1974), 273–307 (French). MR0340258 (49 #5013) ↑2.5

[Fal83] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), no. 3, 349–366 (German). English translation: Finiteness theorems for abelian varieties over number fields, 9–27 in *Arithmetic Geometry (Storrs, Conn., 1984)*, Springer, New York, 1986. Erratum in: Invent. Math. **75** (1984), 381. MR718935 (85g:11026a) ↑1.6, 1.6, 2.7

[HS00] Marc Hindry and Joseph H. Silverman, *Diophantine geometry: an introduction*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York, 2000. MR1745599 (2001e:11058) ↑1.6

[Kim05] Minhyong Kim, *The motivic fundamental group of* $\mathbf{P}^1 \setminus \{0, 1, \infty\}$ *and the theorem of Siegel*, Invent. Math. **161** (2005), no. 3, 629–656. MR2181717 (2006k:11119) ↑1.10

[Kim09] Minhyong Kim, *The unipotent Albanese map and Selmer varieties for curves*, Publ. Res. Inst. Math. Sci. **45** (2009), no. 1, 89–133, DOI 10.2977/prims/1234361156. MR2512779 ↑1.10

[Kob84] Neal Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, 2nd ed., Graduate Texts in Mathematics, vol. 58, Springer-Verlag, New York, 1984. MR754003 (86c:11086) ↑2.1, 2.4

[LV18] Brian Lawrence and Akshay Venkatesh, *Diophantine problems and p-adic period mappings*, August 30, 2018. Preprint, `arXiv:1807.02721v2` . ↑1.6, 3.3

[Mor22] L. J. Mordell, *On the rational solutions of the indeterminate equations of the third and fourth degrees*, Proc. Cambridge Phil. Soc. **21** (1922), 179–192. ↑1.6

[Poo02] Bjorn Poonen, *Computing rational points on curves*, Number Theory for the Millennium, III (Urbana, IL, 2000), 2002, pp. 149–172. MR1956273 (2003k:11105) ↑1.11

[Ser77] Jean-Pierre Serre, *Linear representations of finite groups*, Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott; Graduate Texts in Mathematics, Vol. 42. MR0450380 ↑2.8

[Voj91] Paul Vojta, *Siegel's theorem in the compact case*, Ann. of Math. (2) **133** (1991), no. 3, 509–548. MR1109352 (93d:11065) ↑1.6

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA

Email address: `poonen@math.mit.edu`

URL: `http://math.mit.edu/~poonen/`

RECENT PROGRESS ON MOVING BOUNDARY PROBLEMS

SUNČICA ČANIĆ

ABSTRACT. We give a brief survey of the recent progress in the area of mathematical well-posedness for moving boundary problems describing fluid-structure interaction between incompressible, viscous fluids and elastic, viscoelastic, and rigid solids.

1. INTRODUCTION

In this paper we survey some recent developments and open problems in the mathematical study of moving boundary problems. In particular, the focus is on problems arising from the interaction between incompressible, viscous fluids and elastic, viscoelastic, or rigid solids (also referred to as "structures"). See Fig. 1. Fluid-structure interaction (FSI) problems are ubiquitous in nature, technology and engineering. Examples include the human heart and heart valves interacting with blood flow, biodegradable micro-beads swimming in water to clean up water pollution, a micro camera in the human intestine used for an early colon cancer detection, and design of next generation vascular stents to prop open the clogged arteries, and prevent heart attacks. Numerical simulation and analysis of fluid-structure interaction problems can provide insight into the "invisible" properties of flows and structures, and can be used to advance design of novel technologies and improve the understanding of many biological phenomena.

Interestingly enough, even though the mathematical theory of the motion of bodies in a liquid is one of the oldest and most classical problems in fluid mechanics, mathematicians have only recently become interested in a systematic study of the basic problems related to fluid-structure interaction. One reason for this may be that problems of this type are notoriously difficult to study. In addition to the nonlinearity in the fluid equations (the Navier-Stokes equations) and possibly in the elastic or viscoelastic structure equations, the coupling between the fluid and structure motion may give rise to strong geometric nonlinearities. The study of existence of solutions to the coupled problems must account for the nonlinearities due to the strong energy exchange between the fluid and (elastic) structure motion, and novel compactness arguments have to be designed to deal with such nonlinearities. Due to the fluid domain motion, compactness results holding for *a family of operators* defined on time-dependent function spaces corresponding to moving domains not known *a priori* are needed. The compactness arguments must also account for the fact that the coupled problem involves two sets of equations of different type (parabolic vs. hyperbolic) accounting for the different physics in the problem.

Date: November 2019.

2000 *Mathematics Subject Classification.* Primary 76D05, 76D03; Secondary 74F10, 76D27.

Key words and phrases. Moving boundary problems, fluid-structure interaction.

The author was supported in part by NSF Grants DMS-1853340 and DMS-1613757.

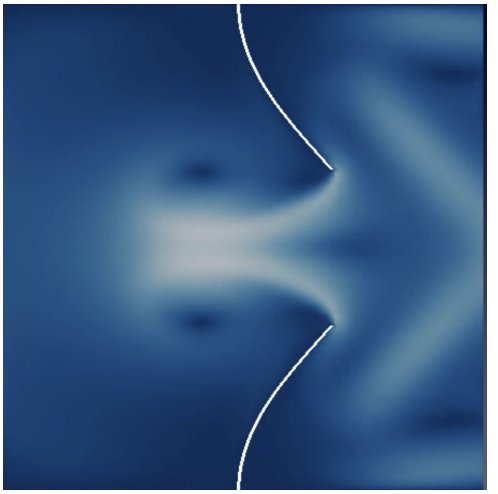


FIGURE 1. Two elastic leaflets "swimming" in fluid. The intensity of blue color is related to the magnitude of fluid velocity. From the work published in [13].

Making use of the smoothing effects by the fluid viscosity, and capturing its role in keeping the high frequency oscillations of the (hyperbolic) structure under control, is crucial for the existence proofs.

In existence proofs, and in numerical schemes, an additional difficulty is imposed by the incompressibility of the fluid. The main difficulty in existence proofs is related to the construction of divergence-free extensions of fluid velocity to a larger domain containing all the moving domains, and obtaining quantitative estimates of the extensions in terms of the changing geometry. Incompressibility is intimately related to the pressure, and pressure is a major component of the load, i.e., contact force, exerted by the fluid onto the solid. Designing constructive existence proofs and numerical schemes that approximate the load "correctly" is a key ingredient for the stability of constructive solution schemes. In particular, the fluid surrounding the structure affects the structure motion as an extra mass that the structure must displace when moving within a fluid. This has long been known in engineering as the "added mass effect", and its negative impact on the stability of partitioned FSI schemes is a well-known problem for FSI problems for which the fluid and structure have comparable densities, i.e., for which the structure is "light" with respect to the fluid. The added mass is a leading order effect in biofluidic FSI problems, since biological tissues (structures) have density which is approximately the same as that of the surrounding fluid. A failure to account for this effect leads to instabilities in partitioned numerical schemes and to the lack of uniform energy estimates in constructive existence proofs.

The question of global-in-time existence of solutions to moving boundary problems is affected by the so called "no collision" paradox. In addition to the problems related to global existence of solutions inherited from the Navier-Stokes equations, global weak solution existence results for moving boundary problems are typically obtained until contact between solids happens. We survey below the results in this area and mention here that several open questions remain, which would shed light

on whether contact of elastic bodies immersed in a viscous incompressible fluid can happen in finite time, and the type of boundary conditions for which the finite-time contact may or may not occur (Navier slip versus no-slip condition).

Nonlinearities in the coupled FSI problem also affect the study of uniqueness of solutions. It is not surprising that uniqueness of weak solutions to the coupled FSI problems is still largely an outstanding open problem, since even in the case of classical 3D Navier-Stokes equations, the uniqueness of the Leray-Hopf weak solutions has not been resolved. However, recent advances in this area are significant, and we summarize those results below.

Thus, the main challenges in the mathematical study of fluid-structure interaction problems can be attributed to the following features of the problem:

(1) *Geometric nonlinearity when fluid and structure are nonlinearly coupled*;
(2) *The problem is of mixed type and defined on moving domains*;
(3) *Incompressibility and the Added Mass Effect*;
(4) *Finite-time contact?*
(5) *Uniqueness.*

To explain the main challenges in more detail, we present a benchmark problem for FSI involving elastic structures, and a benchmark problem for FSI involving rigid solids, and provide a literature review of the recent results.

2. FSI involving elastic structures

Although the development of numerical methods for this class of problems started almost 40 years ago (see e.g., [42, 122, 107, 108, 46, 68, 69, 41, 81, 111, 9, 8, 84, 47, 80, 68, 69, 70, 71, 72] and the references therein), the development of existence theory for FSI problems started less than 20 years ago. We state a benchmark problem in this field, and summarize some recent results and open problems.

2.1. **Description of the Main Problem.**

To describe the interaction between a fluid and an elastic (or viscoelastic) structure across a moving interface mathematically, two types of coupling conditions have to be prescribed. This contrasts classical fluid dynamics problems defined on fixed domains where only one set of boundary conditions, e.g., the no-slip condition, is sufficient to define the problem. The two sets of coupling conditions describe: (1) how the kinematic quantities, such as velocity, are coupled (the *kinematic coupling condition*), and (2) the elastodynamics of the fluid-structure interface (the *dynamic coupling condition*). While the precise form of the kinematic and dynamic coupling conditions depends on the particular application at hand, the most common coupling is done via the no-slip kinematic condition, stating that the fluid and structure velocities are continuous at the moving boundary, and the dynamic coupling condition stating that the fluid-structure interface, namely the moving boundary, is driven by the jump in traction, i.e., normal stress, across the interface. For problems in which one expects small interface displacements and small displacement gradients, the coupling conditions may be evaluated at a fixed interface, without changing the fluid domain, namely, the fluid and structure may be *linearly coupled* [45, 10, 11, 87]. For problems where this may not be a good approximation of reality, the coupling conditions must be evaluated across the moving interface, giving rise to an additional nonlinearity in the problem, which is due to the change of geometry of the moving boundary, namely, the fluid and structure are *nonlinearly coupled*. In the latter case the fluid

domain is a function of time, and additionally, it is not known *a priori* since it depends on the unknowns in the problem, namely, the location of the fluid-structure interface.

The geometric nonlinearity associated with the fact that the fluid equations are defined on a family of time-dependent domains not known a priori, presents one of the major difficulties in studying this class of problems mathematically.

FSI Benchmark Problem with the no-slip Coupling. The simplest example of a moving boundary problem with nonlinear coupling involving a deformable (elastic) structure, is a benchmark problem deriving from modeling blood flow in a segment of an artery. The fluid domain is a cylinder, with an elastic (viscoelastic) lateral boundary. For simplicity, we present the problem in 2D, although 3D versions of the problem have been studied in e.g., [97]. In this benchmark problem, will be assuming that the lateral boundary is thin, with small thickness $h << 1$, and with the *reference configuration* Γ corresponding to a straight cylinder of length L and radius R:

$$\Gamma = \{(z, R) \in \mathbb{R}^2 | z \in (0, L)\}.$$

In most literature involving FSI with thin structures, except for the recent results in [23, 100], the lateral boundary is assumed to displace only in the vertical (normal, transverse) direction, rendering longitudinal displacement negligible. By using η to denote the vertical component of displacement, the fluid domain $\Omega^\eta(t)$ can be described by:

$$\Omega^\eta(t) = \{(z, r) \in \mathbb{R}^2 | z \in (0, L), r \in (0, R + \eta(t, z)\},$$

where R is the radius of the *reference cylinder*. The *reference fluid domain* will be denoted by $\Omega = (0, L) \times (0, R)$.

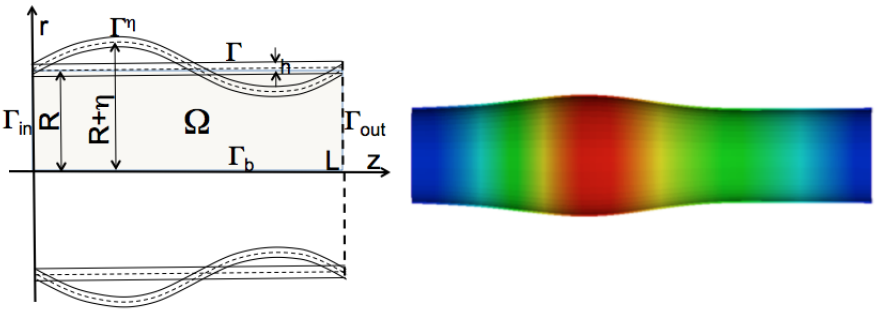


FIGURE 2. Left: A sketch of the fluid domain and the moving lateral boundary. Right: Pressure wave propagation in elastic tube.

The fluid flow is modeled by the Navier-Stokes equations for an incompressible, viscous fluid, defined on a moving domain $\Omega^\eta(t)$:

$$\left.\begin{array}{rcl} \rho_f\big(\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u}\big) &=& \nabla\cdot\boldsymbol{\sigma} \\ \nabla\cdot\mathbf{u} &=& 0 \end{array}\right\} \text{ in } \Omega^\eta(t),\ t \in (0, T), \tag{2.1}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor, ρ_f is the fluid density, and $\mathbf{u} = \mathbf{u}(\boldsymbol{x}, t) = (u_z, u_r)$ is the fluid velocity. For Newtonian fluids

$$\boldsymbol{\sigma} = -p\mathbf{I} + 2\mu\mathbf{D}(\mathbf{u}),$$

where μ is the dynamic viscosity coefficient, and $\mathbf{D}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla^\tau\mathbf{u})$ is the symmetrized gradient of $\mathbf{u}$.

In this benchmark problem the flow is driven by the inlet and outlet dynamic pressure data, and the flow is normal to the inlet and outlet boundary $\Gamma_{in} = \{0\} \times (0, R)$ and $\Gamma_{out} = \{L\} \times (0, R)$:

$$\left.\begin{array}{rcl} p + \dfrac{\rho_f}{2}|u|^2 &=& P_{in/out}(t), \\ u_r &=& 0, \end{array}\right\} \quad \text{on } \Gamma_{in/out}, \tag{2.2}$$

where $P_{in/out} \in L^2_{loc}(0, \infty)$ are given. At the bottom boundary $\Gamma_b = (0, L) \times \{0\}$ the symmetry boundary conditions are prescribed:

$$u_r = \partial_r u_z = 0, \quad \text{on } \Gamma_b. \tag{2.3}$$

Different inlet/outlet boundary conditions have been used in numerical simulations, including Dirichlet data given in terms of the prescribed fluid velocity, and Neumann data given in terms of the prescribed normal stress. The dynamics pressure data (2.2) is a boundary condition "consistent" with the energy of the coupled problem.

Under the fluid loading, and possibly some external loading, the elastic cylinder deforms. See Fig 2. We denote by:

$$\Gamma^{\boldsymbol{\eta}}(t) = \{(t, z, R + \eta(t, z))|z \in (0, L)\} \tag{2.4}$$

the location of the deformed cylinder lateral boundary at time t. The elastic properties of the cylinder's lateral wall can be described by an operator $\mathscr{L}_e$, so that the elastodynamics problem, in Lagrangian formulation, can be written as:

$$\rho_s h\partial_{tt}\eta + \mathscr{L}_e\eta = f, \qquad \text{on } \Gamma,\ t \in (0, T), \tag{2.5}$$

where ρ_s is the structure density, h is the thin structure thickness, and f is the vertical component of the outside loading (force density) experienced by the elastic structure. The loading f in the coupled problem will come from the jump in the normal stress across the structure, i.e., from the fluid load experienced by the structure (assuming that outside loading is zero). The operator $\mathscr{L}_e$ is associated with the elastic energy of the structure, such as the membrane or shell energy, see [96], and is typically continuous, positive-definite, and coercive on some Hilbert space χ.

The coupling. The fluid flow influences the motion of the structure through traction forces, i.e., by the normal stress exerted onto the structure at $\Gamma^{\boldsymbol{\eta}}(t)$, while the structure influences the fluid through its inertial and elastic forces due to the structure motion and stretching/recoil. Additionally, the fluid and structure "feel" each other through the continuity of the fluid and structure velocities at the interface. Thus, the kinematic and dynamic coupling conditions, respectively, are:

$$(\partial_t\eta(t, z), 0) = \boldsymbol{u}(t, z, R + \eta(t, z)), \tag{2.6}$$

$$\rho_s h\partial_t^2\eta + \mathscr{L}_e\eta = -J\ (\boldsymbol{\sigma}\boldsymbol{n})|_{(t,z,R+\eta(t,z))} \cdot \mathbf{e}_r, \tag{2.7}$$

where $J = \sqrt{1 + \left(\frac{\partial\eta}{\partial t}\right)^2}$, is the Jacobian of the transformation from Eulerian coordinates to Lagrangian coordinates, and $\mathbf{e}_r$ is the unit vector in the vertical direction. Here, we have assumed that the outside forcing onto the structure is zero. Generalizations to include outside forcing due to the presence of another elastic structure or other types of forcing can be found in [98].

The geometric nonlinearity due to the fluid domain motion, described by the composite function $\boldsymbol{u}(t,z,R+\eta(t,z))$, is *generally* handled by introducing a family of mappings, parameterized by time, called the Arbitrary Lagrangian-Eulerian (ALE) mappings, discussed below in Sec. 2.2. In terms of ALE mappings, the trace of the fluid velocity on $\Gamma^\eta(t)$ is described by a composite function between the velocity and the ALE mapping.

Equations (2.1)–(2.7) define a nonlinear moving-boundary problem for the unknown functions $\mathbf{u}$ and η. The problem is supplemented with initial conditions:

$$\mathbf{u}(0,.)=\mathbf{u}_0,\ \eta(0,.)=\eta_0,\ \partial_t\eta(0,.)=v_0, \tag{2.8}$$

often times satisfying some extra compatibility conditions, such as:

$$\begin{array}{c} \mathbf{u}_0(z,R+\eta_0(z))=v_0(z)\mathbf{e}_r,\quad z\in(\Gamma),\\ \eta_0(0)=\eta_0(L)=v_0(0)=v_0(L)=0,\\ R+\eta_0(z)>0,\quad z\in[0,L]. \end{array} \tag{2.9}$$

Thus, the **benchmark nonlinear moving-boundary problem**, which exemplifies the main difficulties associated with studying moving boundary problems with nonlinear coupling, can be summarized as follows: Find $\mathbf{u}=(u_z(t,z,r),u_r(t,z,r)),p$, and $\eta(t,z)$ such that:

$$\left.\begin{array}{rcl}\rho_f\big(\partial_t\mathbf{u}+(\mathbf{u}\cdot\nabla)\mathbf{u}\big) &=& \nabla\cdot\boldsymbol{\sigma}(\mathbf{u},p)\\ \nabla\cdot\mathbf{u} &=& 0\end{array}\right\}\ \text{in } \Omega^\eta(t),\ t\in(0,T),$$

$$\left.\begin{array}{rcl}\mathbf{u}|_{\Gamma^\eta(t)} &=& \partial_t\eta\mathbf{e}_r,\\ \rho_s h\partial_t^2\eta+\mathscr{L}_e\eta &=& -J^\eta\boldsymbol{\sigma}\mathbf{n}|_{\Gamma^\eta(t)}\cdot\mathbf{e}_r,\end{array}\right\}\ \text{on } (0,T)\times\Gamma,$$

$$\left.\begin{array}{rcl}u_r &=& 0,\\ \partial_r u_z &=& 0,\end{array}\right\}\ \text{on } (0,T)\times\Gamma_b,$$

$$\left.\begin{array}{rcl}p+\frac{\rho_f}{2}|u|^2 &=& P_{in/out}(t),\\ u_r &=& 0,\end{array}\right\}\text{on } (0,T)\times\Gamma_{in/out},$$

$$\eta(t,0)=\partial_z\eta(t,0)=\eta(t,L)=\partial_z\eta(t,L)=0\quad\text{on } (0,T)$$

$$\left.\begin{array}{rcl}\mathbf{u}(0,.) &=& \mathbf{u}_0,\\ \eta(0,.) &=& \eta_0,\\ \partial_t\eta(0,.) &=& v_0.\end{array}\right\}\text{at } t=0.$$

The energy. This benchmark problem satisfies the following formal energy inequality:

$$\frac{d}{dt}E(t)+D(t)\le C(P_{in}(t),P_{out}(t)), \tag{2.10}$$

where $E(t)$ denotes the sum of the kinetic energy of the fluid and of the structure, and the elastic energy of the membrane shell:

$$E(t) \;=\; \frac{\rho_f}{2}\|\mathbf{u}\|^2_{L^2(\Omega^\eta(t))}+\frac{\rho_s h}{2}\|\partial_t\eta\|^2_{L^2(\Gamma)}+\langle\mathscr{L}_e\eta,\eta\rangle\,, \tag{2.11}$$

where $\langle\mathscr{L}_e\eta,\eta\rangle$ corresponds to the elastic energy of the structure, which for the cylindrical Koiter shell allowing only radial displacement reads:

$$\langle\mathscr{L}_e\eta,\eta\rangle:=\frac{1}{2}\big(C_0\|\eta\|^2_{L^2(\Gamma)}+C_1\|\partial_z\eta\|^2_{L^2(\Gamma)}+C_2\|\partial_z^2\eta\|^2_{L^2(\Gamma)}\big).$$

The term $D(t)$ captures dissipation due to fluid viscosity:

$$D(t) = \mu \|\mathbf{D}(\mathbf{u})\|^2_{L^2(\Omega^\eta(t)))}, \tag{2.12}$$

and $C(P_{in}(t), P_{out}(t)))$ is a constant which depends only on the inlet and outlet pressure data, which are both functions of time.

FSI Benchmark Problem with the Navier-slip Coupling. While the assumption on the continuity of normal velocity components is reasonable for impermeable boundaries:

$$(\partial_t \boldsymbol{\eta} - \boldsymbol{u}) \cdot \mathbf{n} = 0 \quad (non - penetration\ condition), \tag{2.13}$$

the continuity of the tangential velocity component in the no-slip condition is justified only when molecular viscosity is considered [93]. Navier contested the no-slip condition for Newtonian fluids [103] when he claimed that the tangential, slip velocity should be proportional to the shear stress. For moving boundary problems this means that the jump in the tangential components of the fluid and solid velocities at the moving boundary is proportional to the shear stress:

$$(\partial_t \boldsymbol{\eta} - \boldsymbol{u}) \cdot \boldsymbol{\tau} = \alpha \boldsymbol{\sigma} \mathbf{n} \cdot \boldsymbol{\tau} \quad (Navier\ slip), \tag{2.14}$$

where $\mathbf{n}$ and $\boldsymbol{\tau}$ to denote the unit normal and tangent to the fluid domain boundary, respectively, $\boldsymbol{\sigma}$ is the fluid Cauchy stress tensor, and α is the proportionality constant known as the slip length; $1/\alpha$ has the units of friction.

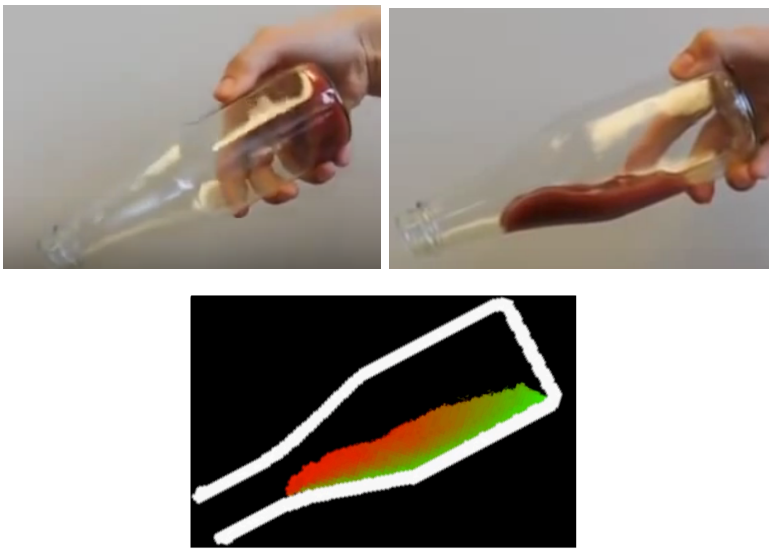


FIGURE 3. Hydrophilic vs. hydrophobic surfaces: no-slip vs. slip condition. Left: Classical Ketchup flow. Middle: Ketchup flowing in a bottle treated with a no-stick coating. Right: Numerical simulation (by Čanić and Wang) of flow with slip boundary condition. The snap-shots were taken at the same time after the bottled had been tilted downwards. The colors in the third panel denote magnitude of fluid velocity.

The benchmark problem defined on the domain shown in Fig. 2, incorporating the Navier-slip condition as the kinematic coupling condition, can be summarized as follows: Find $(\mathbf{u}, p, \boldsymbol{\eta})$ such that the following holds

The fluid equations:

$$\left.\begin{array}{rcl} \rho_F(\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}) &=& \nabla\cdot\boldsymbol{\sigma}(\mathbf{u},p), \\ \nabla\cdot\mathbf{u} &=& 0, \end{array}\right\} \text{ in } \Omega^\eta(t),\ t\in(0,T);$$

The elastic structure (Navier-slip coupling on $(0,L)\times(0,T)$):

$$\begin{array}{rcl} \rho_S h\partial_{tt}\boldsymbol{\eta}(t,z) + \mathscr{L}_e\boldsymbol{\eta}(t,z) &=& -J(t,z)\boldsymbol{\sigma}\big(\boldsymbol{\varphi}(t,z)\big)\mathbf{n}(t,z), \\ \partial_t\boldsymbol{\eta}(t,z)\cdot\mathbf{n}(t,z) &=& \boldsymbol{u}(\boldsymbol{\varphi}(t,z))\cdot\mathbf{n}(t,z), \\ (\partial_t\boldsymbol{\eta}(t,z) - \mathbf{u}(\boldsymbol{\varphi}(t,z))\cdot\boldsymbol{\tau}(t,z) &=& \alpha\boldsymbol{\sigma}\big(\boldsymbol{\varphi}(t,z)\big)\mathbf{n}(t,z)\cdot\boldsymbol{\tau}(t,z), \end{array}$$

with

$$\boldsymbol{\eta}(t,0) = \partial_z\boldsymbol{\eta}(t,0) = \boldsymbol{\eta}(t,L) = \partial_z\boldsymbol{\eta}(t,L) = 0,\ t\in(0,T);$$

Boundary data at the inlet/outlet boundary $\Gamma_{in/out}\times(0,T)$:

$$p + \tfrac{\rho_F}{2}|\mathbf{u}|^2 = P_i,\ \mathbf{u}\cdot\boldsymbol{\tau} = 0.$$

Boundary data at the bottom, symmetry boundary $\Gamma_b\times(0,T)$:

$$\mathbf{u}\cdot\mathbf{n} = 0,\ \partial_{\mathbf{n}}u_\tau = 0.$$

with u_τ denoting the tangental component of velocity $\mathbf{u}$.

Initial conditions:

$$\mathbf{u}(0,.) = \mathbf{u}_0,\ \boldsymbol{\eta}(0,.) = \boldsymbol{\eta}_0,\ \partial_t\boldsymbol{\eta}(0,.) = \mathbf{v}_0. \tag{2.15}$$

The following **energy estimate holds**:

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\big(\rho_F\|\mathbf{u}\|^2_{L^2(\Omega^\eta(t))} + \rho_S h\|\partial_t\boldsymbol{\eta}\|^2_{L^2(\Gamma)} + c\|\boldsymbol{\eta}\|^2_{\chi(\Gamma)}\big) \\ &+\mu\|\mathbf{D}(\mathbf{u})\|^2_{L^2(\Omega^\eta(t))} + \frac{1}{\alpha}\|u_\tau - \partial_t\eta_\tau\|^2_{L^2(\Gamma^\eta(t))} \le \mathbf{C}, \end{aligned} \tag{2.16}$$

where $\mathbf{C}$ depends on the initial and boundary data, and constant c in front of the χ-norm of $\boldsymbol{\eta}$ is associated with the coercivity of the structure operator $\mathscr{L}_e$. The reference configuration of the lateral boundary $\Gamma = (0,L)\times\{R\}$.

The no-slip condition is reasonable for a great variety of problems for which the slip length α is indeed very close to zero. However, in many cases of practical significance no-slip is not adequate. Examples include flows over hydrophobic surfaces or surfaces treated with a no-stick coating, see Fig. 3, flows over "rough" surfaces such as those of, e.g., grooved vascular tissue scaffolds, and problems involving contact of smooth solids immersed in a viscous, incompressible fluid. More precisely, for flows over "rough" (rigid and fixed) surfaces, it has been shown that the Navier slip boundary condition is the appropriate "effective boundary condition" [93, 94]. Instead of using the no-slip condition at the small groove scale, the effective Navier slip boundary condition is applied at the corresponding "groove-free" *smooth* boundary [93, 94]. Regarding contact of smooth bodies immersed in a viscous, incompressible fluid, recent studies have shown that contact is not possible if the no-slip boundary condition is considered [75, 76, 114]. A resolution to this no-collision paradox is to employ a different boundary condition, such as the Navier slip boundary condition, which allows contact between smooth rigid bodies [106].

Problems of this type arise, e.g., in modeling elastic heart valve closure where different kinds of *ad hoc* "gap" conditions with the no-slip boundary condition have been used to get around this difficulty.

2.2. **The Arbitrary Lagrangian-Eulerian (ALE) Mappings.** To deal with the problems associated with the motion of the fluid domain, different approaches have been taken. One approach is to consider the entire moving boundary problem written in Lagrangian coordinates, as was done in [85, 30, 29]. This is possible to do when the fluid domain is contained in a "closed container" and no fluid escapes the fluid domain, which is not the case with the benchmark problem, considered above. Another approach is to map the problems from the moving domain onto a fixed reference domain, using a family of mappings, known as the Arbitrary Lagrangian-Eulerian (ALE) mappings.

The Arbitrary Lagrangian-Eulerian mapping is a family of (diffeomorphic) mappings, parameterized by t, such that

$$\mathscr{A}^{\eta}(t) : \Omega \to \Omega^{\eta}(t). \tag{2.17}$$

ALE mappings have been extensively used in numerical simulations of moving boundary problems, see e.g. [41, 42, 122]. Recently, they have proven to be useful in mathematical analysis as well [96, 23, 97]. In numerics, one of the reasons for the introduction of ALE mappings is the calculation of the the discretized time derivative $\partial_t \mathbf{u}$ since a finite difference approximation of the time derivative, e.g., $(\mathbf{u}^{n+1} - \mathbf{u}^n)/\Delta t$ contains the functions $\mathbf{u}^{n+1}$ and $\mathbf{u}^n$ which are defined on two different domains, one corresponding to the time $t^{n+1} = (n+1)\Delta t$, and the other to $t^n = n\Delta t$. A way to calculate the time derivative of the fluid velocity is then to map the fluid velocities at times t^{n+1} and t^n onto a fixed domain Ω via the ALE mappings corresponding to t^{n+1} and t^n, evaluate the time derivative there, and then map everything back to the physical domain $\Omega^n := \Omega^{\eta^n}(t^n)$ to solve the fluid equations on the "current" domain Ω^n. This introduces an extra advection term in the Navier-Stokes equations, describing the contribution of the motion of fluid domain to the fluid advection, so that the Navier-Stokes equations in ALE form become:

$$\left.\begin{array}{rcl} \rho_f\big(\partial_t \mathbf{u} + ((\mathbf{u} - \mathbf{w}^{\eta^n}) \cdot \nabla)\mathbf{u}\big) &=& \nabla \cdot \sigma, \\ \nabla \cdot \mathbf{u} &=& 0, \end{array}\right\} \text{in } (t^n, t^{n+1}) \times \Omega^n, \tag{2.18}$$

where $\mathbf{w}^{\eta^n} = \partial_t \mathscr{A}^{\eta^n}(t^n)$ describes the fluid domain velocity. In numerical solvers, the ALE mapping is often defined by the harmonic extension of the boundary data onto the fluid domain, i.e., as a solution to the following elliptic problem:

$$\begin{array}{rcl} \Delta \mathscr{A}^{\eta}(t) &=& 0 \text{ on } \Omega, \\ \mathscr{A}^{\eta}(t) &=& \eta(t) \text{ on } \Gamma, \\ \mathscr{A}^{\eta}(t) &=& 0 \text{ on } \partial\Omega \setminus \Gamma, \end{array}$$

calculated at every time step $t = t^n$. Other elastic (elliptic) operators have also been used.

In addition to numerical solvers, the ALE approach has recently been used to study the existence of solutions to this class of problems by either mapping the entire fluid problem onto the fixed domain Ω via an ALE mapping, and analyzing the problem there, as in [96, 97, 98], or by mimicking the approach described above, used in numerical simulations, and working on the "current" domain Ω^n, as in [23]. In the former case, an additional set of nonlinearities is introduced because the

gradient operator ∇ in physical space is mapped into an operator ∇^η, defined on the fixed, reference domain. The mapped fluid velocity is no longer divergence-free in terms of the operator ∇, which presents some difficulties when trying to use known results that hold for divergence-free functions.

Using ALE mappings in analysis requires assumptions on its regularity, which, of course, depends on the regularity of the fluid-structure interface η, not known *a priori*. In most works that use ALE mappings, an assumption on the regularity of the ALE mapping is made *a priori*, which is then justified *a posteriori*, after a proof showing existence of a solution with sufficient regularity of η is obtained.

For the benchmark problem with no-slip coupling presented above, one can introduce a family of ALE mappings, parameterized by η, given explicitly by:

$$\mathscr{A}^\eta(t) : \Omega \to \Omega^\eta(t), \quad \mathscr{A}^\eta(t)(\tilde{z},\tilde{r}) := \begin{pmatrix} \tilde{z} \\ (R+\eta(t,\tilde{z}))\tilde{r} \end{pmatrix}, \quad (\tilde{z},\tilde{r}) \in \Omega, \tag{2.19}$$

where $(\tilde{z},\tilde{r})$ denote the coordinates in the reference domain $\Omega = (0,L)\times(0,R)$. Mapping $\mathscr{A}^\eta(t)$ is a bijection, and its Jacobian is given by

$$\mathscr{J}_{\mathscr{A}^\eta} = |\det \nabla \mathscr{A}^\eta(t)| = |R+\eta(t,\tilde{z})|. \tag{2.20}$$

Composite functions with the ALE mapping will be denoted by

$$\mathbf{u}^\eta(t,.) = \mathbf{u}(t,.)\circ\mathscr{A}^\eta(t) \quad \text{and} \quad p^\eta(t,.) = p(t,.)\circ\mathscr{A}^\eta(t). \tag{2.21}$$

The derivatives of composite functions satisfy:

$$\partial_t \mathbf{u} = \partial_t \mathbf{u}^\eta - (\mathbf{w}^\eta\cdot\nabla^\eta)\mathbf{u}^\eta, \quad \nabla\mathbf{u} = \nabla^\eta\mathbf{u}^\eta,$$

where the ALE domain velocity, $\mathbf{w}^\eta$, and the transformed gradient, ∇^η, are given by:

$$\mathbf{w}^\eta = \partial_t\eta\tilde{r}\mathbf{e}_r, \quad \nabla^\eta = \begin{pmatrix} \partial_{\tilde{z}} - \tilde{r}\dfrac{\partial_z\eta}{R+\eta}\partial_{\tilde{r}} \\ \dfrac{1}{R+\eta}\partial_{\tilde{r}} \end{pmatrix}. \tag{2.22}$$

Note that

$$\nabla^\eta\mathbf{v} = \nabla\mathbf{v}(\nabla A_\eta)^{-1}. \tag{2.23}$$

The following notation will also be useful:

$$\sigma^\eta = -p^\eta\mathbf{I} + 2\mu\mathbf{D}^\eta(\mathbf{u}^\eta), \quad \mathbf{D}^\eta(\mathbf{u}^\eta) = \frac{1}{2}(\nabla^\eta\mathbf{u}^\eta + (\nabla^\eta)^\tau\mathbf{u}^\eta).$$

The resulting problem, defined entirely on the fixed, reference domain, in ALE framework now reads: find $\mathbf{u}(t,\tilde{z},\tilde{r}), p(t,\tilde{z},\tilde{r})$ and $\eta(t,\tilde{z})$ such that

$$\left.\begin{aligned} \rho_f\big(\partial_t\mathbf{u}^\eta + ((\mathbf{u}^\eta-\mathbf{w}^\eta)\cdot\nabla^\eta)\mathbf{u}^\eta\big) &= \nabla^\eta\cdot\boldsymbol{\sigma}^\eta, \\ \nabla^\eta\cdot\mathbf{u}^\eta &= 0, \end{aligned}\right\} \text{in } (0,T)\times\Omega, \tag{2.24}$$

$$\left.\begin{aligned} \mathbf{u}^\eta &= \partial_t\eta\mathbf{e}_r, \\ \rho_s h\partial_t^2\eta + \mathscr{L}_e\eta &= -J\boldsymbol{\sigma}^\eta\mathbf{n}\cdot\mathbf{e}_r, \end{aligned}\right\} \text{ on } (0,T)\times\Gamma,$$

$$\left.\begin{aligned} u_r^\eta &= 0, \\ \partial_r u_z^\eta &= 0 \end{aligned}\right\} \text{on } (0,T)\times\Gamma_b,$$

$$\left.\begin{aligned} p + \tfrac{\rho_f}{2}|\mathbf{u}^\eta|^2 &= P_{in/out}(t), \\ u_r^\eta &= 0, \end{aligned}\right\} \text{on } (0,T)\times\Gamma_{in/out},$$

$$\mathbf{u}^\eta(0,.) = \mathbf{u}_0^\eta, \eta(0,.) = \eta_0, \partial_t\eta(0,.) = v_0.$$

Recent Results and Open Problems. The development of existence theory for this kind of moving boundary, fluid-structure interaction problems started in the late 1990's/early 2000's. The first existence results were obtained for the cases in which the structure is completely immersed in the fluid, and the structure was considered to be either a rigid body, or described by a finite number of modal functions. See e.g., [15, 31, 36, 37, 39, 48, 57, 114], and the references therein. The analysis of the coupling between the $2D$ or $3D$ Navier-Stokes equations and $2D$ or $3D$ linear elasticity started in the early 2000's with the works in which the coupling between the fluid and structure was assumed across a fixed fluid-structure interface (*linear coupling*) as in [45, 10, 11, 87], and then extended to problems with *nonlinear coupling* in the works [14, 90, 64, 25, 65, 96, 97, 89, 34, 35, 85, 30, 29]. More precisely, concerning nonlinear FSI models, the first FSI existence result, locally in time, was obtained in [14], where a **strong** solution for an interaction between an incompressible, viscous fluid in $2D$ and a $1D$ viscoelastic string was obtained, assuming periodic boundary conditions. This result was extended by Lequeurre in [90], where the existence of a unique, local in time, strong solution for any data, and the existence of a global strong solution for small data, was proved in the case when the structure is modeled as a clamped viscoelastic beam. D. Coutand and S. Shkoller proved existence, locally in time, of a unique, regular solution for an interaction between a viscous, incompressible fluid in $3D$ and a $3D$ structure, immersed in the fluid, where the structure was modeled by the equations of linear [34], or quasi-linear [35] elasticity. In the case when the structure (solid) is modeled by a linear wave equation, I. Kukavica et al. proved the existence, locally in time, of a strong solution, assuming lower regularity for the initial data [85, 82]. A similar result for compressible flows can be found in [86]. In [110] Raymod et al. considered a FSI problem between a linear elastic solid immersed in an incompressible viscous fluid, and proved the existence and uniqueness of a strong solution. Most of the above mentioned existence results for strong solutions are local in time. In [83] a global existence result for small data was obtained by Ignatova et al. for a moving boundary FSI problem involving a damped linear wave equation with some additional damping terms in the coupling conditions, showing exponential decay in time of the solution. In the case when the structure is modeled as a $2D$ elastic shell interacting with a viscous, incompressible fluid in $3D$, the existence, locally in time, of a unique regular solution was proved by Shkoller et al. in [30, 29]. We mention that the works of Shkoller et al., and Kukavica at al. were obtained in the context of Lagrangian coordinates, which were used for both the structure and fluid subproblems.

In the context of **weak** solutions, the first existence results came out in 2005 when Chambolle et al. showed the existence of a weak solution for a FSI problem between a $3D$ incompressible, viscous fluid and a $2D$ viscoelastic plate in [25]. Grandmont improved this result in [65] to hold for a $2D$ elastic plate. A constructive existence proof for the interaction between an incompressible, viscous fluid and a linearly elastic Koiter shell with transverse displacement was designed in [96]. The first constructive existence proof for moving boundary problems was presented by

Ladyzhenskaya in 1970 where the interaction between an incompressible, viscous fluid and a *given moving boundary* was constructed using a time-discretization approach, known as Rothe's method, assuming high regularity of the given interface [88]. Muha and Čanić designed a Rothe's-type method in the context of moving boundaries that are *not known a priori* in 2013 [96]. To deal with a moving boundary not known *a priori* they introduced the time discretization via Lie operator splitting, which has been used in numerical schemes, described in [62]. The FSI problem studied in [96] is split into a fluid and a structure subproblem, with the coupling designed so that the resulting scheme is stable. This was achieved by using the results about the "added mass effect", published in [24], which showed the importance of implicit treatment of the fluid and structure inertia in loosely coupled partitioned schemes. The splitting of the coupled problem in [96] was then done in such a way that the fluid and structure inertia terms are kept implicitly together, which provided uniform energy estimate, not otherwise attainable using the "classical" Dirichlet-Neumann" partitioned schemes [109, 112]. A compactness argument, discussed below in Section 2.3 was used to show that subsequence of approximate solutions converge to a weak solution to the coupled problem. After 2013 similar approaches were used to prove existence of a weak solution for a nonlinear FSI problem involving a nonlinear Koiter shell [99], a multi-layered structure [98], and a Koiter shell with the Navier slip condition [100].

To complete the discussion of well-posedness, we mention here a result on continuous dependence of weak solutions on initial data, obtained in [73] for a fluid structure interaction problem with a free boundary type coupling condition.

In all these works existence of a weak solution was proved for as long as the elastic boundary does not touch "the bottom" (rigid) portion of the fluid domain boundary. Recently, Grandmont and Hillairet showed that contact between a rigid bottom of a fluid container, and a *viscoelastic beam*, is not possible in finite time [66]. The finite-time contact involving thin and thick elastic structures interacting with an incompressible, viscous fluid is still **open**.

We conclude this section with a few general remarks related to the geometric nonlinearity in FSI problems for which the coupling across the current location of the moving interface is needed to describe the physical problem. The strong exchange of energy between the fluid and structure motion in the nonlinearly coupled problems gives rise to the various difficulties in the study of mathematical well-posedness. In particular, the functional spaces based on the finite energy considerations may not provide sufficient regularity of the moving interface to even define the trace of the fluid velocity at the fluid-structure interface, and additionally, may lead to various fluid domain degeneracies, as shown in Fig. 4. These problems are particularly evident when elastic structures are thin (modeled by the reduced membrane or shell equations) and the structure model accounts for both transverse and tangential components of displacement, and the structure is interacting in 3D with the flow of an incompressible viscous fluid. In those cases, the weak solution techniques based on finite energy spaces are often times insufficient to guarantee even the Lipschitz regularity of the fluid-structure interface, see [23]. This is one of the reason why most literature on the existence of (weak) solutions to moving boundary problems involving thin elastic structures assumes only the transverse component of displacement to be different from zero [14, 90, 64, 25, 65, 96, 97, 89]. Recently, a 3D FSI problem allowing transverse and tangential displacements of a mesh-supported shell

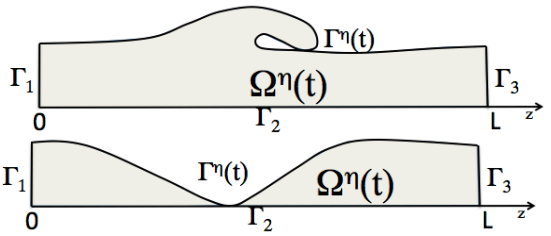


FIGURE 4. Two ways the fluid domain can degenerate. Top: loss of injectivity of the ALE mapping $\mathscr{A}^\eta(t)$. Bottom: loss of injectivity of the ALE mapping $\mathscr{A}^\eta(t)$ and loss of strict positivity of the Jacobian $\mathscr{A}^\eta(t)$.

was studied in [23] where an existence of a weak solution in 3D was obtained under an extra assumption on the uniform Lipschitz property for the fluid-structure interface. FSI problems with elastic structures that are slightly more regular than the Koiter shell, such as, e.g., tripolar materials studied in [16, 113], do not suffer from this difficulty.

The issues related to non-zero transverse displacement cannot be avoided in FSI problems with the Navier-slip coupling. This is the reason why the existence result in 2D for a FSI problem involving a Koiter shell interacting with the flow of an incompressible, viscous fluid via the Navier-slip condition [100], holds only until the fluid domain remains regular, in the sense that degeneracies of the type shown in Fig. 4 do not occur. In problems with slip, compactness may be a an additional problem since the regularizing effects by fluid viscosity are transferred to the structure only via the non-penetration condition holding in the normal direction to the boundary. Never the less, the friction effects in the tangential direction can be used to compensate for the lack of regularization provided by the fluid viscosity. More details about compactness for problems on moving domains are presented next.

2.3. **Compactness.** Compactness results similar to Aubin-Lions-Simon lemma [7, 116] that hold for moving boundary problems are difficult to obtain because, among other things, the function spaces depend on time via the fluid domain motion, and the fluid domains are not known *a priori*. A compactness result in generalized Bochner spaces $L^2(0,T;H(t))$, where $H(t)$ is a family of Hilbert spaces which depend on time, is needed to deal with the fluid flow nonlinearities and with the geometric nonlinearities associated with the fluid domain motion. Such a compactness result should include, among other things, conditions on the dynamic change of the fluid domain geometry, that would guarantee compactness.

To the best of our knowledge, there is no general compactness theory similar to Aubin-Lions-Simon lemma [7, 116] for spaces $L^2(0,T;H(t))$, where $H(t)$ depends on time. There are several compactness results for particular, specific problems for which the spatial domain depends on time, but they were proved using assumptions that hold for that particular problem at hand. The first result of this type was obtained by Fujita and Sauer in [53], where they studied the Navier-Stokes equations in a given, non-cylindrical domain, namely, in a domain which depends on time,

and whose motion is given *a priori*. Similarly, in [33], the authors used a compactness result to study a fluid-rigid body interaction problem, where an assumption on high regularity of the domain motion had to be used to obtain an existence result (see also [50, 114]). In the case of fluid-elastic structure interaction problems, the assumption on high regularity of the interface is typically not satisfied, thus different approaches need to be employed.

In the context of fluid-elastic structure interaction problems, we mention [39, 65] where the authors considered a fluid-elastic structure interaction problem between the flow of a viscous, incompressible fluid and an elastic/viscoelastic plate, in which a compactness argument based on Simon's theorem was used to show L^2-strong convergence of approximate solutions. We also mention [79] where this approach was used in the case of a non-Newtonian fluid. A similar problem, but in a more general geometrical setting, was studied in [89], where compactness of a set of approximate weak solutions, based on a particular linearization and regularization of the problem, was proved by using a modification of the ideas from the proof of Aubin-Lions lemma. Both approaches used the fluid viscosity and kinematic coupling condition to control high frequency oscillations of the structure velocity. Recently, a version of Aubin-Lions lemma for a moving domain problem was proved in the context of *compressible fluids*, see [18], where an existence of a solution to an FSI problem between a compressible fluid and a linearly elastic shell was obtained. The lack of the fluid incompressibility constraint simplifies the compactness argument for the velocity field.

Compactness results in more general frameworks were studied in [104, 105], where the authors developed a functional framework based on the flow method and the Piola transform for problems in *smoothly* moving domains, where the flow causing domain motion was given *a priori*. In those works a version of the Aubin-Lions lemma was obtained within this framework. A different version of Aubin-Lions lemma, in a more general form, was also considered in [95]. The approach in [95] was based on negative Sobolev space-type estimates, defined on non-cylindrical, i.e., time-dependent domains. The latter approach did not require high degree of smoothness of the domain motion.

We also mention the results obtained in [12, 28, 95], where generalizations of the Aubin-Lions-Simon lemma in various types of nonlinear settings were obtained, and the work in [43] where a version of Aubin-Lions-Simon result was obtained in the context of finite element spaces. We also mention the works by Elliott at al. where compactness arguments were developed and used to study parabolic problems on moving surfaces [4, 5, 2, 6, 3].

Most of the works mentioned above were obtained for continuous time, i.e., the time variable was not discretized, and most of them were tailored for a particular application in mind. Working with discretized time brings some additional difficulties in terms of the uniform bounds for the time-shifts (translations in time). In the time-discretized case, namely, for the approaches based on Rothe's semi-discretization method, the uniform bounds on the time-shifts need to be somewhat stronger to guarantee compactness, see Proposition 2 in [44]. In particular, the work in [44] addresses a version of Aubin-Lions-Simon result for piecewise constant functions in time, obtained using Rothe's method, but for a problem defined on a *fixed* Banach space.

The work presented in [101] concerns a generalization of the Aubin-Lions-Simon result involving Hilbert spaces that are *solution dependent*, and not necessarily known *a priori*. This is a significant step forward, since the result can be applied to a large class of moving boundary problems, including numerical solvers. To account for the time dependence of the function spaces associated with the motion of the fluid domains, the authors identified a new set of conditions, which quantify the dependence of the Hilbert spaces on time so that an extension of Aubin-Lions-Simon result can be applied to a sequence of approximate solutions constructed using Rothe's method.

More precisely, the compactness result in [101] is designed for problems which can be described in general as evolution problems:

$$\begin{aligned} \frac{d\mathbf{u}}{dt} &= A^t\mathbf{u}, t \in (0,T), \\ \mathbf{u}(0) &= \mathbf{u}_0, \end{aligned} \tag{2.25}$$

where $A^t : V(t) \to W(t)$ is a family of (nonlinear) spatial differential operators that depend on time as a parameter. For example, $\dfrac{d\mathbf{u}}{dt} = A^t\mathbf{u}$ may correspond to the Navier-Stokes equations for an incompressible, viscous fluid defined on a moving domain $\Omega(t)$. In this case, A^t is a spatial differential operator that associates to each $\mathbf{u}$ the function $\nabla \cdot \boldsymbol{\sigma} - \mathbf{u} \cdot \nabla\mathbf{u}$, where $\boldsymbol{\sigma}$ is the fluid Cauchy stress tensor, and $\mathbf{u}$ is divergence free, satisfying certain boundary conditions on $\Omega(t)$.

A way to "solve" this class of problems is to semi-discretize the problem in time by sub-dividing the time interval $(0,T)$ into N sub-intervals of width $\Delta t = T/N$, and introduce the piecewise constant approximate functions

$$\mathbf{u}_{\Delta t} = \mathbf{u}^n_{\Delta t} \quad \text{for } t \in ((n-1)\Delta t, n\Delta t],\ n = 1, \dots, N, \tag{2.26}$$

which satisfy, e.g., a backward Euler approximation of the problem on (t^n, t^{n+1}):

$$\frac{\mathbf{u}^{n+1}_{\Delta t} - \mathbf{u}^n_{\Delta t}}{\Delta t} = A^{t^{n+1}}\mathbf{u}^{n+1}_{\Delta t} \quad \text{or} \quad \frac{\mathbf{u}^{n+1}_{\Delta t} - \mathbf{u}^n_{\Delta t}}{\Delta t} = A^{t^n}\mathbf{u}^{n+1}_{\Delta t},$$

where the choice of $A^{t^{n+1}}$ or A^{t^n} depends on the problem at hand. For example, if the motion of the domain $\Omega(t)$ is specified *a priori*, $A^{t^{n+1}}$ is typically used, where $A^{t^{n+1}}$ describes an approximation of the spatial differential operator defined on the "current" domain $\Omega(t^{n+1})$. If the motion of the domain $\Omega(t)$ is not know *a priori*, but it depends on the solution of the underlying problem, then A^{t^n} is typically used, where A^{t^n} describes an approximation of the spatial differential operator defined on the "previous" domain $\Omega(t^n)$.

Functions $\mathbf{u}_{\Delta t}$ are defined for all $t \in (0,T)$ and they are piecewise constant on the interval $((n-1)\Delta t, n\Delta t]$, where the constant is defined by its value at $n\Delta t$. This approach to solving the evolution problem (2.25) is sometimes called the Rothe's method.

Rothe's method provides a constructive proof which uses semi-discretization of the continuous problem with respect to time to design approximate solutions $\{\mathbf{u}_{\Delta t}\}$ where $\Delta t = T/N$, for every $N \in \mathbb{N}$. It was first used by Ladyzhenskaya in [88], for a moving-boundary problem in which the motion of a *smooth* moving boundary was known *a priori*. The aim is to prove the existence of a sub-sequence of $\{\mathbf{u}_{\Delta t}\}$ which converges to a weak solution of (2.25) as $\Delta t \to 0$, or equivalently, as $N \to \infty$. Since the problem is nonlinear, weak convergence is not sufficient to show that the limit is a weak solution of the underlying problem. This is why compactness arguments

need to be employed to conclude that there exists a sub-sequence $\{\mathbf{u}_{\Delta t}\}$, which is precompact in a certain generalized Bochner space $L^2(0,T;H(t))$. This will allow passage to the limit in nonlinear terms and show that the limit, as $\Delta t \to 0$, of approximate weak solutions satisfies the weak formulation of the continuous problem.

Employing this strategy to prove the existence of weak solutions to this class of problems is highly nontrivial, and is at the center of the current research in this area [96, 99, 98, 100]. The main source of difficulties is associated with the fact that for every $N \in \mathbb{N}$ and $n \in \{1,\dots,N\}$, the approximate weak solutions $\mathbf{u}^n_{\Delta t}$, which are functions of the spatial variables, belong to different solution spaces $V^n_{\Delta t}$, which are associated with the operators $A^{t^n} : V^n_{\Delta t} \to W^n_{\Delta t}$, and are defined on different domains $\Omega(t^n_{\Delta t})$, thus $V^n_{\Delta t} = V(\Omega(t^n_{\Delta t}))$. We would like to find the conditions under which $\{\mathbf{u}_{\Delta t}\}$ is precompact in some $L^2(0,T;H(\Omega_{\Delta t}(t)))$, where the definition of $L^2(0,T;H(\Omega_{\Delta t}(t)))$ needs to be made precise. Namely, we want to find the conditions under which there exists a sub-sequence, also denoted by $\{\mathbf{u}_{\Delta t}\}$, which converges in $L^2(0,T;H(\Omega_{\Delta t}(t)))$ to a function in $L^2(0,T;H(\Omega(t)))$, as $\Delta t \to 0$.

There are two ways how to make the notion of convergence in $L^2(0,T;H(\Omega_{\Delta t}(t)))$ precise. One way is to introduce a family of *mappings*, which map the domains $\Omega(t^n_{\Delta t})$ onto a fixed domain Ω, and work in the space $L^2(0,T;H(\Omega))$. The other approach is to *extend* the functions $\mathbf{u}^n_{\Delta t}$ onto a larger, fixed domain Ω^M, and work in the space $L^2(0,T;H(\Omega^M))$. In both cases, certain conditions describing the regularity in time of the domain motion need to be satisfied, in order for a compactness argument to hold. In [101] those conditions are identified, and a generalization of the Aubin-Lions-Simon compactness result was obtained, which can be used in both approaches, mentioned above.

The compactness result from [101] was applied to study existence of solutions to FSI with Koiter shell [96, 99], to FSI involving multi-layered structures [98], and to FSI with Koiter shell and Navier-slip coupling [100]. Since the result is based on the backward Euler time discretization approaches to the coupled FSI problem, the compactness result from [101] is a promising tool for proving convergence of numerical schemes that use the backward Euler scheme to discretize the problem in time [19, 20, 21, 22].

3. FSI with rigid bodies

Benchmark problem. We consider an FSI problem between an incompressible, viscous fluid and a motion of a solid in a fluid container. See Fig. 5.

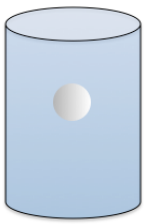


FIGURE 5. A sketch of a falling rigid ball in a fluid container.

We denote by Ω the region corresponding to the fluid container, and we denote by $S_0 \subset \Omega$ a solid (rigid structure) in the container, and by $\boldsymbol{q}_0$ its center of mass. The motion of a rigid body is fully described by the translation of its center of mass, and by the (rigid) rotation about the center of mass, described by the functions

$$\boldsymbol{q} : [0,T] \to \mathbb{R}^3 \text{ and } \boldsymbol{Q} : [0,T] \to SO(3),$$

where $SO(3)$ is the group of rotations in 3D, and $\boldsymbol{q}$ describes the translation of the center of mass. The trajectory of all the points in the rigid body is described by:

$$\boldsymbol{x} = \boldsymbol{B}(t,\boldsymbol{y}) = \boldsymbol{q}(t) + \boldsymbol{Q}(t)\ (\boldsymbol{y} - \boldsymbol{q}(0)),\ \boldsymbol{y} \in S_0,\ t \in [0,T].$$

At time t, the solid occupies the set

$$S(t) = \{\boldsymbol{x} \in \Omega_0 |\ \boldsymbol{x} = \boldsymbol{B}(t,\boldsymbol{y}), \boldsymbol{y} \in S_0 = \boldsymbol{B}(t,S_0),$$

defining the fluid domain at time t, $\Omega_F(t) = \Omega \setminus \overline{S(t)}$. As in the case of elastic structures, the fluid domain is not known *a priori* since it depends on the structural unknowns in the problem, namely the location of the rigid solid at time t.

As before, the fluid flow is described by the Navier-Stokes equations for an incompressible, viscous Newtonian fluid (2.1), while the equations of motion of the rigid body are given by a system of six ordinary differential equations (Euler equations) describing the conservation of linear and angular momentum:

$$\begin{aligned} m\frac{d^2\boldsymbol{q}}{dt^2} &= \boldsymbol{f}, \\ \frac{d(\boldsymbol{J}\boldsymbol{\omega})}{dt} &= \boldsymbol{g}, \end{aligned} \tag{3.1}$$

where m is the mass of the rigid body, $\boldsymbol{\omega}$ corresponds to angular velocity, $\boldsymbol{J}$ is the inertial tensor, and $\boldsymbol{f}$ and $\boldsymbol{g}$ are the total force and torque acting on the rigid body, respectively. The inertial tensor is defined by:

$$\boldsymbol{J} = \int_{S(t)} \rho_S \left(|\boldsymbol{x} - \boldsymbol{q}(t)|^2 \boldsymbol{I} - (\boldsymbol{x} - \boldsymbol{q}(t)) \otimes (\boldsymbol{x} - \boldsymbol{q}(t)) \right) d\boldsymbol{x},$$

where ρ_S is the structure density.

The coupling. The fluid and structure are coupled through two sets of coupling conditions: the kinematic and dynamics coupling conditions. For the *kinematic coupling condition* we take the no-slip, which says that the trace of the fluid velocity $\mathbf{u}$ at the rigid body boundary is equal to the velocity $\boldsymbol{u}_S$ of the rigid body itself:

$$\mathbf{u}(t,\boldsymbol{x}) = \boldsymbol{u}_S(t,\boldsymbol{x}),\ \boldsymbol{x} \in \partial S(t),\ t \in (0,T),$$

where $\boldsymbol{u}_S$ is given by:

$$\boldsymbol{u}_S = \frac{d\boldsymbol{x}}{dt} = \frac{d}{dt}\left(\boldsymbol{q}(t) + \boldsymbol{Q}(t)\ (\boldsymbol{y} - \boldsymbol{q}(0))\right) = \boldsymbol{q}'(t) + \boldsymbol{\omega}(t) \times (\boldsymbol{y} - \boldsymbol{q}(0)).$$

The *dynamic coupling condition* describes the balance of forces and torque. It says that the motion of the rigid structure in the fluid is driven by the contact force exerted by the fluid onto the structure. More precisely, the force and torque $\boldsymbol{f}$ and $\boldsymbol{g}$ in (3.1) are replaced by:

$$\boldsymbol{f} = -\int_{\partial S(t)} \boldsymbol{\sigma}\mathbf{n} ds(\boldsymbol{x}), \quad \boldsymbol{g} = -\int_{\partial S(t)} (\boldsymbol{x} - \boldsymbol{q}(t)) \times \boldsymbol{\sigma}\mathbf{n} ds(\boldsymbol{x}).$$

Thus, the **benchmark nonlinear moving-boundary problem**, describing fluid-structure interaction between an incompressible, viscous fluid and a rigid solid immersed in the fluid, can be summarized as follows: Find $(\mathbf{u}, p, \boldsymbol{q}, \boldsymbol{\omega})$ such that

$$\left.\begin{array}{rcl}\rho_f\big(\partial_t\mathbf{u}+(\mathbf{u}\cdot\nabla)\mathbf{u}\big) &=& \nabla\cdot\boldsymbol{\sigma}(\mathbf{u},p)\\ \nabla\cdot\mathbf{u} &=& 0\end{array}\right\}\ \text{in } \Omega_F(t),\ t\in(0,T),$$

$$\mathbf{u}=\boldsymbol{q}'(t)+\boldsymbol{\omega}(t)\times(\boldsymbol{x}-\boldsymbol{q}(0))\ \text{on}\ (0,T)\times\partial S(t),$$

$$\left.\begin{array}{rcl} m\dfrac{d^2\boldsymbol{q}}{dt^2} &=& -\int_{\partial S(t)}\boldsymbol{\sigma}\mathbf{n}ds(\boldsymbol{x})\\ \dfrac{d(\boldsymbol{J\omega})}{dt} &=& -\int_{\partial S(t)}(\boldsymbol{x}-\boldsymbol{q}(t))\times\boldsymbol{\sigma}\mathbf{n}ds(\boldsymbol{x})\end{array}\right\}\ \text{in } (0,T),$$

$$\mathbf{u}=\mathbf{0}\quad\text{on } \partial\Omega$$

$$\mathbf{u}(0,.)=\mathbf{u}_0,\ \text{in } \Omega,\ \boldsymbol{q}(0,.)=\boldsymbol{q}_0,\ \boldsymbol{q}'(0,.)=\boldsymbol{a}_0,\ \boldsymbol{\omega}(0,.)=\boldsymbol{\omega}_0.$$

Weak solutions. A weak solution is a function $(\mathbf{u},\boldsymbol{B})$ which satisfies the following two conditions:

1. The function $\boldsymbol{B}(t,\cdot):\mathbb{R}^3\to\mathbb{R}^3$, which defines the time-dependent set $S(t)=\boldsymbol{B}(t,S)$ and the corresponding Eulerian velocity $\boldsymbol{u}_S$, is an orientation preserving isometry;
2. The function $\mathbf{u}\in L^2(0,T;V(t))\cap L^\infty(0,T;L^2(\Omega)$ satisfies:

$$\int_0^T\int_{\Omega\setminus\partial S(t)}\{\mathbf{u}\cdot\partial_t\boldsymbol{\psi}+(\mathbf{u}\otimes\mathbf{u}):\mathbf{D}(\mathbf{u}):\mathbf{D}(\boldsymbol{\psi})\}\,d\boldsymbol{x}\,dt-\int_\Omega\mathbf{u}(T)\boldsymbol{\psi}(T)d\boldsymbol{x}$$
$$=-\int_\Omega\mathbf{u}_0\boldsymbol{\psi}(0)d\boldsymbol{x},$$

for all test functions $\boldsymbol{\psi}\in H^2(0,T;V(t))$, where

$$V(t)=\{\mathbf{v}\in H^1_0(\Omega)\mid \text{div } \mathbf{v}=0,\mathbf{D}(\mathbf{v})=0 \text{ in } S(t)\}.$$

Recent Results and Open Problems. Fluid-structure interaction between an incompressible, viscous fluid and an rigid, immersed structure has been extensively studied within the last twenty years. In particular, existence of a unique, local-in-time (or small data) **strong** solution is now known in both two and three space dimensions, and for both the slip [1, 120] and the no-slip coupling [36, 63, 91, 118].

In terms of **weak** solutions of Leray-Hopf type, existence up to collision was obtained by Gérard-Varet and Hillairet in [60] for the slip coupling, and more recently by Chemetov and Nečasová in [26], where they showed global-in-time existence including collision, assuming the Navier slip condition prescribed at the solid boundary, and no-slip at the container boundary. Global existence with the no-slip coupling was established in the works [32, 37, 38, 74, 114].

The question of uniqueness of weak solution is still largely **open**. Even for the classical case of the 3D Navier-Stokes equations the uniqueness of the Leray-Hopf weak solution is an outstanding open problem (see e.g. [56]). However, there are classical results of weak-strong uniqueness type (see e.g. [56, 115, 119]) which state that the strong solution (defined in an appropriate way) is unique in a larger class of weak solutions. For the Navier-Stokes equations the weak solutions that satisfy Serrin's conditions are regular [115]. In the most recent paper by Muha, Nečasová and Radošević [102], these classical weak-strong uniqueness type results

are extended to the case of a fluid-rigid body system under the condition that the rigid body does not touch the boundary of the container. Namely, in the case of contact it has been shown that weak solutions are not unique, see [49, 117], because there are multiple ways of extending the solution after the contact.

While these results discuss uniqueness of strong solutions, the results on uniqueness of weak solutions are sparse. The principal difficulty lies in the fact that different solutions are defined on different domains, so comparison between solutions is difficult. Usually, to handle this difficulty, the problem is mapped onto a fixed domain using a mapping that depends on the regularity of solutions, so strong solutions are easier to deal with. In 2015, however, uniqueness of weak solution for a fluid-rigid body system in the 2D case was obtained by Glass and Sueur in [61] for the no-slip case, and by Bravin in 2019 for the slip case [17], while uniqueness results of weak-strong type were recently published in [27, 40, 54]. In [40] the authors studied a rigid body with a cavity filled with fluid, while in [27] a rather high regularity for strong solutions was required for the uniqueness result to hold (the time derivative and second spatial derivatives of the fluid velocity were required to be in L^2). In [54] the authors studied a rigid body with a cavity filled with a compressible fluid, and showed a weak-strong uniqueness property using a relative entropy inequality. The most resent result by Muha, Nečasová and Radošević [102] generalizes these results, since they prove, for both slip and the no-slip case, a generalization of the well-known weak-strong uniqueness result for the Navier-Stokes equations to the fluid-rigid body system. More precisely, they prove that weak solutions which additionally satisfy the Prodi-Serrin $L^r - L^s$ condition are unique in the class of Leray-Hopf weak solutions.

Finite-time contact. As already addressed in the works related to global-in-time existence of weak solutions mentioned above, global existence of solutions to FSI problems involving incompressible, viscous fluids is affected by the possibility of contact: either the contact between rigid bodies immersed in the fluid, a contact between elastic structures immersed in the fluid, or the contact between an elastic structure with the fluid container (rigid) boundary. While in the case of compressible fluids, contact of rigid bodies is possible in finite time [48], the incompressible, viscous case is different since contact in finite time between rigid bodies is not possible for the scenarios studied in [92, 32, 67, 74, 77, 58, 60, 59]. In a pioneering work in 2009 Nestupa and Penel showed that contact in finite time between rigid bodies immersed in a viscous, incompressible fluid is possible if the Navier-slip boundary condition is used, which became a precursor for a number of existence results involving a slip condition and FSI with rigid solids, described above.

Finite time contact involving elastic structures remains to be an outstanding **open** problem, although a recent result by Grandmont and Hillairet in 2016 [66], indicates that finite time contact with the no-slip condition and deformable structure is impossible. More precisely, Grandmont and Hillairet studied the interaction between a 1D *viscoelastic beam* and a 2D viscous, incompressible fluid, assuming the no-slip coupling, and showed (1) that contact in finite time is not possible, and (2) that strong solutions exist globally in time. Their result is the first no-contact result involving deformable solids, and the first global existence result for FSI problems with an incompressible, viscous fluid and deformable structures.

References

1. H.A. Baba, N.V. Chemetov, Š. Nečasova, and B. Muha. *Strong solutions in L2 framework for fluid-rigid body interaction problem. Mixed case.* Topol. Methods Nonlinear Anal., 52(1):337-350, 2018.
2. A. Alphonse and C.M. Elliott. *A stefan problem on an evolving surface.* Phylosophical Transactions Royal Society A, 373:20140279, 2015.
3. A. Alphonse and C.M. Elliott. *Well-posedness of a fractional porous medium equation on an evolving surface.* Nonlinear Analysis, 137:3–42, 2016.
4. A. Alphonse, C.M. Elliott, and B. Stinner. *An abstract framework for parabolic PDEs on evolving spaces.* Portugaliae Mathematica, 72(1):1–46, 2015.
5. A. Alphonse, C.M. Elliott, and B. Stinner. *On some linear parabolic PDEs on moving hypersurfaces.* Interfaces and Free Boundaries, 17:157–187, 2015.
6. A. Alphonse, C.M. Elliott, and J. Terra. *A coupled ligand-receptor bulk-surface system on a moving domain: Well-posedness, regularity, and convergence to equilibrium.* SIAM J Math Anal, 50:1544–1592, 2018.
7. J-P. Aubin. *Un théoreme de compacité.* CR Acad. Sci. Paris, 256(24):5042–5044, 1963.
8. M. Astorino, J.-F. Gerbeau, O. Pantz, and K.-F. Traoré. *Fluid–structure interaction and multi-body contact: application to aortic valves.* Computer Methods in Applied Mechanics and Engineering, 198(45):3603–3612, 2009.
9. F.P.T. Baaijens. *A fictitious domain/mortar element method for fluid-structure interaction.* International Journal for Numerical Methods in Fluids, 35:743–761, 2001.
10. V. Barbu, Z. Grujić, I. Lasiecka, and A. Tuffaha. *Existence of the energy-level weak solutions for a nonlinear fluid-structure interaction model.* In Fluids and waves, volume 440 of *Contemp. Math.*, pages 55–82. Amer. Math. Soc., Providence, RI, 2007.
11. V. Barbu, Z. Grujić, I. Lasiecka, and A. Tuffaha. *Smoothness of weak solutions to a nonlinear fluid-structure interaction model.* Indiana Univ. Math. J., 57(3):1173–1207, 2008.
12. John W. Barrett and Endre Süli. Reflections on Dubinskiĭ's nonlinear compact embedding theorem. Publ. Inst. Math. (Beograd) (N.S.), 91(105):95–110, 2012.
13. S. Basting, A. Quaini, R. Glowinski, and S. Canic. *An extended ALE method for fluid-structure interaction problems with large structural displacements.* J. Comput. Phys., 331:312–336, 2017.
14. H. Beirão da Veiga. *On the existence of strong solutions to a coupled fluid-structure evolution problem.* J. Math. Fluid Mech., 6(1):21–52, 2004.
15. M. Boulakia. *Existence of weak solutions for the motion of an elastic structure in an incompressible viscous fluid.* C. R. Math. Acad. Sci. Paris, 336(12):985–990, 2003.
16. M. Boulakia. *Existence of weak solutions for an interaction problem between an elastic structure and a compressible viscous fluid.* J. Math. Pures Appl. (9), 84(11):1515-1554, 2005.
17. M. Bravin. *Energy Equality and Uniqueness of Weak Solutions of a "Viscous Incompressible Fluid + Rigid Bod" System with Navier Slip-with-Friction Conditions in a 2D Bounded Domain.* J. Math. Fluid Mech., 21(2):21:23, 2019.
18. D. Breit and S. Schwarzacher. *Compressible fluids interacting with a linear-elastic shell.* Archive for Rational Mechanics and Analysis, Nov 2017.
19. M. Bukač and S. Čanić. *Longitudinal displacement in viscoelastic arteries: a novel fluid-structure interaction computational model, and experimental validation.* Journal of Mathematical Biosciences and Engineering, 10(2):258–388, 2013.
20. M. Bukac, S. Canic, R. Glowinski, B. Muha, and A. Quaini. *A modular, operator-splitting scheme for fluid-structure interaction problems with thick structures.* International Journal for Numerical Methods in Fluids, 74(8):577–604, 2014.
21. M. Bukac, S. Canic, R. Glowinski, J. Tambaca, and A. Quaini. *Fluid-structure interaction in blood flow capturing non-zero longitudinal structure displacement.* J. Comput. Phys., 235:515–541, 2013.
22. M. Bukac, S. Canic, and B. Muha. *A partitioned scheme for fluid-composite structure interaction problems.* J. Comput. Phys., 281:493–517, 2015.
23. S. Čanić, M. Galić, B. Muha. *Analysis of a nonlinear moving-boundary, 3D fluid-mesh-shell interaction problem.* https://arxiv.org/submit/2940144/ Submitted 2019.

24. P. Causin, J. Gerbeau, and F. Nobile. *Added-mass effect in the design of partitioned algorithms for fluid-structure problems.* Comput. Methods Appl. Mech. Eng., 194(42-44):4506–4527, 2005.
25. A. Chambolle, B. Desjardins, M. J. Esteban, and C. Grandmont. *Existence of weak solutions for the unsteady interaction of a viscous fluid with an elastic plate.* J. Math. Fluid Mech., 7(3):368–404, 2005.
26. N.V. Chemetov and Š. Nečasová. *The motion of the rigid body in the viscous fluid including collisions. Global solvability result.* Nonlinear Anal. Real World Appl., 34:416-445, 2017.
27. N.V Chemetov, Š. Nečasová, and B. Muha. *Weak-strong uniqueness for fluid-rigid body interaction problem with slip boundary condition.* Journal of Mathematical Physics, to appear.
28. X.Chen, A. Jüngel, and J.-G. Liu. *A note on Aubin-Lions-Dubinskiĭ lemmas.* Acta Appl. Math., 133:33–43, 2014.
29. C. H. A. Cheng, D. Coutand, and S. Shkoller. *Navier-Stokes equations interacting with a nonlinear elastic biofluid shell.* SIAM J. Math. Anal., 39(3):742–800, 2007.
30. C. H. A. Cheng and S. Shkoller. *The interaction of the 3D Navier-Stokes equations with a moving nonlinear Koiter elastic shell.* SIAM J. Math. Anal., 42(3):1094–1155, 2010.
31. C. Conca, J. San Martín H., and M. Tucsnak. *Motion of a rigid body in a viscous fluid.* C. R. Acad. Sci. Paris Sér. I Math., 328(6):473–478, 1999.
32. C. Conca, J. A. San Martn, and M. Tucsnak. *Existence of solutions for the equations modelling the motion of a rigid body in a viscous fluid.* Comm. Partial Differential Equations, 25(5-6):1019-1042, 2000.
33. Carlos Conca and H Jorge San Martin. *Existence of solutions for the equations modelling the motion of rigid body in a viscous fluid.* Communications in Partial Differential Equations, 25(5-6):99–110, 2000.
34. D. Coutand and S. Shkoller. *Motion of an elastic solid inside an incompressible viscous fluid.* Arch. Ration. Mech. Anal., 176(1):25–102, 2005.
35. D. Coutand and S. Shkoller. *The interaction between quasilinear elastodynamics and the Navier-Stokes equations.* Arch. Ration. Mech. Anal., 179(3):303–352, 2006.
36. P. Cumsille and T. Takahashi. *Well-posedness for the system modelling the motion of a rigid body of arbitrary form in an incompressible viscous fluid.* Czechoslovak Math. J., 58(133)(4):961–992, 2008.
37. B. Desjardins and M. J. Esteban. *Existence of weak solutions for the motion of rigid bodies in a viscous fluid.* Arch. Ration. Mech. Anal., 146(1):59–71, 1999.
38. B. Desjardins and M. J. Esteban. On weak solutions for fluid-rigid structure interaction: compressible and incompressible models. Comm. Partial Differential Equations, 25(7-8):1399-1413, 2000.
39. B. Desjardins, M. J. Esteban, C. Grandmont, and P. Le Tallec. *Weak solutions for a fluid-elastic structure interaction model.* Rev. Mat. Complut., 14(2):523–538, 2001.
40. K. Disser, G. P. Galdi, G. Mazzone, and P. Zunino. *Inertial motions of a rigid body with a cavity filled with a viscous liquid.* Archive for Rational Mechanics and Analysis, 221(1):487-526, 2016.
41. J. Donea, *Arbitrary Lagrangian-Eulerian finite element methods, in: Computational methods for transient analysis*, North-Holland, Amsterdam,1983.
42. J.Donea, A. Huerta, J. P. Ponthot, and A. Rodriguez-Ferran. *Arbitrary Lagrangian-Eulerian Method.* Encyclopedia of Computational Mathematics, Willey, 2004.
43. J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. *The gradient discretisation method.* 2016.
44. M. Dreher and A. Jüngel. *Compact families of piecewise constant functions in lp (0, t; b).* Nonlinear Analysis: Theory, Methods & Applications, 75(6):3072–3077, 2012.
45. Q. Du, M. D. Gunzburger, L. S. Hou, and J. Lee. *Analysis of a linear fluid-structure interaction problem.* Discrete Contin. Dyn. Syst., 9(3):633–650, 2003.
46. L.J. Fauci and R. Dillon. *Biofluidmechanics of reproduction.* Annual Review of Fluid Mechanics, 38:371–394, 2006.
47. C. Figueroa, I. Vignon-Clementel, K.E. Jansen, T. Hughes, and C. Taylor. *A coupled momentum method for modeling blood flow in three-dimensional deformable arteries.* Computer Methods in Applied Mechanics and Engineering, 195:5685–5706, 2006.
48. E. Feireisl. *On the motion of rigid bodies in a viscous compressible fluid.* Arch. Ration. Mech. Anal., 167(4):281–308, 2003.

49. E. Feireisl. *On the motion of rigid bodies in a viscous fluid.* Appl. Math., 47(6):463-484, 2002. Mathematical theory in fluid mechanics (Paseky, 2001).
50. E. Feireisl, M. Hillairet, and Š. Nečasová. *On the motion of several rigid bodies in an incompressible non-newtonian fluid.* Nonlinearity, 21(6):1349, 2008.
51. M.A. Fernández. *Incremental displacement-correction schemes for incompressible fluid-structure interaction : stability and convergence analysis.* Numer. Math., 2012. DOI : 10.1007/s00211-012-0481-9.
52. M. A. Fernández, J-F. Gerbeau, and C. Grandmont. *A projection semi-implicit scheme for the coupling of an elastic structure with an incompressible fluid.* Internat. J. Numer. Methods Engrg., 69(4):794-821, 2007.
53. H. Fujita and N. Sauer. *On existence of weak solutions of the Navier-Stokes equations in regions with moving boundaries.* J. Fac. Sci., Univ. Tokyo, Sect. I, 17:403–420, 1970.
54. G. P. Galdi, V. Mácha, Š. Nečcasová. *On weak solutions to the problem of a rigid body with a cavity filled with a compressible fluid, and their asymptotic behavior.* arXiv:1903.01453
55. G. P. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations. Vol. I*, volume 38 of Springer Tracts in Natural Philosophy. Springer-Verlag, New York, 1994. Linearized steady problems.
56. G. P. Galdi. *An introduction to the Navier-Stokes initial-boundary value problem.* In Fundamental directions in mathematical fluid mechanics, Adv. Math. Fluid Mech., pages 1-70. Birkhäuser, Basel, 2000.
57. G. P. Galdi. *Mathematical problems in classical and non-Newtonian fluid mechanics.* In Hemodynamical flows, volume 37 of Oberwolfach Semin., pages 121–273. Birkhäuser, Basel, 2008.
58. D. Gérard-Varet and M. Hillairet. *Computation of the drag force on a sphere close to a wall: the roughness issue.* ESAIM Math. Model. Numer. Anal., 46(5):1201-1224, 2012.
59. D. Gérard-Varet, M. Hillairet, and C. Wang. *Influence of boundary conditions on the contact problem in a 3d incompressible flow.* http://hal. archives-ouvertes.fr/hal-00795366, 2013.
60. D. Gérard-Varet and M. Hillairet. *Existence of weak solutions up to collision for viscous fluid-solid systems with slip.* Comm. Pure Appl. Math., 67(12):2022-2075, 2014.
61. O. Glass and F. Sueur. *Uniqueness results for weak solutions of two-dimensional fluid-solid systems.* Arch. Ration. Mech. Anal., 218(2):907-944, 2015.
62. R. Glowinski. *Finite element methods for incompressible viscous flow*, in: P.G.Ciarlet, J.-L.Lions (Eds), Handbook of numerical analysis, Volume 9. North-Holland, Amsterdam, 2003.
63. M. Geissert, K. Götze, and M. Hieber. *Lp-theory for strong solutions to fluid-rigid body interaction in Newtonian and generalized Newtonian Fluids.* Trans. Amer. Math. Soc., 365(3):1393-1439, 2013.
64. C. Grandmont, M. Lukáčová-Medvidová, Š. Nečasová. *Mathematical and numerical analysis of some FSI problems. In Fluid-structure interaction and biomedical applications.* (T. Bodnár, G. P. Galdi, Š. Nečasová editors) Advances in Mathematical Fluid Mechanics. Birkhäuser 2014.
65. C. Grandmont. *Existence of weak solutions for the unsteady interaction of a viscous fluid with an elastic plate.* SIAM J. Math. Anal., 40(2):716–737, 2008.
66. Céline Grandmont and Matthieu Hillairet. *Existence of global strong solutions to a beam-fluid interaction system.* Arch. Ration. Mech. Anal., 220(3):1283–1333, 2016.
67. C. Grandmont and Y. Maday. *Existence for an unsteady fluid-structure interaction problem.* M2AN Math. Model. Numer. Anal., 34(3):609-636, 2000.
68. B. Griffith. *On the volume conservation of the immersed boundary method.* Communications in Computational Physics, 12:401–432, 2012.
69. B. Griffith, R. Hornung, D. McQueen, and C. Peskin. *An adaptive, formally second order accurate version of the immersed boundary method.* Journal of Computational Physics, 223:10–49, 2007.
70. B.E. Griffith. *Immersed boundary model of aortic heart valve dynamics with physiological driving and loading conditions.* International Journal for Numerical Methods in Biomedical Engineering, 28(3):317–345, 2012.
71. B.E. Griffith, X.Y. Luo, D.M. McQueen, and C.S. Peskin. *Simulating the fluid dynamics of natural and prosthetic heart valves using the immersed boundary method.* International Journal of Applied Mechanics, 1(1):137–177, 2009.
72. G. Guidoboni, R. Glowinski, N. Cavallini, and S. Čanić. *Stable loosely-coupled-type algorithm for fluid-structure interaction in blood flow.* J. Comput. Phys., 228(18):6916–6937, 2009.

73. G. Guidoboni, M. Guidorzi, and M. Padula. *Continuous dependence on initial data in fluid-structure motions.* J. Math. Fluid Mech., 14(1):1–32, 2012.
74. M. D. Gunzburger, H.-C. Lee, and G. A. Seregin. *Global existence of weak solutions for viscous incompressible flows around a moving rigid body in three dimensions.* J. Math. Fluid Mech., 2(3):219-266, 2000.
75. M. Hillairet. *Lack of collision between solid bodies in a 2D incompressible viscous flow.* Communications in Partial Differential Equations, 32(7-9):1345–1371, 2007.
76. M. Hillairet and T. Takahashi. *Collisions in three-dimensional fluid structure interaction problems.* SIAM Journal on Mathematical Analysis, 40(6):2451–2477, 2009.
77. K.-H. Hoffmann and V. Starovoitov. *On a motion of a solid body in a viscous fluid. Two-dimensional case.* Adv. Math. Sci. Appl., 9(2):633-648, 1999.
78. A. Hundertmark-Zauskova, M. Lukacova-Medvidova, G. Rusnakova. *Fluid-structure interaction for shear-dependent non-Newtonian fluids. Topics in mathematical modeling and analysis.* Necas Center for Mathematical Modeling. Lecture notes, Volume 7, pp. 109-158, 2012.
79. A. Hundertmark-Zaušková, M. Lukáčová-Medviďová, and Š. Nečasová. *On the existence of weak solution to the coupled fluid-structure interaction problem for non-newtonian shear-dependent fluid.* Journal of the Mathematical Society of Japan, 68(1):193–243, 2016.
80. T.J.R. Hughes, J.A. Cottrell, and Y. Bazilevs. *Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement.* Computer Methods in Applied Mechanics and Engineering 194 (39-41) 4135-4195, 2005
81. T. Hughes, W. Liu, and T. Zimmermann. *Lagrangian-Eulerian finite element formulation for incompressible viscous flows.* Computer Methods in Applied Mechanics and Engineering, 29(3):329–349, 1981.
82. M. Ignatova, I. Kukavica, Ir. Lasiecka, and A. Tuffaha. *On well-posedness for a free boundary fluid-structure model.* J. Math. Phys., 53(11):115624, 13, 2012.
83. M. Ignatova, I. Kukavica, I. Lasiecka, and A. Tuffaha. *On well-posedness and small data global existence for an interface damped free boundary fluid–structure model.* Nonlinearity, 27(3):467, 2014.
84. M. Krafczyk, J. Tolke, E. Rank, and M. Schulz. *Two-dimensional simulation of fluid-structure interaction using lattice-Boltzmann methods.* Computers & Structures, 79:2031–2037, 2001.
85. I. Kukavica and A. Tuffaha. *Solutions to a fluid-structure interaction free boundary problem.* DCDS-A, 32(4):1355–1389, 2012.
86. I. Kukavica and A. Tuffaha. *Well-posedness for the compressible Navier-Stokes-Lamé system with a free interface.* Nonlinearity, 25(11):3111-3137, 2012.
87. I. Kukavica, A. Tuffaha, and M. Ziane. *Strong solutions for a fluid structure interaction system.* Adv. Differential Equations, 15(3-4):231–254, 2010.
88. O.A. Ladyzhenskaya. *Initial-boundary problem for Navier-Stokes equations in domains with time-varying boundaries.* In Boundary Value Problems of Mathematical Physics and Related Aspects of Function Theory, pages 35–46. Springer, 1970.
89. D. Lengeler and M. Ružička. *Weak Solutions for an Incompressible Newtonian Fluid Interacting with a Koiter Type Shell.* Arch. Rational Mech. Anal. 211, 205-255 (2014)
90. J. Lequeurre. *Existence of strong solutions to a fluid-structure system.* SIAM J. Math. Anal., 43(1): 389–410, 2011.
91. D. Maity and M. Tucsnak. *Lp-Lq maximal regularity for some operators associated with linearized incompressible fluid-rigid body problems. In Mathematical analysis in fluid mechanics: selected recent results.* Volume 710 of Contemp. Math., pages 175-201. Amer. Math. Soc., Providence, RI, 2018.
92. J. Malek, J. Necas, M. Rokyta, and M. Ruzicka. *Weak and measure-valued solutions to evolutionary PDEs.* Chapman and Hall, London, 1996.
93. A. Mikelić. *Rough boundaries and wall laws.* Qualitative properties of solutions to partial differential equations, Lecture notes of Nečas Center for Mathematical Modeling (Eds. Feireisl, P. Kaplický and J. Málek), 5:103–134, 2009.
94. A. Mikelić, Š. Nečasová, and M. Neuss-Radu. *Effective slip law for general viscous flows over an oscillating surface.* Mathematical Methods in the Applied Sciences, 36(15):2086–2100, 2013.
95. A. Moussa. *Some variants of the classical Aubin–Lions lemma.* Journal of Evolution Equations, 16(1):65–93, 2016.

96. B. Muha and S. Čanić. *Existence of a weak solution to a nonlinear fluid-structure interaction problem modeling the flow of an incompressible, viscous fluid in a cylinder with deformable walls.* Arch. Ration. Mech. Anal., 207(3):919–968, 2013.
97. B. Muha and S. Čanić *A nonlinear, 3D fluid-structure interaction problem driven by the time-dependent dynamic pressure data: a constructive existence proof.* Commun. Inf. Syst., 13(3):357–397, 2013.
98. B. Muha and S. Čanić *Existence of a solution to a fluid-multi-layered-structure interaction problem.* J. Differential Equations, 256(2):658–706, 2014.
99. B. Muha and S. Čanić *Fluid-structure interaction between an incompressible, viscous 3D fluid and an elastic shell with nonlinear Koiter membrane energy.* Interfaces Free Bound., 17(4):465–495, 2015.
100. B. Muha and S. Čanić *Existence of a weak solution to a fluid-elastic structure interaction problem with the Navier slip boundary condition.* J. Differential Equations, 260(12):8550–8589, 2016.
101. B. Muha and S. Čanić *A Generalization of the Aubin-Lions-Simon Compactness Lemma for Problems on Moving Domains.* Journal of Differential Equations, 266(12):8370–8418, October 2019.
102. B. Muha, Š. Nečasová and A. Radošević. *A uniqueness result for 3D incompressible fluid-rigid body interaction problem.* arXiv:1904.05102v1. Submitted 2019.
103. C. Navier. *Sur les lois de léquilibre et du mouvement des corps elastiques.* Mémoires de l'Académie des Sciences de l'Institut de France, 369, 1827.
104. P Nägele, M Růžička, and D Lengeler. *Functional setting for unsteady problems in moving domains and applications.* Complex Variables and Elliptic Equations, 62(1):66–97, 2017.
105. P Nägele. *Monotone operator theory for unsteady problems on non-cylindrical domains [dissertation].* Freiburg: Albert-Ludwigs-Universitt Freiburg, 2015.
106. J. Neustupa and P. Penel. *A weak solvability of the Navier-Stokes equation with Navier's boundary condition around a ball striking the wall.* In Advances in mathematical fluid mechanics, pages 385–407. Springer, Berlin, 2010.
107. C. Peskin. *Numerical analysis of blood flow in the heart.* Journal of Computational Physics, 25:220–252, 1977.
108. C. Peskin and D.M. McQueen. *A three-dimensional computational method for blood flow in the heart: I. Immersed elastic fibers in a viscous incompressible fluid.* Journal of Computational Physics, 81(2):372–405, 1989.
109. S. Piperno, C. Farhat, and B. Larrouturou. *Partitioned procedures for the transient solution of coupled aeroelastic problems. Part I: Model problem, theory and two-dimensional application.* Comp. Meth. Appl. Mech. Engrg., 124:79-112, 1995.
110. J-P. Raymond and M. Vanninathan. *A fluid-structure model coupling the Navier-Stokes equations and the Lamé system.* J. Math. Pures Appl. (9), 102(3):546-596, 2014.
111. A. Quarteroni, M. Tuveri and A. Veneziani, *Computational vascular fluid dynamics*: *problems, models and methods. Survey article*, Comput. Visual. Sci. 2 (2000), pp. 163–197.
112. A. Quarteroni and A. Valli. *Domain decomposition methods for partial differential equations.* Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, 1999. Oxford Science Publications.
113. M. Ružička. *Multipolar materials. In Workshop on the Mathematical Theory of Nonlinear and Inelastic Material Behavior* (Darmstadt, 1992), volume 239 of Bonner Math. Schriften, pages 53-64. Univ. Bonn, Bonn, 1993.
114. J. A. San Martín, V. Starovoitov, and M. Tucsnak. *Global weak solutions for the two-dimensional motion of several rigid bodies in an incompressible viscous fluid.* Arch. Ration. Mech. Anal., 161(2):113–147, 2002.
115. J. Serrin. *The initial value problem for the Navier-Stokes equations.* In Nonlinear Problems (Proc. Sympos., Madison, Wis., 1962), pages 69-98. Univ. of Wisconsin Press, Madison, Wis., 1963.
116. J. Simon. *Compact sets in the space* $L^p(0,T;B)$. Ann. Mat. Pura Appl. (4), 146:65–96, 1987.
117. V. N. Starovoitov. *On the nonuniqueness of the solution of the problem of the motion of a rigid body in a viscous incompressible fluid.* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 306(Kraev. Zadachi Mat. Fiz. i Smezh. Vopr. Teor. Funktsii. 34):199-209, 231-232, 2003.

118. T. Takahashi. *Analysis of strong solutions for the equations modeling the motion of a rigid-fluid system in a bounded domain.* Adv. Differential Equations, 8(12):1499-s1532, 2003.
119. R. Temam. Navier-Stokes equations. Theory and numerical analysis. North-Holland Publishing Co., Amsterdam, 1977. Studies in Mathematics and its Applications, Vol. 2.
120. C. Wang. *Strong solutions for the fluid-solid systems in a 2-D domain.* Asymptot. Anal., 89(3-4):263-306, 2014.
121. Y. Wang, A. Quaini, S. Čanić. *A Higher-Order Discontinuous Galerkin/Arbitrary Lagrangian Eulerian Partitioned Approach to Solving FluidStructure Interaction Problems with Incompressible, Viscous Fluids and Elastic Structures.* Journal of Scientific Computing 76(1), 481-520, 2018
122. T. Wick. *Variational-Monolithic ALE Fluid-Structure Interaction: Comparison of Computational Cost and Mesh Regularity Using Different Mesh Motion Techniques* In: Bock H., Phu H., Rannacher R., Schlöder J. (eds) Modeling, Simulation and Optimization of Complex Processes HPSC 2015. Springer, Cham.

Department of Mathematics, University of California, Berkeley
E-mail address: canics@berkeley.edu

CONVEX INTEGRATION AND FLUID TURBULENCE

VLAD VICOL

ABSTRACT. These are advanced notes for the lecture given by the author as part of the Current Events Bulletin held at the 2020 Joint Mathematics Meeting of the AMS. The lecture is based on the review paper *Convex integration and phenomenologies in turbulence* by T. Buckmaster and the author [10], where the interested reader can find further details.

Our goal is to review some of the recent developments in the field of mathematical fluid dynamics which utilize techniques that go under the umbrella name *convex integration*. In the hydrodynamical context, these ideas go back to the works [20, 22] of C. De Lellis and L. Székelyhidi Jr. A consequence of these works is the existence of rough solutions to the fluid equations which seem to defy the laws of physics. Maybe even more interesting is that these counterintuitive solutions have a number of properties that resemble predictions made by phenomenological theories of fluid turbulence: energy cascades, Reynolds stress, anomalous dissipation of energy, etc. The goal of this lecture is to highlight some of these similarities while maintaining an emphasis on rigorous mathematical statements.

1. INTRODUCTION

We experience turbulent fluids on a day to day basis. The plume rising from a lit candle starts off as smooth and well organized (laminar) and quickly becomes wildly irregular, or chaotic. The air flow around a car in motion is typically laminar around the front of the car, and becomes chaotic (turbulent) in the wake of the car. In order to appreciate the complexity of turbulent flows one needs to look no further than to the fascinating flow visualizations in van Dyke's *Album of fluid motion* [85].

Turbulent flows have received a tremendous amount of attention over the past century [19], not just in the physics literature, but also in the mathematics and engineering one. This topic is too vast to review here and we refer the reader to the books [1, 62, 38, 37]. The phenomenological theories of Reynolds, Prandtl, von Karman, Taylor, Richardson, Heisenberg, Kolmogorov, Onsager, Batchelor, and Kraichnan have been incredibly successful in making predictions about the statistics of turbulent flows (see Section 2). Nonetheless, to date we do not have a mathematically rigorous and unconditional bridge between these phenomenological theories and properties of the solutions to the underlying partial differential equations which are meant to describe the fluid: the Euler and the Navier-Stokes equations. This is why hydrodynamic turbulence is many times referred to as one of the greatest challenges at the intersection of mathematics and physics.

It is widely accepted that the fundamental set of equations governing the motion of incompressible viscous fluid flows are the *Navier-Stokes equations*, written down in the first half of the 19th century [66, 80]. In their homogenous incompressible form these equations predict the evolution of the velocity field v and scalar pressure p of the fluid by

$$\partial_t v + (v \cdot \nabla) v + \nabla p - \nu \Delta v = 0 \,, \tag{1.1a}$$

$$\operatorname{div} v = 0 \,. \tag{1.1b}$$

2010 *Mathematics Subject Classification.* Primary 35Q35.

Here $\nu > 0$ is the kinematic viscosity of the fluid. One may rewrite the nonlinear term in divergence form as $\mathrm{div}\,(v \otimes v) = (v \cdot \nabla)v$, which is important for defining distributional solutions to the system. Formally passing to the *inviscid limit* $\nu \to 0$ we arrive at the *Euler* equations, which are the classical model for the motion of an incompressible homogenous inviscid fluid. Derived in the mid 18th century [31], these equations predict the state of the unknown velocity field v and pressure field p according to

$$\partial_t v + (v \cdot \nabla)v + \nabla p = 0\,, \tag{1.2a}$$

$$\mathrm{div}\, v = 0\,. \tag{1.2b}$$

The Euler and Navier-Stokes equations are to be supplemented with an incompressible initial datum v_0, and with boundary conditions. For simplicity, we only consider here the equations posed on the three-dimensional periodic box $\mathbb{T}^3 = [-\pi, \pi]^3$, with a zero mean initial condition v_0.[1] In order to ensure a nontrivial long-time behavior, it is customary to add a zero mean forcing term f^ν to the right side of the Navier-Stokes equations (1.1a).

The fundamental ansatz of the Kolmogorov/Onsager theories of hydrodynamic turbulence, sometimes called the *zeroth law of turbuelence*, is that in the vanishing viscosity limit solutions of the Navier-Stokes equations do not remain smooth *uniformly with respect to* ν. This prediction has been verified experimentally to a tremendous degree of accuracy [38, 78, 43]. Thus, in the vanishing viscosity limit $\nu \to 0$ it is reasonable to expect that Navier-Stokes solutions converge to non-smooth (distributional), possibly non-unique, solutions of the Euler equations. In an attempt to translate predictions made by turbulence theories into mathematically rigorous questions, it is thus natural to work within the framework of distributional solutions of (1.1) and (1.2) (cf. Definition 3.1 below).

In this context, one of the most celebrated connections between phenomenologies in turbulence and the rigorous mathematical analysis of the Euler equations (1.2) is the *Onsager conjecture* [68], which postulates that weak solutions of the 3D Euler equations with regularity *above* $1/3$ conserve the kinetic energy (the *rigidity* part), while for any regularity level *below* $1/3$ there exists weak solutions which dissipate the kinetic energy (the *flexibility* part). We defer a precise statement to Conjecture 3.2 below. While the rigid part of the conjecture was understood since the mid 1990's due to works of Eyink [33] and Constatin-E-Titi [15], significant progress towards the resolution of the flexible part of this conjecture did not occur until the 2010's and the groundbreaking works [20, 22] of De Lellis and Székelyhidi Jr. These works have developed the mathematical framework and have laid out some of the key ideas which have eventually led to the solution of the flexible part of the Onsager conjecture by Isett [42] (in the context of solutions with compact support in time); and in a subsequent work by Buckmaster-De Lellis-Székelyhidi-V. [9] (for dissipative weak solutions).

The goal of this lecture is to discuss the results in the program of De Lellis, Székelyhidi Jr., and collaborators, and to present some of the analogies between the mathematical machinery which they have developed and counterintuitive constructions in differential geometry on one hand, and phenomenological theories of turbulence on the other hand.

Acknowledgments. The author was partially supported by the NSF grant CAREER DMS 1911413. We *emphasize* that parts of these lecture notes closely follow the review paper *Convex integration and phenomenologies in turbulence* by T. Buckmaster and V.V. [10].

[1] Note that solutions preserve their mean, and thus we have $\int_{\mathbb{T}^3} v(x,t)dx = \int_{\mathbb{T}^3} v_0(x)dx = 0$ for all $t > 0$.

2. Aspects of the Onsager and Kolmogorov theories

While the mathematical results which we will discuss in Sections 3 and 4 do not prove that the Kolmogorov [48, 49, 47] and Onsager [68] theories are indeed manifestations of the statistical behavior of solutions to the Navier-Stokes equations in the infinite Reynolds number limit, it is important to know what these turbulence theories postulate. This will both inform what the meaningful mathematical questions about (1.1) and (1.2) are, and it will allow us to build mathematical objects that mimic the behavior of a real turbulent fluid. Indeed, the theory developed by De Lellis, Székelyhidi Jr., and collaborators has a number of similarities with fluid turbulence: the constructed weak solutions have energy cascades, they arise as a limit of a system with a Reynolds stress, and last but not least, they display an anomalous dissipation of energy. The physics-averse reader is encouraged to skip to the next section.

To fix the notation, let us denote by v^ν a solution of the Cauchy problem for the *forced* version of the Navier-Stokes equations (1.1) with viscosity[2] ν, and initial datum $v_0 \in L^2$ that is incompressible, has zero mean, and is sufficiently smooth. The forcing term f^ν is taken to have zero mean, is stationary, and injects energy into the system at low frequencies.

Given a suitable *observable* F of the solution v^ν, e.g. $F(v^\nu) = \|v^\nu\|_{L^2}^2$, theoretical physicists typically use $\langle F(v^\nu)\rangle$ to denote an *ensemble average* with respect to a putative probability measure μ^ν on L^2 which is time independent[3]: at statistical equilibrium the probability measure μ^ν encodes the macroscopic statistics of the flow. On the other hand, in laboratory experiments a measurement of the turbulent flow is usually a *long time average* at fixed viscosity, in order to reach a stationary regime. In analogy with classical statistical mechanics, turbulence theories deal with the possible discrepancy between ensemble averages and statistical averages by making an impromptu *ergodic hypothesis*. The implication of the ergodic hypothesis is that averages against an ergodic invariant measure are the same as long time averages, giving a meaning to $\langle\cdot\rangle$. We note that $\langle\cdot\rangle$ sometimes includes a spatial average over $\mathbb{T}^3$, justified under the assumption of statistical homogeneity.

2.1. Anomalous dissipation of energy.

The fundamental ansatz of Kolmogorov's 1941 theory of fully developed turbulence [48, 49, 47], sometimes called the *zeroth law of turbulence*, postulates the *anomalous dissipation of energy*: the non-vanishing of the rate of dissipation of kinetic energy of turbulent fluctuations per unit mass, in the limit of zero viscosity (cf. (2.8) below). The zeroth law of turbulence is verified experimentally to a tremendous degree [78, 69, 43], but to date we do not have a single example where it is rigorously proven to hold, directly from (1.1).

To formulate this ansatz, we start with the balance of kinetic energy in the Navier-Stokes equation. By taking an inner product of v^ν with the forced (1.1) system, if the functions v^ν are sufficiently smooth we obtain the pointwise energy balance

$$\partial_t \frac{|v^\nu|^2}{2} + \nabla\cdot\left(v^\nu\left(\frac{|v^\nu|^2}{2}+p^\nu\right) - \nu\nabla\frac{|v^\nu|^2}{2}\right) = f^\nu\cdot v^\nu - \nu|\nabla v^\nu|^2. \tag{2.1}$$

[2] We abuse notation and denote also by ν the inverse of the Reynolds number $\mathrm{Re}^{-1} = \nu/(UL)$, where $L = 2\pi$ is the characteristic length scale of the domain, and U is an average velocity. The infinite Reynolds number limit $\mathrm{Re}\to\infty$ is used interchangeably with the vanishing viscosity limit $\nu\to 0$ (for U and L fixed).

[3] Following the pioneering work of Foias [35, 36], in a deterministic setting one may consider *stationary statistical solutions* to the Navier-Stokes equation. These are probability measures on L^2 which satisfy a stationary Liouville-type equation, when integrated against cylindrical test functions. Their existence may be rigorously established using the concept of a generalized Banach limit from long time averages, but their uniqueness remains famously open. This notion of solution has been explored quite a bit in the past decades [86, 37].

Integrating over the periodic domain we obtain the kinetic energy balance

$$\frac{d}{dt} \fint_{\mathbb{T}^3} \frac{|v^\nu|^2}{2} dx = \fint_{\mathbb{T}^3} f^\nu \cdot v^\nu dx - \nu \fint_{\mathbb{T}^3} |\nabla v^\nu|^2 dx \,, \tag{2.2}$$

where the first term on the right side denotes the total work of the force and the second term denotes the energy dissipation rate per unit mass. Equation (2.2) is the only known coercive a-priori estimate for the 3D Navier-Stokes equations, and it gives an a-priori bound for the solution v^ν in the so-called *energy space* $L^\infty_t L^2_x \cap L^2_t H^1_x$. Leray [52] used the energy balance for a suitable approximating sequence, combined with a compactness argument, to prove the existence of a global in time weak solution to (1.1) which lies in $L^\infty_t L^2_x \cap L^2_t H^1_x$, and obeys (2.2) weakly in time with an inequality instead of the equality.

A posteriori one may ask the question of whether the local energy balance (2.1) may be actually justified when v^ν is a weak solution of equation (1.1) in the energy space (by interpolation v^ν also lies in $L^{10/3}_{x,t} \subset L^3_{x,t}$). To date this question remains open [54, 72, 51]. Instead, as in the work of Duchon-Robert [27], one may prove that for a weak solution v^ν in the energy class the following equality holds in the sense of distributions

$$\partial_t \frac{|v^\nu|^2}{2} + \nabla \cdot \left(v^\nu \left(\frac{|v^\nu|^2}{2} + p^\nu \right) - \nu \nabla \frac{|v^\nu|^2}{2} \right) = f^\nu \cdot v^\nu - \nu |\nabla v^\nu|^2 - D(v^\nu) \tag{2.3}$$

where the (x,t)-distribution $D(v^\nu)$ is defined by a weak form of the Kármán-Howarth-Monin relation [44, 61]

$$D(v^\nu)(x,t) = \lim_{\ell \to 0} \frac{1}{4} \int_{\mathbb{T}^3} \nabla \varphi_\ell(z) \cdot \delta v^\nu(x,t;z) |\delta v^\nu(x,t;z)|^2 dz \,. \tag{2.4}$$

In (2.4) we have denoted the *velocity increment in the direction* z by

$$\delta v^\nu(x,t;z) = v^\nu(x+z,t) - v^\nu(x,t) \tag{2.5}$$

and the approximation of the identity φ_ℓ is given by $\varphi_\ell(z) = \frac{1}{\ell^3} \varphi\left(\frac{z}{\ell}\right)$, where $\varphi \geq 0$ is an even bump function with mass equal to 1. The limit in (2.4) is a limit of $L^1_{x,t}$ objects in the sense of distributions, and it is shown in [27] that $D(v^\nu)$ is independent of the choice of φ. When compared to (2.1), identity (2.3) additionally takes into account the possible dissipation of kinetic energy, due to possible singularities of the flow v^ν, encoded in the *defect measure* $D(v^\nu)$.[4] Similarly to (2.2), once we average the local energy balance (2.3) over $\mathbb{T}^3$, the divergence term on the left side vanishes, and we are left with

$$\frac{d}{dt} \fint_{\mathbb{T}^3} \frac{|v^\nu|^2}{2} dx = \fint_{\mathbb{T}^3} f^\nu \cdot v^\nu dx - \nu \fint_{\mathbb{T}^3} |\nabla v^\nu|^2 dx - \fint_{\mathbb{T}^3} D(v^\nu) dx \,, \tag{2.6}$$

which yields a balance relation between energy input and energy dissipation.

In view of (2.6), we define the *mean energy dissipation rate per unit mass* by

$$\varepsilon^\nu = \nu \left\langle |\nabla v^\nu|^2 \right\rangle + \langle D(v^\nu) \rangle \,, \tag{2.7}$$

where as discussed before, $\langle \cdot \rangle$ denotes a suitable ensemble/long-time and a space average. The zeroth law of turbulence, or the anomalous dissipation of energy, postulates that in the inviscid limit $\nu \to 0$ the mean energy dissipation rate per unit mass does not vanish, and moreover that there exists an $\varepsilon \in (0, \infty)$ such that

$$\varepsilon = \liminf_{\nu \to 0} \varepsilon^\nu > 0. \tag{2.8}$$

[4] Note that if v^ν is sufficiently smooth to ensure that $\lim_{|z| \to 0} |z|^{-1} \int_0^T \int_{\mathbb{T}^3} |\delta v^\nu(x,t;z)|^3 dx dt = 0$, then one may directly show that $D(v^\nu) \equiv 0$. This was proven in [13].

2.2. **Aspects of the Kolmogorov ('41) theory.** Based on the anomalous dissipation of energy and certain scaling arguments, Kolmogorov [48, 49, 47] proposed a theory for homogenous isotropic turbulence, whose key predictions we summarize below. We refer the reader to [38, 62, 34, 75, 10] for further details.

Besides the zeroth law of turbulence (2.8), the assumptions of Kolmogorov's theory are *homogeneity*, *isotropy*, and *self-similarity*. Let $\widehat{z} \in \mathbb{S}^2$ be a unit direction vector and let $\ell > 0$ be a length scale in the *inertial range*, meaning that $\ell_D \ll \ell \ll \ell_I$, where ℓ_I is the integral scale of the system, and $\ell_D = \nu^{3/4}\varepsilon^{-1/4}$ is the *Kolmogorov dissipative length scale*.[5] Homogeneity is the assumption that the statistics of turbulent flows is shift invariant: at large Reynolds numbers the velocity increment $\delta v^\nu(x,t;\ell\widehat{z})$ (recall (2.5)) has the same probability distribution for every $x \in \mathbb{T}^3$. Isotropy is the assumption that the statistics of turbulent flows is locally rotationally invariant: the probability distribution for $\delta v^\nu(x,t;\ell\widehat{z})$ is the same for all $\widehat{z} \in \mathbb{S}^2$. The Kolmogorov theory postulates the existence of an exponent $h > 0$, such that $\delta v^\nu(x,t;\lambda\ell\widehat{z})$ and $\lambda^h \delta v^\nu(x,t;\ell\widehat{z})$ have the same law, for $\lambda > 0$ such that both ℓ and $\lambda\ell$ lie in the inertial range. Based on these assumptions, the theory makes predictions about structure functions and the energy spectrum.

For $p \geq 1$ one may define the p^{th} order *longitudinal structure function*

$$S_p^{\parallel}(\ell) = \langle (\delta v^\nu(x,t;\ell\widehat{z}) \cdot \widehat{z})^p \rangle \ .$$

Note that for p which is odd, $S_p^{\parallel}(\ell)$ need not a-priori have a sign. Instead, one may define the p^{th} order *absolute structure function*

$$S_p(\ell) = \langle |\delta v^\nu(x,t;\ell\widehat{z})|^p \rangle$$

which is intimately related to the definition of a Besov space.[6] $S_p(\ell)$ scales in the same way as $S_p^{\parallel}(\ell)$, and they both have physical units of U^p. Notice that since $\varepsilon\ell$ has units of U^3, it follows that $(\varepsilon\ell)^{p/3}$ has the same physical units as $S_p(\ell)$. Consequently, the only value of the self-similarity exponent which is consistent with physical units as $\ell \to 0$ is $h = 1/3$, and thus the Kolmogorov theory predicts the asymptotic behavior

$$S_p(\ell) \sim (\varepsilon\ell)^{p/3} \tag{2.9}$$

for ℓ in the inertial range, in the infinite Reynolds number limit. Denoting by ζ_p the limiting *structure function exponent*

$$\zeta_p = \lim_{\ell\to 0}\lim_{\nu\to 0} \frac{\log\left(S_p(\ell)\right)}{\log\left(\varepsilon\ell\right)} \ , \tag{2.10}$$

the relation (2.9) indicates that in Kolmogorov's theory we have

$$\zeta_p = \frac{p}{3} \ , \qquad \text{for all} \qquad p \geq 1 \ . \tag{2.11}$$

Except for $p = 3$, when the Kolmogorov prediction $\zeta_3 = 1$ *is indeed supported by all the experimental evidence*, for $p \neq 3$ experiments do indeed deviate from the Kolmogorov prediction [58, 38].

For the third order longitudinal structure function $S_3^{\parallel}$, Kolmogorov derived what is considered an *exact result in turbulence*, the famous $4/5$*-law*, which states that

$$S_3^{\parallel}(\ell) \sim -\frac{4}{5}\varepsilon\ell \tag{2.12}$$

[5] The only object which has the physical unit of length and may be written as $\nu^a\varepsilon^b$.

[6] $v \in B^s_{p,\infty}$ means that $v \in L^p$ and that $\sup_{|z|>0} |z|^{-s} \left\|\delta v(x;z)\right\|_{L^p_x} < \infty$, for $s \geq 0$ and $p < \infty$.

holds in the infinite Reynolds number limit, for $\ell \ll \ell_I$. Identity (2.12) is remarkable because a-priori, there is no good reason for the cubic power of the longitudinal increments to have a sign, on average. Moreover, in addition to claiming that $\zeta_3 = 1$, (2.12) predicts the *universal pre-factor* of $-4/5$. Compelling experimental support of the $4/5$-law is provided for instance by the measurements in [79]. From a mathematical perspective the $4/5$-law is particularly intriguing because under certain uniform in ν regularity and integrality assumptions on the sequence of solutions v^ν the $4/5$-law can be shown to hold; these conditions however have yet to be proven rigorously (only numerical evidence is available [79]). We refer the reader to the results and excellent discussions in [67, 32].

For $p = 2$, from (2.9)–(2.11) the Kolmogorov prediction yields $\zeta_2 = 2/3$. One may translate this scaling of the second order structure function into the famous $-5/3$ energy density spectrum, defined in terms of Fourier projection operators as follows. For $\kappa > 0$ the mean kinetic energy per unit mass carried by wavenumber $\leq \kappa$ in absolute value is given by $\frac{1}{2}\left\langle |\mathbb{P}_{\leq \kappa} v^\nu|^2 \right\rangle$. The energy spectrum is then defined as

$$\mathcal{E}(\kappa) = \frac{1}{2}\frac{d}{d\kappa}\left\langle |\mathbb{P}_{\leq \kappa} v^\nu|^2 \right\rangle \tag{2.13}$$

so that the total kinetic energy may be written as $\frac{1}{2}\left\langle |v^\nu|^2 \right\rangle = \int_0^\infty \mathcal{E}(\kappa) d\kappa$. The Kolmogorov prediction $\zeta_2 = 2/3$ then translates into

$$\mathcal{E}(\kappa) \sim \epsilon^{2/3} \kappa^{-5/3}, \tag{2.14}$$

for κ^{-1} in the inertial range, and in the infinite Reynolds number limit. See [38] for experimental support for (2.14). This power law requires however that velocity fluctuations are uniformly distributed over the three dimensional domain, which is not always justified.

2.3. **Aspects of the Onsager ('49) theory.** In his famous paper on statistical hydrodynamics, Onsager [68] considered the possibility that energy dissipation is not caused by the viscosity, but instead, because the solutions of the Euler equation may not be smooth.

Similarly to (2.2), the kinetic energy balance for smooth solutions v of the Euler equations (1.2) is

$$\frac{d}{dt} \fint_{\mathbb{T}^3} \frac{|v|^2}{2} dx = \fint_{\mathbb{T}^3} f \cdot v dx \,, \tag{2.15}$$

which becomes a conservation law when $f \equiv 0$. Onsager is referring to the fact that if the solution v of (1.2) is not sufficiently smooth, i.e. it is a *weak solution*, then the energy balance/conservation (2.15) cannot be justified. Onsager's remarkable analysis went further and made a precise statement about the regularity of v which is necessary in order to justify (2.15); in mathematical terms this is known as the *Onsager Conjecture* (see Conjecture 3.2 below). We refer to the review articles [34, 75] for a detailed account of the Onsager theory of *ideal turbulence*, and present here only some of the ideas (in terms of Fourier projection operators, as in Onsager's work [68]).

We regularize a weak solution v of the Euler equations (1.2), by a smooth cutoff in the Fourier variables at frequencies $\leq \kappa$, and consider the kinetic energy of $\mathbb{P}_{\leq \kappa} v$. Then, similarly to (2.15) we obtain that

$$\frac{d}{dt} \fint_{\mathbb{T}^3} \frac{|\mathbb{P}_{\leq \kappa} v|^2}{2} dx = \fint_{\mathbb{T}^3} \mathbb{P}_{\leq \kappa} f \cdot \mathbb{P}_{\leq \kappa} v dx - \Pi_\kappa \tag{2.16}$$

where as in [13] we denote by Π_κ the mean energy flux through the sphere of radius κ in frequency space, i.e.

$$\begin{aligned}\Pi_\kappa &= -\fint_{\mathbb{T}^3} \mathbb{P}_{\leq\kappa}(v\otimes v)\colon \nabla\mathbb{P}_{\leq\kappa}v dx\\ &= \fint_{\mathbb{T}^3} \Big((\mathbb{P}_{\leq\kappa}v\otimes\mathbb{P}_{\leq\kappa}v) - \mathbb{P}_{\leq\kappa}(v\otimes v)\Big)\colon \nabla\mathbb{P}_{\leq\kappa}v dx\,.\end{aligned} \tag{2.17}$$

From (2.16) we deduce upon passing $\kappa\to\infty$ that the energy balance (2.15) holds *if and only if* the *total energy flux vanishes*:

$$\Pi = \lim_{\kappa\to\infty}\Pi_\kappa\,. \tag{2.18}$$

Onsager's prediction is that in order for Π to be nontrivial, i.e. for the weak solution v to be non-conservative, it should not obey $|\delta v(x;z)|\lesssim |z|^\theta$ with $\theta > 1/3$.

We emphasize that in 3D turbulent flows the *energy transfer* from one scale/frequency to another is observed to be mainly *local*, i.e. the principal contributions to Π_κ come from $\mathbb{P}_{\approx\kappa'}v$, with $\kappa'\approx\kappa$. A rigorous estimate on the locality of the energy transfer arises in [13], where it is proven that

$$|\Pi_{2^j}|\lesssim \sum_{i=1}^\infty 2^{-2/3|j-i|}2^i\|\mathbb{P}_{\approx 2^i}v\|_{L^3}^3\,. \tag{2.19}$$

Estimate (2.19) gives the best known condition on v which ensures $\Pi = 0$: $v\in L^3_t B^{1/3}_{3,c_0,x}$ (cf. [13]). This condition is sharp in the case of the 1D Burgers equation [75].

It is not an accident that the $1/3$-derivative singularities required by Onsager for a dissipative anomaly $\Pi\neq 0$, matches Kolmogorov's assumed $1/3$ local self-similarity exponent required for $\varepsilon>0$. As already observed by Onsager [68], if v is a weak solution of the Euler equations which is a strong limit of a sequence $\{v^\nu\}$ of Navier-Stokes solutions for which the anomalous dissipation of energy (2.8) holds, then:

$$\varepsilon = \langle\Pi\rangle. \tag{2.20}$$

On the experimental side, the evidence for (2.20) is quite convincing [43]. A physics-style argument could run as follows. Let $f^\nu = f$ be statistically stationary, with $\mathbb{P}_{\kappa_I}f = f$ for some frequency κ_I. Let Π^ν_κ be the energy flux through the frequency ball of radius κ for a solution v^ν of the Navier-Stokes equation, i.e. replace v in (2.17) with v^ν. Then similarly to (2.16), since the ensemble/long-time average $\langle\cdot\rangle$ is stationary, we obtain that

$$\langle\Pi^\nu_\kappa\rangle + \nu\left\langle|\nabla\mathbb{P}_{\leq\kappa}v^\nu|^2\right\rangle = \langle f\cdot\mathbb{P}_{\leq\kappa}v^\nu\rangle \tag{2.21}$$

for $\kappa\geq\kappa_I$. On the other hand, assuming that the Euler solution is statistically stationary, (2.16) yields

$$\langle\Pi_\kappa\rangle = \langle f\cdot\mathbb{P}_{\leq\kappa}v\rangle \tag{2.22}$$

To conclude, we recall that from the definition (2.4) we have $\langle D(v^\nu)\rangle = \lim_{\kappa\to\infty}\langle\Pi^\nu_\kappa\rangle$ (cf. [27]), and with ε^ν as given by (2.7), we pass $\kappa\to\infty$ in (2.21) and (2.22), to arrive at

$$\varepsilon - \langle\Pi\rangle = \lim_{\nu\to 0}(\varepsilon^\nu - \langle\Pi\rangle) = \lim_{\nu\to 0}\langle f\cdot(v^\nu - v)\rangle = 0 \tag{2.23}$$

since we assumed $v^\nu\to v$. The energy flux thus provides a connection between the Kolmogorov and Onsager theories [38].

3. The Euler equations and the conjecture of Onsager

The local in time well-posedness of the 3D Euler equations in spaces of smooth functions (e.g. $C^{1,\alpha}$ with $0 < \alpha < 1$ or H^s with $s > 5/2$) is classical; see [57] and references therein. By the Beale-Kato-Majda criterion [2], global well-posedness holds if and only if the L^∞ norm of the vorticity is L^1 integrable in time. The global well-posedness of the 3D Euler equation is famously unresolved for C^∞ smooth initial datum; nonetheless, very recently Elgindi [28] has displayed solutions with $C^{1,\alpha}$ initial datum for $0 < \alpha \ll 1$, which blow up in finite time (see also [12, 29]).

On the other hand, as mentioned in Section 2, in the inviscid limit, turbulent solutions exhibiting a dissipation anomaly are necessarily *not smooth*, thereby motivating the study of weak solutions to the Euler equations (1.2). For the purpose of this lecture, a weak solution for (1.2) is defined as:

Definition 3.1 (Weak solution)**.** A vector field $v \in L^\infty_t L^2_x$ is called a weak solution of the Euler equations if for any t the vector field $v(\cdot, t)$ is weakly divergence free, has zero mean, and satisfies the Euler equation distributionally:

$$\int_{\mathbb{R}} \int_{\mathbb{T}^3} v \cdot (\partial_t \varphi + (v \cdot \nabla)\varphi) dx dt = 0 \,,$$

for any divergence free test function φ. The pressure can be recovered by the formula $-\Delta p = \operatorname{div} \operatorname{div} (v \otimes v)$ with p of zero mean.

In the regularity class, $L^\infty_t L^2_x$, the *kinetic energy*

$$E(t) = \fint_{\mathbb{T}^3} \frac{|v(x,t)|^2}{2} dx \tag{3.1}$$

is a well-defined bounded function of time, but it need not be the constant function. Compare this with (2.15) when $f = 0$. Note that the definition of a weak solution could be relaxed to $v \in L^2_{t,x}$; moreover, it is standard to modify the above definition in order to take into account an initial value $v_0 \in L^2$.

As discussed in Section 2, motivated by hydrodynamic turbulence, Onsager [68] conjectured the following dichotomy:

Conjecture 3.2 (Onsager's conjecture)**.**

(a) Any weak solution v belonging to the Hölder space $C^\theta_{x,t}$ for $\theta > 1/3$ conserves kinetic energy.

(b) For any $\theta < 1/3$ there exist weak solutions $v \in C^\theta_{x,t}$ which dissipate kinetic energy.

3.1. The conservative part.

Part (a) of this conjecture was partially established by Eyink in [33], and later proven in full by Constantin, E and Titi in [15]; see also Duchon-Robert [27], Cheskidov-Constantin-Friedlander-Shvydkoy [13], and the more recent work of Shvydkoy [77], for refinements.

As shown by [15], the proof follows from a simple commutator argument. For a weak solution v of the Euler equations, let v_ℓ be the spatial mollification of v a length scale ℓ. Then, v_ℓ satisfies

$$\int_{\mathbb{T}^3} \frac{|v_\ell(x,t)|^2}{2} dx - \int_{\mathbb{T}^3} \frac{|v_\ell(x,0)|^2}{2} dx = \int_0^t \int_{\mathbb{T}^3} (v^i v^j)_\ell \, \partial_i v^j_\ell \, dx ds$$

where we have denoted $v = (v^i)_{i=1}^3$, $v_\ell = (v^i_\ell)_{i=1}^3$, $\partial_i = \partial_{x_i}$, and we have used the summation convention on repeated indices. Since v_ℓ is a smooth incompressible function,

i.e. $\nabla \cdot v_\ell = 0$, uniformly in time we have that

$$\int_{\mathbb{T}^3} v_\ell^i \, v_\ell^j \, \partial_i v_\ell^j \, dx = 0 \, .$$

Subtracting the above two statements we obtain that

$$\int_{\mathbb{T}^3} \frac{|v_\ell(x,t)|^2}{2} dx - \int_{\mathbb{T}^3} \frac{|v_\ell(x,0)|^2}{2} dx = \int_0^t \int_{\mathbb{T}^3} \left((v^i \, v^j)_\ell - v_\ell^i \, v_\ell^j \right) \partial_i v_\ell^j \, dx ds \, . \tag{3.2}$$

Applying the Constantin-E-Titi [15] commutator estimate:

Proposition 3.3. *Let $f, g \in C^\infty(\mathbb{T}^3 \times [0,1])$. For any $\theta \in (0,1]$ we have*

$$\left\| (f \, g)_\ell - f_\ell \, g_\ell \right\|_{C^0} \lesssim \ell^{2\theta} \|f\|_{C^\theta} \|g\|_{C^\theta} \, ,$$

where the implicit constant is universal.

and the bound $\|\nabla v_\ell\|_{L^\infty} \lesssim \ell^{1-\theta} \, \|v\|_{C^\theta}$, we deduce from (3.2) that

$$\left| \int_{\mathbb{T}^3} \frac{|v_\ell(x,t)|^2}{2} dx - \int_{\mathbb{T}^3} \frac{|v_\ell(x,0)|^2}{2} dx \right| \lesssim \ell^{3\theta - 1} \, \|v\|_{C^\theta}^3 \, .$$

Thus, if $\theta > 1/3$, the right hand side converges to zero as $\ell \to 0$.

We note that in terms of space integrability, the above estimate is not sharp: the term on the right side of (3.2) is trilinear, and thus should only require that $v \in L_t^3 B^\theta_{3,\infty,x}$, where $B^\theta_{3,\infty}$ denotes the suitable Besov space; this is in fact the statement proven in [15].

A further refinement of this proof was obtained by Cheskidov-Constantin-Friedlander-Shvydkoy [13]. As discussed in Section 2, energy conservation holds if and only if the total energy flux Π_κ defined in (2.17) vanishes as $\kappa \to \infty$, when integrated in time. We note that here κ plays a similar role to ℓ^{-1} in the previous proof. Using a Bony paraproduct decomposition, the authors of [13] prove the detailed bound (2.19) which then immediately shows that if $v \in L_t^3 B^{1/3}_{3,c_0,x}$, then energy conservation holds. This is the strongest known result pertaining to Part (a), the rigid side, of the Onsager conjecture.

3.2. **The flexible part: first paradoxical examples.** In the seminal work [71], Scheffer demonstrated the existence of non-trivial weak solutions of the 2D Euler system (1.2), which lie in $L^2_{x,t}$ and have compact support in time and space! Strictly speaking the weak solutions of Scheffer are not dissipative, as dissipative solutions are required to have non-increasing energy; nonetheless, [71] is considered to be the first result concerning the flexible part, Part (b), of the Onsager conjecture. A different construction of a nontrivial weak solution to the 2D Euler equations, which are periodic in space and have compact support in time, was given by Shnirelman in [74]. The existence of dissipative weak solutions to the Euler equations was first proven by Shnirelman in [73], where he constructions weak solutions which lie in $L^\infty_t L^2_x$.

These results, which were initially referred to as *the Scheffer-Shnirelman paradox*, represent not just a proof of non-uniqueness for weak solutions to the Euler equations, but a drastic failure of determinism within this class of solutions. The proofs in [71, 74, 73] are rather involved, to say the least, they seem quite ad-hoc, and there is no clear path by which one may improve the spatial regularity of the obtained weak solutions, above L^2_x.

3.3. **The $L^\infty_{x,t}$ results: convex integration.** The first example of a bounded (in space and time), dissipative weak solution of the Euler equations (in any dimension $n \geq 2$) was obtained in a groundbreaking work by De Lellis-Székelyhidi Jr. [20]. Their main result is:

Theorem 3.4 (Theorem 4.1 in [20])**.** *For any open bounded space-time domain $\Omega \subset \mathbb{R}^n \times \mathbb{R}$, there exists a weak solution of the Euler equations (1.2) $(v,p) \in L^\infty(\mathbb{R}^n \times \mathbb{R})$, in the sense of Definition 3.1, such that $|v(x,t)| = 1$ for a.e. $(x,t) \in \Omega$, and $v(x,t) = p(x,t) = 0$ for a.e. $(x,t) \in \Omega^c$. Moreover, there exists a sequence of functions $(v_q, p_q, f_q) \in C^\infty_0(\Omega)$ such that:*

- $\partial_t v_q + \operatorname{div}(v_q \otimes v_q) + \nabla p_q = f_q$, *and* $\nabla \cdot v_q = 0$
- $f_q \to 0$ *in* H^{-1} *as* $q \to \infty$
- $\|v_q\|_{L^\infty} + \|p_q\|_{L^\infty}$ *is uniformly bounded in* q
- $(v_q, p_q) \to (v,p)$ *in* L^r *for any* $r < \infty$.

The first part of the theorem establishes the existence of a weak solution which is compactly supported in space and time, while the second part is a manifestation of the proof: the limiting weak solution (v,p) is obtained from a *smooth* approximating sequence (v_q, p_q), which solves a *relaxed Euler system*, whose right side f_q vanishes in a weak sense as $q \to \infty$. This paper also introduced the ideas of a subsolution, and a *Reynolds stress* for the Euler system (1.2). Maybe more important than the result itself, which was later improved by the same authors, is the fact that [20] relates the construction of paradoxical weak solutions of the Euler equations with a classical technique in geometry, *convex integration*, and the notion of *h-principles* for soft partial differential equations.

The method of convex integration can be traced back to the work of Nash, who used it to construct exotic counter-examples to the C^1 isometric embedding problem [65]. The method was later refined by Gromov [39] and it evolved into a general method for solving *soft/flexible* geometric partial differential equations [30]. In the influential paper [64], Müller and Šverák adapted convex integration to the theory of differential inclusions, see also [45], leading to renewed interest in the method as a result of its greatly expanded applicability. Inspired by the works [64, 45], and building on the plane-wave analysis introduced by Tartar [83, 84] and Di Perna [26], De Lellis and Székelyhidi Jr., in [20], applied convex integration in the context of bounded weak solutions to the Euler equations.

We refer the interested reader to the review papers [21, 82, 24] for a detailed discussion connecting convex integration in the context of differential inclusions, and also h-principles, to the type of constructions that were initiated by [20].

The work [20], has since been extended and adapted by various authors to various problems arising in mathematical physics [25, 16, 76, 87, 14]. Here we single out the work [25] of De Lellis-Székelyhidi Jr., who consider the question of whether imposing additional *admissibility criteria* on the weak solutions of the Euler equations could rule out the construction of examples such as those in Theorem 3.4. Physically motivated *admissibility criteria*, based on energetic arguments such as those discussed in Section 2 are (ordered by least restrictive to most restrictive):

(a) *weak energy inequality*: $E(t) \leq E(0)$ for all $t > 0$, where the kinetic energy E is as defined in (3.1);
(b) *strong energy inequality*: $E(t) \leq E(s)$ for all $t > s > 0$;
(c) *local energy inequality*: the distribution $D(v)$ defined weak form of the Kármán-Howarth-Monin relation (2.4) (with v replacing v^ν), is non-negative; note that this definition requires $v \in L^3_{\text{loc}}$.

The main result of [25] may be summarized as follows:

Theorem 3.5 (Theorem 1 in [25]). *For any dimension $n \geq 2$, there exists bounded, compactly supported divergence-free initial data $v_0 \in L^\infty_x \cap L^2_x$, for which there exist infinitely many weak solutions $v \in L^\infty_{x,t} \cap C^0_t L^2_x$ of the Euler equations, such that the admissibility conditions (a), (b), and (c) hold.*

This result shows that "wild" weak solutions of the Euler equations (as constructed by Theorem 3.4 and Theorem 3.5) cannot be ruled by the local energy inequality. Indeed, the Euler equations are not scalar conservation laws!

3.4. **The $C^{0+}_{x,t}$ result: a Nash scheme.** The $L^\infty_{x,t}$ constructions described in the previous subsection are based on writing the Euler equations as a differential inclusion, and then applying a machinery from Lipschitz differential inclusions, which either uses a Baire-category argument, or equivalently, and explicit convex integration approach. These methods face a serious difficulty in constructing *continuous* weak solutions of (1.2), since it seems impossible to extract a uniform continuity estimate for approximating sequences (v_q, p_q). The breakthrough was made by De Lellis and Székelyhidi Jr. in their seminal papers [22, 23], where they developed *a new convex integration scheme*, motivated and resembling in part the earlier schemes of Nash and Kuiper [65, 50]. In [22], De Lellis and Székelyhidi Jr. prove the existence of continuous weak solutions v to the Euler equations satisfying a prescribed kinetic energy profile, which in particular may be decreasing:

Theorem 3.6 (Theorem 1.1, [22]). *Assume $e\colon [0,1] \to (0,\infty)$ is a smooth function. Then there is a continuous vector field $v\colon \mathbb{T}^3 \times [0,1] \to \mathbb{R}^3$ and a continuous scalar field $p\colon \mathbb{T}^3 \times [0,1] \to \mathbb{R}$ which solve the incompressible Euler equations (1.2) in the sense of distributions, and such that*

$$e(t) = \int_{\mathbb{T}^3} |v(x,t)|^2 dx \tag{3.3}$$

for all $t \in [0,1]$.

The works [22, 23] are the first to make significant progress towards resolving Part (b) of Onsager's conjecture by providing the first construction of dissipative Hölder continuous weak solutions to the Euler equations.

The proof of Theorem 3.6 departs from the soft arguments based on functional analysis, which were used to construct bounded weak solutions, and implements a hard analysis scheme, in which the constructions of the building blocks are not plane waves anymore, instead they are adapted to the geometry of steady states of the Euler equations (Beltrami flows), and the estimates involve precise singular integral bounds and Schauder estimates.

Overview of the proof of Theorem 3.6. The proof proceeds via induction on $q \in \mathbb{N}_0$. For each $q \geq 0$ one constructs smooth functions $(v_q, p_q, \mathring{R}_q)$, which solve the *Euler-Reynolds system*

$$\partial_t v_q + \operatorname{div}(v_q \otimes v_q) + \nabla p_q = \operatorname{div} \mathring{R}_q \tag{3.4a}$$

$$\operatorname{div} v_q = 0\,. \tag{3.4b}$$

The pressure p_q is given by $p_q = (-\Delta)^{-1}\operatorname{div}\operatorname{div}(v_q \otimes v_q - \mathring{R}_q)$. The Reynolds stress $\mathring{R}_q$ is symmetric and has zero trace. The goal is to construct the sequence $(v_q, p_q, \mathring{R}_q)$ such that $\mathring{R}_q$ converges uniformly to 0 as $q \to \infty$, and that at the same time the sequence v_q converges uniformly to a weak solution to the Euler equations satisfying (3.3).

The Euler-Reynolds (3.4) system arises naturally in turbulence theories. As mentioned in [38], the concept of eddy viscosity and microscopic to macroscopic stresses may be

traced back to the work of Reynolds [70]. Given a solution v to (1.2), let $\overline{v}$ be the velocity obtained through the application of a filter (or averaging operator) that commutes with derivatives, whose role is to ignore the unresolved small scales. Then, upon defining

$$R = \overline{v \otimes v} - \overline{v} \otimes \overline{v} = \overline{(v - \overline{v}) \otimes (v - \overline{v})},$$

which is thus a the 3×3 symmetric positive definite matrix, we have that the pair $(\overline{v}, R)$ is a solution to (3.4), for a suitably defined pressure.

For comparison, the iterates $(v_q, p_q, \mathring{R}_q)$ constructed via a convex integration scheme are approximately spatial averages of the final solution v at length scales λ_q^{-1}, which are decreasing with q. In view of the analogy to theories in fluid turbulence, one refers to the symmetric tensor $\mathring{R}_q$ as the *Reynolds stress*. Without loss of generality, we will also assume $\mathring{R}_q$ to be traceless, since its trace is absorbed in to the pressure term p_q.

At each inductive step, the goal is to design a perturbation

$$w_{q+1} = v_{q+1} - v_q$$

such that the new velocity v_{q+1} solves the Euler-Reynolds system at level $q+1$

$$\begin{aligned} \partial_t v_{q+1} + \operatorname{div}(v_{q+1} \otimes v_{q+1}) + \nabla p_{q+1} &= \operatorname{div} \mathring{R}_{q+1} \\ \operatorname{div} v_{q+1} &= 0\,, \end{aligned}$$

with *a smaller Reynolds stress* $\mathring{R}_{q+1}$. Using the equation for v_q we obtain the following decomposition of the Reynolds stress at level $q+1$:

$$\begin{aligned} &\operatorname{div} \mathring{R}_{q+1} - \nabla(p_{q+1} - p_q) \\ &= \underbrace{\operatorname{div}(w_{q+1} \otimes w_{q+1} + \mathring{R}_q)}_{\textit{oscillation error}} + \underbrace{\partial_t w_{q+1} + v_q \cdot \nabla w_{q+1}}_{\textit{transport error}} + \underbrace{w_{q+1} \cdot \nabla v_q}_{\textit{Nash error}}\,. \end{aligned} \tag{3.5}$$

Note that not all the terms on the right side of (3.5) are written in divergence form, necessary in order to compute $\mathring{R}_{q+1}$ on the left side of this definition. This is achieved by utilizing a negative one order linear Fourier multiplier operator $\mathcal{R}$ which formally acts as div^{-1} and outputs symmetric traceless matrices.

For an increasing sequence of frequency parameters $\{\lambda_q\}_{q \geq 0}$,[7] the approximate solutions at level $(v_q, \mathring{R}_q)$ are essentially localized at Fourier frequencies $\lesssim \lambda_q$. On the other hand, the perturbation $w_{q+1} = v_{q+1} - v_q$ is constructed as a sum of highly oscillatory *building blocks* (denoted by W_ξ in (3.6) below) which live at the higher frequency λ_{q+1}. The building blocks used in the papers [22, 23], are the so-called *Beltrami waves*. These are families of complex eigenfunctions of the curl operator at the same eigenvalue, λ_{q+1}. These building blocks are used in an analogous fashion to the *Nash twists* and *Kuiper corrugations* employed in the C^1 embedding problem [65, 50]. The perturbation w_{q+1} is designed in order to obtain a cancellation between the low frequencies of the quadratic term $w_{q+1} \otimes w_{q+1}$ and the old Reynolds stress error $\mathring{R}_q$, thereby reducing the size of the oscillation error. Roughly speaking, the principal part of the perturbation, which we label $w_{q+1}^{(p)}$, will be of the form

$$w_{q+1}^{(p)} \sim \sum_\xi a_\xi(\mathring{R}_q) W_\xi\,, \tag{3.6}$$

[7] At first reading it is convenient to think of $\lambda_q = \lambda_0^q$, where λ_0 is a fixed sufficiently large parameter of the problem. In the proof however, we need to take the sequence λ_q to grow super-exponentially, $\lambda_q = \lambda_0^{b^q}$ for some $b > 1$, in order to fight a *derivative loss* issue which arises in the problem. A similar issue arises in more classical constructions, which go under the umbrella name *Nash-Moser* iteration.

where the W_ξ represent the *building blocks* oscillating at a prescribed high frequency λ_{q+1}, and the coefficient functions a_ξ are chosen such that

$$\sum_\xi a_\xi^2(\mathring{R}_q) \fint_{\mathbb{T}^3} W_\xi \mathring{\otimes} W_\xi = -\mathring{R}_q \,. \tag{3.7}$$

Here $\mathring{\otimes}$ denotes the trace-free part of the tensor product. This principal part of the perturbation needs to be modified from the form presented in (3.6) in order to minimize the transport error in (3.5), i.e., to ensure it is the divergence of a small Reynolds stress. This is achieved by flowing the building blocks W_ξ along the ODE flow generated by v_q (we will return to this issue in Section 3.5). Additionally, in order to ensure that w_{q+1} is divergence free, one introduces a corrector $w_{q+1}^{(c)}$ which ensures that $w_{q+1} = w_{q+1}^{(p)} + w_{q+1}^{(c)}$ is divergence free. The size of this incompressibility corrector $w_{q+1}^{(c)}$ is much smaller than the size of $w_{q+1}^{(p)}$, roughly by a factor of $\lambda_q \lambda_{q+1}^{-1}$, because the building blocks W_ξ are divergence-free by definition, and the a_ξ oscillate at the old frequency, λ_q.

In order to ensure that the inductive scheme converges to a Hölder continuous velocity v with Hölder exponent > 0, the perturbation's amplitude is required to satisfy the bound

$$\|w_{q+1}\|_{C^0} \leq \lambda_{q+1}^{-\beta} \tag{3.8}$$

for some $\beta > 0$. Here, we again note that it is convenient to use a super-exponentially growing sequence of frequencies λ_q which obeys $\lambda_{q+1} \approx \lambda_q^b$, for some $b > 1$. In view of (3.7), this necessitates that the Reynolds stress $\mathring{R}_q$ obeys the estimate

$$\left\| \mathring{R}_q \right\|_{C^0} \leq \lambda_{q+1}^{-2\beta} \,. \tag{3.9}$$

Consistent with the definition $v_q = v_0 + \sum_{q'<q} w_{q'}$ and with the bound (3.8) the scheme of [22] also propagates the estimate

$$\|\nabla v_q\|_{C^0} \leq M \lambda_q^{1-\beta} \tag{3.10}$$

for a fixed constant $M > 0$. At this stage it is easy to see that if the bounds (3.8)–(3.10) are propagated throughout the scheme, then as $q \to \infty$ we have that $(v_q, \mathring{R}_q) \to (v, 0)$ uniformly, where v is a Hölder continuous weak solution of the Euler equations. Indeed, for any $\theta \in (0, \beta)$, the following series of increments is summable

$$\begin{aligned} \sum_{q\geq 0} \|w_{q+1}\|_{C^\theta} &\lesssim \sum_{q\geq 0} \|w_{q+1}\|_{C^0}^{1-\theta} \|\nabla w_{q+1}\|_{C^0}^{\theta} \\ &\lesssim M \sum_{q\geq 0} \lambda_{q+1}^{-\beta(1-\theta)} \lambda_{q+1}^{\theta(1-\beta)} \lesssim M \sum_{q\geq 0} \lambda_{q+1}^{\theta-\beta} \lesssim 1 \end{aligned}$$

where the implicit constant is universal. Thus, we may define a limiting function $v = \lim_{q\to\infty} v_q$ which lies in $C^0([0,1]; C^\theta)$. Moreover, v is a weak solution of the Euler equation (1.2), since by (3.9) we have that $\lim_{q\to\infty} \mathring{R}_q = 0$ in $C^0([0,1]; C^0)$.

The main work is now to prove that for a velocity perturbation w_{q+1} of the form (3.6), and with amplitude functions that satisfy (3.7), the bounds stated in (3.8)–(3.10) are indeed attainable inductively for all $q \geq 1$. We note if the building blocks are normalized to $\|W_\xi\|_{C^0} \approx 1$, then it follows from (3.6)–(3.7) and (3.9) that the principal part of the velocity increment already satisfies the bound (3.8). Since the incompressibility corrector is even smaller, we are quite confident about (3.8). The difficult part is to prove (3.9); in view of (3.5) this amounts to bounding the *oscillation error*, the *transport error*, and the *Nash error*. This is the hard analysis part of the construction.

As a demonstration of the typical scalings present in convex integration schemes for the Euler equations, let us consider the Nash error. Heuristically, since $v_q = v_0 + \sum_{q'<q} w_{q'}$ and w_{q+1} is of frequency $\lambda_{q+1} \gg \lambda_{q'}$ for every $q' \leq q$, we have that $w_{q+1} \cdot \nabla v_q$ lives at frequency λ_{q+1}, and thus

$$\|\mathcal{R}\left(w_{q+1} \cdot \nabla v_q\right)\|_{C^0} \lesssim \lambda_{q+1}^{-1} \|w_{q+1}\|_{C^0} \|\nabla v_q\|_{C^0}$$

where we recall that $\mathcal{R}$ is a -1 order linear pseudo-differential operator which inverts the divergence operator. Applying (3.8) and (3.10), for $\beta \in (0,1)$ we obtain

$$\|\mathcal{R}\left(w_{q+1} \cdot \nabla v_q\right)\|_{C^0} \lesssim \lambda_{q+1}^{-1-\beta} \lambda_q^{1-\beta} = \lambda_{q+2}^{-2\beta} \lambda_q^{1-\beta-b(1+\beta)+2\beta b^2}$$

by using that $\lambda_{q+1} \approx \lambda_q^b$. Thus, in order to ensure that $\mathring{R}_{q+1}$ satisfies the bound (3.9) with q replaced by $q+1$, we require that for $\beta \in (0,1)$ and $b > 1$ we have

$$1 - \beta - b(1+\beta) + 2\beta b^2 = (1-b)(1-\beta-2\beta b) < 0\,.$$

Thus, from this simple heuristic, we see that if $b > 1$ is taken to be arbitrarily close to 1, then the Hölder regularity exponent β may be taken to be arbitrarily close to the Onsager-critical Hölder regularity exponent, i.e. $\beta < 1/3$.

The construction described above provides a clear enemy towards reaching the desired Onsager $1/3$ threshold: the transport and oscillation errors in (3.5). Designing a w_{q+1} which minimizes these two errors *simultaneously* turns out to be a very difficult problem. This realization stimulated a series of advancements, through the works of Isett, Buckmaster, De Lellis, Székelyhidi, and Daneri [40, 7, 5, 8, 18], in which the authors incorporated more an more of the specifics of the 3D Euler equation into the convex integration scheme (by designing better W_ξ and a_ξ), in order to obtain higher and higher Hölder regularity exponents. We mention a couple of these developments next.

3.5. **Climbing the Onsager ladder.** The first breakthrough after [22, 23] was to produce a dissipative weak solution of the Euler system with a Hölder regularity exponent θ with $\theta < 1/5$; this was achieved independently by Isett [40] and by Buckmaster [4], De Lellis, and Székelyhidi, resulting in the joint work [7]. The main improvement comes from obtaining a better bound for the transport error in (3.5). In the proof of Theorem 3.6 one did not keep track of precise estimates for the *material derivative of the Reynolds stress* $(\partial_t + v_\ell \cdot \nabla)\mathring{R}_q$. Here v_ℓ is a mollification of v_q at a length scale ℓ which lies in between λ_{q+1}^{-1} and λ_q^{-1}.[8] The realization of the $1/5$ *schemes* is that material derivatives are better behaved than either regular spatial or temporal derivatives: due to classical ODE arguments, a material derivatives should cost a factor proportional to the Lipschitz norm of v_q, i.e. $\lambda_q^{1-\beta}$ in view of (3.10). Compare this with a spatial derivative, whose cost is $\lambda_q \gg \lambda_q^{1-\beta}$. Taking advantage of this observation, one can improve the estimate on the material derivative of $\mathring{R}_q$, and thus improve the bounds for the transport error. Optimizing this new transport bound with the oscillation error yields the improved $1/5-$ Hölder exponent.

The next major result was due to Buckmaster [5] who noted that one can construct infinitely many weak solutions of (1.2) whose x-Hölder regularity exponent, i.e. with respect to the space variable, can be taken to be any θ with $\theta < 1/3$, but only *almost everywhere in*

[8] An inherent issue associated with convex integration schemes is that in order to control nth order derivatives of the perturbation w_{q+1}, one needs control derivatives on v_q of an order strictly greater than n. This *loss of derivative problem* is also intrinsic to Nash-Moser arguments [65, 63]. In order to avoid this loss of derivative, one replaces v_q by a mollified velocity field v_ℓ and the stress $\mathring{R}_q$ by a mollified stress $\mathring{R}_\ell$, where the mollification parameter $\ell \in (\lambda_{q+1}^{-1}, \lambda_q^{-1})$ is to be chosen suitably. This argument was already required in the $C_{x,t}^{0+}$ schemes described in the previous subsection.

time. This new scheme concentrates the transport and oscillation errors on a zero-measure set of times. By taking advantage of this idea and by using a delicate bookkeeping scheme, Buckmaster, De Lellis, and Székelyhidi [8] constructed non-conservative solutions were constructed in the space $L^1_t C^{1/3-}_x$.

3.6. **Resolution of the flexible side of the Onsager conjecture.** The flexible side of the Onsager conjecture was finally resolved by Isett in [42], who proved the existence of non-conservative weak solutions of 3D Euler in the regularity class $C^\theta_{x,t}$, for any $\theta < 1/3$:

Theorem 3.7 (Theorem 1, [42]). *For any $\theta \in (0, 1/3)$ there exists a nonzero weak solution $v \in C^\theta(\mathbb{T}^3 \times \mathbb{R})$, such that v vanishes identically outside of a finite interval.*

The proof of Isett builds upon the ideas in the above mentioned works, and utilizes two new key ingredients. The first, is the usage of Mikado flows which were introduced earlier by Daneri and Székelyhidi [18]. These are a rich family of pressure-less stationary solutions of the 3D Euler equation (straight pipe flows), which have a better (when compared to Beltrami flows) self-interaction behavior in the oscillation error, when they are advected by a mean flow. The second key ingredient is due to Isett: prior to adding the convex integration perturbation w_{q+1}, it is very useful to replace the approximate solution $(v_q, \mathring{R}_q)$ with another pair $(\overline{v}_q, \mathring{\overline{R}}_q)$, which has the property that $\overline{v}_q$ is close to v_q, but more importantly, that $\mathring{\overline{R}}_q(t)$ vanishes on *every other interval* of size $\approx \|\nabla v_q\|^{-1}_{C^0}$ within $[0, T]$. This in turn results in a major improvement of the size of the oscillation error, since different Mikado flows have disjoint supports, and thus do not interact on the time scale dictated by the Lipschitz norm of the mean flow v_ℓ.

The weak solutions constructed by Isett [42] are not strictly dissipative. This issue was resolved in the paper by Buckmaster-De Lellis-Székelyhidi-V. [9], who prove the precise statement of Part b of the Onsager conjecture:

Theorem 3.8 (Theorem 1.1, [9]). *Let $e\colon [0, T] \to \mathbb{R}$ be a strictly positive smooth function. For any $\beta \in (0, 1/3)$ there exists a weak solution $v \in C^\beta(\mathbb{T}^3 \times [0, T])$ of the Euler equations (1.2), whose kinetic energy at time $t \in [0, T]$ equals $e(t)$.*

Finally, we note that in [41], Isett showed that one can further optimize the schemes of [42, 9] in order to construct non-conservative weak solutions to the Euler equations that lie in the intersection of all Hölder spaces C^β for $\beta < 1/3$. It is an open problem to determine whether non-conservative weak solutions to the Euler equations exist that have Hölder exponent *exactly* $1/3$.[9]

4. Concluding comments

The exponent $1/3$ in Onsager's conjecture can be viewed in terms of a larger class of *threshold exponents* at which a dichotomy in the behavior of solutions arises. In a recent expository paper [46], Klainerman considered various threshold exponents in the context of non-linear PDE. For example, the regularity exponent at which the uniqueness of weak solutions to the Euler equations is expected to fail is 1. Indeed, we know that uniqueness holds in $C^{1,\alpha}$ with $\alpha > 0$, and we expect that uniqueness fails below C^1. On the other hand, the Onsager exponent $1/3$ determines the regularity threshold at which the Hamiltonian of the system (the kinetic energy) fails to be conserved. These exponents are not the same. At the moment of writing of this article, it is not known how to construct nonunique

[9] Note that in view of [13], it would be desirable to at least determine whether there exist non-conservative weak solutions which lie in the space $L^3_t B^{1/3}_{3,\infty,x}$.

weak solutions of the 3D Euler equations with $C^{\theta}_{x,t}$ regularity, for some $\theta \in (1/3, 1)$. It appears that such a result would require fundamental new ideas, beyond the ideas provided by the convex integration schemed described earlier.

In view of the discussion in Section 2, and keeping in mind Kolmogorov's $4/5$-law (2.12), it seems that the natural context in which to address energy conservation is not that of L^{∞} based spaces (e.g. Hölder spaces), but rather, L^3-based spaces. In fact, the statistics of the pth order structure functions in turbulent flows motivate an analysis in any L^p with $p \geq 1$. This following question naturally arises: for $p < 3$, what are the maximal deviations from the Kolmogorov prediction (2.11) for the structure function exponent ζ_p, which is sustainable by weak solutions of the 3D Euler equations? For instance, keeping in mind the K41 kinetic energy spectrum (2.14), one may consider Onsager's conjecture in the context of L^2-based Sobolev spaces, and ask to construct dissipative weak solutions of (1.2) with Sobolev regularity *above* $1/3$. Note that the kinetic energy is conserved for any weak solution $v \in H^{5/6}$ (cf. [81, 13]). It is an open problem to determine whether or not this result is sharp. Such a result would have important implications for validating various physical theories of intermittency in fully developed hydrodynamic turbulence.

In closing, we mention that Buckmaster and the author have recently developed in [11] an *intermittent convex integration scheme*, which is able to show the existence of infinitely many distributional solutions $v \in C^0_t L^2_x \cap C^0_t W^{1,1}_x$ of the 3D Navier-Stokes equations (1.1), with a prescribed kinetic energy profile. The principal result of [11] is:

Theorem 4.1 (Theorem 1.2, [11])**.** *There exists $\beta > 0$, such that for any nonnegative smooth function $e(t)\colon [0,T] \to [0,\infty)$, and any $\nu \in (0,1]$, there exists a weak solution of the Navier-Stokes equations (1.1) $v \in C^0([0,T]; H^{\beta}(\mathbb{T}^3)) \cap C^0([0,T]; W^{1,1+\beta}(\mathbb{T}^3))$, such that $\int_{\mathbb{T}^3} |v(x,t)|^2 dx = e(t)$ holds for all $t \in [0,T]$.*

We emphasize that the weak solutions constructed in Theorem 4.1 are *not Leray weak solutions* [52], whose uniqueness remains famously one of the most challenging questions in fluid mechanics.

The ideas of [11] and of the subsequent improvement in Buckmaster-Colombo-V. [6], combined with additional new ideas, has been successfully applied in related contexts in PDEs. Using *intermittent Mikado flows*, Modena and Székeyhidi Jr. have establishied the existence of non-renormalized solutions to the transport and continuity equations with Sobolev vector fields [59, 60]. In [17], Dai demonstrated that these methods can be adapted to prove non-uniqueness of Leray-Hopf weak solutions for the 3D Hall-MHD system. T. Luo and Titi [55] prove that these methods are applicable also to the fractional Navier-Stokes equations with dissipation $(-\Delta)^{\alpha}$, and $\alpha < 5/4$ (the Lions criticality threshold [53]); X. Luo [56] demonstrated the existence of non-trivial stationary solutions to the 4D Navier-Stokes equations, and Beekie-Buckmaster-V. [3] have constructed weak solutions to the ideal 3D MHD equations which do not conserve magnetic helicity.

The emergence of intermittent convex integration schemes is yet another testament to the power and the flexibility of the machinery that De Lellis and Székelyhidi have developed. It is fair to say that the limitations of convex integrations schemes in fluid dynamics are not yet known. A number of physically motivated, very interesting mathematical challenges remain to be explored; we refer the interested reader to the review papers [21, 82, 24, 10] for further details.

REFERENCES

[1] G. K. Batchelor, *The theory of homogeneous turbulence*, Cambridge Monographs on Mechanics and Applied Mathematics, Cambridge at the University Press, 1953. MR 0052268 (14,597a)

[2] J. T. Beale, T. Kato, and A. Majda, *Remarks on the breakdown of smooth solutions for the 3-d euler equations*, Communications in Mathematical Physics **94** (1984), no. 1, 61–66.

[3] R. Beekie, T. Buckmaster, and V. Vicol, *Weak solutions of ideal mhd which do not conserve magnetic helicity*, arXiv preprint arXiv:1907.10436 (2019).

[4] T. Buckmaster, *Onsager's Conjecture*, Ph.D. thesis, Universität Leipzig, 2014.

[5] ———, *Onsager's conjecture almost everywhere in time*, Comm. Math. Phys. **333** (2015), no. 3, 1175–1198. MR 3302631

[6] T. Buckmaster, M. Colombo, and V. Vicol, *Wild solutions of the Navier-Stokes equations whose singular sets in time have hausdorff dimension strictly less than* 1, arXiv:1809.00600 (2018).

[7] T. Buckmaster, C. De Lellis, P. Isett, and L. Székelyhidi, Jr., *Anomalous dissipation for* $1/5$*-Hölder Euler flows*, Annals of Mathematics **182** (2015), no. 1, 127–172.

[8] T. Buckmaster, C. De Lellis, and L. Székelyhidi, Jr., *Dissipative Euler flows with Onsager-critical spatial regularity*, Comm. Pure Appl. Math. **69** (2016), no. 9, 1613–1670. MR 3530360

[9] T. Buckmaster, C. De Lellis, L. Székelyhidi Jr, and V. Vicol, *Onsager's conjecture for admissible weak solutions*, Comm. Pure Appl. Math. **72** (2019), no. 2, 227–448.

[10] T. Buckmaster and V. Vicol, *Convex integration and phenomenologies in turbulence*, arXiv:1901.09023 (2019).

[11] ———, *Nonuniqueness of weak solutions to the Navier-Stokes equation*, Ann. of Math. **189** (2019), no. 1, 101–144.

[12] J. Chen and T.Y. Hou, *Finite time blowup of 2D Boussinesq and 3D Euler equations with* $C^{1,\alpha}$ *velocity and boundary*, arXiv preprint arXiv:1910.00173 (2019).

[13] A. Cheskidov, P. Constantin, S. Friedlander, and R. Shvydkoy, *Energy conservation and Onsager's conjecture for the Euler equations*, Nonlinearity **21** (2008), no. 6, 1233–1252. MR 2422377 (2009g:76008)

[14] Elisabetta Chiodaroli, Camillo De Lellis, and Ondřej Kreml, *Global ill-posedness of the isentropic system of gas dynamics*, Comm. Pure Appl. Math. **68** (2015), no. 7, 1157–1190. MR 3352460

[15] P. Constantin, W. E, and E.S. Titi, *Onsager's conjecture on the energy conservation for solutions of Euler's equation*, Comm. Math. Phys. **165** (1994), no. 1, 207–209. MR 1298949 (96e:76025)

[16] D. Cordoba, D. Faraco, and F. Gancedo, *Lack of uniqueness for weak solutions of the incompressible porous media equation*, Arch. Ration. Mech. Anal. **200** (2011), no. 3, 725–746. MR 2796131

[17] M. Dai, *Non-uniqueness of Leray-Hopf weak solutions of the 3d Hall-MHD system*, arXiv:1812.11311 (2018).

[18] S. Daneri and L. Székelyhidi Jr, *Non-uniqueness and h-principle for Hölder-continuous weak solutions of the Euler equations*, Arch. Ration. Mech. Anal. **224** (2017), no. 2, 471–514.

[19] P.A. Davidson, Y. Kaneda, K. Moffatt, and K.R. Sreenivasan, *A voyage through turbulence*, Cambridge University Press, 2011.

[20] C. De Lellis and L. Székelyhidi, Jr., *The Euler equations as a differential inclusion*, Ann. of Math. (2) **170** (2009), no. 3, 1417–1436. MR 2600877 (2011e:35287)

[21] C. De Lellis and L. Székelyhidi, Jr., *The h-principle and the equations of fluid dynamics*, Bull. Amer. Math. Soc. (N.S.) **49** (2012), no. 3, 347–375. MR 2917063

[22] ———, *Dissipative continuous Euler flows*, Invent. Math. **193** (2013), no. 2, 377–407. MR 3090182

[23] C. De Lellis and L. Székelyhidi, Jr., *Dissipative Euler flows and Onsager's conjecture*, Journal of the European Mathematical Society **16** (2014), no. 7, 1467–1505.

[24] C. De Lellis and L. Székelyhidi Jr, *On turbulence and geometry: from Nash to Onsager*, arXiv:1901.02318 (2019).

[25] Camillo De Lellis and László Székelyhidi, Jr., *On admissibility criteria for weak solutions of the Euler equations*, Arch. Ration. Mech. Anal. **195** (2010), no. 1, 225–260. MR 2564474

[26] Ronald J. DiPerna, *Compensated compactness and general systems of conservation laws*, Transactions of the American Mathematical Society **292** (1985), no. 2, 383–420.

[27] J. Duchon and R. Robert, *Inertial energy dissipation for weak solutions of incompressible euler and navier-stokes equations*, Nonlinearity **13** (2000), no. 1, 249.

[28] T.M. Elgindi, *Finite-time singularity formation for* $C^{1,\alpha}$ *solutions to the incompressible Euler equations on* R^3, arXiv preprint arXiv:1904.04795 (2019).

[29] T.M. Elgindi, T.-E. Ghoul, and N. Masmoudi, *On the stability of self-similar blow-up for* $C^{1,\alpha}$ *solutions to the incompressible Euler equations on* R^3, arXiv preprint arXiv:1910.14071 (2019).

[30] Y. Eliashberg and N. Mishachev, *Introduction to the h-principle*, Graduate Studies in Mathematics, vol. 48, American Mathematical Society, Providence, RI, 2002. MR 1909245

[31] L. Euler, *Principes généraux du mouvement des fluides*, Académie Royale des Sciences et des Belles Lettres de Berlin, Mémoires **11** (1757), 274–315.

[32] G.L. Eyink, *Local 4/5-law and energy dissipation anomaly in turbulence*, Nonlinearity **16** (2002), no. 1, 137.
[33] ———, *Energy dissipation without viscosity in ideal hydrodynamics I. Fourier analysis and local energy transfer*, Physica D: Nonlinear Phenomena **78** (94), no. 3–4, 222–240.
[34] G.L. Eyink and K.R. Sreenivasan, *Onsager and the theory of hydrodynamic turbulence*, Rev. Modern Phys. **78** (2006), no. 1, 87–135. MR 2214822 (2007g:76108)
[35] C. Foias, *Statistical study of Navier–Stokes equations i*, Rend. Sem. Mat. Univ. Padova **48** (1972), no. 219–348.
[36] ———, *Statistical study of Navier–Stokes equations ii*, Rend. Sem. Mat. Univ. Padova **49** (1973), 9–123.
[37] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes equations and turbulence*, Encyclopedia of Mathematics and its Applications, vol. 83, Cambridge University Press, Cambridge, 2001. MR 1855030 (2003a:76001)
[38] U. Frisch, *Turbulence*, Cambridge University Press, Cambridge, 1995, The legacy of A. N. Kolmogorov. MR MR1428905 (98e:76002)
[39] M. L. Gromov, *Convex integration of differential relations. I*, Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya **37** (1973), 329–343.
[40] P. Isett, *Holder continuous Euler flows with compact support in time*, Ph.D. thesis, Thesis (Ph.D.)–Princeton University, 2013, p. 227. MR 3153420
[41] ———, *On the endpoint regularity in Onsager's conjecture*, arXiv:1706.01549 (2017).
[42] ———, *A proof of Onsager's conjecture*, Ann. of Math. **188** (2018), no. 3, 871–963.
[43] Y. Kaneda, T. Ishihara, M. Yokokawa, K. Itakura, and A. Uno, *Energy dissipation rate and energy spectrum in high resolution direct numerical simulations of turbulence in a periodic box*, Physics of Fluids **15** (2003), no. 2, L21–L24.
[44] T. von Kármán and L. Howarth, *On the statistical theory of isotropic turbulence*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **164** (1938), no. 917, 192–215.
[45] Bernd Kirchheim, Stefan Müller, and Vladimír Šverák, *Studying Nonlinear pde by Geometry in Matrix Space*, Geometric Analysis and Nonlinear Partial Differential Equations (Stefan Hildebrandt and Hermann Karcher, eds.), Springer Berlin Heidelberg, 2003, pp. 347–395 (English).
[46] S. Klainerman, *On Nash's unique contribution to analysis in just three of his papers*, Bulletin of the American Mathematical Society **54** (2017), no. 2, 283–305.
[47] A.N. Kolmogorov, *Dissipation of energy in the locally isotropic turbulence*, Dokl. Akad. Nauk SSSR **32** (1941), 16–18.
[48] ———, *Local structure of turbulence in an incompressible fluid at very high reynolds number*, Dokl. Acad. Nauk SSSR **30** (1941), no. 4, 299–303, Translated from the Russian by V. Levin, Turbulence and stochastic processes: Kolmogorov's ideas 50 years on. MR 1124922 (92h:76049)
[49] ———, *On degeneration of isotropic turbulence in an incompressible viscous liquid*, Dokl. Akad. Nauk SSSR, vol. 31, 1941, pp. 538–540.
[50] Nicolaas H. Kuiper, *On C^1-isometric imbeddings. i, II*, Proc. Kon. Acad. Wet. Amsterdam A **58** (1955), 545–556, 683–689.
[51] I. Kukavica, *Role of the pressure for validity of the energy equality for solutions of the Navier-Stokes equation*, J. Dynam. Differential Equations **18** (2006), no. 2, 461–482. MR 2229985 (2007c:35131)
[52] J. Leray, *Sur le mouvement d'un liquide visqueux emplissant l'espace*, Acta Math. **63** (1934), no. 1, 193–248. MR 1555394
[53] J.-L. Lions, *Quelques résultats d'existence dans des équations aux dérivées partielles non linéaires*, Bull. Soc. Math. France **87** (1959), no. 2, 245–273.
[54] ———, *Sur la régularité et l'unicité des solutions turbulentes des équations de Navier Stokes*, Rend. Sem. Mat. Univ. Padova **30** (1960), 16–23.
[55] T. Luo and E.S. Titi, *Non-uniqueness of weak solutions to hyperviscous Navier-Stokes equations - on sharpness of J.-L. Lions exponent*, arXiv:1808.07595 (2018).
[56] X. Luo, *Stationary solutions and nonuniqueness of weak solutions for the Navier-Stokes equations in high dimensions*, arXiv:1807.09318 (2018).
[57] Andrew J. Majda and Andrea L. Bertozzi, *Vorticity and incompressible flow*, Cambridge Texts in Applied Mathematics, vol. 27, Cambridge University Press, Cambridge, 2002. MR 1867882
[58] C. Meneveau and K.R. Sreenivasan, *Simple multifractal cascade model for fully developed turbulence*, Phys. Rev. Lett. **59** (1987), no. 13, 1424.
[59] S. Modena and L. Székelyhidi, Jr., *Non-uniqueness for the transport equation with Sobolev vector fields*, arXiv:1712.03867 (2017).
[60] ———, *Non-renormalized solutions to the continuity equation*, arXiv:1806.09145 (2018).

[61] A.S. Monin, *The theory of locally isotropic turbulence*, Soviet Physics Doklady, vol. 4, 1959, p. 271.
[62] A.S. Monin and A.M. Yaglom, *Statistical fluid mechanics, volume ii: mechanics of turbulence*, vol. 2, Courier Corporation, 2013.
[63] J. Moser, *A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations*, Comm. Pure Appl. Math. **13** (1960), 457–468. MR 0170091 (30 #332)
[64] S. Müller and V. Šverák, *Convex Integration for Lipschitz Mappings and Counterexamples to Regularity*, Annals of Mathematics **157** (2003), no. 3, pp. 715–742 (English).
[65] J. Nash, *C1 isometric imbeddings*, Ann. of Math. (1954), 383–396.
[66] C.L. Navier, *Mémoire sur les lois du mouvement des fluides*, Mémoires de lAcadémie Royale des Sciences de l'Institut de France **6** (1823), 389–440.
[67] Q. Nie and S. Tanveer, *A note on third–order structure functions in turbulence*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **455** (1999), no. 1985, 1615–1635.
[68] L. Onsager, *Statistical hydrodynamics*, Nuovo Cimento (9) **6** (1949), no. Supplemento, 2(Convegno Internazionale di Meccanica Statistica), 279–287. MR 0036116 (12,60f)
[69] B.R. Pearson, P.-Å. Krogstad, and W. Van De Water, *Measurements of the turbulent energy dissipation rate*, Physics of fluids **14** (2002), no. 3, 1288–1290.
[70] O. Reynolds, *On the dynamical theory of incompressible viscous fluids and the determination of the criterion*, Philosophical Transactions of the Royal Society of London. A **186** (1895), 123–164.
[71] V. Scheffer, *An inviscid flow with compact support in space-time*, J. Geom. Anal. **3** (1993), no. 4, 343–401. MR 1231007 (94h:35215)
[72] M. Shinbrot, *The energy equation for the Navier-Stokes system*, SIAM J. Math. Anal. **5** (1974), no. 6, 948–954.
[73] A. Shnirelman, *Weak solutions with decreasing energy of incompressible Euler equations*, Comm. Math. Phys. **210** (2000), no. 3, 541–603. MR 1777341
[74] A.I. Shnirelman, *On the nonuniqueness of weak solution of the Euler equation*, Comm. Pure Appl. Math. **50** (1997), no. 12, 1261–1286. MR 1476315 (98j:35149)
[75] R. Shvydkoy, *Lectures on the Onsager conjecture*, Discrete Contin. Dyn. Syst. Ser. S **3** (2010), no. 3, 473–496.
[76] ———, *Convex integration for a class of active scalar equations*, J. Amer. Math. Soc. **24** (2011), no. 4, 1159–1174.
[77] ———, *Homogeneous solutions to the 3d Euler system*, Trans. Amer. Math. Soc. **370** (2018), no. 4, 2517–2535.
[78] K.R. Sreenivasan, *An update on the energy dissipation rate in isotropic turbulence*, Physics of Fluids **10** (1998), no. 2, 528–529.
[79] K.R. Sreenivasan, S.I. Vainshtein, R. Bhiladvala, I. San Gil, S. Chen, and N. Cao, *Asymmetry of Velocity Increments in Fully Developed Turbulence and the Scaling of Low-Order Moments*, Phys. Rev. Lett. **77** (1996), no. 8, 1488–1491.
[80] G. G. Stokes, *On the theories of the internal friction of fluids in motion and of the equilibrium and motion of elastic solids*, Transactions of the Cambridge Philosophical Society **8** (1845), 287–319.
[81] P.-L. Sulem and U. Frisch, *Bounds on energy flux for finite energy turbulence*, J. Fluid Mech. **72** (1975), no. 3, 417–423.
[82] L. Székelyhidi Jr, *From isometric embeddings to turbulence*, HCDTE lecture notes. Part II. Nonlinear hyperbolic PDEs, dispersive and transport equations **7** (2012), 63.
[83] L. Tartar, *Compensated compactness and applications to partial differential equations*, Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV, Res. Notes in Math., vol. 39, Pitman, Boston, Mass.-London, 1979, p. 136–212.
[84] Luc Tartar, *The compensated compactness method applied to systems of conservation laws*, Systems of nonlinear partial differential equations (Oxford, 1982), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 111, Reidel, Dordrecht, 1983, p. 263–285.
[85] M. van Dyke, *An album of fluid motion*, The Parabolic Press, 1982.
[86] M.I. Vishik, A.I. Komech, and A.V. Fursikov, *Some mathematical problems of statistical hydromechanics*, Uspekhi Mat. Nauk **34** (1979), no. 5(209), 135–210, 256. MR 562801 (83e:35098)
[87] E. Wiedemann, *Existence of weak solutions for the incompressible Euler equations*, Annales de l'Institut Henri Poincare (C) Non Linear Analysis **28** (2011), no. 5, 727–730.

COURANT INSTITUTE FOR MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY.
E-mail address: vicol@cims.nyu.edu

CURRENT EVENTS BULLETIN

Previous speakers and titles

For PDF files of talks, and links to Bulletin of the AMS articles,
see http://www.ams.org/ams/current-events-bulletin.html.

January 18, 2019 (Baltimore, MD)

Bhargav Bhatt, University of Michigan
Perfectoid geometry and its applications

Thomas Vidick, California Institute of Technology
Verifying quantum computations at scale: a cryptographic leash on quantum devices

Stephanie van Willigenburg, University of British Columbia
The shuffle conjecture

Robert Lazarsfeld, Stony Brook University
Tangent Developable Surfaces and the Equations Defining Algebraic Curves

January 12, 2018 (San Diego, CA)

Richard D. James, University of Minnesota
Materials from mathematics

Craig L. Huneke, University of Virginia
How complicated are polynomials in many variables?

Isabelle Gallagher, Université Paris Diderot
From Newton to Navier-Stokes, or how to connect fluid mechanics equations from microscopic to macroscopic scales

Joshua A. Grochow, University of Colorado, Boulder
The Cap Set Conjecture, the polynomial method, and applications (after Croot-Lev-Pach, Ellenberg-Gijswijt, and others)

January 6, 2017 (Atlanta, GA)

Lydia Bieri, University of Michigan
Black hole formation and stability: a mathematical investigation.

Matt Baker, Georgia Tech
Hodge Theory in Combinatorics.

Kannan Soundararajan, Stanford University
Tao's work on the Erdos Discrepancy Problem.

Susan Holmes, Stanford University
Statistical proof and the problem of irreproducibility.

January 8, 2016 (Seattle, WA)

Carina Curto, Pennsylvania State University
What can topology tell us about the neural code?

Lionel Levine, Cornell University and *Yuval Peres, Microsoft Research and University of California, Berkeley
Laplacian growth, sandpiles and scaling limits.

Timothy Gowers, Cambridge University
Probabilistic combinatorics and the recent work of Peter Keevash.

Amie Wilkinson, University of Chicago
What are Lyapunov exponents, and why are they interesting?

January 12, 2015 (San Antonio, TX)

Jared S. Weinstein, Boston University
Exploring the Galois group of the rational numbers: Recent breakthroughs.

Andrea R. Nahmod, University of Massachusetts, Amherst
The nonlinear Schrödinger equation on tori: Integrating harmonic analysis, geometry, and probability.

Mina Aganagic, University of California, Berkeley
String theory and math: Why this marriage may last.

Alex Wright, Stanford University
From rational billiards to dynamics on moduli spaces.

January 17, 2014 (Baltimore, MD)

Daniel Rothman, Massachusetts Institute of Technology
Earth's Carbon Cycle: A Mathematical Perspective

Karen Vogtmann, Cornell University
The geometry of Outer space

Yakov Eliashberg, Stanford University
Recent advances in symplectic flexibility

Andrew Granville, Université de Montréal
Infinitely many pairs of primes differ by no more than 70 million (and the bound's getting smaller every day)

January 11, 2013 (San Diego, CA)

Wei Ho, Columbia University
How many rational points does a random curve have?

Sam Payne, Yale University
Topology of nonarchimedean analytic spaces

Mladen Bestvina, University of Utah
Geometric group theory and 3-manifolds hand in hand: the fulfillment of Thurston's vision for three-manifolds

Lauren Williams, University of California, Berkeley
Cluster algebras

January 6, 2012 (Boston, MA)

Jeffrey Brock, Brown University
Assembling surfaces from random pants: the surface-subgroup and Ehrenpreis conjectures

Daniel Freed, University of Texas at Austin
The cobordism hypothesis: quantum field theory + homotopy invariance = higher algebra

Gigliola Staffilani, Massachusetts Institute of Technology
Dispersive equations and their role beyond PDE

Umesh Vazirani, University of California, Berkeley
How does quantum mechanics scale?

January 6, 2011 (New Orleans, LA)

Luca Trevisan, Stanford University
Khot's unique games conjecture: its consequences and the evidence for and against it

Thomas Scanlon, University of California, Berkeley
Counting special points: logic, Diophantine geometry and transcendence theory

Ulrike Tillmann, Oxford University
Spaces of graphs and surfaces

David Nadler, Northwestern University
The geometric nature of the Fundamental Lemma

January 15, 2010 (San Francisco, CA)

Ben Green, University of Cambridge
Approximate groups and their applications: work of Bourgain, Gamburd, Helfgott and Sarnak

David Wagner, University of Waterloo
Multivariate stable polynomials: theory and applications

Laura DeMarco, University of Illinois at Chicago
The conformal geometry of billiards

Michael Hopkins, Harvard University
On the Kervaire Invariant Problem

January 7, 2009 (Washington, DC)

Matthew James Emerton, Northwestern University
Topology, representation theory and arithmetic: Three-manifolds and the Langlands program

Olga Holtz, University of California, Berkeley
Compressive sensing: A paradigm shift in signal processing

Michael Hutchings, University of California, Berkeley
From Seiberg-Witten theory to closed orbits of vector fields: Taubes's proof of the Weinstein conjecture

Frank Sottile, Texas A & M University
Frontiers of reality in Schubert calculus

January 8, 2008 (San Diego, California)

Günther Uhlmann, University of Washington
Invisibility

Antonella Grassi, University of Pennsylvania
Birational Geometry: Old and New

Gregory F. Lawler, University of Chicago
Conformal Invariance and 2-d Statistical Physics

Terence C. Tao, University of California, Los Angeles
Why are Solitons Stable?

January 7, 2007 (New Orleans, Louisiana)

Robert Ghrist, University of Illinois, Urbana-Champaign
Barcodes: The persistent topology of data

Akshay Venkatesh, Courant Institute, New York University
Flows on the space of lattices: work of Einsiedler, Katok and Lindenstrauss

Izabella Laba, University of British Columbia
From harmonic analysis to arithmetic combinatorics

Barry Mazur, Harvard University
The structure of error terms in number theory and an introduction to the Sato-Tate Conjecture

January 14, 2006 (San Antonio, Texas)

Lauren Ancel Myers, University of Texas at Austin
Contact network epidemiology: Bond percolation applied to infectious disease prediction and control

Kannan Soundararajan, University of Michigan, Ann Arbor
Small gaps between prime numbers

Madhu Sudan, MIT
Probabilistically checkable proofs

Martin Golubitsky, University of Houston
Symmetry in neuroscience

January 7, 2005 (Atlanta, Georgia)

Bryna Kra, Northwestern University
The Green-Tao Theorem on primes in arithmetic progression: A dynamical point of view

Robert McEliece, California Institute of Technology
Achieving the Shannon Limit: A progress report

Dusa McDuff, SUNY at Stony Brook
Floer theory and low dimensional topology

Jerrold Marsden, Shane Ross, California Institute of Technology
New methods in celestial mechanics and mission design

László Lovász, Microsoft Corporation
Graph minors and the proof of Wagner's Conjecture

January 9, 2004 (Phoenix, Arizona)

Margaret H. Wright, Courant Institute of Mathematical Sciences, New York University
The interior-point revolution in optimization: History, recent developments and lasting consequences

Thomas C. Hales, University of Pittsburgh
What is motivic integration?

Andrew Granville, Université de Montréal
It is easy to determine whether or not a given integer is prime

John W. Morgan, Columbia University
Perelman's recent work on the classification of 3-manifolds

January 17, 2003 (Baltimore, Maryland)

Michael J. Hopkins, MIT
Homotopy theory of schemes

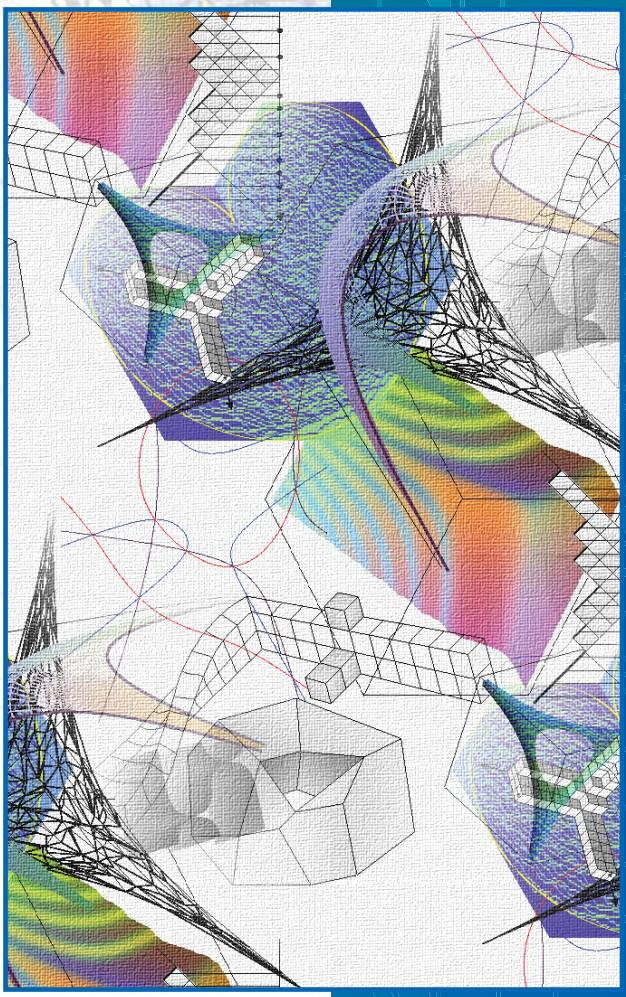
Ingrid Daubechies, Princeton University
Sublinear algorithms for sparse approximations with excellent odds

Edward Frenkel, University of California, Berkeley
Recent advances in the Langlands Program

Daniel Tataru, University of California, Berkeley
The wave maps equation

2020 CURRENT EVENTS BULLETIN

Committee



The back cover graphic is reprinted courtesy of Andrei Okounkov.

Cover graphic associated with Jordan S. Ellenberg's talk courtesy of Greg Herschlag and Jonathan Mattingly.

Cover graphic associated with Bjorn Poonen's talk courtesy of Bjorn Poonen.

Cover graphic associated with Sunčica Čanić's talk courtesy of Sunčica Čanić.

Cover graphic associated with Vlad C. Vicol's talk courtesy of:
Visualization: Amit Chourasia, SDSC/UCSD
Simulation: PK Yeung and Diego Donzis, GaTech
Computation: Datastar cluster at SDSC/UCSD
Link back URL: **visservices.sdsc.edu/projects/pkyeung/**