21-820. Partial Differential Equations Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

0. Introduction.

In memory of Jean LERAY, Nov 7, 1906 - Nov 10, 1998

In teaching any mathematical course where NAVIER-STOKES equations play a role, one must mention the pioneering work of Jean LERAY in the 1930s.

Some of the problems that Jean LERAY left unanswered are still open today, although some improvements were started by Olga LADYZHENSKAYA (*The mathematical theory of viscous incompressible flow*, Gordon and Breach, New York-London, 1963), followed by my advisor, Jacques-Louis LIONS, from whom I learned the basic principles for the mathematical analysis of these equations in the late 60s (J.-L. LIONS, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969).

In the announcement of the course, I had mentioned that I would start by recalling some classical facts about the way to use Functional Analysis for solving Partial Differential Equations of Continuum Mechanics, describe some fine properties of SOBOLEV spaces which are useful, and study in detail the spaces adapted to questions about incompressible fluids. I had stated then that the goal of the course was to describe some more recent mathematical models used in Oceanography, and show how some of them can be solved, and that, of course, I would point out the known defects of these models. I had mentioned that, for the Oceanography part of which I am no specialist, I would follow in particular a book ("*Analyse Mathématique en Océanographie*", Masson, Paris, 1997), written by one of my collaborators, Roger LEWANDOWSKI, who had learned about some of these questions from recent lectures of Jacques-Louis LIONS. I mentioned that I was going to distribute notes, from a course on Partial Differential Equation that I had taught a few years ago, but as I had not written the part that I had taught on STOKES and NAVIER-STOKES equations at the time, I was going to actualize the lecture notes from the graduate course that I had taught in Madison in 1974/75, where I had added small technical improvements from what I had learned ("Nonlinear partial differential equations using compactness method" *Report* #1584, *Mathematics Research Center, University of Wisconsin, Madison,* 1975). Finally, I had mentioned that I would write notes for the parts that I never covered in preceding courses.

I am not good at following plans. I started by reading about Oceanography in a book by A. E. GILL ("*Atmosphere-Ocean Dynamics*", Academic Press, 1982, International Geophysics Series), and I began the course by describing some of the basic principles that I had learned there. Then I did follow my plan of discussing questions of Functional Analysis, but I did not use any of the notes that I had written before. When I felt ready to start describing new models, Roger LEWANDOWSKI visited CMU and gave a talk in the Center for Nonlinear Analysis seminar, and I realized that there were some questions concerning the models and some mathematical techniques which I had not described at all, and I changed my plans. I opted for describing the general techniques for nonlinear partial differential equations that I had developped, Homogenization, Compensated Compactness and H-measures; there are obviously many important situations where they should be useful, and I found more important to teach them than to analyze in detail some particular models for which I do not feel yet how good they are. Regularly, I was trying to explain why what I was teaching had some connection with questions about fluids.

It goes with my philosophy to explain the origin of mathematical ideas when I know about them; perhaps it is because I have had to cope with an organized campaign of misattribution of my ideas to others, and therefore I like to mention why and when I had introduced an original idea. There might be some who use my ideas without quoting my name, and nevertheless do interesting things, and I will be very glad to mention their achievements, but most of the time those who intentionally steal others ideas do not have the right state of mind to create important concepts, and I leave the reader judge in the end.

I have also tried to induce mathematicians to learn more about Continuum Mechanics and Physics, listening to the specialists and then trying to put these ideas into a sound mathematical framework. I hope that some of the discussions in these lecture notes will help in this direction.

Finally, I want to apologize for some of the words which I use, which may have offended some. I have a great admiration for the achievements of physicists and engineers during the last Century (not mentioning Biology, which I never learned, or Chemistry which I can only try to understand in a better way after my program for understanding Physics and Continuum Mechanics has progressed enough), and a lot of the improvements in our lifes owe to their understanding, which is so different than the type of understanding that mathematicians are trained to achieve. If I write that something that they say does not make any sense, it is not a criticism towards physicists or engineers, it is a challenge to my fellow mathematicians that there is something there that mathematicians ought to clarify. I hope that more will understand the challenges, as the result will be that Science has indeed progressed.

Luc TARTAR
University Professor of Mathematics
Department of Mathematical Sciences
CARNEGIE MELLON University
Pittsburgh, May 7, 1999

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

1. Monday January 11.

The course meets at 12.30 pm in Physical Plant 300, on Monday, Wednesday, Friday. No classes on Monday January 18 (Martin Luther KING Jr.'s birthday), or Monday March 1 (Spring Mid-Semester break); Spring Break is March 21-28.

The goal of the course is to teach questions of Partial Differential Equations (PDE) for variants of NAVIER-STOKES equations occuring in Oceanography. My first plan was to follow a book by Roger LEWANDOWSKI, "*Analyse Mathématique en Océanographie*", Masson, Paris, 1997. However, after looking at his bibliography for textbooks on Oceanography, I selected a book by Adrian E. GILL, "*Atmosphere-Ocean Dynamics*", Academic Press, 1982 (International Geophysics Series, vol. 30), and I decided to first follow this book as an introduction to questions of Oceanography. It will have the advantage to start with simpler models, but it will also be a way to learn about the magnitude of different effects. It is indeed an important thing that mathematicians interested in Continuum Mechanics or Physics must learn, as they too often play with Partial Differential Equations without much knowledge of a few obvious facts concerning the "real world".

An example of this fact concerns incompressibility, an assumption that is often made by mathematicians, and which has the unphysical effect that some perturbations may travel at infinite speed. Although water looks difficult to compress, the speed of sound is of the order of 1.5 km/s, and for oceans which are thousands of kilometers wide the information does take some time to travel across. The speed of sound actually varies with temperature and pressure; for a practical salinity S = 35 (i.e. 35 grams of salt per kilogram of water), from 1,439.7 m/s at -2 °C to 1,547.6 m/s at 31 °C at atmospheric pressure, while at 6,000 m deep it varies from 1,542.6 m/s at -2 °C to 1,560.2 at +2 °C.

I will mostly use the metric system, not only because that is the one that I learned in France, but also because it is more natural, and physicists tend to use it anyway. The usual multiplicative prefixes are, deca = 10, hecto = 100, kilo = 1,000, mega = 10^6, giga = 10^9, tera = 10^{12}, peta = 10^{15}, exa = 10^{18}. The usual divisive prefixes are deci = 10^{-1}, centi = 10^{-2}, milli = 10^{-3}, micro = 10^{-6}, nano = 10^{-9}, pico = 10^{-12}, femto = 10^{-15}, atto = 10^{-18}.

The unit of length is the meter m (a kilometer km is 1,000 m, a nautical mile is 1.85318 km; a mile is 1.60932 km). The unit of surface is the square meter m^2 (a kilometer square km^2 is 10^6 m^2; an acre is 4,046.67 m^2). The unit of volume is the cubic meter m^3.

The unit of mass is the kilogram kg (a kilogram is 1,000 grams, a ton is 1,000 kg; a pound is .453 kg). The unit of density is the kilogram per cubic meter kg/m^3. Water has a density of approximately 1,000, while air has a density of approximately 1.29.

The unit of time is the second s (a minute is 60 s, an hour 3,600 s, a day 86,400 s, a year 3.1558 10^7 s).

The unit of velocity is the meter per second m/s (a km/hr is 0.2777 m/s, a knot is 0.51477 m/s). The unit of acceleration is the meter per second square m/s^2. Acceleration of gravity is approximately 9.8 m/s^2.

The unit of force is the Newton N, i.e. 1 $kg.m/s^2$. The unit of pressure is the Pascal Pa, i.e. 1 Newton/m^2 (a bar is 10^5 Pa). Atmospheric pressure is about one bar, and pressure increases of about one bar each time one goes down 10 meters in the ocean.

The unit of energy is the Joule, i.e. 1 $kg.m^2/s^2$ (one calorie is 4.184 J; a calorie is about the amount of energy that one needs for increasing the temperature of a gram of water by one degree, at usual temperatures). The unit of power is the Watt W, i.e. 1 $kg.m^2/s^3$ (a kiloWatt kW is 1,000 W).

The unit of temperature is the degree CELSIUS °C = degree KELVIN K; the temperature in degree C is the absolute temperature -273.15.

Oceanography, of course, is mostly concerned with large scale motions, and it concerns both Air and Sea, and is interested in Motion, Temperature and Salinity of water in the oceans.

Meteorology is the study of Motion, Temperature, Moisture Content and Pressure in the Atmosphere, but one discovers quickly the important role played by the oceans, and the exchanges between Air and Sea appear to be crucial.

Of course, everything starts with the Sun. The Earth moves at an average distance of 140 million kilometers from the Sun, and turns on itself with the inclination of its axis being responsible for the seasons. The average energy flux from the Sun at the mean radius of the Earth, called the "solar constant" S, has the value S = 1.368 kW/m^2. It means that if one was collecting all the energy from the Sun on a panel of one square meter, without any reflection, the panel being oriented perpendicularly to the direction from the Sun, one would get a power of 1.368 kilowatt (an air dryer works between 1 and 1.5 kilowatts, I believe). This energy corresponds to the black body radiation at a temperature of about 6,000 degrees, the surface temperature of the Sun (what happens inside the Sun does not matter for us). The blackbody radiation of a body at absolute temperature T is given by PLANCK's law, which describes the repartition of energy per unit volume of the "photons" having frequency near ν,

$$u(\nu)d\nu = \frac{8\pi\, h\, \nu^3}{c^3(e^{h\nu/k\,T} - 1)}\, d\nu,$$

where h is PLANCK's constant, approximately 6.62 10^{-34} J.s, c is the speed of Light, approximately 3 10^8 m/s, k the BOLTZMANN''s constant, approximately 1.38 10^{-23} J/K. A surface at absolute temperature T emits energy in all directions, and the power radiated by a surface dS in the solid angle d$^2\omega$ making an angle θ with the normal to the surface, between frequency ν and $\nu + d\nu$ is

$$\delta W = u(\nu)\, c\, \cos\theta\, dS \frac{d^2\omega}{4\pi}\, d\nu,$$

(computed as the energy of the photons contained in a cylinder based on dS with length c dt, and therefore having volume dS c dtcos θ), the dependence in $\cos\theta$ is LAMBERT's law. The total power radiated in the half space, obtained by integrating in directions and frequencies is

$$E = \sigma\, T^4,$$

where the STEFAN's constant σ is approximately 5.67 10^{-8} w/m$^2 K^4$. For the Sun, a large part of the energy is in the visible spectrum, betweem 0.4 and 0.8 μ (μ = micron = micrometer), therefore in very short wavelengths. Absorption by the atmosphere, water vapor, carbon dioxyde,..., are specific and vary a lot with frequency.

The average power received from the Sun per unit surface of the Earth is then S/4π = 344 W/m^2, but there is an albedo effect with an average coefficient $\overline{\alpha} = 0.3$, which denotes the proportion of the energy which is reflected; in average, the ground receives then about 240 W/m^2, and 100 W/m^2 is reflected back into space. The albedo is actually varying and depends a lot upon the cloud cover (Venus, covered by clouds, as an albedo of 0.6, while Mars, which has no clouds, has an albedo of 0.15). The albedo of land is about 0.15, and goes up to 0.2/0.3 for deserts, while land covered by snow or ice has an albedo of 0.6. Most of the oceans below latitude 40 has an albedo below 0.1, but the average is between 0.15 and 0.3.

If there was a purely radiative equilibrium between the energy received from the Sun and the energy radiated by the Earth ($(1 - \overline{\alpha})S/4 = \sigma\, T_g^4$), one would observe that the temperature of the ground would only be about 270 K at the equator, 170 K near the North pole and 150 K near the South pole.

It is the fluid cover, water and air, that makes a huge difference from these cold predictions. First radiation can be absorbed in the atmosphere, and the gases present, water vapor, carbon dioxide,...,, are important, creating a "greenhouse" effect. Second, there are convection effects which take the warm waters from the equator to the polar regions. HALLEY had already understood the opposite effects of the Sun and of Gravity in 1686: the Sun creates a horizontal variation, heating the waters near the equator much more than near the poles, while Gravity likes a vertical variation, drawing the cold waters to the bottom and making the warm waters rise to the surface; the competition between these two effects creates convection. MARSIGLI had already studied in 1679 the existence of a countercurrent in the depth of the Bosphorus, created by difference in salinity between the Mediterranean and the Black Sea.

A greenhouse functions from the idea that glass is transparent to the short wavelengths in the solar radiation (which is why glass looks transparent to us, who can only see wavelengths between 0.4 and 0.8 μ), while it is not transparent to the long wavelengths corresponding to the energy radiated from the ground. From PLANCK's law one deduces that the frequency ν_m where u(ν) is maximum is given by WIEN's law $\nu_m = 2.82\,\mathrm{k\,T/h}$, which shows that ν_m is linear in T; therefore if the Sun at 6,000 degrees has its maximum around 0.6 μ, a body at 300 K has its maximum at wavelengths 20 times larger, around 12 μ. If glass absorbs most of these long wavelengths, then it gets hotter (if it was absorbing all the radiations from the ground, it would take the temperature that the ground would have in absence of glass), and it emits its own radiation, both up and down and therefore the ground gets back a part of the energy that it had radiated away. If I is the downward flux coming from the Sun, which is in short wavelengths, U the upward flux corresponding to the temperature $\mathrm{T_g}$ of the ground, e the proportion of U absorbed by the glass and B the flux emitted in each direction (up or down) by the glass, then both U and B correspond to long wavelengths, and the equilibrium equation, under the hypothesis that the glass does not absorb any part of I, is $\mathrm{e\,U} = 2\mathrm{B}$, and $\mathrm{I} = (1\text{-e})\mathrm{U}+\mathrm{B}$, i.e. $\mathrm{I} = (1\text{-e}/2)\mathrm{U}$, and therefore $\sigma\,\mathrm{T_g^4} = \mathrm{U} = \mathrm{I}/(1\text{-e}/2)$, and $\mathrm{T_g}$ is higher by up to 19% in the case $\mathrm{e} = 1$, which gives $\sigma\,\mathrm{T_g^4} = 2\mathrm{I}$, as $2^{1/4}$ is around 1.19.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

2. Wednesday January 13.

An estimate of the radiation balance for the atmosphere is as follows. In order to work with percentages, let us assume that 100 units of incoming solar radiation arrive at the top of the atmosphere, all this energy being in the short wavelengths. 30 units will be sent back to space, corresponding to an albedo of 0.3, but these 30 units can be decomposed in 6 units backscattered by air, 20 units reflected by clouds and 4 units reflected by the ground surface. In the 70 units which are not reflected, 16 are absorbed in the atmosphere by water, dust, ozone, 3 are absorbed by the clouds and the remaining 51 are absorbed by land and oceans.

The surface also absorbs an estimated 98 units of long wave radiation, sent back from the atmosphere (this effect can be understood as in the way the greenhouse functions, rendering it even more efficient by a few layers of glass, as shown below). The net surface emission in the longer wavelengths of the infrared, excess of upward over downward radiation, is 21 units, the remaining upward flux of 30 units being by convection. The temperature of the ground corresponds then to $51 + 98 = 149$ units of radiated energy flux, instead of the 70 units emitted at the top of the atmosphere.

From the 21 units of excess infrared radiation from the surface, 15 are absorbed by water and carbone dioxyde and 6 units end up in space. Space also receives 38 units emitted by water and carbone dioxyde as well as 26 units emitted by clouds, so $6 + 38 + 26 = 70$ units of infrared radiation are sent to space, equilibrating the 70 units of solar radiation which had not been reflected (the albedo is only the fraction of solar radiation sent back). From the 30 units used in convection, 23 are used by latent heat for creating vapor from water and 7 corresponds to sensible heat flux, used to warm directly the atmosphere.

The balance for the atmosphere is then 16 units absorbed in solar radiation, 15 units absorbed in infrared radiation, 7 units received as heat, corresponding to the 38 units emitted.

The balance for the clouds is 3 units absorbed in solar radiation, 23 units invested in latent heat, corresponding to the 26 units emitted.

The processes of absorption and emission of radiation are not simple ones. It is worth noticing that they are highly frequency dependent: roughly speaking, there are frequencies which make a particular type of molecule vibrate and if radiation containing these frequencies goes through a gas containing these molecules, some of the energy at these frequencies will be absorbed by the gas. Conversely, a gas containing these molecules can emit spontaneously at those frequencies which it can absorb.

The generalized greenhouse having p well separated layers of glass which completely absorb the low frequencies but are transparent to the high frequencies is easy to compute: if I is received from solar radiation, $U = B_0$ is emitted by the surface and B_i is emitted on both sides of the glass layer $\#i$, counting from the surface, then the balance equations are

$$I = B_p;\ B_{i+1} + B_{i-1} = 2B_i \text{ for } i = 1, \ldots, p-1,$$

whose solution is

$$B_i = (p+1-i)I \text{ for } i = 0, \ldots, p, \text{ i.e. } U = (p+1)I.$$

Let us consider now the more general situation where the glass layer $\#i$ absorbs a proportion e_i of low frequency radiation but is completely transparent to high frequency radiation. Let B_i be the flux emitted on both sides by the glass layer $\#i$, for $i = 1, \ldots, p$, but now let $U = A_0$ and let A_i denote the ascending flux just above glass layer $\#i$ (this flux includes the ascending B_i) for $i = 1, \ldots, p$, so that $A_p = I$; similarly let D_i denote the descending flux just below glass layer $\#i$ (this flux includes the descending B_i) for $i = 1, \ldots, p$, and for convenience let $D_{p+1} = 0$. The balance equations are

$$\begin{aligned} e_i(A_{i-1} + D_{i+1}) &= 2B_i, i = 1, \ldots, p \\ A_i &= (1-e_i)A_{i-1} + B_i, i = 1, \ldots, p \\ D_i &= (1-e_i)D_{i+1} + B_i, i = 1, \ldots, p, \end{aligned}$$

and eliminating B_i gives

$$A_i = \Big(1 - \frac{e_i}{2}\Big) A_{i-1} + \frac{e_i}{2} D_{i+1}, i = 1, \ldots, p$$
$$D_i = \Big(1 - \frac{e_i}{2}\Big) D_{i+1} + \frac{e_i}{2} A_{i-1}, i = 1, \ldots, p,$$

which by adding gives $A_i + D_i = A_{i-1} + D_{i+1}$ for $i = 1, \ldots, p$, which is easy to see directly ($A_{i-1} + D_{i+1}$ is the amount received by the glass layer $\#i$, while $A_{i-1} + D_{i+1}$ is the amount transmitted by the glass layer $\#i$). Therefore $A_i - D_{i+1}$ is independent of $i = 0, \ldots, p$, and using the value for $i = p$ gives

$$A_i - D_{i+1} = I \text{ for } i = 0, \ldots, p,$$

from which one deduces $A_i = (1 - e_i/2)A_{i-1} + e_i/2(A_i - I)$, $i = 1, \ldots, p$, or

$$A_{i-1} = A_i + \frac{e_i/2}{1 - e_i/2} I, i = 1, \ldots, p,$$

and finally

$$U = \Big(1 + \sum_{i=1}^{p} \frac{e_i}{2 - e_i}\Big) I.$$

One deduces the case of a continuous absorbing media: if the layer between z and $z + dz$ absorbs a proportion $f(z)\,dz$ of low frequency radiation and is transparent to high frequency radiation, one finds $U = \big(1 + \frac{1}{2}\int_0^\infty f(z)\,dz\big) I$.

Of course, the radiative balance described before is not entirely radiative as it relies on observed distribution of water vapor, responsible for a large part of the absorption, and we will certainly need to understand a little more about the Thermodymadics of water and air in order to explain quantitatively the effects of convection, but one can give a quick qualitative explanation.

If there was no water vapor in the air and no other mechanism for absorbing radiation in the atmosphere, the atmosphere would stay cold and a larger amount of radiation would arrive at the surface: the surface would get warmer and the air in contact with the ground would also become warmer by conduction. Warm air is lighter than cold air, and therefore it rises; when one goes up the pressure decreases (the origin of the pressure is mostly the weight of the air above our heads), and when pressure decreases a gas expands and its temperature decreases, so the crucial problem is to compare the decrease in temperature due to expansion and the decrease in temperature due to altitude.

The *lapse rate* denotes the rate at which the temperature of the atmosphere decreases with height; the *dry adiabatic lapse rate* denotes the rate at which the temperature (of dry air) decreases because of expansion, and it is about 10 K/km: as long as the lapse rate is greater than the adiabatic lapse rate, warm air goes up, and this starts convection; convection carries heat upward, and therefore diminishes the lapse rate.

Of course, the real situation gets complicated by the fact that the atmosphere contains water vapor. If it contains a small amount of water vapor, convection will still occur when the dry adiabatic lapse rate is exceeded. However, air at a given temperature and pressure can only hold a certain amount of water vapor, and the amount of water vapor relative to this saturation value is called the *relative humidity*; when the relative humidity reaches 100%, water droplets condense in clouds, releasing the latent heat which had been required for creating the water vapor near the ground. The latent heat of vaporization $L_v(T)$ is given by the formula

$$L_v(T) \simeq 2.5008\,10^6 - 2.3\,10^3\, t\ J/kg,$$

where t denotes the temperature in degrees CELSIUS (so at boiling temparture at atmospheric pressure, $t = 100$, one needs 542 calories for vaporizing one gram of water). Latent heat represent more than 75% (23 units/30 units) of the heat transferred by convection. For saturated air, one must use the *moist adiabatic lapse rate*, which depends on temperature and pressure: in the lower atmosphere it is about 4 K/km at 20° C and 5 K/km at 10° C. Because the saturation amount of water decreases when one goes up, saturated air stays saturated when it goes up and the moist adiabatic lapse rate is used, but air cannot stay saturated when it goes down and the dry adiabatic lapse rate is used.

The preceding qualitative analysis was concerned with vertical differences in temperature, but there are also horizontal differences, as the Sun warms more the equatorial region than the polar regions. Although HALLEY (1686) had proposed the model that warm air rises near the equator and tropics and goes down at higher latitudes, it appears that the rising motion is concentrated in a narrow band called the Inter-Tropical Convergence Zone (ITCZ), usually found between 5° and 10° to the North of the equator (the trade wind, created by CORIOLIS force, push air from the tropics towards the equator); the regions of descending air are dry and include desertic regions found between latitudes 20° and 30°. In mid-latitudes, because of the rotation of the Earth, the motion produced by horizontal density gradients is mainly East-West and there is little meridional circulation; however large disturbances are created, cyclones and anticyclones, which are very efficient at transporting energy towards the poles.

It seems that ocean and atmosphere are equally important in transporting energy, the atmosphere being most important at 50° N and the ocean most important at 20° N, but the error in the estimate of the ocean transport at 20° N could be as high as 77%!

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

3. Friday January 15.

The fraction α of solar radiation reflected by the ocean is a function of the angle of incidence and of the surface roughness; at latitude below 30° it is less than 0.1, but it increases with latitude because of the angle of incidence of rays. Unlike the atmosphere, the ocean absorbs solar radiation rapidly and 80% is absorbed in the top 10 m, and the absorption rate is even greater in coastal areas, where a lot of suspended material exists.

As long wave radiation is rapidly absorbed in the atmposphere because of the presence of water vapor, it is absorbed very rapidly by the ocean, and absorption and emission occurs in a very thin layer, less than 1 mm thick.

Water is about 800 times denser than air (at atmospheric pressure), 1,025 kg/m^3 compared to 1.2-1.3 kg/m^3, and due to the strength of the gravitational restoring force, there is not much mixing, and transfer of properties between the two media takes place near the interface. The atmospheric pressure (1 bar) corresponds to the weight of the atmosphere, but it is just the weight of 10 m of water; the total mass of the ocean is 270 times the total mass of the atmosphere. The *specific heat* (heat capacity per unit mass) of water is 4 times that of air, and therefore the top 2.5 m of ocean has the same heat capacity than the whole atmosphere above (10^7 J/m^{2}K): raising the temperature of the atmosphere by 1 K can be done by lowering the temperature of 2.5 m of ocean by 1 K, or that of 25 m by 0.1 K. Heat can be stored in latent form, and the same amount of heat can be used to evaporate 4 mm of water, or to melt 30 mm of ice (the evaporation rate in the tropics is of order of 4 mm per day). Because of this ability to store heat, the ocean surface temperature changes by much smaller amounts than the land surface, which cannot store much heat. The excess heat gained in Summer is not transported to the Winter hemisphere, but is stored in the surface layers (about 100 m) and returned to the atmosphere in Winter.

So much for heat, now let us consider the balance of momentum and angular momentum. How are the winds produced and what determines their distribution?

HALLEY (1686) had tried to explain the trade winds, which blow from the tropics to the equator (NE to SW in the northern hemisphere and from SE to NW in the southern hemisphere), but his idea only explains the meridional circulation in the Inter-Tropical Convergence Zone described before, which is not called after him now, but after HADLEY who gave a better explanation in 1735. If there was no friction, the equator being longer by 2,083 miles than the tropics (at latitude 23° 27′, i.e. around 23.5°), he argued that air at rest at the tropics would acquire a westward motion of 2.083 miles/day when transported to the equator, but as the observed velocity is not as high as this velocity of almost 140 km/hr, he argued that there was some friction and air had also been given a correcting eastward push from the surface of Earth. He also argued that there must exist opposite winds somewhere in order to compensate the trade winds: this is related to the conservation of *angular momentum*, but it is not valid for the northward/southward component (the moment of a vector parallel to the axis is 0). If the average eastward *force* (or rate of tranfer of eastward momentum) per unit area acting on the surface at latitude φ is $\tau^x(\varphi)$, then the average *torque* (or rate of transfer of angular momentum) per unit area about the axis is $a\,\tau^x(\varphi)\cos\varphi$, where a is the radius of the Earth; the area of the strip between latitudes φ and $\varphi+d\varphi$ is $2\pi\, a^2\cos\varphi\, d\varphi$, so the torque on this strip is $2\pi\, a^3\tau^x(\varphi)\cos^2\varphi\, d\varphi$, and the balance of angular momentum for the Earth is then

$$\int_{-\pi/2}^{+\pi/2}\tau^x(\varphi)\cos^2\varphi\, d\varphi = 0.$$

The force of the atmosphere on the underlying surface is exerted in two ways: one is the force exerted on irregularities in the surface associated with the pressure differences across the irregularities, and the second is by viscous stresses. The irregularities on which forces are exerted may vary in size from mountain ranges down to trees, blades of grass and ocean surface waves. When the irregularities are small enough (as is the case over the ocean), the associated force per unit area added to the viscous stress is called the surface stress or *wind* stress. There is a westward stress in the trade wind zones (latitudes below 30°) and therefore an eastward stress is required at higher latitudes, and one does observe westerly (i.e. eastward) winds at those latitudes. In France there is a definitely dominant wind from the West (I was taught that this is why industrial plants were first built on the East side of Paris so that the wind would push the smoke away from the city; of course the expensive residential areas then developed in the West part of the city!). In the southern hemisphere the wind is indeed from the West, but it is extremely strong, probably because there is no land there to slow it down, and the sailors traveling in latitudes 40° S and 50° S have coined the denominations "Roaring Forties" and "Furious Fifties".

Winds are produced in the atmosphere, a result of the radiative forcing, which creates horizontal and vertical gradients, and it is difficult to understand these effects without writing partial differential equations models, but winds are of the order of 10 m/s. Winds transfer momentum to the ocean, producing currents, but the exact process is not so simple as a shear flow near the interface becomes unstable and turbulent eddies are formed (so there are gusts of wind), and one needs to average over time (a few minutes for points a few meters above the ground): the mean stress τ is equal to the mean value of $\rho\, u\, w$, where u and w are the horizontal and vertical components of the velocity, and ρ is the density. If one measures u at some level, one can induce by dimensional analysis that $\tau = c_D \rho\, u^2$, where c_D is a dimensionless parameter called the *drag coefficient*, which depends upon the roughness of the interface and the lapse rate. The drag coefficient c_D for the ocean surface is found to increase with wind speed: for low speeds it is around $1.1\ 10^{-3}$, but for speeds between 6 m/s and 22 m/s one often uses the relation $10^3 c_D = 0.61 + 0.063\, u$.

There are however other formulas derived, and we will only remember that various effects have to be modeled near the interface and that finding correct boundary conditions is important.

We can now start deriving the basic partial differential equations which describe the variations in space and time of the various physical quantities. We start by investigating the equation of conservation of mass, which is certainly true away from the interface, but not near the interface, where water is lost by evaporation and gained by precipitation.

Notice that the conservation of salt is also important, and salinity increases because of evaporation and decreases because of precipitations. As was first noticed by MARSIGLI in 1681 the difference in salinity between the Black Sea and the Mediterranean is responsible for a deep undercurrent in the Bosphorus, the lighter less salty waters from the Black Sea flowing on the surface towards the Mediterranean, while the heavier more salty waters from the Mediterranean flow below towards the Black Sea.

Although various coordinate systems are used, it is useful to use a Cartesian system of coordinates (with an orthonormal basis) in order to derive more easily the equations. Position is denoted by $x = (x_1, x_2, x_3)$, time by t, velocity by $u = (u_1, u_2, u_3)$, density by ρ. In the Lagrangian point of view, one refers quantities to the initial position ξ of the particle, i.e. the components x_i are expressed in terms of ξ and t, and so $v_i = \frac{\partial x_i}{\partial t}$. In the Eulerian point of view, one refers quantities to the actual position x of the particle and t, and this is the point of view that we will consider, and we will see that conservation of mass takes the form

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3} \frac{\partial(\rho\, u_i)}{\partial x_i} = 0,$$

the second part being abbreviated as $div(\rho\, u)$.

A classical way for deriving this equation, although in a formal way where one assumes that everything is smooth enough, is as follows. One considers a set of material points occupying the domain $\omega(0)$ at time 0 and $\omega(t)$ at time t, and one writes that $\int_{\omega(t)} \rho\, dx$ is independent of t. The variation on the intersection of $\omega(t)$ and $\omega(t + \delta\, t)$ is estimated as $\delta\, t \int_{\omega(t)} \frac{\partial \rho}{\partial t}\, dx$. The part on $\omega(t + \delta\, t) \setminus \omega(t)$ is estimated by a surface integral on $\partial\omega(t)$: as a point x of the surface moves of about $u(x,t)\delta\, t$, it is like if the surface was pushed in the direction of the exterior normal ν of an amount $(u.\nu)\delta\, t$, and therefore the second part is estimated as the surface integral $\delta\, t \int_{\partial\omega(t)} \rho(u.\nu)\, dx'$, and the conservation of mass takes then the form

$$\int_{\omega(t)} \frac{\partial \rho}{\partial t}\, dx + \int_{\partial\omega(t)} \rho(u.\nu)\, dx' = 0.$$

One then transforms the boundary integral $\int_{\partial\omega(t)} \rho(u.\nu)\,dx'$ into $\int_{\omega(t)} div(\rho\,u)\,dx$ by GREEN's formula, which givess

$$\int_{\omega(t)} \Big(\frac{\partial \rho}{\partial t} + div(\rho\,u)\Big)\,dx = 0,$$

and varying $\omega(t)$ gives the result.

Although this type of derivation is common practice among physicists, it is useful for mathematicians to think about the hypotheses needed for carrying out the various steps. One can also try to derive that same basic equation in other ways.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

4. Wednesday January 20.

The argument using GREEN's formula, which is a question of integration by parts, can be shown to hold in SOBOLEV spaces.

For an open set Ω of R^N and $1 \le p \le \infty$, the SOBOLEV space $W^{1,p}(\Omega)$ is defined as

$$W^{1,p}(\Omega) = \Big\{u \in L^p(\Omega) : \frac{\partial u}{\partial x_j} \in L^p(\Omega) \text{ for } j = 1, \ldots, N\Big\},$$

equipped with the norm

$$||u||_{W^{1,p}(\Omega)} = \Big[\int_\Omega \Big(\frac{1}{\lambda^p}|u|^p + \sum_{j=1}^N \Big|\frac{\partial u}{\partial x_j}\Big|^p\Big)\,dx\Big]^{1/p},$$

where λ is a characteristic length (which mathematicians usually take equal to 1, giving then the impression that they can add quantities measured in different units without being the least surprised).

Elements of $L^p(\Omega)$ are classes of LEBESGUE-measurable functions (two functions equal almost everywhere being identified), and in order to restrict functions of $W^{1,p}(\Omega)$ on the boundary $\partial\Omega$ of Ω, which is usually a set of measure zero, one has to be careful.

The derivatives $\frac{\partial u}{\partial x_j}$ are not computed in a classical way, but are weak derivatives, and this idea was introduced by Sergei SOBOLEV, and used by Jean LERAY around 1930 in order to define weak solutions of NAVIER-STOKES equation (which were shown to exist globally in time by Olga LADIZHENSKAYA). LERAY had qualified these weak solutions as "turbulent", but although uniqueness of these weak solutions is still an open question in three dimensions, few people believe that LERAY's ideas about turbulence were right, and the ideas that KOLMOGOROV introduced much later have received much more attention (which does not mean that they were right either). For a function $u \in C^1(\Omega)$ and $\varphi \in C^1_c(\Omega)$, the space of C^1 functions with compact support in Ω, one has $\int_\Omega \frac{\partial u}{\partial x_j}\varphi\,dx = -\int_\Omega u\frac{\partial \varphi}{\partial x_j}\,dx$, and this formula helps in defining the weak derivatives of u: one says that $\frac{\partial u}{\partial x_j} = f \in L^p(\Omega)$ if for all $\varphi \in C^1_c(\Omega)$ one has $-\int_\Omega u\frac{\partial \varphi}{\partial x_j}\,dx = \int_\Omega f\,\varphi\,dx$. The theory of weak solutions was put into a more general framework around 1950 by Laurent SCHWARTZ in his theory of distributions, which helps understand the *linear* partial differential equations with *smooth* coefficients. The way to attack the basic equations of Continuum Mechanics, which can have discontinuous coefficients because of the presence of interfaces, was developed in the 1960s, mostly along the lines that SOBOLEV and LERAY had pioneered.

The case $p = 2$ plays a special role and $W^{1,2}(\Omega)$ is also denoted as $H^1(\Omega)$, and generally one can defined $H^s(R^N)$ for $s \in R$ by FOURIER transform, and $H^s(\Omega)$ for $s \ge 0$ by restriction to Ω (the case $s = 0$ corresponds to $L^2(\Omega)$), and I will use the notation H^s instead of $W^{s,2}$. One should be aware that some other spaces, which are natural in Harmonic Analysis, are also denoted in the same way, the HARDY spaces, and I will write them with $\mathcal{H}$ instead of H when I will encounter them.

For a bounded open set Ω with a LIPSCHITZ boundary, i.e. an open set which is locally on one side of its boundary which has locally an equation $x_N = F(x_1, \ldots, x_{N-1})$ in an orthonormal basis, with F LIPSCHITZ continuous, one can define the restriction of a function in $W^{1,p}(\Omega)$, called its trace, and this is done by an argument of Functional Analysis. One first shows that $C^1(\overline{\Omega})$, the space of restrictions to $\overline{\Omega}$ of functions which are C^1 on an open set containing $\overline{\Omega}$, is dense in $W^{1,p}(\Omega)$ for $1 \le p < \infty$. Then one shows that the linear mapping obtained by restricting a function of $C^1(\overline{\Omega})$ to the boundary (giving a LIPSCHITZ continuous function on the boundary) is continuous if one puts on $C^1(\overline{\Omega})$ the norm of $W^{1,p}(\Omega)$ and on its traces the norm of $L^p(\partial\Omega)$ (for the natural $(N-1)$-dimensional measure on the boundary). The characterization of the space of traces, due to Emilio GAGLIARDO, gives indeed $L^1(\partial\Omega)$ for the case $p = 1$ and the obvious $W^{1,\infty}(\partial\Omega)$ for the case $p = \infty$, and for $1 < p < \infty$ an interpolation space of functions having $1/p'$ derivatives in $L^p(\partial\Omega)$. However one does not need the precise characterization of the space of traces for proving the formula of integration by parts (which implies GREEN's formula)

$$\int_\Omega \frac{\partial u}{\partial x_j}\,dx = \int_{\partial\Omega} u\,\nu_j\,d\sigma \text{ for all } u \in W^{1,1}(\Omega),$$

where $d\sigma$ is the $(N-1)$-dimensional HAUSDORFF measure and ν is the exterior normal to $\partial\Omega$, which exists $d\sigma$ almost everywhere.

The idea for proving the estimate of the integral on $\omega(t+\delta t)\setminus\omega(t)$ is shown on a simpler example. The field u is assumed to be of class C^1; using the hyperplane $x_N=0$ instead of the boundary $\partial\Omega$, one has to integrate on a strip swept by the points $x'+s\,u(x')$ for $x'\in R^{N-1}$ (or a piece of the hyperplane) when s varies from 0 to and δt. One observes that if $u_N(x')\neq 0$, then the mapping $(x',s)\mapsto x'+s\,u(x')$ is a local diffeomorphism, whose Jacobian is precisely $u_N(x')$, and if one integrates a uniformly continuous function φ on the strip, the integral is easily seen to be equivalent to $\delta t\int_{R^{N-1}}\varphi\, u_N(x')\,dx'$.

In the preceding computation, one should have followed the solution of an ordinary differential equation during a time δt and not just followed the tangent (as in the EULER method for approximating solutions of ordinary differential equations). We will see now a different derivation of the equation expressing the conservation of mass, based on the analysis of ordinary differential equations, and it will also make the connection with the Lagrangian point of view.

Let us assume that $u(x,t)$ is of class C^1 in x and t, then the differential equation

$$\frac{Dx}{Dt}=x'(t)=u(x(t),t)\text{ for }t>0;\ x(0)=\xi,$$

describes the position at time t of a material point starting at ξ at time 0; here $\frac{D}{Dt}$ is the partial derivative in t when ξ is fixed, but we reserve $\frac{\partial}{\partial t}$ to denote the partial derivative in t when x is fixed, i.e. in the Eulerian point of view; $\frac{D}{Dt}$ is called the material derivative. It seems natural to ask for uniqueness of a solution, and the classical condition is the local version of the following global LIPSCHITZ condition

$$|u(x,t)-u(y,t)|\leq\lambda(t)|x-y|\text{ for all }x,y\text{ and }t\in(0,T),$$

with $\lambda\in L^1(0,T)$. There is a small improvement due to OSGOOD, which gives uniqueness when one only assumes that

$$|u(x,t)-u(y,t)|\leq\omega(|x-y|)\text{ for all }x,y\text{ and the modulus of uniform continuity }\omega\text{ satisfies }\int_0^1\frac{ds}{\omega(s)}=+\infty.$$

This gives $x=\Phi(\xi,t)$, and with u of class C^1 in (x,t), it is not difficult to prove that $\xi\mapsto\Phi(\xi,t)$ is a local diffeomorphism and that the Jacobian matrix $\frac{\partial\Phi}{\partial\xi}$ satisfies the linear differential equation

$$\frac{D\frac{\partial\Phi}{\partial\xi}}{Dt}=\frac{\partial u}{\partial x}\frac{\partial\Phi}{\partial\xi}\text{ on }(0,T);\ \frac{\partial\Phi}{\partial\xi}(0)=I,$$

so that the Jacobian determinant $det\frac{\partial\Phi}{\partial\xi}$ satisfies

$$\frac{D(det\frac{\partial\Phi}{\partial\xi})}{Dt}=Trace\Big(\frac{\partial u}{\partial x}\Big)det\frac{\partial\Phi}{\partial\xi}=div(u)\,det\frac{\partial\Phi}{\partial\xi}\text{ on }(0,T);\ det\frac{\partial\Phi}{\partial\xi}(0)=1.$$

As $det\frac{\partial\Phi}{\partial\xi}$ represents the increase in volume by the transformation $\xi\mapsto x(t)$, conservation of mass can then be written as

$$\rho(x(t),t)det\frac{\partial\Phi}{\partial\xi}=\rho(\xi,0)\text{ almost everywhere,}$$

and the equation for the Jacobian determinant can therefore be written as

$$\frac{D(\frac{\rho(\xi,0)}{\rho})}{Dt}=div(u)\,\frac{\rho(\xi,0)}{\rho},$$

or equivalently (using $\rho(\xi,0)>0$)

$$\frac{D\rho}{Dt}+\rho\,div(u)=0,$$

which is the desired equation as

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \sum_{j=1}^{N} u_j \frac{\partial}{\partial x_j}.$$

It seems then reasonable to admit the derived form of conservation of mass, but the regularity hypotheses invoked for proving it are a little too strong in some situations. For NAVIER-STOKES equation, under the assumption that the fluid is incompressible and that the viscosity is independent of temperature (so that one just forgets about the equation of conservation of energy), one knows uniqueness of the solution in 2 dimensions, and the solution is smooth enough if the initial data are smooth enough. However, uniqueness is not known in 3 dimensions, and it is only for sufficiently small smooth data that one knows that the solution stays smooth; the dissipation of energy by viscosity gives directly that $u \in L^2\big(0,T;H^1(\Omega;R^3)\big)$, and by improving an argument of Ciprian FOIAS, I proved that $u \in L^1(0,T;Z)$, with Z a little smaller that $W^{1,3}(\Omega;R^3)$ (so that $Z \subset C^0(\overline{\Omega};R^3)$, for example), but that is far from the $W^{1,\infty}$ regularity required for deriving the equation.

For incompressible 2-dimensional EULER equation, there is a global existence (and maybe uniqueness result) due to T. KATO, I believe. The vorticity $\omega = \frac{\partial u_1}{\partial x_2} - \frac{\partial u_2}{\partial x_1}$ is transported by the flow, i.e. satisfies $\frac{D\omega}{Dt} = 0$, and therefore if the initial vorticity is in $L^\infty(R^2)$ it stays in this space. At a given time one has then $curl(u) \in L^\infty$ and $div(u) = 0$, but that does not implies $u \in W^{1,\infty}(R^2,R^2)$ (L^∞ is not a good space for singular integrals).

The singular integrals which often appear in linear systems of partial differential equations with constant coefficients in R^N are convolution equations with a kernel which is homogeneous of degree $-N$ and whose integral on the sphere is 0; very often they are polynomials in the RIESZ operators R_j, which are the natural generalization to R^N of the HILBERT transform in R. Singular operators act on spaces like $C^{k,\alpha}$ (results proved in the 1920s/30s by GIRAUD, I believe), and were extended in the 1950s to L^p with $1 < p < \infty$ by CALDERÓN and ZYGMUND. For the case $p = \infty$, they act on the bigger space BMO (bounded mean oscillations), introduced by JOHN and NIRENBERG. For the case $p = 1$, they act on the smaller HARDY space $\mathcal{H}^1$. Charles FEFFERMAN proved that the dual of $\mathcal{H}^1$ is BMO.

One can then say that $curl(u) \in L^\infty$ and $div(u) = 0$ imply that all the derivatives $\frac{\partial u_i}{\partial x_j}$ belong to BMO, but one can also use another space, the ZYGMUND space Λ_1, which serves as a replacement for the space of LIPSCHITZ functions (as it is an interpolation space between $C^{0,\alpha}$ and $C^{1,\beta}$, it inherits the property that singular integrals act in a continuous way over it). One can they say that $curl(u) \in L^\infty$ and $div(u) = 0$ imply that $u \in \Lambda_1$, and that means that there exists a constant M such that $|u(x+h) + u(x-h) - 2u(x)| \leq M|h|$ for all x, h, and this implies $|u(x+h) - u(x)| \leq C|h| \log(|h|)$ for $|h|$ small, and OSGOOD variant applies.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

5. Friday January 22.

I consider now another approach for deriving the equation of conservation of mass, based on the use of "particles".

One uses partial differential equations in Continuum Mechanics, and one knows that they are not valid at a too small scale, because Matter is made of particles. If one uses the language of physicists, one calls microscopic the level where there are particles and macroscopic our level, and the intermediate levels are called mesoscopic. At usual temperature and pressure, a mole of a gas, either monatomic like one of the rare gases (Ne, Ar, Kr, Xe, Rn) or diatomic (like H^2, N^2, O^2) or a mixture (like air = 79% N^2 + 20% O^2 + 1% Ar) occupies 22.4 liters (one liter = 1 $\mathrm{dm}^3 = 10^{-3}$ m^3) and contains a number of particles around 6.023 10^{23} (the AVOGADRO number). In a crystal interatomic distances are measured in angströms (1 Å$= 10^{-10}$ m). In a liquid, a mole of water (18 grams of H^2O) occupies approximately 18 cm^3, but when transformed into vapour it occupies 22,400 cm^3 so that the average distance between molecules in the vapour is a little more than 10 times that in the liquid, and as the average volume around a molecule in the liquid is about 3 10^{-23} cm^3, it corresponds to an average distance of about 3 angströms.

Of course, once one tries to describe what happens at the level of these particles one discovers that they do not behave like clasical particles, because they are actually waves (but not necessarily described by SCHRÖDINGER equation, which is only an approximation).

In the general theory of Homogenization (intertwined with the Compensated Compactness theory) that I developed in the 70s with François MURAT (extending earlier results of the Italian school, Sergio SPAGNOLO, Ennio DE GIORGI), there are no hypothesis of periodicity (but in the work of Enrique SANCHEZ-PALENCIA and that of Ivo BABUŠKA, who coined the term Homogenization, there were periodicity assumptions). I plan to continue to use the term Homogenization for describing our general approach, i.e. without any periodicity hypotheses, and it would be natural that some of the others who apply our ideas only to periodic situations would at least ackowledge that they use our approach of H-convergence. In that general theory, one starts with partial differential equations at a level which I qualified in the early 70s of microscopic and one tries to discover which constitutive relations and which balance relations (and therefore which effective equations) one should use at our level, for which I used the term macroscopic. I first heard the term mesoscopic in the early 90s, and as our microscopic level is obviously a level for which the equations of Continuum Mechanics apply, it is then one of the various mesoscopic levels that physicists mention. My use was not in opposition with what physicists often do in order to understand an important effect: they select the smallest level where this effect takes place and they seem to neglect the smaller scales, but they actually use an ad hoc theory for summarizing what they know about what happens at smaller scales (or what they believe must happen there); then they try to understand what happens at this particular level in order to derive the effective equations that they will use at the next higher level (which might well be the microscopic level of the atoms or one of these intermediate mesoscopic levels). In the Continuum Mechanics approach, the lower scales are summarized in the laws of "Thermodynamics" which constrain the constitutive relations that one may use, and one of the goals of my program of studying the evolution of microstructures in partial differential equations is to avoid postulating the constitutive laws and instead to show how to deduce them from more basic principles. Although my approach explains in some way why high frequency solutions of some partial differential equations may behave like particles, it still faces a few theoretical obstacles and cannot explain yet how to interpret what physicists say when they use a discrete description, starting from atoms arranged in a crystalline way (or a polycrystalline way with important effects at the grain boundaries), describing defects in the crystalline arrangement and how these defects move around in order to explain the effects of Plasticity which do limit the applicability of theories like Elasticity (Owen RICHMOND described a few years ago his program, very similar to mine, but which deals precisely with the scales that I cannot explain at the moment).

The "particles" that I will now use have nothing to do with the "real" particles that one encounters every few angströms in polycrystalline solids, or a little further apart in liquids or even in gases. The real particles are actually concentrated packets of highly oscillating waves, while the particles that I will use should better be called macroscopic particles; they provide a convenient way for approaching the solutions of partial differential equations, whether smooth or not so smooth, and they are more widely used now in that way as a numerical approach than thirty years ago when I was learning Numerical Analysis, as in those days computers were not so powerful.

At the level of describing conservation of mass, the argument that I will use is quite similar to that which is used in Classical Mechanics, where a rigid solid is replaced by its center of mass, and a point M_i of mass m_i moving at velocity V_i may well represent a body having M_i as its center of mass, m_i as its total mass, and $m_i V_i$ as its total momentum.

From a mathematical point of view, we need to use DIRAC masses and more general objects called RADON measures. However, it will not be enough and in a second step we will need to introduce much more general objects called distributions, but although we will not need so much of the theory of distributions due to Laurent SCHWARTZ, I will give some general definitions.

Physicists describe the DIRAC "function" at the point $a \in R^N$ as the function which is 0 outside a, $+\infty$ at a and has integral 1, and mathematicians are quick to mention that there is no such function, but that is not so important now that Laurent SCHWARTZ has found a mathematical explanation for many (but not for all) strange formulas that physicists had obtained by using these inexistent functions. DIRAC was not the first to use such a function, and G. BIRKHOFF mentions in "A source book in classical analysis" that KIRCHHOFF had used such a function around 1890. Indeed it is just the idea of a point mass, and it is not for that simple idea that DIRAC should be mentioned, but for the much bolder idea that one could use the derivative of that "function".

Let Ω be an open set of R^N. The space $L^1_{loc}(\Omega)$ is the space of (classes of) LEGESGUE-measurable functions such that for every compact $K \subset \Omega$, $M_K = \int_K |f(x)|\,dx < \infty$ (it is not a BANACH space, but a FRÉCHET space). For $f \in L^1_{loc}(\Omega)$ and $\varphi \in C_c(\Omega)$, the space of continuous functions with compact support in Ω, one can define $\int_\Omega f(x)\varphi(x)\,dx$ and one has $|\int_\Omega f(x)\varphi(x)\,dx| \le M_K \max_{x\in K}|\varphi(x)|$ for all functions $\varphi \in C_c(\Omega)$ which have their support in K. A RADON measure μ in Ω is a linear form $\varphi \mapsto \langle \mu, \varphi\rangle$ on $C_c(\Omega)$ (whose elements are called test functions), satisfying similar bounds, i.e. for every compact $K \subset \Omega$ there exists a constant $C(K)$ such that

$$|\langle \mu, \varphi\rangle| \le C(K) \max_{x\in K} |\varphi(x)| \text{ for all functions } \varphi \in C_c(\Omega) \text{ which have their support in } K.$$

The DIRAC mass at $a \in \Omega$ is an example of a RADON measure: it corresponds to $\langle \mu, \varphi\rangle = \varphi(a)$ for all $\varphi \in C_c(\Omega)$ (and $C(K) = 1$ if $a \in K$, $C(K) = 0$ if $a \notin K$).

There is a topology on $C_c(\Omega)$ for which the dual space is $\mathcal{M}(\Omega)$, the space of all RADON measures in Ω; we will not need to know the topology of $C_c(\Omega)$, but there is a useful topology on $\mathcal{M}(\Omega)$, the corresponding weak $\star$ topology $\sigma\big(\mathcal{M}(\Omega), C_c(\Omega)\big)$, also called the *vague* topology. A sequence μ_n converges vaguely to μ_∞ if and only if $\langle \mu_n, \varphi\rangle$ converges to $\langle \mu_\infty, \varphi\rangle$ for all $\varphi \in C_c(\Omega)$; however this does not define the topology, because that topology is not metrizable (but its restrictions to bounded sets, suitably defined, are metrizable). For example, if f_n is a bounded sequence in $L^1(\Omega)$ satisfying $\int_K |f(x)|\,dx \to 0$ for every compact K not containing 0 and $\int_\omega f(x)\,dx \to 1$ for some open set ω containing 0, then the sequence of measures f_n (an abuse of language for $f_n\,dx$) converges vaguely to δ_0, the DIRAC mass at 0. This explains how to handle the DIRAC "function" idea, by using functions sufficiently concentrated near the point and passing to the limit.

In order to understand what the derivative of a DIRAC mass could be, a natural idea is to use a sequence f_n made of smooth functions, and then take the limits of their derivatives, but that require introducing more general objects, the distributions of Laurent SCHWARTZ (RADON measures will appear then to be distributions of order 0).

For distributions in Ω, the test functions are taken in $C_c^\infty(\Omega)$, the space of infinitely differentiable functions with compact support in Ω; this space was denoted $\mathcal{D}(\Omega)$ by SCHWARTZ, so the space of distributions which is its dual is denoted $\mathcal{D}'(\Omega)$. There are plenty of such functions, but just one with a nonzero integral has to be constructed explicitly, for example $u(x) = exp\big(-\frac{1}{1-|x|^2}\big)$ for $|x| < 1$ and $u(x) = 0$ for $|x| \ge 1$ has for support the closed unit ball.

A distribution T is then defined as a linear form $\varphi \mapsto \langle T, \varphi\rangle$ on $C_c^\infty(\Omega)$, such that for every compact $K \subset \Omega$ there exists a constant $C(K)$ and an integer $m(K) \ge 0$ such that

$$|\langle T, \varphi\rangle| \le C(K) \max_{x\in K} \max_{|\alpha|\le m(K)} |D^\alpha \varphi(x)| \text{ for all functions } \varphi \in C_c^\infty(\Omega) \text{ which have their support in } K.$$

For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_N)$, D^α denotes the operator $\left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \ldots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_N}$; the length of α is $|\alpha| = |\alpha_1| + \ldots + |\alpha_N|$. If $m(K)$ can be taken independent of K, the distribution T is said to be of finite order and the smallest possible value of $m(K)$ is called the order of T, so that RADON measures are exactly the distributions of order 0.

By analogy with the formulas for smooth functions, one can multiply a distribution T by a C^∞ function ψ (or by a function of class C^m if the distribution has a finite order $\leq m$), as

$$\langle \psi\, T, \varphi \rangle = \langle T, \psi\, \varphi \rangle \text{ for all } \varphi \in C_c^\infty(\Omega),$$

and one can define derivatives of T as

$$\langle D^\alpha\, T, \varphi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha\, \varphi \rangle \text{ for all } \varphi \in C_c^\infty(\Omega).$$

For example, if H denotes the HEAVISIDE function $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x > 0$, then one quickly checks that $\frac{dH}{dx} = \delta_0$. A simple "paradox" will show that not all formulas extend to distributions: let u be the sign function ($u = -1 + 2H$), so that $\frac{du}{dx} = 2\delta_0$, and notice that $u^2 = 1$ and $u^3 = u$, but the formula $\frac{du^3}{dx} = 3u^2 \frac{du}{dx}$ does not hold as the left side is $2\delta_0$ while the right side is $6\delta_0$; using SOBOLEV imbedding theorem, the N-dimensional formula $\frac{\partial u^3}{\partial x_j} = 3u^2 \frac{\partial u}{\partial x_j}$ is actually valid on $W^{1,p}(R^N)$ for $p \geq \frac{3N}{N+2}$.

Let us consider now a finite number of point masses moving around, the particle # i having mass m_i, position $M_i(t)$ and velocity $V_i(t) = \frac{dM_i}{dt}$ at time t. Conservation of mass is expressed by the fact that m_i is independent of t; although two particles can go through the same point at some time, there is no exchange of mass between them during these "collisions". The analog of a smooth density $\rho(x,t)$ is the measure μ defined by

$$\langle \mu, \varphi \rangle = \sum_i \int_0^T m_i \varphi(M_i(t), t)\, dt,$$

and the analog of the mass density at time t is the measure $\mu_t = \sum_i m_i \delta_{M_i(t)}$. We introduce then a *momentum* measure π by

$$\langle \pi, \varphi \rangle = \sum_i \int_0^T m_i V_i(t) \varphi(M_i(t), t)\, dt,$$

and the analog of the momentum density at time t is the measure $\pi_t = \sum_i m_i V_i(t) \delta_{M_i(t)}$. Notice that V_i are vectors, and therefore π is vector valued measure, and its components will be written as $(\pi)_j$ for $j = 1, \ldots, N$. Then conservation of mass implies that

$$\frac{\partial \mu}{\partial t} + \sum_{j=1}^N \frac{\partial (\pi)_j}{\partial x_j} = 0.$$

Indeed for a test function $\varphi \in C_c^\infty\big(\Omega \times (0,T)\big)$, it means that $\left\langle \mu, \frac{\partial \varphi}{\partial t} \right\rangle + \sum_{j=1}^N \left\langle (\pi)_j, \frac{\partial \varphi}{\partial x_j} \right\rangle = 0$, and this means that $\sum_i m_i \int_0^T \frac{\partial \varphi}{\partial t}(M_i(t), t)\, dt + \sum_{j=1}^N \sum_i m_i \int_0^t (V_i)_j(t)\, \frac{\partial \varphi}{\partial x_j}(M_i(t), t)\, dt = 0$, which follows from the fact that the coefficient of m_i is 0; this coefficient is $\int_0^T \left[\frac{\partial \varphi}{\partial t}(M_i(t), t)\, dt + \sum_{j=1}^N (V_i)_j(t)\, \frac{\partial \varphi}{\partial x_j}(M_i(t), t)\right] dt$, and as the bracket is the total derivative with respect to t of $\varphi(M_i(t), t)$ the integral is indeed 0.

If by a limiting process the measure μ converges vaguely to $\rho(x,t)\, dx\, dt$ and the measure π converges vaguely to $p(x,t)\, dx\, dt$, then one obtains the conservation of mass $\frac{\partial \rho}{\partial t} + div(p) = 0$, and p represents the density of momentum, and the macroscopic velocity is defined by $u = \frac{p}{\rho}$.

Notice that the physical quantities, which are additive, are ρ and p, and not u.

Notice that a particle can leave the domain Ω without any difficulty in the preceding proof, as it stops being taken into account when it goes out of the support of φ and $\varphi(M_i(t), t) = 0$ before the particle exits. However, particles can also enter Ω without any problem, and the conservation of mass is only expressed inside Ω: RADON measures or distributions in Ω do not see the boundary $\partial\Omega$, and in order to treat boundary conditions one will have to use various SOBOLEV spaces and check what is the meaning of boundary conditions.

As mentioned, there is no problem having different particles go through the same point with different velocities, and therefore we have not been following an Eulerian point of view, but we have discovered that the velocity u is actually an average, and in cases where the velocity has oscillations, it will be important to understand which are the physical quantities and what equations they satisfy.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

6. Monday January 25.

Let us now consider the equations describing the conservation of momentum and the conservation of angular momentum.

EULER is credited for writing the equations for an ideal fluid (non viscous), which are

$$\frac{\partial \rho}{\partial t} + \sum_j \frac{\partial(\rho\, u_j)}{\partial x_j} = 0$$

$$\frac{\partial(\rho\, u_k)}{\partial t} + \sum_j \frac{\partial(\rho\, u_k\, u_j)}{\partial x_j} + \frac{\partial p}{\partial x_k} = 0 \text{ for all } k,$$

where p is the pressure.

The equation for the motion of a viscous fluid are attributed to NAVIER and to STOKES, but STOKES only considered the linearized problem, and so one uses the term STOKES equation when inertial terms are neglected but one uses the term NAVIER-STOKES equation when they are taken into account, although NAVIER had discovered it alone. It is unfortunate that so many results are not attributed correctly: the shock conditions expressing the conservation of mass and momentum in gas dynamics, now known after RANKINE and HUGONIOT, were actually first derived in 1848 by STOKES, and then rediscovered in 1860 by RIEMANN for an isentropic gas; STOKES is therefore credited for a discovery of NAVIER but forgotten for some of his discoveries; it could be by his own fault, as when he edited his complete works around 1870 he did not reproduce his derivation of the jump conditions, and he apologized for his mistake, because he had been (wrongly) convinced by Lord KELVIN and Lord RAYLEIGH that his discontinuous solutions were not physical, as they did not conserve energy. It is a quite amazing fact that such great scientists as STOKES, KELVIN and RAYLEIGH did not understand as late as 1870 that heat was a form of energy and that the missing energy had been transformed into heat (CARNOT and WATT did not need partial differential equations to understand that).

The form of the STOKES equation is very similar to that of linearized Elasticity, which CAUCHY had derived, and that involves something more general than pressure, as he had to introduce stress (what we call now the CAUCHY stress tensor, which is symmetric, and appears in the Eulerian point of view, while in the Lagrangian point of view the PIOLA-KIRCHHOFF stress tensor appears, which is not usually symmetric).

Pressure might be considered an easy concept, but I do not think that ARCHIMEDES knew that the reason why a body receives an upward force from the water in which one tries to submerse it is that the body receives a stronger force from below than from above because the hydrostatic pressure is higher below. Even in the beginning of this Century, after people had giggled at the idea of making flying machines that would be heavier than air, it was thought that the reason a plane could fly was that it was sustained from the air below it, while it is more because it is sucked upwards from the air above it, as an important depression is created above the wing by the flow (if the profile of the wing is well designed). The difficulty, of course, was that static questions about pressure had been well understood for some time, while dynamic questions were quite new. For the static question, and some dynamic effects, it is clear from some of his drawings that DA VINCI had well understood what the pressure is, and that should not be so surprising if one remembers that he was first of all an hydraulic engineer. After TORRICELLI had invented the barometer, PASCAL was the first to study the laws governing hydrostatic pressure, and both were remembered when units were chosen, a Torr for a pressure of a millimeter of mercury, and a Pascal for the rather small pressure of 1 Newton per square meter.

One BERNOULLI had studied the movement of a vibrating string by considering the approximation of many smalls masses connected by small springs; he apparently only derived the modes of vibration and it was D'ALEMBERT who first wrote the 1-dimensional wave equation. HUYGHENS had some insight about the wave nature of Light, but it might have been LAPLACE or POISSON who first wrote down the 3-dimensional wave equation.

CAUCHY derived the linearized Elasticity equation using the same idea of masses with small springs, but he only found a one parameter family of isotropic materials, and it was LAMÉ who introduced the two parameter family that we use now for the constitutive equation (strain-stress law) $\sigma_{ij} = 2\mu\,\varepsilon_{ij} + \lambda\delta_{ij}\sum_k \varepsilon_{kk}$, where $\varepsilon_{ij} = \frac{1}{2}\big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\big)$. If one lets λ go to ∞ and μ go to 0, one finds that $\sum_k \varepsilon_{kk}$, which is $div(u)$, tends to 0, and that $\lambda\, div(u)$ tends to a limit, giving the law for an inviscid incompressible fluid $\sigma_{ij} = -p\,\delta_{ij}$ (but this "pressure" for an incompressible fluid is not so physical). It is worth noticing that in his Physics course, FEYNMAN qualified the EULER equation as the equation for dry water, and the NAVIER-STOKES equation as the equation for wet water.

CAUCHY may have understood the force exerted by a part of an elastic body onto its complement as the resultant of all these tiny forces transmitted through these microscopic springs, but if that description might be found convenient for a solid, it does not look so realistic for a liquid or a gas.

The first explanation of what creates the pressure in a gas might have appeared in the work on kinetic theory of BOLTZMANN and MAXWELL (whose name was actually CLERK when he was born, and became CLERK MAXWELL after his father had inherited from an uncle).

In kinetic theory, one considers a gas with so many particles inside that one can take a limit and describe a density $f(x, v, t)$ for particles near the point x, having their velocity near v around the time t (in order to simplify, I assume that all particles have the same mass). If these particles were not interacting and were feeling no exterior forces, the evolution of the density would be given by the free transport equation

$$\frac{\partial f}{\partial t} + \sum_j v_j \frac{\partial f}{\partial x_j} = 0,$$

the density of mass ρ and of momentum p, and the (macroscopic) velocity u being defined by

$$\begin{aligned} \rho(x,t) &= \int_{R^3} f(x,v,t)\,dv \\ p(x,t) &= \int_{R^3} v\,f(x,v,t)\,dv \\ &= \rho(x,t)u(x,t), \end{aligned}$$

so that if one integrates in v the free transport equation, one obtains the equation of conservation of mass

$$\frac{\partial \rho}{\partial t} + \sum_j \frac{\partial(\rho\,u_j)}{\partial x_j} = 0.$$

Actually, there are exterior forces, depending both on position and velocity. For electrically charged particles, one must take into account the LORENTZ force $e(E + v \times B)$ for a particle with charge e; in Oceanography one must take into account Gravity and the CORIOLIS force created by the rotation of the Earth, and the form is similar. If all particles have the same mass m and the same charge e, and we still assume that particles do not interact, the evolution of the density of charged particles would be given by the transport equation

$$\frac{\partial f}{\partial t} + \sum_j v_j \frac{\partial f}{\partial x_j} + \sum_j \frac{e}{m}\Big(E_j(x,t) + \sum_{kl} \varepsilon_{jkl} v_k B_l(x,t)\Big)\frac{\partial f}{\partial v_j} = 0,$$

and integration in v (assuming that f is 0 for large v for example) would give the same form of the conservation of mass because $\sum_k \varepsilon_{jkl}\delta_{jk} = 0$. I will describe another time the CORIOLIS force, and forget about these exterior forces now and concentrate on interior forces, due to "collisions" of particles.

BOLTZMANN equation, in the absence of exterior forces, has the form

$$\frac{\partial f}{\partial t} + \sum_j v_j \frac{\partial f}{\partial x_j} + Q(f,f) = 0,$$

where $Q(f,f)$ is a somewhat complicated term, but for our purpose, because collisions are supposed to conserve mass, momentum and kinetic energy, we will admit that it always satisfies

$$\int_{v\in R^3} Q(f,f)\,dv = 0$$
$$\int_{v\in R^3} v_j Q(f,f)\,dv = 0 \text{ for all } j$$
$$\int_{v\in R^3} |v|^2 |Q(f,f)\,dv = 0.$$

Integrating the BOLTZMANN equation in v gives then again the same equation for conservation of mass, and integrating after multiplicating by v_i will give us the form of the equationm of conservation of momentum, and integrating after multiplicating by $|v|^2$ will give us the form of the equation of conservation of energy. The form is independent of what Q is, as long as Q satisfies the above constraints, but we will have to add constitutive relations. Let us define the symmetric stress tensor σ by

$$\sigma_{ij}(x,t) = -\int_{v\in R^3} f(x,v,t)\Big(v_i - u_i(x,t)\Big)\Big(v_j - u_j(x,t)\Big)\,dv.$$

Then as $v_i v_j = u_i u_j + u_i(v_j - u_j) + u_j(v_i - u_i) + (v_i - u_i)(v_j - u_j)$, and $\int_v f(x,v,t)\big(v_i - u_i(x,t)\big)\,dv = 0$ by definition of u, one deduces that

$$\int_{v\in R^3} v_i v_j f(x,v,t)\,dv = \rho(x,t)u_i(x,t)u_j(x,t) - \sigma_{ij}(x,t),$$

and the equation of conservation of momentum becomes

$$\frac{\partial(\rho\,u_i)}{\partial t} + \sum_j \frac{\partial(\rho\,u_i\,u_j)}{\partial x_j} - \sum_j \frac{\partial\sigma_{ij}}{\partial x_j} = 0 \text{ for all } i.$$

In the case where $\sigma_{ij} = -p\,\delta_{ij}$, p is the pressure, which is nonnegative, as the definition of σ shows that it is a negative definite tensor, because f is a nonnegative function with positive total mass (assuming that all the mass does not move at the same velocity u). This is acceptable in a gas, but not in a solid where extension is possible, and I will rederive the same equation using the point of view of BERNOULLI and CAUCHY based on little springs.

The pressure has a simple explanation if we look at what happens on the boundary. If the normal to the boundary going inside the gas is ν, the usual law of reflection, called specular reflection, is that a particle arriving with velocity v with $(v.\nu) < 0$ is reflected with velocity w given by $w = -2\nu(v.\nu) + v$, so that $(w.\nu) = -(v.\nu) > 0$. Each particle bouncing on the wall receives then a momentum in the direction of ν, and the pressure exerted by the gas is precisely the effect that all the particles transmit to the boundary a momentum in the direction $-\nu$ when they collide the boundary. The specular reflection is not exactly true, because the boundary is also made of particles and if a particle from the gas has enough velocity it may enter slightly into the solid, interact with the particles in the solid, and get back in various direction, after a small delay.

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

7. Wednesday January 27.

Let us look now at the conservation of energy, by multiplying the BOLTZMANN equation by $\frac{|v|^2}{2}$ and then integrating in v; as before, the term $Q(f,f)$ will disappear in this process.

One defines the internal energy per unit of mass e by the formula

$$\rho(x,t)e(x,t) = \int_{R^3} \frac{|v-u(x,t)|^2}{2} f(x,v,t)\,dv,$$

and there is then an automatic relation

$$\rho\, e = -\frac{1}{2} trace(\sigma), \text{ i.e. } \frac{3p}{2} \text{ in the case where } \sigma_{ij} = -p\,\delta_{ij},$$

and this is an obvious defect of the BOLTZMANN equation, that it implies constitutive relations which are not exactly true for real gases. Actually BOLTZMANN equation should only be considered as a model for rarefied gases, in agreement with the way the equation was derived, by assuming that two nearby particles only see each other and none of the other particles in the gas.

One also defines the heat flux q by the formula

$$q_j(x,t) = \int_{R^3} \Big(v_j - u_j(x,t)\Big) \frac{|v-u(x,t)|^2}{2} f(x,v,t)\,dv \text{ for all } j.$$

Apart from the relation already noticed between $\rho\, e$ and σ, there is no other automatic relation between the thermodynamic quantities ρ, σ, e, q, i.e. quantities pertaining to the gas and which therefore do not change in a Galilean transformation (consisting in adding a constant velocity to u). As ρ is the moment of order 0 of f, the moments of order 1 are 0, the moments of order 2 give σ (and $\rho\, e$ is a particular combination of these moments), and a particular combination of moments of order 3 is q, one can show that the only relations between these moments are the nonnegative character of ρ and $-\sigma$.

One has to compute the term $\int_v \frac{|v|^2}{2} f\,dv$, and putting $v = u + \xi$, one finds that $\frac{|v|^2}{2} = \frac{|u|^2}{2} + (u.\xi) + \frac{|\xi|^2}{2}$, and therefore, as $\int_v \xi\, f\,dv = 0$, one finds $\int_v \frac{|v|^2}{2} f\,dv = \frac{\rho\, u^2}{2} + \rho\, e$. Then, for each j, one has to compute the term $\int_v \frac{|v|^2}{2} v_j f\,dv$, and one finds that $\frac{|v|^2}{2} v_j = \frac{|\xi|^2}{2}\xi_j + u_j \frac{|\xi|^2}{2} + \xi_j(\xi.u) + \xi_j \frac{|u|^2}{2} + u_j(\xi.u) + u_j \frac{|u|^2}{2}$, and therefore, $\int_v \frac{|v|^2}{2} v_j f\,dv = q_j + u_j \rho\, e + \sum_k \sigma_{jk} u_k + \rho\, u_j \frac{|u|^2}{2}$. The conservation of energy appears then as

$$\frac{\partial\Big(\frac{\rho\,|u|^2}{2} + \rho\, e\Big)}{\partial t} + \sum_j \frac{\partial\Big[\Big(\frac{\rho\,|u|^2}{2} + \rho\, e\Big)u_j + \sum_k (\sigma_{jk} u_k) + q_j\Big]}{\partial x_j} = 0.$$

One sees from the formula $\int_v \frac{|v|^2}{2} f\,dv = \frac{\rho\,|u|^2}{2} + \rho\, e$ that $\rho\, e$ is the part of the kinetic energy which is hidden at a microscopic level. In BOLTZMANN model all energy is kinetic, i.e. comes from translation effects and none of it comes from rotation effects (as it would if the particles in the gas were also rotating), and the internal energy is that part of the kinetic energy which cannot be explained by looking only at the macroscopic quantities like u. The First Principle of Thermodynamics asserts that Energy is conserved, but one should count all the various forms of energy (in nuclear reactions even mass must be considered a form of energy, with the celebrated EINSTEIN formula $e = m\, c^2$); for a gas made of molecules (i.e. all real gases apart from the rare gases), besides translation and rotation effects of a molecule considered as rigid, there are also vibration effects due to the internal degrees of freedom of the molecule.

BOLTZMANN had also noticed his famous H-theorem (I think that he may have chosen H as the capital letter for η, used for entropy); it follows from the relation

$$\int_v Q(f,f) \log f\,dv \geq 0,$$

which implies

$$\frac{\partial(\int_v f \log f\,dv)}{\partial t} + \sum_j \frac{\partial(\int_v v_j f \log f\,dv)}{\partial x_j} \leq 0.$$

It is a consequence of the symmetric form of the collision operator (and the nonnegativity of the kernel), and equality only occurs for local Maxwellian distributions, i.e.

$$f = \alpha\, exp(-\beta|v-u|^2) \text{ with } \alpha, \beta, u \text{ depending only upon } x, t.$$

One has $\beta = \frac{1}{k\,T}$, where k is the BOLTZMANN constant and T the absolute temperature, and then $\beta^{3/2}\rho = \alpha\pi^{3/2}$, so for locally Maxwellian distributions, one can check that e is proportional to T, that $\sigma = -p\,\delta_{ij}$ and $q = 0$, with p computed as shown before, etc.

For a real gas, there is an equation of state which relates the various thermodynamic quantities, not necessarily the one that comes out of the (formal) computation for BOLTZMANN equation.

The qualitative form of the collision operator is obtained as follows. Two particles with initial velocities v and w "collide" and give two particles of velocities v' and w' and, as the masses are equal, conservation of momentum and conservation of kinetic energy are equivalent to the relations

$$\begin{gathered} v + w = v' + w' \\ |v|^2 + |w|^2 = |v'|^2 + |w'|^2, \end{gathered}$$

which give $|v-w| = |v'-w'|$, and putting $v' = v+z$ and $w' = w-z$ give $(v-w.z) + |z|^2 = 0$, so that by putting $\alpha = \frac{z}{|z|}$ (if $z=0$ one takes for α any unit vector orthogonal to $v-w$), one can parametrize all the solutions by using w and a unit vector α:

$$v' = v + (w - v.\alpha)\alpha;\ w' = w - (w - v.\alpha)\alpha,$$

and if one defines θ by $|v-w|\cos\theta = |(v-w.\alpha)|$, then the deflection is 2θ or $\pi - 2\theta$, i.e. in the Galilean frame of the center of mass (moving at velocity $\frac{v+w}{2}$) the final velocity direction makes and angle 2θ or $\pi-2\theta$ with the initial direction. The kernel only depends upon $|v-w|$ (twice the velocity of approach in the frame of the center of mass), and θ, as in the center of mass there is a symmetry around the direction of the initial velocity. The term $Q(f,f)$ has therefore the form

$$Q(f,f) = \int_{w\in R^3}\int_{\alpha\in S^2} B(|v-w|,\theta)\Big(f(v)f(w) - f(v')f(w')\Big)\,dw\,d\alpha,$$

and the kernel B is nonnegative. Because $\theta = \frac{\pi}{2}$ corresponds to $v' = v$ and $w' = w$ (or $v' = w$ and $w' = v$, as particles are undiscernable), and particle collisions are avoided outside a small effective scattering cross section, $B(|v-w|,\theta)$ tends to $+\infty$ as θ tends to $\frac{\pi}{2}$. That makes BOLTZMANN equation quite difficult, and following GRAD one usually considers an angular cut-off, i.e. one truncates B near $\theta = \frac{\pi}{2}$.

If one notices that the kernel B does not change if one exchanges v and w, or if one exchanges the roles of (v,w) and (v',w') (which is like reversing time so the collision of v' and w' may produce v and w), but $f(v)f(w) - f(v')f(w')$ stays the same for the first transformation and changes sign for the second, then one deduces that

$$\begin{aligned} \int_{v\in R^3} Q(f,f)\log f\,dv = \frac{1}{4}\int_{v,w\in R^3}\int_{\alpha\in S^2} B(|v-w|,\theta)\Big(f(v)f(w) - f(v')f(w')\Big)\Big(\log f(v) + \log f(w) \\ - \log f(v') - \log f(w')\Big)\,dv\,dw\,d\alpha \geq 0, \end{aligned}$$

the last inequality coming from $\log f(v) + \log f(w) - \log f(v') - \log f(w') = \log f(v)f(w) - \log f(v')f(w')$, and the fact that the logarithm is increasing. Equilibrium corresponds to $\int_v Q(f,f)\log f\,dv = 0$, and this is equivalent to $f(v)f(w) - f(v')f(w') = 0$ for all collisions, or $\log f(v) + \log f(w) = \log f(v') + \log f(w') = 0$

for all collisions, and that is certainly true if $\log f(v) = a + (b.v) + c\frac{|v|^2}{2}$ for some a, b, c independent of v, giving the locally Maxwellian functions. That there are no other solutions requires a little care.

In order to find more relations between ρ, e, σ and q, one usually quotes a formal argument of HILBERT, or one of CHAPMAN & ENSKOG, which start by considering a term in $\frac{1}{\varepsilon}Q(f,f)$, relating ε to the mean free path between collisions. The formal argument of HILBERT consists in assuming that $u = u_0 + \varepsilon\, u_1 + \ldots$ and identifying the various terms, the term in $\frac{1}{\varepsilon}$ imposing that u_0 is a local Maxwellian, and the next terms giving EULER equation for an inviscid perfect gas. The argument of CHAPMAN & ENSKOG produces NAVIER-STOKES equation with a viscosity of order ε.

As I mentioned before, letting the mean free path between collisions tend to 0 is in contradiction with the assumption that one deals with a rarefied gas in order to compute the kernel. It does not seem reasonable to assume that BOLTZMANN equation is valid for dense gases and liquids, one reason being that if two many "particles" get nearby, then the only way to deal with them is to consider that they are waves, and not classical particles. Actually, as BOLTZMANN equation (formally) predicts a perfect gas behaviour, and real gases are not perfect gases, either BOLTZMANN equation is not satisfied by real gases, or the formal argument of HILBERT is not valid.

From a philosophical point of view, it is rather curious to observe the efforts made to derive EULER or NAVIER-STOKES equation out of BOLTZMANN equation, as if starting with BOLTZMANN equation was a flawless assumption. On the contrary, BOLTZMANN equation has already postulated some irreversibility, and this is seen by the fact that nonnegative initial data create a nonnegative solutions, a property that is lost after time reversal. Formally this is due to the form of the equation:

$$\frac{\partial f}{\partial t} + \sum_j v_j \frac{\partial f}{\partial x_j} + f\, A(f) = B(f),$$

with $B(f) \geq 0$ a.e. when $f \geq 0$ a.e.; if A and B were locally LIPSCHITZ continuous one could obtain the solution by the iterative process

$$\frac{\partial f^{(n+1)}}{\partial t} + \sum_j v_j \frac{\partial f^{(n+1)}}{\partial x_j} + f^{(n+1)}\, A(f^{(n)}) = B(f^{(n)}),$$

which gives $f^{(n+1)}$ nonnegative when the initial condition and $f^{(n)}$ are nonnegative.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

8. Friday January 29.

Let us look at the way BERNOULLI and D'ALEMBERT were led to discover the 1-dimensional wave equation, and later CAUCHY was led in the same way to discover the equation for linearized Elasticity.

The simplest case, for what concerns its analysis, is that of a 1-dimensional longitudinal wave. The motion of a violin string is different, as it is a transversal wave: the waves propagate along the string but the displacement is mostly perpendicular to the string. A 1-dimensional longitudinal wave corresponds to the experimental situation of a metallic bar which one hits on one end with a hammer. In linearized Elasticity, in two or three dimensions and in an isotropic material, P-waves (pressure waves) are longitudinal waves, while S-waves (shear waves) are transversal waves (they travel at different speeds).

Let us consider the motion of $N-1$ small masses connected with springs, with the purpose of letting N tend to ∞; let $x_0(t)=0$, and $x_N(t)=L$, corresponding to the fixed walls where the first and last masses are attached. Let $x_i(t)$, $i=1,\ldots,N-1$, be the positions at time t of the mass $\#i$, let m_i be its mass. Let us assume that the springs are at equilibrium if the masses are at their rest point, corresponding to ξ_i for the mass $\#i$ (one take $\xi_0=0$ and $\xi_N=L$), and let $\kappa_{i,i+1}$ be the constant of the spring connecting mass $\#i$ and mass $\#(i+1)$ (with $\#0$ and $\#N$ designating the walls), i.e. an increase in length of $\Delta>0$ creates a restoring force $\kappa\Delta$ (and similarly for a compression); of course, this is only realistic if the displacements are small.

The force acting on mass $\#i$ by the spring connecting it to mass $\#(i+1)$ is $\kappa_{i,i+1}(x_{i+1}-x_i-\xi_{i+1}+\xi_i)$, and it is therefore natural to put $x_i(t)=\xi_i+y_i(t)$, and the equation of motion (NEWTON's law) for mass $\#i$ is then

$$m_i\frac{d^2y_i}{dt^2}=\kappa_{i,i+1}(y_{i+1}-y_i)-\kappa_{i-1,i}(y_i-y_{i-1})\text{ for }i=1,\ldots,N-1,$$

and the initial position and velocity of each of the $N-1$ masses must also be given, and as it is a linear differential system there exists a unique solution (global in time). However, we need precise estimates if we want to understand what happens when N tends to ∞.

Before doing that, it is useful to repeat that the reason that one can do the analysis is that one has chosen a linearized problem without saying it expressly: if a spring has size 1 at rest and is elongated of an amount Δ, the restoring force may be of the order of $\kappa\Delta$ if $|\Delta|$ is small, but it makes no sense having Δ go to -1, where the spring is compressed to zero length, or Δ tend to ∞ as no known material can sustain such a deformation without going through permanent plastic deformation before breaking. Of course, these springs are only an idealized classical version of what happens at a microscopic level: electric forces may be attracting of repulsing and both occur in an ionic crystal like salt (NaCl), but forces that bind a metallic crystal are all similar and it is more than the nearby neighbours which play a role in the stability of the crystalline arrangement (at least in liquids, one sometimes invoke LENNARD-JONES potentials, which have a long range attraction potential in $1/r^6$ and a short range repulsion potential in $1/r^{12}$). Then crystals are not very good at Elasticity and cannot support much strain and they change their microstructure to polycrystals, so the idealized description of Elasticity with little springs could have seemed reasonable to CAUCHY (and I am not even sure if that is the way he thought), but is known to contradict our actual knowledge.

However, one may look at this description in another way, and consider it a Numerical Analysis point of view. Indeed, if one uses finite difference schemes or finite elements (where finite is in opposition to infinitesimal and not to infinite), it is quite natural to replace the wave equation by the system that I have written, and interpret it as moving masses connected by springs, and replace the equation of linearized Elasticity by a system very similar in nature. The important difference is that in the Numerical Analysis point of view, mathematicians start from the partial differential equations and want to show that the finite dimensional approximation chosen will indeed approach the solution as the mesh size tends to 0 (while engineers might not even write down the partial differential equations and may only play with the finite dimensional description), and it is not so good for detecting the effects of nonlinearity. We will first follow this point of view, neglecting nonlinearities by pretending that they are small, and later we will try to take them into account.

The system written has an invariant, which is the total energy: multiplying equation $\#i$ by $\frac{dy_i}{dt}$ and summing in i, one obtains

$$\frac{d}{dt}\Big(\sum_{j=1}^{N-1}\frac{m_j}{2}\Big|\frac{dy_j}{dt}\Big|^2+\sum_{j=0}^{N-1}\frac{\kappa_{j,j+1}}{2}|y_{j+1}-y_j|^2\Big)=0,$$

the first part being the kinetic energy, and the second part being the potential energy, i.e. the energy stored inside the springs (and there is one more spring than masses). There is an obvious Hamiltonian framework behind our equation, but one should be aware of the fact that for partial differential equations which are not linear, a Hamiltonian framework is not always so useful, the main reason being that things which are conserved like energy may suddenly start converting to a new form like heat, which may not be described by the same equation, and suddenly the "conserved quantity" starts to change!

Then one wants to let N tend to ∞, and all questions of scaling should be done with care, as there might be different regimes to consider, but here the matter is straightforward. One way to guess the right scaling is to consider that the values y_j are extended by interpolation, filling the intervals in the space variable ξ and time t (that is a Lagrangian point of view), defining a function u, and that the kinetic part should look like $\frac{1}{2}\int_0^L \rho(\xi)\big|\frac{\partial u}{\partial t}\big|^2\,d\xi$ and the potential part like $\frac{1}{2}\int_0^L \kappa(\xi)\big|\frac{\partial u}{\partial \xi}\big|^2\,d\xi$. For example, taking $\xi_j=\frac{j\,L}{N}$ and $m_j=\frac{M}{N}$ so that M is the total mass of the springs, and $\kappa_{j.j+1}=N\kappa$, corresponds to a uniform density of mass $\rho=\frac{M}{L}$ and constant κ, and the equation becomes

$$M\frac{\partial^2 u}{\partial t^2}-\kappa\,L^2\frac{\partial^2 u}{\partial \xi^2}=0,$$

corresponding to a propagation speed

$$c=\sqrt{\frac{\kappa}{M}}L,$$

and its solutions are of the form $f(x-c\,t)+g(x+c\,t)$, as noticed by D'ALEMBERT. One should add the boundary conditions $u(0,t)=u(L,t)=0$, and the initial conditions

$$u(\xi,0)=v(\xi);\ \ \frac{\partial u}{\partial t}(\xi,0)=w(\xi)\text{ a.e. in }(0,L).$$

This can be proved using standard results of Functional Analysis (from any bounded sequence in L^2, one can extract a weakly converging subsequence) and a little use of distributions (for pushing the derivatives to the test functions), but one must be careful that the initial condition should be approached in the right way, i.e. $v\in H^1_0(0,L)$ and $w\in L^2(0,L)$ (I will use this type of method extensively later on, and I will then explain the details of the argument).

As the unit of $\kappa_{j,j+1}$ is mass/time^2, and the mass scales naturally in $m_j=M/N$, the scaling $\kappa_{j,j+1}=N\kappa$ corresponds to a characteristic time in $1/N$, which is quite natural for a characteristic length L/N and a finite propagation speed, but the argument is circular because the discrete system does not have finite propagation speed (a change of position of the first mass is immediately felt at the last one), and it is only the limiting equation that has the finite propagation speed property. However, if the total energy of the initial data is kept fixed and if one takes $\kappa_{j,j+1}=h(N)$ with $h(N)/N$ tending to ∞, then the solution tends to 0 and all the energy goes into vibration, while if $h(N)/N$ tends to 0 there is only kinetic energy at the limit and no interaction between particles; therefore there is only one good scaling!

It seems that BERNOULLI only considered the solutions of the form $y_j(t)=e^{i\,\omega\,t}z_j$, which he found to be $z_j=\gamma\sin\big(\frac{j}{N}m\pi\big)$, corresponding to $\omega^2=\frac{4N^2\kappa}{M}\sin^2\big(\frac{m\pi}{2N}\big)$, which tends to $\frac{K\,m^2\pi^2}{M}$ as N tends to ∞, and that is not as precise as deriving the wave equation. Physicists often find information for special solutions oscillating at a unique frequency, and the result may show that no partial differential equation of a given type may create the same kind of relation, but even if one has to write down a pseudo-differential equation, it is better to understand what all solutions do; actually in a nonlinear setting one cannot expect to reconstruct the solution easily from the knowledge of special solutions, and even in linear situations it does not help much for understanding what the boundary conditions are (as every function in $L^2(0,1)$ can be written as an infinite sum of functions vanishing at 0, one must be careful).

If one considers a 2-dimensional or 3-dimensional array of masses connected by springs, or even the transversal vibrations of a string, the first thing to realize if that without linearization the problem becomes terribly difficult. With linearization, the idea if that if a spring connects points A and B and that these points move of δA and δB, which are small compared to the length of AB, then the new length is $|B-A+\delta B-\delta A| = \sqrt{|B-A|^2+2(B-A,\delta B-\delta A)+|\delta B-\delta A|^2} = |B-A|+(B-A,\delta B-\delta A)/|B-A|+o(|\delta B-\delta A|)$, and therefore only the displacement perpendicular to the initial position of the spring is taken into account. In a two dimensional setting, denoting by x and y the space variables, by u and v the displacement, one sees that a spring parallel to the x axis corresponds to a potential energy involving $|u_x|^2$ (where subscript denotes differentiation), a spring parallel to the y axis corresponds to a potential energy involving $|v_y|^2$, a spring along the first diagonal corresponds to a potential energy involving $|u_x+u_y+v_x+v_y|^2$, and a spring along the second diagonal corresponds to a potential energy involving $|u_x-u_y-v_x+v_y|^2$. One understands then that the notation

$$\varepsilon_{ij} = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\Big),$$

helps in writing the limiting equations as

$$\rho(\xi)\frac{\partial^2 u_i}{\partial t^2} - \sum_j \frac{\partial \sigma_{ij}}{\partial x_j} = 0 \text{ for all i},$$

where the stress σ has the form

$$\sigma_{ij} = \sum_{k,l} C_{ijkl}(\xi)\varepsilon_{kl}.$$

Linearization has the defect of mixing up the Eulerian and the Lagrangian point of views, and the CAUCHY stress, which is symmetric, should appear in the Eulerian point of view while the PIOLA-KIRCHHOFF stress, which is not symmetric, should appear in the Lagrangian point of view. In the isotropic case, CAUCHY had found the relation $\sigma_{ij} = 2\mu\varepsilon_{ij}+\delta_{ij}\sum_k \varepsilon_{kk}$, but with a special relation between μ (the shear modulus) and λ (the LAMÉ parameter), as he had $\lambda=\mu$, and it was LAMÉ who then pointed out that there was a two dimensional family of isotropic materials. Because the tensors ε and σ are symmetric, there is no restriction in assuming that

$$C_{ijkl} = C_{jikl} \text{ and } C_{ijkl} = C_{ijlk} \text{ for all } i,j,k,l,$$

but there is another symmetry relation, for hyperelastic materials, i.e. those materials which have a stored energy function (and this symmetry is probably a necessary condition for the evolution problem to be well posed with the finite propagation speed property, assuming that some kind of ellipticity condition is satisfied),

$$C_{ijkl} = C_{klij} \text{ for all } i,j,k,l.$$

Under this last condition, the conservation of energy becomes

$$\int\Big(\frac{\rho(\xi)}{2}\sum_i\Big|\frac{\partial u_i}{\partial t}\Big|^2 + \frac{1}{2}\sum_{i,j,k,l} C_{ijkl}(\xi)\varepsilon_{ij}\varepsilon_{kl}\Big)\,dx = constant.$$

The preceding discussion was to show the form of the equation, and under an hypothesis of "very strong ellipticity" one can show existence and uniqueness for the evolution problem (and the finite propagation speed property), and the convergence of some natural approximation processes, like the one involving little masses and springs.

However, the CAUCHY stress should be discussed in an Eulerian framework, and the argument of CAUCHY that there must exists a stress tensor used the equilibrium of a small tetrahedron. He assumed that for a domain ω (with LIPSCHITZ boundary!), the force acting on a small set of the boundary of ω by the exterior of ω is a force proportional to the surface of the element and depending upon the position and the normal to $\partial\omega$ (the exterior of ω receiving an opposite force, so that conservation of momentum is satisfied); writing then the equilibrium of a tetrahedron small enough so that the dependence is only in the normal, the following argument was used to deduce the fact that the dependence with respect to the normal must be linear. For $a_1, a_2, a_3 > 0$ and small, the faces of the tetrahedron are the planes $x_i = 0$ and the face T of equation $\frac{x_1}{a_1} + \frac{x_2}{a_2} + \frac{x_3}{a_3} = 1$ with $x_j \geq 0$. Let F_i be the force by unit area on the face $x_i = 0$ and let G be the force per unit area on the face T, then the equilibrium of the tetrahedron is $a_2 a_3 F_1 + a_3 a_1 F_2 + a_1 a_2 F_3 + S\,G = 0$, where S is the area of the triangle T, but as the normal ν to T has the form $\nu_j = \frac{\lambda}{a_j}$ for some $\lambda > 0$, and $S = \frac{a_1 a_2}{\nu_3} = \frac{a_1 a_2 a_3}{\lambda}$, one finds that $G = -\nu_1 F_1 - \nu_2 F_2 - \nu_3 F_3$, and therefore G is linear with respect to ν.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

9. Monday February 1.

Linearization may often seem a reasonable step when some quantities are believed to be small, and a function that one may want to neglect may indeed be small, but the danger comes from the fact that its derivative might not be small. For what concerns hyperbolic equations, which are more or less the partial differential equations for which information travels at finite speed, the difference between the linear and the nonlinear case (actually, the quasilinear case) is quite important.

In the case of an elastic string, taking into account large deformations leads to an equation of the form

$$\frac{\partial^2 w}{\partial t^2} - \frac{\partial}{\partial x}\Big(f\Big(\frac{\partial w}{\partial x}\Big)\Big) = 0,$$

where w denotes the vertical displacement, $\frac{\partial w}{\partial t}$ the velocity, $\frac{\partial w}{\partial x}$ the strain, and $f\big(\frac{\partial w}{\partial x}\big)$ the stress, and the function f is no longer affine, but it satisfies $f' > 0$, and $\sqrt{f'}$ appears to be the local speed of propagation of perturbations.

The first to study such an equation was POISSON, around 1807, but he was concerned with gas dynamics in a simplified form, i.e. the system

$$\begin{aligned}&\frac{\partial \rho}{\partial t} + \frac{\partial(\rho\, u)}{\partial x} = 0\\ &\frac{\partial(\rho\, u)}{\partial t} + \frac{\partial(\rho\, u^2 + p)}{\partial x} = 0,\end{aligned}$$

with p being a nonlinear function of ρ. One of the reasons why POISSON was interested in compressible gases was to compute the exit velocity of a shell out of the barrel of a gun. NEWTON had apparently computed the velocity of sound in air, but his calculation had given a value almost 100 m/s short of the measured velocity (which is a little above 300 m/s under usual conditions). He had certainly not written the wave equation, but he had indeed used what he knew about compressibility of air, i.e. he had used p as a linear function of ρ, according to the law of perfect gases $P\,V = constant$ (as the relation $P\,V = R\,T$ appeared much later). POISSON was using a relation $p = c\,\rho^\gamma$, which LAPLACE may have suggested, and the thermodynamic interpretation came much later: as the wave are fast phenomena, the mechanical energy has no time to be transformed into heat, and the process is therefore adiabatic ($\delta Q = 0$), or equivalently isentropic (as $\delta Q = T\,dS$). Thermodynamics tells us that γ is the ratio $\frac{c_p}{c_v}$, where c_p is the heat capacity per unit mass at fixed pressure, and c_v the heat capacity per unit mass at fixed volume; it is about 5/3 for air. POISSON's solution was not analytical but had an implicit form, and in 1848 CHALLIS found that his formula could not be true for all time, which prompted STOKES to explain that profiles were becoming steeper and steeper, until one had to introduce a discontinuity, for which he computed the velocity, by expressing the conservation of mass and the conservation of momentum.

The basic ideas are more easily explained on the inviscid BURGERS equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$

which will have to be written as

$$\frac{\partial u}{\partial t} + \frac{\partial(\frac{u^2}{2})}{\partial x} = 0,$$

as some solutions will not be smooth (but will not be general distributions, for which one cannot define u^2). In order to be consistent, u must have the dimension of a velocity. In 1948, BURGER had proposed the equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \varepsilon\frac{\partial^2 u}{\partial x^2} = 0,$$

as a 1-dimensional model of turbulence, and apart from pointing out that turbulence was something very different, Eberhard HOPF had been able to study the limiting case $\varepsilon \to 0$ by using a nonlinear transformation which changes the equation into the linear heat equation (that transformation is now known as the HOPF-COLE transformation, as Julian COLE had also found it independently). The work of Peter LAX and of Olga OLEINIK opened then the way for more general cases.

If $a(x,t)$ is LIPSCHITZ continuous in x, the solution of $\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0$ and $u(x,0) = v(x)$ is obtained by the method of characteristic curves, going back to CAUCHY: along the solution of $\frac{dx(t)}{dt} = a(x(t),t)$ and $x(0) = \xi$, the solution u satisfies $\frac{d}{dt}\big(u(x(t),t)\big) = 0$ and so $u(x(t),t) = v(\xi)$. Assuming that the solution u of BURGERS equation is LIPSCHITZ continuous in x for $0 \le t < T$, the characteristic curve is $\frac{dx(t)}{dt} = u(x(t),t)$, and as $u(x(t),t) = v(\xi)$, one finds that the characteristic curve is a line on which u is constant:

$$\begin{aligned} x(t) &= \xi + t\,v(\xi) \\ u(x(t),t) &= v(\xi). \end{aligned}$$

In the spirit of the implicit equation found by POISSON, one could write that $\xi = x(t) - t\,u(x(t),t)$, and therefore for a given t, the function $x \mapsto u(x,t)$ solves the implicit equation

$$v\Big(x - t\,u(x,t)\Big) = u(x,t) \text{ for all } x.$$

I think that CHALLIS's argument was to use $v(x) = \sin x$ (I do not know if he had to use POISSON's formula in a question of Astronomy: CHALLIS was the astronomer in Cambridge and is mentioned in the Encyclopaedia Britannica for quite a negative reason, as he had not found the new planet for which ADAMS had computed the position, and it was then LE VERRIER who got all the fame of the discovery, and the right to call it Neptune). If one believes that u is continuous, the zeros of u must stay at $k\pi$, but looking for the points where $u = 1$ creates a problem, as it gives $v(x-t) = 1$, and therefore $x = t + \frac{\pi}{2} + 2k\pi$, and for $t = \frac{\pi}{2}$ it gives a point where u is known to be 0.

The parametrization $u(x,t) = v(\xi)$ on the line $x = \xi + t\,v(\xi)$ shows more easily the problem: if $\xi < \eta$ but $v(\xi) > v(\eta)$, the lines coming out of ξ and η intersect and there is a conflict between two different values of u at the intersection. More precisely if v is of class C^1 and $v' \ge 0$, the mapping $\xi \mapsto \xi + t\,v(\xi)$ is a global diffeomorphism from R to R, but in the opposite case, if $-\alpha = \inf_\xi v'(\xi)$ with $\alpha > 0$, and $T_c = \frac{1}{\alpha}$ then the solution is of class C^1 for $0 < t < T_c$, but for some t slightly larger than T_c an intersection of two characteristic lines occur.

The computation of STOKES for discontinuous solutions, consists simply in writing that one has a solution of the equation in the sense of distributions, and one calls them weak solutions, but for doing so one must avoid multiplying the derivative of u by a function of u, and use an equation in conservation form, like

$$\frac{\partial u}{\partial t} + \frac{\partial v}{\partial x} + w = 0,$$

where u, v, w are locally integrable functions in an open set Ω of the plane, and v is a function of u. Let us assume that u and v are in $W^{1,1}$ on both sides of a curve $x = Z(t)$ of class C^1 (or just LIPSCHITZ continuous), and so both u and v have limits on each side of the curve, and the equation is satisfied in each of the subdomains $\Omega_- = \{(x,t) : x < Z(t)\}$ and $\Omega_+ = \{(x,t) : x > Z(t)\}$. The speed of propagation of the curve is $s = Z'(t)$. Writing the equation in the sense of distributions means that for every $\varphi \in C_c^\infty(\Omega)$ one has $\int_\Omega \big(-u\frac{\partial \varphi}{\partial t} - v\frac{\partial \varphi}{\partial x} + w\varphi\big)\,dxdt = 0$, and decomposing the integral in two parts, one on Ω_- and one on Ω_+, one can integrate by parts, and transform them in integrals on the curve: $\int_{\Omega_-} \cdots = \int_{x=Z(t)} (u\nu_t^- + v\nu_x^-)\varphi\,d\sigma$ and $\int_{\Omega_+} \cdots = \int_{x=Z(t)} (u\nu_t^+ + v\nu_x^+)\varphi\,d\sigma$, where ν_t and ν_x are the components of the exterior normal ν along x and t, and of course $\nu^- = -\nu^+$. This leads to the jump condition

$$v_+ - v_- = s(u_+ - u_-),$$

which are usually called RANKINE-HUGONIOT conditions, although my preference is to call them after STOKES and RIEMANN.

That is not the end of the story however, because there are too many weak solutions: for example let $a > 0$, and $x_0 \in R$, then the following function u is a weak solution of BURGERS equation with initial data

0: $u(x,t)=0$ if $x<x_0-a\,t$, $u(x,t)=-2a$ if $x_0-a\,t<x<x_0$, $u(x,t)=2a$ if $x_0<x<x_0+a\,t$, $u(x,t)=0$ if $x_0+a\,t<x$. The problem is that some of the discontinuities in a weak solution may not be physical.

It is easy to check that for $t\neq 0$ the function $\frac{x}{t}$ is a particular solution of the BURGERS equation, and one deduces then that for $\varepsilon>0$ and the initial data $v(x)=0$ for $x<0$, $v(x)=\frac{x}{\varepsilon}$ for $0<x<\varepsilon$, $v(x)=1$ for $x>\varepsilon$, the solution (also given by the method of characteristic curves) is $u(x,t)=0$ for $x<0$, $u(x,t)=\frac{x}{\varepsilon+t}$ for $0<x<\varepsilon+t$ and $u(x,t)=1$ for $x>\varepsilon+t$. If one lets ε tend to 0, one sees that the limit of the initial data is the Heaviside function but the limit of the solutions is not a solution with a shock (discontinuity) but a rarefaction wave $u(x,t)=0$ for $x<0$, $u(x,t)=\frac{x}{t}$ for $0<x<t$ and $u(x,t)=1$ for $x>t$.

If then one decides to prefer the (unique) locally LIPSCHITZ solution when it exists, and argue by continuity, one is led to reject all discontinuities for which $u_-<u_+$. There is another way to explain this selection for BURGERS equation: one needs to have $u_->s=\frac{u_-+u_+}{2}>u_+$ because what creates the shock is the fact that the information on the left side travels faster than the shock and is catching upon it, while the information on the right side travels slower than the shock and is caught up by the shock, and as one cannot have the analog of a breaking wave on a beach because one looks for a single valued function, the fast side must help the slow side so that both can move together at an intermediate speed. In the case of an equation with $v=f(u)$ (and $w=0$ for example), LAX's criteria is the analog of this remark: $f'(u_-)\geq\frac{f(u_+)-f(u_-)}{u_+-u_-}\geq f'(u_+)$. However, if f is neither convex nor concave, a more complete analysis shows that one must impose OLEINIK's condition: if $u_-<u_+$, the chord joining $\big(u_-,f(u_-)\big)$ to $\big(u_+,f(u_+)\big)$ should be above the graph of f, while if $u_->u_+$, the chord joining $\big(u_-,f(u_-)\big)$ to $\big(u_+,f(u_+)\big)$ should be below the graph of f.

A more mathematical way to introduce these conditions was found by HOPF, and then extended to the case of systems by LAX, who coined the term "entropy" to designate some functions which appear in supplementary conservation laws in the case of smooth functions; the choice of name is not so good as it is not directly related to the thermodynamic entropy (geometers use the term Casimir for the same idea, I believe). HOPF's idea was that if one multiplies the equation

$$\frac{\partial u_\varepsilon}{\partial t}+\frac{\partial f(u_\varepsilon)}{\partial x}-\varepsilon\frac{\partial^2 u_\varepsilon}{\partial x^2}=0$$

by $\varphi'(u_\varepsilon)$, and if one chooses ψ satisfying $\psi'=f'\varphi'$, one obtains

$$\frac{\partial\varphi(u_\varepsilon)}{\partial t}+\frac{\partial\psi(u_\varepsilon)}{\partial x}-\varepsilon\frac{\partial^2\varphi(u_\varepsilon)}{\partial x^2}+\varphi''(u_\varepsilon)\Big(\frac{\partial u_\varepsilon}{\partial x}\Big)^2=0,$$

and therefore if one knows that u_ε converges almost everywhere to u, then of course u is a weak solution of the equation with $\varepsilon=0$, but it also satisfies the supplementary conditions (called entropy conditions by LAX)

$$\frac{\partial\varphi(u)}{\partial t}+\frac{\partial\psi(u)}{\partial x}\leq 0\text{ for all convex “entropy” }\varphi,$$

and when one tests this condition for smooth convex functions which approximate the following special functions φ_k (which KRUSHKOV used later for the multidimensional scalar case), $\varphi_k(v)=(v-k)_+$, corresponding to $\psi_k(v)=0$ for $v<k$ and $\psi_k(v)=f(v)-f(k)$ for $v>k$, one discovers OLEINIK's condition. The formalism of HOPF/LAX has the advantage of expressing the additional conditions without imposing that the solution must be piecewise smooth and have limits on both sides of the discontinuity lines; for scalar equations, the solution is indeed unique if one imposes these conditions, but the situation for systems is not so clear.

If one writes the equation for a nonlinear string as a system, u being the strain, v the velocity and $\sigma=f(u)$ the stress, then the system is

$$\begin{aligned}&\frac{\partial u}{\partial t}-\frac{\partial v}{\partial x}=0\\&\frac{\partial v}{\partial t}-\frac{\partial f(u)}{\partial x}=0,\end{aligned}$$

and in order to deduce that

$$\frac{\partial\varphi(u,v)}{\partial t}+\frac{\partial\psi(u,v)}{\partial x}=0$$

for all smooth solutions, one requires

$$\begin{aligned}\frac{\partial\psi(u,v)}{\partial v} &= -\frac{\partial\varphi(u,v)}{\partial u}\\ \frac{\partial\psi(u,v)}{\partial u} &= -f'(u)\frac{\partial\varphi(u,v)}{\partial v},\end{aligned}$$

and therefore φ must satisfy the compatibility condition

$$\frac{\partial^2\varphi(u,v)}{\partial u^2} = f'(u)\frac{\partial^2\varphi(u,v)}{\partial v^2},$$

Air is quite compressible, and the speed of sound at atmospheric pressure is a little above the velocity of commercial planes. However, the velocity depends upon temperature and pressure, and the shape of the wings of commercial planes has been chosen so that the flow of air creates a depression above the wing and in this lower pressure the speed of sound is less than the speed of the plane; transsonic flows are therefore important for some practical applications.

For water, the velocities involved are always a small fraction of the speed of sound, which is about 1.5 km/s, but one must remember that the approximation of incompressibility has the unrealistic consequence that a perturbation at one point can be felt immediately very far from it, which is unphysical. The "pressure" appearing in incompressible NAVIER-STOKES equation, for example, should not be mistaken for the real pressure.

One should be aware then of the limitations of most of the approximations used.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

10. Wednesday February 3.

I start introducing now the basic functional spaces that will be used in the proofs of existence of solutions to some model partial differential equations dealing with questions about fluids.

One of the basic functional space that we will use is $H_0^1(\Omega)$. Ω will usually be an open subset of R^N, and of course one should think of $N=3$, but there are sometimes problems which are naturally posed in one or two dimensions only, and mathematicians like to be general and they study problems in R^N without any constraint imposed on N and they want to discover if the dimension matters. Mathematicians also study partial differential equations on manifolds, and sometimes it corresponds to a real question (after all there are some global questions about oceans and it is important to avoid being stuck in technical questions related to the parametrization of the surface of the Earth), but some are not so realistic (periodicity hypotheses for example are a good way to avoid being bothered by what happens on the boundary, but should be considered only as a preliminary step).

Most of the open sets that one encounters are bounded (and anyway the radius of the Earth is only about 6,370 km), but mathematicians do analyze questions in unbounded sets, because it is sometimes useful to consider explicit solutions which can be computed more easily in the whole space; these solutions in the entire space may be the limit of the solutions obtained when one lets the boundary go to infinity, and they may therefore be good approximations when one is far from the boundary. FOURIER transform is a very important technical tool for studying the properties of some functional spaces, but it applies only to functions defined on the entire space. As we will see, some properties of functions inside Ω are similar to what happens for the whole space, while some properties of functions near $\partial\Omega$ are similar to what happens near the boundary of a half space.

The smoothness of the boundary is often not important, for example for the velocity field because the viscosity imposes that the velocity must be 0 on the boundary, but there are important questions for which one should look in a more precise way at what happens near the boundary, for the pressure for example: there may exist thin boundary layers near places where the boundary is not very smooth, and the problem will then be to understand what are good effective boundary conditions to use outside the boundary layer.

The SOBOLEV space $H^1(\Omega)$ is the space of (equivalence classes of measurable) functions in $L^2(\Omega)$, whose all first derivatives are in $L^2(\Omega)$, with a norm

$$||u||_{H^1(\Omega)} = \Big(\int_\Omega \Big(\frac{|u|^2}{L^2} + |grad(u)|^2\Big)\,dx\Big)^{1/2},$$

where L is a characteristic length (which mathematicians usually take equal to 1!).

$H_0^1(\Omega)$ is then defined as the closure of $C_c^\infty(\Omega)$ in $H^1(\Omega)$. In the case where Ω has a compact boundary which is locally defined by a LIPSCHITZ equation, Ω being only on one side of the boundary, one can show that $C^\infty(\overline{\Omega})$ is dense in $H^1(\Omega)$, and that the restriction to the boundary extends into a linear continuous mapping from $H^1(\Omega)$ into $L^2(\partial\Omega)$ called the trace; then $H_0^1(\Omega)$ is exactly the subspace of $H^1(\Omega)$ of functions which have trace 0 on the boundary.

An important property is the POINCARÉ inequality (not unrelated to Oceanography, as I was told that POINCARÉ had introduced that inequality in his studies of the tides), which is

$$\int_\Omega |u|^2\,dx \le C\int_\Omega |grad(u)|^2\,dx.$$

Of course, the constant C has the dimension $length^2$, and one cannot expect to have POINCARÉ inequality in domains like R^N where there is no characteristic length (one does have $H_0^1(R^N) = H^1(R^N)$). The way to express in a mathematical way this argument about units is to replace $u(x)$ by $v(\lambda\,x)$: if POINCARÉ inequality was true in R^N then one would have

$$\int_{R^N} |v(\lambda\,x)|^2\,dx \le C\int_{R^N} \Big|\lambda\,grad\Big(v(\lambda\,x)\Big)\Big|^2\,dx,$$

but using the change of variable $y = \lambda x$ gives

$$|\lambda|^{-N} \int_{R^N} |v|^2\, dy \le C|\lambda|^{2-N} \int_{R^N} |grad(v)|^2\, dy,$$

and taking advantage of the fact that different powers of λ appear on both sides (which is what is meant by the statement that $\int_{R^N} |v|^2\, dy$ and $\int_{R^N} |grad(v)|^2\, dy$ are not measured in the same unit), one gets a contradiction by letting λ tend to 0.

The preceding proof actually shows that if Ω is unbounded and contains balls of arbitrarily large size, then POINCARÉ inequality does not hold: indeed, if $B(x_m, r_m) \subset \Omega$ and $r_m \to \infty$, one uses $u(x) = \varphi\big(\frac{x-x_m}{r_m}\big)$, with $\varphi \in C_c^\infty\big(B(0,1)\big)$ and if POINCARÉ inequality was true one would deduce that $\int_{B(0,1)} |\varphi|^2\, dx = 0$. Therefore such an open set must be considered to have its characteristic length infinite, but the maximum size of balls contained in Ω is not always the right measure for a characteristic length: if Ω is obtained from R^N by removing all the points with integer coordinates, then POINCARÉ inequality does not hold if $N \ge 2$, because one has $H_0^1(\Omega) = H^1(R^N)$ (functions in $H^1(R^N)$ are not necessarily continuous for $N \ge 2$ and therefore points are negligible).

Lemma: i) If Ω lies between two parallel hyperplanes separated by a distance D, then POINCARÉ inequality holds for $H_0^1(\Omega)$ with the constant D^2/π^2 (which is optimal if Ω occupies all the domain between the two hyperplanes).

ii) If $meas(\Omega) < \infty$, then POINCARÉ inequality holds for $H_0^1(\Omega)$ with a constant $C_* meas(\Omega)^{2/N}$.

Proof of i): As both norm being compared are invariant in an orthogonal transformation, one may suppose that one hyperplane has equation $x_N = 0$ and the other $x_N = D$. Then it is enough to prove that

$$\int_0^D |u(x_N)|^2\, dx_N \le \frac{D^2}{\pi^2} \int_0^D |u'(x_N)|^2\, dx_N \text{ for all } u \in C_c^\infty(0,D),$$

and an integration in $x_1, \ldots, x_{N-1}$ gives the result. If one does not care for the best constant, one notices that $|u(x)|^2 = \big|\int_0^x u'(y)\, dy\big|^2$, which by CAUCHY-SCHWARZ inequality is $\le x \int_0^x |u'|^2\, dy \le D \int_0^D |u'|^2\, dy$, and then one integrates in x.

In order to show that D^2/π^2 is the best constant in the above 1-dimensional POINCARÉ inequality, one develops $0 \le \int_0^D \big|u'(y) - u(y)\frac{\varphi'(y)}{\varphi(y)}\big|^2\, dy$ with $\varphi(y) > 0$ in $(0,D)$, and as $-\int_0^D 2u\, u' \frac{\varphi'}{\varphi}\, dy = \int_0^D |u|^2\big(\frac{\varphi''}{\varphi} - \frac{|\varphi'|^2}{\varphi^2}\big)\, dy$, one finds that $\int_0^D |u'|^2\, dy \ge \int_0^D |u|^2\big(\frac{-\varphi''}{\varphi}\big)\, dy$, and then one takes $\varphi(y) = \sin\big(\frac{y\pi}{D}\big)$, and the constant is the best by letting u converge to φ.

As the proof of ii) is based on FOURIER transform, I first recall the basic theory developed by Laurent SCHWARTZ. For a function $f \in L^1(R^N)$, one defines

$$\mathcal{F}f(\xi) = \int_{R^N} f(x) e^{-2i\pi(x.\xi)}\, dx; \ \overline{\mathcal{F}}f(\xi) = \int_{R^N} f(x) e^{+2i\pi(x.\xi)}\, dx.$$

Then one notices that for f a little smoother, one has the two formulas

$$g = \frac{\partial f}{\partial x_j} \text{ implies } \mathcal{F}g(\xi) = 2i\pi\xi_j \mathcal{F}f(\xi)$$
$$g(x) = -2i\pi\, x_j f(x) \text{ implies } \mathcal{F}g = \frac{\partial \mathcal{F}}{\partial \xi_j}.$$

Then one introduces the space $S(R^N)$ of C^∞ functions φ such that $x^\alpha D^\beta \varphi \in L^\infty$ for every multi-indices α, β (it is a FRÉCHET space). This is a natural space where the preceding formulas are true and can iterated as many times as one wishes (as Laurent SCHWARTZ once told me, at the time he did that it was a new idea to introduce a functional space adapted to a given operator). One shows that $\mathcal{F}$ is an isomorphism from $S(R^N)$ onto itself, with inverse $\overline{\mathcal{F}}$, and one proves PLANCHEREL formula

$$\int_{R^N} f(x)\mathcal{F}g(x)\, dx = \int_{R^N} \mathcal{F}f(\xi) g(\xi)\, d\xi \text{ for } f, g \in S(R^N),$$

which one uses in order to define the FOURIER transform of some distributions (called temperate distributions), namely those which are in $S'(R^N)$, the dual space of $S(R^N)$, by

$$\langle \mathcal{F}T, \varphi\rangle = \langle T, \mathcal{F}\varphi\rangle \text{ for all } T \in S'(R^N) \text{ and all } \varphi \in S(R^N),$$

and in order to prove that FOURIER is an isometry of $L^2(R^N)$ onto itself, as well as its inverse $\overline{\mathcal{F}}$,

$$\int_{R^N} |f(x)|^2\,dx = \int_{R^N} |\mathcal{F}f(\xi)|^2\,d\xi \text{ for all } f \in L^2(R^N).$$

Proof of ii): For $u \in C_c^\infty(\Omega)$, extended by 0 outside Ω, one wants to bound $\int_{R^N} |\mathcal{F}u(\xi)|^2\,d\xi$ in terms of $\int_{R^N} 4\pi^2|\xi|^2|\mathcal{F}u(\xi)|^2\,d\xi$.

For $|\xi| \le \rho$, one uses $|\mathcal{F}u(\xi)| \le \int_\Omega |u(x)e^{-2i\pi(x.\xi)}|\,dx \le meas(\Omega)^{1/2}||u||_{L^2(R^N)}$, which gives

$$\int_{|\xi|\le\rho} |\mathcal{F}u(\xi)|^2\,d\xi \le meas\Big(B(0,1)\Big)\rho^N meas(\Omega)\int_{R^N}|u|^2\,dx,$$

and for $|\xi| \ge \rho$, one uses

$$\int_{|\xi|\ge\rho} |\mathcal{F}u(\xi)|^2\,d\xi \le \frac{1}{\rho^2}\int_{|\xi|\ge\rho} |\xi|^2|\mathcal{F}u(\xi)|^2\,d\xi \le \frac{1}{4\pi^2\rho^2}\int_{R^N} 4\pi^2|\xi|^2|\mathcal{F}u(\xi)|^2\,d\xi.$$

Adding these two inequalities and choosing the best ρ (given by $meas\big(B(0,1)\big)\rho^N meas(\Omega) = \frac{2}{N+2}$) gives the result.

The best constant (probably a result of TALENTI, using techniques of radial decreasing rearrangement), is obtained when Ω is a ball and involves the first zero of a BESSEL function.

In the rough proof of i), it was only used that u was zero on one hyperplane, and the best constant for functions which are only zero at 0 is four times larger than for functions which are zero at 0 and at D: it is obtained by taking $\varphi(x) = \sin\frac{y\pi}{2D}$ so that $\varphi'(D) = 0$.

Using the "equivalence lemma", one can show that if the injection from $H^1(\Omega)$ into $L^2(\Omega)$ is compact, then POINCARÉ inequality holds for every closed subspace of $H^1(\Omega)$ which does not contain the constant function 1 (assuming that Ω is connected, of course). The compactness assumption rules out many unbounded domains: if there exists $\alpha > 0$ such that $B(x_m, \alpha) \subset \Omega$ and $|x_m| \to \infty$ then the injection is not compact. Indeed if $\varphi \in C_c^\infty\big(B(0,\alpha)\big)$ is not 0, then $u_n(x) = \varphi(x - x_m)$ is a sequence which converges to 0 in $H^1(\Omega)$ weak (because its support tends to infinity), but which does not converge in $L^2(\Omega)$ strong. The compactness assumption holds for bounded domains (or domains with finite measure) under some smoothness hypothesis concerning the boundary.

Our first incursion in models for fluids will be to consider the stationary incompressible STOKES equation. It would be difficult to prove much on the question of starting with the compressible case and letting the MACH number tend to 0, and therefore I will describe another approach which is to consider these equations as a limit case of the equations of linearized Elasticity. I will not recall the defects of linearized Elasticity, and here we are only interested in the similarities at the level of the equations. As was observed experimentally by Dan JOSEPH, fluids do have some elastic properties, but for the moment I do not want to discuss the limitations of the usual viscosity argument.

I will only consider the isotropic case

$$\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\delta_{ij}\sum_k \varepsilon_{kk}, \text{ with } \varepsilon_{ij} = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big),$$

where the shear modulus $\mu > 0$ and the LAMÉ parameter λ either satisfy the "very strong ellipticity condition" $2\mu + N\lambda > 0$ if coefficients are variable, or the "strong ellipticity condition" $2\mu + \lambda > 0$ if the coefficients are constant and one uses DIRICHLET conditions.

21-820. PDE Models in Oceanography
Luc Tartar, Wean Hall 6212, 268-5734, tartar@andrew.cmu.edu

11. Friday February 5.

We will approach the stationary incompressible Stokes equation by first considering stationary linearized Elasticity for isotropic materials, i.e. study the equilibrium equation

$$-\sum_j \frac{\partial \sigma_{ij}}{\partial x_j} = f_i \text{ in } \Omega \text{ for each } i,$$

with the stress strain relation

$$\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\delta_{ij}\sum_k \varepsilon_{kk},$$

where

$$\varepsilon_{ij} = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big),$$

and u is the displacement. Notice that apart from gravity it is not very realistic to imagine forces acting directly inside Ω, and for a fluid the gravity forces can be incorporated into the pressure term; of course there are electromagnetic forces for conducting fluids, but then one must couple the equations with Maxwell equation, or there could be chemical forces, but one needs then to consider a larger system taking into account all the chemical species present.

We will begin by using Dirichlet boundary conditions, because it is reasonable for a fluid, as the viscosity imposes that the fluid must move at the same velocity than the boundary. However, because that condition is obviously not good at the surface of the ocean, we will have to study other boundary conditions. The use of Dirichlet conditions will also simplify the analysis for the linearized Elasticity equation, as one does not need to prove Korn inequality, because of the following identity.

Lemma: For any open set of R^N, one has

$$\int_\Omega \sum_{jk} |\varepsilon_{jk}|^2\,dx = \frac{1}{2}\int_\Omega \Big(|div(u)|^2 + \sum_j |grad(u_j)|^2\Big)\,dx \text{ for all } u_1,\ldots,u_N \in H^1_0(\Omega).$$

I will give two proofs, very similar, the first one using Fourier transform, the second one using integration by parts.
First proof: The functions u_j are extended by 0 outside Ω and one uses their Fourier transform.

$$\int_\Omega \sum_{jk} |\varepsilon_{jk}|^2\,dx = \int_{R^N} \sum_{jk} |\varepsilon_{jk}|^2\,dx = \int_{R^N} \sum_{jk} \pi^2|\xi_j\mathcal{F}u_k + \xi_k\mathcal{F}u_j|^2\,d\xi =$$
$$= \int_{R^N} \pi^2 \sum_{jk}\Big(|\xi_j|^2|\mathcal{F}u_k|^2 + |\xi_k|^2|\mathcal{F}u_j|^2 + 2\xi_j\xi_k\Re(\mathcal{F}u_k\overline{\mathcal{F}u_j})\Big)\,d\xi =$$
$$= \int_{R^N} 2\pi^2\Big(|\xi|^2\sum_j|\mathcal{F}u_j|^2 + |\sum_j \xi_j\mathcal{F}u_j|^2\Big)\,d\xi = \frac{1}{2}\int_{R^N}\Big(\sum_j|grad(u_j)|^2 + |div(u)|^2\Big)\,dx =$$
$$= \frac{1}{2}\int_\Omega\Big(\sum_j|grad(u_j)|^2 + |div(u)|^2\Big)\,dx.$$

Second proof:

$$\int_\Omega \sum_{jk}|\varepsilon_{jk}|^2\,dx = \frac{1}{4}\int_\Omega \sum_{jk}\Big(\frac{\partial u_j}{\partial x_k} + \frac{\partial u_k}{\partial x_j}\Big)^2\,dx =$$
$$= \frac{1}{2}\int_\Omega \sum_j |grad(u_j)|^2\,dx + \frac{1}{2}\int_\Omega\Big(\sum_{jk}\frac{\partial u_j}{\partial x_k}\frac{\partial u_k}{\partial x_j}\Big)\,dx,$$

and the result is a consequence of the property

$$\int_\Omega \frac{\partial v}{\partial x_j}\frac{\partial w}{\partial x_k}\,dx = \int_{R^N} \frac{\partial v}{\partial x_k}\frac{\partial w}{\partial x_j}\,dx \text{ for every } v,w\in H_0^1(\Omega).$$

This last formula is actually true if v or w belongs to $H_0^1(\Omega)$ and the other belongs to $H^1(\Omega)$, as for any distribution T and $\varphi\in C_c^\infty(\Omega)$ one has

$$\Big\langle \frac{\partial T}{\partial x_j}, \frac{\partial \varphi}{\partial x_k}\Big\rangle = -\Big\langle T, \frac{\partial^2\varphi}{\partial x_j\partial x_k}\Big\rangle = \Big\langle \frac{\partial T}{\partial x_k}, \frac{\partial \varphi}{\partial x_j}\Big\rangle,$$

and therefore the result is true for $v\in H^1(\Omega)$ and $\varphi\in C_c^\infty(\Omega)$, and by density it remains true for $\varphi\in H_0^1(\Omega)$.

The basic existence theorem for variational elliptic problems is the LAX-MILGRAM lemma, or one of its variants.

LAX-MILGRAM lemma: Let V be a real HILBERT space, and $A\in\mathcal{L}(V,V')$ satisfying the V-ellipticity condition

$$\text{there exists } \alpha>0 \text{ such that } (A\,u.u)\geq\alpha||u||^2 \text{ for all } u\in V.$$

Then A is an isomorphism from V onto V'.

From a practical point of view, an equivalent formulation is to have a continuous bilinear form $a(u,v)$ on $V\times V$, and of course $(a(u,v)=(Au.v)$ for every $u,v\in V$; the V-ellipticity condition is $a(u,u)\geq\alpha||u||_2$ for all $u\in V$; the conclusion is that for every linear continuous form $L(v)$ on V, there exists a unique $u\in V$ such that $a(u,v)=L(v)$ for every $v\in V$. One advantage of this formulation is that one does not have to identify what V' is. Another advantage is that the same formulation can be used directly for numerical methods (like finite elements): one first creates a family of finite dimensional spaces $V_h\subset V$, usually made of simple functions on a triangulation of Ω with h related to the mesh size, and one computes $u_h\in V_h$, unique solution of $a(u_h,v_h)=L(v_h)$ for all $v_h\in V_h$ by using techniques of Linear Algebra; the rate of convergence of u_h to the exact solution u is related to the fact that every element of V can be well approximated by sequences from V_h, and this is shown explicitly for smooth functions by estimating an interpolation error (usually the functions in V_h are defined by some values at the vertices of the triangulation, and one must compare a smooth function v to the function in V_h which has the same values than v at the vertices of the triangulation).

For a complex HILBERT space, one may use the hypothesis that a is a sesquilinear form on $V\times V$ such that $\Re\big(a(u,u)\big)\geq\alpha||u||^2$ for all $u\in V$, or that there exists θ such that $\Re\big(e^{i\theta}a(u,u)\big)\geq\alpha||u||^2$ for all $u\in V$, or even more generally that $|(a(u,u)|\geq\alpha||u||^2$ for all $u\in V$, as this condition implies the preceding one by a result of Eduardo ZARANTONELLO on the numerical range of an operator (the set of $\frac{a(v,v)}{||v||^2}$ for $v\neq0$ is a convex set of the complex plane). The same result holds without assuming that A is continuous, as continuity of A can be deduced by using the closed graph theorem. The same result is true for a BANACH space V, as one makes it a HILBERT space by using an equivalent norm corresponding to the scalar product $a(u,v)+a(v,u)$ in the real case, or $a(u,v)+\overline{a(v,u)}$ in the complex case.

Proof of LAX-MILGRAM lemma: As $\alpha||u||^2\leq(A\,u.u)\leq||A\,u||_*||u||$, where $||\cdot||_*$ denotes the dual norm on V', one deduces that

$$||A\,u||_*\geq\alpha||u|| \text{ for all } u\in V,$$

and this means that A is injective and has closed range (if $A\,u_n\to f$, then $A\,u_n$ is a CAUCHY sequence in V', and therefore u_n is a CAUCHY sequence in V, which then converges to some u_∞, and $f=A\,u_\infty$). As $(A\,u,u)=(u,A^Tu)$, one sees that A^T satisfies the same hypothesis, and therefore A^T is injective, which is equivalent to A having a dense range. The range of A being closed and dense is then equal to V'. Therefore A is a bijection from V to V', and its inverse is automatically continuous by the closed graph theorem, which can be avoided here as $||A\,u||_*\geq\alpha||u||$ for all $u\in V$ shows directly that $||A^{-1}||\leq\frac{1}{\alpha}$.

It is useful to know the following variant of the LAX-MILGRAM lemma.

Lemma: Let V be a real HILBERT space, and $A \in \mathcal{L}(V, V')$ such that

$$\text{there exists } \alpha > 0 \text{ such that } (A\,u.u) \geq 0, \text{ and } ||A\,u||_* \geq \beta ||u|| \text{ for all } u \in V.$$

Then A is an isomorphism from V onto V'.

I do not know if this variant was known before R. Tyrrell ROCKAFELLAR proved something analogous for monotone operators; the proof below is the one that I immediately checked with Jean-Claude NEDELEC when we learned about ROCKAFELLAR's result in the late 60s.
Proof of lemma: Let Λ be the canonical isometry from V onto V', i.e. $\Lambda \in \mathcal{L}(V, V')$, satisfies $||\Lambda\,u||_* = ||u||$ and $(\Lambda\,u.u) = ||u||^2$ for all $u \in V$. Then for $\varepsilon > 0$, one can apply LAX-MILGRAM lemma to $A + \varepsilon\Lambda$, with $\alpha = \varepsilon$ and therefore for a given $f \in V'$ there exists a unique $u_\varepsilon \in V$ such that $A\,u_\varepsilon + \varepsilon\Lambda\,u_\varepsilon = f$. Taking the scalar product with u_ε gives the first estimate $\varepsilon||u_\varepsilon|| \leq ||f||_*$, from which one deduces that $||A\,u_\varepsilon||_* \leq ||f||_* + \varepsilon||\Lambda\,u_\varepsilon||_* \leq 2||f||_*$. Therefore $\beta||u_\varepsilon|| \leq 2||f||_*$, and one deduces then that $A\,u_\varepsilon \to f$ in V'. In consequence, $A\,u_{\varepsilon_n}$ is a CAUCHY sequence in V' if $\varepsilon_n \to 0$, so that u_{ε_n} is a CAUCHY sequence in V, and its limit u_0 satisfies $A\,u_0 = f$.

As we will see, the existence theorem for linearized Elasticity will be a simple application of LAX-MILGRAM lemma, and there will not be too much difficulty proving an abstract theorem for describing the limit as the LAMÉ parameter λ tends to infinity: we will find that $div(u) = 0$ at the limit, and the limit problem will correspond to a situation where LAX-MILGRAM lemma applies, but for the space $V = \{u \in H_0^1(\Omega; R^N)$ satisfying $div(u) = 0\}$. Expressing in a concrete way the partial differential equation that the limit satisfies will formally involve a LAGRANGE multiplier, the "pressure", but there will be some technical obstacles to overcome before we can assert where the pressure is.

If one could apply the variant, the obtention of the equation would be straightforward, but checking the hypothesis of the variant will be a technical obstacle equivalent to estimating the pressure in the preceding approach.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

12. Monday February 8.

In the case of stationary linearized Elasticity with variable coefficients (in $L^\infty(\Omega)$), in the general form $\sigma_{ij} = \sum_{kl} C_{ijkl}\varepsilon_{kl}$, one uses the Very Strong Ellipticity condition

$$\text{there exists } \alpha > 0 : \sum_{ijkl} C_{ijkl}(x) M_{kl} M_{ij} \geq \alpha \sum_{ij} M_{ij}^2 \text{ for all symmetric } M, \text{ a.e. in } \Omega.$$

In the case of isotropic materials, i.e. $\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\delta_{ij}\sum_k \varepsilon_{kk}$, it means that $2\mu\sum_{ij} M_{ij}^2 + \lambda(\sum_k M_{kk})^2 \geq \alpha\sum_{ij} M_{ij}^2$, and therefore $\mu(x) \geq \beta > 0$ a.e. in Ω and (in the case $\lambda < 0$) $2\mu(x) + N\lambda(x) \geq \gamma > 0$ a.e. in Ω. Then LAX-MILGRAM lemma applies: the HILBERT space V is $H_0^1(\Omega; R^N)$, the continuous bilinear form a is given by $a(u,v) = \int_\Omega \sum_{ijkl} C_{ijkl}\varepsilon_{kl}(u)\varepsilon_{ij}(v)\,dx$, where $\varepsilon_{ij}(v)$ means $\frac{1}{2}\big(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_i}{\partial x_j}\big)$, the linear continuous form L is given by $L(v) = \int_\Omega \sum_i f_i v\,dx$, with $f_i \in L^2(\Omega)$ for $i = 1, \ldots, N$, and therefore if POINCARÉ inequality holds for $H_0^1(\Omega)$, the Very Strong Ellipticity condition implies the hypothesis of LAX-MILGRAM lemma.

In the case of stationary linearized Elasticity with constant coefficients and DIRICHLET conditions, one uses the Strong Ellipticity condition

$$\text{there exists } \alpha > 0 : \sum_{ijkl} C_{ijkl} a_k \xi_l a_i \xi_j \geq \alpha |a|^2 |\xi|^2 \text{ for all } a, \xi \in R^N, \text{ a.e. in } \Omega,$$

under the symmetry hypothesis $C_{ijkl} = C_{jikl} = C_{ijlk}$. Instead of using a lower bound for the integrand as when the Very Strong Ellipticity Condition holds, one integrates in Ω, but because of DIRICHLET condition it is an integral on R^N, for which one uses FOURIER transform (one could obtain the same result by integration by parts); using the symmetries of the coefficients, one finds $4\pi^2\int_{R^N}\sum_{ijkl} C_{ijkl}\mathcal{F}u_k\xi_l\overline{\mathcal{F}u_i}\xi_j\,d\xi$, and as the hypothesis implies $\Re\big(\sum_{ijkl} C_{ijkl} a_k\xi_l\overline{a_i}\xi_j\big) \geq \alpha|a|^2|\xi|^2$ for all $a \in C^N$ and all $\xi \in R^N$, the integral (which is real) is bounded below by $4\pi^2\int_{R^N}|\xi|^2\sum_k|\mathcal{F}u_k|^2\,d\xi$, which is $\int_\Omega\sum_k|grad(u_k)|^2\,dx$. In the case of isotropic materials, the Strong Ellipticitycondition means $\mu > 0$ and $2\mu + \lambda > 0$.

As we are interested in letting λ tend to infinity, we assume that $\mu(x) \geq \beta > 0$ a.e. in Ω and λ is a nonnegative constant tending to $+\infty$. The equilibrium equation can be written as

$$a^0(u^\lambda, v) + \lambda\, b(u^\lambda, v) = L(v) \text{ for all } v \in V = H_0^1(\Omega; R^N),$$

where L is a linear continuous form on V, and

$$a^0(u,v) = \int_\Omega 2\mu(x)\sum_{i,j=1}^N \varepsilon_{ij}(u)\varepsilon_{ij}(v)\,dx$$
$$b(u,v) = \int_\Omega div(u)\,div(v)\,dx,$$

and it has a unique solution $u^\lambda \in H_0^1(\Omega; R^N)$. This type of problem, related to the method of penalization or to questions of singular perturbations, has been extensively studied by Jacques-Louis LIONS in the late 60s/early 70s. Here the abstract treatment is straightforward: as a^0 is V-elliptic and $b(u,u)$ is the square of the norm of $div(u)$, one uses $v = u^\lambda$ and one obtains $\gamma||u^\lambda||^2 + \lambda|div(u^\lambda)|^2 \leq C||u^\lambda||$ (where $||\cdot||$ is the norm in V, and $|\cdot|$ is the norm in $L^2(\Omega)$), which gives

$$u^\lambda \text{ bounded in } V;\ div(u^\lambda) \to 0 \text{ in } L^2(\Omega),$$

which permits to extract a subsequence converging in V weak to u^∞; the fact that u^∞ will be characterized as the unique solution of an associated problem will show that all the sequence converges weakly (we will also show strong convergence). As $div(u^\infty) = 0$, we introduce the new space

$$W = \{u \in H_0^1(\Omega; R^N), div(u) = 0 \text{ a.e. in } \Omega\},$$

and so $u^\infty \in W$, but as $b(u,v)=0$ for $v\in W$, one has $a^0(u^\lambda,v)=L(v)$ for every $v\in W$, which shows that

$$a^0(u^\infty,w)=L(w) \text{ for all } w\in W, \text{ and } u^\infty\in W,$$

which characterizes u^∞ because a^0 is V-elliptic and therefore W-elliptic. In order to show strong convergence, one notices that $a^0(u^\lambda,u^\lambda)\le L(u^\lambda)$ and therefore

$$\limsup_{\lambda\to\infty} a^0(u^\lambda,u^\lambda)\le L(u^\infty)=a^0(u^\infty,u^\infty),$$

from which one deduces that

$$\limsup_{\lambda\to\infty} a^0(u^\lambda-u^\infty,u^\lambda-u^\infty)=\limsup_{\lambda\to\infty} a^0(u^\lambda,u^\lambda)-a^0(u^\infty,u^\infty)\le 0,$$

and therefore u^λ converges to u^∞ in V strong as $\lambda\to\infty$.

The problem is now to identify what equation u^∞ satisfies, and the difficulty comes from the fact that, because of the constraint $div(v)=0$, one is not allowed to use arbitrary test functions in $C_c^\infty(\Omega)$ which would give us an equation in the sense of distributions.

One could use the fact that our problem is equivalent to minimizing $a^0(v,v)-2L(v)$ on the subspace of V defined by an equation $div(v)=0$, and argue that there will be a LAGRANGE multiplier $q\in L^2(\Omega)$ such that v minimizes $a^0(v,v)-2L(v)+2\big(q.div(v)\big)$ without constraints, but the proof that such a LAGRANGE multiplier exists requires some care.

One could deduce an equation for u^∞ if one knew to what element $\lambda\, div(u^\lambda)$ converges, but for the moment we only know that $\sqrt{\lambda}\, div(u^\lambda)$ is bounded in $L^2(\Omega)$. In a recent discussion with François MURAT, I noticed that if Ω has a LIPSCHITZ boundary, then $\lambda\, div(u^\lambda)$ stays bounded in $L^2(\Omega)$.

In order to prepare for the discussion, we need to become familiar with $H^{-1}(\Omega)$, which is defined as the dual of $H_0^1(\Omega)$; as $C_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, its dual is a space of distributions in Ω, which is characterized as

$$H^{-1}(\Omega)=\Big\{f\in\mathcal{D}'(\Omega): f=g_0+\sum_{j=1}^N \frac{\partial g_j}{\partial x_j} \text{ with } g_0,\ldots,g_N\in L^2(\Omega)\Big\}.$$

Indeed let A be the linear continuous mapping from $H_0^1(\Omega)$ into $L^2(\Omega;R^{N+1})$ defined by

$$A\,u=\Big(u,\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_N}\Big).$$

$H_0^1(\Omega)$ is a closed subspace of $H^1(\Omega)$, which is complete (it is a HILBERT space) because if $u_n\in H^1(\Omega)$ is such that $u_n\to v_0$ in $L^2(\Omega)$ and $\frac{\partial u_n}{\partial x_j}\to v_j$ in $L^2(\Omega)$ then $v_j=\frac{\partial v_0}{\partial x_j}$ by using the definition of derivatives in the sense of distributions: for $\varphi\in C_c^\infty(\Omega)$ one has $\int_\Omega \frac{\partial u_n}{\partial x_j}\varphi\,dx=\langle\frac{\partial u_n}{\partial x_j},\varphi\rangle=-\langle u_n,\frac{\partial\varphi}{\partial x_j}\rangle=-\int_\Omega u_n\frac{\partial\varphi}{\partial x_j}\,dx$, which gives $\int_\Omega v_j\,\varphi\,dx=-\int_\Omega v_0\frac{\partial\varphi}{\partial x_j}\,dx$ for all $\varphi\in C_c^\infty(\Omega)$. By definition of the norm of $H_0^1(\Omega)$, $||A\,u||$ is equivalent to $||u||$, and as $H_0^1(\Omega)$ is complete the range of A is complete and therefore closed in $L^2(\Omega;R^{N+1})$. If L is a linear continuous form on $H_0^1(\Omega)$, it defines a linear continuous form on $R(A)$, which extends (by HAHN-BANACH theorem, or by using orthogonal projection on $R(A)$) to a linear continuous form on $L^2(\Omega;R^{N+1})$, which is of the form $(v_0,\ldots,v_N)\mapsto\sum_{k=0}^N\int_\Omega h_k v_k\,dx$ with $h_0,\ldots,h_N\in L^2(\Omega)$, and therefore $L(u)=\int_\Omega h_0 u\,dx+\sum_{j=1}^N\int_\Omega h_j\frac{\partial u}{\partial x_j}\,dx$. If POINCARÉ inequality holds, one can use $A\,u=(\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_N})$ instead, and one finds that every $g_0\in L^2(\Omega)$ can be written as $\sum_{j=1}^N\frac{\partial g_j}{\partial x_j}$ for some $g_j\in L^2(\Omega)$ for $j=1,\ldots,N$. One notices that each derivative $\frac{\partial}{\partial x_j}$ is a linear continuous operator from $L^2(\Omega)$ into $H^{-1}(\Omega)$.

For $0\le\lambda<\infty$, using $v\in C_c^\infty(\Omega)$ and $L(v)=\sum_i\langle f_i,\varphi\rangle$ with $f_i\in H^{-1}(\Omega)$ for $j=1,\ldots,N$, one deduces the equilibrium equation

$$-\sum_j\frac{\partial}{\partial x_j}\Big[\mu\Big(\frac{\partial u_i^\lambda}{\partial x_j}+\frac{\partial u_j^\lambda}{\partial x_i}\Big)\Big]-\frac{\partial\Big(\lambda\,div(u^\lambda)\Big)}{\partial x_i}=f_i \text{ for } i=1,\ldots,N,$$

and therefore, as u^λ is bounded in $H^1_0(\Omega)$, one finds that $\frac{\partial\big(\lambda\, div(u^\lambda)\big)}{\partial x_i}$ is bounded in $H^{-1}(\Omega)$, so that a subsequence converges to an element $T_i \in H^{-1}(\Omega)$; at this level it is not obvious that $\lambda\, div(u^\lambda)$ stays bounded, even in the space of distributions.

The important property of T_i, $i = 1,\ldots,N$, is that one has the equation

$$-\sum_j \frac{\partial}{\partial x_j}\Big[\mu\Big(\frac{\partial u_i^\infty}{\partial x_j}+\frac{\partial u_j^\infty}{\partial x_i}\Big)\Big] - T_i = f_i \text{ in } \Omega, \text{ for } i = 1,\ldots,N,$$

and

$$\sum_{i=1}^n \langle T_i, w_i\rangle = 0 \text{ for all } w \in W,$$

by taking the limit either for the equations or for the variational formulation.

One has $\frac{\partial T_i}{\partial x_j} = \frac{\partial T_j}{\partial x_i}$ for all i,j, either but noticing that $T_i^\lambda = \frac{\partial\big(\lambda\, div(u^\lambda)\big)}{\partial x_i}$ satisfies the same property as it is a gradient, or (if $i \neq j$) by using $w_i = \frac{\partial \psi}{\partial x_j}$, $w_i = -\frac{\partial \psi}{\partial x_j}$ and $w_k = 0$ for $k \neq i$ and $k \neq j$. By a result of Laurent SCHWARTZ, on each open ball (or any simply connected open subset ω) of Ω there exists a distribution S such that $T_i = \frac{\partial S}{\partial x_i}$. However, there are more functions in W, which can see the topology of Ω, and a result of DE RHAM asserts that if distributions T_i, $i = 1,\ldots,N$, satisfy $\sum_{i=1}^n \langle T_i, \varphi_i\rangle = 0$ for all $\varphi_i \in C_c^\infty(\Omega)$, $i = 1,\ldots,N$, satisfying $div(\varphi) = 0$, then there exists a distribution S such that $T_i = \frac{\partial S}{\partial x_i}$ for $i = 1,\ldots,N$. If one admits this result, the question will be to prove that if all T_i belong to $H^{-1}(\Omega)$ then S belongs to $L^2(\Omega)$, and that requires some smoothness of the boundary: I have noticed a few months ago that the result is not true if the boundary is not LIPSCHITZ continuous, and François MURAT has just mentioned to me a similar counter-example by Giuseppe GEYMONAT and Gianni GILARDI, motivated by showing that KORN's inequality does not always hold if the boundary is not LIPSCHITZ continuous, the connection between the two questions being that they both are related to the space $X(\Omega) = \{u \in H^{-1}(\Omega), \frac{\partial u}{\partial x_i} \in H^{-1}(\Omega), i = 1,\ldots,N\}$ being $L^2(\Omega)$ or not.

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

13. Wednesday February 10.

The idea of considering the space

$$X(\Omega) = \Big\{u \in H^{-1}(\Omega), \frac{\partial u}{\partial x_j} \in H^1(\Omega), j = 1, \ldots, N\Big\},$$

seems due to Jacques-Louis LIONS. We have indeed found that $T_j \in H^{-1}(\Omega)$, $j = 1, \ldots, N$, satisfies $\sum_j \langle T_j, w_j \rangle = 0$ for all $w \in W$, and DE RHAM's result asserts that there exists a distribution S such that $T_j = \frac{\partial S}{\partial x_j}$ for $j = 1, \ldots, N$. Why should S itself belong to $H^{-1}(\Omega)$? That condition is indeed useful, because the information $u \in H^{-1}(\Omega)$ makes the space local, i.e. $u \in X(\Omega)$ implies $\psi u \in X(\Omega)$ for every $\psi \in C_c^\infty(R^N)$, and one can then use partitions of unity in order to study the functions of $X(\Omega)$. The approach of Jacques-Louis LIONS seems to have been to use DE RHAM's result and then prove that $S \in L^2(\Omega)$; he quotes an article of MAGENES and STAMPACCHIA (which I have not read, but LIONS told me that they had mentioned there that they were using one of his results).

I present here the approach which I developed in 1974, which does not rely on DE RHAM's result; I only used it for smooth domains, but the result is indeed true for LIPSCHITZ domains: Olga LADYZHENSKAYA may have done it, but Jindrich NECAS certainly did it; GOBERT is mentioned for KORN's inequality, which is related as follows. KORN's inequality is about proving that if $u \in L^2(\Omega; R^N)$ and $\varepsilon_{ij} \in L^2(\Omega)$ for all $i,j = 1, \ldots, N$, then one has $u \in H^1(\Omega; R^N)$; one notices that

$$2\frac{\partial^2 u_i}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j}\Big(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\Big) - \frac{\partial}{\partial x_i}\Big(\frac{\partial u_k}{\partial x_j} + \frac{\partial u_j}{\partial x_k}\Big) + \frac{\partial}{\partial x_k}\Big(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\Big),$$

and therefore $\frac{\partial^2 u_i}{\partial x_j \partial x_k} \in H^{-1}(\Omega)$; as all $\frac{\partial u_i}{\partial x_j}$ also belong to $H^{-1}(\Omega)$, it shows that all $\frac{\partial u_i}{\partial x_j}$ belong to $X(\Omega)$, and therefore if $X(\Omega) = L^2(\Omega)$, then KORN's inequality holds (GEYMONAT and GILARDI have shown the KORN's inequality does not hold for some non LIPSCHITZ domain, and therefore $X(\Omega) \neq L^2(\Omega)$ for the open set Ω that they considered; my remark that $X(\Omega) \neq L^2(\Omega)$ in many non LIPSCHITZ domains, seems to provide simpler explicit counter-examples, as I checked with François MURAT).

The result we are interested in is the following

Lemma: Assume that Ω is smooth enough. If $T_i \in H^{-1}(\Omega)$ for $i = 1, \ldots, N$ and $\sum_{i=1}^N \langle T_i, w_i \rangle = 0$ for all $w \in W = \{u \in H_0^1(\Omega; R^N), div(w) = 0\}$, then there exists $p \in L^2(\Omega)$ such that $T_i = \frac{\partial p}{\partial x_i}$ for $i = 1, \ldots, N$.

This result is equivalent to the following

Lemma: Assume that Ω is smooth enough. For $g \in L^2(\Omega)$ satisfying $\int_\Omega g\,dx = 0$, there exists $u \in H_0^1(\Omega; R^N)$ such that $div(u) = g$.

Indeed, $A = grad$ operates continuously from $L^2(\Omega)$ into $H^{-1}(\Omega; R^N)$, its transposed is $A^T = -div$ operates continuously from $H_0^1(\Omega; R^N)$ into $L^2(\Omega)$. The kernel of A^T is W, and therefore its orthogonal is the closure of $R(A)$, so the first lemma is equivalent to saying that $R(A)$ is closed. The kernel of A is generated by the constant function 1 (as Ω is always assumed to be connected), its orthogonal is the closure of $R(A^T)$, so the second lemma is equivalent to saying that $R(A^T)$ is closed, which is indeed equivalent to the statement that $R(A)$ is closed.

My method, which I called the equivalence lemma, appeared to be a generalization of an earlier result of Jack PEETRE (sufficient for the present situation); it shows that if the injection of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact (which is the case if $meas(\Omega) < \infty$), and if $X(\Omega) = L^2(\Omega)$ (which requires some smoothness of the boundary), then $R(A)$ is closed.

The other methods that I have heard of are concerned with solving $div(u) = g$. One, which was mentioned to me by Charles GOULAOUIC, uses some results that he had obtained with Salah BAOUENDI, and consists in solving an equation $-div\big(\delta\, grad(v)\big) = g$, where δ is the distance to the boundary, and take

$u = -\delta\, grad(v)$. Their regularity theorem, asserting that $v \in H^1(\Omega)$ and $\delta\, v \in H^2(\Omega)$ was proved using pseudo-differential operators which require a C^∞ boundary (in the late 60s I had derived a simpler proof based on interpolation, which certainly does not require as much smoothness). One should notice that the basic existence result uses the space of functions such that $\sqrt{\delta}\, grad(v) \in L^2(\Omega)$, in which $C_c^\infty(\Omega)$ is dense, and therefore no boundary conditions are added. This approach was analyzed I believe by BOLLEY and CAMUS.

One can construct explicit solutions in R^N_+ and estimates use CALDERÓN-ZYGMUND theorem; I believe that this was the approach taken by Olga LADYZHENSKAYA; Giovanni GALDI reproduces a proof valid for LIPSCHITZ boundaries.

Equivalence lemma: Let E_1 be a BANACH space and E_2, E_3 be normed spaces; let $A \in \mathcal{L}(E_1, E_2)$ and $B \in \mathcal{L}(E_1, E_3)$ satisfy the hypotheses

$$\begin{aligned} &||u||_{E_1} \approx ||A\,u||_{E_2} + ||B\,u||_{E_3} \\ &B \text{ is compact.} \end{aligned}$$

Then the kernel of A is finite dimensional and the range of A is closed.

There exists a constant K such that if $L \in \mathcal{L}(E_1, F)$ for a normed space F satisfies $L\,u = 0$ on the kernel of A, then one has the estimate $||L\,u||_F \leq K||L||_{\mathcal{L}(E_1,F)}\,||A\,u||_{E_2}$ for all $u \in E_1$.

If $p(u)$ is a continuous semi-norm on E_1 which is a norm on the kernel of A, then $||u||_{E_1} \approx ||A\,u||_{E_2} + p(u)$.

The result of Jaak PEETRE assumed reflexive BANACH spaces, and was concerned with the special case where B is the injection from E_1 into E_3, and this is usually the case in applications.

Proof: On $X = ker(A)$, one has $||u||_1 \approx ||B\,u||_3$ and as B is compact one deduces that the closed unit ball of X is compact, which by a theorem of F. RIESZ proves that X is finite dimensional.

X has a topological supplement Y (by HAHN-BANACH theorem), and there exists C such that $||A\,u||_{E_2} \geq C\,||u||_{E_1}$ on Y, which shows that $R(A)$ is closed, because if $A\,u_n \to f$, then $A\,u_n$ is a CAUCHY sequence, and therefore u_n is a CAUCHY sequence, converging to u_∞ and $f = A\,u_\infty$. If the bound was not true there would exist a sequence v_n with norm 1 in Y such that $A\,v_n \to 0$ in E_2, but as $B\,v_n$ belongs to a compact of E_3 one can extract a subsequence v_m such that $B\,v_m$ converges in E_3, and then $A\,v_m$ and $B\,v_m$ being CAUCHY sequences, v_m is a CAUCHY sequence, converging to an element v_∞ satisfying the contradictory properties $A\,v_\infty = 0$ and $||v_\infty||_Y = 1$.

A is a bijection from Y onto its image Z, and its inverse D is continuous (so that Z is a BANACH space even if E_2 is not complete), and as $u - D\,A\,u$ belongs to $ker(A)$, one has $L\,u = L\,D\,A\,u$ for all $u \in E_1$, and one can take $K = ||D||$.

If $u \mapsto ||A\,u||_{E_2} + p(u)$ was not an equivalent norm on E_1, there would exist a sequence w_n with norm 1 in E_1 such that $||A\,w_n||_{E_2} + p(w_n) \to 0$. As $B\,w_n$ belongs to a compact of E_3 one can extract a subsequence w_m such that $B\,w_m$ converges, and using $A\,w_m \to 0$, one deduces that w_m is a CAUCHY sequence and converges to an element w_∞ satisfying the contradictory properties $A\,w_\infty = 0$, $p(w_\infty) = 0$ and $||w_\infty||_{E_1} = 1$.

In our setting, $E_1 = L^2(\Omega)$, $A = grad$ and $E_2 = H^{-1}(\Omega; R^N)$, B is the injection into $E_3 = H^{-1}(\Omega)$, and the hypothesis are satisfied if $meas(\Omega) < \infty$ (so that the injection from $H^1_0(\Omega)$ into $L^2(\Omega)$ is compact), and if $X(\Omega) = L^2(\Omega)$, proving that $R(A)$ is closed.

Some classical inequalities correspond to E_1 being a subspace of $H^1(\Omega)$, $A = grad$ and $E_2 = L^2(\Omega; R^N)$, B is the injection into $E_3 = L^2(\Omega)$: if the injection from $H^1(\Omega)$ into $L^2(\Omega)$ is compact, POINCARÉ inequality holds if (and only if) the constant function 1 does not belong to E_1 (Ω being assumed to be connected), and in the case where $1 \in E_1$ the condition for L is $L(1) = 0$, and the condition for p is $p(1) \neq 0$

Lemma: $X(R^N) = L^2(R^N)$.

Proof: Using FOURIER transform. $u \in H^{-1}(R^N)$ is equivalent to $\frac{\mathcal{F}u}{\sqrt{a^2+|\xi|^2}} \in L^2(R^N)$ (if x is a length, a and ξ have dimension length^{-1}). As $\mathcal{F}\frac{\partial u}{\partial x_j} = 2i\pi\xi_j\mathcal{F}u$, one deduces that $u \in X(R^N)$ is equivalent to $\frac{(1+\sum_j |\xi_j|)\mathcal{F}u}{\sqrt{a^2+|\xi|^2}} \in L^2(R^N)$, and as $\frac{1+\sum_j |\xi_j|}{\sqrt{a^2+|\xi|^2}}$ is bounded above and below, $X(R^N)$ is $L^2(R^N)$.

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

14. Friday February 12.

I will describe a few techniques for analyzing functions in various SOBOLEV spaces, truncation, regularization, localization, extension, and although we want to apply them to $X(\Omega)$, it is useful first to start with the simpler case corresponding to $H^1(\Omega)$.

Truncation: This is used in order to show that functions with compact support are dense, in the case where Ω is not bounded, of course. One chooses $\varphi \in C_c^\infty(R^N)$ such that $\varphi(x) = 1$ for $|x| \leq 1$, and for $u \in L^2(\Omega)$ one defines u_n by $u_n(x) = u(x)\varphi\big(\frac{x}{n}\big)$, which has compact support, is such that $|u_n(x)| \leq K|u(x)|$ and $u_n(x) \to u(x)$ a.e. (as $u_n(x) = u(x)$ for $n \geq |x|$), and therefore $u_n \to u$ in $L^2(\Omega)$ strong by LEBESGUE's dominated convergence theorem. If $u \in H^1(\Omega)$, $\frac{\partial u_n}{\partial x_j}$ is sum of two terms, the first one being $\frac{\partial u}{\partial x_j}\varphi\big(\frac{x}{n}\big)$ which converges to $\frac{\partial u}{\partial x_j}$ by LEBESGUE's dominated convergence theorem, and the second being $\frac{u}{n}\frac{\partial \varphi}{\partial x_j}\big(\frac{x}{n}\big)$ which has a norm in $L^2(\Omega)$ of the order of $1/n$.

Regularization: This is used in order to approach a given function by smoother functions, but the convolution process which is used for that purpose increases the size of the support, and that creates small problems. On R^N, the convolution product h of two integrable functions f and g, denoted $h = f \star g$, is defined by $h(x) = \int_{R^N} f(x-y)g(y)\,dy = \int_{R^N} f(y)g(x-y)\,dy$, and by FUBINI's theorem one has $||f\star g||_{L^1} \leq ||f||_{L^1}||g||_{L^1}$. The convolution product is commutative and associative on $L^1(R^N)$. The convolution product is well defined if $f \in L^1(R^N)$ and $g \in L^p(R^N)$ and gives $f \star g \in L^p(R^N)$ with $||f \star g||_{L^p(R^N)} \leq ||f||_{L^1(R^N)}||g||_{L^p(R^N)}$ for $1 \leq p \leq \infty$, as can be easily seen by applying HÖLDER inequality (or JENSEN inequality, which says that for every convex function Φ one has $\Phi\big(\int u\, f\, dx\big) \leq \int \Phi(u)\, f\, dx$ if $f \geq 0$ and $\int f\, dx = 1$). One also has $||f \star g||_{L^r(R^N)} \leq C(p,q)||f||_{L^p(R^N)}||g||_{L^q(R^N)}$, if $p, q, r \geq 1$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, but using HÖLDER inequality as before gives $C(p,q) = 1$, which is not the best constant if P, q and r are different from 1.

A regularizing sequence is a sequence $\rho_n \in C_c^\infty(R^N)$ such that the support of ρ_n converges to 0, ρ_n is bounded in $L^1(R^N)$ and $\int_{R^N} \rho_n\, dx \to 1$ (usually one chooses $\rho_1 \in C_c^\infty(R^N)$ having its support in the closed unit ball, with $\rho_1 \geq 0$ and $\int_{R^N} \rho_1\, dx = 1$, and then one defines ρ_n by $\rho_n(x) = n^N \rho_1(n\, x)$). For $\varphi \in C_c(R^N)$, one sees easily by using the uniform continuity of φ that $\rho_n \star \varphi$ converges to φ uniformly; then for $1 \leq p < \infty$, using the density of $C_c(R^N)$ in $L^p(R^N)$, one deduces that

$$\rho_n \star g \to g \text{ in } L^p(R^N) \text{ for every } g \in L^p(R^N) \text{ if } 1 \leq p < \infty.$$

Because the convolution product commutes with translations (which are actually convolutions with DIRAC masses), one sees easily that one has

$$\frac{\partial(f \star g)}{\partial x_j} = \frac{\partial f}{\partial x_j} \star g \text{ if } f \in C_c^1(R^N) \text{ and } g \in L^p(R^N), 1 \leq p \leq \infty,$$

by using the uniform continuity of the partial derivatives of f. One deduces then easily for example that the same formula holds if $f \in W^{1,p}(R^N)$ and $g \in L^1(R^N)$, by using the definition of derivatives in the sense of distributions and observing that $\int_{R^N}(f \star g)\varphi\, dx = \int_{R^N} f(\check g \star \varphi)\, dx$ with $\check g(x) = g(-x)$ for $\varphi \in C_c^\infty(R^N)$. If $1 \leq p < \infty$ and $f \in W^{1,p}(R^N)$, one first uses truncation to approach f by functions $f_m \in W^{1,p}(R^N)$ with compact support, and then using regularization one deduces that $\rho_n \star f_m \in C_c^\infty(R^N)$ and converges to f_m in $W^{1,p}(R^N)$, showing that $C_c^\infty(R^N)$ is dense in $W^{1,p}(R^N)$, i.e. $W_0^{1,p}(R^N) = W^{1,p}(R^N)$, which for $p = 2$ is written $H_0^1(R^N) = H^1(R^N)$.

Convolution product has been extended by Laurent SCHWARTZ to some pairs of distributions, but one must be careful with questions of support. For measurable functions one says that $x \notin support(f)$ if $f(y) = 0$ a.e. on $B(x,r)$ for some $r > 0$, and one deduces that in the classical case considered above, one has $support(f \star g) \subset support(f) + support(g)$. For functions in $L^1_{loc}(R)$ with support in $[0,\infty)$ one can define the convolution product, which also has support in $[0,\infty)$, and a theorem of TITCHMARSH states that $f \star g = 0$ if and only if one of the functions f or g is 0, and this was extended to N dimensions by Jacques-Louis LIONS, who has proved (at least for bounded supports) that $\overline{conv[support(f \star g)]} = \overline{conv[support(f)]} +$

$\overline{conv}[support(g)]$, where $\overline{conv}[A]$ is the closed convex hull of A, but it is rarely necessary to use such a refinement. The remark on the support of convolution products is used in order to define $\rho_n \star f$ even though f is not defined everywhere: for example if $f \in W^{1,p}(R^N_+)$ and ρ_n has its support in the strip $\alpha_n \le x_N \le \beta_n$, then $\rho_n \star f$ is well defined in $x_N > \beta_n$, i.e. if g is an extension of f to R^N, then the restriction of $\rho_n \star g$ to the open set $x_N > \beta_n$ is always the same, whatever the extension is. After truncation, let $1 \le p < \infty$ and let $f \in W^{1,p}(R^N_+)$ with compact support, let ρ_n be a regularizing sequence with $\beta_n < 0$, and let S denote the restriction to R^N_+; then $S(\rho_n \star f) \to f$ in $L^p(R^N_+)$, and $\frac{\partial[S(\rho_n \star f)]}{\partial x_j} = S\big(\rho_n \star \frac{\partial f}{\partial x_j}\big) \to \frac{\partial f}{\partial x_j}$ in $L^p(R^N_+)$. As $\rho_n \star f$ is C^∞ for $x_N > \beta_n$, one can mutiply it by a C^∞ function of x_N alone which is 1 for $x_N > 0$ and 0 for $x_N < \frac{\beta_n}{2}$, and therefore $S(\rho_n \star f) = S\varphi_n$ with $\varphi_n \in C^\infty_c(R^N)$, and so $S(\rho_n \star f) \in C^\infty\big(\overline{R^N_+}\big)$ which is by definition the space of restrictions to R^N_+ of functions in $C^\infty_c(R^N_+)$; therefore $C^\infty\big(\overline{R^N_+}\big)$ is dense in $W^{1,p}(R^N_+)$ for $1 \le p < \infty$.

Localization: In order to show that $C^\infty(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$ for $1 \le p < \infty$, one needs some regularity of the boundary of Ω, but the first step is to apply an argument of localization. Assume for example that the boundary $\partial\Omega$ of Ω is compact and that around each point of the boundary there is a small open ball in which the boundary has an equation $x_N = F(x_1, \ldots, x_{N-1})$ in some orthonormal basis, with F continuous, and that Ω is on one side of the boundary, say $x_N > F(x')$, where x' denotes $(x_1, \ldots, x_{N-1})$. The family of all these open balls is an open covering of $\partial\Omega$ from which one extracts a finite covering, to which one adds Ω, to have a finite open covering of $\overline{\Omega}$; let $\omega_j, j = 1, \ldots, J$, be that family of open sets, for which there exists a partition of unity, i.e. functions $\theta_j \in C^\infty_c(\omega_j)$ such that $\sum_{j=1}^J \theta_j = 1$ on $\overline{\Omega}$. One decomposes $u \in W^{1,p}(\Omega)$ as $\sum_{j=1}^J \theta_j u$, and one studies each $\theta_j u$ separately. For an index j corresponding to an open set ω_j which does not intersect the boundary, one uses the techniques developed for R^N, i.e. convolution by a regularizing sequence without paying much attention to its support (as long as it converges to 0), while for the other indices j one uses the techniques developed for R^N_+, i.e. convolution by a regularizing sequence with an adapted support. The problem is the same as for an open set Ω defined globally by an equation $x_N > F(x')$, with F uniformly continuous, and for $f \in W^{1,p}(\Omega)$ having compact support. By uniform continuity of F, for every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that $|x' - y'| \le 2\delta(\varepsilon)$ implies $|F(x') - F(y')| \le \varepsilon$ (one may assume that $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$). One chooses a sequence $\varepsilon_n \to 0$, and one chooses the regularizing sequence ρ_n with its support in $|x'| \le \delta(\varepsilon_n)$ and $x_N \le -\varepsilon_n - \delta(\varepsilon_n)$, and then $\rho_n \star f$ is defined in a domain extending at least a distance $\delta(\varepsilon_n)$ beyond the boundary of Ω, and the method used for R^N_+ applies (with S denoting the restriction to Ω).

Of course, there are open sets Ω for which $C^\infty\big(\overline{\Omega}\big)$ is not dense in $W^{1,p}(\Omega)$ for $1 \le p \le \infty$, the plane R^2 to which one removes the nonnegative x axis, for example: functions of $W^{1,p}(\Omega)$ may have different values on both sides of the removed half axis, while functions of $C^\infty\big(\overline{\Omega}\big)$ must have the same values on both sides.

Extension: For open sets for which $C^\infty\big(\overline{\Omega}\big)$ is dense in $W^{1,p}(\Omega)$ for $1 \le p < \infty$, one cannot always construct a linear continuous extension from $W^{1,p}(\Omega)$ into $W^{1,p}(R^N)$, i.e. a map P such that $S\,P\,u = u$ for all $u \in W^{1,p}(\Omega)$ (S being the restriction to Ω). One must extend each $\theta_j u$, and one is led to add the requirement that each F is LIPSCHITZ continuous; in the case of one open set of equation $x_N > F(x')$, one defines the extension P by

$$P\,u(x', x_N) = u(x', x_N) \text{ if } x_N > F(x')$$
$$P\,u(x', x_N) = u\Big(x', -x_N + 2F(x')\Big) \text{ if } x_N < F(x').$$

For $u \in C^\infty\big(\overline{\Omega}\big)$, one checks that the norm of $P\,u$ is controled by the norm of u, and the important fact is that $P\,u$ is continuous at the boundary of Ω.

The extension of $W^{m,p}(\Omega)$ to $W^{m,p}(R^N)$ for $m \ge 2$ is a little more technical (Alberto CALDERÓN had a proof using CALDERÓN-ZYGMUND theorem, and therefore it does not cover the cases $p = 1$ and $p = \infty$, but STEIN has constructed an extension which serves for all $W^{m,p}$ and is valid for $1 \le p \le \infty$), but for the case of $\Omega = R^N_+$, after proving that $C^\infty\big(\overline{\Omega}\big)$ is dense, one defines P by

$$P\,u(x', x_N) = u(x', x_N) \text{ if } x_N > 0$$
$$P\,u(x', x_N) = a_1 u(x', -b_1 x_N) + a_2 u(x', -b_2 x_N) + \text{ if } x_N < 0,$$

with $b_1, b_2 > 0$ and $b_1 \neq b_2$. The continuity of $P\,u$ (or its partial derivative in x_j for $j < N$) at $x_N = 0$ requires $a_1 + a_2 = 1$. The continuity of $\frac{\partial P\,u}{\partial x_N}$ at $x_N = 0$ requires $-a_1 b_1 - a_2 b_2 = 1$. As $b_1 \neq b_2$, the values of a_1 and a_2 are determined.

The extension property does not hold for some open sets with F only HÖLDER continuous with exponent $\alpha < 1$. The counter-example relies on SOBOLEV imbedding theorem, which I will discuss later, but works as follows. One has $H^1(R^2) \subset L^p(R^2)$ for all $p < \infty$, and therefore if there exists a continuous extension P from $H^1(\Omega)$ into $H^1(R^2)$, it implies that $H^1(\Omega) \subset L^p(\Omega)$ for all $p < \infty$, and the counter-example consists in showing a function of $H^1(\Omega)$ which does not belong to all $L^p(\Omega)$: for example if $\Omega = \{(x,y) : 0 < x < 1$ and $0 < y < x^2$, and $u(x) = x^\beta$, then $u \in L^p(\Omega)$ if and only if $p\beta + 2 > -1$ and $u \in H^1(\Omega)$ if and only if $2(\beta - 1) + 2 > -1$, i.e. $\beta > -1/2$, and therefore one can choose β such that $u \notin L^p(\Omega)$ for $p > 6$.

21-820. PDE Models in Oceanography

Luc Tartar, Wean Hall 6212, 268-5734, tartar@andrew.cmu.edu

15. Monday February 15.

We have seen that $X(R^N) = L^2(R^N)$ by using Fourier transform, and we consider now the case of $X(R^N_+)$. We have noticed that $X(\Omega)$ is a local space, i.e. $u \in X(\Omega)$ implies $\varphi\, u \in X(\Omega)$ for every $\varphi \in C^\infty_c(R^N)$, because $u \in H^{-1}(\Omega)$ implies $\varphi\, u \in H^{-1}(\Omega)$, and this kind of property is seen by duality: for $v \in C^\infty_c(\Omega)$, one has $\langle \varphi\, u, v\rangle = \langle u, \varphi\, v\rangle$ and therefore the key point is that multiplication by φ is continuous from $H^1_0(\Omega)$ into itself (it is sufficient that $\varphi \in W^{1,\infty}(\Omega)$, or if one uses Sobolev imbedding theorem, that $\varphi \in W^{1,N}(\Omega) \cap L^\infty(\Omega)$ for $N \geq 3$, $\varphi \in W^{1,p}(\Omega)$ with $p > 2$ for $N = 2$ and with $p = 2$ for $N = 1$).

After the preceding localization argument, one proves that $C^\infty\big(\overline{R^N_+}\big)$ is dense in $X(R^N_+)$ in a similar way that for $H^1(R^N_+)$, but a little care is useful if one wants to avoid using convolution of distributions. Every $u \in H^{-1}(R^N_+)$ can be written $f_0 + \sum_i \frac{\partial f_i}{\partial x_i}$, with $f_0, \ldots, f_N \in L^2(R^N_+)$, and one can consider that the f_j are extended for $x_N < 0$, but the precise extension will not be used; one chooses a regularizing sequence ρ_n with support in $\alpha_n \leq x_N \leq \beta_n$ with $\beta_n < 0$, and one chooses a C^∞ function ψ_n of x_N alone, which is 1 for $x_N \geq 0$ and 0 for $x_N < \frac{\beta_n}{2}$, and one approaches u by $\psi_n(\rho_n\star u)$ restricted to R^N_+. If $\varphi \in C^\infty_c(\Omega)$, then $\langle \psi_n(\rho_n\star u), \varphi\rangle = \langle(\rho_n\star u), \psi_n\varphi\rangle$, but $\psi_n\varphi = \varphi$, and so it is $\langle u, \check\rho_n\star\varphi\rangle = \big\langle f_0 + \sum_i \frac{\partial f_i}{\partial x_i}, \check\rho_n\star\varphi\big\rangle = \langle \rho_n\star f_0, \varphi\rangle + \sum_i \big\langle \frac{\partial(\rho_n\star f_i)}{\partial x_i}, \varphi\big\rangle$, and as the restriction of $\rho_n\star f_j$ to R^N_+ converges strongly to f_j in $L^2(R^N_+)$, one sees that the restriction of $\psi_n(\rho_n\star u)$ converges to u strongly in $H^{-1}(R^N_+)$. Then one notices that $\big\langle \frac{\partial[\psi_n(\rho_n\star u)]}{\partial x_i}, \varphi\big\rangle = -\big\langle(\rho_n\star u), \frac{\partial(\psi_n\varphi)}{\partial x_i}\big\rangle$, and as $\psi_n\varphi = \varphi$, it is $\big\langle \rho_n\star\frac{\partial u}{\partial x_i}, \varphi\big\rangle$, and therefore $\frac{\partial[\psi_n(\rho_n\star u)]}{\partial x_i}$ converges to $\frac{\partial u}{\partial x_i}$ strongly in $H^{-1}(R^N_+)$.

In order to construct an extension of $X(R^N_+)$ into $X(R^N)$, it is better to work by duality, and we will define an adequate restriction from $H^1(R^N)$ into $H^1_0(R^N_+)$. Let S denote the restriction to R^N_+, which is linear continuous from $H^{-1}(R^N)$ into $H^{-1}(R^N_+)$, as it is the transposed of the extension by 0 (denoted $\tilde{\cdot}$), which is linear continuous from $H^1_0(R^N_+)$ into $H^1(R^N)$. An extension is a map P such that $SP = identity$, then by transposition $P^T\tilde{\cdot} = identity$. We are looking then for a mapping from $H^1(R^N)$ into $H^1_0(R^N_+)$ such that if $\varphi \in H^1_0(R^N_+)$ then $P\widetilde{\varphi} = \varphi$ on R^N_+. Of course, we define the transposed map $Q = P^T$ on $C^\infty_c(R^N)$, which is dense in $H^1(R^N)$, the continuity of Q being easy to check when one uses the norms in H^1:

$$Q\, u(x', x_N) = u(x', x_N) + \sum_{j=1}^2 a_j u(x', -b_j x_N) \text{ for } x_N > 0,$$

where b_1 and b_2 are distinct and positive; Q is obviously linear continuous from $H^1(R^N)$ into $H^1(R^N_+)$ and satisfies $Q\,\widetilde{\varphi} = \varphi$ for every $\varphi \in H^1_0(R^N_+)$, but one needs the condition $1 + a_1 + a_2 = 0$ in order to ensure that Q maps $H^1(R^N)$ into $H^1_0(R^N_+)$. For $j < N$, one has $\frac{\partial(Q\varphi)}{\partial x_j} = Q\big(\frac{\partial\varphi}{\partial x_j}\big)$ for all $\varphi \in C^\infty_c(R^N)$, but for $j = N$ we introduce the operator R defined by

$$R\, u(x', x_N) = u(x', x_N) - \sum_{j=1}^2 \frac{a_j}{b_j} u(x', -b_j x_N) \text{ for } x_N > 0,$$

and one adds the condition $1 - \frac{a_1}{b_1} - \frac{a_2}{b_2} = 0$ (and so a_1 and a_2 are determined), in order to ensure that R maps $H^1(R^N)$ into $H^1_0(R^N_+)$, and one has $\frac{\partial(R\varphi)}{\partial x_N} = Q\big(\frac{\partial\varphi}{\partial x_N}\big)$. Therefore, using $P = Q^T$, if $u \in X(R^N_+)$, one has

$$\begin{aligned}
&\langle P\, u, \varphi\rangle = \langle u, Q\varphi\rangle\\
&\Big\langle \frac{\partial(P\, u)}{\partial x_i}, \varphi\Big\rangle = -\Big\langle P\, u, \frac{\partial\varphi}{\partial x_i}\Big\rangle = -\Big\langle u, Q\frac{\partial\varphi}{\partial x_i}\Big\rangle = -\Big\langle u, \frac{\partial(Q\varphi)}{\partial x_i}\Big\rangle = \Big\langle \frac{\partial u}{\partial x_i}, Q\varphi\Big\rangle\\
&\Big\langle \frac{\partial(P\, u)}{\partial x_N}, \varphi\Big\rangle = -\Big\langle P\, u, \frac{\partial\varphi}{\partial x_N}\Big\rangle = -\Big\langle u, Q\frac{\partial\varphi}{\partial x_N}\Big\rangle = -\Big\langle u, \frac{\partial(R\varphi)}{\partial x_N}\Big\rangle = \Big\langle \frac{\partial u}{\partial x_N}, R\varphi\Big\rangle,
\end{aligned}$$

and this proves that $P\, u$ belongs to $X(R^N)$, which is $L^2(R^N)$, and therefore its restriction u belongs to $L^2(R^N_+)$.

I will postpone the proof that $X(\Omega) = L^2(\Omega)$ for bounded domains with LIPSCHITZ boundary.

I had noticed in the Fall that it is not true for domains of the form $\{(x,y) : 0 < x < 1, 0 < y < x^2\}$, and therefore that one should not expect the pressure to belong to $L^2(\Omega)$ for such domains. François MURAT informed me during his recent visit that it had also been noticed by GEYMONAT and GILARDI that $X(\Omega) \neq L^2(\Omega)$ for a particular domain that they had constructed for showing that KORN's inequality may fail in non LIPSCHITZ domains.

If $X(\Omega) = L^2(\Omega)$ algebraically, then the norms are equivalent by the closed graph theorem as the injection of $L^2(\Omega)$ into $X(\Omega)$ is always continuous. One hypothesis of the Equivalence Lemma is that the injection of $L^2(\Omega)$ into $H^{-1}(\Omega)$ is compact, and by transposition it is equivalent to prove that the injection of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact:

Lemma: If $meas(\Omega) < \infty$, then the injection of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact.
Proof: We extend functions by 0 outside Ω; we have already noticed that POINCARÉ inequality holds for such domains, and we follow a similar proof using FOURIER transform. We consider a bounded sequence in $H_0^1(\Omega)$ and we want to show that one can extract a subsequence converging strongly in $L^2(\Omega)$. Because $L^2(\Omega)$ is separable, the weak topology is metrizable on bounded sets, and one can extract a weakly converging sequence in $L^2(\Omega)$, and by translation, one can assume then that the sequence u_n converges weakly to 0 in $L^2(\Omega)$, that it is bounded in $H_0^1(\Omega)$ and one wants to prove that it converges strongly to 0 in $L^2(\Omega)$.

One has $\mathcal{F}u_n(\xi) = \int_{R^N} u_n(x)e^{-2i\pi(x.\xi)}\,dx = \int_\Omega u_n(x)e^{-2i\pi(x.\xi)}\,dx \to 0$, as it is the $L^2(\Omega)$ scalar product of u_n with a fixed function, which belongs to $L^2(\Omega)$ by the hypothesis $meas(\Omega) < \infty$; on the other hand one has $|\mathcal{F}u_n(\xi)| \le C$, and by LEBESGUE dominated convergence theorem one has $\int_{|\xi|\le\rho} |\mathcal{F}u_n(\xi)|^2\,d\xi \to 0$ for any $\rho < \infty$. Because u_n is bounded in $H_0^1(\Omega)$, one has $\int_{R^N} |\xi|^2|\mathcal{F}u_n(\xi)|^2\,d\xi \le C$, and therefore $\int_{|\xi|\ge\rho} |\mathcal{F}u_n(\xi)|^2\,d\xi \le C/\rho^2$; one deduces that $\limsup \int_{R^N} |u_n - u_m|^2\,dx = \limsup \int_{R^N} |\mathcal{F}u_n - \mathcal{F}u_m|^2\,d\xi \le \limsup \int_{|\xi|\le\rho} |\mathcal{F}u_n - \mathcal{F}u_m|^2\,d\xi + \limsup \int_{|\xi|\ge\rho} |\mathcal{F}u_n - \mathcal{F}u_m|^2\,d\xi \le C/\rho^2$, and therefore $\limsup \int_{R^N} |u_n - u_m|^2\,dx = 0$, so that u_n is a CAUCHY sequence and converges strongly (to 0).

For a bounded open set Ω with LIPSCHITZ boundary, the conditions of the Equivalence Lemma are satisfied with $E_1 = L^2(\Omega)$, $A = grad$, $E_2 = H^{-1}(\Omega; R^N)$, B the injection of $L^2(\Omega)$ into $E_3 = H^{-1}(\Omega)$. The range of A is closed, and therefore equal to its closure, which is always the space of $T \in H^{-1}(\Omega; R^N)$ orthogonal to the kernel of A^T, i.e. $W = \{u \in H_0^1(\Omega; R^N) : div(u) = 0\}$, as $A^T = -div$. Of course, one gets as a corollary that the range of A^T is closed, and therefore equal to its closure, which is always the subspace of $f \in L^2(\Omega)$ with $\int_\Omega f\,dx = 0$, i.e. the orthogonal of the kernel of A, which are the constants.

The Equivalence Lemma says a little more: an equivalent norm on $L^2(\Omega)$ is $||grad(u)||_{H^{-1}(\Omega)} + p(u)$, where p is any semi-norm such that $p(1) \neq 0$, for example $p(u) = \left|\int_\Omega u\,dx\right|$. Therefore, on the subspace of $v \in L^2(\Omega)$ with $\int_\Omega v\,dx = 0$, one has $||v||_{L^2(\Omega)} \approx ||grad(v)||_{H^{-1}(\Omega)}$. An application of this result is that one can now prove that in the limit $\lambda \to \infty$, the sequence $-\lambda\, div(u^\lambda)$ converges in $L^2(\Omega)$ strong to the pressure, as its integral on Ω is 0 (because $u^\lambda \in H_0^1(\Omega; R^N)$) and its gradient converges strongly in $H^{-1}(\Omega; R^N)$ by the equation, as we have already proved that $u^\lambda \to u^\infty$ strongly in $H_0^1(\Omega; R^N)$.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

16. Wednesday February 17.

We have considered stationary STOKES equation by taking advantage of the similarity with the equation of stationary Linearized Elasticity, but the evolution equation are not similar, due to the fact that the unknown u in STOKES equation is a velocity, while for Elasticity it is a displacement. We have also used a coefficient μ, bounded below but variable with x, and that is not physical in general: in the Lagrangian point of view some parameters may depend upon the initial position ξ, but in an Eulerian point of view these properties are transported by the flow, and unlike for Linearized Elasticity where the Lagrangian and Eulerian point of views have been mixed, one mostly uses the Eulerian point of view for fluids.

There is at least one case where μ may reasonably depend upon x: POISEUILLE flows. If different fluids move in an infinite cylinder, there are particular solutions (of STOKES equation as well as of NAVIER-STOKES equation, because the nonlinearity vanishes for these solutions), where the velocities are all parallel to the axis of the cylinder, the gradient of the pressure is constant, parallel to the axis, and the velocity satisfies an equation $-div\big(\mu\, grad(u)\big) = C$ with $u \in H^1_0(\omega)$, where ω is the section of the cylinder. Each choice of μ as a function of x (bounded below and above) gives rise to a POISEUILLE flow, but not all these solutions are stable for the evolution problem; these questions have been investigated by Daniel JOSEPH and Michael RENARDY, for Newtonian fluids (water is used in pipelines carrying oil, and the stable configuration has the water near the boundary, serving as lubricant), or non-Newtonian fluids (for extrusion of molten polymers).

We will consider the viscosity μ and the density ρ_0 as constant (leaving for later the study of mixtures of fluids). As $\sigma_{ij} = 2\mu\varepsilon_{ij} - p\delta_{ij}$, one finds that $-\sum_j \frac{\partial\sigma_{ij}}{\partial x_j} = -\mu\Delta\, u_i + \frac{\partial p}{\partial x_i}$, and the stationary NAVIER-STOKES equations are

$$\begin{aligned}
&\rho_0 \sum_{j=1}^N u_j \frac{\partial u_i}{\partial x_j} - \mu\Delta\, u_i + \frac{\partial p}{\partial x_i} = f_i \text{ in } \Omega,\\
&div(u) = 0 \text{ in } \Omega,\\
&u \in H^1_0(\Omega; R^N).
\end{aligned}$$

As $div(u) = 0$, the nonlinear term $\sum_{j=1}^N u_j \frac{\partial u_i}{\partial x_j}$ can be written as $\sum_{j=1}^N \frac{\partial(u_j u_i)}{\partial x_j}$. Of course, apart from the gravity force which can be incorporated into the pressure, there are usually no body forces, and the fluid is usually put into motion because part of the boundary moves, but at this stage we will only homogeneous DIRICHLET conditions. It is the kinematic viscosity $\nu = \frac{\mu}{\rho_0}$ which really appears in the equation, and it must be noticed that although water is more viscous than air, the kinematic viscosities are in reverse order, 10^{-6} m^2/s for water, and 1.4×10^{-5} m^2/s for air (at atmospheric pressure). Contrary to the case of STOKES equation, which is linear and where the value of N does not matter much, the value of N is important for NAVIER-STOKES equation, because of the nonlinearity; one way the value of N arises is through SOBOLEV imbedding theorem.

SOBOLEV imbedding theorem: If $1 \le p < N$, then $W^{1,p}(R^N) \subset L^{p^*}(R^N)$ with $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}$.
$W^{1,1}(R) \subset C_0(R)$, the space of continuous (bounded) functions tending to 0 at infinity.
For $p = N \ge 2$, $W^{1,N}(R^N) \subset L^q(R^N)$ for all $q \in [N, \infty)$.
For $p > N$, $W^{1,p}(R^N) \subset C^{0,\alpha}(R^N)$ with $\alpha = 1 - \frac{N}{p}$.

The original proof of Sergei SOBOLEV started by using the formula $u = \sum_{i=1}^N \frac{\partial u}{\partial x_i} \star \frac{\partial E}{\partial x_i}$, where E is an elementary solution of the Laplacian; noticing that $\big|\frac{\partial E}{\partial x_i}\big| = O\big(\frac{1}{|x|^{N-1}}\big)$, he extended YOUNG's inequality for convolution to the case of convolution with $\frac{1}{|x|^\lambda}$, using nonincreasing radial rearrangements. His results have been slightly improved: for the case $p < N$ by Jaak PEETRE using the larger family of interpolation spaces known as the LORENTZ spaces, and for the case $p = N$ by Fritz JOHN and Louis NIRENBERG by their introduction of BMO (Bounded Mean Oscillation), and by Neil TRUDINGER. SOBOLEV's method only applies to cases where all the derivatives are in the same functional space, and in the case where the derivatives belong to the same LORENTZ space, the improvement are due to Jaak PEETRE for the case $p < N$ and to

Haïm BREZIS and Stephen WAINGER for the case $p = N$, using a formula by O'NEIL for the nonincreasing rearrangement of a convolution product.

A second method, which applies to cases where the derivatives are in different L^r spaces, has been developed by Emilio GAGLIARDO, and independently by Louis NIRENBERG, and maybe also by Olga LADYZHENSKAYA (there is another method which I have introduced which also applies to the case of derivatives in different LORENTZ spaces). I show this method on the example $H^1(R^3) \subset L^6(R^3)$, proving the estimates for $u \in C_c^\infty(R^3)$. One has $|u|^4(x) = \int_{-\infty}^{x_1} \big(4u^3 \frac{\partial u}{\partial x_1}\big)(t, x_2, x_3)\, dt = -\int_{x_1}^{+\infty} \big(4u^3 \frac{\partial u}{\partial x_1}\big)(t, x_2, x_3)\, dt$, and therefore

$$|u|^4(x) \le 2\int_{-\infty}^{+\infty} \big(|u|^3 \big|\frac{\partial u}{\partial x_1}\big|\big)(t, x_2, x_3)\, dt = F_1(x_2, x_3), \text{ and } \int F_1\, dx_2\, dx_3 \le 2||u||_{L^6}^3 \Big|\Big|\frac{\partial u}{\partial x_1}\Big|\Big|_{L^2}.$$

One has similar inequalities $|u|^4(x) \le F_2(x_1, x_3)$ and $|u|^4(x) \le F_3(x_1, x_2)$, and therefore

$$|u|^6(x) \le G_1(x_2, x_3) G_2(x_1, x_3) G_3(x_1, x_2), \text{ with } G_i = \sqrt{F_i},$$

which one integrates in x. Putting $H(x_1, x_2) = \int_R G_1(x_2, x_3) G_2(x_1, x_3)\, dx_3$ gives by CAUCHY-SCHWARZ inequality $H^2(x_1, x_2) \le \int_R G_1^2(x_2, x_3)\, dx_3 \int_R G_2(x_1, x_3)\, dx_3$, and then by integration $\int_{R^2} H^2(x_1, x_2)\, dx_1\, dx_2 \le \int_{R^2} G_1^2(x_2, x_3)\, dx_2\, dx_3 \int_{R^2} G_2(x_1, x_3)\, dx_1\, dx_3$, and finally by applying CAUCHY-SCHWARZ to $H\, G_3$ one obtains

$$\int_{R^3} |u|^6(x)\, dx \le ||G_1||_{L^2(R^2)} ||G_2||_{L^2(R^2)} ||G_3||_{L^2(R^2)},$$

but as $||G_1||_{L^2(R^2)}^2 = \int_{R^2} F_i\, d\widehat{x_i} \le 2\big|\big|\frac{\partial u}{\partial x_i}\big|\big|_{L^2} \big(\int_{R^3} |u|^6(x)\, dx\big)^{1/2}$, one deduces that

$$||u||_{L^6} \le 2\Big|\Big|\frac{\partial u}{\partial x_1}\Big|\Big|_{L^2}^{1/3} \Big|\Big|\frac{\partial u}{\partial x_2}\Big|\Big|_{L^2}^{1/3} \Big|\Big|\frac{\partial u}{\partial x_3}\Big|\Big|_{L^2}^{1/3}.$$

Similar inequalities in R^N are proved by using HÖLDER inequality, and showing by induction on N that if G_i is independent of x_i and belongs to L^{N-1} for its $N-1$ variables, then $G_1 \ldots G_N \in L^1(R^N)$ and $||G_1 \ldots G_N||_{L^1} \le ||G_1||_{L^{N-1}} \ldots ||G_N||_{L^{N-1}}$.

I will prove the other parts of the theorem, or generalizations, when it will become necessary.

A way to find solutions of the stationary NAVIER-STOKES equation is to use a fixed point argument. There are different ways to define the map for which one seeks a fixed point, and let us start by the following one

$$u = \Phi(v) \text{ is the solution } u \in W \text{ of } \rho_0 \sum_{j=1}^N \frac{\partial(v_j v_i)}{\partial x_j} - \mu \Delta\, u_i + \frac{\partial p}{\partial x_i} = f_i \text{ in } \Omega,$$

for $f \in H^{-1}(\Omega; R^N)$, the condition $div(u) = 0$ being included in the definition of the space W. A natural condition is to take $v \in L^4(\Omega; R^N)$ so that each term $\frac{\partial(v_j v_i)}{\partial x_j}$ belongs to $H^{-1}(\Omega)$, but as one finds $u \in H_0^1(\Omega; R^N)$, it is only for $N \le 4$ that one is sure to find $u \in L^4(\Omega; R^N)$. Using the norm $||grad(u)||_{L^2}$ on $H_0^1(\Omega)$ (so we assume that POINCARÉ inequality holds), the estimate for u becomes

$$\mu \sum_i ||grad(u_i)||_{L^2}^2 \le \sum_i ||f_i||_{H^{-1}} ||grad(u_i)||_{L^2} + \rho_0 \sum_{ij} ||v_j v_i||_{L^2} \Big|\Big|\frac{\partial u_i}{\partial x_j}\Big|\Big|_{L^2}.$$

A simpler inequality is obtained if one uses linearity, adding the bound for f when $v = 0$ and the bound for v when $f = 0$, i.e.

$$\mu\Big(\sum_i ||grad(u_i)||_{L^2}^2\Big)^{1/2} \le \Big(\sum_i ||f_i||_{H^{-1}}^2\Big)^{1/2} + \rho_0 \sum_j ||v_j||_{L^4}^2,$$

as $\sum_{ij} ||v_j v_i||_{L^2}^2 \le \sum_{ij} ||v_j||_{L^4}^2 ||v_i||_{L^4}^2 = \big(\sum_j ||v_j||_{L^4}^2\big)^2$. If

$$||v||_{L^4(\Omega)} \le \gamma ||grad(v)||_{L^2(\Omega)} \text{ for all } v \in H_0^1(\Omega),$$

then putting $X^2=\sum_j ||grad(v_i)||^2_{L^2}$ and $A^2=\sum_i ||f_i||^2_{H^{-1}}$, one obtains $\mu\big(\sum_i ||grad(u_i)||^2_{L^2}\big)^{1/2}\leq A+\rho_0\gamma^2 X^2$, and therefore in order to find a ball which is sent by Φ into itself it is enough to find $X_0>0$ such that $A+\rho_0\gamma^2X_0^2\leq \mu X_0$, and therefore

$$\text{if } \Big(\sum_i ||f_i||^2_{H^{-1}}\Big)^{1/2}\leq \frac{\mu^2}{4\rho_0\gamma^2} \text{ and } C=\Big\{v\in W: \Big(\sum_j ||grad(v_j)||^2_{L^2}\Big)^{1/2}\leq \frac{\mu}{2\rho_0\gamma^2}\Big\} \text{ then } \Phi \text{ maps } C \text{ into } C.$$

In order to check if Φ is a strict contraction on C (or just a contraction, which insures that Φ has at least one fixed point, not necessarily unique, because W is a HILBERT space, as shown below), one take $v'\in C$, and the estimates becomes

$$\mu\sum_i ||grad(u_i-u'_i)||^2_{L^2}\leq \rho_0\sum_{ij}||v_jv_i-v'_jv'_i||_{L^2}\Big|\Big|\frac{\partial(u_i-u'_i)}{\partial x_j}\Big|\Big|_{L^2},$$

so

$$\begin{aligned}\mu\Big(\sum_i ||grad(u_i-u'_i)||^2_{L^2}\Big)^{1/2} &\leq \rho_0\Big[\sum_{ij}||v_jv_i-v'_jv'_i||^2_{L^2}\Big]^{1/2}\\ &\leq \rho_0\Big[\sum_{ij}\Big(||v_j||_{L^4}||v_i-v'_i||_{L^4}+||v'_i||_{L^4}||v_j-v'_j||_{L^4}\Big)^2\Big]^{1/2}\\ &\leq \rho_0\Big[\sum_{ij}||v_j||^2_{L^4}||v_i-v'_i||^2_{L^4}\Big]^{1/2}+\rho_0\Big[\sum_{ij}||v'_i||^2_{L^4}||v_j-v'_j||^2_{L^4}\Big]^{1/2}\\ &\leq \frac{\mu}{\gamma}\Big(\sum_i ||v_i-v'_i||^2_{L^4}\Big)^{1/2},\end{aligned}$$

showing that Φ is a contraction on C; one gets a strict contraction by assuming that $\big(\sum_i ||f_i||^2_{H^{-1}}\big)^{1/2}\leq K<\frac{\mu^2}{4\rho_0\gamma^2}$, so that one can lower the bound in the definition of C.

A second method for looking for a fixed point is to use the function Ψ defined as

$$u=\Psi(v) \text{ is the solution } u\in W \text{ of } \rho_0\sum_{j=1}^N v_j\frac{\partial u_i}{\partial x_j}-\mu\Delta\, u_i+\frac{\partial p}{\partial x_i}=f_i \text{ in } \Omega,$$

which is a linear equation with variable coefficients; it can be treated by LAX-MILGRAM lemma for $N\leq 4$, the variational formulation corresponding to the bilinear form a_v defined by

$$a_v(u,\varphi)=\rho_0\int_\Omega\Big(\sum_{j=1}^N v_j\frac{\partial u_i}{\partial x_j}\varphi_i+\mu\sum_{ij}\frac{\partial u_i}{\partial x_j}\frac{\partial \varphi_i}{\partial x_j}\Big)\,dx,$$

and the important fact is that

$$a_v(u,u)=\mu\sum_i ||grad(u_i)||^2_{L^2},$$

if $v\in W$, as $div(v)$ appears by integration by parts $\sum_{j=1}^N v_j\frac{\partial u_i}{\partial x_j}u_i=\frac{1}{2}\sum_{j=1}^N\frac{\partial(v_ju_i^2)}{\partial x_j}$. If one defines the trilinear form

$$b(v,u,\varphi)=\sum_{ij}\int_\Omega v_j\frac{\partial u_i}{\partial x_j}\varphi_i\,dx \text{ for } v,\varphi\in L^4(\Omega;R^N), u\in H^1(\Omega;R^N),$$

then for $N\leq 4$ (in order to have $H^1_0(\Omega)\subset L^4(\Omega)$)

$$b(v,u,\varphi)+b(v,\varphi,u)=-\int_\Omega div(v)\Big(\sum_i u_i\varphi_i\Big)\,dx \text{ for } u,v,\varphi\in H^1_0(\Omega;R^N),$$

so that $b(v,u,u)=0$ for $u,v\in W$. One finds then a bound for u which is independent of v

$$\mu\Big(\sum_i ||grad(u_i)||_{L^2}^2\Big)^{1/2}\le\Big(\sum_i ||f_i||_{H^{-1}}^2\Big)^{1/2},$$

i.e. $\mu X\le A$ in the preceding notations. In order to check if Ψ is a strict contraction on the ball $X\le\frac{A}{\mu}$, one takes v' in this ball, and by subtracting $\mu\, a(u,\varphi)+\rho_0 b(v,u,\varphi)=L(\varphi)$ and $\mu\, a(u',\varphi)+\rho_0 b(v',u',\varphi)=L(\varphi)$ and taking $\varphi=u-u'$, one obtains

$$\begin{aligned}\mu\, a(u-u',u-u')&=\rho_0 b(v',u',u-u')-\rho_0 b(v,u,u-u')\\&=\rho_0 b(v'-v,u',u-u')-\rho_0 b(v,u-u',u-u')=-\rho_0 b(v'-v,u-u',u')\\&\le\rho_0\sum_{ij}||v'_j-v_j||_{L^4}\Big|\Big|\frac{\partial(u_i-u'_i)}{\partial x_j}\Big|\Big|_{L^2}||u'_i||_{L^4},\end{aligned}$$

and therefore

$$\begin{aligned}\mu\Big(\sum_i ||grad(u_i-u'_i)||_{L^2}^2\Big)^{1/2}&\le\rho_0\Big(\sum_{ij}||v'_j-v_j||_{L^4}^2||u'_i||_{L^4}^2\Big)^{1/2}\\&\le\frac{\rho_0\gamma A}{\mu}\Big(\sum_{ij}||v'_j-v_j||_{L^4}^2\Big)^{1/2}.\end{aligned}$$

One deduces that if $A<\frac{\mu^2}{\rho_0\gamma^2}$ there is a unique solution in W, as Ψ is a strict contraction on the ball $\big(\sum_i ||grad(u_i-u'_i)||_{L^2}^2\big)^{1/2}\le\frac{A}{\mu}$, which contains $\Psi(W)$.

Using SCHAUDER fixed point theorem, we will see that there exists a solution without constraint on A. I conclude with the proof for contractions alluded to before.

Lemma: Let $C\ne\emptyset$ be a closed bounded convex set of a HILBERT space H, and let T be a contraction from C into C; then T has at least one fixed point (the set of fixed points in C is a closed convex set). Let $c_0\in C$, and for $0\le\theta<1$ let $x(\theta)$ be the unique fixed point of $x\mapsto(1-\theta)c_0+\theta\,T(x)$, then as $\theta\to 1$, $x(\theta)$ converges strongly to $z(c_0)$ which is the fixed point of T in C which is the nearest from c_0.
Proof: As $x\mapsto(1-\theta)c_0+\theta\,T(x)$ maps C into C and is a strict contraction, it has a unique fixed point $x(\theta)$ (and $x(0)=c_0$). As $x(\theta)$ is bounded, one can extract a sequence $\theta_n\to 1$ such that $x(\theta_n)\rightharpoonup z$ in H weak (and $z\in C$, as closed convex sets are weakly closed). As $x(\theta_n)=(1-\theta_n)c_0+\theta_n\,T\big(x(\theta_n)\big)$ and C is bounded, one deduces that $x(\theta_n)-T\big(x(\theta_n)\big)\to 0$ strongly, and therefore $T\big(x(\theta_n)\big)\rightharpoonup z$. As $M(x)=x-T(x)$ is monotone continuous, this proves that $M(z)=0$, i.e. $T(z)=z$, as recalled at the end of the proof. If ξ is another fixed point of T in C, then $|T\big(x(\theta_n)\big)-\xi|^2=|T\big(x(\theta_n)\big)-T(\xi)|^2\le|x(\theta_n)-\xi|^2$, i.e. ξ is nearer to $T\big(x(\theta_n)\big)$ than to $x(\theta_n)$, but $x(\theta_n)$ being between c_0 and $T\big(x(\theta_n)\big)$, one deduces that $|c_0-x(\theta_n)|\le|c_0-\xi|$ and therefore $|c_0-z|\le|c_0-\xi|$, so that z is the nearest fixed point to c_0; this shows that all the sequence converges weakly to z, but also that the sequence converges strongly as $\limsup|c_0-x(\theta_n)|\le|c_0-z|$.

M is said to be monotone (from a topological vector space E into its dual E') if $\langle M(b)-M(a),b-a\rangle\ge 0$ for all a,b; here $E=E'=H$ and as $M(x)=x-T(x)$, one has $(M(b)-M(a).b-a)=|b-a|^2-(T(b)-T(a).b-a)\ge|b-a|^2-|T(b)-T(a)|\,|b-a|\ge 0$. M is said to be hemicontinuous if $t\mapsto\langle M(a+t\,c),c\rangle$ is continuous on R for every a,c; here M is LIPSCHITZ continuous. If M is monotone hemicontinuous, if $e_n\rightharpoonup f$ in E weak, $M(e_n)\rightharpoonup g$ in E' weak $\star$, and $\limsup\langle M(e_n),e_n\rangle\le\langle f,g\rangle$, then $M(f)=g$. Indeed, as $\limsup\langle M(e_n)-M(a),e_n-a\rangle\ge 0$, one deduces that $\langle g-M(a),f-a\rangle\ge 0$ for all a; taking $a=f-\varepsilon c$ with $\varepsilon>0$ gives $\langle g-M(f-\varepsilon c),c\rangle\ge 0$ and hemicontinuity gives $\langle g-M(f),c\rangle\ge 0$, so that varying c gives $g=M(f)$. A particular case is that $e_n\rightharpoonup f$ in E weak and $M(e_n)\to g$ in E' strong imply $M(f)=g$.
[In the theory of monotone operators, it is usual to call "MINTY's trick" the preceding argument! If a function is nonnegative and vanishes at a point, the idea that its derivative at this point must be 0 goes back at least to FERMAT in the 17$^{\text{th}}$ Century, and "MINTY's trick" is exactly that: the hypothesis of hemicontinuity is what is needed for the function $b\mapsto\langle M(b),b-a\rangle$ to be GÂTEAUX differentiable at a with derivative $M(a)$.]

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

17. Friday February 19.

For $N\leq 4$ one can show the existence of at least one solution of the stationary NAVIER-STOKES equation without assuming that the data f_i have a small norm in $H^{-1}(\Omega)$, by using SCHAUDER/TIKHONOV fixed point theorem for Ψ.

For $N\leq 4$, Ψ is continuous from $L^4(\Omega;R^N)$ into W, which is a subset of $L^4(\Omega;R^N)$, and Ψ maps W into a bounded set of W.

For $N\leq 3$, the injection from $H^1_0(\Omega)$ into $L^4(\Omega)$ is compact, as $H^1_0(\Omega)\subset L^6(\Omega)$ and the injection of $H^1_0(\Omega)$ into $L^2(\Omega)$ is compact (if $meas(\Omega)<\infty$), and therefore the injection of $H^1_0(\Omega)$ into $L^p(\Omega)$ is compact for $p<6$. Indeed if u^n is bounded in $L^6(\Omega)$ and converges strongly in $L^2(\Omega)$, then by HÖLDER inequality $||u^n-u^m||_{L^p}\leq ||u^n-u^m||_{L^2}^{1-\theta}||u^n-u^m||_{L^6}^{\theta}$, with $\frac{1}{p}=\frac{1-\theta}{2}+\frac{\theta}{6}$, and as $\theta<1$ and u^n is a CAUCHY sequence in $L^2(\Omega)$, one deduces that u^n is a CAUCHY sequence in $L^p(\Omega)$.

For $N=4$, the injection from $H^1_0(\Omega)$ into $L^4(\Omega)$ is not compact. More generally, for $\Omega\subset R^N$ and $p<N$, the injection of $W^{1,p}_0(\Omega)$ into $L^{p^*}(\Omega)$ is not compact. In order to show this, let φ be a nonzero function in $C^\infty_c(R^N)$, and for some $z\in\Omega$ let u^n be defined by $u_n(x)=n^{N/p^*}\varphi\big(n(x-z)\big)$, so that for n large enough $u^n\in C^\infty_c(\Omega)$. One checks easily that u^n is bounded in $L^{p^*}(\Omega)$ while $grad(u^n)$ is bounded in $L^p(\Omega)$ (because $1+N/p^*=N/p$), and therefore if the injection of $W^{1,p}_0(\Omega)$ into $L^{p^*}(\Omega)$ was compact, one could extract from u^n a subsequence converging strongly in $L^{p^*}(\Omega)$, and therefore $|u^n|^{p^*}$ would converge strongly in $L^1(\Omega)$, which is not the case, as $|u^n|^{p^*}$ converges in the sense of distributions (or weakly $\star$ in the sense of measures) to $A\,\delta_z$, with $A=\int_{R^N}|\varphi|^{p^*}\,dx$.

For $N\leq 3$, the closed ball of W containing $\Psi(W)$ is compact in $L^4(\Omega;R^N)$, and as Ψ is continuous from $L^4(\Omega;R^N)$ into W, the SCHAUDER fixed point theorem asserts that Ψ has at least one fixed point.

For $N=4$, the closed ball of W containing $\Psi(W)$ is not compact in $L^4(\Omega;R^4)$ if one uses the strong topology, but as a bounded closed convex set of W is compact if one uses the weak topology, the extension by TIKHONOV of SCHAUDER fixed point theorem to locally convex spaces asserts that Ψ has at least one fixed point when Ψ is (sequentially) weakly continuous from W into itself (the weak topology is not metrizable, but its restriction to a bounded set containing $\Psi(W)$ is metrizable). In order to prove the continuity, one takes a sequence v^n converging weakly to v^∞ in W, and as the corresponding solutions $u^n=\Psi(v^n)$ are bounded in W, one can extract a subsequence such that u^m converges weakly to u^∞ in W, and the problem is to show that $u^\infty=\Psi(v^\infty)$ (which ensures that all the sequence converges). For this purpose one uses the equation $\mu\,a(u^m,\varphi)+\rho_0 b(v^m,u^m,\varphi)=L(\varphi)$ for every $\varphi\in W$, and one notices that if one shows that $b(v^m,u^m,\varphi)\to b(v^\infty,u^\infty,\varphi)$ for all $\varphi\in W$, then one has proved that $u^\infty=\Psi(u^\infty)$. As $b(v^m,u^m,\varphi)=-b(v^m,\varphi,u^m)$, one sees that it is enough to show that $v^m_j u^m_i\rightharpoonup v^\infty_j u^\infty_i$ in $L^2(\Omega)$ weakly for all i,j. We know that v^m_j and u^m_i are bounded in $L^4(\Omega)$ and therefore $v^m_j u^m_i$ is bounded in $L^2(\Omega)$ and a subsequence converges to g in $L^2(\Omega)$ weak, but as $v^m_j\to v^\infty_j$ and $u^m_i\to u^\infty_i$ in $L^p(\Omega)$ strong for $2\leq p<4$, because the injection of $H^1_0(\Omega)$ into $L^p(\Omega)$ is compact for all $2\leq p<4$, one deduces that $v^m_j u^m_i\to v^\infty_j u^\infty_i$ in $L^{p/2}(\Omega)$ strong, and therefore $g=v^\infty_j u^\infty_i$.

I now turn to a method that will only use BROUWER fixed point theorem; it will also provide existence of weak solutions for $N>4$. It uses the (RITZ-) GALERKIN method.

A topological space is separable if it contains a countable dense subset; equivalently a metric space is separable if for every $\varepsilon>0$ it can be covered by a countable number of balls of radius at most ε (and therefore a subset of a separable space is separable). For any (nonempty) open set Ω of R^N, the spaces $L^p(\Omega)$ are separable for $1\leq p<\infty$ (but $L^\infty(\Omega)$ is not separable); for $1\leq p<\infty$, $W^{1,p}_0(\Omega)$ is separable; it suffices to approach functions in $C^\infty_c(\Omega)$ in the norm of $L^p(\Omega)$ or $W^{1,p}_0(\Omega)$ by a family of smooth functions which only depend upon a countable number of parameters: this is the first basic step when one wants to do the Numerical Analysis of solutions of partial differential equations, and there are traditional ways like finite elements to do it, but from a theoretical point of view one checks density by applying HAHN-BANACH theorem. If $\theta_n\in C^\infty_c(\Omega)$ is 1 when the distance to the boundary is more than $1/n$ (and such a θ_n can be obtained by convolution with a regularizing sequence applied to the characteristic function of the points where the distance is more than $1/2n$ for example), then one obtains a dense set by considering the family of all functions $\theta_n P$, for all n and all polynomials P (and one has a countable dense subset by taking only polynomials with rational coefficients). Indeed if f is in the dual of $L^p(\Omega)$ or of $W_0^{1,p}(\Omega)$ and is orthogonal to all $\theta_n P$, then each $\theta_n f$ (which is a distribution with compact support in Ω) is orthogonal to all polynomials, and therefore its FOURIER transform has all its derivatives at 0 equal to 0, but it must be 0 because the FOURIER transform of a distribution with compact support is an analytic function (which extends to C^N with at most exponential growth in the imaginary direction, and the theorem of PALEY-WIENER, extended by Laurent SCHWARTZ to distributions, characterizes these FOURIER transforms); this shows that $\theta_n f = 0$ for all n and therefore $f = 0$. A (RITZ-) GALERKIN basis of a separable topological vector space E is a countable family $e_1, \ldots, e_n, \ldots$ of linearly independent elements which generate a dense subspace of E.

Let $w_1, \ldots, w_n, \ldots$ be any (RITZ-) GALERKIN basis of W, and let W_n be the finite dimensional subspace generated by $w_1, \ldots, w_n$. One considers the approximate equation

$$\mu\, a(u^m, \varphi) + \rho_0 b(u^m, u^m, \varphi) = L(\varphi) \text{ for all } \varphi \in W_m;\ u^m \in W_m,$$

L being a linear continuous form on W and $N \leq 4$. The existence of a solution u^m follows from BROUWER fixed point applied to Ψ_m, where for $v \in W_m$, $u^m = \Psi_m(v)$ is the unique solution of

$$\mu\, a(u^m, \varphi) + \rho_0 b(v, u^m, \varphi) = L(\varphi) \text{ for all } \varphi \in W_m;\ u^m \in W_m.$$

The fact that u^m is defined in a unique way follows from LAX-MILGRAM lemma (and $b(v, u, u) = 0$ for all $u, v \in W$), and provides a bound $\mu||grad(u^m)||_{L^2} \leq C$ independent of v, and independent of m. Using the same method than for Ψ, one sees that the mapping Ψ_m is LIPSCHITZ continuous, and it must have at least one fixed point by BROUWER's theorem, as it maps all W_m into a bounded set of W_m.

As u^m is bounded in W, one can extract a subsequence $u^p \rightharpoonup u^\infty$ in W weak; as soon as $p \geq k$ one may take $\varphi = w_k$ and as before the critical point of the proof is the convergence $b(u^p, u^p, w_k) = -b(u^p, w_k, u^p) \to -b(u^\infty, w_k, u^\infty) = b(u^\infty, u^\infty, w_k)$, and therefore u^∞ satisfies the desired equation for $\varphi = w_k$, i.e. for φ in a dense subspace of W. Because $N \leq 4$ and $\varphi \mapsto b(u, u, \varphi)$ is linear continuous on W for $u \in W$, one obtains then the variational formulation for all $\varphi \in W$ and one has therefore found a solution of NAVIER-STOKES equation for $N \leq 4$ (of course, one then goes through the interpretation of the equation, involving the "pressure" in $L^2(\Omega)$, if Ω is smooth enough).

The preceding method actually gives the existence of a solution in the sense of distributions for $N > 4$. The first step is to take a (RITZ-) GALERKIN basis made of smooth functions, for example by proving that $\mathcal{W} = \{\varphi \in C_c^\infty(\Omega; R^N), div(\varphi) = 0\}$ is dense in W (which is true if Ω is bounded with a LIPSCHITZ boundary). Then $b(v, u, \varphi)$ is trilinear continuous on W_m, so LAX-MILGRAM lemma applies, defining Ψ_m, which by BROUWER's theorem has a fixed point u^m, and as u^m is bounded in W one extracts a subsequence u^p which converges weakly to u^∞, and one can pass to the limit in $b(u^\infty, u^\infty, w_k)$ because w_k is smooth. The interpretation of the variational formulation involves then a "pressure" in some $L^q(\Omega)$, with q depending on the dimension N.

I want to explain now what is behind BROUWER's fixed point theorem, i.e. BROUWER's topological degree theory (which goes back to around 1910, I believe); the extension in the 1930s by Jean LERAY and SCHAUDER to infinite dimension (and for functions of the form *identity+compact*) is rarely necessary, as one usually finds more precise results by using a (RITZ-) GALERKIN basis, applying the methods of topological degree in finite dimension and then letting the dimension tend to infinity.

The idea is that if Ω is a bounded open set of R^N and F is a continuous function from $\overline{\Omega}$ into R^N, then one can make an algebraic count of the number of solutions of $F(x) = p$ which are in Ω, by only looking at the restriction of F on the boundary $\partial\Omega$, assuming that there is no solution of $F(x) = p$ on $\partial\Omega$. This will extend the trivial case of an interval (a, b) in dimension 1, where the count is 1 if $F(a) < p < F(b)$, -1 if $F(a) > p > F(b)$, and 0 otherwise; it will also extend the case of a holomorphic function F for a domain in C bounded by a smooth JORDAN curve Γ, where the number is $\frac{1}{2i\pi}\int_\Gamma \frac{F'(z)}{F(z)-p}\,dz$ (always a nonnegative integer).

The degree $deg(F, \Omega, p)$ is defined for any continuous function F from $\overline{\Omega}$ into R^N which does not take the value p on $\partial\Omega$, and it depends only upon the restriction of F to $\partial\Omega$. It is an integer and if $deg(F, \Omega, p) \neq 0$

then there exists $x \in \Omega$ with $F(x) = p$. If the degree is defined for Ω_1 and for Ω_2, and $\Omega_1 \cup \Omega_2 \subset \Omega \subset \overline{\Omega_1} \cup \overline{\Omega_2}$, then $deg(F, \Omega, p) = deg(F, \Omega_1, p) + deg(F, \Omega_2, p) - deg(F, \Omega_1 \cap \Omega_2, p)$. The degree is invariant by homotopy, i.e. $deg(F, \Omega, p) = deg(G, \Omega, p)$ if there exists a homotopy between F and G, i.e. there exists a continuous function H from $\overline{\Omega} \times [0,1]$ into R^N, such that H does not take the value p on $\partial\Omega \times [0,1]$, and such that $H(x, 0) = F(x)$ and $H(x, 1) = G(x)$ on $\overline{\Omega}$. In the case where F is continuous from $\overline{\Omega}$ into R^N, of class C^1 in Ω, and only takes the value p at a finite number of points $a_j, j = 1, \ldots, r$, of Ω, and if the Jacobian determinant of F is nonzero at each of these points, then $deg(F, \Omega, p) = \sum_{j=1}^r sign\big(det(\nabla F)(a_j)\big)$. For instance, if $F(x) = x$ on $\partial\Omega$ then $deg(F, \Omega, p) = 1$ if $p \in \Omega$, 0 if $p \notin \overline{\Omega}$ (and is not defined if $p \in \partial\Omega$); if $F(x) = -x$ on $\partial\Omega$ then $deg(F, \Omega, p) = (-1)^N$ if $-p \in \Omega$, 0 if $-p \notin \overline{\Omega}$ (and is not defined if $-p \in \partial\Omega$).

A first application is that there exists no nonzero continuous vector field from S^2 into R^3 which is everywhere tangent; more generally every nonzero continuous vector field from S^{2N} into R^{2N+1} is normal at (at least) one point of S^{2N}. Indeed suppose that F is a nonzero continuous vector field from S^{2N} into R^{2N+1} which is nowhere normal; then the homotopy H defined by $H(x, t) = (1-t)F(x) + t\,x$ is not 0 on S^{2N} so $deg(F, B(0,1), 0) = deg(id, B(0,1), 0) = 1$, and similarly the homotopy K defined by $K(x, t) = (1-t)F(x) - t\,x$ is not 0 on S^{2N} so $deg(F, B(0,1), 0) = deg(-id, B(0,1), 0) = -1$, a contradiction.

A second application is that there does not exist a continuous retraction from a bounded open set $\Omega \subset R^N$ onto its boundary $\partial\Omega$, i.e. a continuous function F from $\overline{\Omega}$ into $\partial\Omega$ such that $F(x) = x$ on $\partial\Omega$. Indeed for $p \in \Omega$ one has $deg(F, \Omega, p) = deg(id, \Omega, p) = 1$ and therefore there exists $x \in \Omega$ with $F(x) = p$, contradicting the fact that the range of F is inside $\partial\Omega$. A consequence is BROUWER fixed point theorem: every continuous mapping Φ from the closed unit ball of R^N into itself has at least one fixed point. Let $\Omega = B(0,1)$, and assume that Φ has no fixed point in $\overline{\Omega}$; then for every $x \in \overline{\Omega}$ the line joining x to $\Phi(x)$ is well defined and intersects $\partial\Omega$ in two points, and if one takes $F(x)$ to be that point on the side of x one sees that F is a continuous retraction from Ω onto its boundary. Analytically, $F(x) = (1-t)x + t\Phi(x)$ with $t \leq 0$ and $|(1-t)x + t\,\Phi(x)| = 1$, i.e. $t^2|x - \Phi(x)|^2 + 2t(x - \Phi(x).x) + |x|^2 - 1 = 0$. The theorem extends to any nonempty compact convex set $C \subset R^N$: one first restricts attention to the affine subspace generated by C, so that C has a nonempty interior in that subspace, and one notices that a compact convex set with nonempty interior in R^M is homeomorphic to the closed unit ball in R^M (taking 0 inside C, and using the MINKOWSKI functional p_C, the mapping $c \mapsto \frac{p_C(c)}{||c||}c$ is a homeomorphism of C onto the closed unit ball).

The result does not extend in infinite dimension to the closed unit ball of the HILBERT space l^2: one defines Φ by $\Phi(x) = (\sqrt{1-|x|^2}, x_1, \ldots)$ for $x = (x_1, \ldots, x_n, \ldots)$; then Φ is LIPSCHITZ continuous, maps the closed unit ball into its boundary and has no fixed point; one can deduce that there exists a continuous retraction of the unit ball onto its boundary.

18. Monday February 22.

I had read about topological degree in some lecture notes of the COURANT Institute by J. T. SCHWARTZ "Nonlinear functional Analysis"; I did not find these notes very clear, and in 1974 I had simplified the exposition of the essential results by the following approach, based on the study of functionals $J_\varphi(u) = \int_\Omega \varphi(u)det(\nabla u)\,dx$. [In the Fall of 1975 I learned from John M. BALL about null Lagrangians and the property that Jacobian determinants are sequentially weakly continuous and I immediately linked this kind of robustness to that encountered in the study of topological degree.]

Although topological degree will be defined for some continuous functions, the integrals that we start with assume that the functions are of class C^1, and even to have two derivatives in some proofs; a density argument is then necessary for extending to continuous functions the results obtained.

Let Ω be a bounded regular open set of R^N, i.e. whose boundary is given locally by a LIPSCHITZ function so that the exterior normal n to $\partial\Omega$ is defined almost everywhere on $\partial\Omega$ and the formula of integration by parts is valid (and uses the measure $d\sigma$ on $\partial\Omega$). All our functions are assumed to be continuously extended to $\partial\Omega$, so that they will be defined on $\overline{\Omega}$.

Let u be a C^1 function from $\overline{\Omega}$ into R^N and φ a continuous scalar function on R^N; we define the functional J_φ by the formula

$$J_\varphi(u) = \int_\Omega \varphi(u)det(\nabla u)\,dx,$$

where $\nabla u(x)$ is the Jacobian matrix at the point x, whose entries are the partial derivatives $\frac{\partial u_i}{\partial x_j}$. Remark that the definition makes sense for functions u in the SOBOLEV space $W^{1,N}(\Omega; R^N)$ and φ bounded; Louis NIRENBERG and Haïm BREZIS have recently extended topological degree to functions in VMO (Vanishing Mean Oscillation, the closure of C^∞ functions in BMO). The crucial property of the functional J_φ is the following

Main Lemma: Assume that u is C^2 from $\overline{\Omega}$ into R^N and that φ is C^1 from R^N into R; let v be a C^1 function from $\overline{\Omega}$ into R^N, then

$$\frac{d\Big(J_\varphi(u+\varepsilon\, v)\Big)}{d\varepsilon}\Big|_{\varepsilon=0} = \int_{\partial\Omega} \varphi(u)\Big(\sum_{k=1}^N \psi_k n_k\Big)\,d\sigma,$$

where ψ_k is defined by

$$\psi_k = det\Big(\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_{k-1}}, v, \frac{\partial u}{\partial x_{k+1}},\ldots,\frac{\partial u}{\partial x_N}\Big),$$

i.e. $\psi_k(x)$ is obtained from the Jacobian matrix $\nabla u(x)$ by replacing the k^{th} column by the vector $v(x)$.

A consequence is that if u and w are C^1 functions from $\overline{\Omega}$ into R^N which are equal on the boundary $\partial\Omega$ and if φ is continuous, then $J_\varphi(u) = J_\varphi(w)$. Indeed, by an argument of density, it is enough to prove the corollary when u and w are of class C^2 and φ is of class C^1. The lemma is used for computing the derivative of $J_\varphi\big((1-\theta)u+\theta\, w\big)$ with respect to θ and it says that the derivative is equal to an integral on $\partial\Omega$, and this integral is 0 as the functions ψ_k vanish on $\partial\Omega$ because v is $w-u$, which is assumed to be 0 on the boundary.

More generally one can change the values of u on the boundary without changing the value of $J_\varphi(u)$ if one avoids the support of φ and this gives the following property called invariance by homotopy.

Lemma: Assume that u and w are C^1 functions from $\overline{\Omega}$ into R^N and $\varphi \in C_c(R^N)$. We assume that u and w can be joined by a homotopy having the property that on the boundary $\partial\Omega$ it avoids the support of φ, then $J_\varphi(u) = J_\varphi(w)$. (The hypothesis means that there exists a continuous function F defined on $\overline{\Omega}\times[0,1]$ with values in R^N such that $F(\cdot,0) = u$, $F(\cdot,1) = w$ on $\overline{\Omega}$ and $F(x,\theta) \notin support(\varphi)$ for $(x,\theta) \in \partial\Omega\times[0,1]$).
Proof: Indeed one can regularize u, w, φ and F and still satisfy the same conditions; then one considers $G(\theta) = J_\varphi\big(F(\cdot,\theta)\big)$ and the lemma applies with u replaced by $F(\cdot,\theta)$ and v by $\frac{\partial F}{\partial\theta}$: it says that $G'(\theta)$ is

equal to an integral on the boundary and the integrand is 0 because it contains a term $\varphi\big(F(\cdot,\theta)\big)$ which is 0 on the boundary, and therefore $G(0)=G(1)$ which is our assertion.

The invariance by homotopy enables us to defines $J_\varphi(u)$ when u is only continuous: if u is a continuous function from $\overline{\Omega}$ into R^N satisfying the condition $u(x)\notin support(\varphi)$ when $x\in\partial\Omega$, then one can define $J_\varphi(u)$ by taking any sequence v_n of C^1 functions from $\overline{\Omega}$ into R^N which converges uniformly to u, because for n large $J_\varphi(v_n)$ is constant and $J_\varphi(u)$ is defined as this limiting value. Indeed let $\varepsilon>0$ be small enough so that for $x\in\partial\Omega$ the distance of $u(x)$ to the support of φ is at least 2ε; if 2 functions v_n and v_m of class C^1 are in the ball of center u and radius ε in the C^0 distance, then they can be joined by the homotopy $(1-\theta)v_n+\theta v_m$ and then $J_\varphi(v_n)=J_\varphi(v_m)$, so $J_\varphi(v)$ is constant in a ball around u.

With this extension of the definition to some continuous functions, we can see that the preceding results are true for continuous functions.

Lemma: If u is a continuous function from $\overline{\Omega}$ into R^N such that $J_\varphi(u)\neq 0$ (the condition $u(x)\notin support(\varphi)$ for $x\in\partial\Omega$ being satisfied in order to define $J_\varphi(u)$), then there exists $x\in\Omega$ such that $u(x)\in support(\varphi)$.

Proof: As $J_\varphi(v_n)\neq 0$ for large n, $\varphi(v_n)$ cannot vanish identically and so there exists $x_n\in\Omega$ such that $v_n(x_n)\in support(\varphi)$; every limit point $x\in\overline{\Omega}$ of the sequence x_n is then such that $u(x)\in support(\varphi)$ and $x\notin\partial\Omega$ by hypothesis.

Proof of Main Lemma: The derivative in ε that we are considering is

$$\frac{d\Big(J_\varphi(u+\varepsilon\, v)\Big)}{d\varepsilon}\Big|_{\varepsilon=0}=\int_\Omega\Big[\sum_{i=1}^N\frac{\partial\Big(\varphi(u)\Big)}{\partial u_i}v_i\,det(\nabla\, u)+\varphi(u)\sum_{k=1}^N H_k\Big]\,dx,$$

where the functions H_k are

$$H_k=det\Big(\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_{k-1}},\frac{\partial v}{\partial x_k},\frac{\partial u}{\partial x_{k+1}},\ldots,\frac{\partial u}{\partial x_N}\Big),$$

expressing the multilinearity of the determinant. The Main Lemma will be proved by integration by parts if we show that

$$\sum_{i=1}^N\frac{\partial\Big(\varphi(u)\Big)}{\partial u_i}v_i\,det(\nabla\, u)+\varphi(u)\sum_{k=1}^N H_k=\sum_{k=1}^N\frac{\partial\Big(\varphi(u)\psi_k\Big)}{\partial x_k},$$

and this will follow from the following two identities

$$\sum_{k=1}^N\frac{\partial\psi_k}{\partial x_k}=\sum_{k=1}^N H_k.$$

and

$$\sum_{i=1}^N\frac{\partial u_i}{\partial x_k}\,\psi_k=v_i\,det(\nabla\, u),$$

The first identity requires u to be of class C^2; once again the multilinearity of the determinant is used and we must show that the sum of the terms containing second derivatives of u is 0. Here it is the antisymmetry of the determinant that is needed because there are two terms showing a given second derivative $\frac{\partial^2 u}{\partial x_i\partial x_k}$: one has it in column i with v in column k and the other has it in column k with v in column i, all the other columns being similar. The second identity is linear in v, so we check it in the case where only one component of v, say v_1, is 1 and the others are 0; then the left hand side consists in developping with respect to the first row a determinant obtained from $\nabla\, u$ by replacing the first row by $grad(u_i)$ so it gives $det(\nabla\, u)$ if $i=1$ and, again from antisymmetry, it gives 0 if $i\neq 1$ because two rows are identical.

The usual definition of the topological degree will consist in computing an algebraic number of solutions of $u(x)=p$ for a point $p\in R^N$ and it is obtained by letting the function φ approach the DIRAC mass at the point p. We are led to the following definition.

Definition: If u is a continuous function from $\overline{\Omega}$ into R^N satisfying the condition $u(x) \neq p$ when $x \in \partial\Omega$ then one can define $deg(u, \Omega, p)$ as the limit of the values $J_{\varphi_n}(u)$ for a sequence of functions φ_n whose supports converge to the point p and whose integrals converge to 1.

Obviously one has $u(x) \notin support(\varphi_n)$ for $x \in \partial\Omega$ for large n so that $J_{\varphi_n}(u)$ has a meaning, but one difficulty is to show that this limit exists; the proof actually gives an important property, namely that the degree $deg(u, \Omega, p)$ is always an integer, for the values p for which this degree is defined, i.e. $p \notin u(\partial\Omega)$. The first step is to notice that there is a discrete formula for computing the degree in the case where u is of class C^1 under a slight restriction.

Lemma: Let u be of class C^1 from $\overline{\Omega}$ into R^N such that $\nabla u(z)$ is invertible at every point z solution of $u(z) = p$, none of these solutions being on the boundary, so that there is only a finite number of them; then $deg(u, \Omega, p)$ is an integer, the sum of the signs of the Jacobian at all these points

$$deg(u, \Omega, p) = \sum_{u(z_\alpha)=p} sign\Big(det\Big(\nabla u(z_\alpha)\Big)\Big).$$

Proof: For n large enough φ_n is 0 except in small disjoint neighborhoods of the z_α solutions of $u(z_\alpha) = p$; aroung each z_α one can use a change of variable in the integral by taking $y = u(x)$ as the new variable: one then obtains a contribution $sign\big(det\big(\nabla u(z_\alpha)\big)\big) \int \varphi_n(y)\, dy$ and this gives our formula in the limit $n \to \infty$.

The second step is to notice that every C^1 function from Ω into R^N (and thus every continuous function from Ω into R^N) can be approximated by such special functions u; this is done by adding a small constant vector to u and using SARD's lemma which states that the set of critical values p such that $\nabla u(x)$ is not invertible at some solution of $u(x) = p$ has measure 0 and so has an empty interior.

The third step is to notice that one can extend all the properties of the functionals J_φ to the topological degree $deg(u, \Omega, p)$ and in particular the invariance by homotopy. Another easy consequence is that the degree is continuous in p and, because it is an integer, it is locally constant in each connected component of the complement of $u(\partial\Omega)$ (that one needs to avoid in order to give a meaning to the definition).

It is useful to notice that in practical situations one computes the degree by using a homotopy to a simple C^1 function for which one can obtain the degree explicitly using the formula, and so SARD's lemma is only a technical tool used to show that the degree is defined for every continuous function (the proof of SARD's lemma in the case of spaces of same dimension, which is the one that interests us here, is relatively easy: one covers the initial open set with small cubes and notices that around one critical point the image of the corresponding cube is inside a flat cylinder of much smaller volume; using the uniform continuity of derivatives one concludes that the set of critical values is covered by sets of arbitrarily small volume, and so has measure 0).

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

19. Wednesday February 24.

Let us look now at the evolution equations, first without nonlinearity for the STOKES equation, then with the nonlinearity for the NAVIER-STOKES equation.

I switch now to more traditional notations and denote

$$V = \{u \in H_0^1(\Omega; R^N), div(u) = 0 \text{ in } \Omega\},$$

previously denoted W, and

$$H = \{u \in L^2(\Omega; R^N), div(u) = 0 \text{ in } \Omega, u.n = 0 \text{ on } \partial\Omega\},$$

for which we will have to show that $u.n$ makes sense on the boundary; we will also have to prove that V is dense in H. Before doing so, we start with some abstract results on evolution equations, where the spaces denoted V or H do not necessarily mean those above (which are adapted to the treatment of STOKES equation).

There is an abstract theory for linear evolution equations, the theory of semi-groups, which was develöpped independently in the 40s by Kôsaku YOSIDA in Japan, and by HILLE and then Ralph PHILIPPS in United States, but the theory has proved difficult to generalize to nonlinear equations, apart from situations where the maximum principle plays a role, which is usually not the case for equations of Continuum Mechanics. One advantage of the theory is that it puts into the same framework lots of linear evolution equations with coefficients independent of t, but that is also a defect as it does not take into account the particular properties that the equations may have: transport equations $\frac{\partial u}{\partial t} + \sum_j a_j \frac{\partial u}{\partial x_j} = 0$, diffusion equations like the heat equation $\frac{\partial u}{\partial t} - \sum_{ij} \frac{\partial}{\partial x_i}\big(a_{ij} \frac{\partial u}{\partial x_j}\big) = 0$, SCHRÖDINGER equations $i\frac{\partial u}{\partial t} - \Delta\, u + V\, u = 0$, wave equations $\rho \frac{\partial^2 u}{\partial t^2} - \sum_{ij} \frac{\partial}{\partial x_i}\big(a_{ij} \frac{\partial u}{\partial x_j}\big) = 0$, the systems of linearized Elasticity, MAXWELL, or STOKES, can all be considered in such an abstract framework. The framework uses one BANACH space, which of course changes from an equation to the other, and it is sometimes an important restriction, because a good understanding of some equations from Continuum Mechanics often requires the use of more than one functional space: for STOKES equation, the bound on the kinetic energy corresponds to a bound in $L^\infty(0,T;H)$, while the bound on the energy dissipated by viscosity corresponds to a bound in $L^2(0,T;V)$. Despite these shortcomings, I quickly sketch the main ideas of the semi-group approach.

For an abstract evolution equation $\frac{du}{dt} + A\, u = 0$, where A is a partial differential operator, one cannot define $e^{-t\,A}$ by the usual series $\sum_{k=0}^\infty \frac{(-t)^k}{k!} A^k$, as in the case where $A \in \mathcal{L}(E,E)$ for a BANACH space E, and write then the solution as $u(t) = e^{-t\,A}u(0)$. Nevertheless if one finds a way to define the solution in a unique way, one may expect that the mapping $u(0) \mapsto u(t)$ defines an operator $S(t) \in \mathcal{L}(E,E)$ for $t \geq 0$, satisfying $S(0) = I$ and $S(t_1)S(t_2) = S(t_1 + t_2)$ (the semi-group property), and some sort of continuity in t for $S(t)e$ for each $e \in E$, for example $S(t)e \to e$ in E strong as $t \to 0$. Given such a (strongly continuous) semi-group, the uniform boundedness principle implies that $||S(t)||$ is bounded for $t \in [0,1]$, and then that $||S(t)|| \leq M\, e^{\omega\, t}$ for $t \geq 0$; putting $u(t) = v(t)\, e^{\omega\, t}$ creates a bounded semigroup $S_1(t) = S(t)\, e^{-\omega\, t}$ for v, satisfying $||S_1(t)|| \leq M$ for all $t \geq 0$; in the equivalent norm $||e||_1 = \sup_{t\geq 0} ||S_1(t)e||$, S_1 becomes a semi-group of contractions.

The domain $D(A)$ of the infinitesimal generator A of a (strongly continuous) semi-group S is defined as the subspace of elements $e \in E$ for which $S(t)e$ has a derivative at $t = 0$, denoted $-A\, e$; one deduces that if $e \in D(A)$ then $S(t)e \in D(A)$ and its derivative is $-A\, S(t)e$, so that $S(t)$ does play the role of $e^{-t\,A}$. One shows then that $D(A)$ is dense in E, and that A is closed. If $S(t)$ is a semi-group of contraction, one shows then that $I + \lambda\, A$ is invertible for $\lambda \geq 0$ with $||(I + \lambda\, A)^{-1}|| \leq 1$.

Conversely, if a closed operator A with dense domain is such that $I + \lambda\, A$ is invertible for $\lambda \geq 0$ with $||(I + \lambda\, A)^{-1}|| \leq 1$, then one can construct a semi-group S of contractions, of which A is the infinitesimal generator. Without going into the details (what I am sketching is a simplified view of the HILLE-YOSIDA theorem), the idea is to consider the implicit approximation scheme $\frac{u_{n+1} - u_n}{\Delta\, t} + A\, u_{n+1} = 0$, where u_n serves as an approximation of $u(n\,\Delta\,t)$, and as $u_{n+1} = (I + \Delta\,t\,A)^{-1}u_n$, the way to use the bounds $||(I+\lambda\,A)^{-1}|| \le 1$ for $\lambda \ge 0$ appears easily (the explicit scheme $\frac{u_{n+1}-u_n}{\Delta\,t} + A\,u_n = 0$ requires $u_n \in D(A)$, and therefore one needs $u(0) \in D(A^k)$ for all k just for defining all the u_n, so this scheme is not of great use).

I will present now a different framework, where two HILBERT spaces V and H are used (or three if one count V', H' being identified to its dual); this framework is adapted to solving diffusion equations, or STOKES equation, for example (in semi-group theory it is related to analytic semi-groups, which can be extended for t in a sector of the complex plane. I have learned many of the results that I present from Jacques-Louis LIONS, and I have only improved small technical details.

Let V and H be two (real) HILBERT spaces, with norms $||\cdot||$ for V and $|\cdot|$ for H, V being continuously imbedded in H and being dense in H; H is identified to its dual H', which is continuously imbedded in V' and dense in V' (in some cases the identification of H to its dual H' may create a few problems, and I will consider that question later). Let $A \in \mathcal{L}(V, V')$ be such that there exists $\alpha > 0$ and $\beta \in R$ for which

$$\langle A\,u, u\rangle \ge \alpha||u||^2 - \beta|u|^2 \text{ for all } u \in V,$$

(for simplification, I assume that A is independent of t; in practical situations one may have a bilinear continuous form $a(t, u, v)$ measurable in t).

Lemma: Given $u_0 \in H$, $f_1 \in L^1(0,T;H)$ and $f_2 \in L^2(0,T;V')$, there exists a unique $u \in C([0,T];H) \cap L^2(0,T;V)$ with $\frac{du}{dt} \in L^1(0,T;H) + L^2(0,T;V')$, solution of

$$\frac{du}{dt} + A\,u = f_1 + f_2 \text{ in } (0,T);\ u(0) = u_0,$$

which in variational form means

$$\int_0^T \Big(-\frac{d\varphi}{dt}(u.v) + \varphi\langle A\,u, v\rangle\Big)\,dt = \varphi(0)(u_0, v) + \int_0^T \varphi\Big((f_1.v) + \langle f_2, v\rangle\Big)dt$$
for all $v \in V$ and all $\varphi \in C^\infty([0,T])$ satisfying $\varphi(T) = 0$.

Jacques-Louis LIONS always considered $f \in L^2(0,T;V')$, i.e. the case $f_1 = 0$, which gives $\frac{du}{dt} \in L^2(0,T;V')$; because of the natural bounds $u \in L^\infty(0,T;H) \cap L^2(0,T;V)$, I find natural to take $f \in L^1(0,T;H) + L^2(0,T;V')$.

I will assume that V is separable (and then H is separable as V is dense in H); this is not a restriction for applications, and it avoids some tecnical difficulties about measurability of functions with values in V, H or V'. Let $e_1, \ldots$ be a any (RITZ-) GALERKIN basis of V, and let V_m be the subspace generated by $e_1, \ldots, e_m$. One looks for a function u_m from $[0,T]$ into V_m, i.e. $u_m(t) = \sum_{i=1}^m \xi_{mi}(t)e_i$, and the coefficients ξ_{mi} will belong to $W^{1,1}(0,T)$, which is continuously imbedded in $C([0,T])$; one asks u_m to satisfy

$$\Big(\frac{du_m}{dt}.e_k\Big) + \langle A\,u_m, e_k\rangle = (f_1.e_k) + \langle f_2, e_k\rangle \text{ a.e. in } (0,T) \text{ and } (u_m(0).e_k) = (u_0.e_k) \text{ for } k = 1, \ldots, m.$$

This is an ordinary linear differential equation in R^m, of the form $\xi' + A_m\xi = \eta_m$ in $(0,T)$ and $\xi(0) = \xi_{0m}$, with $\xi_{0m} \in R^m$ and $\eta_m \in L^1(0,T;R^m)$; it has a unique solution in $W^{1,1}(0,T;R^m)$, which is given explicitly by $\xi(t) = e^{-t\,A_m}\xi_{0m} + \int_0^t e^{-(t-s)\,A_m}\eta_m(s)\,ds$ for $t \in [0,T]$; one may prefer to deal with classical C^1 solutions in V_m, and that consists in choosing $u_{0m} \in V_m$, $f_{1m}, f_{2m} \in C([0,T];V_m)$ approaching in a strong or weak way u_0 in H, f_1 in $L^1(0,T;H)$ and f_2 in $L^2(0,T;V)$. Because the equation is linear, we immediately know existence and uniqueness on the whole interval $[0,T]$, but when we will deal with a nonlinear equation like NAVIER-STOKES equation, we will have to start with a local existence result and then we will show that the solution exists on $[0,T]$.

We need now precise bounds (independent of m) in order to take the limit $m \to \infty$, and we will need some technical results.

Lemma: i) $W^{1,1}(0,T) \subset C([0,T])$,

ii) $u, v \in W^{1,1}(0,T)$ imply $u\,v \in W^{1,1}(0,T)$ and $(u\,v)' = u\,v' + u'v$ a.e. in $(0,T)$,

iii) GRONWALL inequality: if $\varphi \in L^\infty(0,T)$ satisfies $\varphi(t) \geq 0$ a.e. on $(0,T)$ and $\varphi(t) \leq \psi(t) = A + \int_0^t (\lambda_1\varphi + \lambda_2)\,ds$ a.e. in $(0,T)$, where $\lambda_1, \lambda_2 \in L^1(0,T)$, then $\psi(t) \leq \big(A + \int_0^t |\lambda_2(s)|\,ds\big)exp\big(\int_0^t |\lambda_1(s)|\,ds\big)$ for $t \in [0,T]$.

Proof: For $f \in L^1(0,T)$, let $f_n \in C([0,T])$ converge to f in $L^1(0,T)$, then $u_n \in C^1([0,T])$ defined by $u_n(t) = \int_0^t f_n(s)\,ds$ converges uniformly to u defined by $u(t) = \int_0^t f(s)\,ds$, and as u_n' converges in the sense of distributions to u', one has $u' = f$ a.e. in $(0,T)$; if $v \in W^{1,1}(0,T)$ has $v' = f$, then $v - u$ has derivative 0 and is therefore a constant, showing that v is continuous and that the constant is $v(0)$.

The proof of i) has shown that $C^1([0,T])$ is dense in $W^{1,1}(0,T)$. If $u_n \in C^1([0,T])$ with $u_n' \to u'$ in $L^1(0,T)$, and $v_n \in C^1([0,T])$ with $v_n' \to v'$ in $L^1(0,T)$, then $u_n v_n$ converges uniformly to $u\,v$ and $(u_n v_n)' = u_n' v_n + u_n v_n'$ converges in $L^1(0,T)$ to $u'v + u\,v'$, and to $(u\,v)'$ in the sense of distributions, and therefore $(u\,v)' = u'v + u\,v'$.

As $\varphi \geq 0$, the inequality stays true if one replaces λ_1 by its absolute value, and one may then assume that $\lambda_1 \geq 0$ a.e. on $(0,T)$. Then $\psi' = \lambda_1\varphi + \lambda_2 \leq \lambda_1\psi + \lambda_2$ a.e. on $(0,T)$. As for the proof of i) and ii), if one defines E by $E(t) = exp\big(-\int_0^t \lambda_1(s)\,ds\big)$ then $E \in W^{1,1}(0,T)$ and $E' = -E\,\lambda_1$. One deduces that $(E\,\psi)' \leq E\,\lambda_2$, so that $\psi(t) \leq A\,exp\big(\int_0^t \lambda_1(s)\,ds\big) + \int_0^t \lambda_2(s) exp\big(\int_s^t \lambda_1(\sigma)\,d\sigma\big)\,ds$, from which the bound follows.

For obtaining estimates on u_m, one replaces e_k bu u_m, which is a linear combination of the e_k, and one obtains

$$\frac{1}{2}\frac{d|u_m|^2}{dt} + \alpha||u_m||^2 - \beta|u_m|^2 \leq |f_{1m}|\,|u_m| + ||f_{2m}||_*||u_m|| \text{ a.e. in } (0,T).$$

Using the inequalities $|f_{1m}|\,|u_m| \leq \frac{1}{2}|f_{1m}| + \frac{1}{2}|f_{1m}|\,|u_m|^2$ and $||f_{2m}||_*||u_m|| \leq \frac{\alpha}{2}||u_m||^2 + \frac{1}{2\alpha}||f_{2m}||_*^2$, gives then

$$\frac{d|u_m|^2}{dt} + \alpha||u_m||^2 \leq (2\beta + |f_{1m}|)|u_m|^2 + \frac{1}{\alpha}||f_{2m}||_*^2 \text{ a.e. in } (0,T),$$

and, forgetting for a while the term $\alpha||u_m||^2$, GRONWALL inequality applies with $\varphi = |u_m|^2$ after integrating in t, and it gives the bound

$$|u_m(t)|^2 \leq \Big(|u_{0m}|^2 + \frac{1}{\alpha}\int_0^t ||f_{2m}(s)||_*^2\,ds\Big) exp\Big(\int_0^t (2\beta + |f_{1m}(s)|)\,ds\Big) \text{ for } t \in [0,T],$$

and then taking into account the term in $\alpha||u_m||^2$ gives

$$\alpha\int_0^T ||u_m(t)||^2\,dt \leq \Big(|u_{0m}|^2 + \frac{1}{\alpha}\int_0^T ||f_{2m}(t)||_*^2\,dt\Big) exp\Big(\int_0^T (2\beta + |f_{1m}(t)|)\,dt\Big) - |u_{0m}|^2.$$

These bounds are good enough for our purpose, but show a strange dependence with respect to the norm of f_{1m}, and one way to avoid it is to use linearity, i.e. to consider first the case $f_{1m} = 0$ for which the above bound is acceptable and then the case where $f_{2m} = 0$, where from the bound

$$\frac{1}{2}\frac{d|u_m|^2}{dt} + \alpha||u_m||^2 - \beta|u_m|^2 \leq |f_{1m}|\,|u_m| \text{ a.e. in } (0,T),$$

one forgets the term $\alpha||u_m||^2$ and one deduces

$$\frac{d|u_m|}{dt} \leq \beta|u_m| + |f_{1m}| \text{ a.e. in } (0,T),$$

giving

$$|u_m(t)| \leq |u_{0m}|e^{\beta\,t} + \int_0^t |f_{1m}(s)|e^{\beta(t-s)}\,ds,$$

and giving the expected affine dependence in the norm of f_{1m}. In our finite dimensional situation one does have $|u_m| \in W^{1,1}(0,T)$ and $\frac{d|u_m|^2}{dt} = 2|u_m|\frac{d|u_m|}{dt}$, but the argument would not work in infinite dimension, where it is better to consider $\sqrt{\varepsilon + |u_m|^2}$ for $\varepsilon > 0$, and therefore

$$\frac{d\sqrt{\varepsilon + |u_m|^2}}{dt} = \frac{1}{2\sqrt{\varepsilon + |u_m|^2}}\frac{d|u_m|^2}{dt} \le (\beta|u_m| + |f_{1m}|)\frac{|u_m|}{\sqrt{\varepsilon + |u_m|^2}} \le \beta\sqrt{\varepsilon + |u_m|^2} + |f_{1m}| \text{ a.e. in } (0,T),$$

and in the bound obtained for $\sqrt{\varepsilon + |u_m|^2}$ one lets ε tend to 0, or one gets the inequality for $\frac{d|u_m|}{dt}$, which shows that $|u_m| \in BV(0,T)$.

This way of getting bounds is not possible if one is dealing with a nonlinear equation, and another way to deal with the bounds is to use the (YOUNG inequality) $||f_{2m}||_*||u_m|| \le \frac{\alpha}{2}||u_m||^2 + \frac{1}{2\alpha}||f_{2m}||_*^2$, which gives

$$\frac{d|u_m|^2}{dt} + \alpha||u_m||^2 \le 2\beta|u_m|^2 + 2|f_{1m}||u_m| + \frac{1}{\alpha}||f_{2m}||_*^2 \text{ a.e. in } (0,T),$$

which, forgetting again the term $\alpha||u_m||^2$ for a while, gives

$$|u_m(t)|^2 \le |u_{0m}|^2 + \frac{1}{\alpha}\int_0^t ||f_{2m}(s)||_*^2\,ds + \int_0^t (2\beta|u_m(s)|^2 + 2|f_{1m}||u_m(s)|)\,ds,$$

and then to use a variant of GRONWALL inequality

$$\varphi(t) \le \psi(t) = A + \int_0^t (2\mu_1\varphi + 2\mu_2\sqrt{\varphi})\,ds$$

which gives

$$\psi' \le 2|\mu_1|\psi + 2|\mu_2|\sqrt{\psi},$$

from which one gets

$$\sqrt{\psi(t)} \le \Big(A + \int_0^t |\mu_2(s)|\,ds\Big)exp\Big(\int_0^t |\mu_1(s)|\,ds\Big),$$

and therefore as $A = |u_{0m}|^2 + \frac{1}{\alpha}\int_0^\tau ||f_{2m}(t)||_*^2\,dt$ as long as $t \le \tau$, one deduces

$$|u_m(t)| \le \Big(\sqrt{|u_{0m}|^2 + \frac{1}{\alpha}\int_0^t ||f_{2m}(s)||_*^2\,ds} + \int_0^t |f_{1m}(s)|)\,ds\Big)e^{\beta t} \text{ on } [0,T].$$

20. Friday February 26.

By taking u_{0n} bounded in H, f_{1n} bounded in $L^1(0,T;H)$ and f_{2n} bounded in $L^2(0,T;V')$, one has obtained a uniform bound for u_n in $C([0,T];H)\cap L^2(0,T;V)$, and one can extract a subsequence u_p converging to u_∞ in $L^\infty(0,T;H)$ weak $\star$ and in $L^2(0,T;V)$ weak, i.e.

$$\int_0^T (u_p.v)\,dt \to \int_0^T (u_\infty.v)\,dt \text{ for all } v\in L^1(0,T;H);\ \int_0^T \langle u_p,v\rangle\,dt \to \int_0^T \langle u_\infty,v\rangle\,dt \text{ for all } v\in L^2(0,T;V).$$

If u_{0n} converges weakly to u_0 in H, f_{1n} converges weakly to f_1 in $L^1(0,T;H)$ and f_{2n} converges weakly to f_2 in $L^2(0,T;V')$, then one can take the limit as $p\to\infty$. For $\varphi\in C^\infty([0,T])$ satisfying $\varphi(T)=0$, one rewrites the term $\int_0^T \varphi\big(\frac{du_p}{dt}.e_k\big)\,dt = -\int_0^T \frac{d\varphi}{dt}(u_p.e_k)\,dt-\varphi(0)(u_{0p}.e_k)$, which enables us to obtain the limit equation

$$-\int_0^T \frac{d\varphi}{dt}(u_\infty.e_k)\,dt+\int_0^T \varphi\langle A\,u_\infty,e_k\rangle=\varphi(0)(u_0.e_k)+\int_0^T\varphi(f_1.e_k)\,dt+\int_0^T\varphi\langle f_2,e_k\rangle\,dt \text{ for all } k,$$

and for all test functions $\varphi\in C^\infty([0,T])$ such that $\varphi(T)=0$. Using linearity one can replace e_k by any linear combination of the elements of the basis, and then by density, by any element $v\in V$. Putting $g_2=f_2-A\,u_\infty\in L^2(0,T;V')$, one has

$$-\int_0^T \frac{d\varphi}{dt}(u_\infty.v)\,dt=\varphi(0)(u_0.v)+\int_0^T\varphi(f_1.v)\,dt+\int_0^T\varphi\langle g_2,v\rangle\,dt \text{ for all } v\in V,$$

and for all $\varphi\in C^\infty([0,T])$ such that $\varphi(T)=0$. This means that $(u_\infty.v)\in W^{1,1}(0,T)$, that its value at 0 is $(u_0.v)$ and that its derivative is $(f_1.v)+\langle g_2,v\rangle$, but if one defines $u_*\in C([0,T];V')$ by $u_*(t)=u_0+\int_0^T\big(f_1(s)+g_2(s)\big)\,ds$, then $\langle u_*,v\rangle$ has the same properties than $(u_\infty.v)$ and therefore they are equal. This shows that $u_\infty\in W^{1,1}(0,T;V')$ with $\frac{du_\infty}{dt}=f_1+g_2$, i.e. u_∞ solves the equation $\frac{du_\infty}{dt}+A\,u_\infty=f_1+f_2$ in $(0,T)$ and $u(0)=u_0$. As we will show that this equation has a unique solution, all the sequence does converge weakly to u_∞.

Uniqueness, which could have been proved before proving existence, follows from the formula

$$\Big\langle\frac{du}{dt},u\Big\rangle=\frac{1}{2}\frac{d|u|^2}{dt} \text{ in } (0,T),$$

as $\frac{du}{dt}+A\,u=0$ implies then $\frac12\frac{d|u|^2}{dt}+\alpha||u||^2-\beta|u|^2\le 0$, and therefore $|u(t)|\le|u(0)|e^{\beta t}$, proving that $u=0$ if $u(0)=0$. The formula is valid if $u\in W_1(0,T)=\{u\in L^2(0,T;V),\,\frac{du}{dt}\in L^1(0,T;H)+L^2(0,T;V')\}$, or in the smaller space $W(0,T)=\{u\in L^2(0,T;V),\,\frac{du}{dt}\in L^2(0,T;V')\}$, used by Jacques-Louis LIONS for the case where $f_1=0$. The formula to be proved is true pointwise if $u\in C^1([0,T];V)$, and in weak formulation it can be written as $\int_0^T\varphi\big\langle\frac{du}{dt},u\big\rangle\,dt=-\frac12\int_0^T\frac{d\varphi}{dt}|u|^2\,dt$ for all $\varphi\in C_c^\infty(0,T)$.

We will first prove that $C^\infty([0,T];V)$ is dense in $W_1(0,T)$; then we will use the density in order to prove that $W_1(0,T)\subset C([0,T];H)$; then we will deduce that the formula is true in $W_1(0,T)$, because both sides of the variational formulation are continuous bilinear forms on $W_1(0,T)$ (the formula implies that $|u|^2\in W^{1,1}(0,T)$ for $u\in W_1(0,T)$). First one notices that the space $W_1(0,T)$ is local, i.e. if $\psi\in C^\infty([0,T])$ and $u\in W_1(0,T)$ then $\psi\,u\in W_1(0,T)$, because u being in $L^2(0,T;V)$ is automatically in $L^1(0,T;H)$ or in $L^2(0,T;V')$. Choosing $\theta\in C^\infty([0,T])$, equal to 1 on $[0,T/3]$ and 0 on $[2T/3,T]$, one can consider $\theta\,u$ as being 0 on $[T,\infty)$ and $(1-\theta)u$ as being 0 on $(-\infty,0]$. One regularizes then $\theta\,u$ by convolution with a regularizing sequence with support on $(-\infty,0)$, and similarly one regularizes $(1-\theta)u$ by convolution with a regularizing sequence with support on $(0,\infty)$, and the usual properties of regularization show that $C^\infty([0,T],V)$ is dense in $W_1(0,T)$.

We want to prove now that $W_1(0,T)$ is continuously imbedded in $C([0,T];H)$, and for that one only needs to show that there exists C such that $||u||_{C([0,T];H)} \leq C||u||_{L^2(0,T;V)} + C\big|\big|\frac{du}{dt}\big|\big|_{L^1(0,T;H)+L^2(0,T;V')}$ for all functions $u \in C^\infty([0,T];V)$. The norm in $L^1(0,T;H)+L^2(0,T;V')$ is the infimum of $||h_1||_{L^1(0,T;H)} + ||h_2||_{L^2(0,T;V')}$ over all the decompositions $\frac{du}{dt} = h_1 + h_2$ with $h_1 \in L^1(0,T;H)$ and $h_2 \in L^2(0,T;V')$. By reasoning on θu and then $(1-\theta)u$, one may assume that u is 0 at one end of the interval; for example, assuming that $u(0) = 0$ one has $|u(t)|^2 = 2\int_0^t \big(\frac{du}{dt}.u\big)\,ds = 2\int_0^t \big\langle \frac{du}{dt}, u\big\rangle\,ds = 2\int_0^t \big((h_1.u) + \langle h_2, u\rangle\big)\,ds$, from which one deduces

$$\begin{aligned}||u||^2_{C([0,T];H)} &\leq 2||h_1||_{L^1(0,T;H)}||u||_{C([0,T];H)} + 2||h_2||_{L^2(0,T;V')}||u||_{L^2(0,T;V)}\\ &\leq 2||h_1||^2_{L^1(0,T;H)} + \frac{1}{2}||u||^2_{C([0,T];H)} + 2||h_2||_{L^2(0,T;V')}||u||_{L^2(0,T;V)},\end{aligned}$$

and therefore

$$||u||^2_{C([0,T];H)} \leq 4||h_1||^2_{L^1(0,T;H)} + 4||h_2||_{L^2(0,T;V')}||u||_{L^2(0,T;V)},$$

and by taking the infimum on the decompositions $\frac{du}{dt} = h_1 + h_2$ with $h_1 \in L^1(0,T;H)$ and $h_2 \in L^2(0,T;V')$, it proves the continuous imbedding of $W_1(0,T)$ into $C([0,T];H)$.

This being done, all the terms of the weak formulation are now seen to be continuous on $W_1(0,T)$, and the formula is true by density.

It must be noticed that one does not have in general $\frac{du_n}{dt}$ bounded in $L^1(0,T;H)+L^2(0,T;V')$. In order to obtain bounds for $\frac{du_n}{dt}$, one can either make time regularity hypotheses on f_1 and f_2 and a regularity hypothesis on u_0 (which corresponds to regularity in space variables in applications to partial differential equations), or use a special (Ritz-) Galerkin basis.

The first idea consists in noticing that formally $u' = \frac{du}{dt}$ satisfies $\frac{du'}{dt} + A\,u' = f_1' + f_2'$ and $u'(0) = f_1(0) + f_2(0) - A\,u_0$, which suggests that if $f_1 \in W^{1,1}(0,T;H)$, $f_2 \in W^{1,1}(0,T;V')$ and $u_0 \in V$ with $A\,u_0 - f_2(0) \in H$, then $u' \in W_1(0,T)$ and one can expect a bound on $\frac{du_n}{dt}$. Indeed, one can easily choose $f_{1n} \in C^\infty([0,T];H)$ and $f_{2n} \in C^\infty(0,T;V')$ converging respectively to f_1 and f_2, so that $u_n \in C^\infty(0,T;V_n)$, but one must be a little careful for the bound on $u_n'(0)$; Jacques-Louis Lions taught the trick of taking u_0 as the first element of the basis (if it is not 0), so that one can take $u_{0n} = u_0$ and then one asks that $f_{2n}(0) - f_2(0)$ converges weakly to 0 in H.

It is useful to notice that if $f_1 \in W^{1,1}(0,T;H)$, $f_2 \in W^{1,1}(0,T;V')$, but $u_0 \in H$ only, then one does not obtain $u' \in W_1(0,T)$ by lack of the needed regularity on u_0, but one has $t\,u' \in W_1(0,T)$, as $v = t\,u'$ satisfies $\frac{dv}{dt} + A\,v = t\frac{df_1}{dt} + t\frac{df_2}{dt} + u'$ and $v(0) = 0$. This is a form of regularization effect for the solutions of the equation, already apparent from the fact that one does not need $u_0 \in V$ in order to have the solution taking its values in V. For obtaining the corresponding estimate for $t\,u_n'$, it is better to use $w = t\,u' - u$, which satisfies $\frac{dw}{dt} + A\,w = t\frac{df_1}{dt} + t\frac{df_2}{dt} - A\,u$ and $v(0) = 0$, and the corresponding bounds for $t\frac{du_n}{dt} - u_n$ are obtained easily.

The choice of a special (Ritz-) Galerkin basis is a different trick, and we will use it for Navier-Stokes equation (at least in dimension 3, as the case of dimension 2 can be handled more easily), but it requires the symmetry of A, or simply $A^T - A \in \mathcal{L}(V,H)$, and the compact injection of V into H. One assumes that $A = A_0 + B$ with A_0 symmetric V-elliptic and $B \in \mathcal{L}(V,H)$. As A_0 is an isomorphism from V onto V', its inverse A_0^{-1} maps V' and therefore H into V, so A_0^{-1} is a compact operator on H, and as it is symmetric, Riesz theory asserts that H has an orthonormal basis made of eigenvectors of A_0^{-1}, $e_n, n \geq 1$, with real positive eigenvalues μ_n converging to 0. Therefore $A_0\,e_n = \lambda_n e_n$ with $\lambda_n = \frac{1}{\mu_n}$ tending to $+\infty$, and if one replaces the norm $||u||$ on V by the equivalent norm $\sqrt{\langle A_0 u, u\rangle}$, then the basis is also orthogonal in V, and therefore it is also orthogonal in V'. The estimate for $\frac{du_n}{dt}$ comes easily once one has observed that for a finite linear combination $v = \sum_i v_i e_i$, one has $||v||^2 = \sum_i \lambda_i |v_i|^2$, $|v|^2 = \sum_i |v_i|^2$, and $||v||_*^2 = \sum_i \frac{1}{\lambda_i}|v_i|^2$.

It is not necessary to take this special basis of eigenvectors in order to deduce estimates on $\frac{du_n}{dt}$, and a more general condition is obtained in the following way. Let P_n be the orthogonal projection of H onto the (closed finite dimensional) subspace V_n, where orthogonality is understood in the scalar product of H, so that P_n is a contraction if one uses the norm of H for V_n; let C_n be the norm of P_n considered as a mapping from V onto V_n equipped with the norm of V; then the basis is special enough in order to obtain a bound for $\frac{du_n}{dt}$ if C_n is bounded (the choice of eigenvectors of A_0 gives $C_n = 1$ if one uses $\sqrt{\langle A_0\cdot,\cdot\rangle}$ for the norm on V). Indeed, if $k \in V'$ and $k_n \in V_n$ is defined by $(k_n.v) = \langle k, v\rangle$ for all $v \in V_n$, then for $w \in V$, one has $\langle k_n, w\rangle = (k_n, w) = (k_n.P_n w) = \langle k, P_n w\rangle = \langle P_n^T k, w\rangle$, so that $||k_n||_* \le C_n ||k||_*$.

There are interesting results of uniqueness which are not based on the ellipticity of A but on its symmetry, and they can be used as well for A or $-A$ with data at 0, or for A with either initial data at 0 or final data at T (one talks of backward uniqueness then). One approach, due to Shmuel AGMON and Louis NIRENBERG consists in proving that if u is a non vanishing solution of $u' + A\,u = 0$, then $|u|$ is log-convex, i.e. $t \mapsto \log|u(t)|$ is convex. Indeed, the derivative of $\log|u|$ is $\frac{(u'.u)}{|u|^2}$, whose derivative is $\frac{(u''.u)+|u'|^2}{|u|^2} - 2\frac{(u'.u)^2}{|u|^4}$, and as $(u''.u) = (A^2u.u) = |A\,u|^2 = |u'|^2$, one concludes by using CAUCHY-SCHWARTZ inequality. In the early 70s I extended this method with Claude BARDOS, and we could apply it to NAVIER-STOKES equation in two dimensions. I have noticed that the log-convexity property is true if A is normal (i.e. A commutes with A^T, or $|A^T v| = |A\,v|$ for all v), and it is equivalent in finite dimension, while Shmuel FRIEDLAND has noticed that a sufficient condition is $|A^T v| \le |A\,v|$ for all v, which may happen without equality in infinite dimension. There is a second approach, by Jacques-Louis LIONS and Bernard MALGRANGE, based on CARLEMAN estimates. I have also introduced another approach, which is useful for improving the localization of the trajectory which results from the log-convexity property: if the solution exists on $[0,T]$, then for $0 \le \tau_1 < \tau_2 \le T$, the trajectory for $t \in (\tau_1, \tau_2)$ lies inside the closed ball with diameter the segment $[u(\tau_1), u(\tau_2)]$; this could certainly be more useful if one knew how to prove similar results for nonlinear equations.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

21. Wednesday March 3.

For applying the preceding abstract framework to STOKES equation, there are questions about the functional spaces. We are already familiar with V (which was denoted W in the stationary case), but we have to identify its closure in $L^2(\Omega; R^N)$, which is the space H of the abstract theory. As we will see later, $H = \{u \in L^2(\Omega; R^N), div(u) = 0, \text{ and } u.n = 0 \text{ on } \partial\Omega\}$, and we will have to explain the meaning of $u.n$ on the boundary, the normal trace of u (physically, $u.n = 0$ means that the flow is tangent, the so called slip condition).

There is however another question which is about the "pressure": have we really solved the STOKES equation

$$\begin{aligned}&\frac{\partial u_i}{\partial t} - \nu\Delta\, u_i + \frac{\partial p}{\partial x_i} = f_i, i = 1, \ldots, N, \text{ in } \Omega\\ &div(u) \text{ in } \Omega\\ &u(\cdot, 0) = u_0 \text{ in } \Omega?\end{aligned}$$

Certainly, if $f \in L^2\big(0, T; H^{-1}(\Omega; R^N)\big)$ and if the solution satisfies $u \in L^2(0, T; V) \cap C^0([0, T]; H)$, $\frac{\partial u}{\partial t} \in L^2\big(0, T; H^{-1}(\Omega; R^N)\big)$ and $p \in L^2\big(0, T; L^2(\Omega)\big)$, then using a test function $v \in V$, one deduces that

$$\frac{d(u.v)}{dt} + \nu\, a(u, v) = \langle f, v\rangle \text{ in } (0, T),$$

and with the data $u_0 \in H$, one has a solution of the abstract problem, solution which we know to be unique. The question is: can we deduce from the abstract formulation that $\frac{\partial u}{\partial t} \in L^2\big(0, T; H^{-1}(\Omega; R^N)\big)$ and $p \in L^2\big(0, T; L^2(\Omega)\big)$? Of course, as one can add to p an arbitrary function of t without changing the equation (which is quite unphysical, but is the price to pay for the unrealistic hypothesis of incompressibility), one must normalize p by asking for example that $\int_\Omega p(x, t)\, dx = 0$ in $(0, T)$.

If $f_1 = 0$, the abstract formulation has given $\frac{du}{dt} = g \in L^2(0, T; V')$, and the problem comes from the fact that V' is not a space of distributions in Ω, as $C_c^\infty(\Omega; R^N)$ is certainly not dense in V, as it is not even included in V because of the constraint $div(u) = 0$. I have mentioned that for a bounded open set Ω with LIPSCHITZ boundary the elements of $H^{-1}(\Omega; R^N)$ orthogonal to V have the form $grad(q)$ with $q \in L^2(\Omega)$ (I have deduced it in the case where $meas(\Omega) < \infty$ from $X(\Omega) = L^2(\Omega)$, but I have not proved that last assertion). For any $L \in V'$, one can solve for $w_L \in V$, unique solution of $a(w_L, v) = L(v)$ for every $v \in V$, and even without the interpretation of this equation using the gradient of a pressure, one sees that $-\Delta\, w_L \in H^{-1}(\Omega; R^N)$, that it defines the same linear form than L on V, and that $||w_L|| = ||L||_*$. The element $g \in L^2(0, T; V')$ can be transformed in this way into a $w_g \in L^2(0, T : V)$ and for every $v \in V$ and every $\varphi \in C_c^\infty(0, T)$, one has

$$-\int_0^T (u(t).v)\frac{d\varphi}{dt}\, dt = \int_0^T \Big\langle \frac{du(t)}{dt}.v\Big\rangle\varphi\, dt = \int_0^T a(w_g, v)\varphi\, dt.$$

One defines then $W_g \in H^1(0, T; V)$ by

$$W_g(t) = \int_0^t w_g(s)\, ds,$$

and as $\int_0^T a(w_g, v)\varphi\, dt = -\int_0^T a(W_g, v)\frac{d\varphi}{dt}\, dt$ for $\varphi \in C_c^\infty(0, T)$, one deduces that $(u(t).v) - a(W_g, v)$ is a constant, and therefore taking $t = 0$ one has

$$(u(t).v) = (u_0.v) + a(W_g, v) \text{ in } (0, T), \text{ for every } v \in V.$$

As $u - u_0 + \Delta\, W_g \in C^0\big(0, T; H^{-1}(\Omega; R^N)\big)$, it shows that

$$u - u_0 + \Delta\, W_g = grad(q) \text{ for some } q \in C^0\Big(0, T; L^2(\Omega)\Big),$$

Taking the derivative in t (in the sense of distributions), one finds that

$$\frac{\partial u}{\partial t} + \Delta\, w_g = grad(p), \text{ but } p = \frac{\partial q}{\partial t}.$$

In order to avoid having the pressure in a space of distributions, one can use a regularity theorem.

Lemma Assume that $A^T = A$. If $u_0 \in V$ and $f \in L^2(0,T;H)$, then $\frac{\partial u}{\partial t}, A\,u \in L^2(0,T;H)$ and $u \in C^0([0,T];V)$. If $u_0 \in H$ and $\sqrt{t}\,f \in L^2(0,T;H)$, then $\sqrt{t}\,\frac{\partial u}{\partial t}, \sqrt{t}\,A\,u \in L^2(0,T;H)$ and $\sqrt{t}\,u \in C^0([0,T];V)$.
Proof: Formally, one can multiply the equation by $\frac{\partial u}{\partial t}$ or by $A\,u$, and one gets either $|u'|^2 + \frac{1}{2}a(u,u)' = (f.u')$ or $\frac{1}{2}a(u,u)' + |A\,u|^2 = (f.A\,u)$, and each implies $a(u,u)' \le \frac{1}{2}|f|^2$, giving the bound of u in $L^\infty(0,T;V)$; then one gets either a bound for u' or a bound for $A\,u$ in $L^2(0,T;H)$, the other bound being given by the equation. Multiplying by u' can be done on the (RITZ-) GALERKIN approach, but not multiplying by $A\,u$ unless one uses a special basis; the same estimates are obtained in the finite dimensional case, and the limit inherits of these bounds, but there is a little work necessary in order to improve $u \in L^\infty(0,T;V)$ into $u \in C^0([0,T];V)$. For example, if $f \in H^1(0,T;H)$ and $u_0 \in D(A)$, then the hypotheses for time regularity are satisfied and $u \in H^1(0,T;V) \subset C^0([0,T];V)$, one concludes by a density argument.

The regularizing effect in the case where one only has $u_0 \in H$, is obtained by multiplying by $t\,u'$ or $t\,A\,u$, the first one being more adapted to the (RITZ-) GALERKIN approach.

In the case where $A^T \neq A$, one has the same result by replacing $u_0 \in V$ by $u_0 \in [D(A),H]_{1/2} = (D(A),H)_{1/2,2}$; however, Jacques-Louis LIONS has shown that if $D(A^T) = D(A)$, then the interpolation space mentioned is actually V.

For the application to STOKES equation (or to NAVIER-STOKES equation), one must be careful about the strange consequences of having identified H and its dual, as this identification is not compatible with the usual basic identification of $L^2(\Omega)$ with its dual. The hypothesis $f \in L^2(0,T;H)$ actually means $f \in L^2(0,T;H')$, so that $f \in L^2\big(0,T;L^2(\Omega;R^N)\big)$ is actually possible, without imposing $div(f) = 0$, which is one condition for taking values in H. One way to think about that question is to remember that gradients have no effect on V or H (if $p \in H^1(\Omega)$) and therefore any element of the form $h + grad(p)$ with $h \in H$ and $p \in H^1(\Omega)$ belongs to H'; when we will study the space H, we will actually prove that the orthogonal of H in $L^2(\Omega;R^N)$ is $\{grad(p), p \in H^1(\Omega)\}$. When we interpret $A\,u \in L^2(0,T;H)$, it means $L^2(0,T;H')$, because $A\,u$ is only defined through the bilinear form $a(u,v)$ with $v \in V$ (although with a constant viscosity, the divergence of $\Delta\,u$ is 0, the normal trace is not 0 in general). However, when we interpret $u' \in L^2(0,T;H)$, it does mean H and not H', as u takes values in $V \subset H$, and u' is a limit of $\frac{u(t+h)-u(t)}{h}$, which take values in V.

If $f \in L^2\big(0,T;L^2(\Omega;R^N)\big)$ and $u_0 \in V$, then one has $u' \in L^2(0,T;H) \subset L^2\big(0,T;L^2(\Omega;R^N)\big)$, and therefore $S = \frac{\partial u}{\partial t} - \nu\Delta\,u - f \in L^2\big(0,T;H^{-1}(\Omega;R^N)\big)$. As $\int_0^T \varphi\langle S(t),v\rangle\,dt = 0$ for all $v \in V$ and all $\varphi \in C_c^\infty(0,T)$, one deduces that for almost every $t \in (0,T)$, $S(t)$ is orthogonal to V (using the separability of V), and therefore $S(t) = -grad\big(p(t)\big)$ with $p(t) \in L^2(\Omega)$; if one normalizes $p(t)$ by adding a constant so that its integral in Ω is 0, one has $||p(t)||_{L^2(\Omega)} \le C||S(t)||_{H^{-1}(\Omega;R^N)}$ and therefore $S = -grad(p)$ with $p \in L^2\big(0,T;L^2(\Omega)\big)$.

Another case where the pressure can be estimated easily is the case $\Omega = R^N$, where one can use FOURIER transform (in x alone): the STOKES equation becomes

$$\begin{aligned}
&\frac{\partial \mathcal{F}u}{\partial t} + 4\nu\pi^2|\xi|^2\mathcal{F}u + 2i\pi\xi\mathcal{F}p = \mathcal{F}f \text{ in } R^N \times (0,T)\\
&(\mathcal{F}u.\xi) = 0 \text{ in } R^N \times (0,T)\\
&\mathcal{F}u(\cdot,0) = \mathcal{F}u_0,
\end{aligned}$$

and taking the scalar product with ξ gives

$$\mathcal{F}p = \frac{1}{2i\pi}\frac{(\mathcal{F}f.\xi)}{|\xi|^2} \text{ in } R^N \times (0,T).$$

Of course, as POINCARÉ inequality does not hold for R^N, one must be careful: if $f \in L^2\big(0,T;L^2(R^N;R^N)\big)$, then one finds a bound for $grad(p)$, but not for p, and therefore one does not find that p takes values in

$H^1(R^N)$, but in a different space (which has been studied first by Jacques DENY and Jacques-Louis LIONS, and present a particular difficulty for $N=2$); however if

$$f_i = \sum_{j=1}^N \frac{\partial g_{ij}}{\partial x_j} \text{ with } g_{i,j} \in L^2\big(0,T; L^2(R^N)\big), i,j = 1,\ldots,N, \text{ in } R^N \times (0,T),$$

then

$$p \in L^2\Big(0,T; L^2(R^N\Big) \text{ with } ||p(\cdot,t)||_{L^2(R^N)} \le C \sum_{i,j=1}^N ||g_{ij}(\cdot,t)||_{L^2(R^N)} \text{ a.e. } t \in (0,T).$$

For NAVIER-STOKES equation, the dimension N will play a very important role, more important than for the stationary case. For $N=2$, we will be able to prove an existence and uniqueness result. For $N \ge 3$, we will prove existence of weak solutions defined on $(0,T)$, but the uniqueness of weak solutions is an open question; for smooth data, we will also prove that strong solutions exist locally, and that they are unique, but it is an open question to show that they can be extended up to T; however, for small smooth data the strong solution will exist globally.

It should be noticed, however, that the approach for proving existence goes absolutely against physical intuition: there is a transport operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \sum_j u_j \frac{\partial}{\partial x_j},$$

and there are various physical quantities transported along the flow, like mass, momentum, energy (or vorticity, helicity, thermodynamic entropy); each component of the velocity satisfies an equation

$$\Big(\frac{D}{DT} - \nu\Delta\Big) u_i = f_i - \frac{\partial p}{\partial x_i} \text{ in } \Omega \times (0,T),$$

and the operator $\frac{D}{DT} - \nu\Delta$ which is applied to each u_i has good properties, some of the bound using the maximum principle and requiring little smoothness of the coefficients u_j, but as $\frac{\partial p}{\partial x_i}$ is needed and the equations are coupled via $div(u) = 0$, it would be useful to have an equation for p; taking the divergence of the equation gives

$$-\Delta\, p = -div(f) + \sum_{i,j=1}^N \frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i},$$

where one has used $div(u) = 0$ for simplifying the divergence of the nonlinear term. The difficulty comes from the fact that one does not have adequate boundary conditions for p. The nonlinearity appearing in the equation for p is actually a little special, with slightly better bounds than expected.

The usual approach, however, does not work with the operator $\frac{D}{DT} - \nu\Delta$, but cuts the operator $\frac{D}{DT}$ into two parts: sending the nonlinear term to play with f, one considers NAVIER-STOKES equation as a perturbation of STOKES equation, and this is obviously not a good idea, but no one has really found how to do better yet.

We have seen in studying the stationary NAVIER-STOKES equations that the nonlinear operator B defined by

$$\langle B(u,v), w\rangle = \int_\Omega u_j \frac{\partial v_i}{\partial x_j} w_i\, dx,$$

is continuous from $V \times V$ into V' for $N \le 4$, and satisfies $\langle B(u,v), w\rangle + \langle B(u,w), v\rangle = 0$ (in particular $\langle B(u,v), v\rangle = 0$, but for the evolution problem we will need more precise bounds, and as $||B(u,v)||_{V'} \le C\sum_{ij} ||u_j v_i||_{L^4(\Omega)}$, one deduces

$$||B(u,u)||_* \le C||u||^2 \text{ in dimension } N = 4,$$
$$||B(u,u)||_* \le C||u||^{3/2}|u|^{1/2} \text{ in dimension } N = 3,$$
$$||B(u,u)||_* \le C||u||\,|u| \text{ in dimension } N = 2.$$

The first two inequalities follow from SOBOLEV imbedding theorem $H_0^1(\Omega) \subset L^4(\Omega)$ in dimension $N = 4$, $H_0^1(\Omega) \subset L^6(\Omega)$ in dimension $N = 3$, the second using also HÖLDER inequality $||v||_{L^4} \le ||v||_{L^6}^{3/4} ||u||_{L^2}^{1/4}$. The third inequality, attributed to Olga LADYZHENSKAYA, uses the same method with which Emilio GAGLIARDO and Louis NIRENBERG proved SOBOLEV imbedding theorem: $|u|^2 \le F(x_2) = \int_R |u| \left|\frac{\partial u}{\partial x_1}\right| dx_1$ and $|u|^2 \le G(x_1) = \int_R |u| \left|\frac{\partial u}{\partial x_2}\right| dx_2$, and therefore $\int_{R^2} |u|^4\,dx \le C \int_{R^2} |u| \left|\frac{\partial u}{\partial x_1}\right| dx \int_{R^2} |u| \left|\frac{\partial u}{\partial x_1}\right| dx$, which gives the desired result by using CAUCHY-SCHWARZ inequality.

The natural bounds for a solution are $u \in L^\infty(0,T;H)$, which corresponds to the fact that the kinetic energy is bounded, and $u \in L^2(0,T;V)$, which corresponds to the fact that the energy dissipated by viscosity between time 0 and T is bounded. The dependence with N becomes then

$$u \in L^2(0,T;V) \cap L^\infty(0,T;H) \text{ imply } B(u,u) \in \begin{cases} L^1(0,T;V') \text{ in dimension } N = 4, \\ L^{4/3}(0,T;V') \text{ in dimension } N = 3, \\ L^2(0,T;V') \text{ in dimension } N = 2, \end{cases}$$

and therefore it is only for $N = 2$ that $B(u,u)$ falls into a space which is allowed for the (abstract) STOKES equation; for $N \ge 3$, the nonlinearity is then a much too strong nonlinear operator, and NAVIER-STOKES equation is then not a mere perturbation of STOKES equation.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

22. Friday March 5.

For $\Omega\subset R^N$, $V=\{u\in H^1_0(\Omega;R^N), div(u)=0 \text{ in } \Omega\}$ and $H=\{u\in L^2(\Omega;R^N), div(u)=0 \text{ in } \Omega$ and $(u.n)=0$ on $\partial\Omega\}$. We will see a little later the meaning of $(u.n)$, where n is the exterior normal to $\partial\Omega$. If V is dense in H (which we will show if Ω is bounded with $\partial\Omega$ smooth enough), then one can consider the (incompressible) NAVIER-STOKES equation as an abstract evolution equation $u'+B(u,u)+A\,u=f$ in $(0,T)$, where $A\in\mathcal{L}(V,V')$ is given by $\langle A\,u,v\rangle=\nu\int_\Omega\big(\sum_{ij}\frac{\partial u_i}{\partial x_j}\frac{\partial v_i}{\partial x_j}\big)\,dx$ for all $u,v\in V$, and B is the bilinear continuous mapping from $V\times V$ into V' (for $N\le 4$), given by $\langle B(u,v),w\rangle=\int_\Omega\big(\sum_{ij}u_j\frac{\partial v_i}{\partial x_j}w_i\big)\,dx$ for all $u,v,w\in V$, and it satisfies $\langle B(u,v),w\rangle+\langle B(u,w),v\rangle=0$, and in particular $\langle B(u,v),v\rangle=0$. As mentioned before, $u'+B(u,u)$ is $\frac{Du}{Dt}$ where $\frac{D}{Dt}=\frac{\partial}{\partial t}+\sum_j u_j\frac{\partial}{\partial x_j}$ is the operator of transport along the flow, and it is not a good idea to handle this operator by cutting it in two parts (geometers use a formalism involving affine connections, but up to now it has not helped understand more on that question of transport, and it is therefore not clear yet what is the right way to handle this operator). However this bad way of treating the nonlinearity does not hurt for $N=2$ and one can prove uniqueness, or more precisely the following continuous dependence with respect to the data.

Proposition: For $N=2$, Ω being any open set in R^2, if

$$f_j=\sum_{k=1}^2\frac{\partial g_{jk}}{\partial x_k}, j=1,2, \text{ with } g_{jk}\in L^2\Big(0,T;L^2(\Omega;R^2)\Big),$$

and if $u_j\in L^2(0,T;V)\cap C^0([0,T];H)$ solve

$$u_j'+B(u_j,u_j)+A\,u_j=f_j, j=1,2, \text{ in } (0,T); u_j(0)=u_{0j}\in H, j=1,2,$$

then one has

$$||u_2-u_1||^2_{C^0([0,T];H)}+\nu||u_2-u_1||^2_{L^2(0,T;V)}\le C\Big(|u_{02}-u_{01}|^2_H+\frac{1}{\nu}\sum_{k=1}^2||g_{2k}-g_{1k}||^2_{L^2(0,T;L^2(\Omega;R^2))}\Big)K$$
$$K(u_1,u_2)=exp\Big(\frac{C}{\nu}\int_0^T\min\{|grad(u_1)|^2_{L^2},|grad(u_2)|^2_{L^2}\}\,dt\Big),$$

where C is a universal constant.

Proof: As $u_j\in L^2(0,T;V)\cap C^0([0,T];H)$ implies $B(u_j,u_j)\in L^2(0,T;V')$ in dimension 2, one has $u_j\in W(0,T)$. Subtracting the two equations and multiplying by u_2-u_1 gives

$$\frac{1}{2}\frac{d|u_2-u_1|^2_H}{dt}+\langle B(u_2,u_2)-B(u_1,u_1),u_2-u_1\rangle+\nu|grad(u_2-u_1)|^2_{L^2}=-\sum_{k=1}^2\Big(g_{2k}-g_{1k}.grad(u_2-u_1)\Big).$$

One has $-\sum_{k=1}^2\Big(g_{2k}-g_{1k}.grad(u_2-u_1)\Big)\le\frac{\nu}{3}|grad(u_2-u_1)|^2_{L^2}+\frac{3}{4\nu}\sum_{k=1}^2|g_{2k}-g_{1k}|^2_{L^2}$ and $\langle B(u_2,u_2)-B(u_1,u_1),u_2-u_1\rangle=\langle B(u_2,u_2-u_1)+B(u_2-u_1,u_1),u_2-u_1\rangle=\langle B(u_2-u_1,u_1),u_2-u_1\rangle$, or $\langle B(u_2,u_2)-B(u_1,u_1),u_2-u_1\rangle=\langle B(u_2-u_1,u_2)+B(u_1,u_2-u_1),u_2-u_1\rangle=\langle B(u_2-u_1,u_2),u_2-u_1\rangle$, and therefore

$$|\langle B(u_2,u_2)-B(u_1,u_1),u_2-u_1\rangle|\le C\min\{|grad(u_1)|_{L^2},|grad(u_2)|_{L^2}\}|u_2-u_1|_H\,|grad(u_2-u_1)|_{L^2},$$

where C is a universal constant, independent of Ω (one does not assume here that POINCARÉ inequality holds, which is the reason for the restriction on f_j). Using the bound $|\langle B(u_2,u_2)-B(u_1,u_1),u_2-u_1\rangle|\le\frac{\nu}{3}|grad(u_2-u_1)|^2_{L^2}+\frac{3C^2}{4\nu}\min\{|grad(u_1)|^2_{L^2},|grad(u_2)|^2_{L^2}\}|u_2-u_1|^2_H$, one obtains

$$\frac{1}{2}\frac{d|u_2-u_1|^2_H}{dt}+\frac{\nu}{3}|grad(u_2-u_1)|^2_{L^2}\le\frac{3}{4\nu}\sum_{k=1}^2|g_{2k}-g_{1k}|^2_{L^2}+\frac{3C^2}{4\nu}\min\{|grad(u_1)|^2_{L^2},|grad(u_2)|^2_{L^2}\}|u_2-u_1|^2_H,$$

and one concludes by GRONWALL inequality.

The same type of proof applies in dimension $N = 3$ or higher, if the solutions are regular enough.

It would be better if one did not cut the operator $\frac{D}{Dt}$ into two pieces, and if one could use the fact that the maximum principle applies to $\frac{D}{Dt} - \nu\Delta$, but a difficulty appears because of the pressure. As a way to handle a similar situation, I want to show a uniqueness result of Michel ARTOLA for an equation

$$\frac{\partial u}{\partial t} - div\big(A(x,u)grad(u)\big) = f,$$

with DIRICHLET conditions (there is an extension to the case $f(x,u)$ that we worked out together, with the type of proof that I show, which is a little more general than Michel's original proof, which extended a result of Neil TRUDINGER). Of course A is a CARATHÉODORY function; one assumes that $|A(x,u)| \leq \beta$ and $(A(x,u).u) \geq \alpha|u|^2$ for all $u \in R^N$ with $\alpha > 0$, and $|A(x,u) - A(x,v)| \leq \omega(|v-u|)$ with ω nondecreasing and satisfying $\int_0^1 \frac{ds}{\omega^2(s)} = +\infty$. Under these conditions, if $f = div(g)$ with $g \in L^2\big(0,T;L^2(\Omega;R^N)\big)$ and $u_0 \in L^2(\Omega)$ there is a unique solution $u \in C^0\big([0,T];L^2(\Omega)\big) \cap L^2\big(0,T;H^1_0(\Omega)\big)$ (there is actully a contraction property in $L^1(\Omega)$ also). If u_1 and u_2 are two such solutions, one subtract the two equations and one multiplies by $\varphi'(u_2-u_1)$ where φ is convex, $\varphi'(0) = 0$ and φ' is bounded. One obtains

$$\frac{d(\int_\Omega \varphi(u_2-u_1)\,dx)}{dt} + \int_\Omega \varphi''(u_2-u_1)\big(A(x,u_2)grad(u_2) - A(x,u_1)grad(u_1).grad(u_2-u_1)\big)\,dx = 0,$$

and using $A(x,u_2)grad(u_2) - A(x,u_1)grad(u_1) = A(x,u_2)grad(u_2-u_1) + \big(A(x,u_2) - A(x,u_1)\big)grad(u_1)$, one deduces

$$\begin{aligned}\frac{d(\int_\Omega \varphi(u_2-u_1)\,dx)}{dt} &+ \alpha\int_\Omega \varphi''(u_2-u_1)|grad(u_2-u_1)|^2\,dx \leq \\ &\int_\Omega \varphi''(u_2-u_1)\omega(|u_2-u_1|)|grad(u_1)|\,|grad(u_2-u_1)|\,dx,\end{aligned}$$

from which one deduces

$$\begin{aligned}\frac{d(\int_\Omega \varphi(u_2-u_1)\,dx)}{dt} &\leq C\int_\Omega \varphi''(u_2-u_1)\omega^2(|u_2-u_1|)|grad(u_1)|^2\,dx \\ &\leq C\max_{s\in R}\{\varphi''(s)\omega^2(|s|)\}\int_\Omega |grad(u_1)|^2\,dx.\end{aligned}$$

One chooses then $0 < \varepsilon < \eta$, and $\varphi_{\varepsilon\eta}$ even and defined by $\varphi_{\varepsilon\eta}(0) = \varphi'_{\varepsilon\eta}(0) = 0$ and $\varphi''_{\varepsilon\eta}(s) = \frac{1}{\omega^2(s)}$ for $\varepsilon < s < \eta$ and $\varphi''_{\varepsilon\eta} = 0$ in $(0,\varepsilon)$ and on $(\eta,+\infty)$; this gives after integration

$$\int_\Omega \varphi_{\varepsilon\eta}(|u_2(x,t) - u_1(x,t)|)\,dx \leq C\int_0^t\int_\Omega |grad(u_1)|^2\,dx\,dt \text{ for } t \in [0,T].$$

Fixing $\eta > 0$ one let ε tend to 0, and as $\varphi_{\varepsilon\eta} \to +\infty$ on $(\eta,+\infty)$, one deduces that $|u_2(x,t) - u_1(x,t)| \leq \eta$ for almost every $x \in \Omega$, and letting then η tend to 0, one deduces that $u_2 = u_1$.

Our next step is to study the functional space H, and show that there is a notion of normal trace $(u.n)$ on the boundary, so that the definition makes sense. Then we will check that V is dense in H.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

23. Monday March 8.

I think that it was Jacques-Louis LIONS who introduced the space $H(div;\Omega) = \{u \in L^2(\Omega;R^N), div(u) \in L^2(\Omega)\}$ and proved that one can give a meaning to $(u.n)$ on the boundary $\partial\Omega$ if Ω has a LIPSCHITZ boundary. First, one notices that $H(div;\Omega)$ is a local space, i.e. $\theta\, u \in H(div,\Omega)$ for all $u \in H(div;\Omega)$ and $\theta \in C^\infty(R^N)$ as $div(\theta\, u) = \theta\, div(u) + (grad(\theta).u)$ (notice that we plan to use the results for the case $div(u) = 0$, but that property is lost by multiplication by smooth functions). Then one shows that $C^\infty(\overline{\Omega};R^N)$ is dense in $H(div,\Omega)$ if the boundary is smooth enough: after using a partition of unity for localizing the problem, one regularizes each $\theta_i\, u$ by convolution with a suitable regularizing sequence adapted to the support of θ_i, and this is possible if Ω is an open set with compact boundary and if near each point of the boundary Ω is only on one side of the boundary and the boundary has an equation $x_N = F(x')$ with F continuous (one can have unbounded boundaries if one asks for a global equation of the corresponding piece with F uniformly continuous). Then, assuming that Ω has a LIPSCHITZ (compact) boundary, so that one can define the normal to the boundary and traces on the boundary for functions in $H^1(\Omega)$, one has the formula

$$\int_\Omega \Big[\Big(u.grad(\varphi)\Big) + div(u)\varphi\Big]\,dx = \int_{\partial\Omega} (u.n)\varphi\, d\sigma,$$

for $u \in C^\infty(\overline{\Omega};R^N)$ and $\varphi \in H^1(\Omega)$. The left side of the equation is a bilinear continuous form on $H(div;\Omega)\times H^1(\Omega)$ and therefore the right side is also continuous for that topology, but the right side is 0 for $\varphi \in H^1_0(\Omega)$ and therefore it is actually defined on the quotient $H^1(\Omega)/H^1_0(\Omega)$; here a natural choice is to use $T(\Omega)$, the space of traces of functions of $H^1(\Omega)$, equipped with the quotient norm

$$||v||_{T(\Omega)} = \inf\{||u||_{H^1(\Omega)}, trace(u) = v\},$$

and then the right side is continuous for the norm of $H(div;\Omega)\times T(\Omega)$, and therefore by density $(u.n)$ is defined on $H(div;\Omega)$ as a linear continuous form on $T(\Omega)$.

If Ω has a compact LIPSCHITZ boundary, then $T(\Omega) = H^{1/2}(\partial\Omega)$, and the proof (and definition of the space) is easily derived from the property for R^N_+, which I review below using FOURIER transform. The interest of the preceding result is that it applies even if the boundary is not so smooth and the trace space $T(\Omega)$ has not been characterized.

It is important to notice that one cannot define each of the terms $u_j n_j$ on the boundary, but only their sum. There is a framework using differential forms which is also useful to know (Jacques-Louis LIONS was not aware of this aspect when he worked on the preceding question). For smooth functions, one considers the $(N-1)$-form $\omega = \sum_i (-1)^{i-1} u_i dx_1 \wedge \ldots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \ldots \wedge dx_N$ (also written $\sum_i (-1)^{i-1} u_i \widehat{dx_i}$), whose exterior derivative is $d\omega = div(u)dx$; in the case of smooth (coefficients and) boundary, one can restrict a p-form to a manifold, as it is a p-linear alternating form and therefore it needs p vectors to act upon and its restriction on the manifold uses only vectors from the tangent space to the manifold; if one restricts the $(N-1)$-form ω to the (smooth) boundary one obtains a form which has only one coefficient (as the dimension of the boundary is $N-1$) and that coefficient is $(u.n)$; it is natural, but not straightforward, that one can relax the hypotheses of regularity and still be able to define the intrinsic quantity $(u.n)$.

There is another space which is important in applications (to Electromagnetism, but also to fluids once one considers the vorticity), i.e. $H(curl;\Omega) = \{u \in L^2(\Omega;R^3), curl(u) \in L^2(\Omega;R^3)\}$; here one should consider the 1-form $\omega = \sum_i u_i dx_i$ and $d\omega = \sum_i \big(curl(u)\big)_i \widehat{dx_i}$ and of course $div\big(curl(u)\big) = 0$ expresses the fact that $d\,d = 0$; in the smooth case the exterior derivative commutes with the restriction and therefore the restriction is a 1-form on the boundary has its exterior derivative well defined. It is the tangential component of u which is well defined on $H(curl;\Omega)$, with a differential restriction corresponding to writing the exterior derivative, and this has been extended to the smooth case by L. PAQUET (I did and taught the LIPSCHITZ case a few years ago).

For $u \in H^1(R^N_+)$ its trace on $x_N = 0$ has been shown to belong to $L^2(R^{N-1})$ (after extending u to a function in $H^1(R^N)$); we want to show that it actually belongs to $H^{1/2}(R^{N-1})$, and that all elements of

$H^{1/2}(R^{N-1})$ can be traces. Of course for $s \geq 0$, the space $H^s(R^m)$ is defined by FOURIER transform as

$$H^s(R^m) = \{u \in L^2(R^m), |\xi|^s |\mathcal{F}u| \in L^2(R^N)\},$$

but for $s < 0$ it is $\{u \in \mathcal{S}'(R^N) : (1 + |\xi|^2)^{s/2} \mathcal{F}u \in L^2(R^N)\}$.

For $u \in C_c^\infty(R^N)$, let $v \in C_c^\infty(R^{N-1})$ be the restriction of u to $x_N = 0$; one defines the FOURIER transforms of u and v

$$\begin{aligned}\mathcal{F}u(\xi', \xi_N) &= \int_{R^N} u(x', x_N) e^{-2i\pi(x'.\xi') - 2i\pi\, x_N \xi_N}\, dx'\, dx_N \\ \mathcal{F}v(\xi') &= \int_{R^{N-1}} u(x', 0) e^{-2i\pi(x'.\xi')}\, dx',\end{aligned}$$

and the critical relation is

$$\mathcal{F}v(\xi') = \int_R \mathcal{F}u(\xi', \xi_N)\, d\xi_N.$$

Indeed for a given ξ', if one defines w by $w(x_N) = \int_{R^{N-1}} u(x', x_N) e^{-2i\pi(x'.\xi')}\, dx'$, then $w \in \mathcal{S}(R)$ and therefore $w(0) = \int_R \mathcal{F}w(\xi_N) d\xi_N$, because $w = \overline{\mathcal{F}}\mathcal{F}w$; after putting back explicitly the dependence in ξ', it is exactly our relation. Using CAUCHY-SCHWARTZ inequality, one deduces

$$\begin{aligned}|\mathcal{F}v(\xi')| &\leq \int_R \frac{1}{\sqrt{1 + |\xi'|^2 + |\xi_N^2}} \sqrt{1 + |\xi'|^2 + |\xi_N|^2}\, |\mathcal{F}u(\xi', \xi_N)|\, d\xi_N \\ &\leq \Big(\int_R \frac{d\xi_N}{1 + |\xi'|^2 + |\xi_N|^2}\Big)^{1/2} \Big(\int_R (1 + |\xi'|^2 + |\xi_N|^2) |\mathcal{F}u(\xi', \xi_N)|^2\, d\xi_N\Big)^{1/2} \\ &= \Big(\frac{\pi}{\sqrt{1 + |\xi'|^2}}\Big)^{1/2} \Big(\int_R (1 + |\xi'|^2 + |\xi_N|^2) |\mathcal{F}u(\xi', \xi_N)|^2\, d\xi_N\Big)^{1/2},\end{aligned}$$

and therefore $(1 + |\xi'|^2)^{1/4} \mathcal{F}v \in L^2(R^{N-1})$.

Conversely, given $v \in H^{1/2}(R^{N-1})$, one must find $u \in H^1(R^N)$ such that $\mathcal{F}v(\xi') = \int_R \mathcal{F}u(\xi', \xi_N)\, d\xi_N$, and one chooses

$$\mathcal{F}u(\xi', \xi_N) = \mathcal{F}v(\xi') \varphi\Big(\frac{\xi_N}{\sqrt{1 + |\xi'|^2}}\Big) \frac{1}{\sqrt{1 + |\xi'|^2}},$$

where $\varphi \in C_c^\infty(R)$ satisfies $\int_R \varphi(s)\, ds = 1$. It remains to check that $u \in H^1(R^N)$, and this follows from

$$\begin{aligned}\int_{R^N} (1 + |\xi|^2) |\mathcal{F}u(\xi)|^2\, d\xi &= \int_{R^N} (1 + |\xi'|^2 + |\xi_N|^2) |\mathcal{F}v(\xi')|^2 \varphi^2\Big(\frac{\xi_N}{\sqrt{1 + |\xi'|^2}}\Big) \frac{1}{1 + |\xi'|^2}\, d\xi'\, d\xi_N \\ &= \int_{R^{N-1}} |\mathcal{F}v(\xi')|^2 \Big(\int_R \frac{1 + |\xi'|^2 + |\xi_N|^2}{1 + |\xi'|^2} \varphi^2\Big(\frac{\xi_N}{\sqrt{1 + |\xi'|^2}}\Big)\, d\xi_N\Big)\, d\xi' \\ &= \Big(\int_R (1 + s^2) \varphi^2(s)\, ds\Big) \int_{R^{N-1}} \sqrt{1 + |\xi'|^2} |\mathcal{F}v(\xi')|^2\, d\xi'.\end{aligned}$$

One defines $H = \{u \in L^2(\Omega; R^N), div(u) = 0 \text{ in } \Omega, (u.n) = 0 \text{ on } \partial\Omega\}$; as one imposes $div(u) = 0$, one has $u \in H(div; \Omega)$ and therefore $(u.n)$ has a meaning; more precisely $(u.n) = 0$ means that for all $\varphi \in H^1(\Omega)$ one has $\int_\Omega \big((grad(\varphi).u) + \varphi\, div(u)\big)\, dx = 0$ for all $\varphi \in H^1(\Omega)$, and $u \in H$ implies then $\int_\Omega (grad(\varphi).u)\, dx = 0$ for all $\varphi \in H^1(\Omega)$, as $div(u) = 0$. One sees then that H is orthogonal to the subspace of gradients of functions in $H^1(\Omega)$.

Lemma: If the injection of $H^1(\Omega)$ into $L^2(\Omega)$ is compact, then the orthogonal of H in $L^2(\Omega; R^N)$ is the (closed) subspace of $grad(\varphi)$ for $\varphi \in H^1(\Omega)$.

Proof: One can apply the equivalence lemma to the case where $E_1 = H^1(\Omega)$, $A = grad$ with $E_2 = L^2(\Omega; R^N)$, and B is the (compact) injection of $H^1(\Omega)$ into $E_3 = L^2(\Omega)$; the equivalence lemma asserts that the range of A is closed. Let us check that the orthogonal of $R(A)$ is H, which proves that the orthogonal of H is the closure of $R(A)$, i.e. $R(A)$ itself.

Assume that $u \in L^2(\Omega; R^N)$ is orthogonal to $R(A)$; taking $\varphi \in C_c^\infty(\Omega)$ and noticing that $\langle div(u), \varphi\rangle = -\langle u, grad(\varphi)\rangle = 0$ shows that $div(u) = 0$ in the sense of distributions. This proves that $u \in H(div; \Omega)$, and using already the information that $div(u) = 0$ in Ω, one deduces that $\int_\Omega (grad(\varphi).u)\, dx = \langle (u.n), trace(\varphi)\rangle$ for all $\varphi \in H^1(\Omega)$, and as the left side is 0 by definition of u, the right side is 0 and therefore $(u.n) = 0$ (as a linear continuous form on the space of traces of functions in $H^1(\Omega)$), i.e. $u \in H$.

We can now look at the important question of density of V into H, as this is basic to the framework used.

Lemma: If $meas(\Omega) < \infty$ and if $L^2(\Omega) = X(\Omega)$ (which is $\{u \in H^{-1}(\Omega), \frac{\partial u}{\partial x_j} \in H^1(\Omega)$ for $j = 1, \ldots, N\}$), then V is dense in H.
Proof: Let $h \in L^2(\Omega; R^N)$ belong to the orthogonal of V in $L^2(\Omega; R^N)$; then h can be considered an element of $H^{-1}(\Omega; R^N)$, orthogonal to V for the duality product, and therefore of the form $grad(p)$ with $p \in L^2(\Omega)$, and as $grad(p) \in L^2(\Omega; R^N)$ it means that $p \in H^1(\Omega)$. Therefore $h = grad(p)$ with $p \in H^1(\Omega)$ and so h is orthogonal to H, which proves that V is dense in H.

The fact that $X(\Omega) = L^2(\Omega)$ requires more regularity of the boundary that the simple compactness of $H^1(\Omega)$ into $L^2(\Omega)$: I had noticed in the Fall that it is not true in a plane domain of the form $\{(x, y) : 0 < x < 1, 0 < y < x^2\}$, and as pointed out by François MURAT during his recent visit, it had been observed a little earlier for similar domains by Giuseppe GEYMONAT and Gianni GILARDI.

One may avoid this problem by taking a definition of H which does not mention a trace on the boundary: $H = \{u \in L^2(\Omega; R^N), div(u) = 0$ in Ω, and $\int_\Omega (grad(\varphi).u)\, dx = 0$ for all $\varphi \in H^1(\Omega)\}$. Then as in a previous lemma, if the injection of $H^1(\Omega)$ into $L^2(\Omega)$ is compact, then the subspace of gradients of functions of $H^1(\Omega)$ is closed and is the orthogonal of H in $L^2(\Omega; R^N)$. The definition of V involves $H_0^1(\Omega)$, which is defined without any reference to the regularity of the boundary, as the closure of $C_c^\infty(\Omega)$ into $H^1(\Omega)$. If $h \in L^2(\Omega; R^N)$ is orthogonal to V, then without knowing anything on the regularity of the boundary one has $h = grad(S)$ for a distribution S, if one invokes a theorem of DE RHAM, or for $S \in L^2_{loc}(\Omega)$ is one uses $X(\omega_k) = L^2(\omega_k)$ for an increasing sequence ω_k of connected open set with smooth boundaries, whose union is Ω (however if a distribution S has its gradient in $L^2(\Omega; R^N)$, one also deduces that $S \in H^1_{loc}(\Omega)$ by classical methods of partial differential equations). It remains to look for hypotheses which imply that every $S \in L^2_{loc}(\Omega)$ with $grad(S) \in L^2(\Omega; R^N)$ actually belongs to $H^1(\Omega)$, and that can be proved for a bounded open set which is locally on one side of the boundary, with a local equation $x_N > F(x')$ with F continuous, by using the techniques already described.

Then, in interpreting the solution of STOKES equation with $f \in L^2\big(0, T; L^2(\Omega; R^N)\big)$ and $u_0 \in V$, one finds $\frac{\partial u}{\partial t} \in L^2(0, T; H)$ and $u \in C^0([0, T]; V)$, and therefore there exists $g \in L^2\big(0, T; L^2(\Omega; R^N)\big)$ such that $a(u(t), v) = (g(t).v)$ for almost every $t \in (0, T)$, and for every $v \in V$. For t outside a set of measure 0, one has then $-\Delta\, u + grad(p) = g$ and $div(u) = 0$, and one deduces $\Delta\, p = div(g)$ in Ω and therefore $p \in H^1_{loc}(\Omega)$, and $u \in H^2_{loc}(\Omega)$. Even if Ω is a bounded open set with LIPSCHITZ boundary, one cannot always deduce that $p \in H^1(\Omega)$, as the $H^2(\Omega)$ regularity for u which is implied is known to be false for some open sets with LIPSCHITZ boundary (in the case $N = 2$, it can be checked in polar coordinates).

The application to existence of weak solutions to NAVIER-STOKES equation is then almost straightforward, but we will have to prove a compactness result. One takes a special basis in the case where $f \in L^2\big(0, T; H^{-1}(\Omega; R^N)\big)$ and one solves the approximate equation for u_n a combination of $e_1, \ldots, e_n$,

$$\Big(\frac{du_n}{dt}.e_k\Big) + \langle B(u_n, u_n), e_k\rangle + \langle A\, u_n, e_k\rangle = \langle f, e_k\rangle \text{ and } (u_n(0).e_k) = (u_0.e_k), \text{ for } k = 1, \ldots, n.$$

The solution exists on an interval $(0, T_c)$ with $T_c \le T$, and one deduces that $T_c = T$ from the bound obtained by taking the combination of e_k corresponding to u_n, i.e.

$$\frac{1}{2}\frac{d|u_n|^2}{dt} + \nu |grad(u_n)|^2 = \langle f, u_n\rangle,$$

and using GRONWALL inequality, one obtains a bound independent of t, so that $T_c = T$, and independent of n: u_n stays in a bounded set of $C^0([0,T];H)$ and in a bounded set of $L^2(0,T;V)$ so that one can extract a weakly converging subsequence. In order to pass to the limit in the nonlinear term $B(u_n,u_n)$, one will use a compactness argument which needs information on the derivative in t of u_n, and this is the reason for the special choice of the basis: $B(u_n,u_n)$ stays bounded in $L^p(0,T;V')$ with $p=2$ for $N=2$, $p=4/3$ for $N=3$, and $p=1$ for $N=4$, and therefore $\frac{du_n}{dt}$ stays bounded in that space too.

The question of regularity mentioned previously has to do with corners (it is a question which has been extensively studied by Pierre GRISVARD). In R^2, if $v = r^\alpha f(\theta)$ then from $dv = \frac{\partial v}{\partial r}dr + \frac{\partial v}{\partial \theta}d\theta = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy$, together with $r\,dr = x\,dx + y\,dy$ and $r^2\,d\theta = x\,dy - y\,dx$, one finds that $\frac{\partial v}{\partial x} = r^{\alpha-1}(\alpha\, f(\theta)\cos\theta - f'(\theta)\sin\theta)$ and $\frac{\partial v}{\partial y} = r^{\alpha-1}(\alpha\, f(\theta)\sin\theta + f'(\theta)\cos\theta)$, and then $\Delta\, v = r^{\alpha-2}\big(f''(\theta) + \alpha^2 f(\theta)\big)$. We look for

$$p = r^\alpha g(\theta),\ u_1 = r^{\alpha+1} f_1(\theta),\ u_2 = r^{\alpha+1} f_2(\theta),$$

and as $\Delta\, p = 0$, it means $g''(\theta) + \alpha^2 g(\theta) = 0$, i.e.

$$g(\theta) = A\cos\alpha\theta + B\sin\alpha\theta,$$

and the equation $-\Delta\, u + grad(p) = 0$ gives

$$\begin{aligned} f_1''(\theta) + (\alpha+1)^2 f_1(\theta) &= \alpha\, g(\theta)\cos\theta - g'(\theta)\sin\theta = \alpha\, A(\cos\alpha\theta\cos\theta + \sin\alpha\theta\sin\theta)\\ &\quad + \alpha\, B(\sin\alpha\theta\cos\theta - \cos\alpha\theta\sin\theta) = \alpha\, A\cos(\alpha-1)\theta + \alpha\, B\sin(\alpha-1)\theta\\ f_2''(\theta) + (\alpha+1)^2 f_2(\theta) &= \alpha\, g(\theta)\sin\theta + g'(\theta)\cos\theta = \alpha\, A(\cos\alpha\theta\sin\theta - \sin\alpha\theta\cos\theta)\\ &\quad + \alpha\, B(\sin\alpha\theta\sin\theta + \cos\alpha\theta\cos\theta) = -\alpha\, A\sin(\alpha-1)\theta + \alpha\, B\cos(\alpha-1)\theta. \end{aligned}$$

If $\alpha \neq 0$, this gives

$$\begin{aligned} f_1(\theta) &= \frac{A}{2}\cos(\alpha-1)\theta + \frac{B}{2}\sin(\alpha-1)\theta + C_1\cos(\alpha+1)\theta + D_1\sin(\alpha+1)\theta\\ f_2(\theta) &= -\frac{A}{2}\sin(\alpha-1)\theta + \frac{B}{2}\cos(\alpha-1)\theta + C_2\cos(\alpha+1)\theta + D_2\sin(\alpha+1)\theta. \end{aligned}$$

The condition $div(u) = 0$ means

$$(\alpha+1)f_1(\theta)\cos\theta - f_1'(\theta)\sin\theta + (\alpha+1)f_2(\theta)\sin\theta + f_2'(\theta)\cos\theta = 0.$$

The coefficient of $\frac{A}{2}$ is $(\alpha+1)\cos(\alpha-1)\theta\cos\theta + (\alpha-1)\sin(\alpha-1)\theta\sin\theta - (\alpha+1)\sin(\alpha-1)\theta\sin\theta - (\alpha-1)\cos(\alpha-1)\theta\cos\theta = 2\cos\alpha\theta$, the coefficient of $\frac{B}{2}$ is $(\alpha+1)\sin(\alpha-1)\theta\cos\theta - (\alpha-1)\cos(\alpha-1)\theta\sin\theta + (\alpha+1)\cos(\alpha-1)\theta\sin\theta - (\alpha-1)\sin(\alpha-1)\theta\cos\theta = 2\sin\alpha\theta$, the coefficient of $(\alpha+1)C_1$ is $\cos(\alpha+1)\theta\cos\theta + \sin(\alpha+1)\theta\sin\theta = \cos\alpha\theta$, the coefficient of $(\alpha+1)D_1$ is $\sin(\alpha+1)\theta\cos\theta - \cos(\alpha+1)\theta\sin\theta = \sin\alpha\theta$, the coefficient of $(\alpha+1)C_2$ is $\cos(\alpha+1)\theta\sin\theta - \sin(\alpha+1)\theta\cos\theta = -\sin\alpha\theta$, the coefficient of $(\alpha+1)D_2$ is $\sin(\alpha+1)\theta\sin\theta + \cos(\alpha+1)\theta\cos\theta = \cos\alpha\theta$. Therefore the coefficient of $\cos\alpha\theta$ is $A + (\alpha+1)C_1 + (\alpha+1)D_2$ and the coefficient of $\sin\alpha\theta$ is $B + (\alpha+1)D_1 - (\alpha+1)C_2$ and these two coefficients must be 0. If one works for $0 < \theta < \theta_0$, then f_1 and f_2 must be 0 for $\theta = 0$ and for $\theta = \theta_0$; for $\theta = 0$ one finds $\frac{A}{2} + C_1 = 0$ and $\frac{B}{2} + C_2 = 0$; this gives $A = -2(\alpha+1)E, B = -2(\alpha+1)F, C_1 = (\alpha+1)E, C_2 = (\alpha+1)F, D_1 = (\alpha+3)F, D_2 = (1-\alpha)E$, and writing that f_1 and f_2 are 0 for $\theta = \theta_0$ gives two equations for the two unknowns E, F and a nontrivial solution requires a determinant equal to 0, which gives a transcendental equation relating θ_0 to α, which I am not courageous enough to check.

24. Wednesday March 10.

Almost all compactness results use a variant of ASCOLI's theorem, but one also needs to learn a few technical tricks to add to the classical theorems; one trick is about approximation questions, so that one can replace spaces of continuous functions by spaces of integrable functions; another trick, which I show first, permits to transfer compactness from one space to another.

Lemma: (Jacques-Louis LIONS) If E_1, E_2, E_3, are three normed spaces with $E_1 \subset E_2$ with continuous and compact injection, and $E_2 \subset E_3$ with continuous injection, then for every $\varepsilon > 0$ there exists $C(\varepsilon)$ such that

$$||u||_{E_2} \leq \varepsilon ||u||_{E_1} + C(\varepsilon)||u||_{E_3} \text{ for all } u \in E_1.$$

Proof: If it was not true, there would exist $\varepsilon_0 > 0$ such that for every n one could find $u_n \in E_1$ with $||u_n||_{E_2} > \varepsilon_0 ||u_n||_{E_1} + n||u_n||_{E_3}$. By homogeneity, one may normalize u_n in order to have $||u_n||_{E_1} = 1$, and by continuity one obtains $||u_n||_{E_2} \leq C$, and the inequality implies then that $||u_n||_{E_3} \leq C/n$, so that $u_n \to 0$ in E_3. The sequence u_n is bounded in E_1, and therefore belongs to a compact subset of E_2; one can then extract a subsequence u_m which converges (strongly) in E_2 to a limit z, but one finds a contradiction because one must have $||z||_{E_2} \geq \varepsilon_0 > 0$ and $z = 0$ because u_n must converge to z in E_3.

Lemma: Let E_1, E_2, E_3 be three BANACH spaces with $E_1 \subset E_2$ with continuous and compact injection, and $E_2 \subset E_3$ with continuous injection, and let $p \in [1, \infty]$. If a sequence is bounded in $L^p(0, T; E_1)$ and belongs to a compact of $L^p(0, T; E_3)$, then it belongs to a compact of $L^p(0, T; E_2)$.
Proof: One extracts a subsequence u_n which converges in $L^p(0, T; E_3)$ and therefore it is a CAUCHY sequence in $L^p(0, T; E_3)$. From the preceding lemma one has $||u_n - u_m||_{E_2} \leq \varepsilon ||u_n - u_m||E_1 + C(\varepsilon)||u_n - u_m||_{E_3}$, and taking the norms in $L^p(0, T)$ gives

$$||u_n - u_m||_{L^p(0,T;E_2)} \leq \varepsilon ||u_n - u_m||_{L^p(0,T;E_1)} + C(\varepsilon)||u_n - u_m||_{L^p(0,T;E_3)}.$$

Taking the lim sup as n, m tend to infinity gives then $\limsup_{n,m\to\infty} ||u_n - u_m||_{L^p(0,T;E_2)} \leq 2M\varepsilon$, where M is a bound for the sequence in $L^p(0, T; E_1)$, and letting then ε tend to 0 shows that u_n is a CAUCHY sequence in $L^p(0, T; E_2)$.

The inequality of the first lemma also occurs in the case $E_1 \subset E_2 \subset E_3$ as a consequence of an interpolation inequality, i.e. if there exists $\theta \in (0, 1)$ such that $||u||_{E_2} \leq M\, ||u||_{E_1}^{1-\theta} ||u||_{E_3}^{\theta}$ for all $u \in E_1$ (it is not necessary to know what interpolation theory is, but a result of Jacques-Louis LIONS and Jaak PEETRE asserts that the preceding inequality is equivalent to $(E_1, E_3)_{\theta,1} \subset E_2$ with continuous injection). In that case the first lemma is just YOUNG's inequality, and the second lemma can be strengthened.

Lemma: Let E_1, E_2, E_3 be three BANACH spaces with $E_1 \subset E_2 \subset E_3$ and assume that there exists $\theta \in (0, 1)$ such that $||u||_{E_2} \leq M\, ||u||_{E_1}^{1-\theta} ||u||_{E_3}^{\theta}$ for all $u \in E_1$. Let $p_1, p_3 \in [1, \infty]$. If a sequence is bounded in $L^{p_1}(0, T; E_1)$ and belongs to a compact of $L^{p_3}(0, T; E_3)$, then it belongs to a compact of $L^{p_2}(0, T; E_2)$ with p_2 defined by $\frac{1}{p_2} = \frac{1-\theta}{p_1} + \frac{\theta}{p_3}$.
Proof: This is just HÖLDER inequality applied to the inequality $||u_n - u_m||_{E_2} \leq M\, ||u_n - u_m||_{E_1}^{1-\theta} ||u_n - u_m||_{E_3}^{\theta}$, and the choice of p_2 gives $||u_n - u_m||_{L^{p_2}(0,T;E_2)} \leq M\, ||u_n - u_m||_{L^{p_1}(0,T;E_1)}^{1-\theta} ||u_n - u_m||_{L^{p_3}(0,T;E_3)}^{\theta}$, showing that $\limsup_{n,m\to\infty} ||u_n - u_m||_{L^{p_2}(0,T;E_2)} = 0$ because $\theta < 1$.

Notice that there is no compact injection hypothesis in this last lemma, and one often applies it to the case $E_1 = E_2 = E_3$ in order to change the value of p.

In his lectures, Jacques-Louis LIONS used a compactness lemma in which a sequence is bounded in $L^p(0, T; E_1)$ while its derivative with respect to t is bounded in $L^p(0, T; E_3)$ and the conclusion is that the sequence belongs to a compact of $L^p(0, T; E_2)$ (the injection of E_1 into E_2 being compact); he used $1 < p < \infty$ and reflexive spaces, and he referred to Jean-Pierre AUBIN, but as I never read the corresponding article, I do not know what each of them did; in his book on nonlinear problems he also refers to Roger

TEMAM for some variants. As I had heard Pascal MARONI mention that the hypothesis of reflexivity was not necessary, I taught the lemma for $1 < p < \infty$ without that hypothesis in Madison in 1974/75; a few years later, I remember discussing the case $p = 1$ with François MURAT and Lucio BOCCARDO (I remember writing the scenario of a proof on the paper tablecloth of a restaurant in Roma), but I am not sure if it is written anywhere. Putting down different ideas together leads to the following results.

Lemma: Let E be a BANACH space. Assume that that there exists $p \in [1, \infty)$, $\eta > 0$ and a constant M such that a sequence u_n is bounded in $L^p(0,T;E)$ with $\left(\int_0^{T-h} ||u_n(t+h) - u_n(t)||^p\,dt\right)^{1/p} \le M\,|h|^\eta$ for all $h \in (0, T/2)$. Then u_n is bounded in $L^q(0,T;E)$ with $q < p/(1-\eta\,p)$ if $\eta < 1/p$, and $q < \infty$ if $\eta \ge 1/p$.

Proof: Let $\varphi(t) = ||u_n(t)||$, then $\varphi \in L^1(0,T)$ and $\left(\int_0^{T-h} |\varphi(t+h) - \varphi(t)|^p\,dt\right)^{1/p} \le M\,|h|^\eta$ for all $h \in (0, T/2)$. The lemma is just a part of the equivalent of SOBOLEV imbedding theorem for BESOV spaces, and the natural spaces here are not the $L^p(0,T)$ spaces but the MARCINKIEWICZ spaces ($L^{p,\infty}(0,T)$ in the notations of LORENTZ spaces). One may assume that the support of u is included on $(0, T/2)$ by localization, and one uses then nonnegative regularizing sequences $\rho_\varepsilon(x) = \frac{1}{\varepsilon}\rho\left(\frac{x}{\varepsilon}\right)$ with support in $(-1, 0)$.

Assume that $u \in L^a(0,T;E)$ and $\int_0^{T-h} ||u(t+h) - u(t)||_E^p\,dt \le M^p\,|h|^{\eta\,p}$ for all $h < \frac{T}{2}$; then one writes $u = (\rho_\varepsilon \star u) + \big(u - (\rho_\varepsilon \star u)\big)$, and one has $\big(u - (\rho_\varepsilon \star u)\big)(x) = \frac{1}{\varepsilon}\int \rho\left(\frac{y}{\varepsilon}\right)\big(u(x) - u(x-y)\big)\,dy$ because $\int_R \rho(x)\,dx = 1$, and therefore $||u - (\rho_\varepsilon \star u)||_{L^p(0,T/2)} \le \frac{M}{\varepsilon}\int \rho\left(\frac{y}{\varepsilon}\right)|y|^\eta\,dy \le M\varepsilon^\eta$. On the other hand, $(\rho_\varepsilon \star u)(x) = \frac{1}{\varepsilon}\int \rho\left(\frac{y}{\varepsilon}\right)u(x-y)\,dy$, so $||\rho_\varepsilon \star u||_{L^\infty(0,T/2)} \le ||u||_{L^a}||\rho_\varepsilon||_{L^{a'}} \le ||u||_{L^a}\left(\frac{K}{\varepsilon}\right)^{1/a}$, where $K = ||\rho||_{L^\infty}$.

For $\lambda > 0$, one wants an estimate of the measure of $\omega_{2\lambda} = \{t \in (0, T/2) : |u(t)| \ge 2\lambda\}$, and one chooses $\varepsilon > 0$ such that $||u||_{L^a}\left(\frac{K}{\varepsilon}\right)^{1/a} = \lambda$, so that $u = v + w$ with $v = \rho_\varepsilon \star u$ bounded by λ; therefore $\omega_{2\lambda}$ is included in the set where $|w| = |u - (\rho_\varepsilon \star u)| \ge \lambda$, which gives the estimate $\lambda^p meas(\omega_{2\lambda}) \le M^p\varepsilon^{\eta\,p}$, and using the choice of ε, i.e. $K\lambda^{-a}||u||_{L^a}^a$, one obtains the estimate $\lambda^{p+a\eta\,p} meas(\omega_{2\lambda}) \le M^p K^{\eta\,p}||u||_{L^a}^{a\eta\,p}$.

This estimate shows that $u \in L^b$ for $b < p(1 + a\,\eta)$. In the case $\eta\,p < 1$, let $a^* = p(1 + a^*\,\eta)$, then starting with $a = p$, one sees that for $p \le a < a^*$ one has $a < p(1 + a\eta)$ and one finds that $u \in L^{a_k}$ with a_k converging to a^*, and therefore one can find a bound for $||u||_{L^q}$ for any $q \in [p, a^*)$ in terms of M and $||u||_{L^p}$. In the case $\eta\,p \ge 1$, one always has $a < p(1 + a\,\eta)$, and one finds that $u \in L^{a_k}$ with a_k tending to $+\infty$, and therefore one can find a bound for $||u||_{L^q}$ for any $q \in [p, +\infty)$ in terms of M and $||u||_{L^p}$.

Lemma: Let E_1 and E_3 be two BANACH spaces with $E_1 \subset E_3$, the injection being continuous and compact. Assume that for some $p \in [1, \infty]$ a sequence u_n is bounded in $L^p(0,T;E_1)$, and that there exists $\eta > 0$ and a constant M such that $\left(\int_0^{T-h} ||u_n(t+h) - u_n(t)||^p\,dt\right)^{1/p} \le M\,|h|^\theta$ for all $h \in (0, T/2)$, then u_n belong to a compact set of $L^p(0,T;E_3)$.

Proof: After localization so that the support of all u_n is in $(0, T/2)$, one chooses $h > 0$ small and one defines $v_n(t) = \frac{1}{h}\int_t^{t+h} u_n(s)\,ds$ and $w_n = u_n - v_n$. As before, $w_n(t) = \frac{1}{h}\int_t^{t+h}\big(u_n(t) - u_n(s)\big)\,ds$, giving the bound $||w_n||_{L^p(0,T;E_3)} \le C\,|h|^\theta$. For a fixed $h > 0$, v_n takes its values in a bounded set of E_1, and therefore in a compact set of E_3, but as from the previous lemma u_n is bounded in some $L^q(0,T;E_3)$ with $q > 1$, one sees that v_n has its derivative bounded in $L^q(0,T;E_3)$ and is thereforeuniformly HÖLDER continuous with values in E_3; by ASCOLI's theorem a subsequence v_m converges uniformly. One deduces that $\limsup_{m,m'\to\infty} ||u_m - u_{m'}||_{L^p(0,T;E_3)} \le \limsup_{m,m'\to\infty} ||w_m - w_{m'}||_{L^p(0,T;E_3)} \le 2C\,|h|^\theta$, and letting h tend to 0 shows that u_m is a CAUCHY sequence in $L^p(0,T;E_3)$.

In our application to NAVIER-STOKES equation using a special (RITZ-) GALERKIN basis, one has u_n bounded in $L^2(0,T;V)$ and in $L^\infty(0,T;H)$, and $\frac{du_n}{dt}$ is bounded in $L^p(0,T;V')$, with $p = 2$ for $N = 2$, $p = 4/3$ for $N = 3$, and $p = 1$ for $N = 4$; moreover V is continuously and compactly imbedded into H (and therefore into V'). One can take $\theta = 1$ and one first deduces that u_n belongs to a compact of $L^p(0,T;V')$; but as it is bounded in $L^\infty(0,T;V')$, it belongs to a compact of $L^q(0,T;V')$ for all $q < \infty$; then it belongs to a compact of $L^2(0,T;H)$, the limitation by 2 being due to the estimate of u_n in $L^2(0,T;V)$, and one can extract subsequences which converge almost everywhere in $\Omega \times (0,T)$. In dimension $N = 2$, using an interpolation inequality, each component of u_n is bounded in $L^4\big(\Omega \times (0,T)\big)$ and therefore the term $(u_m)_j(u_m)_i$ for which one needs the limit (in order to compute the limit of the term $\langle B(u_m, u_m), e_k\rangle$) is bounded in $L^2\big(\Omega \times (0,T)\big)$ and converges almost everywhere to $(u_\infty)_j(u_\infty)_i$; in dimension $N = 3$, each component of u_n is bounded in $L^{8/3}\big(0,T;L^4(\Omega)\big)$ and therefore the term $(u_m)_j(u_m)_i$ is bounded in $L^{4/3}\big(0,T;L^2(\Omega)\big)$ and converges then almost everywhere to $(u_\infty)_j(u_\infty)_i$. The case $N \ge 4$ uses interior regularity for the e_k.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

25. Friday March 12.

We have obtained the existence of a weak solution for an abstract formulation of NAVIER-STOKES equation; our solution is defined on $(0,T)$, and one may take $T=+\infty$ without much change in the proof. Even in dimension $N=2$, where we know the solution to be unique, it is useful to know whether or not it is regular when the data are more regular. For $N=3$ (or $N\geq 4$ for purely mathematical reasons), one may wonder if one can find a strong solution, i.e. a solution having a better regularity so that the solution would be unique, for example, or if the pressure would be found in a space of locally integrable functions in (x,t). Some of the regularity results depend upon the smoothness of the boundary $\partial\Omega$, for example in order to use the fact that $D(A)\subset H^2(\Omega;R^N)$, which is not always true for LIPSCHITZ domains.

Lemma: If $N=2$, if $f\in L^2\big(0,T;L^2(\Omega;R^2)\big)$ and $u_0\in V$, if Ω is smooth enough so that POINCARÉ inequality holds and $D(A)\subset H^2(\Omega;R^2)$, then the solution of $u'+B(u,u)+A\,u=f$ and $u(0)=u_0$ satisfies $u\in C^0([0,T];V)$, $u\in L^2\big(0,T;H^2(\Omega;R^2)\big)$, and $u'\in L^2(0,T;H)$.

Proof: One proves the estimates for the approximation with the special basis, multiplying by $A\,u_n$, and one needs to bound terms like $(u_n)_j\frac{\partial(u_n)_i}{\partial x_j}$ in $L^2(\Omega)$, for example from a bound of $(u_n)_j$ in $L^\infty(\Omega)$; one may use a bound $||v||_{L^\infty(\Omega)}\leq C||v||_{L^2(\Omega)}^{1/2}||v||_{H^2(\Omega)}^{1/2}\leq C|v|^{1/2}|A\,v|^{1/2}$ (valid in dimension 2), and one obtains

$$\begin{aligned}\frac{1}{2}\frac{d(||u_n||^2)}{dt}+|A\,u_n|^2&\leq |f|\,|A\,u_n|+C|u_n|^{1/2}||u_n||\,|A\,u_n|^{3/2}\\&\leq \varepsilon|A\,u_n|^2+C(\varepsilon)|f|^2+C(\varepsilon)|u_n|^2||u_n||^4,\end{aligned}$$

where one has used YOUNG's inequality $ab\leq|\varepsilon\,a|^p/p+|b/\varepsilon|^{p'}/p'$, with $p=4/3,p'=4$, and one deduces

$$||u_n(t)||^2\leq||u_0||^2+C\int_0^t|f(s)|^2\,ds+C\int_0^t\lambda_n(s)||u_n(s)||^2\,ds,\ \text{with}\ \lambda_n=C|u_n|^2||u_n||^2,$$

from which a uniform bound for $||u_n||$ is deduced by applying GRONWALL inequality, as λ_n is bounded in $L^1(0,T)$, and this implies that $|A\,u_n|$ is bounded in $L^2(0,T)$.

In the preceding proof, one can use different estimates for the bound in L^∞; for example, trying to get the power of $|A\,u_n|$ as low as possible, one can use $||v||_{L^\infty(\Omega)}\leq C(\eta)||v||_{L^2(\Omega)}^{1/(1+\eta)}||v||_{H^{1+\eta}(\Omega)}^{\eta/(1+\eta)}$ (valid in dimension 2), and if one uses $\theta=\eta/(1+\eta)\in(0,1/2)$, it gives a bound for $B(u_n,u_n)$ in $L^2(\Omega;R^2)$ of the form $C(\theta)|u_n|^\theta||u_n||^{2-2\theta}|A\,u_n|^\theta$. The application of YOUNG's inequality, with $p=2/(1+\theta),p'=2/(1-\theta)$, gives a term in $\varepsilon|A\,u_n|^2+C(\varepsilon,\theta)|u_n|^{2\theta/(1-\theta)}||u_n||^4$, and therefore one gains on the power of $|u_n|$ but not on the power of $||u_n||$.

Lemma: If $N=3$, if $f\in L^2\big(0,T;L^2(\Omega;R^3)\big)$ and $u_0\in V$, if Ω is smooth enough so that POINCARÉ inequality holds and $D(A)\subset H^2(\Omega;R^3)$, then there exists $T_c\in(0,T]$ depending upon the norms of the data such that there exists a unique solution of $u'+B(u,u)+A\,u=f$ and $u(0)=u_0$ on $[0,T_c]$ which satisfies $u\in C^0([0,T_c];V)$, $u\in L^2\big(0,T_c;H^2(\Omega;R^3)\big)$, and $u'\in L^2(0,T_c;H)$.

Proof: Same type of proof than before, but now one has $||v||_{L^\infty(\Omega)}\leq C||v||^{1/2}|A\,v|^{1/2}$ (valid in dimension 3), giving a bound for $B(u_n,u_n)$ in $L^2(\Omega;R^3)$ of the form $||u_n||^{3/2}|A\,u_n|^{1/2}$. The application of YOUNG's inequality, with $p=4/3,p'=4$, gives a term in $\varepsilon|A\,u_n|^2+C(\varepsilon)||u_n||^6$, and the exponent is too large for obtaining a bound by GRONWALL's inequality, and one can only obtain a local bound from the inequality

$$\frac{d(||u_n||^2)}{dt}+\alpha|A\,u_n|^2\leq C\,|f|^2+C\,||u_n||^6.$$

After integration, and omission of the term in $|A\,u_n|^2$, one obtains

$$||u_n(t)||^2\leq||u_0||^2+C\int_0^t|f(s)|^2\,ds+C\int_0^t||u_n(s)||^6\,ds\leq K+C\int_0^t||u_n(s)||^6\,ds,\ \text{in}\ (0,T),$$

where $K = ||u_0||^2 + C\int_0^T |f(s)|^2\,ds$, and if one defines φ by $\varphi(t) = \int_0^t ||u_n(s)||^6\,ds$, then one has $\varphi' \le (K + C\,\varphi)^3$, which implies $\big((K + C\,\varphi)^{-2}\big)' = -2C(K + C\,\varphi)^{-3}\varphi' \ge -2C$ and therefore $\big(K + C\,\varphi(t)\big)^{-2} \ge K^{-2} - 2C\,t$, which is only useful on $[0, T_c]$ if $K^{-2} - 2C\,T_c > 0$, in which case it gives the desired bounds.

For proving uniqueness one assumes that one has two solutions u^1, u^2, one subtracts the equations and one multiplies by $A(u^2 - u^1)$, and one has to estimate the norm in $L^2(\Omega; R^3)$ of $B(u^2, u^2) - B(u^1, u^1)$, which one writes as $B(u^2, u^2 - u^1) + B(u^2 - u^1, u^1)$; the first term can be bounded as $||u^2||_{L^6(\Omega;R^3)}||u^2 - u^1||_{W^{1,3}(\Omega;R^3)}$, bounded by $C||u^2||\,||u^2 - u^1||^{1/2}|A\,u^2 - A\,u^1|^{1/2}$; the second term can be bounded as $||u^2 - u^1||_{L^\infty(\Omega;R^3)}||u^1||$, bounded by $C||u^1||\,||u^2 - u^1||^{1/2}|A\,u^2 - A\,u^1|^{1/2}$. One obtains then $(||u^2 - u^1||^2)' + 2|A\,u^2 - A\,u^1|^2 \le C(||u^1|| + ||u^2||)||u^2 - u^1||^{1/2}|A\,u^2 - A\,u^1|^{3/2} \le |A\,u^2 - A\,u^1|^2 + C(||u^1|| + ||u^2||)^4\,||u^2 - u^1||^2$, and one concludes by using GRONWALL inequality.

If the data are small enough, one can take $T_c = T$, but one can even obtain global existence on $[0, \infty)$ if the data are small enough; for simplicity, I consider first the case $f = 0$.

Lemma: If $N = 3$, if Ω is smooth enough so that POINCARÉ inequality holds and $D(A) \subset H^2(\Omega; R^3)$, then if $u_0 \in V$, and if $|u_0|\,||u_0||$ is small enough, then the solution of $u' + B(u, u) + A\,u = 0$ with $u(0) = u_0$ exists for all $t \in [0, \infty)$ and satisfies $u \in C^0([0, \infty); V)$, $u \in L^2\big(0, \infty; H^2(\Omega; R^3)\big)$, and $u' \in L^2(0, \infty; H)$.
Proof: One bounds the norm of $v_j \frac{\partial v_i}{\partial x_j}$ in $L^2(\Omega)$ by $||v_j||_{L^3(\Omega)}\big|\big|\frac{\partial v_i}{\partial x_j}\big|\big|_{L^6(\Omega)}$, $\big|\big|\frac{\partial v_i}{\partial x_j}\big|\big|_{L^6(\Omega)}$ by $C||v||_{H^2(\Omega)}$ and $||v_j||_{L^3(\Omega)}$ by $C|v|^{1/2}||v||^{1/2}$, so that

$$|\langle B(u_n, u_n), A\,u_n\rangle| \le C_0|u_n|^{1/2}\,||u_n||^{1/2}\,|A\,u_n|^2 \text{ for every } u_n \in D(A).$$

Because $f = 0$, one obtains

$$\begin{aligned}&\frac{1}{2}\frac{d(|u_n|^2)}{dt} + ||u_n||^2 \le 0\\ &\frac{1}{2}\frac{d(||u_n||^2)}{dt} + |A\,u_n|^2 \le C_0|u_n|^{1/2}\,||u_n||^{1/2}\,|A\,u_n|^2 \text{ on } (0, T),\end{aligned}$$

and therefore as long as $C_0|u_n|^{1/2}\,||u_n||^{1/2} < 1$, both the quantities $|u_n|$ and $||u_n||$ are nonincreasing and their product is less than its value at time 0; consequently, if the initial data $u_0 \in V$ is chosen so that

$$C_0|u_0|^{1/2}\,||u_0||^{1/2} \le 1,$$

then one has $|u_n(t)| \le |u_0|$ and $||u_n(t)|| \le ||u_0||$ on $(0, \infty)$, and a global solution exists on $(0, \infty)$.

In the case $f \ne 0$, one has $|u_n(t)| \le |u_0| + \int_0^t |f(s)|\,ds$, but if one wants to avoid assuming $f \in L^1\big(0, \infty; L^2(\Omega; R^3)\big)$, one may use POINCARÉ inequality $||v||^2 \ge \lambda_1|v|^2$ for all $v \in H_0^1(\Omega)$, and the inequality $(|u_n|^2)' + 2\lambda_1|u_n|^2 \le 2|f|\,|u_n| \le 2\lambda_1|u_n|^2 + |f|^2/2\lambda_1$ gives $|u_n(t)|^2 \le |u_0|^2 + \frac{1}{2\lambda_1}\int_0^t |f(s)|^2\,ds$. If one can enforce the condition $C_0|u_n|^{1/2}||u_n||^{1/2} \le 1/2$, then one has $(||u_n||^2)' + |A\,u_n|^2 \le 2|f|\,|A\,u_n| \le |A\,u_n|^2 + |f|^2$ and therefore $||u_n||^2 \le ||u_0||^2 + \int_0^t |f|^2\,dt$, and the condition to enforce is satisfied if one asks that

$$\Big(|u_0|^2 + \frac{1}{2\lambda_1}\int_0^\infty |f(s)|^2\,ds\Big)\Big(||u_0||^2 + \int_0^\infty |f(s)|^2\,ds\Big) \le \frac{1}{16C_0^4}.$$

In the case where $f = 0$, instead of putting conditions on $|u_0|\,||u_0||$ one can impose a more natural condition that u_0 be small in the domain of $A^{1/4}$; this is done by multiplying by $A^{1/2}u_n$, and using the estimate $||v||_{L^3(\Omega;R^3)} \le C|A^{1/4}v|$, which implies $|(B(u_n, u_n), A^{1/2}u_n)| \le C_1|A^{1/4}u_n|\,|A^{3/4}u_n|^2$, and therefore if $C_1|A^{1/4}u_0| \le 1$, then the norm of $|A^{1/4}u_n|$ is nonincreasing and stays then $\le \frac{1}{C_1}$; one easily extends this idea to the case $f \ne 0$.

All these types of inequalities are quite standard, and although I may have improved on details, I had learned most of these techniques in lectures of Jacques-Louis LIONS in the late 70s; I had taught these tecgniques in my 1974/75 course in Madison (the lecture notes were written by graduate students). After that I advocated following a little more the Physics of the fluid flows in order to get better results, and I still insist that one should not cut the transport term into two pieces, and I thought that everything important had been found out of these differential inequalities. I was wrong; in January 1980, Colette GUILLOPÉ showed me some handwritten pages by Ciprian FOIAS, and I took a copy which I looked at during the following month, which I spent at the TATA Institute in Bangalore; I made an improvement on Ciprian FOIAS's original computation, but the idea is his.

In the case $N=3$, taking $f=0$ in order to simplify, one starts from the already mentioned differential inequality $(||u_n||^2)' + |A\,u_n|^2 \leq C||u_n||^6$, together with $(|u_n|^2)' + 2||u_n||^2 = 0$, which gives the existence on $(0,T)$ of the approximate solution u_n; the idea of Ciprian FOIAS was to divide by $1+||u_n||^6$, while my improvement is to divide only by $1+||u_n||^4$! One obtains

$$\frac{d\Big(\arctan(||u_n||^2)\Big)}{dt} + \frac{|A\,u_n|^2}{1+||u_n||^4} = \frac{1}{1+||u_n||^4}\Big((||u_n||^2)' + |A\,u_n|^2\Big) \leq \frac{C||u_n||^6}{1+||u_n||^4} \leq C||u_n||^2.$$

Integrating from 0 to T (which can be $+\infty$), one obtains

$$\int_0^T \frac{|A\,u_n|^2}{1+||u_n||^4}\,dt \leq \arctan(||u_n(0)||^2) + C\int_0^T ||u_n||^2\,dt \leq \frac{\pi}{2} + \frac{C|u_0|^2}{2}.$$

One has $\frac{|A\,u_n|^{1/2}}{1+||u_n||}$ bounded in $L^4(0,T)$, but as $||u_n||^{1/2}(1+||u_n||)$ is bounded in $L^{4/3}(0,T)$, one has

$$||u_n||^{1/2}|A\,u_n|^{1/2} \text{ bounded } inL^1(0,T),$$

from which one obtains

$$u_n \text{ is bounded in } L^1\Big(0,T;L^\infty(\Omega;R^3)\Big).$$

One deduces the same properties for the limit.

Notice how far this estimate is from that which would give well defined curves followed by particles along the flow.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

26. Monday March 15.

If U_0 is a characteristic velocity and L_0 is a characteristic length of a flow, then the corresponding REYNOLDS number $\frac{U_0\,L_0}{\nu}$ is adimensional; large ν or small R correspond to laminar flows, while small ν or large R correspond to turbulent flows. In the ocean, the characteristic lengths L_0 are large.

As u denotes a velocity, it has dimension $L\,T^{-1}$, where L denotes length and T denotes time, and the kinematic viscosity $\nu = \mu/\rho$ has the dimension L^2T^{-1} (while μ has dimension $M\,L^{-1}T^{-1}$, where M denotes mass). In dimension $N = 3$, the norm $|u|$ has the dimension $L^{5/2}T^{-1}$ (and therefore as ρ has dimension $M\,L^{-3}$, $\rho_0|u|^2$ has dimension $M\,L^2T^{-2}$, i.e. energy), the norm $||u||$ has the dimension $L^{3/2}T^{-1}$ (and $\mu||u||^2$ has the dimension $M\,L^2T^{-3}$, as energy dissipated per unit of time), and $|u|^{1/2}||u||^{1/2}$ has the dimension L^2T^{-1}, as the does $||u||_{L^3}$.

The term $\langle B(u,u), A\,u\rangle$, as an integral $\nu\int u\,\partial u\,\partial^2 u\,dx$ has the dimension L^5T^{-4}, while $|u|^{1/2}||u||^{1/2}|A\,u|^2$ has the dimension L^7T^{-5}, and therefore C_1 has the dimension $L^{-2}T$, i.e. $1/C_1$ has the same dimension L^2T^{-1} as ν. As the only parameter in the equation is ν (the term $1/\rho$ in front of $grad(p)$ is hidden, and as ρ is assumed constant, it is p/ρ_0 which we have called pressure!), it is natural to compare norms to that number, but that only makes sense for norms whose dimension is a power of L^2T^{-1}.

The limitations of the estimates shown before are due in part to the fact that one uses norms which give a global information on the solution and not a local information; the total kinetic energy at time t is seen by $\rho_0|u(t)|^2$ and the energy dissipated by viscosity between time 0 and T is seen by $\mu\int_0^T ||u(t)||^2\,dt$, but these norms do not tell if some regions corresponds to large velocities or to a large dissipation of energy (as we have assumed that ρ_0 and μ are independent of temperature, the energy dissipated by viscosity appears in the equation of balance of energy, which is decoupled from the equation of motion that we have been dealing with up to now).

A different approach, which I initiated in 1979 for a different class of equations, consists in avoiding the semi-group approach where one deals with functional spaces which are functions in x alone and one defines the domain of a nonlinear operator, and instead deals with functional spaces in (x,t) adapted to the equation. In the class of discrete velocity models in kinetic theory, which are supposed to simplify the BOLTZMANN equation, there is a particular model attributed to BROADWELL (Renée GATIGNOL attributes this kind of model to MAXWELL); in two dimensions (x,y) it is

$$\begin{aligned}
&\frac{\partial u_1}{\partial t} + \frac{\partial u_1}{\partial x} + (\alpha\,u_1u_2 - \beta\,u_3u_4) = 0 \text{ in } R^2\times(0,T), u_1(x,y,0) = u_{01}(x,y) \text{ in } R^2\\
&\frac{\partial u_2}{\partial t} - \frac{\partial u_2}{\partial x} + (\alpha\,u_1u_2 - \beta\,u_3u_4) = 0 \text{ in } R^2\times(0,T), u_2(x,y,0) = u_{02}(x,y) \text{ in } R^2\\
&\frac{\partial u_3}{\partial t} + \frac{\partial u_3}{\partial y} - (\alpha\,u_1u_2 - \beta\,u_3u_4) = 0 \text{ in } R^2\times(0,T), u_3(x,y,0) = u_{03}(x,y) \text{ in } R^2\\
&\frac{\partial u_4}{\partial t} - \frac{\partial u_4}{\partial y} - (\alpha\,u_1u_2 - \beta\,u_3u_4) = 0 \text{ in } R^2\times(0,T), u_4(x,y,0) = u_{04}(x,y) \text{ in } R^2,
\end{aligned}$$

where u_1,u_2,u_3,u_4 denote the density of particles at (x,y,t); these particles have all the same mass but their velocities are respectively $(+1,0),(-1,0),(0,+1),(0,-1)$ (Jim GREENBERG denotes the unknowns l,r,u,d, for left, right, up, down); α,β are positive parameters related to probability of collisions, which are usually taken equal to $1/\varepsilon$, where ε is related to a mean free path between collisions (BROADWELL was actually interested in the formal fluid limit $\varepsilon \to 0$, and there are plenty of open questions in that direction). Local existence for data in $L^\infty(R^2)$ is standard (locally LIPSCHITZ perturbation of a linear semigroup), and if the data are nonnegative the solution is nonnegative. The model conserves mass and momentum (density of mass is $u_1+u_2+u_3+u_4$, density of momentum is (u_1-u_2, u_3-u_4)), and also kinetic energy as it is proportional to mass (so there is no temperature for this model). An analog of the H-theorem of BOLTZMANN holds, and there is an entropy which decreases (density of entropy is $u_1\log(u_1)+u_2\log(u_2)+u_3\log(u_3)+u_4\log(u_4)$). Takaaki NISHIDA and I (independently) have noticed that there is a global existence of a solution for small nonnegative data in $L^2(R^2)$, and it is useful to notice that the $L^2(R^2)$ norm in invariant by scaling: if $U = (u_1, u_2, u_3, u_4)$ is a solution, then V defined by $V(x, y, t) = \lambda U(\lambda x, \lambda y, \lambda t)$ is also a solution for any $\lambda > 0$, and the norm of the initial data in $L^2(R^2; R^4)$ is the same for U or V; my argument generalized what I had done for the one dimensional case, which I describe now.

If u_3 and u_4 are equal at time 0 and independent of y then they stay equal and independent of y for $t > 0$ (also for $t < 0$ as long as the solution exists, but nonnegativity is only conserved when t increases); one considers then the simplified model where $\alpha = \beta = 1$

$$\begin{aligned} &u_t + u_x + u\,v - w^2 = 0 \text{ in } R \times (0, \infty), u(x, 0) = u_0(x) \text{ in } R \\ &v_t - v_x + u\,v - w^2 = 0 \text{ in } R \times (0, \infty), v(x, 0) = v_0(x) \text{ in } R \\ &w_t - u\,v + w^2 = 0 \text{ in } R \times (0, \infty), w(x, 0) = w_0(x) \text{ in } R. \end{aligned}$$

In 1975, in collaboration with Michael CRANDALL, we had proved global existence for bounded nonnegative data by using finite propagation speed, the entropy estimate as a compactness argument in L^1, and a crucial result that MIMURA and Takaaki NISHIDA had just published, where they had shown that for small nonnegative data in L^1 and arbitrary bound in L^∞, the L^∞ estimate was controlled for all t. As our argument could only be used in one dimension (as the generalization of the estimate of MIMURA and NISHIDA to more than one dimension was unlikely), I thought of using more physical spaces than L^∞, and I thought that BMO was a good substitute, as it controls the portion of the mass which is out of equilibrium, but I could not get my colleague Yves MEYER to help me, and I never went forward with this idea. In 1979, I was wondering which discrete velocity models of kinetic theory were stable by weak convergence (as I had noticed that the CARLEMAN model was not, although it is not really a model of kinetic theory as it does not conserve momentum), and I found that (apart from the affine case) it only happened in one space dimension for models of the form

$$\frac{\partial u_i}{\partial t} + C_i \frac{\partial u_i}{\partial x} + \sum_{j,k} A_{ijk} u_j u_k + affine(u) = 0 \text{ in } R \times (0, \infty), u_i(x, 0) = u_{0i}(x) \text{ in } R,$$

where the interaction coefficients (with $A_{ijk} = A_{ikj}$ for all i, j, k) satisfy the condition

$$C_j = C_k \text{ implies } A_{ijk} = 0 \text{ for all } i.$$

I looked into the existence of solutions for these models, as I had found them in connection with a question of Compensated Compactness, for which better bounds were known or conjectured (a topic which I call now Compensated Integrability in order to point out the difference with Compensated Compactness, because I had noticed that in an interesting article of Ronald COIFMAN, Pierre-Louis LIONS, Yves MEYER and Stephen SEMMES using HARDY spaces, they had wrongly claimed to improve the Compensated Compactness theory, while they were actually improving one of my argument of Compensated Integrability based on using LORENTZ spaces); I noticed a simple trick, which gave global existence (from $-\infty$ to $+\infty$) for small data in $L^1(R)$.

Let $V_c = \{u : u_t + c\,u_x \in L^1(R^2), u(\cdot, 0) \in L^1(R)\}$ and $W_c = \{u : |u(x, t)| \le U(x - c\,t) \text{ a.e.}, U \in L^1(R)\}$, then $V_c \subset W_c$ and if $u \in V_c, v \in V_{c'}$ with $c \ne c'$, then $u\,v$ belongs to $L^1(R^2)$, and $|c - c'|\,||u\,v||_{L^1(R^2)} \le ||u||_{W_c} ||v||_{W_{c'}}$. Then using an iterative scheme in $V = \prod_i V_{c_i}$ one finds a strict contraction in a small ball centered at 0, and this gives the global existence for small data in $L^1(R)$. Therefore one does not try to define the domain of the nonlinear operator, one finds that all products $u_j u_k$ appearing in the equation belong to $L^1(R^2)$ and therefore by FUBINI's theorem, for almost every t the product $u_j u_k$ belong to $L^1(R)$.

For the two dimensional case, it is u_i^2 which belongs to a space like V_c, and for example the analog of the space W_c are $|u_1(x, y, t)| \le U_1(x - t, y), |u_2(x, y, t)| \le U_2(x + t, y), |u_3(x, y, t)| \le U_3(x, y - t), |u_4(x, y, t)| \le U_4(x, y + t)$, with $U_1, U_2, U_3, U_4 \in L^2(R^2)$; one has then to show that $u_1 u_3 u_4 \in L^1(R^3)$, and this is analogous to the trick used in the proof of SOBOLEV imbedding theorem in the methods of Emilio GAGLIARDO and of Louis NIRENBERG.

I have not found how to use this idea for NAVIER-STOKES equation, but there has been some application to BOLTZMANN equation or FOKKER-PLANCK equation by my student Kamel HAMDACHE, and if the idea to use $|f(x,v,t)| \leq F(x - v\,t, v)$ was clear, it was not obvious how to choose F, and Kamel HAMDACHE extended an initial result of Reinhard ILLNER and SHINBROT, who had taken $F(\xi,v) = M\,e^{-\alpha|\xi|^2}$.

A second idea, is to use pointwise estimates with maximal functions, and the possibility of using that idea for STOKES equation only occurred to me two years ago, but I have not found a way to handle the question of transport and extend it to NAVIER-STOKES equation. I had learned the trick in an argument of Lars HEDBERG, reproduced by Haïm BREZIS and Felix BROWDER for a question of truncation (which some justly call the HEDBERG truncation method); after finding how to use the trick for the heat equation or STOKES equation two years ago (during a meeting dedicated to Jindrich NEČAS in Lisbon), I exchanged e-mail with Lars HEDBERG in order to learn about the origin of the idea, which is his but he pointed out an earlier result of Lennart CARLESON in the late 60s and a result of Elias STEIN in his book from the early 70s.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

27. Wednesday March 17.

At the moment, uniqueness of weak solutions of 3-dimensional incompressible NAVIER-STOKES equation is an open problem. Jean LERAY had conjectured that there could be point singularities of the equation and he had imagined that they could be self similar solutions, of the form $u(x,t) = \frac{1}{\sqrt{2a(T-t)}}U\big(\frac{x}{\sqrt{2a(T-t)}}\big)$; he also thought that it was related to turbulence, but the prevailing ideas on turbulence now are those imagined much later by KOLMOGOROV (I do not think that KOLMOGOROV was absolutely right, but LERAY's idea does not fit well with what I understand about effective properties of microstructures). Recently Jindřich NEČAS, M. RUŽICKA and Vladimir ŠVERÁK have shown that the self similar solutions imagined by LERAY cannot have $U \in L^3(R^3; R^3)$.

Measuring the HAUSDORFF dimension of the singular set of a solution has been a way to determine how far it is from being smooth, and at the moment the best result has been obtained by Luis CAFFARELLI, Robert KOHN and Louis NIRENBERG; they showed that the 1-dimensional HAUSDORFF measure of the singular set is 0, and therefore it cannot be a point singularity moving along a nice curve; Michael STRUWE has obtained a similar result for the stationary case in 5 dimensions.

Jean LERAY had already obtained results bounding the HAUSDORFF dimension of the singular set in t alone; I have not read his argument, but I think that it is based on the already mentioned differential inequality $(||u||^2)' \leq C||u||^6$ as follows. Let $\varphi = ||u||^2$, so we start with the information $\varphi \in L^1(0,T)$ and $\varphi' \leq a\varphi^3$; the differential inequality implies $(\varphi^{-2})' \leq -2a$ and therefore $\varphi(t) \leq \varphi(0)\big(1 - 2a\,t\,\varphi(0)^2\big)^{-1/2}$ as long as $1 - 2a\,t\,\varphi(0)^2 > 0$, and therefore the blow up time satisfies $T_c \geq \frac{1}{2a\,\varphi(0)^2}$. One divides $(0,T)$ into N equal intervals $I_1, \ldots, I_N$, of length $\tau = T/N$; if $j < N$ and $\int_{I_j} \varphi(t)\,dt < \sqrt{\tau/4a}$, then there is a point $x_j \in I_j$ such that $\varphi(x_j) < \sqrt{1/4\tau\,a}$; one deduces that the blow up time after x_j is at least $\frac{1}{2a\,\varphi(x_j)^2} > 2\tau$, and therefore the next interval I_{j+1} is free of singularities. The singular set is then contained in I_1 and the union of all the I_{j+1} for an index j such that $\int_{I_j} \varphi(t)\,dt > \sqrt{\tau/4a}$, but as $\sum_j meas(I_{j+1})^{1/2} \leq 2\sqrt{a}\sum_j \int_{I_j} \varphi(t)\,dt \leq 2\sqrt{a}||f||_{L^1(0,T)}$, one deduces that the $1/2$ HAUSDORFF dimension of the singular set (in t alone) is finite by letting N tend to infinity; by applying the argument to a family of intervals containing the set where φ takes large values, one deduces that the $1/2$ HAUSDORFF dimension of the singular set of $||u||$ is 0.

Estimating the HAUSDORFF dimension of the singular set in (x,t) relies on local regularity results, but CAFFARELLI, KOHN and NIRENBERG used the regularizing effect of the heat kernel, considering the pressure as given by an equation; for scaling, instead of balls in the (x,t) space, they used flat cylinders, scaling in ε in x and ε^2 in t. Taking the divergence of the equation, one has

$$-\Delta\,p = \sum_{i,j=1}^{3} \frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i},$$

and there are special results, which I call compensated integrability results, for that equation. Before reviewing some of that information, I want to describe a different approach, which is based on using local estimates in terms of maximal functions, a subject which I had understood from an example used by Lars HEDBERG.

For a function $f \in L^1_{loc}(R^N)$, the maximal function of f, denoted Mf, is defined by

$$(Mf)(x) = \sup_{r>0} \frac{\int_{B(x,r)} |f(y)|\,dy}{\int_{B(x,r)} dy}.$$

At every LEBESGUE point, and therefore almost everywhere, one has $|f(x)| \leq (Mf)(x)$. If $f \in L^\infty$, one has $||Mf||_{L^\infty} \leq ||f||_{L^\infty}$, but if $f \in L^1(R^N)$ and $f \neq 0$, then $Mf \notin L^1(R^N)$. However, using a simple covering argument, one can show that $meas\{x : Mf(x) \geq t\} \leq \frac{C||f||_{L^1}}{t}$ for every $t > 0$, and using an interpolation argument one deduces that for $p > 1$ one has $||Mf||_{L^p} \leq C_p||f||_{L^p}$ for every $f \in L^p(R^N)$, and $C_p \to \infty$ as

$p \to 1$ (HARDY & LITTLEWOOD, WIENER). In 1972, Lars HEDBERG used an inequality giving a pointwise estimate of a convolution product by a radial function in terms of the maximal function, i.e.

$$\text{for a radial function } f, |(f \star g)(x)| \leq C\, Mg(x).$$

His function f was special ($1/r^\lambda$ for $0 < r < r_0$, 0 for $r > r_0$), and he used a dyadic decomposition, but after studying his proof a few years ago I realized that one could write the inequality above for any f radial and nonincreasing, in which case $C = ||f||_{L^1(R^N)}$, and more generally, assuming f radial and smooth in order to avoid technical details, with $C = ||r\, grad(f)||_{L^1(R^N)}$. Two years ago I thought of using this type of inequality for estimating various norms of solutions of the heat equation or STOKES equation, and I thought that if one knew how to extend this type of inequality when transport terms are present, it could be quite useful for improving the abstract approach to solutions of NAVIER-STOKES equation. I had then an e-mail exchange with Lars HEDBERG in order to learn about the origin of the idea that he had used, and he said that it was his program to show that many classical global inequalities can actually be improved into pointwise inequalities using maximal functions, but he saw the first example of this kind in a result of Lennart CARLESON of 1967, showing that the solution of $\Delta\, u = 0$ in the unit disc, $u = \varphi$ on the boundary satisfied a bound $|u(x)| \leq C\, M\varphi\big(\frac{x}{|x|}\big)$ for $x \neq 0$ (maybe with $C = 1$). When I pointed out the inequalities which I had shown for the heat kernel, it reminded him of another result, shown by Elias STEIN in his 1970 book on singular integrals, where he shows the idea for the POISSON integrals (same as CARLESON but for a half space instead of a disc), but he does add a remark showing that my bound $C = ||r\, grad(f)||_{L^1(R^N)}$ is not the good one: he does not assume f to be radial but that $|f| \leq \psi$ with ψ radial as I had done, but he only considered ψ nonincreasing and integrable, and he proved $C = \inf_\psi ||\psi||_{L^1(R^N)}$, where the infimum is taken on the radial nonincreasing ψ larger or equal than $|f|$ almost everywhere (I had not taken advantage of the remark that the bound $||r\, grad(f)||_{L^1(R^N)}$ may decrease while replacing f by a larger function, which automatically gives the bound used by STEIN).

The proof is easy once one realizes that a radial nonincreasing function is an integral with nonnegative coefficient of characteristic functions of balls centered at 0, or simply that it is a limit in $L^1(R^N)$ norm of finite combinations with positive coefficients of characteristic functions of balls centered at 0; by linearity it suffices then to prove the result for f being the characteristic function χ of a ball $B(0,\rho)$ centered at 0, but then the convolution by χ does compute $\int_{B(x,\rho)} |g(x)|\, dx$ which is bounded by $Mg(x)$ multiplied by the volume of $B(0,\rho)$.

If one applies the idea to the solution of the heat equation

$$\frac{\partial u}{\partial t} - \Delta\, u = 0 \text{ in } R^N \times (0,T);\; u(x,0) = v(x) \text{ in } R^N,$$

then the solution is

$$u(x,t) = \int_{R^N} E(x-y,t)v(y)\, dy, \text{ with } E \text{ given by } E(z,t) = C_N t^{-N/2} e^{-|z|^2/4t} \text{ on } R^N \times (0,\infty),$$

and C_N is such that the elementary solution E satisfies $\int_{R^N} E(x,t)\, dx = 1$ for any (or all) $t > 0$. As $E(\cdot,t)$ is radial decreasing with integral 1, one deduces that

$$|u(x,t)| \leq Mv(x) \text{ a.e. } x \in R^N, \text{ for all } t > 0.$$

As $|u(\cdot,t)| \leq E(\cdot,t) \star |v|$, if χ is the characteristic function of the ball of radius ρ, then $\chi \star E(\cdot,t)$ is also radial decreasing and one deduces the more precise inequality

$$Mu(x,t) \leq Mv(x) \text{ a.e. } x \in R^N, \text{ for all } t > 0.$$

This inequality cannot be deduced from global bounds like $\int_{R^N} \varphi\big(u(x,t)\big)\, dx \leq \int_{R^N} \varphi\big(v(x)\big)\, dx$ for every convex function φ, or simply the inequalities $||u(\cdot,t)||_{L^p(R^N)} \leq ||v||_{L^p(R^N)}$ for all $p \in [1,\infty]$; the maximal function changes if one replaces v by an equimeasurable function. If one applies the idea to derivatives, then one obtains

$$M(D^\alpha u)(x,t) \leq C_\alpha t^{-|\alpha|/2} Mv(x) \text{ a.e. } x \in R^N, \text{ for all } t > 0, \text{ for all derivarive } D^\alpha \text{ of order } |\alpha|,$$

and STEIN had noticed the analogous inequality for the POISSON integrals.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

28. Friday March 19.

A few years ago, Victor MIZEL had asked me if I knew about HEDBERG's truncation method. At that time I did not know what was meant by that term, and as François MURAT is a good friend of Lars HEDBERG, I had asked him about it, and he had sent me a few pages of an article of Haïm BREZIS and Felix BROWDER who had reproduced HEDBERG's proof in an appendix (they had read it in an article by WEBB, I believe, to whom HEDBERG had taught the method); I carried the copy for a while before looking at it, a Summer on the beach, and saw what the key ideas were. I did explain in the lounge afterward what the key convolution idea was, but later I also looked at a paper of Emilio ACERBI and Nicola FUSCO, because I had understood from a comment of Irene FONSECA that there might be similarities between the methods; the key point in their argument was a result of (Fon Che) LIU, which I could derive easily by HEDBERG's trick, and I had asked my student Sergio GUTIERREZ to work on an extension. Although these results are not directly related to the questions of fluids, I discuss them briefly as they may be useful in order to obtain a clear picture of what the methods are.

HEDBERG's truncation method permits to approach a function in $u \in W^{m,p}(R^N)$ by a sequence $u_n \in L^\infty(R^N) \bigcap W^{m,p}(R^N)$ such that $u_n(x)u(x) \geq 0$ a.e., and as it uses CALDERÓN-ZYGMUND theorem one must have $p > 1$ (for $p > N/m$ there is nothing to prove as $W^{m,p}(R^N) \subset L^\infty(R^N)$). In order to simplify, I show how it works for approaching a function with second derivatives in $L^p(R^N)$ with $1 < p < N/2$. One solves $-\Delta\, v = |\Delta\, u|$, which gives $\partial_j\partial_k v \in L^p(R^N)$ for all j,k by CALDERÓN-ZYGMUND theorem, and $v \geq |u|$ by the maximum principle (in general one takes convolutions by powers of $1/r$), and one defines u_n by $u_n(x) = u(x)\varphi\big(v(x)/n\big)$, where φ is smooth and is equal to 1 on $[0,1/2]$ and 0 on $[1,\infty)$, showing that $|u_n| \leq n$. Then $\partial_j\partial_k u_n = \partial_j\partial_k u\,\varphi\big(v(x)/n\big) + \partial_j u\,\varphi'\big(v(x)/n\big)\partial_k v/n + \partial_k u\,\varphi'\big(v(x)/n\big)\partial_j v/n + u\,\varphi''\big(v(x)/n\big)\partial_j v\partial_k v/n^2 + u\,\varphi'\big(v(x)/n\big)\partial_j\partial_k v/n$, and the first term converges to $\partial_j\partial_k u$ and the last term converges to 0 by LEBESGUE dominated convergence theorem (using $|u| \leq v$); HEDBERG's method is based on the fact that both $\partial_j u$ and $\partial_j v$ are bounded (pointwise) by $C\,v^{1/2}(M|\Delta\, u|)^{1/2}$, so that the other terms can also be treated in the same way. Indeed one has $\partial_j u = \partial_j E \star \Delta\, u$ and $\partial_j v = \partial_j E \star |\Delta\, u|$, both bounded by $C/r^{N-1} \star |\Delta\, u|$, which is cut into two parts; the first one is $f \star |\Delta\, u|$ with $f = C/r^{N-1}$ for $0 < r < \delta$ and 0 for $r > \delta$, and this term is bounded by $C\,\delta(M|\Delta\, u|)$ by using the argument on convolution with radial functions, and the second is bounded by $(C/\delta)E \star |\Delta\, u| = C\,v/\delta$; then the best δ is chosen (depending upon x).

The estimate of LIU is about $|u(x)-u(y)| \leq C|x-y|\big(M|grad(u)|(x)+M|grad(u)|(y)\big)$ for a.e. $x,y \in R^N$; one starts from $u(x) - u(y) = \int_0^1 \big(grad(u)\big(x+t(y-x)\big).y-x\big)\,dt$, from which one deduces

$$\int_{B(x,\rho)} \frac{|u(y)-u(x)|}{|x-y|}\,dy \leq \int_0^1 \int_{B(x,\rho)} \Big|grad(u)\Big(x+t(y-x)\Big)\Big|\,dt\,dy,$$

and one uses the change of variable $z = -t(y-x)$; the variable t varies from $|z|/\rho$ to 1, and the last integral is $(N-1)\int_{B(x,\rho)} |grad(u)(x-z)|(|z|^{1-N}-1)\,dz$, which is a convolution of $|grad(u)|$ by a radial decreasing function and is therefore bounded by $M|grad(u)|(x)$ multiplied by the L^1 norm of the radial function, which is the value obtained when one replaces $|grad(u)|$ by 1, i.e. the volume of $B(0,\rho)$. One integrates then $\frac{|u(y)-u(x_1)|}{|y-x_1|} + \frac{|u(y)-u(x_2)|}{|y-x_2|}$ on $A = B(x_1,\rho)\bigcap B(x_2\rho)$, and it is bounded by the integral on $B = B(x_1,\rho)\bigcup B(x_2\rho)$, i.e. by $M|grad(u)|(x_1) + M|grad(u)|(x_2)$ multiplied by twice the volume of $B(0,\rho)$; one choose $\rho = |x_1 - x_2|$ for example and one finds a $y \in A$ such that $\frac{|u(y)-u(x_1)|}{|y-x_1|} + \frac{|u(y)-u(x_2)|}{|y-x_2|} \leq C\big(M|grad(u)|(x_1) + M|grad(u)|(x_2)\big)$ and as $|y-x_1|,|y-x_2| \leq |x_1-x_2|$ it implies the desired inequality.

The inequality that Lars HEDBERG used in his truncation method is reminiscent of GAGLIARDO-NIRENBERG inequality, and indeed one can find a pointwise version of GAGLIARDO-NIRENBERG inequality, as I checked last December, only to discover a week or two after that Patrick GÉRARD had made the same observation; however his proof is different from mine, relying on a dyadic decomposition in the style of LITTLEWOOD-PALEY, while mine is more elementary and uses a parametrix. Let E be the usual elementary solution of $-\Delta$, i.e. $E(x) = C_N/|x|^{N-2}$ if $N > 2$, or $E(x) = C\log(|x|)$ for $N = 2$; let $\varphi \in C_c^\infty(R^N)$ be such that $\varphi(x) = 1$ for $|x| \le 1$ and $\varphi(x) = 0$ for $|x| \ge 2$, and for $\alpha > 0$ let us consider the parametrix P_α defined by $P_\alpha(x) = E(x)\varphi(x/\alpha)$, so that $-\Delta\, P_\alpha = \delta_0 + g$ with $g(x) = \psi(x/\alpha)/|x|^N$, with $\psi \in C_c^\infty(R^N)$. Taking the convolution by $\partial_j u$ gives $\partial_j P_\alpha \star (-\Delta\, u) = \partial_j u + \partial_j g \star u$, and as $||r\, grad(\partial_j P_\alpha)||_{L^1(R^N)} = C\,\alpha$ and $||r\, grad(\partial_j g)||_{L^1(R^N)} = C/\alpha$, one deduces $|\partial_j u| \le C\, M(\Delta\, u)\alpha + C\, M(u)/\alpha$, and taking then the best α (depending on x) gives $|grad(u)|^2 \le C\, Mu\, M(\Delta\, u)$.

I want to finish with some remarks on compensated integrability.

In the Summer 1982, at a meeting in Oxford, I heard about a result of WENTE, and back in Paris I derived a proof by interpolation which I mentioned around. If $u, v \in H^1(R^2)$ then one cannot assert that $u_x v_y - u_y v_x$ belongs to $H^{-1}(R^2)$ because $L^1(R^2)$ is not imbedded in $H^{-1}(R^2)$ as $H^1(R^2)$ is not imbedded into $L^\infty(R^2)$; it does not follow either from writing that quantity as $(u\, v_y)_x - (u\, v_x)_y$ or $(u_x v)_y) - (u_y v)y$, which is the key to the sequential weak lower semicontinuity observed by MORREY and which I learned from by John M. BALL before extending this kind of property into the Compensated Compactness method with François MURAT. However, it is indeed true that $u_x v_y - u_y v_x \in H^{-1}(R^2)$, and WENTE also proved that if one solves the equation $-\Delta\, w = u_x v_y - u_y v_x$, then one also has $w \in H^1(R^2) \bigcap C_0(R^N)$.

I do not know what the original proof of WENTE was, but my first proof used interpolation and LORENTZ spaces. It would take us too far if I explained what interpolation of BANACH spaces is according to the theory developed by Jacques-Louis LIONS and Jaak PEETRE, and how the theory applied to L^1 and L^∞ creates the family of LORENTZ spaces; therefore I will just state the ideas for those readers who know these tools. First one uses the fact that $H^{1/2}(R^2) \subset L^{4,2}(R^2)$, as noticed by Jaak PEETRE (a result used was $||u||_{L^4(R^2)} \le C\, ||u||_{L^2(R^2)}^{1/2} ||grad(u)||_{L^2(R^2)}^{1/2}$, and it is not as precise because a theorem of LIONS and PEETRE asserts that this statement is equivalent to the fact that the interpolation space $\big(H^1(R^2), L^2(R^2)\big)_{1/2,1}$, which is smaller than $H^{1/2}(R^2)$, is included in $L^4(R^2)$, which is bigger than $L^{4,2}(R^2)$). Then one uses the fact that the product of two functions in $L^{4,2}(R^2)$ is in $L^{2,1}(R^2)$. This shows that $B(u,v) = u_x v_y - u_y v_x$ is a sum of derivatives of functions in $L^{2,1}(R^2)$, in the case where $u \in H^{1/2}(R^2)$ and $v \in H^{3/2}(R^2)$ by using the formula $(u\, v_y)_x - (u\, v_x)_y$, or in the case where $u \in H^{3/2}(R^2)$ and $v \in H^{1/2}(R^2)$ by using the formula $(u_x v)_y) - (u_y v)y$; by another theorem of LIONS and PEETRE on bilinear mappings the same property is then true for $u, v \in H^1(R^2)$. Then by CALDERÓN-ZYGMUND theorem and interpolation, one finds that w has its two partial derivatives in $L^{2,1}(R^2)$ (a smaller space than $L^2(R^2)$), and this implies that $w \in C_0(R^2)$.

In 1984, I described a second method which extends immediately to more general situations similar to those found in Compensated Compactness theory for the quadratic forms which are sequentially weakly continuous; the method uses FOURIER transform and interpolation, but not CALDERÓN-ZYGMUND theorem, and the results are slightly different. The example that I had chosen was the equation for the pressure in NAVIER-STOKES equation in 2 dimensions, but that is similar to the previous example. Using x_1, x_2, instead of x, y, one has $|\xi|^2 \mathcal{F}w(\xi) = \int_{R^2} B(\xi-\eta,\eta)\mathcal{F}u(\xi-\eta)\mathcal{F}v(\eta)\, d\eta$, where $B(\zeta,\eta) = \zeta_1\,\eta_2 - \zeta_2\,\eta_1$, but using the fact that $B(\xi,\xi) = 0$ for all ξ, one has $B(\xi-\eta,\eta) = B(\xi,\eta) = B(\xi,\eta-\xi)$, so that one has the two bounds $|B(\xi-\eta,\eta)| \le C\,|\xi|\,|\eta|$ and $|B(\xi-\eta,\eta)| \le C\,|\xi|\,|\xi-\eta|$ and therefore $|B(\xi-\eta,\eta)| \le C\,|\xi|\,|\eta|^{1/2}|\xi-\eta|^{1/2}$. One deduces that $|\xi|\,|\mathcal{F}w| \le C\,|\xi|^{1/2}|\mathcal{F}u| \star |\xi|^{1/2}|\mathcal{F}v|$, but as $|\xi|^{-1/2} \in L^{4,\infty}(R^2)$, one deduces that $\xi\,\mathcal{F}u \in L^2(R^2)$ implies $|\xi|^{1/2}|\mathcal{F}u| \in L^{4/3,2}(R^2)$ and the convolution product of two functions in $L^{4/3,2}(R^2)$ is in $L^{2,1}(R^2)$. In particular $|\xi|\,|\mathcal{F}w| \in L^2(R^2)$ but I do not know how to compare the informations $grad(w) \in L^{2,1}(R^2; R^2)$ and $|\xi|\mathcal{F}w \in L^{2,1}(R^2; R^2)$. As $|\xi|^{-1} \in L^{2,\infty}(R^2)$, one deduces that $\mathcal{F}w \in L^1(R^2)$, and then $w \in \mathcal{F}L^1(R^2) \subset C_0(R^2)$.

My approach has been slightly improved by Ronald COIFMAN, Pierre-Louis LIONS, Yves MEYER and Stephen SEMMES, using the HARDY spaces $\mathcal{H}^1$; their result has the advantage of showing that the second derivatives of w belong to $\mathcal{H}^1(R^2)$, and therefore $w \in W^{2,1}(R^2)$; however, contrary to what they have claimed, many applications do not require their improvement and can be obtained by using my second method.

So much for technical details around NAVIER-STOKES equation. Let us go back to Oceanography!

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

29. Monday March 29.

Up to now, I have not mentioned the effect of CORIOLIS force due to the rotation of the Earth; it is small but it does have some effect.

Assume that we have a first frame, called fixed, in which NEWTON's law of Classical Mechanics, *force = mass × acceleration* applies, and let us see what it implies for the equation in a moving frame. Let $x(t)$ be the position of a material point in the fixed frame, and let $\xi(t)$ be the position of the same point in the moving frame; let $a(t)$ be the position of the origin of the moving frame and let $P(t)$ be the rotation which maps the basis of the initial frame into the basis of the moving frame, so that one has

$$x(t) = a(t) + P(t)\xi(t).$$

As $P(t)^T P(t) = I$, if $'$ denotes the derivative with respect to t, one has $P'(t)^T P(t) + P(t)^T P'(t) = 0$, and so if one defines $B(t)$ by $P'(t) = P(t)B(t)$, one obtains $B(t)^T + B(t) = 0$, and therefore, as we work in R^3, there exists a vector $\Omega(t)$ such that $B(t)x = \Omega(t) \times x$ for every $x \in R^3$; one deduces

$$x'(t) = a'(t) + P(t)\Big(\xi'(t) + \Omega(t) \times \xi(t)\Big),$$

and

$$x''(t) = a''(t) + P(t)\Big(\Omega'(t) \times \xi(t)\Big) + P(t)\Big[\xi''(t) + 2\Omega(t) \times \xi'(t) + \Omega(t) \times \Big(\Omega(t) \times \xi(t)\Big)\Big].$$

In the case of the rotation of the Earth, one considers that $\Omega'(t) = 0$. The term $2\Omega(t) \times \xi'(t)$ is the CORIOLIS acceleration (although LAGRANGE had introduced it in 1778-79 in his studies of tides, while CORIOLIS's work dates from 1835). If one uses the formula $a \times (b \times c) = (a.c)b - (a.b)c$, one deduces that $\Omega \times (\Omega \times \xi) = (\Omega.\xi)\Omega - |\Omega|^2\xi$, and therefore the term $\Omega \times (\Omega \times \xi)$, which is related to the centrifugal acceleration with $a''(t)$, derives from a potential, which changes slightly the gravitation potential, creating the *geopotential*. For the rotation of the Earth, $|\Omega| = \frac{2\pi}{86400} \approx 3.6\,10^{-5}$, so that at the equator the centrifugal acceleration is about $8.3\,10^{-3}$, less than one thousandth of the acceleration of gravity.

Because the term $\Omega \times (\Omega \times \xi)$ is a gradient, it changes only what p is, and therefore adding the CORIOLIS term $\Omega \times u$ in NAVIER-STOKES equation does not change much in the proofs that we have seen, because this term is orthogonal to u and therefore does not work, and the basic estimates are the same as before.

The CORIOLIS force depends upon the velocity, in a way that reminds of Electromagnetism, where the LORENTZ force acting on a charge ρ moving with velocity v in an electric field E and magnetic induction field B is $\rho\big(E + v \times B\big)$. The analogy goes further and it has been used in connection with MHD (Magnetohydrodynamics), at least by MOFFATT: in MHD the fluid is a plasma, which has electrical charges moving around, but the forces acting on a neutral fluid are very similar, as we will see by computing $u \times curl(u)$ in a domain of R^3.

Let ε_{ijk} be the totally antisymmetric tensor, which is 0 if two of the indices i, j, k, are equal, and equal to the signature of the permutation $123 \mapsto ijk$ in other cases, i.e. $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$ and $\varepsilon_{321} = \varepsilon_{213} = \varepsilon_{132} = -1$. Then the definition of the exterior product of two vectors in R^3 is

$$c = a \times b \text{ means } c_i = \sum_{j,k} \varepsilon_{ijk} a_j b_k,$$

and the *curl* of a function u, sometimes denoted $\nabla \times u$, is defined by

$$\Big(curl(u)\Big)_i = \sum_{j,k} \varepsilon_{ijk} \frac{\partial u_k}{\partial x_j}.$$

One has

$$\Big(u\times curl(u)\Big)_i=\sum_{j,k}\varepsilon_{ijk}u_j\Big(\sum_{l,m}\varepsilon_{klm}\frac{\partial u_m}{\partial x_l}\Big)=\sum_{j,k}\varepsilon_{ijk}u_j\Big(\varepsilon_{kij}\frac{\partial u_j}{\partial x_i}+\varepsilon_{kji}\frac{\partial u_i}{\partial x_j}\Big)$$
$$=\sum_{j,k}\varepsilon_{ijk}^2u_j\Big(\frac{\partial u_j}{\partial x_i}-\frac{\partial u_i}{\partial x_j}\Big)=\sum_{j\neq i}u_j\Big(\frac{\partial u_j}{\partial x_i}-\frac{\partial u_i}{\partial x_j}\Big)=\sum_j u_j\frac{\partial u_j}{\partial x_i}-\sum_j u_j\frac{\partial u_i}{\partial x_j},$$

and therefore

$$\sum_j u_j\frac{\partial u_i}{\partial x_j}=\Big(u\times curl(-u)\Big)_i+\frac{1}{2}\frac{\partial|u|^2}{\partial x_i},$$

so that NAVIER-STOKES equation becomes

$$\frac{\partial u}{\partial t}-\nu\Delta\,u+u\times curl(-u)+grad\Big(\frac{p}{\rho_0}+\frac{|u|^2}{2}\Big)=0,\ div(u)=0,$$

and CORIOLIS acceleration justs adds 2Ω to $curl(-u)$.

In the case $\nu=0$, corresponding to EULER equation, one sees that a stationary irrotational flow (i.e. satisfying $curl(u)=0$), corresponds to $\frac{p}{\rho_0}+\frac{|u|^2}{2}=constant$ (BERNOULLI's law); one also sees that in the all space R^3 the helicity $\big(u.curl(u)\big)$ is conserved ($curl$ is a symmetric operator); this was first observed by Jean-Jacques MOREAU (and also by someone else, whose name I do not remember), and MOFFATT has given an interpretation of this quantity in terms of linking of vorticity lines (in order to avoid boundary conditions, the result is considered in the whole space, as I do not care much for unrealistic periodic conditions). As the quantity integrated is not positive, the conservation of helicity has not helped for questions of global existence or smoothness of solutions of NAVIER-STOKES equation in 3 dimensions.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

30. Wednesday March 31.

I want to derive now the equation describing the evolution of the vorticity in two dimensions, and then in three dimensions, as a consequence of NAVIER-STOKES equation.

In two dimensions the NAVIER-STOKES equation, with zero exterior forces (the gravitational force being included in the pressure term) is

$$\begin{aligned}
&\frac{\partial u_1}{\partial t} + u_1\frac{\partial u_1}{\partial x_1} + u_2\frac{\partial u_1}{\partial x_2} - \nu\Delta\, u_1 + \frac{1}{\rho_0}\frac{\partial p}{\partial x_1} = 0\\
&\frac{\partial u_2}{\partial t} + u_1\frac{\partial u_2}{\partial x_1} + u_2\frac{\partial u_2}{\partial x_2} - \nu\Delta\, u_2 + \frac{1}{\rho_0}\frac{\partial p}{\partial x_2} = 0\\
&div(u) = 0,
\end{aligned}$$

and the vorticity is the scalar quantity

$$\omega = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}.$$

Applying $-\frac{\partial}{\partial x_2}$ to the first equation, $\frac{\partial}{\partial x_1}$ to the second equation and adding, one finds that the vorticity ω satisfies the equation

$$\frac{\partial \omega}{\partial t} + u_1\frac{\partial \omega}{\partial x_1} + u_2\frac{\partial \omega}{\partial x_2} - \nu\Delta\omega = 0,$$

because the pressure disappears and the supplementary terms coming from the first equation are $-\frac{\partial u_1}{\partial x_2}\frac{\partial u_1}{\partial x_1} - \frac{\partial u_2}{\partial x_2}\frac{\partial u_1}{\partial x_2} = -\frac{\partial u_1}{\partial x_2}div(u)$, while the supplementary terms coming from the second equation are $\frac{\partial u_1}{\partial x_1}\frac{\partial u_2}{\partial x_1} + \frac{\partial u_2}{\partial x_1}\frac{\partial u_2}{\partial x_2} = \frac{\partial u_2}{\partial x_1}div(u)$.

In three dimensions the computation is a little more involved; the NAVIER-STOKES equation, with zero exterior forces, is

$$\begin{aligned}
&\frac{\partial u_i}{\partial t} + \sum_{j=1}^3 u_j\frac{\partial u_i}{\partial x_j} - \nu\Delta\, u_i + \frac{1}{\rho_0}\frac{\partial p}{\partial x_i} = 0 \text{ for } i = 1,2,3,\\
&div(u) = 0,
\end{aligned}$$

and the vorticity is the vector valued quantity

$$\omega = curl(u), \text{ i.e. } \omega_i = \sum_{j,k=1}^3 \varepsilon_{ijk}\frac{\partial u_k}{\partial x_j} \text{ for } i = 1,2,3.$$

The equation for ω is

$$\frac{\partial \omega_i}{\partial t} + \sum_{j=1}^3 u_j\frac{\partial \omega_i}{\partial x_j} - \sum_{j=1}^3 \omega_j\frac{\partial u_i}{\partial x_j} - \nu\Delta\,\omega_i = 0 \text{ for } i = 1,2,3.$$

Indeed, the pressure disappears and the supplementary term in the equation for ω_i is $\sum_{j,k,l=1}^3 \varepsilon_{ijk}\frac{\partial u_l}{\partial x_j}\frac{\partial u_k}{\partial x_l}$. As only the terms where $\varepsilon_{ijk} \neq 0$ are useful, l takes the values i, j and k, and the sum is $\sum_{j,k=1}^3 \varepsilon_{ijk}\big(\frac{\partial u_i}{\partial x_j}\frac{\partial u_k}{\partial x_i} + \frac{\partial u_j}{\partial x_j}\frac{\partial u_k}{\partial x_j} + \frac{\partial u_k}{\partial x_j}\frac{\partial u_k}{\partial x_k}\big)$, and using $div(u) = 0$ it is $\sum_{j,k=1}^3 \varepsilon_{ijk}\big(\frac{\partial u_i}{\partial x_j}\frac{\partial u_k}{\partial x_i} - \frac{\partial u_i}{\partial x_i}\frac{\partial u_k}{\partial x_j}\big)$, which one writes $-\omega_i\frac{\partial u_i}{\partial x_i} + \sum_{j,k=1}^3 \varepsilon_{ijk}\frac{\partial u_i}{\partial x_j}\big(\frac{\partial u_k}{\partial x_i} - \frac{\partial u_i}{\partial x_k}\big)$, and for $j \neq i$ and k being the third index, the term $\frac{\partial u_k}{\partial x_i} - \frac{\partial u_i}{\partial x_k}$ is indeed $-\omega_j$.

Except in the whole space or the unrealistic peridoc case, there are no clear boundary conditions for the vorticity (vorticity is created at the boundary).

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

31. Friday April 2.

In a talk by Roger LEWANDOWSKI we have seen a model used in Oceanography: an horizontal fixed boundary is used to model the interface between Ocean and Atmosphere and a boundary condition is chosen there, which is supposed to take into account the turbulent kinetic energy (TKE) arising in a neighbourhood of the interface. Before looking at questions of averaging, I want to discuss the question of which types of boundary conditions are natural.

In studying the stationary STOKES equation, I have mentioned the approach of considering Linearized Elasticity and letting the LAMÉ coefficient λ tend to ∞, which forces the constraint $div(u) = 0$ at the limit, and the limit p of $-\lambda\, div(u)$ plays the role of a pressure. This similitude disappears as soon as one considers the evolution problems, because in Linearized Elasticity u denotes a displacement (whose gradient is supposed to be small), while for STOKES equation u denotes a velocity; the acceleration involves then a term in $\frac{\partial^2 u}{\partial t^2}$ in the first case and a term in $\frac{\partial u}{\partial t}$ in the second case.

I have initially discussed the homogeneous DIRICHLET condition $u = 0$ on $\partial\Omega$, and one may also consider the case of a nonhomogeneous DIRICHLET condition, $u = g$ on $\partial\Omega$: one first chooses a function equal to g on the boundary, and the difference satisfies the homogeneous DIRICHLET condition; one must then have characterized the space of traces of functions in $H^1(\Omega)$ (which is $H^{1/2}(\partial\Omega)$ in the good cases), but in the limiting case $\lambda \to \infty$, one needs to add a constraint. If $u \in H^1(\Omega; R^N)$ satisfies $div(u) = 0$, and $u = g$ on $\partial\Omega$, then integrating $div(u)$ in Ω gives $\int_{\partial\Omega}(g.n)\,dx = 0$, where n denotes the exterior normal to Ω; conversely if $g \in H^{1/2}(\partial\Omega; R^N)$ satisfies $\int_{\partial\Omega}(g.n)\,dx = 0$, then one first chooses $v \in H^1(\Omega; R^N)$ equal to g on the boundary and it remains to add a function $u \in H^1_0(\Omega; R^N)$ with $div(u) = -div(v)$, but as $\int_\Omega div(v)\,dx = 0$ because of the condition on g, a function u exists if Ω is smooth enough (bounded with $X(\Omega) = L^2(\Omega)$ for example).

The case of NEUMANN condition over all the boundary of Ω is of the form

$$-\sum_{j=1}^N \frac{\partial \sigma_{ij}}{\partial x_j} = f_i \text{ in } \Omega$$
$$\sum_{j=1}^N \sigma_{ij} n_j = g_i \text{ on } \partial\Omega,$$

and it requires the compatibility conditions

$$\int_\Omega f_i\,dx + \int_{\partial\Omega} g_i\,d\sigma = 0 \text{ for all } i$$
$$\int_\Omega \Big(\sum_{j,k} \varepsilon_{ijk} x_j f_k\Big)\,dx + \int_{\partial\Omega} \Big(\sum_{j,k} \varepsilon_{ijk} x_j g_k\Big)\,d\sigma = 0 \text{ for all } i,$$

which express the fact that the total force and the total torque acting on $\overline{\Omega}$ are 0. It is important to notice that this follows from the equilibrium equation and the symmetry of the stress tensor, so that it is true for Linearized Elasticity as well as for the general (nonlinear) Elasticity in the deformed configuration, where the symmetric CAUCHY stress tensor appears. Indeed, the variational formulation is

$$\int_\Omega \sum_{ij} \sigma_{ij} \frac{\partial v_i}{\partial x_j}\,dx = \int_\Omega \sum_i f_i v_i\,dx + \int_{\partial\Omega} g_i v_i d\sigma \text{ for all } v \in H^1(\Omega; R^3),$$

and by the symmetry of the stress tensor one has $\sum_{ij} \sigma_{ij} \frac{\partial v_i}{\partial x_j} = \sum_{ij} \sigma_{ij}\varepsilon_{ij}(v)$, where as usual $\varepsilon_{ij}(v) = \frac{1}{2}\big(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\big)$, and therefore the left side is 0 if v is such that $\varepsilon_{ij}(v) = 0$ for all i,j; this is the case if $v_i = a_i + \sum_j M_{ij} x_j$ for all i with M antisymmetric, and in three dimensions it means $M\,x = m \times x$ for some $m \in R^3$, and writing that the right side is 0 for all these v gives the necessary conditions on f and g, corresponding to the physical interpretation of total force and total torque. In Linearized Elasticity, i.e. $\sigma_{ij} = \sum_{k,l} C_{ijkl}\varepsilon_{kl}(u)$ for all i,j, with $C_{ijkl} = C_{jikl} = C_{ijlk}$ for all i,j,k,l, and under the hypothesis of Very Strong Ellipticity (i.e. there exists $\alpha > 0$ such that $\sum_{ijkl} C_{ijkl}A_{ij}A_{kl} \geq \alpha \sum_{ij} |A_{ij}|^2$ for all symmetric A), then the necessary conditions are sufficient if the injection of $H^1(\Omega)$ into $L^2(\Omega)$ is compact and if KORN's inequality holds, as a consequence of the Equivalence Lemma. This requires that one identifies all the $v \in H^1(\Omega; R^3)$ satisfying $\varepsilon_{ij}(v) = 0$ for all i,j, and it follows from the identity

$$2\frac{\partial^2 u_i}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j}\Big(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\Big) - \frac{\partial}{\partial x_i}\Big(\frac{\partial u_k}{\partial x_j} + \frac{\partial u_j}{\partial x_k}\Big) + \frac{\partial}{\partial x_k}\Big(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\Big) \text{ for all } i,j,k,$$

so that $\varepsilon_{ij}(v) = 0$ for all i,j implies that all second derivatives are 0, so that $v = a + M\,x$ and M must be antisymmetric. The solution u exists then and is defined up to the addition of $a + m \times x$ for $a, m \in R^3$, and it must be pointed out that these are not rigid displacements but linearized rigid displacements (the antisymmetric matrices appear as the tangent space at I for the manifold of all rotations $SO(3)$, which is compact). If the necessary conditions are not satisfied, the evolution equation will still have a solution and the body will move away in the direction of those linearized rigid displacements.

Let us imagine now, in the approximation of Linearized Elasticity, an elastic body with a flat part of its boundary put on an horizontal table, and assume that the system of forces applied to it does not take it away from the table (or consider the purely mathematical problem that the displacement satisfies $u_3 = 0$ on this flat part of the boundary); the body is allowed to slide horizontally on the table, and one expects to have less stringent compatibility conditions, corresponding to the horizontal part of the total force being 0 (there is no friction on the table and so the table will give a vertical reaction which will cancel the vertical component of the total force), and the torque along the x_3 axis must be 0 (the reactions of the table being able to compensate for the rest of the total torque). Mathematically, the condition $u_3 = 0$ on a piece of the boundary sitting in the plane $x_3 = H$, is imposed in the definition of the functional space, and v is constrained to be in this space, so only the elements $a + m \times x$ satisfying this constraint are allowed, i.e. one must choose $a_3 = m_1 = m_2 = 0$, and the necessary conditions corresponding to a_1, a_2 imply that the horizontal part of the total force is 0, while the necessary condition corresponding to m_3 implies then that the total torque around any vertical axis is 0.

Mathematically one can study nonhomogeneous conditions, like imposing u_3 on a piece of the boundary which is not necessarily flat, and the natural boundary conditions implied by the variational formulation will involve the traction T defined by $T_i = \sum_j \sigma_{ij} n_j$ (as for normal traces in $H(div;\Omega)$), and T_1 and T_2 can be imposed, with natural compatibility conditions

Mathematically, one could also impose the displacements u_1 and u_2 on a piece of the boundary, and the natural boundary condition implied by the variational formulation will involve T_3.

For (Newtonian) fluids, one has

$$\sigma_{ij} = 2\mu\varepsilon_{ij} - p\,\delta_{ij} = \mu\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big) - p\,\delta_{ij} \text{ for all } i,j,$$

so that

$$\sum_j \sigma_{ij} n_j = \mu\frac{\partial u_i}{\partial n} + \mu\sum_j \frac{\partial u_j}{\partial x_i} n_j - p\,n_i \text{ for all } i.$$

For an horizontal boundary, like the fixed interface separating Ocean from Atmosphere in the model considered by Roger LEWANDOWSKI, one has $n_1 = n_2 = 0$, and $n_3 = 1$ for Ocean and $n_3 = -1$ for Atmosphere, so that in the Ocean one has $T_1^O = \mu\frac{\partial u_1^O}{\partial x_3} + \mu\frac{\partial u_3^O}{\partial x_1}, T_2^O = \mu\frac{\partial u_2^O}{\partial x_3} + \mu\frac{\partial u_3^O}{\partial x_2}, T_3^O = 2\mu\frac{\partial u_3^O}{\partial x_3} - p^O$, and similarly for Atmosphere $T_1^A = -\mu\frac{\partial u_1^A}{\partial x_3} - \mu\frac{\partial u_3^A}{\partial x_1}, T_2^A = -\mu\frac{\partial u_2^A}{\partial x_3} - \mu\frac{\partial u_3^A}{\partial x_2}, T_3^A = -2\mu\frac{\partial u_3^A}{\partial x_3} + p^A$, and it is usually the jump of these quantities which appears in the variational formulations.

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

32. Monday April 5.

Modelization of turbulent flows is an important scientific and technological question, and although engineers may say that they are able to control turbulent flows, it is mainly because adaptive control ideas seem to work even in situations where no one knows what the right equations are for describing the phenomena which one wants to control. From a scientific point of view, not so much is understood about turbulence. For what concerns Oceanography, some modelization of turbulent flows is necessary in order to describe correctly what goes on at "small scales", remembering that the scales used for the Ocean, or the Atmosphere, are quite large.

It is quite common to experience the presence of microstructures in some fluid flows, but it is a very arduous task to propose a model that would describe accurately the important effects occuring in these flows. My first experimental evidence concerns the structure of the interface in front of a rainstorm, as I had observed many times long before becoming a mathematician, after a hot Summer day in the French countryside: one knows that a storm is coming, although the air is still, perhaps because the pressure is higher than usual, and then one starts to hear the leaves of the trees moving while the branches stay still; soon after the small branches start to move too, followed by the large branches a little after and the whole trees are in motion when the rain arrives. It clearly suggests that the classical idea of a sharp interface with some partial differential equations being satisfied on each side and with some boundary conditions being imposed on the "interface" might not be so efficient for describing the effects occuring in that living layer, with small vortices on the dry side and large vortices on the wet side. My second experimental evidence concerns the structure of the "wind", as I had observed twenty years ago, on a week end where I had expected to sail between La Rochelle and Ile de Ré, but the morning had provided us with what one calls "calme plat" in French: there was no wind, and the surface of the sea was extremely smooth and only showing a long swell ("houle" in French), which combined with the steady movement sustained by the small engine of the boat to produce a beginning of seasickness; fortunately, it did not last two long, because after a while we saw what one calls "une risée" in French (light squall in English): the wind waiting for us; it is an amazing fact to come from the windless side with a smooth sea surface to the place where the wind is, with the surface of the sea all wrinkled with wavelengths of the order of 5 to 10 centimeters, and when one crosses the transition line (which seemed stationary, but it might have been moving at a much slower pace that the boat, which was carried by its small engine), the sails inflated, and sailing started.

It was around twenty years ago too that I had heard Joe KELLER mention that at one time there had been a lot of articles about the statistical distribution of wavelengths of the waves at the surface of the sea, until one had been able to measure this distribution and it had appeared that all the theories had been wrong, as one had observed much more energy that any theorist had expected in the small capillary waves, those which I had observed as the signature of the wind waiting for us.

In other words, many like to imagine that natural phenomena obey the probabilistic processes or the statistical laws that are already known, and these people usually do not care that the phenomena that they are trying to study are described by complicated systems of partial differential equations for which their standard processes are obviously not adapted. In another meeting, Joe KELLER had mentioned the evolution from ideas about three dimensional turbulence by KOLMOGOROV, the two dimensional turbulence ideas used in meteorology (where the stratification by gravity simplifies the full three dimensional aspects), some one dimensional ideas that were not so good, and the zero dimensional ideas of iterating maps, followed by continually improving numerical simulations in dimensions one, two and even three, but he emphasized that something important had been lost in the way: in the 40s, turbulence specialists talked about velocities, pressure, kinetic energy, temperature, heat flux, while now they talk about statistics without reference to any important physical quantity related to fluids.

The only thing about turbulence that everybody agrees with is that it is created by oscillations in the velocity field, and REYNOLDS might have been the first to notice that if the "average" of u_i is denoted $\overline{u_i}$, then the average of $u_i u_j$ is $\overline{u_i}\,\overline{u_j} + R_{ij}$, where the symmetric REYNOLDS tensor R with entries R_{ij} is not necessarily 0.

Probabilists like to imagine that all functions in the fluid depend upon a parameter ω belonging to a space endowed with a probability measure, and integration with respect to this probability measure, the expectation, plays the role of the intuitive averaging technique.

Some specialists of asymptotic expansions like to plug functions like $u_0(x) + \varepsilon_n u_1\big(x, \frac{x}{\varepsilon_n}\big) + \ldots$ into the system of equations governing fluids, where the functions $u_j(x,y)$ are periodic in y, the vague idea of averaging becoming the precise technique of averaging in y, and this deterministic approach is sometime useful, although it is not able to explain some multiple scale effects that turbulent fluids are believed to show.

For about twenty five years, I have been developping a mathematical approach to the study of "oscillations" in solutions of partial differential equations, partly in collaboration with François MURAT, and various notions of weak convergence appear in this approach, which definitely has an advantage on all the others, that it does not postulate anything about oscillations but tries to determine what kind of oscillations are compatible with linear differential balance laws and nonlinear constitutive relations. First, I should point out that I use the term "oscillations" to englobe also "concentration effects", i.e. the meaning used is to consider weakly convergent sequences which are not strongly convergent, but convergences of a weak type but different from the usual weak convergence are also used. Second, I should point out that the use of sequences is a purely mathematical trick whose object is to identify the correct topology (usually related to some kind of weak convergence) that one should use for various physical or nonphysical quantities (it is similar to the description of R by starting from CAUCHY sequences in Q, and once R is understood a real number is not related to a sequence of rationals any more!).

The classical weak convergence appears to be natural for some quantities and not for others, and the notion of differential forms will clarify this question. In the equation expressing conservation of mass, $\frac{\partial\rho}{\partial t} + div(\rho\, u) = 0$, the quantities ρ and $q_i = \rho\, u_i$ are coefficients of differential forms, but u_i only appears as a quotient of two quantities for which the adapted topology is weak convergence; therefore density and momentum are more easy to handle than velocity. It will be useful then to describe some properties of H-convergence (introduced with François MURAT, and generalizing the notion of G-convergence introduced by Sergio SPAGNOLO, with some ideas from Ennio DE GIORGI), and I will describe some properties of weakly converging sequences of solutions of equations like $div\big(A_n\, grad(u_n)\big) = f$. It will then be natural to consider sequences of operators of the form $\frac{\partial}{\partial t} + \sum_i u_i^n \frac{\partial}{\partial x_i}$, and as nothing general is known in the case when the coefficients only converge weakly, I will describe in detail some special cases.

It is worth mentioning that geometers like to think that they know how to write equations for fluid flows in intrinsic forms, but as long as one does not know how to pass to the limit in weakly convergent sequences of solutions of these equations, one cannot assert that geometers have or have not introduced the correct framework (my guess is that they have not!).

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

33. Wednesday April 7.

In questions of asymptotic expansions, one considers sequences of functions like $v^n(x) = u_0(x) + \varepsilon_n u_1\big(x, \frac{x}{\varepsilon_n}\big) + \ldots$, where the functions $u_j(x,y)$ are periodic in y (and smooth enough in (x,y)), and ε_n tends to 0. If ε_n is a small characteristic length, and if the solution of a physical problem has this form, then if one measured the value of v^n at a few points, quite far apart compared to the characteristic length ε_n, then one would find $u_0(x) + O(\varepsilon_n)$ (plus some eventual errors due to the measuring process), and one might well believe that the measured solution is u_0, considering the little discrepancies as systematic errors.

It is usual in Physics courses to be told that one term is small and that it will be neglected (it is not always clear if these terms are indeed small, as they may be small in the real world, but if the equation used is not a good model of the physical world the corresponding term might not be so small); having neglected some terms one performs some formal computations with the simplified equation, like taking derivatives, and the first remark, which seems to infuriate Physics teachers, is that the derivative of a small term might not be small; actually, in our example one has $\frac{\partial v^n}{\partial x_i} = \frac{\partial u_0}{\partial x_i} + \frac{\partial u_1}{\partial y_i} + O(\varepsilon_n)$, and $\frac{\partial u_1}{\partial y_i}$ is not small when u_1 does depend upon y.

Fortunately, in some cases like linear partial differential equations with smooth coefficients, the procedure can be shown to work, because of the generalized framework of the theory of distributions of Laurent SCHWARTZ for example: if a sequence v^n converges to v^∞, then $\frac{\partial v^n}{\partial x_i}$ converges to $\frac{\partial v^\infty}{\partial x_i}$ for every i, but in this statement it must be realized that the meaning of convergence is not that the differences are uniformly small; therefore we do have $v^n = u_0(x) + O(\varepsilon_n)$, while $\frac{\partial v^n}{\partial x_i} \neq \frac{\partial u_0}{\partial x_i} + O(\varepsilon_n)$, but there is no contradiction as long as one only considers linear questions.

If one wants to avoid the too general framework of distributions, one can instead mention classical results of Functional Analysis concerning weak topologies; weak convergence appears then natural for quantities which are integrated against test functions, or integrated on certain sets; in Continuum Mechanics it is often the case that such quantities are coefficients of differential forms (and it is probably only for those that the weak convergence should be used).

For example, in the equation of conservation of mass $\frac{\partial \rho}{\partial t} + \sum_i \frac{\partial(\rho\, u_i)}{\partial x_i} = 0$, ρ and $\rho\, u_i$, $i = 1,2,3$, are the coefficients of a 3-differential form in space-time, namely

$$\omega = \rho\, dx_1 \wedge dx_2 \wedge dx_3 - \rho\, u_1 dt \wedge dx_2 \wedge dx_3 + \rho\, u_2 dt \wedge dx_1 \wedge dx_3 - \rho\, u_3 dt \wedge dx_1 \wedge dx_2,$$

and as

$$d\,\omega = \Big(\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3} \frac{\partial(\rho\, u_i)}{\partial x_i}\Big) dt \wedge dx_1 \wedge dx_2 \wedge dx_3,$$

the equation of conservation of mass is $d\,\omega = 0$. One must notice that the components u_i of the velocity field are not themselves coefficients of differential forms (and the weak convergence is not adapted for them), and it is the momentum which is the correct physical quantity, which has an additive character. The velocity is not always mentioned when one deals with conservation of electric charge, and it is written as $\frac{\partial \rho}{\partial t} + div(j) = 0$, and one does not even bother to define a velocity as $\frac{j}{\rho}$, because it would usually be meaningless: indeed the electric charge is transported by light electrons and by heavy ions, and an average velocity would be of little use (it is better to think of two interacting populations, one of electrons and one of ions, eventually having their own temperature).

I will show later an example where a quantity which is not a coefficient of a differential form necessitates a different type of weak topology (the H-convergence, which I have introduced with François MURAT; it generalizes to nonsymmetric operators the G-convergence introduced by Sergio SPAGNOLO, but it also introduces a quite different point of view).

Although I am using the framework of differential forms, my motivation is quite different from that of geometers, and it is worth describing the differences of points of view.

The exterior calculus is purely algebraic: one considers the p-linear alternated forms on a finite dimensional vector space E, i.e. f is multilinear and satisfies $f(e_{s(1)}, \ldots, e_{s(p)}) = \varepsilon(s) f(e_1, \ldots, e_p)$ for all

$e_1, \dots, e_p \in E$ and all permutations s of p elements, where $\varepsilon(s)$ is the signature of the permutation s. One defines then the exterior product $\wedge$: if f is p-linear alternated and g is q-linear alternated then $f \wedge g$ is the $(p+q)$-linear alternated defined by $(f \wedge g)(e_1, \dots, e_{p+q}) = \frac{1}{p!q!} \sum_s \varepsilon(s) f(e_{s(1)}, \dots, e_{s(p)}) g(e_{s(p+1)}, \dots, e_{s(p+q)})$, where s runs through the permutations of $p+q$ elements; one checks easily that $g \wedge f = (-1)^{pq} f \wedge g$. The exterior product is associative.

A differential form of order p, or a p-form, on an open set Ω of E, is a (smooth enough) mapping from Ω into the space of p-linear alternated forms; a 0-form is a function, and the derivative of a function is a 1-form. Then one defines the exterior derivative d, which maps p-forms into $(p+1)$-forms, with the rules that $d(f \wedge g) = (df) \wedge g + (-1)^p f \wedge (dg)$ if f is a p-form, and $df = \sum_i \frac{\partial f}{\partial x_i} dx_i$ if f is a function. One shows that $d \circ d = 0$, and POINCARÉ's lemma asserts that if $df = 0$ then locally $f = dh$ for a $(p-1)$-form h (asking for global results leads to questions of Algebraic Topology).

One can restrict a differential form to a submanifold by considering its action only on vectors tangent to the submanifold, and actually one can develop all the theory of differential forms on abstract manifolds (not necessarily orientable), with or without boundary, and one proves the STOKES formula $\int_\Omega d\omega = \int_{\partial\Omega} \omega$. As a student, before learning this framework, I was taught about formulas by GREEN, STOKES, and OSTROGRADSKI, where *curl* and *div* appear: a vector field V in R^3 can be attached to a 1-form (because vectors and covectors are identified) $\omega(V) = V_1 dx_1 + V_2 dx_2 + V_3 dx_3$ but also to a 2-form $\pi(V) = V_1 dx_2 \wedge dx_3 + V_2 dx_3 \wedge dx_1 + V_3 dx_1 \wedge dx_2$ (again because one uses the Euclidean structure), and $d\omega(V) = \pi\big(curl(V)\big)$ and $d\pi(V) = div(V) dx_1 \wedge dx_2 \wedge dx_3$ (one usually suppresses the $\wedge$ and one replaces $dx_1 \wedge dx_2 \wedge dx_3$ by dx). From a practical point of view, *curl* appears for 1-forms and *div* appears for $(N-1)$-forms in dimension N.

I suppose that all this beautiful theory was developped by Henri POINCARÉ and Elie CARTAN, but I have also heard the name of PFAFF being mentioned.

If differential forms are natural for geometers, as they are the right objects which transform well under change of variables, the reason why I am using them is different: they are adapted to weak convergence. In the early 70s, I worked with François MURAT on questions that were not yet called Homogenization, and we had understood from reading some work of Henri SANCHEZ-PALENCIA (who was using asymptotic expansions for problems with periodic microstructures), that what we had done was related to effective properties of mixtures (we also discovered that Sergio SPAGNOLO had solved earlier the first step of our program): we were considering a sequence of elliptic problems $-div\big(A_n\, grad(u_n)\big) = f$ in Ω, together with some natural boundary conditions, and when A_n converged weakly we could extract a subsequence u_m converging weakly in $H^1(\Omega)$ to u_∞, but the limit of $A_m\, grad(u_m)$ could not be defined easily, and we had introduced an adapted notion (later called H-convergence, and generalizing the G-convergence introduced by Sergio SPAGNOLO). Using notations from Electrostatics, with $E_n = -grad(u_n)$ and $D_n = A_n E_n$, I was considering the weak convergence natural for the electric field E_n, interpreting its weak limit E_∞ as a macroscopic field, and similarly for the polarization field D_n, but the right limit for A_n was to relate D_∞ to E_∞ by a different physical process where there was no averaging of A_n: one created a macroscopic field E_∞ by choosing correctly f (which is ρ in Electrostatics) and one measured the limit D_∞ and that gave a partial information of a tensor A^{eff} such that $D_\infty = A^{eff} E_\infty$; therefore one does not "measure" A^{eff} by computing averages, but one "identifies" A^{eff} from averages of the electric and polarization fields. We had also discovered the Div-Curl lemma, which I will describe below, and during the year 1974/75 which I spent in Madison, Joel ROBBIN had explained to me that our results became quite clear when expressed in the framework of differential forms (using HODGE decomposition; he had also taught me how to write MAXWELL equation using differential forms). In the Fall of 1975, I met John BALL and learned about the sequential weak continuity of Jacobians (which I thought he had proved, but understood later that MORREY had done that in the 50s), and in dimension 2 or 3 I could derive easily these results from the Div-Curl lemma; the general framework of Compensated Compactness appeared the year after, again with participation of François MURAT.

The Div-Curl lemma states that if $\Omega \subset R^N$, if $E_n \rightharpoonup E_\infty$ in $L^2_{loc}(\Omega; R^N)$ weak, $D_n \rightharpoonup D_\infty$ in $L^2_{loc}(\Omega; R^N)$ weak, $div(D_n) \to div(D_\infty$ in $H^{-1}_{loc}(\Omega)$ strong, and $curl(E_n) \to curl(E_\infty$ in $H^{-1}_{loc}(\Omega; X)$ strong (where X has the right dimension), then $(E_n.D_n)$ converges to $(E_\infty.D_\infty)$ in the sense of measures (i.e. integrated against test functions in $C_c(\Omega)$). In the case where $E_n = grad(u_n)$, this is integration by parts and uses the compactness of the injection of $H^1_{loc}(\Omega)$ into $L^2_{loc}(\Omega)$ by writing $(E_n.D_n) = -(grad(u_n).D_n) = -div(u_n D_n)+$

$u_n\, div(D_n)$, which one integrates against $\varphi \in C_c^\infty(\Omega)$.

The Compensated Compactness quadratic theorem considers a general framework of linear differential equations with constant coefficients, $U^n \rightharpoonup U^\infty$ in $L^2_{loc}(\Omega; R^p)$ weak and $\sum_{jk} A_{ijk} \frac{\partial U_j^n}{\partial x_k} \to f_i$ in $H^{-1}_{loc}(\Omega)$ strong for $i = 1, \ldots, q$, and identifies the possible limits (in the sense of measures) of all quadratic functions in U^n: if Q is quadratic and $Q(U^n)$ converges to $Q(U^\infty) + \nu$ in the sense of measures then $\nu \geq 0$ if $Q(\lambda) \geq 0$ for all $\lambda \in \Lambda = \{\lambda \in R^p :$ there exists $\xi \in R^N \setminus 0, \sum_{jk} A_{ijk}\lambda_j \xi_k = 0$ for $i = 1, \ldots, q\}$; this result is optimal. In particular $Q(U^n) \rightharpoonup Q(U^\infty$ in the sense of measures if $Q(\lambda) = 0$ for all $\lambda \in \Lambda$.

The Compensated Compactness method adds the use of "entropies" (which geometers call CASIMIRs) in order to deduce which YOUNG measures could be associated to the sequence U^n, assumed also to satisfy the constraints $U^n(x) \in K$ a.e. $x \in \Omega$ (which corresponds to constitutive relations, while the differential equations corresponds to balance equations for problems in Continuuum Mechanics).

If one did not know about differential forms, one would discover them by looking at sequences U^n which converge strongly to U^∞ in $L^2_{loc}(\Omega; R^p)$ but only weakly in $H^1_{loc}(\Omega; R^p)$, and wonder if one could compute the limit of some functions of $grad(U^n)$; indeed the quadratic theorem of Compensated Compactness would show that $dU_i^n \wedge dU_j^n$ converges to $dU_i^\infty \wedge dU_j^\infty$ in the sense of measures; by reiteration, using the entropy conditions following from the formula for $d(f \wedge g)$, one could recover MORREY's result about Jacobians. The Compensated Compactness framework is of course more general than MORREYs result, and can be used for any system that one encounters in Continuum Mechanics, but the Compensated Compactness method still needs to be improved. For example, when applied to MAXWELL equation, one find three independent quadratic quantities which are sequentially weakly continuous, and if one knows the framework with differential forms, they come from exterior products of forms which have a good exterior derivarive. More generally, let f^n be a sequence of p-forms converging to f^∞ in $L^2_{loc}(\Omega)$ (for its coefficients) and such that df^n has its coefficients staying in a compact of $H^{-1}_{loc}(\Omega)$ strong, let g^n be a sequence of q-forms converging to g^∞ in $L^2_{loc}(\Omega)$ (for its coefficients) and such that dg^n has its coefficients staying in a compact of $H^{-1}_{loc}(\Omega)$ strong, then $df^n \wedge dg^n$ converges to $df^\infty \wedge dg^\infty$ in the sense of measures (of course one has better convergences if one improves the hypotheses). Of course, the Compensated Compactness method must sometimes be used in conjunction with the ideas of H-convergence that I developed with François MURAT for Homogenization.

I always wonder why there is still a group of people who pretend to be interested in Elasticity and wants to ignore this framework which I have taught more than 20 years ago, where one can naturally introduce the equilibrium equation; I have heard so many talks by mathematicians who pretend to be interested in Elasticity and never mention the word stress, that I wonder if it is really so hard for them to learn about Continuum Mechanics (maybe it is hard for them to quote my results when they need them, and they often prefer to quote some others who have used my methods and have forgotten to mention where they had learned about them).

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

34. Friday April 9.

In one dimension, one can solve explicitly all Homogenization problems by computing various weak limits, and the same is true in more than one dimension when the oscillating coefficients only depend upon one variable; the general case of a diffusion equation was solved by François MURAT in the early 70s, then I learned in 1975 about the computation by MCCONNELL of the general case for Linearized Elasticity, and I derived the general approach shown below a few years after; in 1979, having been asked how to compute the effective properties of a material layering steel and rubber, I also explained how to carry out the computations in a nonlinear setting, although there is no general theory of Homogenization for (nonlinear) Elasticity (despite the claim of those who have fallen into the trap of the Γ-convergence approach, this is still the situation today).

The basic idea is an application of the Div-Curl lemma: if $\Omega \subset R^N$ and $D^n \rightharpoonup D^\infty$ in $L^2_{loc}(\Omega; R^N)$ weak with $div(D^n)$ staying in a compact of $H^{-1}_{loc}(\Omega)$ strong, then "D^n_1 does not oscillate in x_1", i.e. whenever f_n only depends upon x_1 and $f_n \rightharpoonup f_\infty$ in $L^2_{loc}(\Omega)$ weak, one has $D^n_1 f_n \rightharpoonup D^\infty_1 f_\infty$ in the sense of measures (the precise definition that a sequence is not oscillating in x_1 says that the corresponding H-measures do not charge the point e^1 of the unit sphere); of course this follows from the fact that $E^n = f_n e^1$ is a gradient.

For a diffusion equation, if $E^n = grad(u_n)$ and $D^n = A^n E^n$ satisfies $div(D^n) \to f$ in $H^{-1}_{loc}(\Omega)$ strong, with A^n only depending upon x_1, one remarks that D^n_1 does not oscillate in x_1 as well as $E^n_2, \ldots, E^n_N$, because of the equation $curl(E^n) = 0$; from the components of E^n and D^n, one creates a good vector G^n whose components are $D^n_1, E^n_2, \ldots, E^n_N$, and a bad vector B^n whose components are $E^n_1, D^n_2, \ldots, D^n_N$, and one has $B^n = \Phi(A^n) G^n$, where $\Phi(A^n)$ is obtained by algebraic computations from A^n, and these computations (which start by eliminating E^n_1 in the equation giving D^n_1) only require that A^n_{11} stay away from 0; as A^n only depends upon x_1, so does $\Phi(A^n)$ and one can pass to the limit in $\Phi(A^n)G^n$, so that $B^\infty = $ [weak limit $\Phi(A^n)]G^\infty$, i.e. $\Phi(A^{eff})$ is the weak limit of $\Phi(A^n)$. For Linearized Elasticity, the good vector uses the components σ_{i1} (and σ_{1i} which is equal to σ_{i1} because the CAUCHY stress tensor is used), and the ε_{ij} for $i,j \geq 2$, while the bad vector uses the other components; starting from $\sigma_{ij} = \sum_{kl} C_{ijkl}\varepsilon_{kl}$, the algebraic computations only require that the acoustic tensor $A(e^1)$ be invertible (one defines $A_{ik}(\xi) = \sum_{jl} C_{ijkl}\xi_j\xi_l$). For (nonlinear) Elasticity, the good vector uses the components σ_{i1} (but not σ_{1i} which is different from σ_{i1} because the PIOLA-KIRCHHOFF stress tensor is used), and the $\frac{\partial u_i}{\partial x_j}$ for $j \geq 2$, while the bad vector uses the other components (in the case of HyperElasticity where there is a stored energy function W, the computations only require a uniform rank-one convexity for W).

If one considers now to the general problem of Homogenization, and I recall that I do not imply any restriction to periodic structures like so many do when using this term (probably because they have not understood the general framework that I had developed with François MURAT), one imposes a uniform ellipticity condition, which for the diffusion case is that there exists $0 < \alpha \leq \beta < \infty$ such that $(A^n(x)\xi.\xi) \geq \alpha|\xi|^2$ and $(A^n(x)\xi.\xi) \geq \frac{1}{\beta}|A^n(x)\xi|^2$ for all $\xi \in R^N$ and a.e. $x \in \Omega$ (if A^n is symmetric, it means that $\alpha\, I \leq A^n \leq \beta\, I$ almost everywhere).

In the G-convergence approach, developed in the late 60s by Sergio SPAGNOLO (helped by the insight of Ennio DE GIORGI), one only considers symmetric A^n and one extracts a subsequence such that for every $f \in H^{-1}(\Omega)$ the solution $u_m \in H^1_0(\Omega)$ of the equation $-div\big(A^m grad(u_m)\big) = f$ converges weakly to u_∞, and one shows that there exists A^{eff} (symmetric with $\alpha\, I \leq A^{eff} \leq \beta\, I$ almost everywhere) such that $-div\big(A^{eff} grad(u_\infty)\big) = f$ (this is the convergence of the GREEN kernels, and explains the choice of the prefix G).

In the H-convergence approach, which I developed in the early 70s with François MURAT without knowing at the time what Sergio SPAGNOLO had already done, one can consider nonsymmetric A^n and one extracts a subsequence such that for every $f \in H^{-1}(\Omega)$ the solution $u_m \in H^1_0(\Omega)$ of the equation $-div\big(A^m grad(u_m)\big) = f$ converges weakly to u_∞, but also $A^m grad(u_m)$ converges weakly to D^∞, and one shows that there exists A^{eff} (with $(A^{eff}(x)\xi.\xi) \geq \alpha|\xi|^2$ and $(A^{eff}(x)\xi.\xi) \geq \frac{1}{\beta}|A^{eff}(x)\xi|^2$ for all $\xi \in R^N$ and a.e. $x \in \Omega$) such that $D^\infty = A^{eff} grad(u_\infty)$ and therefore $-div\big(A^{eff} grad(u_\infty)\big) = f$ (it is equivalent to

G-convergence in the symmetric case, and the choice of the prefix H, chosen in the late 60s, reminds of the term Homogenization introduced by Ivo BABUŠKA).

It is important to realize that A^{eff} cannot be computed using the YOUNG measures associated with the sequence A^n in dimension at least 2; using YOUNG measures is the mathematical way of dealing with one-point statistics which physicists use in their probabilistic framework, and therefore the preceding statement says that one cannot deduce the effective properties of a mixture by using only the proportions of the different constituents used (something we had known since the early 70s).

The preceding statement contradicts a few other claims for the following reasons. Some mathematicians claim that YOUNG measures are the right objects to study microstructures in crystals for example, and it is because they have not understood what I had already taught in my 1978 lectures at HERIOT-WATT University, where I had used (YOUNG) paraletrized measures for describing the limits of sequences constrained (in a pointwise way) by constitutive relations, and the Compensated Compactness method for describing the constraints due to the (linear differential) balance relations; I had shown the importance of characterizing which YOUNG measures are compatible with a given set of linear differential equations and a nonlinear constitutive relation, but those who have used much later terms like "gradient YOUNG measures" usually forget to mention that I had introduced that notion for a general system because the laws of Continuum Mechanics cannot be expressed using only gradients (cf. the widespread disease of pretending to work on Elasticity without ever mentioning stress); even for questions like twinning, which are akin to the method for computing effective properties of layers which I have described before, the mistake consists in not understanding that one uses in a crucial way the directions of the twins and therefore the statement that YOUNG measures are the right objects should be restricted to one dimensional geometries. Physicists do write formulas for effective properties of mixtures, but they are only approximations, or bounds, and I have initially developped the technique of H-measures to explain why some formulas guessed by physicists are good in situations where the properties of the constituents are very similar. There are other situations where phycisists might be right, because they only observe the result of an evolution, like for mixtures of gases or liquids, and it might be that the evolution dissipates energy and ends up at a stable equilibrium, which they can compute (there are no good mathematical methods yet for studying the evolution of mixtures).

YOUNG measures have been introduced in the 30s by Laurence C. YOUNG (son of William and Grace YOUNG, who were both mathematicians and collaborated extensively so that it is not clear if some of the famous results attributed to YOUNG are due to his father or his mother); I learned about these measures as parametrized measures in seminars on control theory in the late 60s (without attribution to YOUNG) and I first used them under that name; for a sequence of measurable functions U^n on $\Omega \subset R^N$, taking values in a closed bounded set $K \subset R^p$, there is a subsequence and a measurable family ν_x of probability measures on K such that for every continuous function φ on K the subsequence $\varphi(u_m)$ converges in $L^\infty(\Omega)$ weak $\star$ to a limit l_φ such that $l_\varphi(x) = \langle \nu_x, \varphi \rangle$ for a.e. $x \in \Omega$. If K is unbounded, one may lose information at infinity (one may use a compactification of K), and this corresponds to concentration effects; notice that the Compensated Compactness quadratic theorem, or the theory of H-measures which generalizes it, can deal with oscillations and concentration effects simultaneously; losing mass at infinity was a classical question in problems of theoretical Physics, and observing concentration effects in minimizing sequences was a well known fact for geometers, and when Pierre-Louis LIONS studied these questions he might not have been aware of all the earlier results, but he told me that had choosen to call his approach the Concentration-Compactness method with the goal of inducing people in error because of the similarity in name with the Compensated Compactness method (which he had obviously not understood well himself even a few years after telling me that).

If one mixes two isotropic materials of conductivity (or permittivity) α, β, it means that $A^n = (\chi_m \alpha + (1-\chi_n)\beta) I$, and if $\chi_n \rightharpoonup \theta$ in $L^\infty(\Omega)$ weak $\star$, then $\theta(x)$ is the local proportion of the first material near x; the YOUNG measure in this case is $\nu_x = \theta(x)\delta_{\alpha\, I} + \big(1-\theta(x)\big)\delta_{\beta\, I}$. One can construct layers in x_1 or layers in x_2 for two sequences having the same YOUNG measure by taking θ constant, but the effective properties are different: if a_+ is the arithmetic average $\theta\alpha + (1-\theta)\beta$, and a_- is the harmonic average $\big(\frac{\theta}{\alpha} + \frac{1-\theta}{\beta}\big)^{-1}$, then layering in the direction x_j corresponds to A^{eff} being diagonal with $A^{eff}_{ii} = a_-\delta_{ij} + a_+(1-\delta_{ij})$; it is a little more technical to construct sequences for which A^{eff} is of the form γI and show that the value γ can be different for two sequences using the same proportions.

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

35. Monday April 12.

For questions of Oceanography, we have to face the extremely difficult problem of passing to the limit in NAVIER-STOKES in situations where the REYNOLDS number gets large; if one maintains the size of the domain and the size of the velocities it corresponds to letting the kinematic viscosity ν tend to 0. We have seen that for ρ and $\rho\, u$, the density of mass and the density of momentum, the weak convergence is well adapted as they are coefficients of differential forms, but as turbulence is related to situations where the velocity u fluctuates, we must understand how to average u, in the sense of finding the right topology adapted to that quantity. The velocity u appears inside the differential operator

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \sum_{i=1}^{3} u_i \frac{\partial}{\partial x_i},$$

which is a transport operator, and what is transported is mass, momentum, angular momentum, temperature, salinity, pollutants, etc... It is natural to ask the same question that was solved for equations of the form $div\big(A^n\, grad(u_n)\big) = f$ for equations of the form $\frac{\partial v^n}{\partial t} + \sum_{i=1}^{3} u_i^n \frac{\partial v_n}{\partial x_i} = f$, where u^n is now a given oscillating field and the solution v^n may represent any one of the transported quantities. If v^n is a scalar quantity, we are considering a first order operator, which is hyperbolic, but it can also be considered as a degenerate elliptic operator and this could give some hope that the results for elliptic operators could extend to degenerate cases, but this extension is not straightforward, as we will see on much simpler examples, because nonlocal effects appear.

That nonlocal effects may appear by Homogenization had first been noticed by Henri SANCHEZ-PALENCIA (using asymptotic expansions in a periodic setting), for questions like ViscoElasticity or for some memory effects in Electricity corresponding to the fact that some coefficients depend upon frequency, and Jacques-Louis LIONS had invented examples where one needed to introduce pseudo-differential operators (with an interpretation as memory effects). I started thinking about this question in 1980 because I guessed that these effects were the main reason behind the strange rules of absorption and reemission used by physicists, and I looked at the very simplified following model

$$\frac{\partial u_n(x,t)}{\partial t} + a_n(x)u_n(x,t) = f(x,t) \text{ in } \Omega \times (0,T);\ u_n(x,0) = v(x),$$

where a_n takes values between α and β and converges to a_∞ in $L^\infty(\Omega)$ weak $\star$ (the YOUNG measure of a_n contains all the information that we will need). I guessed that the limiting equation would have a convolution term

$$\frac{\partial u_\infty(x,t)}{\partial t} + a_\infty(x)u_\infty(x,t) - \int_0^t K(x,t-s)u_\infty(x,s)\,ds = f(x,t) \text{ in } \Omega \times (0,T);\ u_\infty(x,0) = v(x),$$

and I expected $K \geq 0$ for reasons related to the maximum principle. If one defines $B_n(x,t) = e^{-t\,a_n(x)}$, then one has $u_n(x,t) = B_n(x,t)v(x) + \int_0^t B_n(x,t-s)f(x,s)\,ds$, and if $B_n \rightharpoonup B_\infty$ in $L^\infty\big(\Omega \times (0,T)\big)$ weak $\star$, then u_∞ satisfies the same equation with B_n replaced by B_∞; as $B_\infty(x,t) \neq e^{-t\,a_\infty(x)}$ except if a_n converges strongly to a_∞ (in $L^1_{loc}(\Omega)$ for example), one cannot have $K = 0$, and this gives the simplest example of a sequence of semi-groups whose limit is not a semi-group. Of course, the kernel K must satisfy the equation $\frac{\partial B_\infty(x,t)}{\partial t} + a_\infty(x)B_\infty(x,t) = \int_0^t K(x,t-s)B_\infty(x,s)\,ds$ for a.e. $x \in \Omega$; this was the first approach I had taken after trying the approach by LAPLACE transform which I will follow now.

One defines the LAPLACE transform of a function g defined on $(0,\infty)$ by $\mathcal{L}g(p) = \int_0^\infty g(t)e^{-p\,t}\,dt$, and usually the LAPLACE transform is holomorphic in some half space $\Re p > \gamma$, and the theory has been extended by Laurent SCHWARTZ to some distributions; the important fact is that $\mathcal{L}(g \star h) = \mathcal{L}g\,\mathcal{L}h$, and $\mathcal{L}\frac{dg}{dt} = p\mathcal{L}g + g(0)$. One has

$$\Big(p + a_n(x)\Big)\mathcal{L}u_n(x,p) = \mathcal{L}f(x,p) + v(x),$$

and

$$\Big(p + a_\infty(x) - \mathcal{K}(x,p)\Big)\mathcal{L}u_\infty(x,p) = \mathcal{L}f(x,p) + v(x),$$

so that K is characterized by

$$p + a_\infty - \mathcal{K}(\cdot,p) = \Big(\text{weak limit } \frac{1}{p+a_n}\Big)^{-1}.$$

Of course, using the YOUNG measures associated with the sequence a_n, the weak limit of $\frac{1}{p+a_n}$ is $\int_{[\alpha,\beta]}\frac{d\nu_x(a)}{p+a}$, and the key to the formula for K is a property about functions $F(z)$ which satisfy $\Im F(z) > 0$ when $\Im z > 0$ (I had first heard of it in talks by David BERGMAN, and the idea is attributed to various authors such as HELMHOLTZ, PICK, NEVANLINA or STIELTJES): suppose as a simplification that F is defined in the complex plane except for a bounded closed interval I on the real axis, and satisfies $\Im F(z)\Im z > 0$ for $\Im z \neq 0$, then there exists $A \geq 0$, $B \in R$ and a nonnegative RADON measure μ with support in I such that

$$F(z) = A\,z + B + \int_I \frac{d\mu(\lambda)}{\lambda - z} \text{ for all } z \notin I.$$

One deduces that

$$\frac{1}{\int_I \frac{d\nu_x(a)}{p+a}} = p + a_\infty(x) + \int_{[-\beta,-\alpha]} \frac{d\mu_x(\lambda)}{\lambda - p} \text{ for all } p \notin [-\beta,-\alpha], \text{ a.e. } x \in \Omega,$$

as a TAYLOR expansion near $p = \infty$ gives $A = 1$ and $B = a_\infty(x) = \int_I a\,d\nu_x(a)$; the inverse LAPLACE transform is then easily performed and gives

$$K(x,t) = \int_{-I} e^{\lambda\,t}\,d\mu_x(\lambda).$$

If a_n takes only k different values, then ν_x is a combination of at most k DIRAC masses, and μ_x is a combination of at most $(k-1)$ DIRAC masses which are the roots of a polynomial for which there is no simple formula in general.

In the preceding example, we found a solution u_∞ and we looked then for an equation that it satisfies, and the reason that the one obtained is natural is that the operator $\frac{d}{dt} + a_n$ is linear and commutes with translation in t, and a theorem of Laurent SCHWARTZ says that every linear operator which commutes with translation is a convolution operator (with a distribution kernel), and the only kernel that works here is $\frac{d\delta_0}{dt} + a_\infty\delta_0 - K$; we will use again this argument below, but in nonlinear settings the situation is not as clear.

Although the preceding example is not of great interest from a physical point of view (one could use it for a mixture of materials decaying at different rates for example), it shows something important from a philosophical point of view: the memory effect term is not related to any probabilistic argument! It is actually possible to invent a probabilistic game, with particles absorbed and particles reemitted, which will create the equation that we have found; there is absolutely no reason other than ideology to give a better status to the probabilistic approach than to any other way of considering the preceding equation (Probability is a part of Analysis, but from the point of view of Analysis without Probability integral equations with smooth kernels are treated as mere perturbations, in semi-group theory for example).

The model explains qualitatively something about irreversibility. One may start from an equation for which one can reverse time and a limiting process may make an irreversible equation appear: diffusion equations arrive naturally in certain situations by letting the velocity of Light c tend to ∞, but the equation that one starts from has already incorporated a modelisation of scattering which is not reversible. A more puzzling question is asked by people who start from a finite dimensional Hamiltonian system and let the number of degrees of freedom tend to ∞, as numerical simulations show that something like entropy increases, but the system is reversible and the same occurs for the reversed equation. The answer provided by the example is that one might have to consider memory effects in order to describe well what is going on, and an observer using time in a backward way will then do the same analysis and get an integral term from t

to ∞ instead in his equations; it is when one wants to get rid of the nonlocal effects and only use partial differential equations that the problems occur. In the previous example, one can approach $d\mu$ by a finite combination of DIRAC masses and transform the equation obtained into a system of differential equations.

The method which I have shown above was applied by my former student Kamel HAMDACHE (with AMIRAT and AbdelHamid ZIANI) to a question which is more relevant to the questions of fluids that we are interested in (but from a pedagogical point of view I prefer to start with the simpler problem which was done first); their motivation was in flows in porous media, and they considered

$$\frac{\partial u_n}{\partial t} + a_n(y)\frac{\partial u_n}{\partial x} = f(x,y,t) \text{ in } R \times \Omega \times (0,T);\ u_n(x,y,0) = v(x,y) \text{ in } R \times \Omega.$$

The method is essentially the same, using LAPLACE transform in t, but also FOURIER transform in x, due to the fact that the partial differential operator that we are dealing with commutes with translations in t but also in x; this gives

$$\Big(p + 2i\pi\xi\, a_n(y)\Big)\mathcal{L}\mathcal{F}(\xi,y,p) = \mathcal{L}\mathcal{F}f(\xi,y,p) + \mathcal{F}v(\xi,y),$$

and if one uses the YOUNG measures ν_y associated with a subsequence, one needs the weak $\star$ limit of $\frac{1}{p+2i\pi\xi\, a_n(y)}$, which is $\int_I \frac{d\nu_y(a)}{p+2i\pi\xi\, a} = \frac{1}{2i\pi\xi}\int_I \frac{d\nu_y(a)}{q+a}$ where $q = \frac{p}{2i\pi\xi}$, and the same formula that we used before appears, so one can perform the inverse LAPLACE-FOURIER transform easily and one obtains the only convolution equation is (x,t) (independent of f and v) that the limit solution may satisfy

$$\frac{\partial u_\infty}{\partial t} + a_\infty(y)\frac{\partial u_\infty}{\partial x} - \int_0^t \int_{[-\beta,-\alpha]} \frac{\partial^2 u_\infty(x+\lambda(t-s),y,t-s)}{\partial x^2}\, d\mu_y(\lambda)\, ds = f(x,y,t) \text{ in } R \times \Omega \times (0,T),$$

with $u_\infty(x,y,0) = v(x,y)$ in $R\times\Omega$. Notice that the second derivatives are not computed at the point (x,y,t) but on lines approaching the point with a velocity $-\lambda$, with a weight depending upon λ; the equation obtained has of course the finite propagation speed property and AMIRAT, HAMDACHE and ZIANI checked that this is true for any nonnegative measure $d\mu_y$ with bounded support (the ones coming from the formula have a constraint on their mass for example); they also proposed a way to look at this equation as a possibly infinite hyperbolic system, by using the auxilliary functions

$$\varphi(x,y,t;V) = \int_0^t \frac{\partial u_\infty(x - V(t-s),y,t-s)}{\partial x}\, ds,$$

for $V \in [\alpha,\beta]$, so that

$$\frac{\partial\varphi(x,y,t;V)}{\partial t} + V\frac{\partial\varphi(x,y,t;V)}{\partial x} = \frac{\partial u_\infty(x,y,t)}{\partial x} \text{ in } R \times \Omega \times (0,T),$$

and the equation becomes

$$\frac{\partial u_\infty}{\partial t} + a_\infty(y)\frac{\partial u_\infty}{\partial x} - \frac{\partial}{\partial x}\Big(\int_{[\alpha,\beta]} \varphi(x,y,t;V)\, d\mu_y(-V)\Big) = f(x,y,t) \text{ in } R \times \Omega \times (0,T),$$

with the initial conditions $u(x,y,0) = v(x,y)$ and $\varphi(x,y,0;V) = 0$ in $R\times\Omega$ for $V \in [\alpha,\beta]$.

This example suggests that if a general transport operator with oscillating coefficients is used, one may expect nonlocal effects, but as we lose the commutations properties, one has to find other methods of proofs. In the example, the coefficients are divergence free, so that one could write the equation in conservation form, and the transport operator applied to the coefficients give 0; for fluids the coefficients are the components of u, which is divergence free, but the transport operator applied to u does not give 0, as the gradient of the pressure and the viscous term appear.

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

36. Wednesday April 14.

Physicists often describe some properties of matter at a given frequency, and in the relations that they obtain then the frequency often occurs explicitly; they usually do not bother to explain what could be a general equation valid for all solutions. In linear cases, one can usually give a meaning to these computations, and something like pseudo-differential operators, or nonlocal effects do appear, but not much is understood for nonlinear equations. For example, if one considers MAXWELL equation

$$div(D) = \rho;\ -\frac{\partial D}{\partial t} + curl(H) = j;\ div(B) = 0;\ \frac{\partial B}{\partial t} + curl(E) = 0,$$

one usually assumes that there are relations $D = \varepsilon(x)E$ between the polarization field D and the electric field E (ε is the electric permittivity), and $B = \mu(x)H$ between the induction field B and the magnetic field H (μ is the magnetic susceptibility), and one often adds the relation $j = \sigma(x)E$ (σ is the conductivity), and in that case one forgets about the equation $div(D) = \rho$, which is automatically satisfied if it is true at time 0 because of the relation $\frac{\partial\rho}{\partial t} + div(j) = 0$. The physicists' point of view is to look at solutions of the form $B(x,t) = e^{i\omega\, t}b(x), D(x,t) = e^{i\omega\, t}b(x), \ldots$, so that for example one has $(\sigma + i\omega\,\varepsilon)e + curl(h) = 0$ and one sees a complex conductivity $\sigma + i\omega\,\varepsilon$ appear. The mathematians' point of view is to use LAPLACE transform, and the same equation becomes $(\sigma + p\varepsilon)\mathcal{L}E + curl(\mathcal{L}H) = \varepsilon\, E(\cdot, 0)$. Whatever the point of view, if one considers a mixture of such materials, an Homogenization process usually creates coefficients which depend upon ω or p in a non polynomial way, but in the second case one can look for a convolution equation for linking D and j to E and its history (of course, one imposes the principle of causality, i.e. nonlocal effects must only use the past), and this was done using asymptotic expansions in a periodic framework by Henri SANCHEZ-PALENCIA. So the physicists say that ε depends upon the frequency ω, but the mathematicians go further and try to identify a memory kernel for a convolution equation valid for all solutions and not only for those of the form $e^{i\omega\, t}f(x)$, and this is what we have done on the model examples. However, physicists do use the same approach for nonlinear problems, but mathematicians do not have a general theory for these cases, but one can start following the approach shown below, but I have not solved the bookkeeping problem and the convergence problem.

One could in principle use pseudo-differential operators, which Joseph KOHN and Louis NIRENBERG had introduced for developping a calculus that one can use for expressing the solutions of elliptic equations, or the theory of FOURIER integral operators, which Lars HÖRMANDER developped for similar questions for hyperbolic equations, but these theories have unfortunately been developped only with smooth coefficients, and this is a serious handicap even for linear problems originating in Continuum Mechanics or Physics.

It is not known how to extend to more realistic questions of fluid dynamics the results obtained for the models that I have shown, but it is useful to derive the same results with different methods for which there is more hope for an extension; one of these methods, which physicists often use, is a perturbation method. I consider now a time dependent model problem

$$\frac{\partial u_n}{\partial t} + a_n(x,t)u_n = f(x,t) \text{ in } \Omega\times(0,T);\ u_n(x,0) = v(x) \text{ in } \Omega.$$

Under the assumption that a_n is globally LIPSCHITZ in t, this was first considered by my former student Luisa MASCARENHAS; she had used a time discretization, but the following method is more easy to apply, and although I assumed equicontinuity in t, the result seems valid without such an assumption (it simplifies that having extracted a subsequence such that $a_n(x,t)a_n(x,s)$ converges in $L^\infty(\Omega)$ weak $\star$ for s,t belonging to a countable dense set of Ω, it is then true for all $s,t \in (0,T)$, but only some integrals of the limits is really needed). Assuming that $a_n \rightharpoonup a_\infty$ in $L^\infty(\Omega)$ weak $\star$, one defines $b_n = a_n - a_\infty$ and one considers for a parameter γ the equation

$$\frac{\partial U^n(x,t;\gamma)}{\partial t} + \big(a_\infty(x,t) + \gamma\, b_n(x,t)\big)U^n(x,t;\gamma) = f(x,t) \text{ in } \Omega\times(0,T);\ U^n(x,0;\gamma) = v(x) \text{ in } \Omega,$$

so that the preceding problem corresponds to $\gamma = 1$. Obviously U^n is analytic in γ and we can consider the TAYLOR expansion at $\gamma = 0$,

$$U^n(x,t;\gamma) = \sum_{k=0}^{\infty} \gamma^k V_k^n(x,t) \text{ in } \Omega \times (0,T),$$

and one finds immediately that V_0^n is independent of n, and solution of

$$\frac{\partial V_0(x,t)}{\partial t} + a_\infty(x,t)V_0(x,t) = f(x,t) \text{ in } \Omega \times (0,T);\ V_0(x,0) = v(x) \text{ in } \Omega,$$

and that for $k \geq 1$, V_k^n is solution of

$$\frac{\partial V_k^n(x,t)}{\partial t} + a_\infty(x,t)V_k^n(x,t) + b_n(x,t)V_{k-1}^n(x,t) = 0 \text{ in } \Omega \times (0,T);\ V_k^n(x,0) = 0 \text{ in } \Omega.$$

As $b_n \rightharpoonup 0$ in $L^\infty(\Omega)$ weak $\star$, one sees that V_1^n converges weakly to 0, but one needs the limit of $b_n V_1^n$ in order to compute the limit of V_2^n, and more generally one needs the explicit form of each V_k^n for $k \geq 1$,

$$V_k^n(x,t) = -\int_0^t exp\Big(-\int_s^t a_\infty(x,\sigma)\,d\sigma\Big) b_n(x,s)V_{k-1}^n(x,s)\,ds \text{ for } (x,t) \in \Omega \times (0,T),$$

so that if one defines $R(x,s,t) = exp\big(-\int_s^t a_\infty(x,\sigma)\,d\sigma\big)$, one has $V_1^n(x,t) = -\int_0^t R(x,s,t)b_n(x,s)V_0(x,s)\,ds$ and therefore the limit of $b_n V_1^n$ involves limits of $b_n(x,t)b_n(x,s)$ and has the form $\int_0^t C(x,t,s)V_0(x,s)\,ds$; one can deal similarly with the following terms and integral terms having appeared naturally, one may look for a kernel having the analytic form

$$K(x,t,s,\gamma) = \sum_{k=2}^{\infty} \gamma^k K_k(x,t,s),$$

and it is not difficult to obtains bounds for the functions V_k^n and K_k, and these bounds show that the TAYLOR expansions written have an infinite radius of convergence, and one can take $\gamma = 1$ safely. The formula obtained for the kernel is quite different from the one which was obtained by using the representation formula for PICK functions.

In principle one could do the same type of expansions for some nonlinear problems, but the bookkeeping is quite arduous (and FEYNMANN seems to have introduced his famous diagrams for a similar purpose), and the convergence questions are not so clear (and it is for similar reasons that physicists like to use PADÉ approximants, or other ways to sum divergent series).

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

37. Monday April 19.

In the preceding analysis, I was studying questions of Homogenization for first order differential equations with oscillating coefficients, but the reality of fluid dynamics is a little different, in particular because of viscosty and pressure. Of course, the questions we would like to understand are related to small viscosities, and as this problem is far from being understood now, it is useful to derive simpler models retaining as much as possible of the qualitative properties that we are interested in.

I started in this direction in 1976, and my analysis was based on the fact that the nonlinear term in NAVIER-STOKES equation could be written as $u\times curl(-u)+grad(|u|^2/2)$: I knew from Electromagnetism that force terms in $u\times b$ have the effect of making particles turn, and as I had heard turbulence to be associated with vorticity, I decided to replace $curl(-u)$ with a given oscillating function in order to study its effect. Not knowing what to expect, I decided to start with the stationary case, and I first used the formal method of asymptotic expansions in a periodic setting, so that my problem was

$$-\nu\Delta\, u_\varepsilon+u_\varepsilon\times\frac{1}{\varepsilon}b\Big(\frac{x}{\varepsilon}\Big)+grad(p_\varepsilon)=f,\ div(u_\varepsilon)=0 \text{ in } \Omega;\ u\in H^1_0(\Omega;R^3),$$

for a periodic vector field b. I did the formal computations with Michel FORTIN who was visiting Orsay that year and sharing my office, and the first thing that we noticed was that the average of b must be 0 or we were looking at a different question; in that case we derived an equation satisfied by the first term of the formal expansion. Using then my method of oscillating test functions in Homogenization, I did not have difficulties proving the formal result; it had the interesting feature that although the force is perpendicular to the velocity u_ε and therefore does not work, it induces oscillations in $grad(u_\varepsilon)$ and therefore more energy is dissipated by viscosity (per unit of time, as we are looking at a stationary problem), but the added dissipation which appeared in our equation was not quadratic in $grad(u)$, but quadratic in u, contrary to the usual belief about turbulent viscosity. There was not much difference avoiding the periodicity hypothesis and considering terms of the form $u_\varepsilon\times curl(v_\varepsilon)$ with v_ε converging weakly, but when I wrote it down for a meeting in 1984 I noticed something else, a quadratic effect in a strength parameter λ; the new problem was

$$-\nu\Delta\, u_n+u_n\times curl(v_0+\lambda\, v_n)+grad(p_n)=f,\ div(u_n)=0 \text{ in } \Omega,$$

with $v_0\in L^3(\Omega;R^3)$ and $v_n\rightharpoonup 0$ in $L^3(\Omega;R^3)$ weak, and I did not impose boundary conditions but I assumed that $u_n\rightharpoonup u_\infty$ in $H^1(\Omega;R^3)$ weak (it is a classical requirement in Homogenization that if one wants to speak about the effective properties of a mixture one should obtain a result which is independent of the boundary conditions; if one does not do this, one can only mention the global properties of the mixture and the container). I showed that there exists a symmetric nonnegative matrix M, depending only upon a subsequence of v_n that one may have to extract, such that u_∞ satisfies the equation

$$-\nu\Delta\, u_\infty+u_\infty\times curl(v_0)+\lambda^2 M\, u_\infty+grad(p_\infty)=f,\ div(u_\infty)=0 \text{ in } \Omega,$$

and of course a more precise convergence result is

$$u_n\times curl(v_n)\rightharpoonup \lambda\, M\, u_\infty \text{ in } H^{-1}_{loc}(\Omega;R^3) \text{ weak},$$

and

$$\nu|grad(u_n)|^2\rightharpoonup \nu|grad(u_\infty)|^2+\lambda^2(M u_\infty.u_\infty) \text{ in the sense of measures.}$$

The way M is defined follows my approach to Homogenization, but the quadratic dependence in λ and a particular formula in the case where $div(v_n)=0$ was my first hint about the possibility of defining H-measures, and after I was led to introduce H-measures for another purpose I checked that M could indeed be computed from the H-measures associated to the sequence v_n.

One extracts a subsequence from v_n and one constructs M in the following way: for $k\in R^3$ one solves

$$-\nu\Delta\, w_n+k\times curl(v_n)+grad(q_n)=0,\ div(w_n)=0 \text{ in } \Omega,$$

adding boundary conditions which imply that $w_n \rightharpoonup 0$ in $H^1(\Omega; R^3)$ weak (DIRICHLET conditions, or periodic conditions in the case where v_n is defined in a periodic way, for example); one can indeed extract a subsequence such that this occurs for three independent vectors k and one defines M by

$$w_n \times curl(v_n) \rightharpoonup M\, k \text{ in } H^{-1}_{loc}(\Omega; R^3) \text{ weak.}$$

The first remark is that $u \mapsto u \times curl(v)$ maps continuously $H^1(\Omega; R^3)$ into $H^{-1}(\Omega; R^3)$ if $v \in L^3(\Omega; R^3)$, if the boundary of Ω is smooth, using SOBOLEV imbedding theorem $H^1(\Omega) \subset L^6(\Omega)$; indeed if $u, \varphi \in H^1(\Omega)$, then $u\,\varphi \in W^{1,3/2}(\Omega)$; if $f \in H^{-1}(\Omega; R^3)$ one finds that, after eventually adding a constant, p_n is bounded in $L^2_{loc}(\Omega)$, and one can assume that $p_n \rightharpoonup p_\infty$ in $L^2_{loc}(\Omega)$ weak. Similarly, in the problem for w_n one can assume that $q_n \rightharpoonup 0$ in $L^2_{loc}(\Omega)$ weak. Using elliptic regularity theory (and CALDERÓN-ZYGMUND theorem), $grad(w_n)$ is bounded in $L^3_{loc}(\Omega; R^3)$ and therefore $w_n \to 0$ in $L^p_{loc}(\Omega; R^3)$ strong for every $p < \infty$; one has then a better convergence for $w_n \times curl(v_n)$, which converges to $M\,k$ in $H^{-1}_{loc}(\Omega; R^3)$ strong, because of writing products of the form $(w_n)_i \frac{\partial (v_n)_j}{\partial x_k}$ as $\frac{\partial [(w_n)_i (v_n)_j]}{\partial x_k} - \frac{\partial (w_n)_i}{\partial x_k}(v_n)_j$ and terms like $(w_n)_i (v_n)_j$ converge strongly to 0 in $L^2_{loc}(\Omega)$ and terms like $\frac{\partial (w_n)_i}{\partial x_k}(v_n)_j$ are bounded in $L^{3/2}_{loc}(\Omega)$ and converge strongly to 0 in $L^q_{loc}(\Omega)$ for every $q > 3/2$ and therefore in $H^{-1}_{loc}(\Omega)$ strong. One applies the method of oscillating test functions, multiplying the equation for u_n by $\varphi\, w_n$ and the equation for w_n by $\varphi\, u_n$, with $\varphi \in C^1_c(\Omega)$, and noticing that $div(\varphi\, w_n) = (grad(\varphi).w_n)$ and $div(\varphi\, u_n) = (grad(\varphi).u_n)$ and therefore the estimates on the pressures p_n and q_n are needed. One assumes that $u_n \times curl(v_n) \rightharpoonup g$ in $H^{-1}(\Omega; R^3)$ weak and one wants to identify g; one finds

$$\lim_{n\to\infty} \int_\Omega \nu\varphi\Big(grad(u_n).grad(w_n)\Big)\,dx + \lambda\langle \varphi\, u_n \times curl(v_n), w_n\rangle = 0,$$

and

$$\lim_{n\to\infty} \int_\Omega \nu\varphi\Big(grad(u_n).grad(w_n)\Big)\,dx = \langle \varphi\, g, k\rangle,$$

but as

$$\langle \varphi\, u_n \times curl(v_n), w_n\rangle = -\langle \varphi\, w_n \times curl(v_n), u_n\rangle \to -\langle \varphi\, M\,k, u_\infty\rangle,$$

one has shown that

$$g = \lambda\, M^T\, u_\infty.$$

The fact that M is symmetric follows easily by the same method, w'_n being the solution for k', multiplying the equation for w'_n by $\varphi\, w_n$, the equation for w_n by $\varphi\, w'_n$, and comparing. The limit of $\nu|grad(u_n)|^2$ is obtained by multiplying the equation for u_n by $\varphi\, u_n$.

In the case where $div(v_n) = 0$, which is the case for fluid dynamics, one can take

$$w_n = \Big(k.grad(z_n)\Big),$$

where z_n solves

$$-\nu\Delta\, z_n = v_n,$$

and therefore

$$\nu \sum_{l,m} \frac{\partial^2 (z_n)_l}{\partial x_i \partial x_m} \frac{\partial^2 (z_n)_l}{\partial x_j \partial x_m} \rightharpoonup M_{ij} \text{ in } L^{3/2}(\Omega) \text{ weak.}$$

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

38. Wednesday April 21.

The same analysis can be done for the evolution problem if one can obtain a bound for the pressure, so I initially did it for the whole space; VON WAHL later told me that he had shown by semi-group methods that one can obtain estimates on the pressure for any smooth bounded open set. In my original proof for the evolution case, I had not seen how to prove that the matrix M corresponding to the added dissipation is symmetric; two years ago, I worked with Chun LIU and Konstantina TRIVISA about extending the formula using H-measures to the evolution case, and we first checked the symmetry, but then we noticed that one needed a new variant of H-measures, with a parabolic scaling, i.e. instead of identifying rays $s\xi$ through a nonzero ξ with $s > 0$, one had to identify curves $(s\xi, s^2\tau)$ through a nonzero (ξ, τ).

Of course, we should not lose sight of the reasons why the preceding models were chosen and the previous computations were done: the initial purpose was to understand what was the adapted weak type topology for the velocity, appearing as coefficients of a transport equation, with or without viscosity. Starting from a model without viscosity, we saw that various nonlocal terms could appear, and this analysis could also be useful for correcting the defects of NAVIER-STOKES equations, but the class to consider should then at least contain some equations with memory effects. Starting with a model with viscosity but magnifying the ocillations possible for the velocity field, we have seen some lower order terms appear, and it could have some analogy with the framework using affine connections which geometers have advocated. Obviously one should improve the model, but it is worth mentioning that the formula using H-measures which gives an explicit form for M has a $\frac{1}{\nu}$ in front of an integral on the unit sphere, and although it is tempting to rescale the equation, one should remember that turbulence is supposed to show an infinite number of length scales, and that an object like H-measures which mixes different frequencies cannot reasonable describe that.

Before being able to describe the tool of H-measures and the various formulas that one can deduce from it, it is worth starting with the previous theory, which I had developped with François MURAT in the late 70s, Compensated Compactness. I make a distinction between the basic Compensated Compactness theorem, described below, and the Compensated Compactness Method, which I developpeed after, and which is more general. The basic theorem, which I call the quadratic theorem, is the following.

Theorem: Let $U^n \rightharpoonup U^\infty$ in $L^2_{loc}(\Omega; R^p)$ weak, and assume that

$$\sum_{jk} A_{ijk} \frac{\partial U_j^n}{\partial x_k} \text{ stays in a compact of } H^{-1}_{loc}(\Omega) \text{ for } i = 1, \ldots, q.$$

Define the two characteristic sets $\mathcal{V}$ and Λ

$$\mathcal{V} = \Big\{(\lambda, \xi) \in R^p \times (R^N \setminus 0) : \sum_{jk} A_{ijk}\lambda_j \xi_k = 0 \text{ for } i = 1, \ldots, q\Big\},$$

$$\Lambda = \Big\{\lambda \in R^p : \text{ there exists } \xi \in R^N \setminus 0, (\lambda, \xi) \in \mathcal{V}\Big\}.$$

Let Q be a quadratic form on R^p satisfying

$$Q(\lambda) \geq 0 \text{ for all } \lambda \in \Lambda,$$

then

$$Q(U^n) \rightharpoonup Q(U^\infty) + \nu \text{ in the sense of measures implies } \nu \geq 0.$$

I have already mentioned the Div-Curl lemma, which I had found with François MURAT in 1974 in connection with Homogenization: it is the particular case where $U = (E, D)$ with the list of differential information corresponding to $div(D)$ and the components of $curl(E)$; then $\mathcal{V} = \{(E, D, \xi)$ with ξ parallel to E and orthogonal to $D\}$ and $\Lambda = \{(E, D)$ with $(E.D) = 0\}$; then the quadratic form Q_0 defined by

$Q_0(E,D) = (E.D)$ is 0 on Λ, and therefore by applying the theorem to $\pm Q_0$ one deduces that $(E^n.D^n)$ converges to $(E^\infty.D^\infty)$ in the sense of measures. The proof of the quadratic theorem mimicks the one we had found for the Div-Curl lemma, using FOURIER transform and PLANCHEREL formula.

I spent the year 1974/75 in Madison, and Joel ROBBIN taught me how to translate some of my results in the language of differential forms, and he showed me a new proof of the Div-Curl lemma using HODGE theory. We did not notice that the same method was giving the properties of sequential weak continuity of Jacobian determinants, which MORREY had actually proved in the 50s, and I only learned these results from John M. BALL in the Fall of 1975, mistakenly believing that he had proved them.

Later, Jacques-Louis LIONS asked François MURAT to extend the Div-Curl lemma (and he gave him an article by SCHULENBERGER and WILCOX which he thought related), and François extended it first to a bilinear setting, i.e. $U = (V,W)$ with some differential list for V and a differential list for W, and then he looked at the bilinear forms $B(V,W)$ which are sequentially weakly continuous. I pointed out that the splitting of U and the restriction to bilinear forms was not natural, and therefore François MURAT proved the above theorem in the case where $Q(\lambda) = 0$ for all $\lambda \in \Lambda$ (deducing that $\nu = 0$ in this case); however, his proof was a little different than the one we had followed for the Div-Curl lemma, probably because he had also extended the Div-Curl lemma itself to a (L^p, L^q) setting (for which he used the MIKHLIN-HÖRMANDER theorem on $\mathcal{F}L^p$ multipliers, and one needs to check the smoothness of the multiplier), and a similar approach forced him to impose an hypothesis of constant rank: if for each $\xi \in R^N \setminus 0$ one denotes $\Lambda_\xi = \left\{\lambda \in R^p : \sum_{jk} A_{ijk}\lambda_j\xi_k = 0 \text{ for } i = 1,\ldots,q\right\}$, so that Λ is the union of the subspaces Λ_ξ, the constant rank hypothesis imposes that the dimension of Λ_ξ is independent of ξ. It was Jacques-Louis LIONS who then coined the term Compensated Compactness for the type of result that François had obtained, as it looked like a compactness argument because one could deduce the weak limit of a nonlinear quantity, but it was the result of a compensation effect.

I extended then the result as shown above, and it has the following consequence: if $U_i^n U_j^n \rightharpoonup U_i^\infty U_j^\infty + R_{ij}$ in the sense of measures, then this defines a symmetric matrix R whose entries may be RADON measures: if all R_{ij} are integrable functions, one has

$$R(x) \text{ belongs to the closed convex hull of } \{\lambda \otimes \lambda, \lambda \in \Lambda\} \text{ a.e. } x \in \Omega,$$

and the general case is similar once one uses RADON-NIKODYM theorem: if $\tau = \sum_k R_{kk}$, then $R_{ij} = \rho_{ij}\tau$ with the functions ρ_{ij} being τ-integrable, and it is $\rho(x)$ which belongs to the convex hull of the elements of the form $\lambda \otimes \lambda$ for $\lambda \in \Lambda$, and this property holds τ-almost everywhere. A point in the convex hull of a set K is the center of mass of a probability measure with support on K, and the theory of H-measures, which I developed in the late 80s, extends the Compensated Compactness theorem and gives explicitly a way to describe these probability measures; the interest is that the H-measures are measures in (x,ξ) and that they permit to extend the Compensated Compactness theorem to the case of differential equations with variable coefficients, and in some situations the H-measures satisfy partial differential equations in (x,ξ). There are however parts of the Compensated Compactness Method which still require improvements, as there is not yet a characterization of which pairs (YOUNG measures/H-measures) can be created, although I have obtained some information in this direction with François MURAT).

21-820. PDE Models in Oceanography
Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

39. Friday April 23.

My analysis of problems of Continuum Mechanics in the 70s was that there was a dichotomy between the constitutive relations which are possibly nonlinear pointwise constraints of the form

$$U(x) \in K \text{ a.e. } x \in \Omega,$$

with K eventually depending upon x (with possible oscillations requiring techniques from Homogenization), and the balance equations which are linear differential constraints of the form

$$\sum_{jk} A_{ijk} \frac{\partial U_j}{\partial x_k} = f_i \text{ in } \Omega, \text{ for } i = 1, \ldots, q.$$

Perhaps because I had learned some Continuum Mechanics as a student, I knew that Elasticity meant

$$\rho(x) \frac{\partial^2 u_i}{\partial t^2} - \sum_j \frac{\partial \sigma_{ij}}{\partial x_j} = f_i \text{ in } \Omega,$$

and in that case U would contain the components of the momentum $\rho \frac{\partial u}{\partial t}$, the strain ∇u, the stress σ (as I do not remember hearing the term PIOLA-KIRCHHOFF stress from my studies, it might be that I was told mostly the Eulerian point of view, where the symmetric CAUCHY stress appears), and the constitutive relations relate σ and ∇u, while the list of balance equations contain the equilibrium equation above and the compatibility conditions related to using gradients (after I developped the Compensated Compactness Method in 1977, I added "entropies" to that description; it may be useful to point out that "entropies" have nothing to do with the fact that one considers an evolution equation, or that one is interested in hyperbolic problems, as it is just a name for designing supplementary differential equations which are consequences of those already written for smooth solutions; certainly Peter LAX could have chosen a better name, and geometers call them Casimirs).

It is of course a handicap that the Compensated Compactness theorem cannot handle variable coefficients, but H-measures do not suffer from this defect: each first order partial differential equation written in conservative form $\sum_{jk} \frac{\partial (A_{jk}(x) U_j^n)}{\partial x_k} \to f$ in $H^{-1}_{loc}(\Omega)$ strong, with the coefficients A_{jk} being continuous, can be seen by H-measures (and the Localization Principle implies $\sum_{jk} \xi_k A_{jk}(x) \mu^{jl} = 0$ for every l).

YOUNG measures cannot take into account partial differential equations, but it might be because I used them in order to describe some constraints that they must satisfy as a consequence of the quadratic Compensated Compactness theorem, that some may have misunderstood their role. Suppose for example that a sequence U^n is bounded in L^∞, corresponds to a YOUNG measure $\nu_x, x \in \Omega$, and satisfies a partial differential equation with constant coefficient $\sum_{jk} A_{jk} \frac{\partial U_j^n}{\partial x_k} = 0$; decompose R^N as a union of cubes of size $1/n$ and for each of these cubes chose a rigid displacement mapping the cube onto itself and transport the values of U^n accordingly, and let V^n be any of the new functions obtained this way; it is not difficult to check that V^n corresponds to the same YOUNG measure $\nu_x, x \in \Omega$, than U^n; however V^n is unlikely to solve the same partial differential equation than U^n, and therefore YOUNG measures cannot feel if the sequence that they analyze satisfies or not a given partial differential equation. A different way to express the same idea is to notice that in defining YOUNG measures the only important property of Ω is to be endowed with a nonnegative measure without atoms, like the LEBESGUE measure, and therefore the structure of differentiable manifold is not seen by YOUNG measure. However what the Compensated Compactness theorem says can be expressed in terms of YOUNG measures, as it says that if Q is quadratic and satisfies $Q(\lambda) \geq 0$ for all $\lambda \in \Lambda$, then the limit of $Q(U^n)$ is $\langle \nu_x, Q \rangle$, while the limit of U^n is $\langle \nu_x, id \rangle$ (where id is the identity mapping), and therefore one has the following inequality (reminiscent of JENSEN's inequality, which says that is one replaces Q by any convex function the following inequality is true)

$$\langle \nu_x, Q \rangle \geq Q(\langle \nu_x, id \rangle) \text{ a.e. } x \in \Omega.$$

If $\Lambda = \{0\}$, one is looking at an elliptic system or some overdetermined system; the ellipticity of the system corresponds to saying that for every $\xi \neq 0$, the linear mapping $U \mapsto V$ with $V_i = \sum_{jk} A_{ijk} U_j \xi_k$ for $i = 1, \ldots, q$, is invertible (so that $q = p$); if it is not the case then one necessarily has $q > p$, but that by itself is not enough to imply that $\Lambda = \{0\}$. In the case $\Lambda = \{0\}$ one has $U^n \to U^\infty$ in $L^2_{loc}(\Omega)$ strong, because one can take Q positive definite, and as Q is 0 on Λ one finds that $Q(u^n) \rightharpoonup Q(U^\infty)$ in the sense of measures; if B is the symmetric bilinear form associated to Q one has then $Q(U^n - U^\infty) = Q(U^n) - 2B(U^n, U^\infty) + Q(U^\infty)$, and therefore $Q(U^n - U^\infty) \rightharpoonup 0$ in the sense of measures. The case $\Lambda = \{0\}$ corresponds then to using a compactness argument.

The case $\Lambda = R^p$ corresponds to using a convexity argument, as $Q \geq 0$ is the same as Q convex for quadratic forms; this happens if there is no differential equation ($q = 0$), but also for some list of differential equations that are not constraining enough: if U^n consists of the list or k vector fields whose divergence is controlled, then one has $\Lambda = R^p$ if $k < N$, but $\Lambda \neq R^p$ if $k \geq N$.

In the case of the Div-Curl lemma, the information about $curl(E)$ gives rises to the equations $\xi_i E_j - \xi_j E_i = 0$ for all i, j, i.e. ξ parallel to E, and the information about $div(D)$ gives ξ orthogonal to D, so that Λ is the set of (E, D) with D orthogonal to E. This case is related to the monotonicity method.

The case of MAXWELL equation is more intrincate, the dual variable is (τ, ξ) with $|\tau| + |\xi| \neq 0$ and the equations are $(\xi.D) = 0$, $-\tau\, D + \xi \times H = 0$, $(\tau.B) = 0$, $\tau\, B + \xi \times E = 0$; if $\tau \neq 0$, one may assume $\tau = 1$, i.e. one has $D = \xi \times H$ and $B = -\xi \times E$, so that D is orthogonal to H, B is orthogonal to E, and $(E.D) = (B.H)$; if $\tau = 0$, then H and E must be parallel to E and H, and one still has $(E.B) = (D.H) = (E.D) - (B.H) = 0$; one can check that this is exactly the description of Λ. We will see later why these quantities are natural, using the framework of differential forms.

It is easy to see that the condition on Q in the theorem is necessary. More generally, let F be a (continuous) function on R^p: if one wants that for all $U^n \rightharpoonup U^\infty$ in $L^\infty(\Omega; R^p)$ weak $\star$, satisfying the equations with $f_i = 0$ for $i = 1, \ldots, q$, and such that $F(u^n) \rightharpoonup F(U^\infty) + \nu$ in $L^\infty(\Omega)$ weak $\star$, one can deduce that $\nu \geq 0$, then it is necessary that F be convex in every direction of Λ, i.e. for every $a \in R^p$ and every $\lambda \in \Lambda$ the mapping $t \mapsto F(a + t\,\lambda)$ should be convex in $t \in R$ (this is not always sufficient, but the theorem says that it is if the function is quadratic). Indeed, for $(\lambda, \xi) \in \mathcal{V}$ the function U defined by $U(x) = a + \lambda\varphi\big((\xi.x)\big)$ satisfies $\sum_{jk} A_{ijk} \frac{\partial U_j}{\partial x_k} = (\sum_{jk} A_{ijk} \lambda_j \xi_k)\varphi'\big((\xi.x)\big) = 0$, and this stays true if φ is replaced by a characteristic function χ, so that U takes only the values a and $b = a + \lambda$; if one uses a sequence of characteristic functions χ_n converging to $\theta \in (0,1)$ in $L^\infty(R)$ weak $\star$, then $U^n \rightharpoonup U^\infty = a + \theta\,\lambda = (1-\theta)a + \theta\, b$ in $L^\infty(\Omega; R^p)$ weak $\star$ and $F(U^n) = (1 - \chi_n)F(a) + \chi_n F(b) \rightharpoonup (1-\theta)F(a) + \theta\, F(b)$ in $L^\infty(\Omega)$ weak $\star$, and the result follows easily. If one wants to deduce that $F(u^n) \rightharpoonup F(U^\infty)$ in $L^\infty(\Omega)$ weak $\star$, it is necessary that F be affine in every direction of Λ; this is sufficient for quadratic functions but not always sufficient in general; in the case where Λ spans all R^p then F must be a combination of multilinear forms and in particular it should be a polynomial of degree at most p, but there are other necessary conditions which imply that the degree can be at most N.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

40. Monday April 26.

We have seen that there are three linearly independent quadratic forms which are 0 on the characteristic set Λ for MAXWELL equation, $(E.B)$, $(H.D)$ and $(E.D)-(H.B)$; Λ is actually the intersection of the zero sets of these three quadratic forms. In discussing the interpretation of these quantities in terms of differential forms, it is worth recalling the analogy between MAXWELL equation and the equation for fluids. The transport term $\frac{\partial u}{\partial t}+u.\nabla u$ can be written as $\frac{\partial u}{\partial t}+u\times curl(-u)+grad(|u|^2/2)$, and there may also exist a CORIOLIS term $u\times 2\Omega$ so that a part of it has the same form than a part of the LORENTZ force $E+u\times B$ that one encounters in ElectroMagnetism; if one denotes $e=\frac{\partial u}{\partial t}+grad\big(\Phi+\frac{p}{\rho_0}+\frac{|u|^2}{2}\big)$ and $b=curl(-u)+2\Omega$, then one has $div(b)=0$ and $\frac{\partial b}{\partial t}+curl(e)=0$, in the case where the viscosity is 0 and there is no exterior force other than those related to the geopotential Φ. In ElectroMagnetism one introduces the scalar and vector potentials V,A, such that $B=-curl(A)$ and $E=\frac{\partial A}{\partial t}-grad(V)$, and therefore one has similar relations if one denotes $a=u-u^*$ with u^* being a velocity field such that $curl(u^*)=2\Omega$, and $v=-\Phi-\frac{p}{\rho_0}-\frac{|u|^2}{2}$.

In terms of differential forms, let the 1-form α be defined by $\alpha=V\,dt+\sum_j A_j\,dx_j$, and therefore $d\alpha=\sum_j\frac{\partial V}{\partial x_j}\,dx_j\wedge dt+\sum_j\frac{\partial A_j}{\partial t}\,dt\wedge dx_j+\sum_{jk}\frac{\partial A_j}{\partial x_k}\,dx_k\wedge dx_j$, so that the 2-form $\beta=d\alpha$ satisfies $\beta=\sum_i E_i\,dt\wedge dx_i-B_1\,dx_2\wedge dx_3-B_2\,dx_3\wedge dx_1-B_3\,dx_1\wedge dx_2$; the equations $div(B)=0$ and $\frac{\partial B}{\partial t}+curl(E)=0$ simply mean $d\beta=0$. Similarly, let the 2-form γ be defined by $\gamma=\sum_i H_i\,dt\wedge dx_i+D_1\,dx_2\wedge dx_3+D_2\,dx_3\wedge dx_1+D_3\,dx_1\wedge dx_2$, then $d\gamma=\sum_{i,j}\frac{\partial H_i}{dx_j}\,dx_j\wedge dt\wedge dx_i+\frac{\partial D_1}{\partial t}\,dt\wedge dx_2\wedge dx_3+\frac{\partial D_2}{\partial t}\,dt\wedge dx_3\wedge dx_1+\frac{\partial D_3}{\partial t}\,dt\wedge dx_1\wedge dx_2+div(D)\,dx_1\wedge dx_2\wedge dx_3$, and therefore if one defines the 3-form δ by $\delta=\rho\,dx_1\wedge dx_2\wedge dx_3-j_1\,dt\wedge dx_2\wedge dx_3-j_2\,dt\wedge dx_3\wedge dx_1-j_3\,dt\wedge dx_1\wedge dx_2$, then the equations $div(D)=\rho$ and $-\frac{\partial D}{\partial t}+curl(H)=j$ simply mean $d\gamma=\delta$. The conservation of charge is $d\delta=0$ and therefore there must exist a 2-form γ with $d\gamma=\delta$, and naming the six coefficients of γ leads us to introduce the components of D and H; studying the movement of charged particles leads to discover the LORENTZ force $E+u\times B$, and the components of E and B appear to be the coefficients of an exact two form β, and therefore the existence of the 1-form α. [I learned about this interpretation of MAXWELL equations in terms of differential forms from Joel ROBBIN, and then I heard Laurent SCHWARTZ mention it in a talk, where he condidered the vacuum with $\varepsilon_0=\mu_0=1$, and instead of $d\gamma=\delta$ he wrote $d^*\beta=0$; it is actually important not to identify E and D or H and B, even if that is possible in the vacuum, because in presence of matter they do play different roles].

Now we can identify easily what the quanties $(E.B)$, $(H.D)$ and $(E.D)-(H.B)$ mean in terms of the 2-forms $\beta=\sum_i E_i\,dt\wedge dx_i-B_1\,dx_2\wedge dx_3-B_2\,dx_3\wedge dx_1-B_3\,dx_1\wedge dx_2$ and $\gamma=\sum_i H_i\,dt\wedge dx_i+D_1\,dx_2\wedge dx_3+D_2\,dx_3\wedge dx_1+D_3\,dx_1\wedge dx_2$: one has $\beta\wedge\beta=-2(E.B)\,dt\wedge dx_1\wedge dx_2\wedge dx_3$, and similarly $\gamma\wedge\gamma=2(H.D)\,dt\wedge dx_1\wedge dx_2\wedge dx_3$, and $\beta\wedge\gamma=[(E.D)-(H.B)]\,dt\wedge dx_1\wedge dx_2\wedge dx_3$.

Of course, the quadratic theorem of Compensated Compactness says that the situation encountered for MAXWELL equation is general: if a sequence of p-forms a^n converges weakly to a^∞ and a sequence of q-forms b^n converges weakly to b^∞ then $a^n\wedge b^n$ converges weakly to $a^\infty\wedge b^\infty$ if the convergences hold for $L^2_{loc}(\Omega)$ weak for the coefficients of a^n and b^n, and if $d\,a^n$ and $d\,b^n$ have their coefficients staying in compact sets of $H^{-1}_{loc}(\Omega)$. For proving this result one must compute what the characteristic set Λ is, and one sees easily that if for $\xi\neq 0$ one defines the (alternated) linear form $\eta=\sum_i\xi_i\,dx_i$, then $\Lambda=\{a,b\colon a$ is an alternated p-linear form with $a\wedge\eta=0$, and b is an alternated q-linear form with $b\wedge\eta=0\}$; it is an elementary result that $a\wedge\eta=0$ if and only if $a=\eta\wedge c$ for some alternated $(p-1)$-linear form and $b\wedge\eta=0$ if and only if $b=\eta\wedge d$ for some alternated $(q-1)$-linear form and therefore $\eta\wedge\eta=0$ implies that $a\wedge b=0$ on Λ.

In order to reiterate the result, it is useful to notice that $d(a^n\wedge b^n)=(d\,a^n)\wedge b^n+(-1)^p a^n\wedge(d\,b^n)$. It is also useful to notice that the preceding result holds if the coefficients of a^n converge in $L^r_{loc}(\Omega)$ weak with the coefficients of $d\,a^n$ staying in a compact of $W^{-1,r}_{loc}(\Omega)$, if the convergence of b^n converge in $L^s_{loc}(\Omega)$ weak with the coefficients of $d\,b^n$ staying in a compact of $W^{-1,s}_{loc}(\Omega)$, with $1<r,s<\infty$ and $\frac{1}{r}+\frac{1}{s}\leq 1$. François MURAT proved the Div-Curl lemma in such a situation, and the proof involves either CALDERÓN-ZYGMUND theorem or the theorem on FOURIER multipliers of MIKHLIN-HÖRMANDER (which require a constant rank hypothesis, satisfied in our case), but the general case can also be proved along the lines proposed earlier by Joel ROBBIN, using HODGE decomposition.

One finds as a particular case the result of MORREY that Jacobian determinants are sequentially weakly continuous: if $a_j^n = d\,u_j^n$ converges to $a_j^\infty = d\,u_j^\infty$ for $j = 1, \ldots, k$, (with obvious constraints on the values of $p_j, j = 1, \ldots, k$, if the convergence of the coefficients of a_j^n holds in $L_{loc}^{p_j}(\Omega)$ weak), then $a_1^n \wedge \ldots \wedge a_k^n$ converges weakly to $a_1^\infty \wedge \ldots \wedge a_k^\infty$. This result is more easy to prove because the 1-forms used are exact and a simple proof can be obtained by integration by parts, as the formula $d(a \wedge b) = (d\,a) \wedge b - a \wedge (d\,b)$ means here $\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_k} - \frac{\partial u}{\partial x_k}\frac{\partial v}{\partial x_j} = \frac{\partial}{\partial x_j}\big(u\frac{\partial v}{\partial x_k}\big) - \frac{\partial}{\partial x_k}\big(u\frac{\partial v}{\partial x_j}\big)$. Using SOBOLEV imbedding theorem one can deduce that Jacobian determinants of size k are not only defined for functions in $L_{loc}^k(\Omega)$ but also for $L_{loc}^p(\Omega)$ with $p \geq k\,N/(N+1)$, and the sequential weak continuity holds if $p > k\,N/(N+1)$. One can of course take the functions in various spaces $W_{loc}^{1,p_j}(\Omega)$, with a corresponding relation for the exponents p_j, but another improvement is sometime useful and involves HARDY spaces, an idea which I believe is due to Pierre-Louis LIONS, with a basic theorem obtained with Ronald COIFMAN, Yves MEYER and Stephen SEMMES (although their false claim to have improved the Compensated Compactness Method shows that they did not understand at the time what it was about).

In the Div-Curl lemma, one cannot replace the weak convergence in the sense of measures by a convergence in $L_{loc}^1(\Omega)$ weak (for $N \geq 2$, of course); I had constructed a counter-example in the following way. Let Ω be a smooth bounded open set of R^N, and let ω be an open set whose closure is contained in Ω ($\omega \neq \emptyset$, of course); one chooses then a sequence f_n converging weakly to 0 in $H^{1/2}(\partial\Omega)$ but not strongly (it is here that the hypothesis $N \geq 2$ is used, so that $H^{1/2}(\partial\Omega)$ is indeed an infinite dimensional HILBERT space); one solves $-\Delta\,u_n = 0$ in ω with the trace of u_n on $\partial\omega$ being f_n, and the sequence u_n converges weakly to 0 in $H^1(\omega)$ but not strongly. One uses a linear continuous extension P from $H^1(\omega)$ into $H_0^1(\Omega)$ and one takes $E_n = grad(P\,u_n)$. One also solves the equation $-\Delta\,v_n = 0$ in $\Omega \setminus \overline{\omega}$ with $v_n = 0$ on $\partial\Omega$ and $\frac{\partial v_n}{\partial\nu} = \frac{\partial u_n}{\partial\nu}$ on $\partial\omega$, where ν is the normal to $\partial\omega$, which in variational formulation means $\int_{\Omega\setminus\overline{\omega}}\big(grad(v_n).grad(w)\big)\,dx = -\int_\omega\big(grad(u_n).grad(w)\big)\,dx$ for every $w \in H_0^1(\Omega)$, and one takes $D_n = grad(u_n)$ in ω and $D_n = grad(v_n)$ in $\Omega \setminus \omega$. Then E_n and D_n converge weakly to 0 in $L^2(\Omega; R^N)$, and although $curl(E_n) = 0$ and $div(D_n) = 0$, one has $\lim_{n\to\infty}\int_\omega(E_n.D_n)\,dx > 0$. The quadratic theorem of Compensated Compactness says that for $\varphi \in C_c(\Omega)$ one has $\int_\Omega \varphi(E_n.D_n)\,dx \to 0$, and the preceding counter-example shows that one does not have the same result if $\varphi = \chi_\omega$, the characteristic function of ω.

The case of Jacobian determinants gives an example where there are some polynomials of degree more than 2 which are sequentially weakly continuous (and $\Lambda \neq 0$, of course). We have seen that a necessary condition that a (continuous) real function F on R^p be such that $F(U^n) \rightharpoonup F(U^\infty)$ in $L^\infty(\Omega)$ weak $\star$ for all sequences U^n converging to U^∞ in $L^\infty(\Omega; R^p)$ weak $\star$ and satisfying the equations $\sum_{jk} A_{ijk}\frac{\partial U_j^n}{\partial x_k} = 0$ for $i = 1, \ldots, q$, is that F must be affine in all directions of Λ, but there are in general other necessary conditions. In the case where Λ spans R^p, the preceding necessary condition implies that F is a combination of multilinear forms, of degree at most p, while the new conditions will imply that the degree is at most N. The basic idea can be shown on the following example, which I encountered while studying oscillating sequences of solutions of discrete kinetic velocity models like the BROADWELL model. Let $N = 2$, $p = 3$ and the list of equations being $\frac{\partial U_1^n}{\partial x_1} = 0, \frac{\partial U_2^n}{\partial x_2} = 0, \frac{\partial U_3^n}{\partial x_1} + \frac{\partial U_3^n}{\partial x_2} = 0$, so that the characteristic set Λ is defined by $\xi_1 U_1 = \xi_2 U_2 = (\xi_1 + \xi_2)U_3 = 0$, and as ξ must be different from 0, one finds the three axes; the only candidates for sequential weak continuity are then U_1U_2, U_1U_3, U_2U_3, and $U_1U_2U_3$; the first three are quadratic and therefore they are sequentially weakly continuous, but the fourth one is not, and this is seen by taking the sequence $U_1^n(x) = \cos(n\,x_2), U_2^n(x) = \cos(n\,x_1), U_3^n(x) = \cos\big(n(x_1 - x_2)\big)$, so that U^n converges weakly to 0 but $U_1^nU_2^nU_3^n(x) = \cos^2(n\,x_1)\cos^2(n\,x_2) - \sin(2n\,x_1)\sin(2n\,x_2)/4$, and therefore $U_1^nU_2^nU_3^n$ converges weakly to $1/4$.

In the general case, the new necessary conditions uses the characteristic set $\mathcal{V} = \{(\lambda,\xi) \in R^p \times (R^N \setminus 0) : \sum_{jk} A_{ijk}\lambda_j\xi_k = 0$ for $i = 1, \ldots, q\}$, and not only its projection Λ; taking combinations of m functions of the form $\lambda\,f\big(n(\xi.x)\big)$ with $(\lambda,\xi) \in \mathcal{V}$, one finds that if $(\lambda^1,\xi^1), \ldots, (\lambda^m,\xi^m) \in \mathcal{V}$ with $rank(\xi^1, \ldots, \xi^m) \leq m-1$, then $F^{(m)}(a)(\lambda^1, \ldots, \lambda^m) = 0$ for all $a \in R^p$ (the case $m = 2$ corresponds to the previously used necessary condition).

Coming back to MAXWELL equations, there are other sequentially weakly continuous quantities if one uses the 1-form $\alpha = V\,dt + \sum_i A_i\,dx_1$, as for example $\alpha \wedge \beta = -(A.B)dx_1 \wedge dx_2 \wedge dx_3 - C_1\,dt \wedge dx_2 \wedge dx_3 - C_2\,dt \wedge dx_3 \wedge dx_1 - C_3\,dt \wedge dx_1 \wedge dx_2$, where $C = V\,B + A \times E$. In the case of fluids with viscosity $\nu = 0$,

A is replaced by $u - u^*$, B by $curl(-u) + 2\Omega$, and $(A.B)$ by $(u - u^*, curl(-u) + 2\Omega)$, i.e the helicity in the case where the CORIOLIS force is neglected. This shows then that the helicity is a robust quantity, not too sensitive to oscillations, and therefore useful even in turbulent flows.

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

41. Wednesday April 28.

It is not always easy to characterize the quadratic forms which are nonnegative on a characteristic cone Λ, but in the case where $\Lambda = \{(E,D) \in R^N \times R^N : (E.D) = 0\}$, $Q \geq 0$ on Λ is equivalent to the existence of $c \in R$ such that $Q(E,D) + c(E.D) \geq 0$ for all $(E,D) \in R^N \times R^N$.

As was first pointed out to me by Joel ROBBIN, the natural generalization of this result is that if Q_0 is a nondefinite nondegenerate quadratic form and Q is a quadratic form such that $Q(\lambda) \geq 0$ whenever $Q_0(\lambda) = 0$, then there exists $c \in R$ such that $Q + c\,Q_0$ is nonnegative everywhere, i.e. convex (I think that Denis SERRE proved something similar, a few years afterward). Indeed, one first defines α to be the minimum of $Q(\lambda)$ for all $\lambda \in R^p$ satisfying $Q_0(\lambda) = 0$ and $|\lambda| = 1$ (and such λ exist as Q_0 is assumed to be nondefinite); as $Q(\lambda) - \alpha|\lambda|^2$ satisfies the hypothesis, one may therefore assume that $\alpha = 0$. Let e_1 be a unit vector where $Q_0(e_1) = 0$ and $Q(e_1) = 0$; as Q_0 is nondegenerate, one has $Q_0'(e_1) \neq 0$, and let $Q_0'(e_1) = 2\beta\,e_2$ for a unit vector e_2 and $\beta \neq 0$ (e_2 is orthogonal to e_1 because $(Q_0'(e_1).e_1) = 2Q_0(e_1) = 0$); one has $Q_0(x) = 2\beta\,x_1x_2 + R_0(x_2,\ldots,x_p)$ with R_0 quadratic, and $Q(x) = 2x_1L(x_2,\ldots,x_p) + R(x_2,\ldots,x_p)$ with L linear and R quadratic. For every $(y_2,\ldots,y_p)$ with $y_2 \neq 0$, one chooses $y_1 = -R_0(y_2,\ldots,y_p)/2\beta\,y_2$ so that $Q_0(y) = 0$ and therefore $Q(y) \geq 0$, i.e. $R(y_2,\ldots,y_p) - R_0(y_2,\ldots,y_p)L(y_2,\ldots,y_p)/\beta\,y_2 \geq 0$; letting y_2 tend to 0 shows that $L(y_2,\ldots,y_p)$ must be $\gamma\,y_2$ for some $\gamma \in R$ (if this was not true $R_0(y_2,\ldots,y_p)$ would be divisible by y_2, but as Q_0 is nondegenerate, it could only happen if $p = 2$, in which case L could only be of the form $\gamma\,y_2$); as $Q - \gamma\,Q_0/\beta$ satisfies the hypothesis, one may assume that $\gamma = 0$, and one finds that $R(y_2,\ldots,y_p) \geq 0$ for every $(y_2,\ldots,y_p)$ with $y_2 \neq 0$, and therefore for every $(y_2,\ldots,y_p)$.

As far as I know, no one has found in the general case a simple characterization of the set of quadratic forms which are nonnegative on Λ, or at least the extreme rays of this convex cone; as I will explain later, one must introduce a list of entropies before doing that.

It is worth looking at simpler problems in order to explain what the ideas of the Compensated Compactness Method are. In N-dimensional Elasticity (with $N = 2$ or 3 for applications), one must deal with a strain tensor $F = \nabla\,u$ (in the evolution case, the derivatives with respect to t, i.e. the velocities, must be added to the list), and the PIOLA-KIRCHHOFF stress tensor σ (I think that I had described clearly my approach in the late 70s, and the fact that the proponents of a stressless Elasticity theory have received more attention is not directly my fault, although my pointing out that it is nonsense to deal with Elasticity in terms of F alone may have actually boosted some unscientific support for the other camp). The constitutive relations relate the stress σ to F; in the hyperelastic case one often prefers to deal with an energy functional. The list of differential information consists in the compatibility conditions for gradients, and NEWTON's law of Mechanics, the equilibrium equation in the stationary case. The characteristic set $\mathcal{V}$ is the set of F, σ, ξ, with $\xi \in R^N \setminus 0$, $F = a \otimes \xi$ for some $a \in R^N$ and $\sigma\,\xi = 0$; the characteristic set Λ is then the set of F, σ, such that the rank of F is less or equal than 1, σ is singular, and $\sigma\,F^T = 0$; the list of quantities which are sequentially weakly continuous include the subdeterminants extracted from F, but also all the components of $\sigma\,F^T$, and if $N = 2$ the determinant of σ. In the hyperelastic case, i.e. if there exists a real function W such that $\sigma_{ij} = \frac{\partial W}{\partial F_{ij}}$ for all $i,j = 1,\ldots,N$, one can derive new equations for smooth solutions, and these go under the general term of "entropies", choosen by Peter LAX (geometers call them Casimirs): in the stationary case, without exterior forces, one has

$$\sum_j \frac{\partial(\sigma_{ij}F_{ik})}{\partial x_j} = F_{ik}\sum_j \frac{\partial \sigma_{ij}}{\partial x_j} + \sum_j \frac{\partial W}{\partial F_{ij}}\frac{\partial^2 u_i}{\partial x_j \partial x_k} = \frac{\partial\Big(W(F)\Big)}{\partial x_k} \text{ for all } i,k.$$

One can reiterate the application of the quadratic theorem after introducing as new components for U the quantities which appear in the new conserved quantities; I show the idea on the next example, dealing with solutions of the wave equation.

It is an important effect in fluids that waves can transport momentum and energy without transporting mass; one can learn something important on this question by looking at weakly converging sequences of solutions of linear wave equations (there is not so much proved at the moment for semi-linear or quasi-linear cases). For simplicity, let us consider a scalar wave equation with constant coefficients

$$\rho_0 \frac{\partial^2 u}{\partial t^2} - a\,\Delta\, u = f \text{ in } R^N \times (0,T);\; u(\cdot,0) = v \text{ in } R^N;\; \frac{\partial u}{\partial t}(\cdot,0) = w \text{ in } R^N,$$

which can arise in various ways, often with $f = 0$, in general after having linearized a complex nonlinear system near a trivial solution; for example u can be a vertical displacement, or a variation in pressure. Using methods of Functional Analysis, similar to the ones we used for the abstract framework for STOKES equation for example, one can show that if $v \in H^1(R^N)$, $w \in L^2(R^N)$, and $f \in L^1\big(0,T;L^2(R^N)\big)$, then there is a unique solution $u \in C^0\big([0,T];H^1(R^N)\big) \bigcap C^1\big([0,T];L^2(R^N)\big)$, and a very important property of the preceding equation, apart from describing an isotropic medium where information travels at velocity $\sqrt{a/\rho_0}$, is the balance of energy (conservation if $f = 0$)

$$\frac{\partial}{\partial t}\Big(\frac{\rho_0}{2}\Big|\frac{\partial u}{\partial t}\Big|^2 + \frac{a}{2}\sum_j\Big|\frac{\partial u}{\partial x_j}\Big|^2\Big) - a\sum_j \frac{\partial}{\partial x_j}\Big(\frac{\partial u}{\partial x_j}\frac{\partial u}{\partial t}\Big) = f\frac{\partial u}{\partial t} \text{ in } R^N \times (0,T),$$

and the density of energy $\frac{\rho_0}{2}\big|\frac{\partial u}{\partial t}\big|^2 + \frac{a}{2}\sum_j\big|\frac{\partial u}{\partial x_j}\big|^2$ is the sum of the density of kinetic energy and of the density of potential energy (related to the elastic properties or to the compressibility properties of the medium, according to the interpretation given to u); if $f = 0$, the integral of the density of energy is constant, equal to the total energy of the entire medium, and therefore equal to its value at time 0, $\int_{R^N}\big(\frac{\rho_0}{2}|w|^2 + \frac{a}{2}|grad(v)|^2\big)\,dx$.

If one considers a sequence v_n converging to v_∞ in $H^1(R^N)$ weak, and w_n converging to w_∞ in $L^2(R^N)$ weak, the limit of $\int_{R^N}\big(\frac{\rho_0}{2}|w_n|^2 + \frac{a}{2}|grad(v_n)|^2\big)\,dx$ will be strictly greater than $\int_{R^N}\big(\frac{\rho_0}{2}|w_\infty|^2 + \frac{a}{2}|grad(v_\infty)|^2\big)\,dx$, in the case where the convergences are not strong convergences, and that means than part of the initial energy is put in high frequencies. It is a common fact that if one observes the ocean from a plane, the surface of the ocean does looks flat, and that is similar to the case where $v_\infty = w_\infty = 0$, but there is some energy (and momentum) moving around at the surface. Although the solution u_n will converge weakly to the solution u_∞ corresponding to the initial data v_∞, w_∞, part of the energy will be missing in the description using only u_∞, and it is important to know where this missing energy has gone, transported around in high frequencies.

The approach of Compensated Compactness only gives an interesting but incomplete information about this energy traveling around in high frequencies, the Equipartition of Energy: it is the "action" $\frac{\rho_0}{2}\big|\frac{\partial u}{\partial t}\big|^2 - \frac{a}{2}\sum_j\big|\frac{\partial u}{\partial x_j}\big|^2$ which is sequentially weakly continuous. This result does not tell what the density of energy transported by the high frequencies is, but it tells that half of it is in kinetic form and half of it is in potential form. The proof is just an application of the Div-Curl lemma, where E^n in the full gradient of u_n, i.e. in all the variables (x,t), and D^n is $\big(-a\frac{\partial u_n}{\partial x_1},\ldots,-a\frac{\partial u_n}{\partial x_N},\rho_0\frac{\partial u_n}{\partial t}\big)$, whose divergence is f.

One can describe where the energy goes with a mathematical tool which I introduced in the late 80s, H-measures (I introduced H-measures for a different purpose, and then I proved a general propagation theorem for H-measures; the same objects were also introduced independently by Patrick GÉRARD for still another purpose; taking into account the problem of initial data for the wave equation was done by Gilles FRANCFORT and François MURAT, with the technical help of Patrick GÉRARD).

If the gradient of u_n in all the variables (x,t) stays bounded in L^3_{loc}, one can use the Div-Curl lemma with D^n being replaced by $\big(-a\frac{\partial u_n}{\partial x_1}\frac{\partial u_n}{\partial t},\ldots,-a\frac{\partial u_n}{\partial x_N}\frac{\partial u_n}{\partial t},\frac{\rho_0}{2}\big|\frac{\partial u_n}{\partial t}\big|^2 + \frac{a}{2}\sum_j\big|\frac{\partial u_n}{\partial x_j}\big|^2\big)$, and in that case $(E^n.D^n) = \frac{\partial u_n}{\partial t}\big(\frac{\rho_0}{2}\big|\frac{\partial u_n}{\partial t}\big|^2 - \frac{a}{2}\sum_j\big|\frac{\partial u_n}{\partial x_j}\big|^2\big)$, i.e. $\frac{\partial u_n}{\partial t}\,action_n$; one deduces that if $u_\infty = 0$, then not only $action_n$ converges weakly to 0, but also $\frac{\partial u_n}{\partial t}\,action_n$ converges weakly to 0 (and actually $\frac{\partial u_n}{\partial x_j}\,action_n$ also converges weakly to 0 for every j, as a consequence of other balance relations). However, one should not conclude that one has found a new sequentially weakly continuous function, because if $u_\infty \neq 0$, then E^∞ is the full gradient of u_∞, but D^∞ is not necessarily the corresponding quantity associated to u_∞, because the components of D^n are quadratic quantities which are not from the list of sequentially weakly continuous functions (which only contains the action); actually the functions $\frac{\partial u}{\partial t}\,action$ (as well as $\frac{\partial u}{\partial x_j}\,action$) are such that their weak limits can be computed from the full gradient of u_∞ and the H-measure associated with the full gradient of $u_n - u_\infty$.

I have almost described the idea of the Compensated Compactness Method: one looks for "entropies", consequence of the differential equations and eventually also of the nonlinear constitutive relations, and one applies the quadratic theorem of Compensated Compactness to that extended system; an obvious difficulty is that one may find an enormous system, and one may have to make a choice of which "entropies" to use, but there is another difficulty that can be described on the example of BURGERS equation.

One considers a weakly converging sequence u_n of solutions of the equation $\frac{\partial u_n}{\partial t} + u_n \frac{\partial u_n}{\partial x} = 0$, written in conservative form $\frac{\partial u_n}{\partial t} + \frac{1}{2}\frac{\partial (u_n)^2}{\partial x} = 0$. If $(u_n)^k$ converges in L^∞_{loc} weak $\star$ to U_k for $k \geq 1$, one deduces that $\frac{\partial U_1}{\partial t} + \frac{1}{2}\frac{\partial U_2}{\partial x} = 0$, but the Compensated Compactness theory does not help, apart from $U_2 \geq U_1^2$ which one deduces from a convexity argument (by taking the weak $\star$ limit of $(u_n - U_1)^2 \geq 0$). If one assumes now that u_n is smooth enough so that one can multiply the equation by u_n and write the result in conservation form $\frac{1}{2}\frac{\partial (u_n)^2}{\partial t} + \frac{1}{3}\frac{\partial (u_n)^3}{\partial x} = 0$, then one can use the Div-Curl lemma and one deduces that $u_n \frac{(u_n)^3}{3} - \frac{(u_n)^2}{2}\frac{(u_n)^2}{2}$ converges weakly to $U_1 \frac{U_3}{3} - \frac{U_2}{2}\frac{U_2}{2}$, which must then be equal to $\frac{U_4}{12}$, and therefore one has $U_4 = 4U_1U_3 - 3U_2^2$. Because $(u_n - U_1)^4 \geq 0$ gives at the limit $U_4 - 4U_1U_3 + 6U_1^2U_2 - 3U_1^4 \geq 0$, one deduces $-3(U_2 - U_1^2)^2 \geq 0$, and therefore $U_2 = U_1^2$, which implies that $u_n \to U_1$ in L^2_{loc} strong (and therefore in L^p_{loc} strong for every $p < \infty$).

One sees that the use of "entropies" together with the Compensated Compactness theorem has created some kind of ellipticity for an enlarged system, forbidding then oscillations, and that it is the main idea of the Compensated Compactness Method; however, one must complement it with a technical remark, because the sequence u_n is not usually smooth enough: if one starts with oscillations in the initial data, the derivative in x at time 0 must be large and negative somewhere and shocks will therefore appear after a very short time. Actually, as a consequence of an explicit formula valid for $\frac{\partial u}{\partial t} + \frac{\partial \big(f(u)\big)}{\partial x} = 0$ with f convex (used by Peter LAX, but Sergei GODUNOV also claims priority), or as a consequence of my argument (valid for more general f), these shocks will interact and decay rapidly, all this in such a short time that for every $\tau > 0$ the sequences converges strongly in $t \geq \tau$. For doing that, one must restrict attention to physically realistic sequences of approximating solutions, those which satisfy an "entropy condition" like $\frac{1}{2}\frac{\partial (u_n)^2}{\partial t} + \frac{1}{3}\frac{\partial (u_n)^3}{\partial x} \leq 0$ in the sense of measures (as noticed by Eberhard HOPF in the scalar case, and extended to systems of conservation laws by Peter LAX), and although measures are not necessarily in H^{-1}_{loc}, here the ones which appear do stay in a compact of H^{-1}_{loc} strong, as a consequence of a theorem that François MURAT had proved for another purpose: if a sequence is bounded in $W^{-1,p}_{loc}(\Omega)$ and stays bounded in the space of measures, then it stays in a compact of $W^{-1,q}_{loc}(\Omega)$ strong if $1 \leq q < p$.

Of course, it remains to apply these ideas to realistic questions of fluids. I had asked my former student Luisa MASCARENHAS to work on the case where one considers the gradient of a divergence free velocity field, in the case where the potential energy is a function of $F + F^T$, and she had obtained some results in this direction in the early 80s; I had in mind to improve some earlier results by Olga LADIZHENSKAYA and by Shmuel KANIEL, but Dan JOSEPH had then mentioned to me a defect of this kind of model, and I therefore did not pursue in that direction. I was not aware at the time of models coming from Oceanography, and there is obviously some potential applications of the methods which I had developped in the late 70s for these models, but instead of describing models I have made the choice of describing first the mathematical tools available for analyzing these models, and it remains to describe quickly what H-measures are, as it would be quite unrealistic to ignore the question of propagation of momentum or energy by waves (one should not either postulate too much about these waves, as probabilist sometimes do).

21-820. PDE Models in Oceanography

Luc TARTAR, WEAN Hall 6212, 268-5734, tartar@andrew.cmu.edu

42. Friday April 30.

In the late 70s, I was trying to improve YOUNG measures by adding a direction variable ξ in order to prove propagation results (because Lars HÖRMANDER had proved results of "propagation of singularities" where bicharacteristic rays played a role); I had mentioned the question to George PAPANICOLAOU and he had mentioned to me the WIGNER transform, but I could not find a way to use it (later Pierre-Louis LIONS and T. PAUL used it in order to give a different definition of the semi-classical measures introduced by Patrick GÉRARD, but these objects use a characteristic length and my vision of the physical world being one where there are plenty of different scales, this was not what I was looking for). I tried to use the limit of functions $F\big(x, u_n, \frac{grad(u_n)}{|grad(u_n)|}\big)$, but if u_n is not smooth and one approaches it by smooth functions, the limit might depend upon the approximating sequence, and I did not pursue that idea. I thought of using Homogenization for various elliptic systems with coefficients being general functions of u_n, but I could not find a direct way of using that idea, until I understood that a simpler problem was to mix materials with similar properties (what I later called Small Amplitude Homogenization). The first hint occured in 1984, when I discovered that in the problem modeled on stationary STOKES with a force field in $u \times curl(v_0 + \lambda v_n)$, the correcting matrix M had a factor λ^2, and that I could almost compute the correction from the behaviour at infinity of quadratic quantities in the FOURIER transform. The second hint occured in 1986, when I heard a talk of Stephen COWIN on bone evolution, and it seemed to me that the tensors that he was computing using methods from stereometry should instead appear as second or fourth derivatives of functions defined in a similar way than the ones that David BERGMAN had been using in an isotropic setting (again Small Amplitude Homogenization was behind). The third hint occured in the Fall of 1986, when I tried to check what LANDAU and LIFSHITZ had written about the conductivity of mixtures, which I had been aware of twelve years before (and at that time I had dismissed their computations as nonsense); I realized that their formula was just one of the bounds of Zvi HASHIN and S. SHTRIKMAN (for which I had given the first mathematical proof in 1980 by using a method based on Compensated Compactness that I had introduced earlier and which is now often called the translation method; actually there is something like H-measures hidden in the formal argument of HASHIN and SHTRIKMAN showing that the bounds must hold, while their argument that these bounds are attained required little change and using that part of their argument was clear to me), but as they proposed to apply it to mixtures where the variations in conductivities are small, I checked their result with Gilles FRANCFORT and François MURAT (again using an idea from Compensated Compactness), and their formula appeared to be accurate; it was quite a miracle if one considers the lack of any logical inference in their derivation, and I understood then what a framework for Small Amplitude Homogenization should be, and from the previous hints I could now guess easily how I was going to prove a mathematical version of their result, using these H-measures which I immediately knew how to use, before I had a clear idea of how I was going to define them correctly.

For a scalar sequence u_n converging weakly to 0 in $L^2_{loc}(\Omega)$, and $\varphi \in C_c(\Omega)$, the FOURIER transform of $\varphi\, u_n$ is bounded in $C_0(R^N)$ and converges pointwise to 0, and therefore $\mathcal{F}(\varphi\, u_n)$ converges to 0 in $L^2_{loc}(R^N)$ strong by LEBESGUE dominated convergence theorem; from the hints, I knew that I needed to describe how the information contained in $|\mathcal{F}(\varphi\, u_n)|^2$ was going away to infinity, and therefore for a continuous function ψ defined on the sphere S^{N-1}, I looked at the limit

$$L(\varphi, \psi) = \lim_{n\to\infty} \int_{R^N} |\mathcal{F}(\varphi\, u_n)(\xi)|^2 \psi\Big(\frac{\xi}{|\xi|}\Big)\, d\xi.$$

Of course, using separability arguments and a diagonal procedure, one can extract a subsequence such that the preceding limit exists for every $\varphi \in C_c(\Omega)$ and every $\psi \in C(S^{N-1})$, and for φ given there exists a nonnegative RADON measure μ_φ on S^{N-1} such that the limit is $\langle \mu_\varphi, \psi\rangle$, but although the dependence with respect to φ is not straightforward, I was sure from what I knew about the local character of Homogenization and the hints, that the limit was given by a nonnegative RADON measure μ on $\Omega \times S^{N-1}$,

$$L(\varphi, \psi) = \langle \mu, |\varphi|^2 \otimes \psi\rangle, \text{ written formally as } \int_{\Omega\times S^{N-1}} |\varphi(x)|^2 \psi(\xi)\, d\mu(x, \xi).$$

Of course, the hints had also told me that these measures in (x,ξ) could handle partial differential equations, and in order to compare with the Compensated Compactness theorem, there was an obvious generalization for the case of a sequence U^n converging weakly to 0 in $L^2_{loc}(\Omega;R^p)$; for two indices j and k, I found more natural to take two different test functions in x, $\varphi_1,\varphi_2\in C_c(\Omega)$, and to consider the limit

$$L_{jk}(\varphi_1,\varphi_2,\psi)=\lim_{n\to\infty}\int_{R^N}\mathcal{F}(\varphi_1\,U^n_j)(\xi)\overline{\mathcal{F}(\varphi_2\,U^n_k)(\xi)}\psi\Big(\frac{\xi}{|\xi|}\Big)\,d\xi,$$

which I expected to be given by a RADON measure μ_{jk} on $\Omega\times S^{N-1}$,

$$L_{jk}(\varphi_1,\varphi_2,\psi)=\langle\mu_{jk},\varphi_1\overline{\varphi_2}\otimes\psi\rangle,\text{ written formally as }\int_{\Omega\times S^{N-1}}\varphi_1(x)\overline{\varphi_2(x)}\psi(\xi)\,d\mu_{jk}(x,\xi).$$

As L_{jk} is linear in φ_1, antilinear in φ_2, and linear in ψ, it was reasonable to think of a RADON measure ν_{jk} on $\Omega\times\Omega\times S^{N-1}$ with $L_{jk}(\varphi_1,\varphi_2,\psi)=\int_{\Omega\times\Omega\times S^{N-1}}\varphi_1(x)\overline{\varphi_2(y)}\psi(\xi)\,d\nu_{jk}(x,y,\xi)$, and therefore it was important to show that the support of ν_{jk} was in $\{(x,y,\xi):x=y\}$. However, a general linear continuous form from $C_c(\Omega)$ (with the sup norm) into $\mathcal{M}(S^{N-1})$, the space of RADON measures on S^{N-1}, is given by an operator with a distribution kernel, according to the kernel theorem of Laurent SCHWARTZ, and I expected this kernel to be a measure by a positivity argument: if $\varphi_2=\varphi_1$ and $\psi\geq 0$, then $L_{jk}(\varphi_1,\varphi_2,\psi)\geq 0$. Jacques-Louis LIONS had told me that he had obtained with Lars GÅRDING a simple proof of SCHWARTZ's kernel theorem, which I could then avoid, and the crucial step was to show that $L_{jk}(\varphi_1,\varphi_2,\psi)$ only depended on $\varphi_1\overline{\varphi_2}$ and ψ. This was obtained by a commutation lemma, and some kind of pseudo-differential calculus. I do not like to use the classical theory of pseudo-differential operators introduced by Joseph KOHN and Louis NIRENBERG, or the theory of FOURIER integral operators introduced by Lars HÖRMANDER, because these theories require smooth coefficients, and when one is interested in Continuum Mechanics or Physics, one must avoid any unnecessary hypothesis of smoothness for coefficients, and therefore I developped the theory which I needed.

I assume that all functions are extended by 0 outside Ω, so that one works on R^N. To any $b\in L^\infty(R^N)$ one associates the operator M_b of multiplication by b, i.e. $(M_bv)(x)=b(x)v(x)$ a.e. $x\in R^N$; M_b is linear continous from $L^2(R^N)$ into itself, and its norm is the $L^\infty(R^N)$ norm of b. To any $a\in L^\infty(R^N)$ one associates the operator P_a defined by $\mathcal{F}P_a=M_a\mathcal{F}$, i.e. $\mathcal{F}P_av(\xi)=a(\xi)\mathcal{F}v(\xi)$ a.e. $\xi\in R^N$; P_a is linear continous from $L^2(R^N)$ into itself, and its norm is the $L^\infty(R^N)$ norm of a. In the quantity which I had considered, I only used functions ψ defined on the sphere S^{N-1} and extended to $R^N\setminus 0$ as homogeneous functions of order 0, and I wanted then the limit of $\int_{R^N}\mathcal{F}(P_\psi M_{\varphi_1}U^n_j)\overline{\mathcal{F}M_{\varphi_2}U^n_k}\,d\xi$, which by PLANCHEREL formula is $\int_{R^N}P_\psi M_{\varphi_1}U^n_j\overline{M_{\varphi_2}U^n_k}\,d\xi$; as the limit of $\int_{R^N}M_{\varphi_1}P_\psi U^n_j\overline{M_{\varphi_2}U^n_k}\,d\xi$ obviously depends only upon $\varphi_1\overline{\varphi_2}$ and ψ, it remained to show that $P_\psi M_{\varphi_1}U^n_j-M_{\varphi_1}P_\psi U^n_j$ converges strongly to 0 in $L^2(R^N)$, and as U^n_j converges weakly to 0 in $L^2(R^N)$, this would be a consequence of the commutator $P_\psi M_{\varphi_1}-M_{\varphi_1}P_\psi$ being a compact operator from $L^2(R^N)$ into itself. It was not too difficult to prove that for $a\in C(S^{N-1})$ and $b\in C_0(R^N)$ the commutator $[P_a,M_b]=P_aM_b-M_bP_a$ is indeed a compact operator, using the fact that HILBERT-SCHMIDT operators (which have a kernel in $L^2(R^N\times R^N)$) are compact, and that uniform limits of compact operators are compact; as a consequence of a commutation lemma of COIFMAN, ROCHBERG and WEISS, this result is actually true for $b\in VMO(R^N)$, and one can therefore extend my theory to use functions in $L^\infty\bigcap VMO$. The "pseudo-differential" operators of order 0 which I use have symbols of the form $s(x,\xi)=\sum_k a_k(\xi)b_k(x)$ with $a_k\in C(S^{N-1})$ and $b_k\in C_0(R^N)$ for all k and $\sum_k||a_k||\,||b_k||<\infty$, where the norms are sup norms; I define the standard operator S of symbol s by $S=\sum_k P_{a_k}M_{b_k}$, which corresponds to $\mathcal{F}Sv(\xi)=\int_{R^N}s(x,\frac{\xi}{|\xi|})u(x)e^{-2i\pi(x.\xi)}\,dx$ a.e. $\xi\in R^N$ when $v\in L^2(R^N)\bigcap L^1(R^N)$, and I say that a linear continuous operator L from $L^2(R^N)$ into itself has symbol s if $L-S$ is compact; this is the case for the operator $L_0=\sum_k M_{b_k}P_{a_k}$, which corresponds to $L_0v(x)=\int_{R^N}s(x,\frac{\xi}{|\xi|})\mathcal{F}u(\xi)e^{+2i\pi(x.\xi)}\,d\xi$, a.e. $x\in R^N$ when $v\in L^2(R^N)\bigcap\mathcal{F}L^1(R^N)$, which specialists of linear partial differential equations prefer; in my framework it is more natural to apply first an operator of multiplication in x in order to have a function defined on all R^N, so that one can apply FOURIER transform, but it is for the second commutation lemma that the choice of S is more crucial.

The H-measure μ associated to the chosen subsequence of U^n is a $p \times p$ matrix whose entries are (complex) RADON measures on $\Omega \times S^{N-1}$, and μ is Hermitian nonnegative. By taking $\psi = 1$, one sees that the integral of μ_{jk} in ξ gives the limit of $U_j^n \overline{U_k^n}$, i.e. if $U_j^n \overline{U_k^n} \rightharpoonup \pi_{jk}$ in the sense of measures, then for every $\varphi \in C_c(\Omega)$ one has $\langle \pi_{jk}, \varphi \rangle = \langle \mu_{jk}, \varphi \otimes 1 \rangle$. With this calculus modulo compact operators at hand, one can improve the Compensated Compactness theorem by the Localization Principle: if the functions A_{jk} are continuous and if U^n satisfies $\sum_{jk} \frac{\partial (A_{jk} U_k^n)}{\partial x_j} \to 0$ in $H^{-1}_{loc}(\Omega)$ strong, then one has $\sum_{jk} \xi_j A_{jk} \mu_{kl} = 0$ for all l; the converse is actually true: if $\sum_{jk} \xi_j A_{jk} \mu_{kl} = 0$ for all l, then $\sum_{jk} \frac{\partial (A_{jk} U_k^n)}{\partial x_j} \to 0$ in $H^{-1}_{loc}(\Omega)$ strong (if R_j is the RIESZ operator which has symbol $i\xi_j/|\xi|$, then the condition is equivalent to $\sum_{jk} R_j A_{jk} U_k^n \to 0$ in $L^2_{loc}(\Omega)$ strong). One sees then that what H-measures do is to compute the limits of sequences $v_n \overline{w_n}$, where v_n and w_n are obtained from U^n by applying "pseudo-differential" operators of the class introduced; of course they do not make YOUNG measures obsolete as H-measures cannot see the limits of nonquadratic quantities, and therefore the Compensated Compactness theorem has been improved (as one can consider equations with continuous coefficients if they are written in conservation form), but not so much the Compensated Compactness method, which is only strengthened by the addition of the theory of H-measures.

The question of Small Amplitude Homogenization consists for example in looking at elliptic problems of the form $div\big(A^n grad(u_n)\big) = f$, where $A^n = A^\infty + \gamma B^n$ and B^n converges weakly to 0, in which case the effective coefficient A^{eff} is analytic in γ (as was first noticed by Sergio SPAGNOLO for the symmetric case), and $A^{eff} = A^\infty + \gamma^2 C + O(\gamma^3)$. The H-measure of B^n permits to compute the coefficient of γ^2 in the expansion; the reason is that if one takes DIRICHLET conditions for a slighly larger open set for example, for $\lambda \in R^N$ one can choose f (depending on γ) so that $grad(u_n) = \lambda + \gamma\, grad(v_n) + \gamma^2\, grad(w_n) + O(\gamma^3)$ in Ω, then $grad(v_n)$ converges weakly to 0 and $div(A^\infty grad(v_n) + B^n \lambda) = 0$, and $grad(w_n)$ converges weakly to 0 and $div\big(A^\infty grad(w_n) + B^n grad(v_n)\big) = 0$ and $C\,\lambda$ is the weak limit of $B^n grad(v_n)$; if we were on R^N, the map $B^n \mapsto grad(u_n)$ would be given by a "pseudo-differential" operator, and the limit of $B^n grad(v_n)$ could be computed using the H-measure associated to B^n, but another way to prove the result is to use the Localization Principle.

A similar method enters the computation of the correction M in the problem modeled on stationary STOKES with a force in $u \times curl(v_0 + \lambda\, v_n)$.

Although the Small Amplitude Homogenization is important in many instances, a crucial step is to realize that H-measures can describe transport properties (of oscillations / concentration effects, which are the usual words involved when one looks at the difference between strong and weak convergence). I first considered a first order differential operator

$$\sum_{j=1}^N b_j \frac{\partial u_n}{\partial x_j} = f_n \text{ in } R^N,$$

with $u_n \rightharpoonup 0$ weakly in $L^2(R^N)$ and $f_n \to 0$ strongly in $H^{-1}_{loc}(R^N)$, and $b_j \in C_0^1(R^N)$ for $j = 1, \ldots, N$; if the sequence is associated with the H-measure μ, then the Localization Principle asserts that $P(x,\xi)\mu = 0$ with $P(x,\xi) = \sum_{j=1}^N b_j(x)\xi_j$ (notice that if the b_j are complex there may well be no points in the zero set of P). I assume that the coefficients b_j are real, so that multiplying the equation by $\overline{u_n}$ and taking the real part, one has

$$\sum_{j=1}^N b_j \frac{\partial |u_n|^2}{\partial x_j} = 2\Re(f_n \overline{u_n}) \text{ in } R^N.$$

One assumes then that $f_n \rightharpoonup 0$ weakly in $L^2(R^N)$, and for a real $a \in C^1(S^{N-1})$, one applies the operator P_a to the equation and one obtains $P_a f_n = \sum_j P_a \partial_j (M_{b_j} u_n) - \sum_j P_a (\partial_j M_{b_j}) u_n = \sum_j \partial_j \big((P_a M_{b_j} - M_{b_j} P_a) u_n\big) + \sum_j M_{b_j} \partial_j P_a u_n + \sum_j \big((\partial_j M_{b_j}) P_a - P_a (\partial_j M_{b_j})\big) u_n$. One needs then a second commutation lemma, which requires a little more smoothness either on all the b_j or on a (and one uses a commutation result of Alberto CALDERÓN in that case): $P_a M_{b_j} - M_{b_j} P_a$ maps $L^2(R^N)$ into $H^1(R^N)$ and $\partial_j (P_a M_{b_j} - M_{b_j} P_a)$ has symbol $\xi_j \sum_k \frac{\partial a}{\partial \xi_k} \frac{\partial b_j}{\partial x_k} = \xi_j \{a, b_j\} = \{a, \xi_j b_j\}$, where the POISSON bracket $\{g, h\}$ of two functions on $R^N \times S^{N-1}$ is

$\sum_k \frac{\partial g}{\partial \xi_k}\frac{\partial h}{\partial x_k} - \frac{\partial g}{\partial x_k}\frac{\partial h}{\partial \xi_k}$. Therefore one has

$$\sum_{j=1}^{N} b_j \frac{\partial (P_a u_n)}{\partial x_j} + K\, u_n = P_a f_n \text{ in } R^N, \text{ and the symbol of } K \text{ is } \{a, P(x,\xi)\},$$

and one deduces that

$$\sum_{j=1}^{N} b_j \frac{\partial \Big((P_a u_n)\overline{u_n}\Big)}{\partial x_j} + (K\, u_n)\overline{u_n} = 2\Re(P_a f_n)\overline{u_n} \text{ in } R^N.$$

One assumes then that $U^n = (u_n, f_n)$ corresponds to a H-measure ν, so that $\nu_{11} = \mu$, and one applies the last equation to a test function $\varphi \in C^1_c(R^N)$, and one gets $-\langle a\mu, \sum_j \partial_j(b_j\varphi)\rangle + \langle \{a,P\}\mu, \varphi\rangle = 2\langle \Re(a\nu_{12}), \varphi\rangle$, so that if one defines $\Phi(x,\xi) = a(\xi)\varphi(x)$, one has

$$\langle \mu, \{\Phi, P\} - div(b)\Phi\rangle = 2\langle \Re\nu_{12}, \Phi\rangle,$$

which extends by linearity and density to all $\Phi \in C^1_c(R^N \times S^{N-1})$, and this is a first order differential equation (i.e. a transport equation) for μ, written in weak formulation.

The method applies to linear differential systems endowed with a sesquilinear balance relation for their complex solutions (even if u_n is real, $P_a u_n$ takes complex values in general); in principle it applies also to semilinear equations, but what the source term ν_{12} is in these cases is not clear.

I failed to find the way to use H-measures for proving theorems of compactness by averaging, but it is precisely for this purpose that Patrick GÉRARD introduced independently the same objects (actually he introduced them for functions taking values in HILBERT spaces); he called these objects microlocal defect measures, and if I agree with the qualificative microlocal, I do not like the qualificative defect : the transport theorem for the wave equation shows that H-measures give the way to describe what a beam of Light is for example (and it is polarized Light if one uses MAXWELL equation) and the important physical quantities carried along it, nothing that looks like a defect.

H-measures use no characteristic lengths, and if they are useful it is for phenomena where the frequency is not so important as long as it is high. The transport equation for the wave equation says that in the limit of high frequency one obtains Geometrical Optics, but refraction effects or grazing rays in the Geometric Theory of Diffraction of Joe KELLER are frequency dependent, and one needs other objects. My idea for taking care of one characteristic length was to add a new coordinate and consider H-measures in R^{N+1}, but Patrick GÉRARD had another idea, quite related, and he introduced the semi-classical measures on $\Omega \times R^N$ by looking for a sequence ε_n tending to 0 at the quantities

$$\lim_{n\to\infty} \int_{R^N} |\mathcal{F}(\varphi\, u_n)(\xi)|^2 \psi(\varepsilon_n \xi)\, d\xi.$$

There are various technical improvements of this idea, and known deficiencies of this or other variants, but it is not well understood yet how to handle situations with many length scales; obviously this is of importance for fluids in general, and for Oceanography in particular.

I have preferred to sketch the existing mathematical tools before looking at more precise mathematical models in Oceanography, but time was a little short for doing that. I hope that nevertheless these lecture notes will stimulate many to investigate more on all the questions which I have only sketched.