Advanced Mathematics for Engineers and Scientists/Scale Analysis

Contents

Scale Analysis

In the chapter on nondimensionalization, variables (both independent and dependent) were nondimensionalized and at the same time scaled so that they ranged from something like 0 to 1. "Something like 0 to 1" is the mentality.

Scale analysis is a tool that uses nondimensionalization to:

- Understand what's important in an equation and, more importantly, what's not.
- Gain insight into the size of unknown variables, such as velocity or temperature, before (even without) actually solving.
- Simplify solution process (nondimensional variables ranging for 0 to 1 are very amiable).
- Reduce dependence of the solution on physical parameters.
- Allow higher quality numeric solution since variables that are of the same range maintain accuracy better on a computer.

Scale analysis is very common sense driven and not too systematic. For this reason, and since it is somewhat unnatural and hard to describe without endless example, it may be difficult to learn.

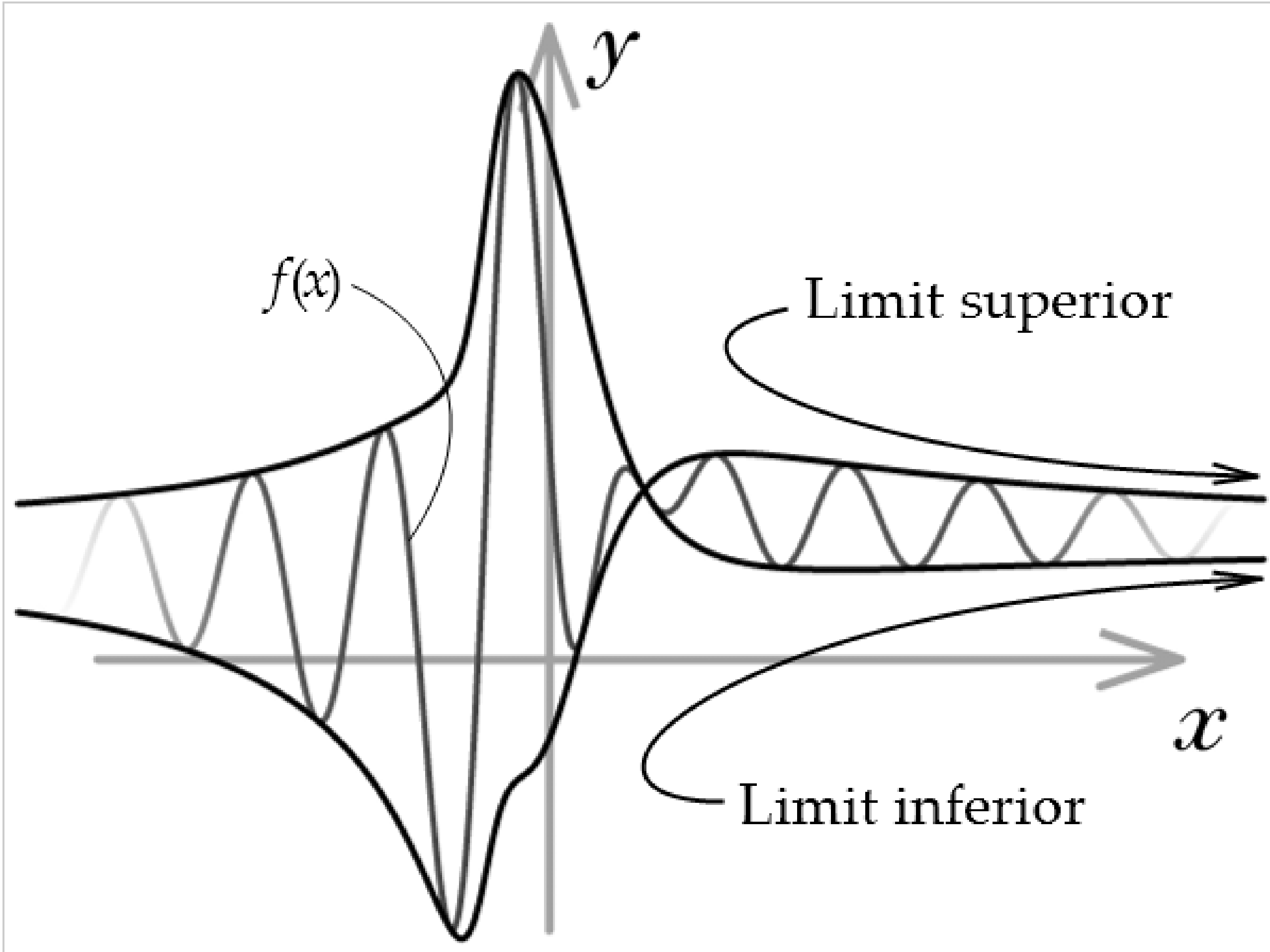
Before going into the concept, we must discuss orders of magnitude.

Orders of Magnitude and Big O Notation

Suppose that there are two functions $f(x)$ and $g(x)$. It is said (and notated) that:

$$f(x) \quad \text{is} \quad \mathcal{O}(g(x)) \quad \text{as} \quad x \to a \qquad \text{if} \qquad \limsup_{x \to a} \left| \frac{f(x)}{g(x)} \right| < \infty$$

It's worth understanding fully this possibly obscure definition. $\limsup$, short for limit superior, is similar to the "regular" limit, only it is the limit of the upper bound. This concept, alongside limit inferior, is illustrated at right. This intuitive analysis will have to suffice here, as the precise definition and details of these special limits are rather complicated.



Visualization of the limit superior and limit inferior of f(x) as x increases without bound.

As a further example, the limits of the cosine function as x increases without bound are:

$$\limsup_{x\to\infty} \cos(x) = 1$$

$$\liminf_{x\to\infty} \cos(x) = -1$$

$$\lim_{x\to\infty} \cos(x) \quad \text{does not exist}$$

With this somewhat off topic technicality hopefully understood, the statement that:

$$f(x) \quad \text{is} \quad \mathcal{O}(g(x)) \quad \text{as} \quad x \to a$$

Is saying that near $x = a$, the **o**rder (or size, or magnitude) of $f(x)$ is bounded by $g(x)$. It's saying that $|f(x)|$ isn't crazily bigger then $|g(x)|$ near $x = a$, and this is precisely notated by saying that the limit superior is bounded (the "regular" limit wouldn't work since oscillations would ruin everything). The notation involving the big O is rather surprisingly called "big O notation", it's also known as Landau notation.

Take, for example, $f(x) = 3x^2 - 100x + 2$ at different points:

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(x^2) \text{ as } x \to \infty$$

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(x^0) \text{ as } x \to 0$$

In the first case, the x^2 term will easily dominate for large x. Even if the coefficient on that term is very near zero, for large enough x that term will dominate. Hence, the function is of order x^2 for large x.

In the second case, near $x = 0$ the first two terms are limiting to zero while the constant term, 2, isn't changing at all. It is said to be of order 1, notated as order x^0 above. Why O(1) and not O(2)? Both are correct, but O(1) is preferred since it is simpler and more similar to x^0.

This may put forth an interesting question: what would happen if the constant term was dropped? Both of the remaining terms would limit to zero. Since we are looking at x *near* zero and not *at* zero,

$$3x^2 - 100x \text{ is } \mathcal{O}(x) \text{ as } x \to 0$$

This is because as x approaches zero, the quadratic term gets smaller much faster then the linear term. It would also be correct, though kind of useless, to call the quantity O(1). It would be incorrect to state that the quantity is of order zero since the limit would not exist, not under any circumstance.

As implied above, $g(x)$ is by no means a unique function. All of the following statements are true, simply because the limit superior is bounded:

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(\sinh(x)) \text{ as } x \to \infty$$

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(500 \cdot 2^{2^x} - \sin(x)) \text{ as } x \to \infty$$

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(x) \text{ as } x \to 0$$

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(x^2) \text{ as } x \to 0$$

While technically correct, these are very misleading statements. Normally, the simplest, smallest magnitude function g(x) is selected.

Before ending the monotony, it should also be mentioned that it's not necessary for $f(x)$ to be *smaller* then $g(x)$ near $x = a$, only the limit superior must exist. The following two statements are also true:

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(0.00001) \text{ as } x \to 0$$

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(1000000000) \text{ as } x \to 0$$

But again, these are misleading and it's most proper to state that:

$$3x^2 - 100x + 2 \text{ is } \mathcal{O}(1) \text{ as } x \to 0$$

A relatively simple concept has been beaten to death, to the point of being confusing. It'll be more clear in context, and it'll be used more in later chapters for different purposes.

Scale Analysis on a Two Term ODE

Previously, the following BVP was considered:

$$\frac{d^2u}{dy^2} = \frac{P_x}{\nu\rho}$$

$$u(0) = 0$$

$$u(D) = 0$$

Wipe away any memory of solving this simple problem, the concepts of this chapter do not look at the actual solution. The variables are nondimensionalized by defining new variables:

$$u = u_s\hat{u} \ \ , \ \ y = D\hat{y}$$

So that y **is scaled by** D, and u **is scaled by an unknown scale** u_s. Now note that, thanks to the scaling:

$$\hat{y} \ \text{ is } \ \mathcal{O}(1) \text{ , and } \ \hat{u} \ \text{ is } \ \mathcal{O}(1)$$

These are both true near zero. u will be O(1) (this is read "of order one") when its scale is properly chosen. Using the chain rule, the ODE was turned into the following:

$$\frac{u_s}{D^2}\frac{d^2\hat{u}}{d\hat{y}^2} = \frac{P_x}{\nu\rho}$$

Now, if both u and y are of order one, then it is reasonable to *assume* that, at least at some point in the domain of interest:

$$\frac{d^2\hat{u}}{d\hat{y}^2} \ \text{ is } \ \mathcal{O}(1)$$

This is by no means guaranteed to be true, however it is reasonable.

To identify the velocity scale, we can set the derivative **equal** to one and solve. There is nothing "illegal" about purposely setting the derivative equal to one since all we need is *some equation* to specify an unknown constant, u_s. There is much freedom in defining this scale, because *what this constant is and how it's found has no effect on the validity of the solution of the BVP* (as long as it's not something stupid like 0).

$$\frac{u_s}{D^2} \cdot 1 = \frac{P_x}{\nu\rho} \Rightarrow u_s = \frac{D^2 P_x}{\nu\rho}$$

Since:

$$\hat{u} \text{ is } \mathcal{O}(1)$$

It follows that:

$$u \text{ is } \mathcal{O}(u_s) \Rightarrow u \text{ is } \mathcal{O}\left(\frac{D^2 P_x}{\nu\rho}\right)$$

This velocity scale may be thought of as a characteristic velocity. It's a number that shows us what to expect the velocity to be like. The velocity could actually be larger or smaller, but this gives a general idea. Furthermore, this scale tells us how chaging various physical parameters will affect the velocity; *there are four of them summarized into one constant*.

Compare this result to the coefficient (underlined) on the complete solution, with u dimensional and y nondimensional:

$$u(\hat{y}) = \underline{\frac{D^2 P_x}{2\nu\rho}}(\hat{y}^2 - \hat{y})$$

They differ by a factor of 2, but they are of the same order of magnitude. So, indeed, u_s **characterizes** the velocity.

Words like "reasonable" and "assume" were used a few times, words that would normally lead to the uglier word "approximate". Relax: the BVP itself hasn't been approximated or otherwise violated in any way. We just used scale analysis to *pick* a velocity scale that:

- Turned the ODE into something very easy to look at: $\frac{d^2\hat{u}}{d\hat{y}^2} = 1 \quad ; \quad u(0) = u(1) = 0$
- Gained good insight into what kind of velocity the solution will produce without finding the actual solution.

Note that a zero pressure gradient can no longer show itself in the ODE. This is by no means a restriction, since a zero pressure gradient would result in a zero velocity scale which would unconditionally result in zero velocity.

Scale Analysis on a Three Term PDE

The last section was still more of nondimensionalization then it was scale analysis. To just begin getting deeper into the subject, we'll consider the pressure driven transient parallel plate IBVP, identical to the above only with a time component:

$$\frac{\partial u}{\partial t} = \nu\frac{\partial^2 u}{\partial y^2} - \frac{P_x}{\rho}$$

$$u(0,t) = 0$$

$$u(D,t) = 0$$

$$u(x,0) = 0$$

See the change of variables chapter to recall the origins of this problem. Scales are defined as follows:

$$u = u_s\hat{u} \ \ , \ \ y = D\hat{y} \ \ , \ \ t = t_s\hat{t}$$

Again, the scale on y is picked to make it an order one quantity (based on the BCs), and the scales on u and t are just letters representing unknown quantities.

The chain rule has been used to define derivatives in terms of the new variables. Instead of taking this path, recall that, given variables x and y (for the sake of example) and their respective scales x_s and y_s:

$$\frac{\partial^n y}{\partial x^n} = \frac{\partial^n \hat{y}}{\partial \hat{x}^n}\frac{y_s}{x_s^n} \qquad ; \quad x = x_s\hat{x} \ , \ y = y_s\hat{y}$$

So that makes things much easier. Performing the change of variables:

$$\frac{\partial \hat{u}}{\partial \hat{t}}\frac{u_s}{t_s} = \nu\frac{\partial^2 \hat{u}}{\partial \hat{y}^2}\frac{u_s}{D^2} - \frac{P_x}{\rho}$$

In the previous section, there was one unknown scale and one equation, so the unknown scale could be easily and *uniquely* isolated. Now, there are two unknown scales but only one equation (no, the BCs/IC will not help). What to do?

The physical meaning of scales may be taken into consideration. Ask: "What should the scales represent?"

There is no unique answer, but **good** answers for this problem are:

- u_s characterizes the steady state velocity.
- t_s characterizes the response time: the time to establish steady state.

Once again, these are **picked** (however, for this problem there really aren't any other choices). In order to determine the scales, the physics of each situation is considered. There may not be unique choices, but there are best choices, and these are the "correct" choices. An understanding of what each term in the PDE represents is vital to identifying these "correct" choices, and this is notated below:

$$\underbrace{\frac{\partial \hat{u}}{\partial \hat{t}}\frac{u_s}{t_s}}_{\text{acceleration}} = \underbrace{\nu\frac{\partial^2 \hat{u}}{\partial \hat{y}^2}\frac{u_s}{D^2}}_{\text{viscosity}} - \underbrace{\frac{P_x}{\rho}}_{\substack{\text{driving} \\ \text{force}}}$$

For the velocity scale, a steady state condition is required. In that case, the time derivate (acceleration) must *small*. We could obtain the characteristic velocity associated with a steady state condition by *requiring* that the acceleration be something small (read: zero), stating that the second derivative is $O(1)$, and solving:

$$0 = \frac{u_s}{D^2}\cdot 1 - \frac{P_x}{\nu\rho} \Rightarrow u_s = \frac{D^2 P_x}{\nu\rho}$$

This is the same as the velocity scale found in the previous section. This is expected since both situations are describing the same steady state condition. The neglect of acceleration equates to what's called a **balance between driving force and viscosity** since driving force and viscosity are all that remain.

Getting the time scale may be a little more elusive. The time associated with achieving steady state is dictated by the acceleration and the viscosity, so it follows that the time scale may be obtained by considering **a balance between acceleration and viscosity**. Note that this statement has nothing to do with pressure, so it should apply to a variety of disturbances. To balance the terms, pretend that the derivatives are $O(1)$ quantities and disregard the pressure:

$$1 \cdot \frac{u_s}{t_s} = \nu \cdot 1 \cdot \frac{u_s}{D^2} + 0 \Rightarrow t_s = \frac{D^2}{\nu}$$

This is a statement that:

- The smaller the viscosity, the longer you wait for steady state to be achieved.
- The smaller the separation distance, the less you wait for steady state to be achieved.

Hence, the scale describes what will affect the transient time and how. The results may seem counterintuitive, but they are verified by experiment if the pressure is truly a constant capable of combating possibly huge viscosity forces for a high viscosity fluid.

Compare these scales to constants seen in the full, dimensional solution:

$$u(y,t) = \frac{D^2 P_x}{2\rho\nu}\left(\frac{y^2}{D^2} - \frac{y}{D} - \sum_{n=1}^{\infty} e^{-\frac{\nu(n\pi)^2}{D^2}\cdot t} \cdot \frac{4(-1)^n - 4}{n^3\pi^3}\sin(\frac{n\pi}{D}y)\right)$$

The velocity scales match in order of magnitude, nothing new there. But examine the time constant (extracted from the exponential factor) and compare to the time scale:

$$\text{time constant} = \frac{D^2}{\nu(n\pi)^2} \quad ; \quad t_s = \frac{D^2}{\nu}$$

They are of the same order with respect to the physical parameters, though they'll differ by nearly a factor of 10 when $n = 1$. This result is more useful then it looks. Note that after determining the velocity scale, all three terms of the equation may have been considered to isolate a time scale. This would've been a poor choice that wouldn't have agreed with the time constant above since it wouldn't be describing the required settling between viscosity and acceleration.

Suppose that, for some problem, a time dependent PDE is too hard to solve, but the steady state version is easier and it is what you're interested in. A natural question would be: "How long do I wait until steady state is achieved?"

The time scale provided by a proper scale analysis will at least give an idea. In this case, assuming that the first term of the sum in the solution is dominant, the time scale will overestimate the response time by nearly a factor of 10, which is priceless information if you're otherwise clueless. This overestimate is actually a good (safe) overestimate, it's always better to wait longer and be certain of the steady state condition. Scales in general have a tendency to overestimate.

Before closing this section, consider the actual nondimensionalization of the PDE. During the scale analysis, the coefficients of the last two terms were equated and later the coefficients of the first two terms were equated. This implies that the nondimensionalized PDE will be:

$$\frac{\partial \hat{u}}{\partial \hat{t}} = \frac{\partial^2 \hat{u}}{\partial \hat{y}^2} - 1$$

And this may be verified by substituting the expressions found for the scales into the PDE. This dimensionless PDE, too, turned out to be completely independent of the physical parameters involved, which is very convenient.

Heat Flow Across a Thin Wall

Now, an important utility of scale analysis will be introduced: determining what's important in an equation and, better yet, what's not.

As mentioned in the introduction to the Laplacian, steady state heat flow in a homogeneous solid may be described by, in three dimensions:

$$\nabla^2 = 0 \qquad \xrightarrow{\text{It's 3D.}} \qquad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

Now, suppose we're interested in the heat transfer inside a large, relatively **thin** wall, with differing temperatures (not necessarily uniform) on different sides of the wall. The word 'thin' is crucial, write it down on your palm right now. You should suspect that if the wall is indeed thin, the analysis could be simplified somehow, and that's what we'll do.

Not caring about what happens at the edges of the wall, a BVP may be written:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

$$u(x, y, 0) = f(x, y)$$

$$u(x, y, D) = g(x, y)$$

D is the thickness of the wall (implication: z is the coordinate across the wall). Suppose that the wall is a boxy object with dimensions $s \times s \times D$. Using the box dimensions as scales,

$$x = s\hat{x}\ , \quad y = s\hat{y}\ , \quad z = D\hat{z}\ , \quad u = u_s\hat{u}$$

Only the scale of u is unknown. Substituting into the PDE,

$$\frac{\partial^2 \hat{u}}{\partial \hat{x}^2}\frac{u_s}{s^2} + \frac{\partial^2 \hat{u}}{\partial \hat{y}^2}\frac{u_s}{s^2} + \frac{\partial^2 \hat{u}}{\partial \hat{z}^2}\frac{u_s}{D^2} = 0$$

$$\left(\frac{\partial^2 \hat{u}}{\partial \hat{x}^2} + \frac{\partial^2 \hat{u}}{\partial \hat{y}^2}\right)\left(\frac{D}{s}\right)^2 + \frac{\partial^2 \hat{u}}{\partial \hat{z}^2} = 0$$

Note that the scale on u divided out — so a logical choice must be made for it's scale; in this case it'd be an extreme boundary value (ie, the maximum value of $\max(f, g)$), let's say it's chosen and taken care of. Thanks to this scaling and the rearrangement that followed, we may get a good idea of the *magnitude* of each term in the equation:

$$\left(\underbrace{\frac{\partial^2 \hat{u}}{\partial \hat{x}^2}}_{\mathcal{O}(1)} + \underbrace{\frac{\partial^2 \hat{u}}{\partial \hat{y}^2}}_{\mathcal{O}(1)}\right) \cdot \underbrace{\left(\frac{D}{s}\right)^2}_{?} + \underbrace{\frac{\partial^2 \hat{u}}{\partial \hat{z}^2}}_{\mathcal{O}(1)} = 0$$

Each derivative is approximately $O(1)$. But what about the squared ratio of dimensions? This is called a *dimensionless parameter*. Look at your palm now (the one you don't write with), recall the word "thin". "Thin" in this case means exactly the same thing as:

$$\frac{D}{s} \ll 1$$

And if the ratio above is much smaller then 1, then the square of this ratio is even smaller. Our dimensionless parameter is called a *small parameter*. When a parameter is small, there are many opportunities to simplify analysis; the simplest would be to state that it's too small to matter, so that:

$$\left(\underbrace{\frac{\partial^2 \hat{u}}{\partial \hat{x}^2}}_{\mathcal{O}(1)} + \underbrace{\frac{\partial^2 \hat{u}}{\partial \hat{y}^2}}_{\mathcal{O}(1)}\right) \cdot \underbrace{\left(\frac{D}{s}\right)^2}_{\substack{\text{really} \\ \text{small}}} + \underbrace{\frac{\partial^2 \hat{u}}{\partial \hat{z}^2}}_{\mathcal{O}(1)} = 0 \quad \Rightarrow \quad \frac{\partial^2 \hat{u}}{\partial \hat{z}^2} = 0$$

What was just done couldn't have been justified without scaling variables so that their derivatives are (likely) $O(1)$, since you have no idea what order they are otherwise. We know that each derivative is hopefully $O(1)$, but some of these $O(1)$ derivatives carry a very small factor. Only then can terms be righteously dropped. The dimensionless BVP becomes:

$$\frac{\partial^2 \hat{u}}{\partial \hat{z}^2} = 0$$

$$\hat{u}(s\hat{x}, s\hat{y}, 0) = \hat{f}(s\hat{x}, s\hat{y})$$

$$\hat{u}(s\hat{x}, s\hat{y}, 1) = \hat{g}(s\hat{x}, s\hat{y})$$

Note that it's still a partial differential equation (the x and y varialbes haven't been made irrelevant – look at the BCs). Also note that scaling on u is undone since it cancels out anyway (the scale could've still been picked as, say, a maximum boundary value). This problem may be solved very simply by integrating the PDE twice with respect to z, and then considering the BCs:

$$\int \frac{\partial^2 \hat{u}}{\partial \hat{z}^2} \, d\hat{z} = 0$$

$$\int \frac{\partial \hat{u}}{\partial \hat{z}}\, d\hat{z} = C_1(s\hat{x}, s\hat{y})$$

$$\hat{u}(x, y, z) = zC_1(s\hat{x}, s\hat{y}) + C_2(s\hat{x}, s\hat{y})$$

C_1 and C_2 are integration "constants". The first BC yields:

$$\hat{u}(x, y, 0) = \hat{f}(s\hat{x}, s\hat{y})$$

$$0 \cdot C_1(s\hat{x}, s\hat{y}) + C_2(s\hat{x}, s\hat{y}) = \hat{f}(s\hat{x}, s\hat{y}) \quad \Rightarrow C_2(s\hat{x}, s\hat{y}) = \hat{f}(s\hat{x}, s\hat{y})$$

And the second:

$$\hat{u}(x, y, 1) = \hat{g}(s\hat{x}, s\hat{y})$$

$$1 \cdot C_1(s\hat{x}, s\hat{y}) + \hat{f}(s\hat{x}, s\hat{y}) = \hat{g}(s\hat{x}, s\hat{y}) \quad \Rightarrow C_1(s\hat{x}, s\hat{y}) = \hat{g}(s\hat{x}, s\hat{y}) - \hat{f}(s\hat{x}, s\hat{y})$$

The solution is:

$$\hat{u}(x, y, z) = z \cdot (\hat{g}(s\hat{x}, s\hat{y}) - \hat{f}(s\hat{x}, s\hat{y})) + \hat{f}(s\hat{x}, s\hat{y})$$

It's just saying that the temperature varies linearly from one wall face to the other. It's worth noting that in practice, once scaling is complete, the hats on variables are "dropped" for neatness and to prevent carpal tunnel syndrome.

Words of Caution

"Extreme caution" is more fitting.

In the wall heat transfer problem, we took the partial derivatives in x and y to be $O(1)$, and this was justified by the scaling: $\hat{x}$, $\hat{y}$ and $\hat{u}$ are $O(1)$, so the derivatives must be so as well. Right?

Not necessarily. That they're $O(1)$ is a *linear approximation*, however if the function $u(x, y, z)$ is significantly nonlinear with respect to a variable of interest, then the derivatives may not be as $O(1)$ as thought. In this problem, one way that this can happen is if the temperature at each wall face (the functions $f(x, y)$ and $g(x, y)$) have large and differing Laplacians. This **will** result in three dimensional heat conduction.

Examine carefully the image at right. Suppose that side length is ten times the wall thickness; $f(x, y)$ and $g(x, y)$ have zero Laplacians everywhere except along circles where temperatures suddenly change. At these locations, the Laplacian can be huge (unbounded if the sudden changes are discontinuities). This will suggest that the derivatives in question are *not* O(1) but much greater, so that these terms become important even though in this case:

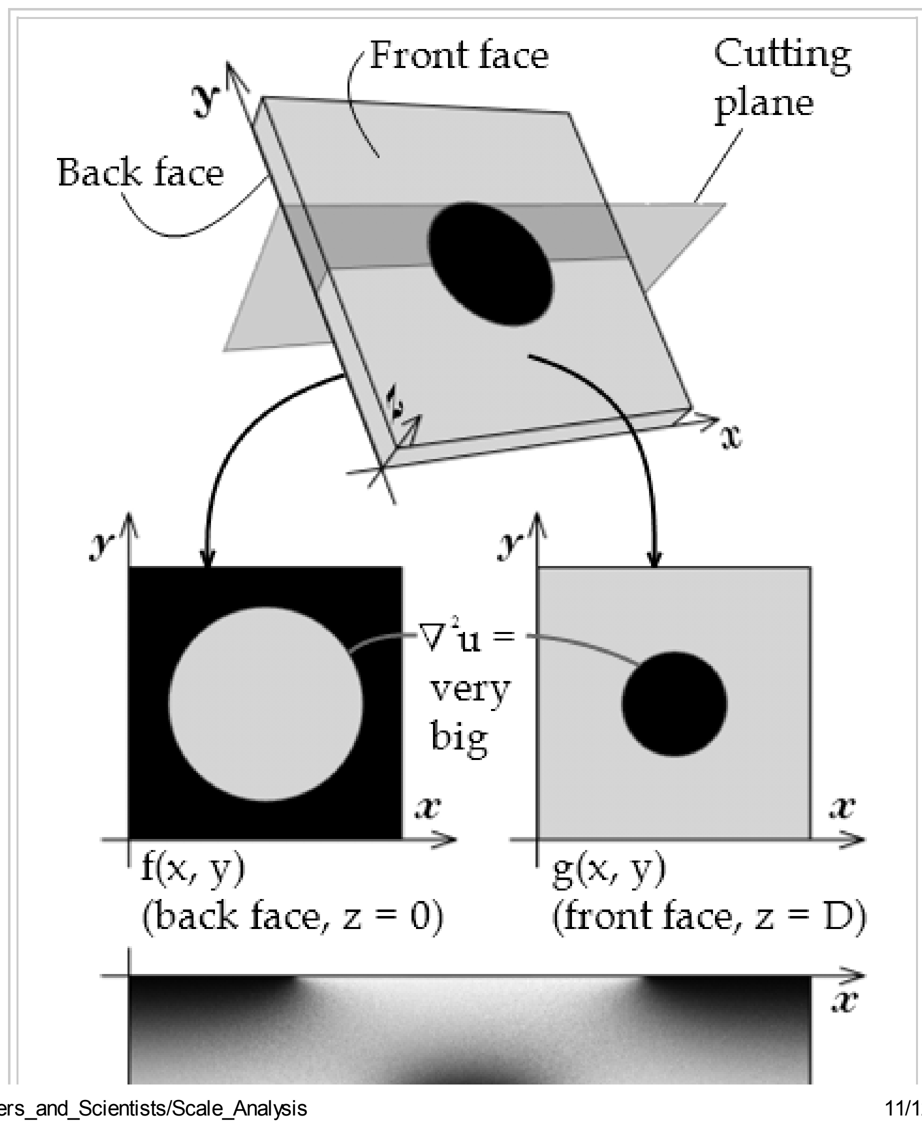
$$\frac{D}{s} = 0.1 \Rightarrow \left(\frac{D}{s}\right)^2 = 0.01 \ll 1$$

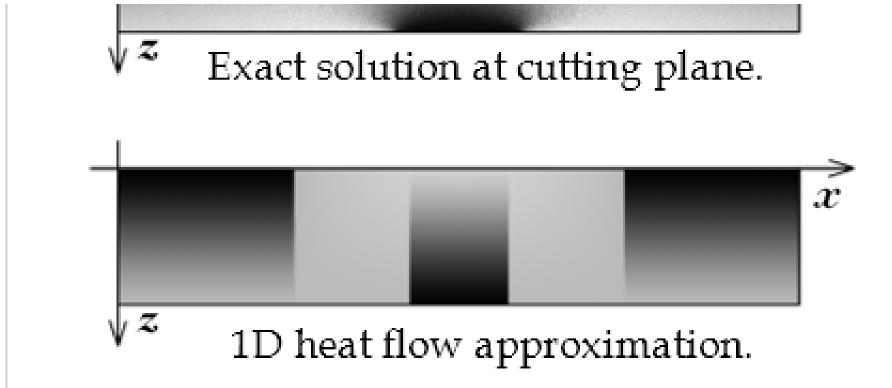
Which is as required by the scale analysis: the wall is clearly thin. But apparently, the small thinness ratio multiplied by the large derivatives leads to significant quantities.

Both the exact solution and the solution to the problem approximated through scaling are shown at the location of a cutting plane. The exact solution shows at least two dimensional heat transfer, while the solution of the simplified solution shows only one dimensional heat transfer and is substantially different.

It's easy to see why the 1D approximation fails even without knowing what a Laplacian is: this is a heat transfer problem involving the diffusion of temperature, and the temperature will clearly need to diffuse along x near the sudden changes within the wall (can't say the same about the BCs since they're fixed).

The caption of the figure starts with the word "failure". Is it really a failure? That depends on what you're looking for, it may or may not be. Note that if the wall were even thinner and the sudden jumps not discontinuities, the exact and 1D solutions could again eventually become indistinguishable.





Failure of one dimensional flow approximation.