Structured Backward Error for the WKB method

Robert M. Corless

University of Geneva, 2025-04-15

Western University, Canada

Slides available at `rcorless.github.io`; please download them

Joint work with Nic Fillion

Announcing Maple Transactions

a "Diamond" class open access journal with no page charges

Now listed by DBLP

mapletransactions.org

An exemplary paper:

Some Instructive Mathematical Errors by Richard P. Brent,

(we might remark on this paper later)

Perturbation Methods using Backward Error

A book, by Nic Fillion and myself, to be published by SIAM.

(break to look at the table of contents)

There must be "a few books" already on Perturbation Methods

Figure 1: RMC and Don Quixote, in Alcalá de Henares, 2017

Why, on Earth, write another?

Fools rush in where angels fear to tread.
—Alexander Pope, *An essay on criticism*, written 1709

- There are only two other books that use backward error [4, 5]
- We claim backward error is *very* useful for perturbation methods*
- We think computer algebra is still under-utilized nowadays, although there are some works that use it systematically
- Even though scientific computing has progressed *far* beyond perturbation methods, there is still a need for them.

* This fact may seem obvious in retrospect. We contend that the obstacle of hand labour has discouraged full use of backward error in practice till now. We will see its advantages!

The goal: short lucid formulae

Numerical solution and graphs (and animations) are truly valuable, but sometimes a short lucid formula can tell you just as much as an hour with a simulator and visualization tools can.

This *depends* on the scientist (or student!) understanding the terms in the formula, of course!

Backward error is not purely mathematical

Although backward analysis is a perfectly straightforward concept there is strong evidence that **a training in classical mathematics leaves one unprepared to adopt it**. . . . *I have even detected a note of moral disapproval in the attitude of many to its use and there is a tendency to seek a forward error analysis even when a backward error analysis has been spectacularly successful.*

—J. H. Wilkinson, in [**Wilkinson1985**]

What is “Backward Error?”

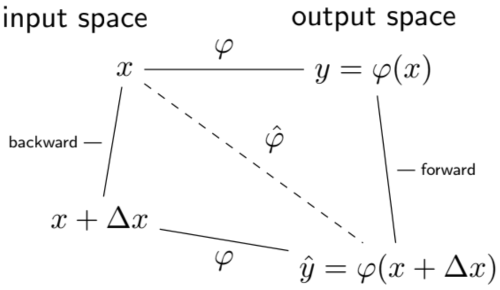

Figure 2: We want to compute $\varphi(x)$ but we cannot, for some reason. We *can* compute $\widehat{y} = \widehat{\varphi}(x)$. This has forward error $y - \widehat{y}$. But perhaps $\widehat{y} = \varphi(x + \Delta x)$ exactly; Δx is a “backward error” (this need not be unique). Or perhaps $\widehat{y} = (\varphi + \Delta\varphi)(x)$; then $\Delta\varphi$ is another kind of backward error.

That's not mathematics

Changes in the input data x, to $x + \Delta x$, are usual in science (engineering, economics, psychology, anything). Changes in the mathematical model φ are also usual: one normally neglects terms and effects that are considered to be "small" or "unimportant."

If we can put our errors-in-solution in the same context as these kinds of data or modelling errors, then we can *reuse* the tools that we have to use for such (e.g. the "sensitivity" or "conditioning" of the problem).

Aging spring, Lengthened pendulum

Some physical problems have natural "secular" (slowly-varying) terms in them. For instance, consider the "aging spring" [1]:

$$\ddot{y} + e^{-\varepsilon t} y = 0\,. \tag{1}$$

Cheng and Wu claimed to have used the "two-scale" method to get the solution $\exp(\varepsilon t/4)\sin(2(1-\exp(-\varepsilon t/2))/\varepsilon)$. The "WKB method" gets this solution directly. Its residual is

$$\frac{1}{16}\varepsilon^2 \mathrm{e}^{\varepsilon t/4} \sin\left(\frac{2(1-\mathrm{e}^{-\varepsilon t/2})}{\varepsilon}\right)\,. \tag{2}$$

Is that a "small" residual? It's a bit hard to tell.

Better backward error

But! Notice that the residual in equation (2) is just $\varepsilon^2 Y(t)/16$ where $Y(t)$ is the computed solution. This means that $Y(t)$ is the *exact* solution to

$$y'' + \left(e^{-\varepsilon t} - \frac{\varepsilon^2}{16}\right) y(t) = 0\,. \tag{3}$$

This is an equation that we can *directly* interpret in terms of the original model.

Notice that the spring constant becomes *zero* when $\exp(-\varepsilon t) = \varepsilon^2/16$, or $t = -2\ln(\varepsilon/4)/\varepsilon$. We thus learn that the approximate solution is likely not valid for t larger than this, *in a way that is consonant with the mathematical modelling*. [Cheng and Wu say that this equation is used in some kind of quantum application.]

The aging spring is sensitive to some changes

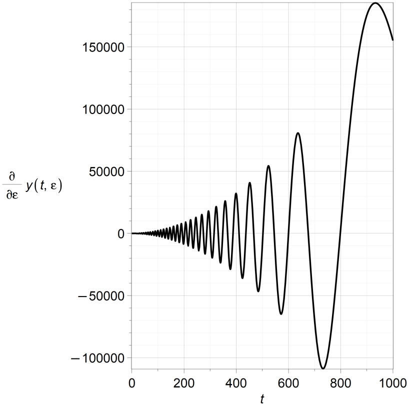

Figure 3: Taking the derivative with respect to ε shows that the solution is sensitive to changes in ε. $\varepsilon = 1/100$ here.

The context matters

Both of those details matter. Changing $e^{-\varepsilon t}$ to $e^{-\varepsilon t} - \varepsilon^2/16$ introduces a *spurious turning point* into the equation. This is likely "not physical" and demonstrates that for t large enough the Cheng–Wu solution will not be valid.

The fact that the solution varies strongly when tiny ε is changed by even a tinier amount is also a *kind* of ill-conditioning (but somehow it's "under control" in the model because we can see its consequences directly).

For both cases, to draw conclusions we need to know the physical context.

Perturbation vs Exact Solution

The analysis of the aging spring just performed—exhibiting an approximate solution that is the exact solution of a nearby problem of similar type, together with a residual and a condition number—tells us at least as much information as the exact solution in terms of Bessel functions would have.

We have identified an important issue, namely the sensitivity of the solution to changes in the problem, that will still be important for the exact (reference) solution.

WKB and backward error

The WKB (Wentzel–Kramers–Brillouin) method (or WKBJ method where the J is for Jeffreys, or LG method for Liouville–Green, or the "phase integral" method, even) gives the "solution of physical optics" of $\varepsilon^2 y'' = Q(x)y$ as

$$y_{WKB} = c_1 Q(x)^{-1/4} e^{S(x)/\varepsilon} + c_2 Q(x)^{-1/4} e^{-S(x)/\varepsilon} \tag{4}$$

where $S(x) = \int_{x_0}^{x} \sqrt{Q(\xi)}\, d\xi$. It's amazingly simple (once you get used to it); it's inspired by the integrating factor for $\varepsilon y' = P(x)y$ which is $I(x) = \int^{x} P(\xi)\, d\xi/\varepsilon$.

How good is the solution?

Backward error for WKB

y_{WKB} gives the *exact solution* to $\varepsilon^2 y'' = \widehat{Q}(x) y$ where

$$\widehat{Q}(x) = Q(x) + \varepsilon^2 \left(\frac{Q''}{4Q} - 5 \left(\frac{Q'}{4Q} \right)^2 \right) . \tag{5}$$

There is no further approximation there. That's a finite formula for the exact backward error $r(x) = \varepsilon^2 Q_2(x)$. The WKB method gives an exact solution to a nearby equation (provided $Q(x) \neq 0$—places where $Q(x) = 0$ are called *turning points*).

We have not seen this fact mentioned in any other textbook.

Green's functions and forward error

The forward error is then

$$\int_{x_0}^{x} G(x,\xi) r(\xi) y_{WKB}(\xi)\, d\xi \qquad (6)$$

where $G(x,\xi)$ is the Green's function. We can compute it (pretty easily) for the WKB solution; it is $O(1/\varepsilon)$ in size, so the forward error will be $O(\varepsilon)$ as $\varepsilon \to 0$.

The Green's function *also* (and more importantly) measures the sensitivity to changes in the equation or model, such as added noise.

“Small” vs “Small enough”

Since Backward Error Analysis requires the *context* of the original problem to be taken into account, this explicitly allows us to consider whether the computed residual (or other backward error form) is actually small compared to other neglected effects.

This is *not* mathematics! Mathematics abstracts, as far as possible, with the goal of making its results and predictions independent of context.

This is the crux of the matter.

Once this is settled, we can consider conditioning: are such small effects amplified to the point where we lose all predictability or control, or is the solution useful?

Morrison's counterexample

In [2, pp. 192–193], we find a discussion of the equation

$$y'' + y + \varepsilon(y')^3 + 3\varepsilon^2(y') = 0\,. \tag{7}$$

O'Malley's solution, there and in [3], is incorrect. We claim that had he computed a residual, he would have identified the blunder*.

* By "blunder" we mean arithmetic error, or algebra error, no more. It's just that the word "error" is a bit overused in this field already. Also, I feel some worry* in pointing out this blunder: O'Malley was a giant of perturbation methods. But we are certain that our solution is correct.

* inquietud, intranquilo, desasogado, ...

RG Solution

All we need do is to change the differential equation in our Jupyter notebook script, and alter the interrogations of the solution afterward. At $N = 2$, we get

$$\begin{aligned} z(t) =& 2R(t)\cos(t+\theta(t)) + \frac{\varepsilon R(t)^3 \sin(3t+3\theta(t))}{4} \\ &+ \varepsilon^2 \left(\frac{27R(t)^5\cos(3t+3\theta(t))}{32} - \frac{3R(t)^5\cos(5t+5\theta(t))}{32} \right) \end{aligned} \tag{8}$$

with

$$\dot{R}(t) = -\frac{3\varepsilon}{2} R(t) \left(R(t)^2 + \varepsilon \right) \tag{9}$$

and

$$\dot{\theta}(t) = \frac{9}{16} R^4(t) \varepsilon^2 \,. \tag{10}$$

With this, we get a uniformly small residual, which is small even compared to the decaying amplitude.

An important tangent: Published blunders

Off the top of my head, blunders in published perturbation computations have been exhibited by

- John P. Boyd (a hyperasymptotic expansion)
- Robert E. O'Malley (Morrison's counterexample)
- Émile Mathieu (in his 1868 paper which defined what are now called Mathieu functions)
- Bender & Orszag (a plain multiple scales computation, fixed in later editions)

at least. I claim that had they computed a final residual, they would have detected their blunders. Given that *all* of the above are/were experts, and therefore it's clear that the rest of us make blunders at least as frequently, I claim that residuals are even more necessary for us.

To be fair

Richard Brent was fair enough to include some of his own errors in the paper "Some instructive mathematical errors" I mentioned previously and so I should say explicitly that I make blunders, too. In my paper (with David Jeffrey and Donald Knuth) "A Sequence of Series for the Lambert W function " I claimed a certain series had infinite radius of convergence. Richard Crandall pointed out that I was wrong and the series had radius of convergence $\sqrt{2\pi}$.

So I am guilty, too!

Iterative WKB

If the WKB method gets the solution to the problem with $Q + \varepsilon^2 Q_2$, why not *try* to get the solution for $Q - \varepsilon^2 Q_2$? This should get the answer to the problem with $Q + O(\varepsilon^4)$!

This *works*. But the integrals get complicated. The error terms contain high order derivatives of Q.

If it works, do it again! This amounts to solving the following iteratively:

$$\widetilde{Q}(x) + \varepsilon^2 \left(5 \left(\frac{\widetilde{Q}'}{4\widetilde{Q}} \right)^2 - \frac{\widetilde{Q}''}{4\widetilde{Q}} \right) = Q(x) \tag{11}$$

Sir Michael Berry pointed out that this process does not converge in general, but that does not bother us: we stop when the residuals stop decreasing.

A problem for WKB

The "bottleneck" in the WKB method is the integral

$$S(x) = \frac{1}{\varepsilon}\int_{x_0}^{x} \sqrt{Q(\xi)}\, d\xi\,. \tag{12}$$

If this is complicated, then the answer is not as useful as it might be. For instance,

$$S_0 = \int_0^x \sqrt{1+\xi^8}\, d\xi = xF\left(\left.\begin{matrix} -1/2,\, 1/8 \\ 1+1/8 \end{matrix}\right| -x^8\right) \tag{13}$$

Here F is a hypergeometric function. It works well in Maple, though. But maybe you don't want one in your code. And there are many potentials for which no formula for the answer is known, at all.

Enter Chebfun (link) by Nick Trefethen and his group at Oxford, since 2004. This does for approximation of functions what floating-point does for arithmetic.

(show examples in Matlab)

Perturbation is just taking derivatives

The simplicity of a perturbation computation hides its importance. We are investigating what happens if a small part of the model changes.

This is itself a fundamental question of science. It's not surprising that the old techniques are still valuable; maybe it's a surprise just how valuable they can be.

That said, nowadays one can do a heck of a lot with a simulation window and a slider bar.

Thank you for listening.

This work supported by NSERC, and by the Spanish MICINN. I also thank CUNEF University for the opportunity to give this talk.

I am also happy to announce that SIAM has offered Nic and me a contract for this book, and we are to deliver it to them by December. Your feedback today will help to improve the book, and we will acknowledge you all.

Book text available at `https://github.com/rcorless/rcorless.github.io/blob/main/PerturbationBEABook.pdf`. Please download it and read it and send me (or Nic) your comments, by Nov 10 if possible.

Let's open the topic for discussion.

References

[1] Hung Cheng and Tai Tsun Wu. **"An aging spring".** In: *Studies in applied Mathematics* 49.2 (1970), pp. 183–185.

[2] Robert E. O'Malley. **Historical Developments in Singular Pertubations.** Springer, 2014.

[3] Robert E. O'Malley and Eleftherios Kirkinis. **"A Combined Renormalization Group-Multiple Scale Method for Singularly Perturbed Problems".** In: *Studies in Applied Mathematics* 124.4 (2010), pp. 383–410.

[4] Anthony John Roberts. **Model emergent dynamics in complex systems.** SIAM, 2014.

[5] Donald R. Smith. **Singular-perturbation Theory.** Cambridge University Press, 1985.

Using computation to illustrate formulae

Let's try to understand which is bigger, the term $\exp(-1/\varepsilon)$ or any algebraic term ε^j. L'Hopital's rule shows that as $\varepsilon \to 0^+$ the exponential is transcendentally smaller than any ε^j. But what happens if we ask when the two are equal?

$$e^{-1/\varepsilon} = \varepsilon^j \tag{14}$$

exactly when $\varepsilon_{-1} = e^{W_{-1}(-1/j)}$ (on the left) and when $\varepsilon_0 = e^{W_0(-1/j)}$. Here W_{-1} and W_0 are the two real branches of the Lambert W function. [Short, lucid formulae, just what we want*.]

So ε^j is *smaller* than $\exp(-1/\varepsilon)$ if $\varepsilon_{-1} < \varepsilon < \varepsilon_0$. Paradoxically, this is most of the interval, for large j!

* Heh. $\varepsilon_{-1} \sim 1/(j \ln j)$ and $\varepsilon_0 \sim 1 - 1/j$ might be easier to understand!

When is that?

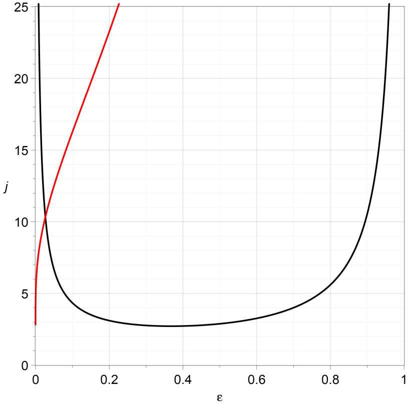

Figure 4: For values of j above this curve, $\varepsilon^j < \exp(-1/\varepsilon)$. That is, the "exponentially small" term is more important! Left of the red line is lost to rounding error in double precision. Note $\exp(-1/\varepsilon) = 2^{-54}$ already when $\varepsilon = \ln(2)/54 \approx 0.0267$.