An Alternative Approach of Evaluating the Validity of Perturbative Calculations with Respect to Convergence of Power Series based on Polynomials for the Inverse Function.

Andrej Liptaj*

Study of the Convergence of Perturbative Power Series

Abstract

Results of perturbative calculations in quantum physics have the form of truncated power series in a coupling constant. In order to evaluate the uncertainty of such results, the power series of the inverse function are constructed. These are inverted and the difference between the outcome of this procedure and the initial power series is taken as uncertainty.

Introductory note: This text was previously published on Scribd[1].

1 Introduction

Results in quantum field theories are often based on perturbative calculations and have the form of power series built in a coupling constant. The coefficients of these series come from evaluation of corresponding Feynman diagrams. This is in general a difficult task and therefore only few first terms of the series are usually available. One might trust the result if the coupling is small, one might trust the result less if the coupling is rather big. The latter situation unfortunately occurs in present quantum field theories, an example is the quantum chromodynamics (QCD). Here the coupling becomes large for processes with small momentum transfer. One can hardy apply standard mathematical criteria to rigorously evaluate the convergence of such series since these criteria require some general knowledge about the coefficient behavior (for all coefficients) which is not available. The standard criteria to evaluate the convergence if only a truncated result[2] $\sigma_N(\alpha) = \sum_{n=0}^{n=N} c_n \alpha^n$ is available are

*Institute of Physics, Bratislava, Slovak Academy of Sciences, andrej.liptaj@savba.sk
I am willing to publish any of my ideas presented through free-publishing services in a journal, if someone (an editor) judges them interesting enough. Journals in the "*Current Contents*" database are strongly preferred.

[1] *https://www.scribd.com/document/13699759/About-inverse-derivatives-and-convergence-of-perturbative-series*

[2] The Greek letter σ here stands for a general function.

- The size of the parameter α. If $\alpha \ll 1$ then one usually trusts the result.
- Ratios of consecutive terms $r_n = (c_n\alpha^n)/(c_{n-1}\alpha^{n-1})$. If for all n one has $|r_n| \ll 1$ then one usually trusts the result.
- The influence of the last known term $L_N = \left|(c_N\alpha^N)/\sigma_N\right|$. If $L_N \ll 1$ then one usually trusts the result. One often assigns an error of the order of $\left|c_N\alpha^N\right|$ to σ.

In this text I would like to present an additional, alternative criterion. The criterion exploits the idea of Taylor series, a coefficient c_n in series can be related to the n-th derivative of $\sigma(\alpha) = \sigma_\infty(\alpha)$ at $\alpha = 0$: $c_n = \frac{1}{n!}\frac{d^n\sigma}{d\alpha^n}|_{\alpha=0}$. This approach is justified because of the one-to-one correspondence, the evaluation of Feynman graphs can be (equivalently) regarded as the calculation of the derivatives of σ at $\alpha = 0$ (free theory).

2 Inverse Function Series

To avoid the confusion between higher derivatives and powers it is convenient to adopt the notation

$$fn = \frac{d^n}{dx^n}f(x)|_{x=x_0}, \quad gn = \frac{d^n}{dy^n}g(y)|_{y=y_0},$$

where f is a differentiable function, g is the inverse function of f: $g(f(x)) = f(g(x)) = x$ and $y_0 = f(x_0)$. Higher order derivatives of g can be expressed in terms of the derivatives of f by mean of the formula

$$g1 = \frac{1}{f1},$$

$$gn = \lim_{\Delta x \to 0} n! \frac{\Delta x - \Sigma_{i=1}^{n-1}\frac{1}{i!}(gi)\left[\Sigma_{j=1}^{n}\frac{1}{j!}(fj)(\Delta x)^j\right]^i}{\left[\Sigma_{i=1}^{n}\frac{1}{i!}(fi)(\Delta x)^i\right]^n}, \tag{1}$$

which is in a recursive and limit form. The explicit expressions for the first four derivatives of g are

$$g1 = \frac{1}{f1}, \quad g2 = -\frac{f2}{(f1)^3}, \quad g3 = \frac{3(f2)^2 - (f1)(f3)}{(f1)^5},$$

$$g4 = -\frac{15(f2)^3 - 10(f1)(f2)(f3) + (f1)^2(f4)}{(f1)^7}.$$

The Taylor expansion for the inverse function is then

$$g(y_1) = g(y_0) + (g1)(y_1 - y_0) + \frac{1}{2!}(g2)(y_1 - y_0)^2 + \ldots = x_0 + \sum_{i=1}^{\infty}\frac{1}{i!}(gi)(y_1 - y_0)^i.$$

One may consult [1] for more details.

3 An Alternative Evaluation of the Series Convergence

Let $\tau(\beta)$ be the inverse function of $\sigma(\alpha)$. With respect to the previous section one can establish the correspondence $\sigma \leftrightarrow f$, $\tau \leftrightarrow g$, $\alpha \leftrightarrow x$, $\beta \leftrightarrow y$, $\alpha_0 = 0 \leftrightarrow x_0$ and $\beta_0 = \sigma(0) \leftrightarrow y_0$. In addition let us adopt the following notations (still keeping the notation from the previous section)

$$\sigma_N(\alpha) = \sigma(0) + \sum_{i=1}^{N} \frac{1}{i!}(\sigma i)\alpha^i, \quad \tau_N(\beta) = \sum_{i=1}^{N} \frac{1}{i!}(\tau i)(\beta - \beta_0)^i$$

and

$$s_N(\alpha) = [\tau_N(\alpha)]^{-1},$$

the last expression meaning that s_N is the inverse function of τ_N. In general[3] $s_N \neq \sigma_N$, however one expects[4] $s_\infty = \sigma_\infty = \sigma$ because the inverse of the inverse function is the initial function. The comparison of s_N and σ_N is the main idea of this text. The first N derivatives of both functions at $\alpha = 0$ are equal (by construction). In the context where one relates the derivatives at $\alpha = 0$ with the Feynman diagrams, one may interpret both functions as corresponding to the same (truncated) series of Feynman diagrams. Both function s_N and σ_N are supposed to approximate the "true" function σ and the difference between them comes for higher orders. The functions σ_N have vanishing derivatives of the order greater then N at $\alpha = 0$ (by construction). Higher-then-N order derivatives of functions s_N do not need to vanish, since only derivatives of polynomials vanish at some order and s_N are not polynomials. Actually, one may think of many possibilities how to construct arbitrary functions t_N such that they have the same derivatives as σ_N at $\alpha = 0$ up to the order N . However in these cases one usually implicitly introduces some arbitrary behavior of higher-then-N order derivatives at $\alpha = 0$. I would like to stress that in the presented approach there is no assumption on the behavior of these higher derivatives, and if they come out not to be zero it is entirely due to the the procedure of inverting the Taylor polynomial for the inverse function, expecting that the result approximates σ. In this sens the procedure is "canonical".

To study the convergence of the series σ_N one can do the following

- Plot the graphs of s_N and σ_N and see for which $\alpha = \alpha_x$ the curves start to significantly split one from another. There is (obviously) no objective criterion to make this judgment, a suggestion would be to plot both functions in an orthonormal coordinate system[5] and judge by eye. If one in his physics calculations uses a value of $\alpha = \alpha_1$ and if $\alpha_x < \alpha_1$ then one

[3] σ_N is a polynomial whereas s_N is the inverse function of a polynomial, which is in general not a polynomial.

[4] On condition that the appropriate Taylor series converge.

[5] To avoid the judgment to be biased by an inappropriate coordinate system, for example logarithmic.

should be careful about the result. If $\alpha_1 < \alpha_x$ than one might want to trust the result.

- Calculate the difference $d_N = |s_N - \sigma_N|$ at $\alpha = \alpha_1$ which one uses in his physics calculations. If the difference is small compared to the values of the functions s_N and σ_N (for example $d_N/(0.5 \times s_N + 0.5 \times \sigma_N) \ll 1$) then one might trust the result. If not, one may be careful about the result.

- One might assign an error $\Delta\sigma = d_N$ to the calculated value: $\sigma(\alpha_1) = \sigma_N(\alpha_1) \pm d_N$.

It is not straightforward to invert the function τ_N in order to obtain s_N that one needs to plot the corresponding graph and to calculate $s_N(\alpha_1)$. One might use numerical methods for that purpose.

The presented approach offers one additional interesting option: It might happen that the series τ_N have better convergence then the series σ_N and thus one would prefer the value $s_N(\alpha_1)$ as the result rather then $\sigma_N(\alpha_1)$. It is possible to illustrate such a situation on basic mathematical functions. Let $f(x) = \ln(x)$, $g(y) = f^{-1}(y)$, $x_0 = 1$, $y_0 = f(x_0) = 0$ and let us evaluate $f(x)$ at $x = x_1 = 3$ using Taylor series. It is a mathematical fact that the Taylor series for $\ln(x)$ with the development at $x_0 = 1$ do not converge for $x = 3$ because the radius of convergence is 1. Thus the problem is well-posed (the value for ln(3) exists and is finite), but the method of solving it fails (series diverge). However if one here applies the “method of inverse function” one gets, thanks to the formula 1, the development of $f^{-1}(y)$ at $y_0 = 0$. It happens that $f^{-1}(y) \equiv \exp(y)$ and it is well-known that the Taylor polynomials for the exponential function converge to this function for each $y \in \Re$. Numerically inverting the (convergent) series for $g = f^{-1}(y)$ one obtains a prediction for $f(x) = \left(f^{-1}\right)^{-1}(x)$ and thus one can predict the value for $f(x_1) = \ln(3)$.

This example demonstrates that for certain functions the Taylor series of the function itself might diverge for a distant argument, however the series of the inverse function might converge everywhere. And thus, inverting one more time, one gets prediction for the initial function that converges on a bigger interval then the function’s Taylor series.

It may be however difficult to put this method into practice for truncated series because, unlike in the presented example, one has not the theoretical control over the function and its inverse function. The suggestion would be to study the series σ_N and τ_N using the criteria given in the *Introduction* and with respect to these criteria use either $\sigma_N(\alpha_1)$ or $s_N(\alpha_1)$ as the result.

I will try to demonstrate all mentioned techniques on the concrete example that follows.

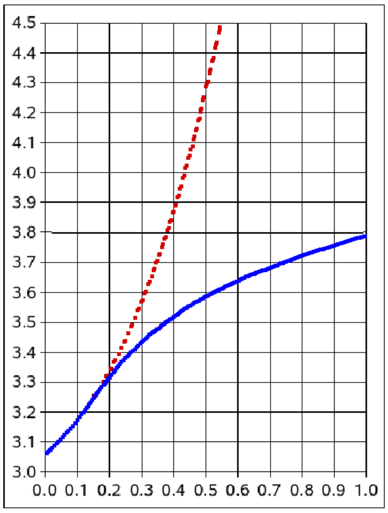

Figure 1: Functions σ_3 (dashed line) and s_3 (full line).

4 Example

The perturbative QCD contributions to the semi leptonic branching ratio of the tau lepton can be written [2]

$$R_\tau^{pQCD} = 3.058\left[1 + \frac{\alpha_S}{\pi} + 5.2\left(\frac{\alpha_S}{\pi}\right)^2 + 26.4\left(\frac{\alpha_S}{\pi}\right)^3 + \dots\right].$$

To avoid changing once more the notation let us use the notation from previous section with $\sigma(\alpha) \equiv R_\tau^{pQCD}(\alpha)$. Thus one has

$$\sigma_3(\alpha) = 3.058 + 0.973 \times \alpha + 1.611 \times \alpha^2 + 2.604 \times \alpha^3$$

and the corresponding derivatives at $\alpha = 0$ are $\sigma 1 = 0.973$, $\sigma 2 = 3.222$ and $\sigma 3 = 15.622$. Using the formula 1 it is possible to calculate the derivatives of the inverse function $\tau_N(\beta)$ at $\beta_0 = \sigma(\alpha = \alpha_0 = 0) = 3.058$. One gets $\tau 1 = 1.027$, $\tau 2 = -3.494$, $\tau 3 = 18.246$ and

$$\tau_3(\beta) = 1.027 \times (\beta - 3.058) - 1.747 \times (\beta - 3.058)^2 + 3.041 \times (\beta - 3.058)^3.$$

Finally one needs to numerically invert the function τ_3 to get s_3. The functions $\sigma_3(\alpha)$ and $s_3(\alpha)$ are depicted on the Figure 1 . The formula for R_τ^{pQCD} was in the ref. [2] used to extract the value of α_S from the measurement. Here I will consider the value of α_S to be fixed, $\alpha_S = \alpha_1 = 0.35$. One has $\sigma_3(0.35) = 3.708$ and $s_3(0.35) = \beta_1 = 3.479$. Following the suggestions given in the previous section one observes

- The graphs for $\sigma_3(\alpha)$ and $s_3(\alpha)$ start to split one from another around $\alpha_x = 2.5$ (judging by eye). One has $\alpha_x < \alpha_1$ and thus one may expect a non-negligible error on the result.

- Subtracting $\sigma_3(0.35)$ and $s_3(0.35)$ one gets $d_3 = 0.229$. One observes $d_3/(0.5 \times s_3 + 0.5 \times \sigma_3) = 0.064$, the error does not seem to exceed 10%.
- The result could be written $R_\tau^{pQCD}(\alpha_S = 0.35) = 3.708 \pm 0.229$.

One can finally study the convergence of σ_N and τ_N using the criteria from the *Introduction*.

- The size of the parameter in which the series are built: One has $\alpha_1 = 0.35 < 0.421 = (\beta_1 - 3.058)$, so the series σ_N seem to be better convergent then τ_N.
- Ratios of consecutive terms: One gets $|r_{\sigma 2}| = 0.579 < 0.716 = |r_{\tau 2}|$ and $|r_{\sigma 3}| = 0.566 < 0.733 = |r_{\tau 3}|$. This criterion also suggests that the series σ_N converge better then τ_N.
- The influence of the last term: The calculations lead to $L_{\sigma 3} = 0.030$ and $L_{\tau 3} = 0.649$. This criterion strongly favors[6] the result based on σ_N.

Using the techniques described in this text one would use σ_N to make prediction for R_τ^{pQCD} and one would claim the result to be $R_\tau^{pQCD}(\alpha_S = 0.35) = 3.708 \pm 0.229$. One can notice that the error is much smaller then what one gets using the "last known term" criterion. This might seem dangerous, one could however argue that the error of the order of the last known term is overestimated.

5 Summary and Conclusion

In this text I presented an alternative way of evaluating the convergence of the power series coming from perturbative calculations in quantum field theories. It might be regarded as an additional way of assigning the error on the perturbative result besides existing approaches. The idea is to construct Taylor series for the inverse function which behave, after inverting them, similar to the initial function at the point of the series development. They however differ when the argument gets distant from the development point. This difference is taken as the uncertainty. The nice feature of this method is that it does not require any further ad hoc assumptions on (unknown) higher order contributions and is in this sens canonical.

The presented approach also opens an interesting possibility to use for the prediction inverted Taylor polynomials of the inverse function instead of the original power series. I have shown on a simple mathematical function (logarithm) that at least from the mathematical point of view such a situation can occur. It is an opened question whether such situation can happen within the framework of perturbative calculations in physics and if yes how to recognize it.

[6] The series σ_N are however strongly "helped" by the absolute term $c_0 = 3.058$ which is missing in τ_N and which increases the denominator in the calculation of $L_{\sigma 3}$.

References

[1] Andrej Liptaj, Higher Order Derivatives of the Inverse Function, *https://www.scribd.com/document/113474210/Higher-order-derivatives-of-the-inverse-function-version-2,* *http://www.scribd.com/doc/13699758/Higher-order-derivatives-of-the-inverse-function* *http://vixra.org/abs/1703.0295*

[2] S. Eidelman et al., *Review of Particle Physis*, Phys. Lett. B592, 1 (2004).