Comment on "An alternative implementation of the 'principle of minimal sensitivity'"

P. M. Stevenson

CERN, CH-1211 Geneva 23, Switzerland

(Received 26 October 1982)

I explain that the difficulties encountered by Wrigley in the preceding paper arise only because he is attempting to find the stationary value of an undefined function.

I wish to comment on the preceding paper by Wrigley,¹ which itself is a comment on an earlier paper of mine.² The value of Wrigley's work, in my opinion, is that it outlines a procedure for finding an approximate solution to the "optimization equations"² which will be useful in third and higher orders. The advantage of this algorithm is that in n th order it reduces $(n-1)$ simultaneous transcendental equations to $(n-2)$ algebraic equations and just one transcendental equation.

However, it should be emphasized that Wrigley's is an *approximate*, and not an *alternative*, implementation of the "principle of minimal sensitivity." He deviates from that philosophy in a small but critical way, and consequently encounters an embarrassing problem. I want to explain that this problem is spurious, and does not arise in a proper formulation.

For brevity, I take second order as an example. The procedure of Ref. 2 is *not* to take

$$\mathcal{R}^{(2)} = a(1 + r_1 a) \quad (1)$$

then seek its stationary point, and then set the higher-order β -function coefficients $c_2, c_3, \dots$ to zero. That procedure would make no sense. In fact, (1) has no stationary point—and this is the spurious problem alluded to above. The analysis in Ref. 2 actually begins with

$$\mathcal{R}^{(2)} \equiv a^{(2)}(1 + r_1 a^{(2)}) \quad (2)$$

in which $c_2, c_3, \dots$ are already set to zero as part of the definition of the second-order approximant (in just the same way that the higher-order coefficients $r_2, r_3, \dots$ are set to zero by definition). This way of proceeding is not a sly means of evading a difficulty; it is a logical necessity—simply because the second-order approximant $\mathcal{R}^{(2)}$ must be well defined in each fixed scheme, before one can begin to discuss its renormalization-scheme (RS) dependence.

Equation (1) is not well defined, given the limited information we have available at this order of perturbation theory. Even if we know the numerical values of $r_1(A)$ and $\tilde{\Lambda}_{(A)}$ in some scheme " A ," we cannot find the numerical value of the couplant $a_{(A)}$. We are therefore forced to approximate a , and the conventional procedure is to work with a truncated β

function. It should be stressed that this is a *definition*—it merely formalizes what is conventionally meant by "second-order perturbation theory", within a fixed scheme. [Any objection to setting $c_2, c_3, \dots$ to zero could be made with equal force to setting $r_2, r_3, \dots$ to zero: both truncations are necessary, practically, and have the same (lack of) justification.] This issue has nothing to do with "optimization" itself: It is merely a matter of defining what one is intending to "optimize."

Equation (2) is well defined and calculable within each RS—but of course it varies from one RS to another. That is entirely due to the truncations made in defining a usable approximant. The exact result, before truncations, was RS invariant. The degree of sensitivity to RS changes thus indicates how severely the result has been affected by the truncations. One is therefore motivated to seek the point at which the approximant is least sensitive to small variations in RS.² The approximant of interest is Eq. (2), and one is concerned with the effect of *both* truncations (in $\mathcal{R}$ and β) on the RS dependence of $\mathcal{R}^{(2)}$.

By beginning with Eq. (1), Wrigley generates an artificial problem. Setting $c_i = 0$ at the last stage will not make the derivatives with respect to c_i disappear. Consequently, he is still left with some highly embarrassing $\partial \mathcal{R}^{(2)} / \partial c_i$ expressions, which cannot be made to vanish, and which include terms of order a^3 —i.e., the *same* order as that to which he is supposedly working. To merely ignore these terms, as Wrigley is forced to do, is highly unsatisfactory. If this problem were real, then no genuine stationary point would exist, and the minimal-sensitivity approach would be a failure. However, Eq. (1) is ill defined even within a fixed scheme, and so cannot be the right starting point. It is not, in any useful sense, an approximant at all. Consequently, its RS dependence is of no interest. Provided we begin with a well-defined approximant, such as (2), the issue of c_i dependences simply does not arise. (See Ref. 3 for a related discussion.)

Locating the stationary point of (2), considered as a function of RS, is a well-defined mathematical problem,² and its solution leads to

$$r_{1, \text{opt}} = -\frac{1}{2}c(1 + c\bar{a})^{-1} \quad (3)$$

If one discards the $(1 + c\bar{a})^{-1}$ factor, then one is not at the stationary point of anything in particular. One has lost contact with the original motivation, and is no longer implementing the principle of minimal sensitivity.

Of course, it may be that approximating $r_{1, \text{opt}}$ by $-\frac{1}{2}c$ (or the analogous procedure in higher orders^{1,4}) is not *numerically* significant in that the error involved is much less than one's best estimate of the numerical error in the "optimum" result itself. *However, this depends on the particular application.* One cannot make any general statement. This point is especially relevant in third and higher orders, because the coefficients of the allegedly negligible terms involve $\bar{c}_2, \bar{c}_3$, etc., whose values are not fixed (unlike c), but become determined in the "optimization" process. Whether or not they are of order unity depends entirely on the application in question. For the same reason, I do not agree with the suggestion^{1,4,5} that optimization is always very similar to "fastest apparent convergence" (FAC). (Indeed, there are reasons for expecting the two methods to differ markedly in very high orders.⁶) It is not clear

to me why Wrigley considers "ease of comparison with FAC"—an ill-motivated and empirically unreliable^{2,3} method—to be an advantage.

In conclusion, Wrigley's work provides an *approximation* to the optimum result. It does not give an *alternative* implementation of the principle of minimal sensitivity, since it departs from the motivating philosophy in a way which destroys the logical coherence of the approach. Given that one starts from the generally accepted definition of perturbation theory, the principle of minimal sensitivity leads unambiguously to the full optimization procedure of Ref. 2. Of course, since the full procedure becomes algebraically complicated beyond second order, it is clearly valuable to have an approximation to it which is easier to use. (One such approximation, the so-called "improvement formula," was mentioned in Ref. 2, but I regard Wrigley's method as much superior to this.) The only proviso is that the extra error introduced by using a simplified form of optimization should be much less than the estimated error of the optimum result itself. Whether this is true or not will depend on the particular application.

¹J. C. Wrigley, Phys. Rev. D **27**, 1965 (1983) (preceding paper).

²P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981).

³P. M. Stevenson, Nucl. Phys. **B203**, 472 (1982).

⁴M. R. Pennington, Phys. Rev. D **26**, 2048 (1982).

⁵J. Kubo and S. Sakakibara, Phys. Rev. D **26**, 3656 (1982).

⁶P. M. Stevenson, CERN Report No. TH-3358, 1982 (unpublished).