STARK ESSENTIALS

of the determinantal approach to
time-independent spectral perturbation theory

Nicholas Wheeler, Reed College Physics Department
October 2000

Introduction. My primary intent in a manuscript completed and distributed yesterday[1] was to expose more clearly the anatomy of an idea explored in an earlier effort,[2] and to demonstrate that the "determinantal approach to spectral perturbation theory" is capable of producing useful results of higher order than can be achieved without excessive labor by the Rayleigh-Schrödinger method. I feel, however, that the utter simplicity of the essential idea, and the rationale for some curious aspects of its implementation, have yet to be made plain. My intent here is to see if I can rectify that state of affairs.

My approach will take me back—almost but not quite coincidentally—to the birthplace[3] of the determinantal method: I will suppose quantum mechanical state space to be (not ∞-dimensional but) 2-dimensional. We recognize that in such a toy world spectral perturbation theory is *not necessary*: the roots E_1 and E_2 (also the eigenvectors) of

$$\det\big\{\,\mathbb{H}^0 + \lambda\mathbb{V} - E\,\mathbb{I}\,\big\} = 0$$

can be found exactly, by the most elementary of means; it is not necessary to develop them "by adjustment"

$$E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \cdots \quad : \quad n = 1, 2$$

of the roots E_1^0 and E_2^0 of $\det\big\{\,\mathbb{H}^0 - E\,\mathbb{I}\,\big\} = 0$. But that is not to say that perturbation theory is *not possible* in such a world: one can—we will—use toy

[1] "Higher-order spectral perturbation by a new determinantal method," distributed on 3 October 2000. I will refer to this as Part B.

[2] "Perturbed spectra without $\overset{\text{it says here}}{\wedge}$ pain: New approach to time-independent perturbation theory" (April 2000).

[3] See pages 39/40 of Chapter I in Advanced Quantum Topics (Spring 2000).

perturbation theory to illustrate features of real-world theory. Our results will even serve to provide independent confirmation of the *accuracy* of some of the unfamiliar results reported in Part B.

I have organized my remarks around *Mathematica* commands illustrative of the simple points at issue, and it is my hope/expectation that my reader will be *executing* those commands as he/she works through my text. I will distribute a notebook "DeterminantalMethod2D.nb" to make that easy.

Basic plan of attack. Define[4]

$$\mathbb{H}^0 = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix} \quad : \quad \texttt{unpert} = \begin{pmatrix} \texttt{U}_1 & 0 \\ 0 & \texttt{U}_2 \end{pmatrix}$$

and[5]

$$\mathbb{V} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \quad : \quad \texttt{pert} = \begin{pmatrix} \texttt{a} & \texttt{b} \\ \texttt{c} & \texttt{d} \end{pmatrix}$$

$$\texttt{unit} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Command

$$\texttt{Det[unpert} + \lambda\,\texttt{pert} - \texttt{W unit]} \tag{1}$$

and get

$$W^2 - aW\lambda - dW\lambda - bc\lambda^2 + ad\lambda^2 - WU_1 + d\lambda U_1 - WU_2 + a\lambda U_2 + U_1U_2$$

Pretend not to notice that that the roots of that polynomial (of second order in W) are given by the quadratic formula. Announce to *Mathematica* our decision to work in (say) sixth order, and to concern ourselves with the induced displacement

$$U_1 \longrightarrow U_1 + \lambda W_1 + \lambda^2 W_2 + \cdots$$

of—specifically—U_1 (rather than U_2) by commanding

$$\%/.\texttt{W} \to \texttt{U}_1 + \lambda\texttt{W}_1 + \lambda^2\texttt{W}_2 + \lambda^3\texttt{W}_3 + \lambda^4\texttt{W}_4 + \lambda^5\texttt{W}_5 + \lambda^6\texttt{W}_6$$

Group terms according to the power of their λ-dependence by commanding

$$\texttt{Series[}\%,\{\lambda, 0, 6\}\texttt{]}$$

The output[6] is of such a form

$$D_1\lambda + D_2\lambda^2 + \cdots + D_6\lambda^6$$

as to motivate the definitions

---

[4] I use `U` to denote `U`nperturbed spectral values because *Mathematica* has preempted `E`.

[5] *Mathematica* has a curious aversion to doubly indexed matrix elements.

[6] Notice that we have been careful to discard terms of orders $\lambda^7, \ldots, \lambda^{12}$ which, though present, are misleading.

$$\begin{aligned}
D_1 &= -aU_1 + aU_2 + U_1W_1 - U_2W_1\\
D_2 &= -bc + ad - aW_1 - dW_1 + W_1^2 + U_1W_2 - U_2W_2\\
D_3 &= -aW_2 - dW_2 + 2W_1W_2 + U_1W_3 - U_2W_3\\
D_4 &= W_2^2 - aW_3 - dW_3 + 2W_1W_3 + U_1W_4 - U_2W_4\\
D_5 &= 2W_2W_3 - aW_4 - dW_4 + 2W_1W_4 + U_1W_5 - U_2W_5\\
D_6 &= W_3^2 + 2W_2W_4 - aW_5 - dW_5 + 2W_1W_5 + U_1W_6 - U_2W_6
\end{aligned}$$

These we set equal to zero and proceed to solve serially for $W_1, \ldots, W_6$.[7] We command

$$\texttt{Solve}[D_1 == 0,\ W_1]$$

and get back the response $W_1 \to a$. To translate from *Mathematica*'s notation to our orthodox notation, command

$$a\ /.\{a \to V_{11}, b \to V_{12}, c \to V_{21}, d \to V_{22}\}$$

and obtain

$$W_1 = V_{11} \tag{2.1}$$

To proceed to the next domino, command

$$D_2\ /.\ W_1 \to a$$

$$\texttt{Solve}[\% == 0,\ W_2]$$

and get

$$W_2 \to \frac{bc}{U_1 - U_2}$$

Translate

$$\frac{bc}{U_1 - U_2}\ /.\ \{a \to V_{11}, b \to V_{12}, c \to V_{21}, d \to V_{22}\}$$

and obtain

$$W_2 = \frac{V_{12}V_{21}}{U_1 - U_2} \tag{2.2}$$

Repeat the cycle

$$D_3\ /.\ \{W_1 \to a, W_2 \to \frac{bc}{U_1 - U_2}\}$$

$$\texttt{Solve}[\% == 0,\ W_3]$$

$$W_3 \to \frac{abc - bcd}{(U_1 - U_2)^2}$$

$$\frac{abc - bcd}{(U_1 - U_2)^2}\ /.\ \{a \to V_{11}, b \to V_{12}, c \to V_{21}, d \to V_{22}\}$$

---

[7] The Rayleigh-Schrödinger formalism—indeed: every perturbation theory I can think of—presents a similar "row of dominoes."

and obtain

$$W_3 = -\frac{V_{12}V_{21}(V_{11}-V_{22})}{(U_1-U_2)^2} \tag{2.3}$$

Similarly

$$W_4 = +\frac{\boxed{\text{numerator}}_4}{(U_1-U_2)^3} \tag{2.4}$$

$$W_5 = -\frac{\boxed{\text{numerator}}_5}{(U_1-U_2)^4} \tag{2.5}$$

$$W_6 = +\frac{\boxed{\text{numerator}}_6}{(U_1-U_2)^5} \tag{2.6}$$

where `Simplify[%]` supplies

$$\begin{aligned}
\boxed{\text{numerator}}_4 &= V_{12}V_{21}(V_{11}^2 - V_{12}V_{21} - 2V_{11}V_{22} + V_{22}^2)\\
\boxed{\text{numerator}}_5 &= V_{12}V_{21}(V_{11}-V_{22})(V_{11}^2 - 3V_{12}V_{21} - 2V_{11}V_{22} + V_{22}^2)\\
\boxed{\text{numerator}}_6 &= V_{12}V_{21}\big(V_{11}^4 + 2V_{12}^2V_{21}^2 - 4V_{11}^3V_{22} - 6V_{12}V_{21}V_{22}^2 + V_{22}^4\\
&\qquad + 6V_{11}^2(V_{22}^2 - V_{12}V_{21}) + 4V_{11}(3V_{12}V_{21}V_{22} - V_{22}^3)\big)
\end{aligned}$$

Comparison with the implications of exact analysis. In 2-dimensional theory—exceptionally—it is possible/easy to proceed *exactly*: one has only to solve a quadratic polynomial to obtain[3]

$$\begin{aligned}
W_1 = \tfrac{1}{2}\Big\{&\big[(U_1+\lambda V_{11}) + (U_2+\lambda V_{22})\big] \qquad (3.1)\\
&+\sqrt{\big[(U_1+\lambda V_{11}) - (U_2+\lambda V_{22})\big]^2 + 4\lambda^2 V_{12}V_{21}}\;\Big\}\\
W_2 = \tfrac{1}{2}\Big\{&\big[(U_1+\lambda V_{11}) + (U_2+\lambda V_{22})\big] \qquad (3.2)\\
&-\sqrt{\big[(U_1+\lambda V_{11}) - (U_2+\lambda V_{22})\big]^2 + 4\lambda^2 V_{12}V_{21}}\;\Big\}
\end{aligned}$$

Expansion of the right side of (4.1)—the work of an instant for *Mathematica*—precisely reproduces (2).

Formulæ for the perturbation of U_2 can be obtained substitutionally from (2): $1 \rightleftarrows 2$, and agree with results obtained by expansion of (3.2).

Comparison with the results of the general theory. The constrained sums present in the formulæ (B35) in which the determinantal method (also, of course, the Rayleigh-Schrödinger method) culminates…collapse into single terms in the

2-dimensional case. We have, in the present notation,

$$
\begin{aligned}
W_1 &= V_{11}\\
W_2 &= -\frac{V_{12}V_{21}}{U_2-U_1}\\
W_3 &= \frac{V_{12}V_{22}V_{21}}{(U_2-U_1)^2} - V_{11}\cdot\frac{V_{12}V_{21}}{(U_2-U_1)^2}\\
&= -\frac{V_{12}V_{21}(V_{11}-V_{22})}{(U_2-U_1)^2}\\
W_4 &= \Big[\frac{V_{12}V_{21}}{U_2-U_1}\Big]\Big[\frac{V_{12}V_{21}}{(U_2-U_1)^2}\Big] - V_{11}^2\frac{V_{12}V_{21}}{(U_2-U_1)^3}\\
&\qquad + V_{11}\Big[\frac{2}{(U_2-U_1)^3}\Big]V_{12}V_{22}V_{21} - \frac{1}{(U_2-U_1)^3}V_{12}V_{22}^2V_{21}\\
&= \frac{V_{12}V_{21}(V_{12}V_{21}-V_{11}^2+2V_{11}V_{22}-V_{22}^2)}{(U_2-U_1)^3}
\end{aligned}
$$

These specialized implications of general formulæ are readily seen to be in precise agreement with the results (2) of direct calculation. . . which is gratifying, and inspires confidence in the accuracy of the conclusions reached by intricate analysis in Part B. I leave to my industrious reader the pleasure of showing that the agreement is precise also in 5^{th} and 6^{th} order.

Comparison with the methods of the general theory. Notice first of all how much more complicated the argument would have been if—with Rayeigh-Schrödinger —we had had to concern ourselves, on a parallel track, with the (at each step renormalized) eigen*vectors*. All of that, in contexts where we have physical interest only in the perturbed spectrum, would have been extraneous effort. . . which touches upon the main selling point of the determinantal method.

The direct/elementary calculation so casually taken at (1) becomes increasingly awkward as the dimension of $\mathbb{H}^0$ becomes large, and is not feasible (or even meaningful) in the ∞-dimensional limit. If any step toward creation of the determinantal method can be said to approach dim brilliancy, it is surely the step where I introduce[8] this "remarkable identity which deserves to be better known:"

$$
\det(\mathbb{I}+\lambda\mathbb{M}) = 1+\lambda T_1+\tfrac{1}{2!}\lambda^2\begin{vmatrix}T_1 & T_2\\ 1 & T_1\end{vmatrix}+\tfrac{1}{3!}\lambda^2\begin{vmatrix}T_1 & T_2 & T_3\\ 1 & T_1 & T_2\\ 0 & 2 & T_1\end{vmatrix}
+\tfrac{1}{4!}\lambda^2\begin{vmatrix}T_1 & T_2 & T_3 & T_4\\ 1 & T_1 & T_2 & T_3\\ 0 & 2 & T_1 & T_2\\ 0 & 0 & 3 & T_1\end{vmatrix}+\cdots \tag{4}
$$

[8] See (13) in Part A, (6.1) in Part B.

Here $T_p \equiv \operatorname{tr}\mathbb{M}^p$. If $\mathbb{M}$ is $N\times N$ then the series (3) can be shown to terminate at order λ^N by virtue of the Cayley-Hamilton theorem.

Though the statements

$$\det(\mathbb{I}+\lambda\mathbb{M}) = 1+\lambda\operatorname{tr}\mathbb{M} + \lambda^2\det\mathbb{M}$$
$$\det\mathbb{M} = \tfrac{1}{2}\big\{(\operatorname{tr}\mathbb{M})^2 - \operatorname{tr}\mathbb{M}^2\big\}$$

are quite *frequently* encountered, I have in fact *never encountered the general proposition* (4) *in the literature*. But it has played a recurrent role in my own writing for more than four decades. I take this opportunity to describe how (4) came to my attention, to sketch what I know of its history, and to allude to a few of its diverse applications.

Historical note. In 1958 I was at work on a document best forgotten,[9] and had encountered a nest of combinatorial problems rooted in what might be called "the chain rule to the n^{th}": the problem of constructing the n^{th} derivative $F^{(n)}(x)$ of a composite function $F(x)\equiv f[g(x)]$. Proceeding in naive ignorance of the classical literature, I had "discovered" that

$$F^{(n)}(x) = \sum_{m=0}^{n} f^{(m)}[g]\cdot\sum(n;a_1,a_2,\ldots,a_n)\big\{g'\big\}^{a_1}\big\{g''\big\}^{a_2}\cdots\big\{g^{(n)}\big\}^{a_n} \tag{5}$$

where $(n;a_1,a_2,\ldots,a_n) \equiv n!/(1!)^{a_1}a_1!(2!)^{a_2}a_2!\cdots(n!)^{a_n}a_n!$ and where $\sum$ is subject to the constraints $a_1+a_2+\cdots+a_n = m$ and $a_1+2a_2+\cdots na_n = n$. But the latter circumstance made (5) almost useless for my purposes. While on a trip to the math library to consult an article on the subject by A. Dresden[10] I happened—entirely by accident—upon

> **Advanced Problem 4782**. Proposed by V. F. Ivanoff, San Carlos, California: Given a composite function $F(x)\equiv f[g(x)]$. Denoting the n^{th} derivative of $f[g]$ by $D^n f$, and the derivatives of $g(x)$ by $g', g'', \ldots, g^{(n)}$, show that

$$F^{(n)} = \begin{vmatrix} g'D & g''D & g'''D & g''''D & \ldots & g^{(n)}D \\ -1 & g'D & 2g''D & 3g'''D & \ldots & \binom{n-1}{1}g^{(n-1)}D \\ 0 & -1 & g'D & 3g''D & \ldots & \binom{n-2}{1}g^{(n-2)}D \\ 0 & 0 & -1 & g'D & \ldots & \binom{n-3}{1}g^{(n-3)}D \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \ldots \;\; -1 & g'D \end{vmatrix} f$$

[9] "Foundations & applications of the Schwinger action principle," (Brandeis University doctoral dissertation, February 1960).

[10] Amer. Math. Monthely **50**, 9 (1943).

I devised several alternative proofs of "Ivanoff's formula,"[11] and found that it served my needs—of which more in a moment—quite admirably.

Ultimately, a 2-page solution of Problem 4782 was published by one Frank Schmittroth of Oregon State University,[12] who mentions that "other formulas for the n^{th} derivative of a compound function are given in [Dresden, cited above]; M. McKiernan, Monthly **63**, (1956);E. P. Adams, *Smithsonian Misc. Collection*, **74**, 157 (1923); I. M. Ryzhik, *Tables of Series, Products and Integrals*, p. 20 (3^{rd} edition 1957) [see §0.43 "The n^{th} derivative of a composite function" in I. S. Gradshteyn & I. M. Ryzhik (4^{rd} edition 1965)]."

But unmentioned by Schmittroth is Faà di Bruno, to whom the authors of "Chapter 20: Combinatorial Analysis" in M. Abramowitz & I. Stegun's *Handbook of Mathematical Functions* (1965) attribute (5). Only recently have I been able to discover who Faà di Bruno (1825–1888) was,[13] and that he wrote a variant of (5) into his *Traite Elementaire du Calcul* (1869).

Abramowitz & Stegun present "Faà di Bruno's formula" at the bottom of page 823. At the top of page 824 one encounters a determinant of Ivanoff's design; i.e., of the design encountered in (4). But I have been unable to discover the *point* of that insertion, which appears to have passed over the heads also of Abramowitz & Stegun's other many readers. One encounters determinants of roughly that design also in Thomas Muir's *A Treatise on the Theory of Determinants* (1933): see especially his Chapter XXI. But nowhere in that compendious volume do they appear—so far as I have been able to determine—in connection with the description of the characteristic polynomial.

At various times I have used Ivanoff's formula to develop aspects of the statistical theory of cummulants, properties of Bell polynomials, properties of statistical mechanical partition functions, other things. I have used it to show, for example, that the Hermite polynomials

$$H_n(x) = (-)^n e^{\frac{1}{2}x^2}\left(\tfrac{d}{dx}\right)^n e^{-\frac{1}{2}x^2}$$

[11] It is my impression that V. F. Ivanoff was a one of the community of retired mathematicians who find it amusing to submit problems/solutions to the Monthly; he had, in any event, been doing so several times a year for many years, but after 1961 his contributions ceased. Problem 4782 appears on page 212 of Amer. Math. Monthly **65** (1958).

[12] Amer. Math. Monthly **68**, 69 (1961). Apparently Schmittroth was the only person (other than myself) who was attracted by Ivanoff's problem; the usual notice that "solutions were submitted also by…" is absent. I was at CERN, thinking about other matters in 1961, and did not learn of Schmittroth's solution until many years later.

[13] See footnote 8 in Part A.

can (for some purposes usefully) be described

$$H_n(x) = \begin{vmatrix} x & 1 & & & & & \\ 1 & x & 2 & & & & \\ & 1 & x & 3 & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & x & n{-}1 \\ & & & & & 1 & x \end{vmatrix}$$

But the most striking applications have, in my experience, all stemmed from (4), which is got by using Ivanoff's formula, Taylor's theorem and the lovely identity

$$\det\big\{\mathbb{I} + \lambda\,\mathbb{M}\big\} = e^{\text{tr}\,\log\{\mathbb{I}+\lambda\mathbb{M}\}}$$

in combination. Details can be found in "Some applications of an elegant formula due to V. F. Ivanoff."[14] For application of (4) to a problem area which has (so far as I am aware) nothing to do with perturbation theory, see "A Mathematical Note: Algorithm for the efficient evaluation of the trace of the inverse of a matrix" (1996); this material was developed in response to a question posed by Richard Crandall, and was used by him in a paper "On the quantum zeta function."[15]

Such, then, are the tangled roots of the idea most basic to determinantal perturbation theory, from which the method acquires its most distinctive features, and to which it owes its success. The recollection that I became acquainted with that idea in happy days, long ago, when I shared the office of Sam Schweber and the companionship of some good people. . . the thought that it seems now to have yielded a little bit of fruit. . . fill me with an October contentment.

[14] These are notes for a seminar presented 28 May 1969 to the Applied Math Club at Portland State University, and can be found in COLLECTED SEMINARS 1963–1970.

[15] J. Phys. A: Math. Gen. **29**, 6795 (1996).