Larry Spruch and Variational Principles

A. R. P. Rau *

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803

Abstract

The development and use of variational principles for various problems in atomic physics by Larry Spruch and his collaborators over the last forty years is reviewed and a roadmap provided for their publications.

Typeset using REVTEX

I. INTRODUCTION

Variational principles are among the most important tools in physics. It is a powerful idea that one can write an expression so constructed that it improves on an initial guessed estimate of the solution or physical quantity being sought. An initial error in the guess is reduced quadratically, thus allowing analytical or numerical improvements, including iterative application so as to converge towards the exact answer.

The Rayleigh-Ritz principle for the ground state energy is the best known and most widely used of all variational principles. It can be said safely that a very large fraction of all theoretical calculations in physics use this principle. The next major contributions are the Kohn, Hulthen and Schwinger variational principles formulated in scattering theory but also applicable elsewhere. After these major contributions of over fifty years ago, the most significant contributions have been those made by Larry Spruch and his group who have both provided new variational expressions and clarified and made systematic the basics of the subject.

Their work provides a general method for constructing variational principles for almost any given problem in mathematical physics, thus removing the need for ingenuity in writing down the variational expression. This work also disabuses a feeling some (many?) had that only a few, specific problems admit variational principles. All the known principles, including those listed above, can be recovered through application of this method which is at the same time very simple to apply. Many new variational expressions have been developed as well, including by others who have adopted the general method of construction. At the same time, Spruch has also emphasized the distinctions between terms often used interchangeably, such as variational principles, variational bounds, extremum principles and non-variational bounds.

This brief survey confines itself to the work on variational principles. It should be read in conjunction with the companion article on variational bounds by Yukap Hahn.

II. DEFINITIONS AND WORK OF THE EARLY 1960S

Spruch's early work, starting with a 1959 paper [1] , actually dealt with variational bounds on the scattering length through the Kohn and Hulthen variational principles. Since a succession of papers over the next half-dozen years dealt with demonstrating that these principles actually provided bounds and under increasingly less-restrictive conditions, it is useful to start with the distinction that he has always emphasized. By a variational principle, we mean that the error in the final estimate is quadratic in the error of the input functions, that is, the expression is stationary with respect to departures of the input functions from their exact values. The input functions are called trial functions and their departure from the exact values called first-order errors. The variational expression itself will contain only second- and higher-order errors. Equivalently, such principles could be termed stationary principles, a term Spruch would actually favor [2] , reserving the usage variational only when the expression and the functions therein contain parameters that can be varied. However, current usage of ”variational principles” is sufficiently widespread that we will use the terms variational and stationary interchangeably.

But we will preserve the next distinction, namely when one can also guarantee that the expression is not only stationary but the second-order error is of well defined sign. In such a case, we have an ”extremum principle” or, equivalently, a ”variational bound”, upper or lower bound depending on the sign of the second-order error. The Rayleigh-Ritz result is, therefore, actually an extremum principle or a variational upper bound, its value both stationary and guaranteed to lie above the exact ground state energy. Variational bounds, when available, are of course much more powerful than ”mere” variational principles, the guaranteed monotonic convergence from one direction towards the exact answer being superior to possibly oscillatory convergence, particularly in numerical applications in a many parameter space. But the problems for which one can write down variational bounds form a small subset, depending on ”the physics of the problem” (for example, that the state is the ground state of the system and therefore has the lowest energy, or that the entropy is an extremum), whereas variational principles can always be constructed as first recognized by Spruch and his collaborators. As one example, a specific claim that no variational principle obtains for the temperature distribution in heat transfer problems was shown to be incorrect by explicit construction of such a principle through the general procedure [3] . This ability to construct variational principles for almost any given problem, and the fact that the stationary aspect is itself useful even when not an extremum, makes variational principles also of great interest.

Much of Spruch's work in the 1960s was concerned with demonstrating that the Kohn, Hulthen and Schwinger variational expressions were actually variational bounds in many contexts. Since these are covered in Hahn's article, I will mention here only two reviews [4,5] which are concerned with various variational bounds and the contexts in which they apply but which also necessarily have much useful material on variational principles. Two items are particularly noteworthy. One is the Kato identity [6] which became a powerful tool in the hands of the Spruch group to derive from it an expression that is an extremum principle. The second is an identity for the inner product (a,x), where a is a known function or vector, and x an unknown one defined through an equation $Kx = a$, where K is some linear operator on x. Identities are expressions for exact quantities, such as the (a,x) in this example, which involve both known quantities such as K, a, and a trial or guessed x_t, as well as the unknown x itself. Although, therefore, only a formal expression, a recurrent theme in Spruch's work is that such identities are the starting point for deriving variational or extremum principles. In [5], the focus was on kernels K that are non-negative so that a minimum principle could be obtained, but this same procedure with a general K leads to a variational principle. Indeed, this is the simplest route to the general Kohn and Schwinger variational principles.

In the late 1960s, the group developed variational principles for three-body break-up problems [7] . Since this subject is treated in more detail in the companion article by Mike Lieber, I merely note here that the first paper developed a Kohn variational principle for short range potentials. Since this requires as input time-reversed states as well, one needs the three-body asymptotic wave function, generally difficult to obtain. However, this was circumvented by modifying the derivation so as to involve only the plane wave part of the three-body wave function. It was also necessary to ensure that divergent integrals do not occur in the final variational expression. The next paper gave an alternative, more rigorous derivation based on the Faddeev formalism. This paper also gave an adjoint version of the Kohn principle in which the operator $H-E$ acts on the final state instead of the initial, besides deriving a Schwinger variational principle as well.

III. GENERAL CONSTRUCTION OF VARIATIONAL PRINCIPLES

When I joined Larry Spruch as a post-doctoral associate in 1970, he was interested in variational principles for matrix elements of operators. It is standard practice, even today, that physical properties of a system, whether in nuclear, atomic or condensed matter physics, are evaluated by sandwiching the operator of interest between ”variationally derived” wave functions. By this one means that the wave function (or, one-particle density in today's widely used density functional methods) has been obtained through a Rayleigh-Ritz calculation, giving an accurate value for the energy of the state. But, it is immediate and well recognized that this only ensures that the energy correction is of second order, the wave function itself containing first-order errors. Therefore, matrix elements of other operators will also have first-order corrections when evaluated with such wave functions. It would be much nicer were one to have a variational principle for the matrix element itself, guaranteed to differ from the exact value by second-order errors, when evaluated with trial functions.

With this as motivation, we examined the few results that were available on variational principles for matrix elements. A variational principle for diagonal matrix elements of a self-adjoint operator W in a state ϕ was known [8–10] as also limited results on off-diagonal matrix elements [11] . There was also an unpublished report by Borowitz and Gerjuoy [12] , and Ed Gerjuoy was a frequent visitor to Spruch's group at that time. In examining these, we saw in them the seeds of a very general construction, applicable much more widely.

The expression for a diagonal matrix element, denoted by $\langle W \rangle \equiv \phi^{\dagger} W \phi$, serves to make the point. We use the dagger sign to denote the Hermitian adjoint along with matrix multiplication or integration over relevant coordinates. With ϕ defined by

$$H\phi = E\phi, \tag{1}$$

and

$$\phi^\dagger \phi = 1, \tag{2}$$

and an auxiliary or ”adjoint” function L defined by

$$(H-E)L = -W\phi + \langle W\rangle \phi, \tag{3}$$

and

$$L^\dagger \phi = 0, \tag{4}$$

the following is a variational expression for $\langle W\rangle$:

$$\langle W\rangle_{\rm var} = \phi_t^\dagger W \phi_t + L_t^\dagger (H-E)\phi_t + [(H-E)\phi_t]^\dagger L_t. \tag{5}$$

With trial solutions ϕ_t and L_t of Eqs. (1-4), the expression in Eq.(5) differs from the exact $\langle W\rangle$ only in second order terms such as $(\delta\phi)^2$ and $\delta L \delta\phi$, where

$$\delta\phi \equiv \phi_t - \phi,\ \delta L \equiv L_t - L. \tag{6}$$

With a background that had me familiar with null Lagrangians for deriving, for instance, the Dirac equation or elsewhere in field theory, I was struck by the structure of Eq.(5), that the right-hand side had a first term which we would write down naturally as the desired matrix element evaluated with the trial function and the other terms were like ”adding a zero” since they vanish when $\phi_t = \phi$. This suggested that quite generally, when one seeks some quantity that depends on an unknown ϕ, which may even stand collectively for a set of functions, a variational expression is given by writing that quantity in terms of ϕ_t and adding the defining equations for ϕ such as Eqs.(1-2) as constraints incorporated through Lagrange multipliers. The notation L above is deliberate, to stand for Lagrange multipliers. Given the nature of the defining equations and what is being sought, these multipliers may take different forms: constants, functions as in Eq.(5), Green’s functions, vectors, matrices, etc. Thus, starting from Eqs.(1-2), one writes down Eq.(5), with L_t at this point a trial solution of some exact Lagrange function that is not yet known. But, already at this point, regardless of what L or L_t are, the expression satisfies the requirement that if by some lucky circumstance, ϕ_t equals the exact ϕ, then the expression gives the exact $\langle W\rangle$ being sought. Next, as a matter of ”construction”, when one is not so lucky and there are first-order errors as defined in Eq.(6), one *requires* that the right-hand side of Eq.(5) have no first-order quantities which gives at this point the defining equations for the Lagrange functions L. Thus, one ”engineers” or constructs a variational principle.

In our first papers [13,14], we presented this general scheme with several illustrations as well as a construction of related identities. These identities are just a rewrite of the equations that set first-order quantities in $\delta\phi$ equal to zero. They involve the exact Lagrange functions which are, of course, unknown, being often at least as difficult to get as the exact ϕ. But, as with the Kato identity, which is an example of such identities, they are useful starting points for examining the form of the second-order error and whether the variational principle can be made a variational bound. As an illustration, an identity associated with the variational principle in Eq.(5) and to which it reduces upon replacing all exact functions by their trial values is [13]:

$$(\phi_t^\dagger\phi+\phi^\dagger\phi_t)\langle W\rangle=\phi_t^\dagger W\phi+\phi^\dagger W\phi_t+L^\dagger(H-E)\phi_t+[(H-E)\phi_t]^\dagger L. \tag{7}$$

The general construction has the merit that it applies to defining equations of a very broad class, whether differential, integral, integro-differential, linear or nonlinear. There is no restriction on the form of the equations satisfied by ϕ. The Lagrange multipliers satisfy in general a different set of equations than those for the original ϕ, an ”adjoint” set of equations. Only in special cases, for self-adjoint problems, will the two coincide, as in the Rayleigh-Ritz problem for the ground state energy when W equals the Hamiltonian H. Since the Lagrange multipliers are introduced only linearly, the equations for L are always linear. This is an important feature, and was explored in applications to nonlinear and integral equations in [15–18].

An example of a variational principle in a time-dependent problem, besides the one on heat transfer already mentioned [3], was charge transfer in proton-hydrogen collisions as treated in the impact parameter approximation. A time-dependent potential experienced by the electron causes transitions, and Shakeshaft and Spruch [19] showed how to make expressions that use Sturmian basis functions variational. See the companion article by Robin Shakeshaft. Another extension was the paper on variational principles and identities along with so-called supervariational principles which are engineered to have only errors of order higher than the second [20] . This paper also provided the ”ultimate variational principle in quantum physics”, for the wave function itself!

IV. USE OF THE VARIATIONAL PRINCIPLES

As sketched in the above section, the general construction of variational principles for any well-defined problem in mathematical physics turned out to be simple and straightforward. Indeed, I recall clearly Larry’s pleasure at the simplicity of the procedure, that Lagrange’s method of incorporating the defining equations of the system as constraints afforded a variational procedure. He said that ”ingenuity should not be necessary” [21], that it is progress in physics when we do not have to rely on the ingenuity of a Rayleigh or a Schwinger but reduce procedures to such a routine that any one of us can apply them. In [21], Larry also made the point that variational principles are not only useful for specific calculations but also for establishing analytical results, as indeed Schwinger [22] did in deriving the form of the effective range expansion. See the companion article by Lenny Rosenberg on modified effective range theories.

But there are subtleties in the choice of the trial functions which are inputs into the variational expressions. Once again, the Rayleigh-Ritz principle is unique in that the trial function can be chosen ”almost blindly”, satisfying only some very broad constraints such as being normalizable. However, more care is necessary in using an expression such as Eq.(5), particularly in the choice of L_t. We realized that even the few extant variational principles for matrix elements had not sufficiently appreciated that too casual a choice of L_t as a solution of Eqs.(3-4) with trial functions in place of exact ones throughout can lead to a loss of the variational principle. The expression in Eq.(5) reduces then just to the first term with its first-order errors. As he said in his review [21], authors have been ”caught with their parameters down”. In [23], we introduced an auxiliary extremum principle for obtaining L_t which required modifying the Hamiltonian in Eq.(3) so that no singularities would be encountered in inverting $H-E$. An alternative, involving a sequence of modified Hamiltonians, was given in [24]. Without these prescriptions, the formally stationary expression cannot be used as a variational principle except in very limited circumstances. A short review article [25] summarized the above developments.

Already, even without the additional extremum principle, an early application of the variational principle for matrix elements was made in [26] to the calculation of expectation values of powers of r for atoms from He through Ar. The first application of the modified Hamiltonian was in [27] to diagonal matrix elements, the diamagnetic susceptibility and form factor of He, calculated with Hylleraas wave functions as trial functions. This was extended in [28] to off-diagonal matrix elements when the generalized oscillator strength for the excitation of the 2^1P state of He was calculated. An application to scattering was in [29] which established a variational principle for the scattering length when the target wave function is imprecisely known (as it is in almost any atomic physics problem except for the hydrogen atom!). This was tested in a numerical example for scattering of electrons and positrons by hydrogen, assuming an approximate description of the hydrogen atom [2] .

A major review article [30] on the unified construction of variational principles gave many illustrations, including pedagogical ones such as Newton's method for finding roots of a function, inversion of a matrix, solving a resistor network, etc. It also spelled out questions regarding phases of complex quantities, time reversal invariance, and others which render variational principles practically useful.

V. SUMMARY

The work of Larry Spruch and his collaborators over the last forty years on the construction and use of variational principles in physics has been summarized and the relevant publications listed with commentary.

ACKNOWLEDGMENTS

It is with great pleasure that I look back on a thirty-year association with Larry Spruch, starting with the two years spent at New York University with him and the many interactions since, discussing varied topics in and out of physics. I have learnt much about variational principles, and more, from him.

* e-mail: arau@rouge.phys.lsu.edu

REFERENCES

[1] Larry Spruch and L. Rosenberg, Phys. Rev. **116**, 1034 (1959) and **117**, 143 (1960).

[2] I. Aronson, R. Blau, C.J. Kleinman, L. Rosenberg and Larry Spruch, Phys. Rev. A **19**, 1568 (1979).

[3] K. Kalikstein and Larry Spruch, Phys. Rev. A **29**, 832 (1984).

[4] Larry Spruch, in *Lectures in Theoretical Physics*, edited by W.E. Brittin, W.B. Downs and J. Downs (John Wiley, New York, 1962), Vol.4, p.161.

[5] Larry Spruch, in *Lectures in Theoretical Physics*, edited by S. Geltman, K.T. Mahantappa and W.E. Brittin (Gordon and Breach, New York, 1969), Vol. XI-C, p.77.

[6] T. Kato, Phys. Rev. **80**, 475 (1950) and Progr. Theo. Phys. **6**, 295 and 394 (1951).

[7] M. Lieber, L. Rosenberg and Larry Spruch, Phys. Rev. D **5**, 1330 and 1347 (1972).

[8] A. Dalgarno and A.L. Stewart, Proc. Roy. Soc. **A257**, 534 (1960).

[9] Charles Schwartz, Ann. Phys. (N.Y.) **2**, 156 and 170 (1959).

[10] L.M. Delves, Nucl. Phys. **41**, 497 (1963).

[11] L.M. Delves, Nucl. Phys. **45**, 313 (1963).

[12] S. Borowitz and E. Gerjuoy, JILA Report No. 36, University of Colorado, Boulder, 1965 (unpublished).

[13] E. Gerjuoy, A.R.P. Rau and Larry Spruch, J. Math. Phys. **13**. 1797 (1972).

[14] E. Gerjuoy, A.R.P. Rau and Larry Spruch, Phys. Rev. A **8**, 662 (1973).

[15] A.R.P. Rau, Mon. Not. R. astr. Soc. **168**, 273 (1974).

[16] A.R.P. Rau, J. Math. Phys. **17**, 1338 (1976).

[17] A.R.P. Rau, Mitio Inokuti and D.A. Douthat, Phys. Rev. A **18**, 971 (1978).

[18] Mitio Inokuti, D.A. Douthat and A.R.P. Rau, Phys. Rev. A **22**, 445 (1980).

[19] R. Shakeshaft and Larry Spruch, Phys. Rev. A **10**, 92 (1974).

[20] E. Gerjuoy, A.R.P. Rau, L. Rosenberg and Larry Spruch, J. Math. Phys. **16**, 1104 (1975).

[21] Larry Spruch, in *The Physics of Electronic and Atomic Collisions*, Invited Papers IX ICPEAC, edited by J.S. Risley and R. Geballe (University of Washington Press, Seattle, 1976), p.685.

[22] J. Schwinger, Phys. Rev. **72**, 742 (1947).

[23] E. Gerjuoy, A.R.P. Rau, L. Rosenberg and Larry Spruch, Phys. Rev. A **9**, 108 (1974).

[24] E. Gerjuoy, L. Rosenberg and Larry Spruch, J. Math. Phys. **16**, 455 (1975).

[25] R. Blau, A.R.P. Rau, L. Rosenberg and Larry Spruch, in *Electron and Photon Interactions with Atoms*, edited by H. Kleinpoppen and M.R.C. McDowell (Plenum, New York, 1976), p.601.

[26] R.O. Mueller, A.R.P. Rau and Larry Spruch, Phys. Rev. A **10**, 1511 (1974).

[27] R. Shakeshaft, L. Rosenberg and Larry Spruch, Phys. Rev. A **14**, 1989 (1975).

[28] J.M. Wadehra, R. Shakeshaft and Larry Spruch, Phys. Rev. A **18**, 344 (1978).

[29] R. Blau, L. Rosenberg and Larry Spruch, Phys. Rev. A **15**, 1475 (1977).

[30] E. Gerjuoy, A.R.P. Rau and Larry Spruch, Rev. Mod. Phys. **55**, 725 (1983).