No-Go Theorem for Energy Surfaces

B. G. Giraud
Institut de Physique Théorique, Centre d'Etudes Saclay, 91190 Gif-sur-Yvette, France[*]

S. Karataglidis
Department of Physics, University of Johannesburg, P. O. Box 524, Auckland Park, 2006, South Africa[†]

(Dated: June 24, 2011)

Concavity properties prevent the existence of significant landscapes in energy surfaces obtained by energy minimizations under constraints. The contradiction holds at finite temperatures T as well as at zero temperature.

PACS numbers: 21.60.De, 21.60.Ev, 31.15.A-, 71.15.-m

arXiv:1106.4702v1 [nucl-th] 23 Jun 2011

The concept of collective [1] degrees of freedom has been of central importance in nuclear physics, if only to generate models involving much fewer degrees of freedom than the true number, $3A$, needed for a proper microscopic description of a nucleus of mass number A. Under adiabatic assumptions, or even proofs, the description of nuclear dynamics can be compressed into slow motions of a few collective degrees of freedom B, while the other degrees, faster, can be averaged out. Often, such collective degrees can be one-body operators, $B = \sum_{i=1}^{A} \beta(\mathbf{r}_i, \mathbf{p}_i, \sigma_i, \tau_i)$, where $\mathbf{r}_i, \mathbf{p}_i, \sigma_i, \tau_i$ refer to the position, momentum, spin, and isospin, respectively, of nucleon i. The summation over all nucleons provides, intuitively at least, a motivation for more inertia in B than in the individual (nucleonic) degrees of freedom β_i.

The concept of energy surfaces [2] has been almost as important. Given a "coordinate-like" collective operator B and its expectation value $b \equiv \langle B \rangle$, every collective model accepts the common wisdom that there exists an energy function, $e(b)$, and an inertia parameter, $\mu(b)$, that drive the collective dynamics. This concept is at the core of any theory of fission or fusion, for instance. Keywords like "saddles", "valleys", "barriers", etc., flourish.

Simultaneously, it is implicitly assumed that the function, $e(b)$, results from an energy *minimization* under constraint. Namely, while the system evolves through various values of b, it is believed to tune its energy to achieve a local minimum. To illustrate, consider the nuclear Hamiltonian, $H = \sum_i t_i + \sum_{i<j} v_{ij}$, where t and v denote the usual kinetic and interaction operators. Given a suitable trial set of density operators $\mathcal{D}$ in many-body space, normalized by $\mathrm{Tr}\mathcal{D} = 1$, a theory for an energy surface most often reads,

$$e(b) = \inf_{\mathcal{D} \Rightarrow b} \mathrm{Tr}\left\{H\mathcal{D}\right\}, \tag{1}$$

where Tr is a trace in the many-body space for A nucleons and the constraint, $\mathcal{D} \Rightarrow b$, enforces $\mathrm{Tr}\left\{B\mathcal{D}\right\} = b$. In practice, this requires a diagonalization, or at least a minimization, of the constrained operator, $\mathcal{H} \equiv H - \lambda B$, where λ is a Lagrange multiplier. The purpose of this work is to show that this constrained minimization can only return *concave* [3] functions $e(b)$. Maxima are impossible. In the trivial generalization where several collective operators $B_1, \ldots, B_N$, are involved, concavity stills holds, hence saddle points are excluded as well as maxima, whether local or not. Only an absolute minimum is available.

Consider, indeed, the exact diagonalization at $T = 0$, or an exact partition function at finite T. However, if λ is a positive number and the spectrum of B is not bounded from above, there is a risk that the spectrum of $\mathcal{H}$ extends to $-\infty$, hence no ground state of $\mathcal{H}$ exists to provide a minimal energy. A similar risk occurs if λ is negative and B is not bounded from below. To start, we shall assume that either $\mathcal{H}$ is always bounded from below or both H and B have been projected in a finite subspace of A-body basis states; in the latter case, the problem reduces to handling finite matrices. Let $\psi(\lambda)$ be the ground state of $\mathcal{H}$. (For the sake of simplicity, we assume that there is no degeneracy.) The corresponding eigenvalue, $\varepsilon(\lambda)$, is stationary with respect to variations of ψ, among which is the "online" variation, $d\lambda\,(d\psi/d\lambda)$, leading to the well-known first derivative, $d\varepsilon/d\lambda = -b \equiv -\langle\psi\,|B|\,\psi\rangle$. Let $P = |\psi\rangle\langle\psi|$ and $Q = 1 - P$ be the projector and the complementary projector, respectively. Brillouin-Wigner perturbation theory yields the first derivative of ψ, *viz.*

$$\frac{d\,|\psi\rangle}{d\lambda} = -\frac{Q}{\varepsilon - Q\mathcal{H}Q} B\,|\psi\rangle\,. \tag{2}$$

This provides the second derivative of ε,

$$-\frac{db}{d\lambda} \equiv -\frac{d}{d\lambda}\langle\psi\,|B|\,\psi\rangle = 2\left\langle \psi \middle| B\frac{Q}{\varepsilon - Q\mathcal{H}Q}B \middle| \psi \right\rangle. \tag{3}$$

Since the operator, $(\varepsilon - Q\mathcal{H}Q)$, is clearly negative-definite, the energy, ε, is a convex function of λ. It is trivial to prove that the same convexity holds for the ground state eigenvalue $\varepsilon(\lambda_1, \ldots, \lambda_N)$ if several constraints, $B_1, \ldots, B_N$, are used. When a temperature T is considered, a partition function is calculated from an $\mathcal{H}$ bounded from below, and one obtains a free energy,

$\varepsilon(\lambda_1,\ldots,\lambda_N;T)$, that also contains the entropy contribution, $-TS$, where $S=-\mathrm{Tr}\left\{\mathcal{D}\ln\mathcal{D}\right\}$. A proof of the convexity of an exact $\varepsilon(\lambda_1,\ldots,\lambda_N;T)$ is also easy [4].

At $T=0$, the traditional Legendre transform expresses the energy, $e\equiv\langle\psi\left|H\right|\psi\rangle$, in terms of the constraint value(s) rather than the Lagrange multiplier(s). For simplicity, set one constraint only, the generalization to $N>1$ being easy. Since $e\equiv\varepsilon+\lambda b$, and $d\varepsilon=-bd\lambda$, then $de=\lambda db$, giving $de/db=\lambda$, a most traditional result for conjugate variables. Furthermore, the second derivative, d^2e/db^2, reads, $d\lambda/db=1/(db/d\lambda)$. From Eq. (3), the derivative, $db/d\lambda$, is positive-definite. Accordingly, e is a concave function of b. Now, if $T>0$, the Legendre transform instead generates a reduced free energy, $\eta\equiv(e-TS)$, which is a concave function of the constraint value(s). An additional Legendre transform, between T and S, returns e alone, as a concave function of both the constraint(s) and S.

Let b_- and b_+ be the lowest and highest eigenvalues of B, respectively. It is expected that b spans the interval, $[b_-,b_+]$, when λ runs from $-\infty$ to $+\infty$. In every exact diagonalization of $\mathcal{H}$, or exact partition function calculation, concavity sets a one-to-one mapping between b in this interval and λ. More generally, with exact calculations, there is a one-to-one mapping between the set of Lagrange multipliers, $\{\lambda_1,\ldots,\lambda_N\}$, and that of obtained values, $\{b_1,\ldots,b_N\}$, of the constraints. Concavity, *in the whole obtained domain of constraint values*, imposes a very poor landscape: one valley only.

Consider, now, situations where the energy surface does not result from a diagonalization, or the exact calculation of a partition function. Typically, the minimization could be estimated from a Hartree-Fock(-Bogoliubov) approximation, whether at zero or finite T. The trial states involved in such nonlinear approximations do not correspond to a linear manifold; indeed, for instance, a sum of two determinants usually does not make a determinant. Let $\mathcal{D}(\lambda)$ denote one A-body density operator where, within such nonlinear approximations, a minimum of $\mathrm{Tr}\left\{\mathcal{H}\mathcal{D}\right\}$ or, at finite T, a minimum of $(\mathrm{Tr}\left\{\mathcal{H}\mathcal{D}\right\}-TS)$ is reached. Let $\varepsilon(\lambda)$ denote this minimum value. It may be degenerate, incidentally, because of symmetry breaking but, whether degenerate or not, this value is stationary with respect to arbitrary variations $\delta\mathcal{D}$ within the manifold of trial states. As a consequence of this stationarity, the first derivative again reads, $d\varepsilon/d\lambda=-b\equiv-\mathrm{Tr}\left\{B\mathcal{D}\right\}$. Then, if a Legendre transform is possible, defining $\eta\equiv\varepsilon+\lambda b$, in terms of b, the same argument that was used for the exact case again yields, $d\eta/db=\lambda$. With several constraints, the gradient of η in the vector space spanned by $\{b_1,\ldots,b_N\}$ is again the vector $\{\lambda_1,\ldots,\lambda_N\}$.

There remains to discuss second derivatives. Consider, for instance, Hartree-Fock (HF) calculations, where A-body density operators are dyadics of Slater determinants, $\mathcal{D}=\left|\varphi\right\rangle\left\langle\varphi\right|$. Norm-conserving variations of an HF solution, φ, can be parametrized as, $\left|\delta\varphi\right\rangle=\exp(iX\delta\alpha)\left|\varphi\right\rangle$, with X an arbitrary particle-hole Hermitian operator, and $\delta\alpha$ an infinitesimal coefficient. Under such a variation in the neighborhood of an HF solution, the first and second order variations of the energy, $\varepsilon\equiv\mathrm{Tr}\left\{\mathcal{H}\exp(iX\delta\alpha)\mathcal{D}\exp(-iX\delta\alpha)\right\}$, read,

$$\delta\varepsilon=i\delta\alpha\mathrm{Tr}\left\{\left[\mathcal{H},X\right]\mathcal{D}\right\}=0, \tag{4}$$

and

$$\delta^2\varepsilon=-\left(\delta\alpha^2/2\right)\mathrm{Tr}\left\{\left[\left[\mathcal{H},X\right],X\right]\mathcal{D}\right\}\geq0, \tag{5}$$

respectively. Since we are dealing with an HF solution, the first order vanishes $\forall X$, and since only those solutions that give minima are retained, the second order variation of ε is semi-positive-definite, $\forall X$ again. Now, when $\mathcal{H}$ receives the variation, $-Bd\lambda$, there exists a particle-hole operator Y, a special value of X, that, with a coefficient $d\lambda$, accounts for the modification of the HF solution. This reads $\left|\Phi\right\rangle=\exp(iYd\lambda)\left|\varphi\right\rangle$. Simultaneously, those particle-hole operators that refer to this new Slater determinant Φ become $\mathcal{X}=\exp(iYd\lambda)X\exp(-iYd\lambda)$. The new energy is thus,

$$\mathcal{E}=\mathrm{Tr}\left\{\exp(-iYd\lambda)(\mathcal{H}-Bd\lambda)\exp(iYd\lambda)\mathcal{D}\right\}. \tag{6}$$

The stationary condition, Eq. (4), becomes,

$$\begin{aligned}0&=\mathrm{Tr}\left\{\exp(-iYd\lambda)\left[(\mathcal{H}-Bd\lambda),\mathcal{X}\right]\exp(iYd\lambda)\mathcal{D}\right\}\\&=\mathrm{Tr}\left\{\left[\exp(-iYd\lambda)(\mathcal{H}-Bd\lambda)\exp(iYd\lambda),X\right]\mathcal{D}\right\}.\end{aligned} \tag{7}$$

The zeroth order in $d\lambda$ of this, Eq. (7), reads, $\mathrm{Tr}\left\{\left[\mathcal{H},X\right]\mathcal{D}\right\}$, and vanishes $\forall X$, because of Eq. (4). Then the first order in $d\lambda$ gives, again $\forall X$,

$$\mathrm{Tr}\left\{\left[B,X\right]\mathcal{D}\right\}=i\mathrm{Tr}\left\{\left[\left[\mathcal{H},Y\right],X\right]\mathcal{D}\right\}. \tag{8}$$

The second derivative is,

$$\begin{aligned}d^2\varepsilon/d\lambda^2&=-(d/d\lambda)\mathrm{Tr}\left\{\exp(-iYd\lambda)B\exp(iYd\lambda)\mathcal{D}\right\}\\&=-i\mathrm{Tr}\left\{\left[B,Y\right]\mathcal{D}\right\}.\end{aligned} \tag{9}$$

Upon taking advantage of Eq. (8), for Y as a special case of X, this becomes,

$$d^2\varepsilon/d\lambda^2=\mathrm{Tr}\left\{\left[\left[\mathcal{H},Y\right],Y\right]\mathcal{D}\right\}, \tag{10}$$

the right-hand side of which is semi-negative-definite, see Eq. (5). The solution branch obtained when λ runs is, therefore, convex. Its Legendre transform is concave.

Beyond HF, the proof can be generalized as follows. Consider a solution branch $\mathcal{D}(\lambda)$, and expand $\mathcal{D}(\lambda)$ up to second order, assuming that the manifold of solutions is suitably analytic,

$$\mathcal{D}(\lambda+d\lambda)=\mathcal{D}(\lambda)+d\lambda(d\mathcal{D}/d\lambda)+(d\lambda^2/2)(d^2\mathcal{D}/d\lambda^2). \tag{11}$$

Let λ receive a variation $d\lambda$ in $\mathcal{D}(\lambda)$ alone, not in $\mathcal{H}$. The stationarity and minimality of $\mathrm{Tr}\{\mathcal{H}\mathcal{D}\}$ with respect to any variation of $\mathcal{D}$ induce,

$$\mathrm{Tr}\{\mathcal{H}d\mathcal{D}/d\lambda\} = 0,$$
$$\mathrm{Tr}\left\{\mathcal{H}d^2\mathcal{D}/d\lambda^2\right\} \geq 0. \tag{12}$$

There is also a stationary property for the solution, $\mathcal{D}(\lambda + d\lambda)$, but the Hamiltonian is now, $\mathcal{H}(\lambda) - Bd\lambda$, and the derivative of the state becomes, $d\mathcal{D}/d\lambda + d\lambda(d^2\mathcal{D}/d\lambda^2) + \mathcal{O}(d\lambda^2)$, and therefore,

$$0 = \mathrm{Tr}\left\{(\mathcal{H} - Bd\lambda)\left[d\mathcal{D}/d\lambda + d\lambda(d^2\mathcal{D}/d\lambda^2) + \mathcal{O}(d\lambda^2)\right]\right\}. \tag{13}$$

The zeroth order of this, Eq. (13), reads, $\mathrm{Tr}\{\mathcal{H}d\mathcal{D}/d\lambda\}$. It vanishes, because of the first of Eqs. (12). The first order, once divided by $d\lambda$, gives,

$$-\mathrm{Tr}\{Bd\mathcal{D}/d\lambda\} = -\mathrm{Tr}\left\{\mathcal{H}d^2\mathcal{D}/d\lambda^2\right\}. \tag{14}$$

The left-hand side of Eq. (14) is nothing but the the second derivative, $d^2\varepsilon/d\lambda^2$. The right-hand side is semi-negative-definite, because of the second of Eqs. (12). Concavity of $e(b)$ is again a property of such approximations with strict minimizations.

A landscape can be recovered, however, in a somewhat contrived way, if one accepts the crossing of various solution branches. Exact diagonalization normally prevents level crossing, unless distinct quantum numbers are considered, but approximations, usually nonlinear, often generate multiple solutions corresponding to local minima. Special solutions may also occur if, in special cases, variational spaces have edges because the energy gradient does not need to vanish: it just has to be orthogonal to the edge. We assume, without proof, that such exceptional edge solutions occur either only for isolated values of λ or, if they make true functions $\varepsilon(\lambda)$, that they retain enough analyticity for a concavity of the corresponding $e(b)$. All told, various branches of $e(b)$ may intersect, and the pattern formed by their lowest parts is, piecewise, concave. This is schematically illustrated by Figs. 1 and 2. Fig. 1 shows the plots of three convex parabolas, with equations, $\varepsilon = -2-6\lambda-\lambda^2$, $\varepsilon = -1-\lambda^2$, and $\varepsilon = -2+4\lambda-\lambda^2$, respectively. Then Fig. 2 shows, from their Legendre transforms, the plots of $e(b)$, namely $e = 7+3b+b^2/4$, $e = -1+b^2/4$, $e = 2-2b+b^2/4$, respectively. Their lowest segments make a pattern with interesting structures. The physical significance of the spikes, however, is doubtful. A further theory for the rounding of such singularities seems to be in demand. A calculation of collective inertia parameters smoothly through them does not seem to be obvious either. One would have to face the contradiction with the "one valley syndrome" of an exact solution. How could an approximation yield better physics than the exact solution?

One may argue that the zoo of stationary solutions of such approximate methods is usually rich enough to accommodate inflections, where the curvature signature, $\mathrm{Tr}\left\{\mathcal{H}d^2\mathcal{D}/d\lambda^2\right\}$, can change sign. "Phase transitions", a somewhat incorrect wording for a finite system, are sometimes advocated to accept continuing branches of energy minima into metastable branches. But this means dropping the hypothesis of strict energy minimization.

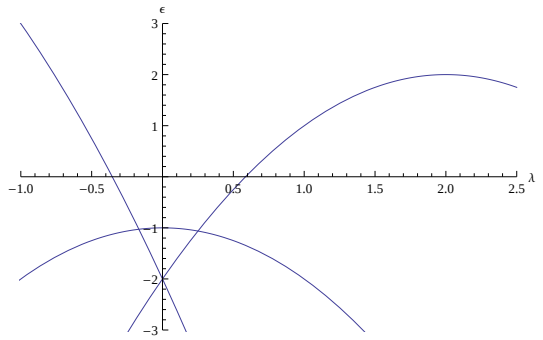

FIG. 1: Toy example with three branches $\varepsilon(\lambda)$, as discussed in the text.

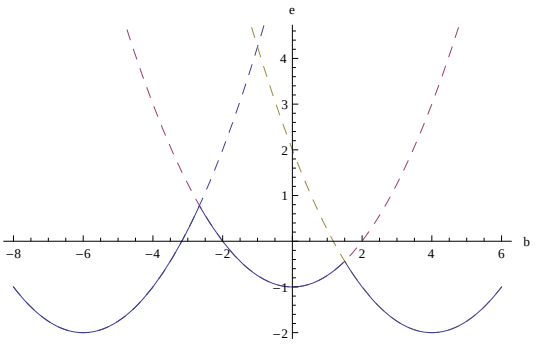

FIG. 2: Legendre transforms, $e(b)$, of the branches in Fig. 1.

For unbounded constraint operators, a further paradox may occur, as seen from the following example. Assume that the interaction v present in H has a strictly finite range, namely $v_{ij} = 0$ if $|\mathrm{r}_i - \mathrm{r}_j| > r_0$. Consider a set of orbitals, $\varphi_1, \ldots, \varphi_A$, with wave functions $\varphi_i(\mathrm{r}, \sigma, \tau) = c_i \exp\left[-1/(r - R_i) - 1/(S_i - r)\right]/r$ if $R_i < r < S_i$, and $\varphi_i = 0$ otherwise. Here $r \equiv |\mathrm{r}|$, and c_i is the normalization coefficient. Given a large positive number L, set $S_1 = R_1 + L, \ldots, S_A = R_A + L$, so that such orbitals are very flat, and, therefore, carry a minuscule kinetic energy, $\propto 1/L^2$, as small as wanted. Set, also, $R_2 \geq S_1 + r_0, \ldots, R_A \geq S_{A-1} + r_0$, so that such orbitals are so separated that, not only their orthogonality is trivial, but also any antisymmetrized interaction matrix element, $\langle\varphi_i\varphi_j\,|v|\,\varphi_k\varphi_\ell\rangle$, vanishes. The Slater determinant χ made of such orbitals has an energy, $\langle\chi\,|H|\,\chi\rangle$, as close to zero as wanted. Simultaneously, for a collective operator such as, $B_\nu = \sum_{i=1}^{A} r_i^\nu$, with, typically, $\nu \geq 2$, it is trivial to show that $AR_1^\nu < \langle\chi\,|B_\nu|\,\chi\rangle < AS_A^\nu$. The situation is illustrated in Fig. 3 for a chosen set of parameters. Now, given the ground state Ψ_0 of H, with its eigenvalue

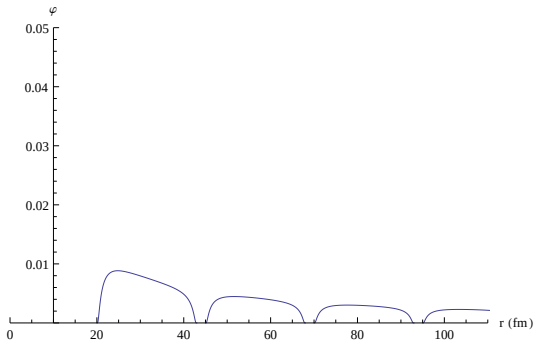

FIG. 3: Example, with $R_1 = 20$ fm, $L = 23$ fm and $r_0 = 2$ fm, see text, of orbitals φ_i for a Slater determinant χ that, if mixed with the ground state, can leave the energy hardly changed while allowing a very large collective moment. True ground state nuclear densities, halos and tails included, occupy the first 10 fm at the very most.

E_0, define an A-body state by the normalized mixture of density matrices, $\mathcal{D}_{mix} = (|\Psi_0\rangle\langle\Psi_0| + \omega|\chi\rangle\langle\chi|)/(1+\omega)$, where the positive number ω will eventually become infinitesimal. When $L \to +\infty$, the energy of this mixture can be made as close to $E_0/(1+\omega)$ as wanted. Simultaneously, if $R_1 \to +\infty$, the expectation value, $b_\nu \equiv \mathrm{Tr}\{B_\nu \mathcal{D}_{mix}\}$, of the collective operator diverges at least as fast as $\omega A R_1^\nu/(1+\omega)$. Maintain constant the product, ωR_1, and the ratio, L/R_1, while R_1 grows to infinity. It is clear that the limit of the energy is E_0, while b_ν can grow arbitrarily, like $\propto R_1^{\nu-1}$ at least. This makes the right-hand side of the energy surface, $e(b_\nu)$, flat.

The paradox still holds with interactions v that do not strictly vanish beyond r_0. For any tail of v, the separations, $R_2 - S_1, \ldots, R_A - S_{A-1}$ can be adjusted to make the matrix elements, $\langle\varphi_i\varphi_j|v|\varphi_k\varphi_\ell\rangle$, as small as wanted.

For traditional nuclear multipole operators, a spherical harmonic multiplies the radial factor, r^ν. We insert the same harmonic in the φ's to extend the paradox; an arbitrary value of a multipole is compatible with the ground state energy. A choice of collective operators that are bounded seems to be mandatory if the concept of a non-flat landscape for a collective energy has to be salvaged.

Even under such a precaution, this work shows that concavity, piecewise at least, is a major property of any energy surface obtained by a strict minimization of the energy under constraint(s). If an energy landscape with "mountains" and "saddles" is needed, this concavity contradicts the intuition that energy transits through local minima. The success of collective models that use a non trivial landscape is too strong to be rejected as physically and/or mathematically unsound, but its validation must now be searched through other methods such as, likely, resonating group methods [5], generator coordinate ones [6], Born-Oppenheimer approximations, influence functionals [7], deconvolutions of wave packets in collective coordinate spaces [8], etc. Except for anharmonic vibrations, where one valley is sufficient, one might have to accept that metastability may be as important as energy minimization. Or could it happen that collective operators have to be λ-dependent, in order to convert metastable branches for constant operators into strict minimization branches for λ-dependent ones? Since fluctuations are important at "phase transitions", combinations of the form, $\mathcal{K} = H - \lambda B - f(\lambda)B^2$, with at first very simple forms of $f(\lambda)$, deserve some attention. Such operators $\mathcal{K}$ govern both a constraint and its fluctuation, but obviously differ from a double constraint form, $\mathcal{H} = H - \lambda_1 B - \lambda_2 B^2$, with two independent parameters, λ_1, λ_2. A one-dimensional path with non trivial structures, because of an additional term, $-d^2 f/d\lambda^2 \left\langle (B + B^2 df/d\lambda)\right\rangle$, in the second derivative, $d^2\varepsilon/d\lambda^2$, might be induced by $\mathcal{K}$ inside that more trivial two-dimensional landscape corresponding to $\mathcal{H}$. It must be kept in mind that H is often not very well suited to ensure a localisation of B's, a necessary condition for the exploration of an energy surface parametrized by b's. A review of those landscapes obtained by constrained Hartree-Fock and/or Hartree-Bogoliubov, whether at zero or finite temperature, is in order, if only to analyze the role of fluctuations and also verify whether edges of variational spaces, and additional solutions, trapped by such edges, might exist.

SK acknowledges support from the National Research Foundation of South Africa. BG thanks the University of Johannesburg for its hospitality during part of this work.

* Electronic address: bertrand.giraud@cea.fr
† Electronic address: stevenka@uj.ac.za

[1] A. Bohr and B. Mottelson, *Nuclear Structure* (World Scientific, 1998).
[2] J. R. Nix and W. J. Swiatecki, Nucl. Phys. **71**, 1 (1964) .
[3] B. R. Barrett, B. G. Giraud, B. K. Jennings, and N. Toberg, Nucl. Phys. **A828**, 267 (2009).
[4] R. Balian, *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics* (Springer, Berlin, 2007).
[5] K. Wildemuth and Y. C. Tang, *A Unified Theory of the Nucleus* (Academic, New York, 1977).
[6] L. Hill and J.A. Wheeler, Phys. Rev. **89**, 112 (1953); J. J. Griffin and J.A. Wheeler, Phys. Rev. **108**, 311 (1957).
[7] R. P. Feynman and F. L. Vernon, Jr., Ann. Phys. (N.Y.) **24**, 118 (1963).
[8] B. Giraud and B. Grammaticos, Nucl. Phys. **A233**, 373 (1974).