NUMERICAL QUANTUM FIELD THEORY USING THE SOURCE GALERKIN METHOD*

G.S. Guralnik

Physics Department
Brown University
Providence RI 02912

INTRODUCTION

This talk describes a program, still in its early stages, to develop a new numerical approach that we call the Source Galerkin method [1]. This method has little in common with the Monte Carlo methods which have been successful in demonstrating that numerical techniques can be an important tool for obtaining results from quantum field theories. Source Galerkin appears to have strengths which will allow its use for problems (such as scattering predictions) that are not readily accessible to conventional methods with the currently available computers. It is, unlike Monte Carlo, equally suited for studying fermionic or bosonic problems. Further, at least in lower dimensions, a continuum formulation of Source Galerkin is straightforward. While we have made considerable progress in understanding the issues in defining and applying these ideas, we can not yet been able to complete enough calculations to evaluate the chances that Source Galerkin could actually replace Monte Carlo techniques in any of their traditional applications.

The basic idea of Source Galerkin is to formulate a field theory in the presence of external sources and then to construct approximate solutions to the linear source functional differential equations for the vacuum-vacuum amplitude Z of the theory. These solutions are constructed by (enlightened) guessing. It is possible to utilize

* Supported in part by DOE Grant DE-FG09-91-ER-40588—Task D

internal (charge, statistical) and external (spatial) symmetries expected of the exact answer in the construction of the approximate solutions. This possibility, is one of the great advantages of our method. In contrast, it is difficult or impossible to insert knowledge obtained from analytic techniques into the construction of solutions using Monte Carlo techniques.

A straightforward way to make a Galerkin guess is to pick a linear combination of a subset of a complete set of functions of the source variables appropriate for the system being studied. This choice has the advantage of having a clear iteration path defined by extending the number of functions from the complete set used in successive approximations. This guess will not satisfy the original equations.The deviation from the exact solution will be characterized by the numerical coefficients multiplying each member of the complete set contained in the guess. The idea of Galerkin is to fix these coefficients by adjusting them so that the guess is true as an appropriately defined weighted average over some region of source space. The optimal definition of average is part of the problem of defining a rapidly convergent approximation scheme. Such a scheme allows the reduction of the error between the approximate solution and the exact answer as the number of iterations is increased. Galerkin is a weighted residue procedure. These methods guarantee convergence in a mean. That is, the integral of the square of the local error vanishes in an appropriate limit. In this context, it should be noted that the more familiar variational Hamiltonian approximation techniques are in this same general class of approximation. However, the Source Galerkin methods do not have a direct correspondence to such methods as they have been used in physics.

EXAMPLE OF GALERKIN METHOD:

Before we examine a field theory, we outline the application of Galerkin to a generic linear differential equation

$$L\, f = 0$$

in some domain $D\left(x_1, x_2 - x_N\right)$ with boundary conditions on a surface $S\left(x_1 - x_N\right)$. Guess a solution:

$$f^* = \varphi_0 (x_1 - x_N) + \sum_{j=1}^{M} a_j \varphi_j (x_1 - x_N) .$$

Assume the φ_ℓ are linearly independent and that they satisfy the appropriate boundary conditions but are otherwise unrestricted.

Now make the definition

$$\begin{aligned} R(a_1, a_2 - a_M; x_1 - s_N) &\equiv L(f^*) \\ &= L(\varphi_0) + \sum_{j=1}^{M} a_j L(\varphi_j) \end{aligned} .$$

Next, pick M linearly independent test functions $\rho_i (x_1 - x_N)$. We do not confine our considerations to the usual Galerkin choice of picking $\rho_i = \varphi_i$. The coefficients $a_1 - a_M$ are fixed by requiring that

$$\langle R, \rho_i \rangle = 0 \quad i = 1, 2 \cdots M.$$

We define this generalized dot product as follows:

$$\langle R, \rho_i \rangle \equiv \int dx_1 - dx_N \; R \, \rho_i d (x_1 - x_N)$$

To make sure that this is sensible we choose d so it dampens integrals. Options that we have used in field theory problems include restricting d to be composed of δ functions, so as to confine the integral to a finite volume or choosing it to be of the form $\rho^{-\sum_{i,j} x_i A_{ij} x_j}$.

After choosing d and performing the above integrations, we obtain M linear equations in M unknowns. If these are independent, we can solve for all of the unknowns. If not, we must change our choice of test functions or increase their number. Under very general circumstances, these approximations converge (weakly) to the exact answer as $M \to \infty$. Obviously the convergence is best if the φ_i form a complete set.

SCALAR FIELD THEORY

Now we can look at the example of Euclidean quartic field theory coupled to a source $J(x)$ for the scalar field [1-3]. This source will be the variable used in our Galerkin solution and hence the name Source Galerkin. The action is given by

$$S_E = \int d^4x \left[\varphi(x)\left[-\Delta + m^2\right]\varphi(x) + \frac{g}{4}\varphi^4(x) - J(x)\rho(x)\right]$$

which leads to the field equations

$$\left(-\Delta + m^2\right)\varphi(x) + g\varphi(x)^3 - J(x) = 0\ .$$

With the usual definition $Z = \langle 0|0\rangle$, it follows that

$$\left(-\Delta + m^2\right)\frac{\delta}{\delta J}Z + g\left(\frac{\delta}{\delta J}\right)^3 Z - JZ = 0.$$

The Galerkin method appears to be a self consistent method on the continuum because of the regularization properties that can be directly built into the test functions and guess solutions. However, for the very simple considerations of this paper we will confine our examples to the lattice. On the lattice we have

$$\left(-\Delta + m^2\right) \equiv \left(2D + m^2\right)\delta_{x,y} - \sum_a \delta_{y,x+\hat{u}}.$$

The $\hat{u}$ sum is over nearest neighbors and the lattice spacing a is absorbed. Consequently, the lattice equations for this model are

$$\left(-\Delta + m^2\right)_{x,y}\frac{\partial Z}{\partial J_y} + g\frac{\partial^3 Z}{\partial J_x^3} - J_x Z = 0.$$

Note, this is a very large set of coupled differential equations, since there is one equation for each point on the lattice. For the transition to a lattice to be a sensible analogue to the continuum equations, we must define finite lattices at their boundary points. In what follows, periodic boundary conditions will always be implicit.

In order to solve this set of equations, we make an initial Galerkin guess. An uninspired (but, nevertheless, interesting) way to proceed is to just try a Taylor series expansion. This has the advantage of being an expansion in a complete set of functions.

$$Z\left(J_1 - J_N\right) = \sum_{m_1, n_2 - n_N} G_{m_1 n_2 - n_N} \, J_1^{n_1} J_2^{n_2} - J_N^{n_N}$$

Substituting this expansion into the differential equations above results in complicated sets of recursion relations particularly for a large lattice. Realistically, the appropriate application of boundary conditions and the size of the sets involved makes direct solution impossible.

One alternative way to proceed is to truncate the series at some maximal power of the sources. Truncated series are definitely not solutions to our equations. It is reasonable to hope that for small sources that the contributions from higher terms in the expansion decrease rapidly. In any event, we can minimize the error from truncation if we determine the coefficients of the Taylor expansion by a Galerkin procedure in the source co-ordinates. This means using the definition of Galerkin given above with the identification $\int dx_1 - dx_N \rightarrow \int dJ_1 \cdots dJ_N$. In low source dimensions this is using excessive force for the outcome, but as the dimension increases the use of this method to guarantee consistency is not trivial.

To solve the truncated equations we use lattice symmetry (translation invariance). If, for example, we examine a 3 site lattice in 1 dimension keeping terms to quartic polynomial order in lattice invariant polynomials the test solution is

$$Z_4^* = Z_4^*\left(J_1, J_2 J_3\right) = 1 + a_0^2 P_0^2(J)^\dagger a_1^2 P_1^2(J) + a_0^4 P_0^4(J) + \cdots a_3^4 P_3^4(J).$$

Here, we have introduce the definitions:

$$\begin{aligned} P_0^2(J) &= J_1^2 + J_2^2 + J_3^2 \\ P_1^2(J) &= J_1 J_2 + J_2 J_3 + J_3 J_1 \\ P_0^4(J) &= J_1^4 + J_2^4 + J_3^4 \end{aligned}$$

Altogether there are six independent P_m^n to this order . We Galerkin the results of inserting this guess into the lattice equations by using

$$T_m^n \equiv \frac{\partial P_m^n}{\partial J_x}$$

as test functions. By lattice symmetry. all lattice points are equivalent. Consequently, we need only examine equations for one lattice site. This problem can easily be solved using a computer algebra program (we use Aljabr which is very robust and powerful for this type of problem). We can easily compare our method with Monte Carlo Methods on one dimensional systems without using any substantial amount of computer time. In the following table we calculate the two point function as we increase the degree of the polynomial and compare the results with Monte Carlo. We also include the results of a Shanks interpolation of the Galerkin results. The coupling is relatively small but even for much larger coupling the results remain good, despite the small number of polynomials used. This reflects the fact that the Galerkin procedure forces the guess to be reasonable in the average we have defined. The computer time required for the Galerkin solution is orders of magnitude less than that required for the Monte Carlo of comparable accuracy.

Table 1. 1-dimensional two-point function: 11 site lattice; $M = 1$, $g = 0.5$

$i - j$	J^4	J^6	J^8	Shanks	MC
0	0.326538	0.360178	0.347223	0.35082	0.3510±0.0005
1	0.094508	0.117673	0.108121	0.11091	0.1110±0.0004
2	0.026768	0.038823	0.033422	0.03509	0.0351±0.0003
3	0.007454	0.012940	0.010269	0.01114	0.0112±0.0004
4	0.002075	0.004461	0.003197	0.00363	0.0036±0.0003
5	0.000694	0.001922	0.001218	0.00147	0.0011±0.0004

For the purposes of this talk, I have kept the examples very simple. We can calculate much bigger systems than this tiny example even in 4 dimensions. However, the number of representatives of an invariance class of polynomials on the lattice grows exponentially. Unless a low order polynomial gives a good approximation (which happens in the example), simple Taylor series are not a great starting point and we should pick other expansion functions which more closely approximate the actual answer. While, we do not usually know what this is, we do know a lot about the structure of an exact

answer from symmetry and the spectral representations of quantum field theory. Consequently, it is easy to make a better guess.

NON–LINEAR METHODS

Our initial examinations of the Galerkin method concentrated on expansions that were simple for Z. However, because of the disconnected Greens functions content of Z it is often much better to deal with $\ln Z$. The difficulty with this is that the resulting equations are non-linear in the Galerkin parameters. For example, from the previous equations, it is straightforward to show that the approximation for a solution to quartic scalar field theory of the form

$$Z^* = e^{J_i A_{ij} J_j + J_1 J_m A_{pm_{pq}} J_p J_q + \cdots}$$

yields non-linear equations after being acted on by the source differential equations and then Galerkin averaged. These can be handled numerically. It is possible to develop iteration schemes based on the above guess which keep the problems associated with non-linearity under control.

An effective way to construct leading approximations is to build Z in terms of spectral functions (propagators). A good combination can be guessed by writing down the perturbation theory or an effective theory for the action being studied. However, the only information we use from these structures is the general spectral forms. The masses and couplings are left arbitrary and set by Galerkin. Another effective and rapidly convergent guess can be made using polynomials weighted with propagators. There are a huge number of reasonable choices. In practice, we have found remarkable stability (per computational time) between very different types of iterations schemes when used on simple problems.

As a special case of the above non-linear form, we have been examining a spectral iteration scheme that we would think has much of the structure of an actual solution. The initial ansatz is the form

$$\log \mathcal{Z} = \int dw\, dx \frac{1}{2} J(w) G_2^*(w,x) J(x) + \int dw\, dx\, dy\, dz\, \frac{1}{4} J(w) J(x) G_4^*(w,x,y,z) J(y) J(z)$$

with

$$\tilde{G}_2^*(p) = \frac{1}{p^2 + \mu^2}$$

and

$$\tilde{G}_4^*(p,q,r) = \tilde{G}_2^*(p)\tilde{G}_2^*(q)\frac{B}{(p+q)^2 + \mu'^2}\tilde{G}_2^*(r)\tilde{G}_2^*(p+q-r).$$

We are generating results based on this starting form in 4 dimensions on fairly big systems. We are finding that this method consumes small amounts of CPU (relatively) but at the cost of large amounts of memory.

BOUNDARY CONDITIONS

The discussion up to this point has avoided a fundamental issue. One of the most complicated and interesting problems associated with this type of numerical solution lies in the analysis of the appropriate boundary conditions imposed on the source differential equations. If we look at the special case of zero space time dimensions (ultralocal model) there is only one differential equation. It is a third order differential equation so its solution involves the specification of three a-priori arbitrary constants. In particular, the normalization of Z, and the one and two field Green's functions are unspecified. The situation for arbitrary space-time dimension is more complex with a correspondingly greater number of unspecified constants. This situation is not entirely unfamiliar. In this discussion we have excluded terms involving odd powers of the sources. This constitutes a boundary condition on the first derivative of Z. We know that, if we do not exclude odd terms, we can under the correct conditions write solutions that have spontaneous symmetry breaking. Dealing with the boundary condition on the second derivative terms is a much more complicated and a very interesting problem

that leads to an enhanced understanding of phase structure and the possibility of solutions to quantum field theory which are extremely singular for small bare couplings. We analyze this elsewhere [5]. I have avoided direct contact with this problem in the example detailed here by truncating the Taylor series solution in Z. It is easy to see that truncation forces the approximation to agree with the smooth $g = 0$ limit of the theory that corresponds to perturbation theory. Therefore, truncation imposes an implicit boundary condition on the second derivative terms.

FERMIONS

An interesting feature of this scheme is that fermions can be included in a natural manner. Fermions present difficulty in Monte Carlo approaches for two reasons. The first is because of the "fermion determinant". The problem occurs because after an approximate bosonic configuration is generated, the fermionic information contained in a determinant involving every point of the lattice system must be evaluated. This evaluation requires a large amount of computer time and it must be done for many bosonic configurations. There are of course many schemes to decrease the amount of computation this requires. However, the difficulty is intrinsic to how path integration forces us to deal with fermions. Source Galerkin avoids this difficulty because it allows making approximations for fermions without first solving a bosonic problem. I will outline how this is done in the following. The second problem is the well known "fermion doubling problem". This difficulty has nothing to do with numerical methods, but is an artifact of lattice formulation. As mentioned, it is possible that we can avoid this as well. We have worked several fermionic problems in low dimensions on the continuum. This has been easy because the divergences we dealt with were trivial. We are currently working on a "proof" that Galerkin has enough flexibility to regularize all calculations without being forced to resort to a lattice.

The key difficulty in developing a Source Galerkin procedure for fermions is that the naive inner product definition is not appropriate because of the Grassmann nature of the sources.

Direct integration of Grassmann polynomials is defined by

$$\int dc = 0, \qquad \int dc\ c = 1.$$

Fermionic polynomials are decimated if we base the inner product on this type of integration. Instead we define a similar inner product to that developed for bosons. For Majorana fermions, we take this to be

$$\langle A[\eta], B[\eta] \rangle \equiv \int [d\eta] \exp[-\eta\beta\mathbf{q}\eta] A[\eta] B[\eta].$$

Since det $\beta\mathbf{q} = 1$, this leads us to the normal Gaussian integration, which in turn gives us the Wick contraction in η. The formulation for Dirac fermions is

$$\langle A[\bar{\eta}, \eta], B[\bar{\eta}, \eta] \rangle \equiv \int [d\bar{\eta}][d\eta] \exp[-\bar{\eta}\eta] A[\bar{\eta}, \eta] B[\bar{\eta}, \eta] \, .$$

To test these definitions, we have studied simple quartic fermion interacting systems on 4 sites (one dimension) and compared our solutions to the exact result [3]. Convergence is rapidly achieved by the Galerkin averaging.

SUMMARY

Previous published work in developing the Source Galerkin method has been focussed primarily on free scalars and ϕ^4, in various spacetime dimensions [1–3]. Current work targets robust methods for gauge and fermionic theories. We are also attempting to improve the convergence rate of our approximations by using the full Schwinger action principle (this is somewhat like imposing renormalization group restrictions). We have tested these ideas on the Thirring and Schwinger models and have shown that nothing in principle precludes working with models with many degrees of freedom and gauge structure. In these simple solvable models convergence to the exact answer is almost instant. We have looked at leading order in Gross Neveu model and find results which correspond to the leading order $\frac{1}{N}$. As we develop our techniques we anticipate to be able to answer questions about more realistic theories. At this stage, I am very optimistic that the Source Galerkin model will be important in making a contribution to our understanding of the solutions of quantum field theory.

Acknowledgement

Results presented here were obtained in collaborations with S. García (IBM), Z. Guralnik (Princeton), S. Hahn (Brown), J. Lawson (ICTP Trieste) and K. Platt (Brown).

REFERENCES

1. "A New Approach to Numerical Quantum Field Theory", (with S. Garcia and J. Lawson),*Phys. Lett.* **B22**, 119 (1994).
2. "New Numerical Method for Fermion Field Theory (with John Lawson),to be published (1995).
3. "Source Galerkin Calculations in Scalar Field Theory (with John Lawson),to be published (1995).
4. "Multiple Solutions in Quantum Field Theory (with Santiago Garcia),to be published (1995)
5. " Boundary conditions of the Schwinger Dyson equations and multiple solutions of Quantum Field Theory" (with Santiago Garcia and Z. Guralnik), Brown Preprint (1996).