Statistical errors in stochastic perturbation theory

Martin Lüscher July 2015

1. Introduction

The renormalizability of the Langevin equation [1,2] implies that the autocorrelation times in numerical stochastic perturbation theory [3–5] grow approximately like $1/a^2$ as the lattice spacing a is taken to zero. Since the variances of the calculated coefficients cannot, in general, be expressed through the expansion coefficients of physical observables, the renormalizability of the Langevin equation alone however does not allow to predict the scaling behaviour of the statistical errors.

Numerical experiments suggest that the variances of physical quantities grow only slowly towards the continuum limit (see ref. [6], for example). The goal in this note is to show that, up to any fixed order in perturbation theory, the variances of the fields evolved by the Langevin equation do in fact increase at most logarithmically.

2. Stochastic perturbation theory

The theoretical analysis presented in this note is expected to apply to any renormalizable theory in four space-time dimensions. For simplicity, and in order to bring out the essence of the argument most clearly, the theory considered is taken to be the one of a real scalar lattice field $\phi(x)$ with action

$$S = a^4 \sum_x \left\{ \frac{1}{2}\partial_\mu\phi(x)\partial_\mu\phi(x) + \frac{1}{2}(m^2+\delta m^2)\phi(x)^2 + \frac{g_0}{4!}\phi(x)^4 \right\}. \tag{2.1}$$

Here g_0 denotes the bare coupling, m the additively renormalized mass parameter and

$$\delta m^2 = \sum_{k=1}^{\infty} (\delta m^2)^{(k)} g_0^k \tag{2.2}$$

the (quadratically divergent) additive mass counterterm. The operators ∂_μ are the forward nearest-neighbour lattice derivatives.

A suitable normalization condition for the renormalized mass is assumed, but does not need to be specified. Since the focus in this note is exclusively on the cancellation of the power divergences, logarithmic divergences and the associated counterterms will not be discussed. It is thus understood that the lattice regularization is never actually removed.

2.1 Langevin equation

In stochastic perturbation theory, the field $\phi(x)$ is evolved as a function of the Langevin time t according to

$$\partial_t \phi = (\Delta - m^2 - \delta m^2)\phi - \frac{g_0}{3!}\phi^3 + \eta, \qquad \Delta = \sum_{\mu=0}^{3} \partial_\mu^* \partial_\mu, \tag{2.3}$$

where η is a Gaussian random field with mean zero and variance

$$\langle \eta(t,x)\eta(s,y)\rangle = 2a^{-4}\delta_{xy}\delta(t-s). \tag{2.4}$$

The equation is solved in powers of the coupling by substituting

$$\phi = \sum_{k=0}^{\infty} g_0^k \phi_k \tag{2.5}$$

and equating the terms of equal order in g_0. As a result a tower of equations

$$\partial_t \phi_0 = (\Delta - m^2)\phi_0 + \eta, \tag{2.6}$$

$$\partial_t \phi_1 = (\Delta - m^2)\phi_1 - (\delta m^2)^{(1)}\phi_0 - \frac{1}{3!}\phi_0^3, \tag{2.7}$$

$$\partial_t \phi_2 = (\Delta - m^2)\phi_2 - (\delta m^2)^{(2)}\phi_0 - (\delta m^2)^{(1)}\phi_1 - \frac{1}{2!}\phi_0^2\phi_1, \tag{2.8}$$

etc., is obtained.

In order to simplify the notation, it is helpful to introduce the operator

$$\mathcal{D} = \partial_t - \Delta + m^2 \tag{2.9}$$

and the interaction term

$$\mathcal{R}_k = -\sum_{j=0}^{k-1} (\delta m^2)^{(k-j)} \phi_j - \frac{1}{3!} \sum_{j_1,j_2,j_3=0}^{k-1} \delta_{k,j_1+j_2+j_3+1} \phi_{j_1} \phi_{j_2} \phi_{j_3}. \tag{2.10}$$

The tower of equations then assumes the form

$$\mathcal{D}\phi_0 = \eta, \tag{2.11}$$

$$\mathcal{D}\phi_k = \mathcal{R}_k \quad \text{for all} \quad k \geq 1. \tag{2.12}$$

Since the terms on the right of these equations depend only on the lower-order fields, the equations can be solved recursively.

2.2 Frequency-momentum space

Passing to frequency-momentum space,

$$\tilde{\phi}_k(\omega, p) = \int_{-\infty}^{\infty} \mathrm{d}t \, a^4 \sum_x \mathrm{e}^{i\omega t - ipx} \phi_k(t, x), \tag{2.13}$$

the inverse of the operator $\mathcal{D}$ is given by the kernel

$$\tilde{K}(\omega, p) = \{\hat{p}^2 + m^2 - i\omega\}^{-1}, \tag{2.14}$$

where $\hat{p}_\mu = \frac{2}{a}\sin(\frac{1}{2}ap_\mu)$ as usual. For the solutions of equations (2.11),(2.12), the expressions

$$\tilde{\phi}_0(\omega, p) = \tilde{K}(\omega, p)\tilde{\eta}(\omega, p), \tag{2.15}$$

$$\tilde{\phi}_k(\omega, p) = \tilde{K}(\omega, p)\tilde{\mathcal{R}}_k(\omega, p), \quad k \geq 1, \tag{2.16}$$

are thus obtained.

The kernel $\tilde{K}(\omega, p)$ is non-singular and so is the two-point function

$$\langle \tilde{\phi}_0(\omega, p)\tilde{\phi}_0(\nu, q)\rangle = (2\pi)^5 \delta(\omega + \nu)\delta_P(p + q)\tilde{G}(\omega, p), \tag{2.17}$$

$$\tilde{G}(\omega, p) = 2\tilde{K}(\omega, p)\tilde{K}(-\omega, -p) = 2\{(\hat{p}^2 + m^2)^2 + \omega^2\}^{-1}. \tag{2.18}$$

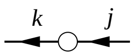
Here and below, the bracket $\langle \ldots \rangle$ stands for the average over the noise field η and

$$\delta_P(p) = \sum_{n\in\mathbb{Z}^4} \delta(p + 2\pi n/a) \tag{2.19}$$

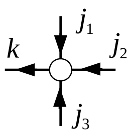
for the periodic δ-function.

2.3 Feynman diagrams

The fields $\tilde{\phi}_l$, $l = 0, 1, 2, \ldots$, can be represented through rooted tree diagrams with the random field $\tilde{\eta}$ attached to their leaves. In these diagrams, the vertices

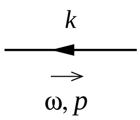


$$= \begin{cases} -(\delta m^2)^{(k-j)} & \text{if } k > j, \\ 0 & \text{otherwise,} \end{cases} \tag{2.20}$$



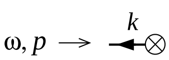
$$= -\delta_{k,j_1+j_2+j_3+1}, \tag{2.21}$$

correspond to the interaction terms in eq. (2.10), while the lines



$$= \tilde{K}(\omega, p), \tag{2.22}$$

represent the propagation of the fields ϕ_k. The random sources at the leaves,



$$= \delta_{k0}\tilde{\eta}(\omega, p), \tag{2.23}$$

are represented by a crossed circle. Since the random fields only occur in the leading-order equation (2.11), the attached line must have label $k = 0$. Examples of diagrams

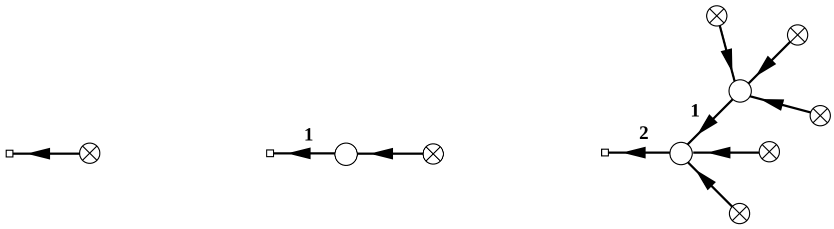


Fig. 1. Feynman diagrams contributing to $\tilde{\phi}_0(\omega, p)$, $\tilde{\phi}_1(\omega, p)$ and $\tilde{\phi}_2(\omega, p)$, respectively. For simplicity, the line label is suppressed if it is equal to 0. The frequency-momentum (ω, p) flows into the diagrams at the external line ending in a little square (see subsect. 2.3 for the values of the propagators and vertices).

are shown in fig. 1. Clearly, for a diagram to be non-zero, the line labels must be such that the expressions (2.20) and (2.21) for the vertices do not vanish.

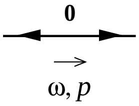
All lines in these diagrams carry a frequency-momentum. There is a frequency-momentum conservation δ-function at each vertex and the frequency-momenta assigned to the internal lines (including the ones from the sources to the vertices) are integrated over. As usual, the contributions of the diagrams must be divided by their symmetry factors. The field $\tilde{\phi}_k$ is given by the sum of all tree diagrams whose root line (the one ending in the little square) has label k.

2.4 Correlation functions

The n-point functions

$$\langle \tilde{\phi}_{k_1}(\omega_1, p_1) \ldots \tilde{\phi}_{k_n}(\omega_n, p_n) \rangle \tag{2.24}$$

are obtained by contracting the random source fields at the leaves of the tree diagrams contributing to the fields $\tilde{\phi}_{k_j}(\omega_j, p_j)$. Each contraction combines the directed lines attached to the sources and yields a bidirected line



$$= \tilde{G}(\omega, p) \tag{2.25}$$

with label 0. The label may be omitted since there are no lines of this type with non-zero label. Since the random fields all get contracted, the diagrams contributing to the correlation functions (2.24) have no directed lines with label 0, i.e. all lines propagating the field ϕ_0 become bidirected ones. The Feynman rules are otherwise the same as before.

3. Problem description

The renormalizability of the theory and the Langevin dynamics [1,2] guarantees that the fixed-order linear combinations

$$\sum_{k_1,\ldots,k_n=0}^{\infty} \delta_{k_1+\ldots+k_n,k} \langle \tilde{\phi}_{k_1}(\omega_1,p_1)\ldots\tilde{\phi}_{k_n}(\omega_n,p_n)\rangle, \qquad k=0,1,2,\ldots, \tag{3.1}$$

of the correlation functions (2.24) are at most logarithmically divergent, i.e. the mass counterterm in the action (2.1) removes all power divergences. However, while the variances of the fixed-order combinations of fields are again given by the correlation functions (2.24), they cannot, in general, be expressed through the complete fixed-order combinations (3.1). The statistical errors of any observable computed in stochastic perturbation theory are therefore not obviously free of power-divergent contributions†.

The aim in the present note is to prove that all correlation functions (2.24) are in fact only logarithmically divergent in the continuum limit. First the structure of the possible power divergences is determined by representing the correlation functions through a functional integral in 4+1 dimensions. Using the fact that the fixed-order correlation functions (3.1) are known to be free of power divergences, the coefficients of the power-divergent counterterms (other than δm^2) are then shown to vanish.

4. Field theory in 4+1 dimensions

In the following, the theory is considered up to some fixed order n in the coupling. Equations (2.11),(2.12) may then be regarded as a stochastic equation for a multiplet of $n+1$ fields. Following ref. [1], the associated field theory in $4+1$ dimensions is introduced, which allows the correlation functions (2.24) to be studied using standard tools.

† Beyond the lowest few orders in the coupling, the statistical errors in instantaneous stochastic perturbation theory [7], for example, appear to grow like a power of the lattice spacing [6].

4.1 Fields and action

The expectation value of any product $\mathcal{O}[\phi]$ of the fields $\phi_0, \ldots, \phi_n$ is given by

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathrm{D}[\eta]\mathrm{D}[\phi_0] \ldots \mathrm{D}[\phi_n]\, \mathrm{e}^{-S_\eta}$$

$$\times\, \delta(\mathcal{D}\phi_0 - \eta)\delta(\mathcal{D}\phi_1 - \mathcal{R}_1) \ldots \delta(\mathcal{D}\phi_n - \mathcal{R}_n)\mathcal{O}[\phi], \tag{4.1}$$

where

$$S_\eta = \frac{1}{4} \int \mathrm{d}t\, a^4 \sum_x \eta^2. \tag{4.2}$$

Up to a constant factor, which is canceled by the normalization factor $\mathcal{Z}$, the functional integral (4.1) yields the expectation value of $\mathcal{O}[\phi]$ as determined by the stochastic equations (2.11),(2.12) and the average over the random field. The functional integral representation is thus completely equivalent to the stochastic equations.

The Dirac δ-functions in eq.(4.1) may now be eliminated by introducing a set of *purely imaginary* Lagrange multiplier fields $L_0(t, x), \ldots, L_n(t, x)$. After integrating out the random field, the functional integral then becomes

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathrm{D}[\phi_0] \ldots \mathrm{D}[\phi_n]\mathrm{D}[L_0] \ldots \mathrm{D}[L_n]\, \mathrm{e}^{-\hat{S}}\mathcal{O}[\phi], \tag{4.3}$$

$$\hat{S} = \int \mathrm{d}t\, a^4 \sum_x \left\{ L_0(\mathcal{D}\phi_0 - L_0) + \sum_{k=1}^{n} L_k(\mathcal{D}\phi_k - \mathcal{R}_k) \right\}. \tag{4.4}$$

It is understood here that an infinitesimal term proportional to $\sum_{k=1}^{n} L_k^2$ is added to the action so as to guarantee the absolute convergence of the integral.

4.2 Perturbation expansion

The functional integral (4.3) may be expanded in powers of the interaction terms $\mathcal{R}_k$. As will become clear in a moment, the expansion effectively ends after a finite number of terms, since all higher-order terms vanish identically.

To lowest order, the action $\hat{S}$ reduces to a non-degenerate quadratic form in the fields $\phi_0, \ldots, L_n$. The theory is thus Gaussian at this order, the non-zero two-point functions being given by

$$\langle\phi_k(t,x)L_j(s,y)\rangle_{\rm LO} = \delta_{kj}\int_{\omega,p} {\rm e}^{i\omega(t-s)-ip(x-y)}\tilde{K}(\omega,p), \quad k,j=0,\ldots,n, \tag{4.5}$$

$$\langle\phi_0(t,x)\phi_0(s,y)\rangle_{\rm LO} = \int_{\omega,p} {\rm e}^{i\omega(t-s)-ip(x-y)}\tilde{G}(\omega,p). \tag{4.6}$$

These two-point functions exactly correspond to the propagators that appear in the diagrams contributing to the correlation function (2.24) [cf. eqs. (2.22),(2.25)].

At higher orders in the expansion, a series of Feynman diagrams is obtained, as usual, with the above propagators and interaction vertices given by the interaction terms $\mathcal{R}_k$. The Feynman rules for the vertices are the same as in sect. 2.4, apart from the fact that the diagrams have inward-directed external lines if the correlation function considered includes a product of the Lagrange-multiplier fields. Since there is only finite number of diagrams that contributes to a given correlation function, the expansion terminates (as already mentioned). The functional integral is thus a rather peculiar one.

Algebraically the finiteness of the expansion derives from the fact that the label of the outgoing line of each vertex must be larger than the labels of the ingoing lines. In particular, there cannot be any closed loops of directed lines, since the line label increases from one vertex to the next. Each vertex is thus connected to an outward-directed external lines through a path of directed lines. The labels of the external lines therefore set an upper limit on the number of vertices in the diagram, the labels of their outgoing lines and, consequently, on the number of non-zero diagrams.

5. Absence of power divergences

The discussion in the previous section reveals that the correlation functions (2.24) are those of local field theory in $4+1$ dimensions. Since the perturbation expansion of this theory has the standard form, with well-behaved propagators and vertices, the singularities of the associated Feynman diagrams can be analyzed and subtracted as usual, using power-counting and local counterterms.

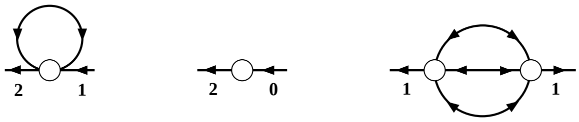


Fig. 2. Diagrams contributing to the $L_2\phi_1$, $L_2\phi_0$ and L_1L_1 vertex functions (from left to right). Amputated external lines are graphically distinguished from ordinary external lines by omitting the little square at their outer end (cf. fig. 1).

5.1 Vertex functions

The correlation functions of the fields ϕ_k and L_j can be decomposed into connected, 1-particle irreducible parts, the vertex functions of the theory, in the standard manner. In perturbation theory, the vertex functions are given by sums of connected, 1-particle irreducible diagrams with amputated external lines. Examples of such diagrams are drawn in fig. 2.

The (amputated) external lines of the vertex diagrams are either outward or inward directed. There are no undirected external lines. In the diagrams contributing to the full correlation functions, an outward directed line of a vertex subdiagram with label k is the L_k-end of the $\phi_k L_k$ propagator (4.5). Inward directed lines are the ϕ_k-end of the same propagator or of the $\phi_0\phi_0$ propagator (4.6) (if $k = 0$).

Diagrams contributing to a vertex function must have at least one vertex and therefore at least one outward directed line with label $k > 0$ (cf. discussion at the end of sect. 4). Moreover, the diagrams have no outward directed lines with label 0, since there is no vertex proportional to L_0.

5.2 Structure of the power-divergent terms

Power-counting now shows that the primitive degree of divergence of the vertex diagrams with n_1 external L_k and n_2 external ϕ_j lines is equal to $6 - 3n_1 - n_2$. Since the action (4.4) is invariant under a change of sign of all fields, the non-zero diagrams must have an even number of external lines. Moreover, as noted above, n_1 must be positive. The only power-divergent diagrams are therefore those with $n_1 = n_2 = 1$, all other diagrams being at most logarithmically divergent.

It follows from the discussion so far that the power divergences in the correlation functions of the L_k and ϕ_j fields, if any, can be cancelled by adding the counterterms

$$\int \mathrm{d}t\, a^4 \sum_x \sum_{k>j=0}^{n} c_{kj} L_k \phi_j \tag{5.1}$$

to the action (4.4) with the appropriate quadratically divergent coefficients c_{kj}.

5.3 Fixed-order two-point functions

As already mentioned, the fixed-order two-point correlation functions

$$\sum_{j=0}^{k} \langle \tilde{\phi}_{k-j}(\omega, p) \tilde{\phi}_j(\nu, q) \rangle = (2\pi)^5 \delta(\omega + \nu) \delta_P(p+q) \tilde{G}_k(\omega, p) \tag{5.2}$$

are at most logarithmically divergent before the counterterms (5.1) are added to the theory. Once the counterterms are included, their contribution to these combinations of correlation functions must be finite or at most logarithmically divergent, since there are no power-divergent terms to be canceled.

The structure of the counterterms (5.1) is such that they make no contribution to the fixed-order correlation function (5.2) if $k = 0$. For $k = 1$ their contribution is

$$-c_{10} \left\{ \tilde{K}(\omega, p) \tilde{G}(-\omega, -p) + \tilde{G}(\omega, p) \tilde{K}(-\omega, -p) \right\}, \tag{5.3}$$

which shows that the coefficient c_{10} must vanish.

In essentially the same way, the vanishing of the coefficients c_{kj} can be proved recursively. Assuming the coefficients c_{lj} are known to vanish for all $l < k$, the only two-point functions in eq. (5.2) to which the counterterms (5.1) can contribute are $\langle \tilde{\phi}_k(\omega, p) \tilde{\phi}_0(\nu, q) \rangle$ and $\langle \tilde{\phi}_0(\omega, p) \tilde{\phi}_k(\nu, q) \rangle$. Explicitly, their contributions are

$$-\sum_{j=0}^{k-1} c_{kj} \left\{ \tilde{K}(\omega, p) \tilde{H}_j(-\omega, -p) + \tilde{H}_j(\omega, p) \tilde{K}(-\omega, -p) \right\}, \tag{5.4}$$

where $\tilde{H}_j(\omega, p)$ denotes the $\phi_0 \phi_j$ two-point function in frequency-momentum space. Since the index j is always less than k, the recursion hypothesis implies that these two-point functions are at most logarithmically divergent. Moreover, the functions multiplying the coefficients c_{kj} in eq. (5.4) are linearly independent (at $p = 0$, for example, their poles at $\omega = \pm i m^2$ have different multiplicities). In order to guarantee that $\tilde{G}_k(\omega, p)$ remains free of power divergences, the coefficients c_{kj} must therefore all be equal to zero.

5.4 Remarks

The proof of the absence of power divergences given in this section rests on the fact that the correlation functions of the fields ϕ_k are those of a local field theory. This property strongly constrains the structure the power divergent parts of the Feynman diagrams.

Another property of the theory, which was heavily used, is the time-ordering that derives from the Langevin equation. On the level of the Feynman diagrams, the time-ordering is associated with the directed lines and the particular ordering of the vertices along paths of directed lines. The theory in 4+1 dimensions is, for this reason, of a rather special kind, where the perturbation expansion terminates and the counterterms (5.1) cannot be arbitrarily inserted in the Feynman diagrams.

References

[1] J. Zinn–Justin, *Renormalization and stochastic quantization*, Nucl. Phys. B275 [FS17] (1986) 135

[2] J. Zinn–Justin, D. Zwanziger, *Ward identities for the stochastic quantization of gauge fields*, Nucl. Phys. B295 [FS21] (1988) 297

[3] F. Di Renzo, G. Marchesini, P. Marenzoni, E. Onofri, *Lattice perturbation theory on the computer*, Nucl. Phys. Proc. Suppl. 34, 795 (1994).

[4] F. Di Renzo, E. Onofri, G. Marchesini, P. Marenzoni, *Four loop result in SU(3) lattice gauge theory by a stochastic method: Lattice correction to the condensate*, Nucl. Phys. B426, 675 (1994).

[5] F. Di Renzo, L. Scorzato, *Numerical stochastic perturbation theory for full QCD*, JHEP 10 (2004) 073.

[6] M. Dalla Brida, M. Garofalo, A. D. Kennedy, *Numerical stochastic perturbation theory and gradient flow in ϕ^4 theory*, PoS LATTICE2015 (2016) 309

[7] M. Lüscher, *Instantaneous stochastic perturbation theory*, JHEP 1504 (2015) 142.